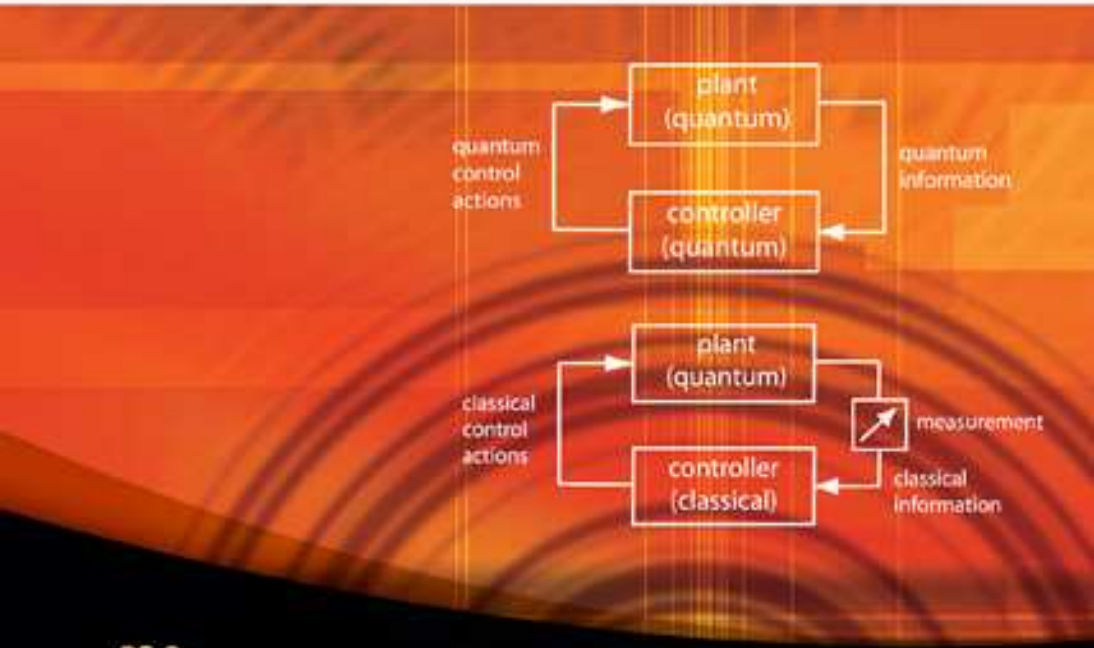
QUANTUM STOCHASTICS AND INFORMATION

Statistics, Filtering and Control

Editors

V. P. Belavkin & M. Guţă



World Scientific

QUANTUM STOCHASTICS AND INFORMATION

Statistics, Filtering and Control

This page intentionally left blank

QUANTUM STOCHASTICS AND INFORMATION

Statistics, Filtering and Control

University of Nottingham, UK 15 – 22 July 2006

Editors

V. P. Belavkin
M. Guţă
University of Nottingham, UK

World Scientific
NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data

Quantum stochastics and information : statistics, filtering, and control : University of Nottingham, UK, 15–22 July 2006 / edited by V. Belavkin & M. Guţă.

p. cm.

Includes index.

ISBN-13: 978-981-283-295-5 (hardcover : alk. paper)

ISBN-10: 981-283-295-5 (hardcover : alk. paper)

1. Quantum statistics--Congresses. 2. Control theory--Congresses. 3. Mathematical physics--Congresses. I. Belavkin, V. P. II. Guţă, M. (Madalin), 1973–

QC174.4.A1Q36 2008

530.13'3--dc22

2008024091

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Printed in Singapore.

PREFACE

Quantum Probability was initiated by von Neumann as a new, intrinsically probabilistic Hilbert space theory of operator observables. Starting with the 80's of the last century the field was greatly enlarged and developed into a non-commutative analogue of Kolmogorov's theory of stochastic processes. One of its aims is to clarify the probabilistic foundations of quantum theory and its statistical interpretation. The basic problems centre around non-commutative independence (including freeness), quantum dynamic dependence (entanglement) and quantum stochastic dependence (non-commutative Markov processes). Quantum Probability is now a heterogeneous subject with many connections to other areas of mathematics and physics. On the pure mathematics side, it interacts with infinite dimensional functional and non-commutative harmonic analysis, operator algebra and operator space theory, non-commutative dynamics, non-commutative K-theory.

The first results on Quantum Information and Quantum Statistics date back to Stratonovich in the 60's and Helstrom, Belavkin and Holevo in the 70's. In the 80's Belavkin developed the dynamical theory of Quantum Feedback Control based on quantum probabilistic foundations. Currently, these fields enjoy a rapid expansion and have a direct impact on technological advances in Quantum Physics and Engineering. Significant applications to physics include quantum metrology, control of quantum networks, validation of quantum state preparation through statistical inference. The most recent advances along these lines are based on quantum stochastic filtering theory as application of quantum stochastic calculus, all developed in Nottingham and presented in this volume. Quantum filtering describes the interface of classical events with quantum mechanics in the new Eventum Mechanics, the information dynamics based on constructive quantum stochastic models of the unified quantum-classical mechanics. It gives the dynamical solution of the long standing quantum measurement problem, resolving many famous paradoxes of quantum physics.

The *Quantum Probability, Information and Control Symposium* was organised during July 15-24, 2006 as the reunion of two closely related conferences. Quantum Probability and Applications - QP 27, was the last of a series of network events of the EU Research Training Network 'Quantum Probability and Applications to Physics, Information and Biology'. The second event was dedicated to Quantum Information, Quantum Statistics, Filtering and Control - QIC and was supported financially by the London Mathematical Society and the Engineering and Physical Sciences Research Council. The 60's birthday of Professor Belavkin was celebrated in a specially dedicated session.

This proceedings volume contains a broad and representative selection of papers on subjects ranging from Quantum Control, Statistics and Probability, to Operator Algebras and Non-commutative Analysis. The unifying paradigm is the further integration between well established classical fields such as Information Theory, Filtering and Control, Statistical Inference, and their quantum counterparts, using the language of Hilbert space operators and the probabilistic interpretation of quantum mechanics. This synthesis is catalysed by recent developments and promising applications in Quantum Engineering.

We thank John Gough, Martin Lindsay and Joachim Zacharias for their support in the organisation of the Symposium. We also thank the members of the Scientific Committee for their help in drawing up the conference program.

V.P. Belavkin
M. Guţă

Nottingham, U.K.
1 May 2008

ORGANIZING COMMITTEES

Scientific Advisory Committee

L. Accardi, S. Albeverio, M. D'Ariano, O. Barndorff-Nielsen, M. Bożejko, U. Franz, O. Hirota, R.L. Hudson, M. James, A. Khrennikov, H. Maassen, D. Petz, M. Schürmann, H. Yuen

Local organisers

V.P. Belavkin, J. Gough, M. Guţă, M. Lindsay, J. Zacharias

This page intentionally left blank

CONTENTS

This page intentionally left blank

PART A

Quantum Probability and Analysis

This page intentionally left blank

APPROXIMATION VIA TOY FOCK SPACE — THE VACUUM-ADAPTED VIEWPOINT

ALEXANDER C. R. BELTON

School of Mathematical Sciences,
University College, Cork, Ireland
E-mail: a.belton@ucc.ie

Dedicated to Slava Belavkin on the occasion of his 60th birthday

After a review of how Boson Fock space (of arbitrary multiplicity) may be approximated by a countable Hilbert-space tensor product (known as toy Fock space) it is shown that vacuum-adapted multiple quantum Wiener integrals of bounded operators may be expressed as limits of sums of operators defined on this toy space, with strong convergence on the exponential domain. The vacuum-adapted quantum Itô product formula is derived with the help of this approximation and a brief pointer is given towards the unbounded case.

Keywords: Discrete approximation; quantum stochastic integration; vacuum adaptedness.

1. Introduction

The idea of using discrete approximations in quantum stochastic calculus goes back at least as far as Meyer's notes [16], where he gave credit to Jean-Lin Journé. Around the same time, articles by Parthasarathy [18] and Lindsay and Parthasarathy [15] showed that certain quantum flows (which are generalisations of classical diffusions) may be approximated by so-called spin random walks, while Accardi and Bach produced (in an unpublished preprint — see [1,17]) a central-limit theorem which may be viewed as a result on toy-Fock-space approximation. These ideas have recently been the subject of renewed interest.

Attal revisited and extended the Journé–Meyer ideas in [2], giving a heuristic derivation of the quantum Itô product formula using the approximation, and this was followed by further work of Attal and Pautrat [3] and of Pautrat [19]. Their point of view may be considered as physical, rather

than probabilistic; in [9], Gough examined the physics of this set-up and explained its connexion with Holevo's time-ordered exponentials.

Meanwhile, Sinha [21] revived the ideas of [15], emphasising that, in many cases, sufficiently strong convergence holds to enable one to deduce that the limit flows are $*$-homomorphic. Further work in this direction has done by Sahu [20], who moved away from the spin approach of Lindsay–Parthasarathy–Sinha to adopt the same type of coupling between system and noise as Attal–Pautrat; it is this direct (as opposed to spin) coupling which is used below.

Many other people have worked with these concepts, including Bouten, van Handel and James [6] (in quantum filtering), Brun [7] and Gough and Sobolev [10] (who view the situation as physicists), Franz and Skalski [8] (for constructing random walks on quantum groups), Kümmerer ([12] gives a detailed physical interpretation of discrete models and is an excellent introduction to his earlier work in this area) and Leitz-Martini [13] (who expressed many of these approximation ideas rigorously using non-standard analysis; for example, the discrete Itô table of Attal [2, Section VII] agrees with the continuous-time version only in the limit, but in the non-standard setting the anomalous terms are infinitesimal [13, (2.2.8)]).

Here, a vacuum-adapted approach to approximation is adopted and, as might be expected, a very straightforward theory results. After revising the embedding of toy Fock space into Boson Fock space in Section 2, modified versions of the vacuum-adapted Wiener integral are defined in Section 3. The natural 'discrete integral' (which is, of course, a sum) is examined in Section 4 and is shown to be given, up to an error term, by the modified integral previously defined. Section 5 extends this working to the case of multiple integrals, Section 6 shows how the quantum Itô product formula arises naturally from the discrete approximation and Section 7 points the way to further developments involving unbounded operators. Applications of these results will appear elsewhere [5].

1.1. *Conventions and Notation*

All sesquilinear inner products are conjugate linear in the first variable. We follow [14] for the most part, although the ordering of certain objects is changed: for us, the initial space always appears first (the 'usual' convention, to quote Lindsay [14, p.183]).

The vector space of linear operators between vector spaces U and V is denoted by $\mathcal{L}(U;V)$, or $\mathcal{L}(U)$ if U and V are equal; the identity operator on a vector space V is denoted by I_V. The operator space of bounded operators

between Hilbert spaces H_1 and H_2 is denoted by $\mathcal{B}(\mathsf{H}_1;\mathsf{H}_2)$, abbreviated to $\mathcal{B}(\mathsf{H}_1)$ if H_1 equals H_2. The tensor product of Hilbert spaces and bounded operators is denoted by $\otimes$; the algebraic tensor product is denoted by $\odot$. The restriction of a function f to a set A is denoted by $f|_A$; the indicator function of A is denoted by 1_A. If P is a proposition then the expression $\mathbb{1}_P$ has the value 1 if P is true and 0 if P is false.

2. Toy and Boson Fock spaces

> Men more frequently require to be reminded than informed.
> – Samuel Johnson, *The Rambler*, No. 2 (1749–50).

Notation 2.1. Let k be a complex Hilbert space (called the *multiplicity space*) and let $\widehat{\mathsf{k}} := \mathbb{C} \oplus \mathsf{k}$ be its one-dimensional extension. Elements of $\widehat{\mathsf{k}}$ will be written as column vectors, with the first entry a complex number and the second a vector in k; if $x \in \mathsf{k}$ then $\widehat{x} := \left(\begin{smallmatrix}1\\x\end{smallmatrix}\right)$.

Definition 2.1. *Toy Fock space* is the countable tensor product

$$\Gamma := \bigotimes_{n=0}^{\infty} \widehat{\mathsf{k}}_{(n)} \tag{1}$$

with respect to the stabilising sequence $\left(\omega_{(n)} := \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\right)_{n=0}^{\infty}$, where $\widehat{\mathsf{k}}_{(n)} := \widehat{\mathsf{k}}$ for each n; the subscript (n) is used here and below to indicate the relevant copy. (For information on infinite tensor products of Hilbert spaces, see, for example [11, Exercise 11.5.29].)

For all $n \in \mathbb{Z}_+ := \{0, 1, 2, \ldots\}$, let

$$\Gamma_{n)} := \bigotimes_{m=0}^{n-1} \widehat{\mathsf{k}}_{(m)} \quad \text{and} \quad \Gamma_{[n} := \bigotimes_{m=n}^{\infty} \widehat{\mathsf{k}}_{(m)}, \tag{2}$$

where $\Gamma_{0)} := \mathbb{C}$. The identity $\Gamma = \Gamma_{n)} \otimes \Gamma_{[n}$ is the analogue of the continuous tensor-product structure of Boson Fock space.

Notation 2.2. For any interval $A \subseteq \mathbb{R}_+$, let $\mathcal{F}_A$ denote Boson Fock space over $L^2(A;\mathsf{k})$ and let $\mathcal{F} := \mathcal{F}_{\mathbb{R}_+}$. For further brevity, let $\mathsf{K} = L^2(\mathbb{R}_+;\mathsf{k})$.

If $\tau := \{0 = \tau_0 < \tau_1 < \cdots < \tau_n < \cdots\}$ is a partition of $\mathbb{R}_+$ (so that $\tau_n \to \infty$ as $n \to \infty$) then there exists an isometric isomorphism

$$\Pi_\tau : \mathcal{F} \xrightarrow{\cong} \mathcal{F}_\tau := \bigotimes_{n=0}^{\infty} \mathcal{F}_{[\tau_n,\tau_{n+1}[};\ \varepsilon(f) \mapsto \bigotimes_{n=0}^{\infty} \varepsilon(f|_{[\tau_n,\tau_{n+1}[}), \tag{3}$$

where the tensor product is taken with respect to the stabilising sequence $\left(\Omega_{[\tau_n,\tau_{n+1}[} := \varepsilon(0|_{[\tau_n,\tau_{n+1}[})\right)_{n=0}^{\infty}$ and $\varepsilon(g)$ denotes the exponential vector in

$\mathcal{F}_A$ corresponding to the function $g \in L^2(A;\mathsf{k})$. The set of all such partitions of $\mathbb{R}_+$ is denoted by T.

Definition 2.2. For all $\tau \in \mathrm{T}$ and $n \in \mathbb{Z}_+$, define the natural isometry

$$j[\tau]_n : \widehat{\mathsf{k}} \to \mathcal{F}_{[\tau_n,\tau_{n+1}[};\ \tbinom{\lambda}{x} \mapsto \lambda\Omega_{[\tau_n,\tau_{n+1}[} + x\widetilde{1}_{[\tau_n,\tau_n+1[}, \tag{4}$$

where $\widetilde{1}_A := 1_A/\|1_A\|_{L^2(\mathbb{R}_+)}$ is the normalised indicator function of the interval $A \subseteq \mathbb{R}_+$, viewed as an element of the one-particle subspace of $\mathcal{F}_A$. These give an isometric embedding

$$J_\tau : \Gamma \to \mathcal{F}_\tau;\ \bigotimes_{n=0}^{\infty}\theta_n \mapsto \bigotimes_{n=0}^{\infty} j[\tau]_n(\theta_n). \tag{5}$$

Note that $Q_\tau := \Pi_\tau^* J_\tau J_\tau^* \Pi_\tau$ is an orthogonal projection on $\mathcal{F}$ and

$$J_\tau^*\Pi_\tau\varepsilon(f) = \bigotimes_{n=0}^{\infty}\widehat{f_\tau(n)} \qquad \forall\, f \in \mathsf{K}, \tag{6}$$

where

$$f_\tau(n) := \frac{1}{\sqrt{\tau_{n+1}-\tau_n}}\int_{\tau_n}^{\tau_{n+1}} f(t)\,\mathrm{d}t \qquad \forall\, n \in \mathbb{Z}_+. \tag{7}$$

Notation 2.3. Let T be the directed set of all partitions of $\mathbb{R}_+$, ordered by inclusion; the expression '$f_\tau \to f$ as $|\tau| \to 0$' means that the net $(f_\tau)_{\tau\in\mathrm{T}}$ converges to f. For all $\tau \in \mathrm{T}$, let $P_\tau \in \mathcal{B}(\mathsf{K})$ be the orthogonal projection given by

$$P_\tau f := \sum_{n=0}^{\infty}\frac{1}{\tau_{n+1}-\tau_n}\int_{\tau_n}^{\tau_{n+1}} f(t)\,\mathrm{d}t\, 1_{[\tau_n,\tau_{n+1}[} \qquad \forall\, f \in \mathsf{K}. \tag{8}$$

Lemma 2.1. *The projection P_τ converges strongly to I_{K} as $|\tau| \to 0$.*

Proof. If $f \in \mathsf{K}$ is continuous and compactly supported, a uniform-continuity argument may be used to show that $P_\tau f \to f$ uniformly; the density of such functions in K completes the proof. □

Theorem 2.1. *As $|\tau| \to 0$, the projection Q_τ converges strongly to $I_{\mathcal{F}}$.*

Proof. By (6) and (8), if f, $g \in \mathsf{K}$ then (compare [18, (2.10)])

$$\begin{aligned}\langle \varepsilon(f), Q_\tau \varepsilon(g)\rangle &= \prod_{n=0}^{\infty}\Big(1 + \langle f_\tau(n), g_\tau(n)\rangle\Big)\\ &= \exp\Big(\sum_{n=0}^{\infty}\log\Big(1 + \int_{\tau_n}^{\tau_{n+1}}\langle f(t), P_\tau g(t)\rangle\,\mathrm{d}t\Big)\Big)\\ &\sim \exp\Big(\sum_{n=0}^{\infty}\int_{\tau_n}^{\tau_{n+1}}\langle f(t), P_\tau g(t)\rangle\,\mathrm{d}t\Big) \qquad (9)\\ &\to \exp\Big(\int_0^{\infty}\langle f(t), g(t)\rangle\,\mathrm{d}t\Big) \qquad \text{as } |\tau| \to 0,\end{aligned}$$

by Lemma 2.1, so $Q_\tau \to I_{\mathcal{F}}$ weakly on $\mathcal{E}$, the linear span of the set of exponential vectors; the asymptotic identity (9) holds because $\log(1+z) = z + O(z^2)$ as $z \to 0$. Since each Q_τ is an orthogonal projection, strong convergence on $\mathcal{E}$, so on $\mathcal{F}$, follows. □

3. Modified QS integrals

Natura abhorret vacuum.
– François Rabelais, *Gargantua et Pantagruel*, Bk. 1, Ch. 5 (1534).

To examine the behaviour of the discrete approximations which will be constructed in the following sections, it is useful first to introduce a slight extension of the iterated QS integral (QS being, of course, an abbreviation for quantum stochastic).

Notation 3.1. Let h be a fixed complex Hilbert space (the *initial space*) and let $\widetilde{\mathcal{F}} := \mathsf{h} \otimes \mathcal{F}$, $\widetilde{\Gamma} := \mathsf{h} \otimes \Gamma$ and $\widetilde{\mathcal{E}} := \mathsf{h} \odot \mathcal{E}$. As is customary, the tensor sign will be omitted before exponential vectors: $u\varepsilon(f) := u \otimes \varepsilon(f)$.

Definition 3.1. Given a Hilbert space H, an H-*process* $X = (X_t)_{t\in\mathbb{R}_+}$ is a weakly measurable family of linear operators with common domain $\mathsf{H} \odot \mathcal{E}$, *i.e.*,

$$X_t \in \mathcal{L}(\mathsf{H} \odot \mathcal{E}; \mathsf{H} \otimes \mathcal{F}) \qquad \forall\, t \in \mathbb{R}_+ \qquad (10)$$

and $t \mapsto \langle u\varepsilon(f), X_t v\varepsilon(g)\rangle$ is measurable for all u, $v \in \mathsf{H}$ and f, $g \in \mathsf{K}$.

An H-process X is *vacuum adapted* if

$$\langle u\varepsilon(f), X_t v\varepsilon(g)\rangle = \langle u\varepsilon(1_{[0,t[}f), X_t v\varepsilon(1_{[0,t[}g)\rangle \qquad (11)$$

for all $t \in \mathbb{R}_+$, u, $v \in \mathsf{H}$ and f, $g \in \mathsf{K}$. Equivalently, the identity $X_t = (I_{\mathsf{H}} \otimes \mathbb{E}_t)X_t(I_{\mathsf{H}} \otimes \mathbb{E}_t)$ holds for all $t \in \mathbb{R}_+$, where $\mathbb{E}_t \in \mathcal{B}(\mathcal{F})$ is the second quantisation of the multiplication operator $f \mapsto 1_{[0,t[}f$ on K.

An H-process X is *semi-vacuum-adapted* if $(I_{\mathsf{H}} \otimes \mathbb{E}_t)X_t = X_t$ for all $t \in \mathbb{R}_+$; clearly every vacuum-adapted process is semi-vacuum-adapted.

If $M \in \mathcal{B}\big(\mathsf{K}\big)$ then an $\mathsf{h} \otimes \widehat{\mathsf{k}}$-process X is *M-integrable* if

$$\|X\widehat{\nabla}^M\theta\|^2_{L^2([0,t[;\mathsf{h}\otimes\widehat{\mathsf{k}}\otimes\mathcal{F})} = \int_0^t \|X_s\widehat{\nabla}_s^M\theta\|^2\,\mathrm{d}s < \infty \qquad \forall\,\theta \in \widetilde{\mathcal{E}},\ t \in \mathbb{R}_+, \tag{12}$$

where the *modified gradient* $\widehat{\nabla}^M : \widetilde{\mathcal{E}} \to \widetilde{\mathcal{F}} \oplus (\mathsf{h}\otimes\mathsf{K}\otimes\mathcal{F})$ is the linear operator such that

$$u\varepsilon(f) \mapsto [u \otimes \widehat{Mf}]\varepsilon(f) = \begin{pmatrix} u\varepsilon(f) \\ [u \otimes Mf]\varepsilon(f)\end{pmatrix} \tag{13}$$

and the definition $\widehat{\nabla}_s^M u\varepsilon(f) := [u \otimes \widehat{Mf(s)}]\varepsilon(f)$ is extended by linearity.

Notation 3.2. Let $\Delta \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}} \otimes \mathcal{F})$ be the orthogonal projection onto $\mathsf{h} \otimes \mathsf{k} \otimes \mathcal{F}$ and $\Delta^\perp := I_{\mathsf{h}\otimes\widehat{\mathsf{k}}\otimes\mathcal{F}} - \Delta$ the projection onto its complement, $\widetilde{\mathcal{F}}$.

Theorem 3.1. *Let $M \in \mathcal{B}\big(\mathsf{K}\big)$. If X is an M-integrable, semi-vacuum-adapted $\mathsf{h}\otimes\widehat{\mathsf{k}}$-process, there exists a unique semi-vacuum-adapted h-process $\Lambda_\Omega(X;M)$, the* modified QS integral of X, *such that, for all $t \in \mathbb{R}_+$,*

$$\|\Lambda_\Omega(X;M)_t\theta\| \le c_t\|X\widehat{\nabla}^M\theta\|_{L^2([0,t[;\mathsf{h}\otimes\widehat{\mathsf{k}}\otimes\mathcal{F})} \qquad \forall\,\theta \in \widetilde{\mathcal{E}}, \tag{14}$$

where $c_t := \sqrt{2\max\{t,1\}}$, and

$$\langle u\varepsilon(f), \Lambda_\Omega(X;M)_t v\varepsilon(g)\rangle = \int_0^t \langle [u \otimes \widehat{f(s)}]\varepsilon(f), X_s\big([v \otimes \widehat{Mg(s)}]\varepsilon(g)\big)\rangle\,\mathrm{d}s \tag{15}$$

for all u, $v \in \mathsf{h}$ and f, $g \in \mathsf{K}$.

Proof. This follows from the behaviour of the Bochner integral and the abstract Itô integral: for all $t \in \mathbb{R}_+$ and $\theta \in \widetilde{\mathcal{E}}$ let

$$\Lambda_\Omega(X;M)_t\theta := \int_0^t \Delta^\perp X_s\widehat{\nabla}_s^M\theta\,\mathrm{d}s + \mathcal{I}_t(\Delta X\widehat{\nabla}^M\theta), \tag{16}$$

where $\mathcal{I}_t$ is the Itô integral on $[0,t[$ (the adjoint of the adapted gradient). As $s \mapsto \Delta X_s\widehat{\nabla}_s^M\theta$ is an adapted vector process, *i.e.*, $\Delta X_s\widehat{\nabla}_s^M\theta \in \mathsf{h}\otimes\mathsf{k}\otimes\mathcal{F}_{[0,s[}$ for (almost) all $s \in \mathbb{R}_+$, this is a good definition, and the isometric nature of the Itô integral [4, Proposition 3.28] implies that

$$\begin{aligned}\|\Lambda_\Omega(X;M)_t\theta\|^2 &\le 2t\int_0^t \|\Delta^\perp X_s\widehat{\nabla}_s^M\theta\|^2\,\mathrm{d}s + 2\int_0^t \|\Delta X_s\widehat{\nabla}_s^M\theta\|^2\,\mathrm{d}s \\ &\le c_t^2\int_0^t \|X_s\widehat{\nabla}_s^M\theta\|^2\,\mathrm{d}s,\end{aligned} \tag{17}$$

as claimed. The identity (15) follows immediately and yields semi-vacuum-adaptedness. □

Remark 3.1. As may be seen from (15), the modified integral preserves semi-vacuum-adaptedness but need not preserve vacuum-adaptedness. This identity also shows that if $A \in \mathcal{B}(\mathsf{h})$ commutes with X, in the sense that

$$X_t(A \otimes I_{\widehat{\mathsf{k}\otimes\mathcal{F}}}) = (A \otimes I_{\widehat{\mathsf{k}\otimes\mathcal{F}}})X_t \qquad \forall\, t \in \mathbb{R}_+, \tag{18}$$

then A commutes with $\Lambda_\Omega(X;M)$ in the same sense: $\Lambda_\Omega(X;M)_t(A\otimes I_{\mathcal{F}}) = (A \otimes I_{\mathcal{F}})\Lambda_\Omega(X;M)_t$ for all $t \in \mathbb{R}_+$.

Notation 3.3. For all $n \geq 1$ and $t \in \mathbb{R}_+$, let

$$\Delta_n(t) := \{\mathbf{t} := (t_1, \ldots, t_n) \in [0,t[^n : t_1 < \cdots < t_n\} \subseteq \mathbb{R}_+^n \tag{19}$$

and, given $M \in \mathcal{B}(\mathsf{K})$, define $(\widehat{\nabla}^M)^n \in \mathcal{L}\big(\widetilde{\mathcal{E}}; L^2(\Delta_n(t); \mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n} \otimes \mathcal{F})\big)$ such that

$$(\widehat{\nabla}^M)^n_{\mathbf{t}} u\varepsilon(f) := \big((\widehat{\nabla}^M)^n u\varepsilon(f)\big)(\mathbf{t}) := [u \otimes \widehat{Mf}^{\otimes n}(\mathbf{t})]\varepsilon(f) \tag{20}$$

for all $u \in \mathsf{h}$, $f \in \mathsf{K}$ and $\mathbf{t} \in \Delta_n(t)$, where $\widehat{g}^{\otimes n}(\mathbf{t}) := \widehat{g(t_1)} \otimes \cdots \otimes \widehat{g(t_n)}$ for all $g \in \mathsf{K}$ and $\mathbf{t} \in \mathbb{R}_+^n$.

Theorem 3.2. *Let $M \in \mathcal{B}(\mathsf{K})$. If $n \geq 1$, $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$ and Y is a locally uniformly bounded, semi-vacuum-adapted $\mathbb{C}$-process then there exists a unique semi-vacuum-adapted h-process $\Lambda^n_\Omega(X \otimes Y; M)$, the* modified n-fold QS integral*, such that, for all $t \in \mathbb{R}_+$,*

$$\|\Lambda^n_\Omega(X \otimes Y; M)_t\theta\|^2 \leq c_t^{2n} \int_{\Delta_n(t)} \|(X \otimes Y_{t_1})(\widehat{\nabla}^M)^n_{\mathbf{t}}\theta\|^2 \,\mathrm{d}\mathbf{t} \tag{21}$$

for all $\theta \in \widetilde{\mathcal{E}}$ and

$$\begin{aligned}&\langle u\varepsilon(f), \Lambda^n_\Omega(X \otimes Y; M)_t v\varepsilon(g)\rangle \\ &\qquad = \int_{\Delta_n(t)} \langle u \otimes \widehat{f}^{\otimes n}(\mathbf{t}), X[v \otimes \widehat{Mg}^{\otimes n}(\mathbf{t})]\rangle \langle \varepsilon(f), Y_{t_1}\varepsilon(g)\rangle \,\mathrm{d}\mathbf{t}\end{aligned} \tag{22}$$

for all u, $v \in \mathsf{h}$ and f, $g \in \mathsf{K}$.

Proof. If $n = 1$ then the result follows by applying Theorem 3.1 to the process $X \otimes Y : t \mapsto X \otimes Y_t$. Now suppose the theorem holds for a particular $n \geq 1$, let $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n+1})$ and define $X' := \widetilde{R}^*_{n+1} X \widetilde{R}_{n+1}$, where the unitary map $\widetilde{R}_{n+1} : \mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n+1} \to \mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n+1}$ implements the permutation

$$u \otimes x_1 \otimes x_2 \otimes \cdots \otimes x_{n+1} \mapsto u \otimes x_2 \otimes \cdots \otimes x_{n+1} \otimes x_1.$$

By replacing h with $\mathsf{h} \otimes \widehat{\mathsf{k}}$, this assumption yields a semi-vacuum-adapted $\mathsf{h} \otimes \widehat{\mathsf{k}}$-process $\Lambda^n_\Omega(X' \otimes Y; M)$ such that, for all $t \in \mathbb{R}_+$, u, $v \in \mathsf{h}$, x, $y \in \widehat{\mathsf{k}}$ and f, $g \in \mathsf{K}$,

$$\begin{aligned}
&\langle (u \otimes x)\varepsilon(f), \Lambda^n_\Omega(X' \otimes Y; M)_t (v \otimes y)\varepsilon(g)\rangle \\
&\quad = \int_{\Delta_n(t)} \langle u \otimes x \otimes \widehat{f}^{\otimes n}(\mathbf{t}), X'[v \otimes y \otimes \widehat{Mg}^{\otimes n}(\mathbf{t})]\rangle \langle \varepsilon(f), Y_{t_1}\varepsilon(g)\rangle \,\mathrm{d}\mathbf{t} \\
&\quad = \int_{\Delta_n(t)} \langle u \otimes \widehat{f}^{\otimes n}(\mathbf{t}) \otimes x, X[v \otimes \widehat{Mg}^{\otimes n}(\mathbf{t}) \otimes y]\rangle \langle \varepsilon(f), Y_{t_1}\varepsilon(g)\rangle \,\mathrm{d}\mathbf{t}.
\end{aligned}$$

Letting

$$\Lambda^{n+1}_\Omega(X \otimes Y; M) := \Lambda_\Omega(\Lambda^n_\Omega(X' \otimes Y; M); M) \tag{23}$$

gives the result: if $t \in \mathbb{R}_+$, u, $v \in \mathsf{h}$ and f, $g \in \mathsf{K}$ then

$$\begin{aligned}
&\langle u\varepsilon(f), \Lambda^{n+1}_\Omega(X \otimes Y; M)_t v\varepsilon(g)\rangle \\
&= \int_0^t \langle [u \otimes \widehat{f(s)}]\varepsilon(f), \Lambda^n_\Omega(X' \otimes Y; M)_s\big([v \otimes \widehat{Mg(s)}]\varepsilon(g)\big)\rangle \,\mathrm{d}s \\
&= \int_0^t \int_{\Delta_n(s)} \langle u \otimes \widehat{f}^{\otimes n+1}(\mathbf{t}, s), X[v \otimes \widehat{Mg}^{\otimes n+1}(\mathbf{t}, s)]\rangle \langle \varepsilon(f), Y_{t_1}\varepsilon(g)\rangle \,\mathrm{d}\mathbf{t}\,\mathrm{d}s \\
&= \int_{\Delta_{n+1}(t)} \langle u \otimes \widehat{f}^{\otimes n+1}(\mathbf{t}), X[v \otimes \widehat{Mg}^{\otimes n+1}(\mathbf{t})]\rangle \langle \varepsilon(f), Y_{t_1}\varepsilon(g)\rangle \,\mathrm{d}\mathbf{t};
\end{aligned}$$

the norm estimate (21) (and M-integrability of $\Lambda^n_\Omega(X' \otimes Y; M)$) may be shown similarly. □

Proposition 3.1. *If $n \geq 1$, $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$ and $X \otimes \mathbb{E}$ is the vacuum-adapted $\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n}$-process given by setting $(X \otimes \mathbb{E})_t := X \otimes \mathbb{E}_t$ then*

$$\Lambda^n_\Omega(X) := \Lambda^n_\Omega(X \otimes \mathbb{E}; I_{\mathsf{K}}) \tag{24}$$

is a vacuum-adapted, bounded process: each $\Lambda^n_\Omega(X)_t$ extends uniquely to an element of $\mathcal{B}(\widetilde{\mathcal{F}})$, the vacuum-adapted n-fold quantum Wiener integral *of X.*

Proof. Note first that if Z is a locally uniformly bounded, vacuum-adapted $\mathsf{h} \otimes \widehat{\mathsf{k}}$-process and $\theta \in \widetilde{\mathcal{E}}$ then

$$\begin{aligned}
\|Z\widehat{\nabla}\theta\|^2_{L^2([0,t[;\mathsf{h}\otimes\widehat{\mathsf{k}}\otimes\mathcal{F})} &\leq \|Z\|^2_{\infty,t} \int_0^t \|(I_{\mathsf{h}\otimes\widehat{\mathsf{k}}} \otimes \mathbb{E}_s)\widehat{\nabla}_s\theta\|^2 \,\mathrm{d}s \\
&= \|Z\|^2_{\infty,t} \int_0^t \big(\|\mathbb{E}_s\theta\|^2 + \|\mathcal{D}_s\theta\|^2\big) \,\mathrm{d}s \\
&\leq \|Z\|^2_{\infty,t}(t+1)\|\theta\|^2,
\end{aligned} \tag{25}$$

where $\|\cdot\|_{\infty,t}$ is the essential-supremum norm on $[0,t[$, $\widehat{\nabla} := \widehat{\nabla}^{I_{\mathsf{K}}}$ and $\mathcal{D}$ is the adapted gradient on $\widetilde{\mathcal{F}}$ [4, Proposition 3.27]. Hence $\Lambda_\Omega(Z;I_{\mathsf{K}})_t$ extends to a unique element of $\mathcal{B}(\widetilde{\mathcal{F}})$ for all $t \in \mathbb{R}_+$ and $\Lambda_\Omega(Z;I_{\mathsf{K}})$ is a locally uniformly bounded, vacuum-adapted h-process. The result now follows from the inductive construction of $\Lambda^n_\Omega(X)$ given in the proof of Theorem 3.2. □

Proposition 3.2. *Let M, $N \in \mathcal{B}\big(\mathsf{K}\big)$. If $n \geq 1$, $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$, Y is a locally uniformly bounded, semi-vacuum-adapted $\mathbb{C}$-process and $t \in \mathbb{R}_+$ then*

$$\big\|\big(\Lambda^n_\Omega(X\otimes Y;M)_t - \Lambda^n_\Omega(X\otimes Y;N)_t\big)u\varepsilon(f)\big\|^2 \leq 2^{n-1}c_t^{2n}\|Y\|^2_{\infty,t}\|\varepsilon(f)\|^2\sum_{m=1}^n L_m, \quad (26)$$

for all $u \in \mathsf{h}$ and $f \in \mathsf{K}$, where $\|\cdot\|_{\infty,t}$ is the essential-supremum norm on the interval $[0,t[$,

$$L_m := \int_{\Delta_n(t)} \|X\big(u\otimes\widehat{Mf}^{\otimes m-1}(\mathbf{t}_{m)})\otimes[(M-N)f](t_m)\otimes\widehat{Nf}^{\otimes n-m}(\mathbf{t}_{(m)})\big)\|^2\,\mathrm{d}\mathbf{t} \quad (27)$$

$\mathbf{t}_{m)} := (t_1,\ldots,t_{m-1})$ and $\mathbf{t}_{(m} := (t_{m+1},\ldots,t_n)$.

Proof. Note that, with notation as in the proof of Theorem 3.2,

$$\begin{aligned}\Lambda^{n+1}_\Omega(X\otimes Y;M) - \Lambda^{n+1}_\Omega(X\otimes Y;N) &= \Lambda_\Omega(\Lambda^n_\Omega(X'\otimes Y;M);M) - \Lambda_\Omega(\Lambda^n_\Omega(X'\otimes Y;M);N)\\ &\quad + \Lambda_\Omega\big(\big(\Lambda^n_\Omega(X'\otimes Y;M) - \Lambda^n_\Omega(X'\otimes Y;N)\big);N\big).\end{aligned}$$

Now use induction, together with (14), (21) and the fact that

$$\|(\Lambda_\Omega(Z;M)_t - \Lambda_\Omega(Z;N)_t)u\varepsilon(f)\|^2 \leq c_t^2\int_0^t \|Z_s\big([u\otimes[(M-N)f](s)]\varepsilon(f)\big)\|^2\,\mathrm{d}s, \quad (28)$$

by (16). □

Definition 3.2. If A is an ordered set and $n \geq 1$ then $A^{n,\uparrow}$ is the collection of strictly increasing n-tuples of elements of A. Given $\tau \in \mathrm{T}$, let

$$[\tau_{\mathbf{p}},\tau_{\mathbf{p}+1}[:= \{\mathbf{t}\in\mathbb{R}^n_+ : \tau_{p_i} \leq t_i < \tau_{p_i+1}\ (i=1,\ldots,n)\} \quad (29)$$

for all $\mathbf{p} = (p_1,\ldots,p_n) \in \mathbb{Z}^{n,\uparrow}_+$ and, for all $t \in \mathbb{R}_+$, let

$$\Delta^\tau_n(t) := \bigcup_{\mathbf{p}\in\{0,\ldots,m-1\}^{n,\uparrow}} [\tau_{\mathbf{p}},\tau_{\mathbf{p}+1}[\qquad \text{if } t\in[\tau_m,\tau_{m+1}[. \quad (30)$$

Theorem 3.3. *Let $M \in \mathcal{B}(\mathsf{K})$. If $n \geq 1$, $\tau \in \mathrm{T}$, $X \in \mathcal{B}(\mathsf{h}\otimes\widehat{\mathsf{k}}^{\otimes n})$ and Y is a locally uniformly bounded, semi-vacuum-adapted $\mathbb{C}$-process then there exists a unique semi-vacuum-adapted h-process $\Lambda^n_\Omega(X\otimes Y;M)^\tau$, the* modified n-fold QS integral subordinate to τ, *such that, for all $t \in \mathbb{R}_+$,*

$$\|\Lambda^n_\Omega(X\otimes Y;M)^\tau_t\theta\|^2 \leq c_t^{2n}\int_{\Delta^\tau_n(t)} \|(X\otimes Y_{t_1})(\widehat{\nabla}^M)^n_{\mathbf{t}}\theta\|^2\,\mathrm{d}\mathbf{t} \tag{31}$$

for all $\theta \in \widetilde{\mathcal{E}}$ and

$$\begin{aligned}&\langle u\varepsilon(f), \Lambda^n_\Omega(X\otimes Y;M)^\tau_t v\varepsilon(g)\rangle\\ &\qquad= \int_{\Delta^\tau_n(t)} \langle u\otimes \widehat{f}^{\otimes n}(\mathbf{t}), X[v\otimes \widehat{Mg}^{\otimes n}(\mathbf{t})]\rangle\langle\varepsilon(f), Y_{t_1}\varepsilon(g)\rangle\,\mathrm{d}\mathbf{t}\end{aligned} \tag{32}$$

for all u, $v \in \mathsf{h}$ and f, $g \in \mathsf{K}$.

Proof. When $n = 1$, apply Theorem 3.1 to the process $X\otimes Y : t \mapsto X\otimes Y_t$ and let

$$\Lambda^1_\Omega(X\otimes Y;M)^\tau_t := \sum_{m=0}^\infty \mathbb{1}_{t\in[\tau_m,\tau_{m+1}[}\Lambda_\Omega(X\otimes Y;M)_{\tau_m}.$$

Now suppose the theorem holds for a particular $n \geq 1$, let $X \in \mathcal{B}(\mathsf{h}\otimes\widehat{\mathsf{k}}^{\otimes n+1})$ and define $X' := \widetilde{R}^*_{n+1}X\widetilde{R}_{n+1}$ as in the proof of Theorem 3.2. The semi-vacuum-adapted $\mathsf{h}\otimes\widehat{\mathsf{k}}$-process $\Lambda^n_\Omega(X'\otimes Y;M)^\tau$ is such that, for all $t \in \mathbb{R}_+$, u, $v \in \mathsf{h}$, x, $y \in \widehat{\mathsf{k}}$ and f, $g \in \mathsf{K}$,

$$\begin{aligned}&\langle (u\otimes x)\varepsilon(f), \Lambda^n_\Omega(X'\otimes Y;M)^\tau_t (v\otimes y)\varepsilon(g)\rangle\\ &\qquad= \int_{\Delta^\tau_n(t)} \langle u\otimes \widehat{f}^{\otimes n}(\mathbf{t})\otimes x, X[v\otimes \widehat{Mg}^{\otimes n}(\mathbf{t})\otimes y]\rangle\langle\varepsilon(f), Y_{t_1}\varepsilon(g)\rangle\,\mathrm{d}\mathbf{t},\end{aligned}$$

so, as

$$\{0,\dots,m-1\}^{n+1,\uparrow} = \bigcup_{k=0}^{m-1}\big\{(p_1,\dots,p_n,k) : \mathbf{p}\in\{0,\dots,k-1\}^{n,\uparrow}\big\} \tag{33}$$

and therefore $\Delta^\tau_{n+1}(\tau_m) = \bigcup_{k=0}^{m-1}\big(\Delta^\tau_n(\tau_k)\times[\tau_k,\tau_{k+1}[\big)$, letting

$$\Lambda^{n+1}_\Omega(X\otimes Y;M)^\tau := \sum_{m=0}^\infty \mathbb{1}_{t\in[\tau_m,\tau_{m+1}[}\Lambda_\Omega(\Lambda^n_\Omega(X'\otimes Y;M)^\tau;M)_{\tau_m} \tag{34}$$

gives the result. □

Proposition 3.3. *Let $M \in \mathcal{B}(\mathsf{K})$. If $n \geq 1$, $\tau \in \mathrm{T}$, $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$ and Y is a locally uniformly bounded, semi-vacuum-adapted $\mathbb{C}$-process then*

$$\left\|\left(\Lambda_\Omega^n(X \otimes Y; M)_t - \Lambda_\Omega^n(X \otimes Y; M)_t^\tau\right)\theta\right\|^2 \leq 2^{n-1} c_t^{2n} \int_{\Delta_n(t)\setminus\Delta_n^\tau(t)} \|(X \otimes Y_{t_1})(\widehat{\nabla}^M)_{\mathbf{t}}^n \theta\|^2 \, \mathrm{d}\mathbf{t} \quad (35)$$

for all $t \in \mathbb{R}_+$ and $\theta \in \widetilde{\mathcal{E}}$.

Proof. This follows by induction and the fact that if $t \in [\tau_m, \tau_{m+1}[$ then the set

$$\{(\mathbf{t}, s) : s \in [0, \tau_m[,\ \mathbf{t} \in \Delta_n(s) \setminus \Delta_n^\tau(s)\} \cup \{(\mathbf{t}, s) : s \in [\tau_m, t[,\ \mathbf{t} \in \Delta_n(s)\}$$

is contained in $\Delta_{n+1}(t) \setminus \Delta_{n+1}^\tau(t)$. □

4. The toy integral

I see salvation in discrete individuals
– Anton Chekhov, Letter to I.I. Orlov (22nd February, 1899).

Definition 4.1. For all $n \in \mathbb{Z}_+$ let $\widetilde{s}_n : \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}) \to \mathcal{B}(\widetilde{\Gamma})$ be the normal $*$-homomorphism such that $B \otimes C \mapsto B \otimes I_{\Gamma_{n)}} \otimes C \otimes P_{[n+1}^\omega$, where

$$P_{[n+1}^\omega : \Gamma_{[n+1} \to \Gamma_{[n+1}; \bigotimes_{m=n+1}^\infty x_n \mapsto \bigotimes_{m=n+1}^\infty \langle \omega_{(m)}, x_n \rangle \omega_{(m)} \quad (36)$$

is the orthogonal projection onto the one-dimensional subspace of $\Gamma_{[n+1}$ spanned by the vector $\otimes_{m=n+1}^\infty \omega_{(m)}$.

Notation 4.1. For all $\tau \in \mathrm{T}$, let

$$D_\tau := I_\mathsf{h} \otimes J_\tau^* \Pi_\tau : \widetilde{\mathcal{F}} \to \widetilde{\Gamma};\ u\varepsilon(f) \mapsto u \otimes \bigotimes_{n=0}^\infty \widehat{f_\tau(n)} \quad (37)$$

and note that, as $|\tau| \to 0$,

$$D_\tau^* D_\tau = I_\mathsf{h} \otimes \Pi_\tau^* J_\tau J_\tau^* \Pi_\tau = I_\mathsf{h} \otimes Q_\tau \to I_{\widetilde{\mathcal{F}}} \quad (38)$$

in the strong operator topology and

$$D_\tau D_\tau^* = I_\mathsf{h} \otimes J_\tau^* \Pi_\tau \Pi_\tau^* J_\tau = I_{\widetilde{\Gamma}}, \quad (39)$$

since Π_τ is an isometric isomorphism and J_τ an isometry.

Remark 4.1. Let $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}})$ and $t \in \mathbb{R}_+$ be fixed. For all $\tau \in \mathrm{T}$, let $n = n(\tau) \in \mathbb{Z}_+$ be such that $t \in [\tau_n, \tau_{n+1}[$ and note that

$$\begin{aligned}\langle u\varepsilon(f), D_\tau^* \widetilde{s}_n(X) D_\tau v\varepsilon(g)\rangle \\ = \prod_{m=0}^{n-1} \big(1 + \langle f_\tau(m), g_\tau(m)\rangle\big) \langle u \otimes \widehat{f_\tau(n)}, X[v \otimes \widehat{g_\tau(n)}]\rangle \end{aligned} \tag{40}$$

for all u, $v \in \mathsf{h}$ and f, $g \in \mathsf{K}$. As $|\tau| \to 0$, $\tau_n \nearrow t$ and

$$\begin{aligned}\prod_{m=0}^{n-1} \big(1 + \langle f_\tau(m), g_\tau(m)\rangle\big) &= \langle J_\tau^* \Pi_\tau \varepsilon(f), J_\tau^* \Pi_\tau \varepsilon(1_{[0,\tau_n[} g)\rangle \\ &= \langle Q_\tau \varepsilon(f), \mathbb{E}_{\tau_n} \varepsilon(g)\rangle \to \langle \varepsilon(f), \mathbb{E}_t \varepsilon(g)\rangle. \end{aligned} \tag{41}$$

To analyse the second term in the right-hand side of (40), let $X = \left(\begin{smallmatrix} E & F \\ G & H \end{smallmatrix}\right)$, where $E \in \mathcal{B}(\mathsf{h})$, $F \in \mathcal{B}(\mathsf{h} \otimes \mathsf{k}; \mathsf{h})$, $G \in \mathcal{B}(\mathsf{h}; \mathsf{h} \otimes \mathsf{k})$ and $H \in \mathcal{B}(\mathsf{h} \otimes \mathsf{k})$. Then

$$\begin{aligned}\langle u \otimes \widehat{f_\tau(n)}, X[v \otimes \widehat{g_\tau(n)}]\rangle = \langle u, Ev\rangle + \langle u, F[v \otimes g_\tau(n)]\rangle + \langle u \otimes f_\tau(n), Gv\rangle \\ + \langle u \otimes f_\tau(n), H[v \otimes g_\tau(n)]\rangle; \end{aligned} \tag{42}$$

this equation shows the necessity of scaling the components of X in order to obtain non-trivial limits. Replacing X by $X_{\tau,n}$, where

$$\begin{pmatrix} E & F \\ G & H \end{pmatrix}_{\tau,n} := \begin{pmatrix} (\tau_{n+1} - \tau_n)E & (\tau_{n+1} - \tau_n)^{1/2} F \\ (\tau_{n+1} - \tau_n)^{1/2} G & H \end{pmatrix} \tag{43}$$

$$= \begin{pmatrix} (\tau_{n+1} - \tau_n)^{1/2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} \begin{pmatrix} (\tau_{n+1} - \tau_n)^{1/2} & 0 \\ 0 & 1 \end{pmatrix}, \tag{44}$$

the right-hand side of (42) becomes

$$\begin{aligned}\int_{\tau_n}^{\tau_{n+1}} \big(\langle u, Ev\rangle + \langle u, F[v \otimes g(t)]\rangle \\ + \langle u \otimes f(t), Gv\rangle + \langle u \otimes f(t), H[v \otimes P_\tau g(t)]\rangle\big)\, \mathrm{d}t \\ = \int_{\tau_n}^{\tau_{n+1}} \big(\langle u \otimes \widehat{f(t)}, X[v \otimes \widehat{g(t)}]\rangle + \langle u \otimes f(t), H[v \otimes (P_\tau g - g)(t)]\rangle\big)\, \mathrm{d}t. \end{aligned} \tag{45}$$

Theorem 4.1. *For all $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}})$ and $t \in \mathbb{R}_+$,*

$$\Sigma_\tau(X)_t := \sum_{n=0}^{\infty} \mathbb{1}_{\tau_{n+1} \in [0,t]} D_\tau^* \widetilde{s}_n(X_{\tau,n}) D_\tau \to \Lambda_\Omega(X)_t \quad \text{as } |\tau| \to 0 \tag{46}$$

strongly on $\widetilde{\mathcal{E}}$.

Proof. It follows from (45) that $\Sigma_\tau(X)_t$ can be written as the sum of a semi-vacuum-adapted QS integral and an Itô-integral remainder term. Let

$$\mathbb{E}_t^\tau := \sum_{n=0}^{\infty} \mathbb{1}_{t\in[\tau_n,\tau_{n+1}[}\mathbb{E}_{\tau_n} \qquad \forall\, t \in \mathbb{R}_+ \tag{47}$$

and note that $t \mapsto X \otimes Q_\tau \mathbb{E}_t^\tau$ is vacuum-adapted. If $t \in [\tau_m, \tau_{m+1}[$ then

$$\begin{aligned}
&\langle u\varepsilon(f), \Sigma_\tau(X)_t v\varepsilon(g)\rangle \\
&\quad= \sum_{n=0}^{m-1} \langle Q_\tau\varepsilon(f), \mathbb{E}_{\tau_n}\varepsilon(g)\rangle \langle u \otimes \widehat{f_\tau(n)}, X_{\tau,n}[v \otimes \widehat{g_\tau(n)}]\rangle \\
&\quad= \int_0^{\tau_m} \langle [u \otimes \widehat{f(s)}]\varepsilon(f), (X \otimes Q_\tau\mathbb{E}_s^\tau)\big([v \otimes \widehat{g(s)}]\varepsilon(g)\big)\rangle \,\mathrm{d}s \\
&\qquad+ \int_0^{\tau_m} \langle [u \otimes f(s)]\varepsilon(f), H[v \otimes (P_\tau g - g)(s)] \otimes Q_\tau\mathbb{E}_s^\tau\varepsilon(g)\rangle \,\mathrm{d}s. \qquad (48)
\end{aligned}$$

If $\mathcal{I}_s$ denotes the abstract Itô integral on $[0, s[$ then this shows that

$$\begin{aligned}
&(\Sigma_\tau(X)_t - \Lambda_\Omega(1_{[0,\tau_m[}X \otimes Q_\tau\mathbb{E}^\tau; I_\mathsf{K})_t)v\varepsilon(g) \\
&\qquad\qquad = \mathcal{I}_{\tau_m}\big(H[v \otimes (P_\tau g - g)(\cdot)] \otimes Q_\tau\mathbb{E}_\cdot^\tau\varepsilon(g)\big); \qquad (49)
\end{aligned}$$

as the Itô integral is an isometry, the norm of this quantity is bounded above by

$$\|H\|\,\|v\|\,\|P_\tau g - g\|_{L^2([0,t[;\mathsf{k})}\|\varepsilon(g)\| \to 0 \quad \text{as } |\tau| \to 0.$$

Finally, since $Q_\tau\mathbb{E}_s^\tau \to \mathbb{E}_s$ strongly as $|\tau| \to 0$ for all $s \in \mathbb{R}_+$, Theorem 3.1 and the dominated-convergence theorem imply that

$$\Lambda_\Omega(X)_t - \Lambda_\Omega(1_{[0,\tau_m[}X \otimes Q_\tau\mathbb{E}^\tau; I_\mathsf{K})_t = \Lambda_\Omega(X \otimes (\mathbb{E} - 1_{[0,\tau_m[}Q_\tau\mathbb{E}^\tau); I_\mathsf{K})_t \tag{50}$$

tends to 0 strongly on $\widetilde{\mathcal{E}}$ as $|\tau| \to 0$, as required. □

5. Multiple integrals

O, thou hast damnable iteration
– William Shakespeare, *Henry IV, Part 1*, Act I, Scene ii (1596).

Remark 5.1. For all $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes 2})$, $t \in \mathbb{R}_+$ and $\tau \in \mathrm{T}$, let

$$\Sigma_\tau^2(X)_t := \sum_{n=0}^{\infty}\sum_{m=0}^{n-1} \mathbb{1}_{\tau_{n+1}\in[0,t]} D_\tau^* \widetilde{s}_{m,n}(X_{\tau,m,n}) D_\tau, \tag{51}$$

where $\widetilde{s}_{m,n} : \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes 2}) \to \mathcal{B}(\widetilde{\Gamma})$ is the normal $*$-homomorphism such that

$$B \otimes C_1 \otimes C_2 \mapsto B \otimes I_{\Gamma_{m)}} \otimes C_1 \otimes P^\omega_{[m+1,n)} \otimes C_2 \otimes P^\omega_{[n+1}, \tag{52}$$

with $P^{\omega}_{[m+1,n)}$ and $P^{\omega}_{[n+1}$ the orthogonal projections onto $\bigotimes_{k=m+1}^{n-1}\mathbb{C}\omega_{(k)}$ and $\bigotimes_{k=n+1}^{\infty}\mathbb{C}\omega_{(k)}$, respectively.

To find the correct scaling for $X_{\tau,m,n}$, note that if

$$\Psi[\tau]_n := \begin{pmatrix} (\tau_{n+1}-\tau_n)^{1/2} & 0 \\ 0 & I_{\mathsf{k}} \end{pmatrix} \in \mathcal{B}(\widehat{\mathsf{k}}) \qquad \forall\, n \in \mathbb{Z}_+ \tag{53}$$

and $Y \in \mathcal{B}(\mathsf{h}\otimes\widehat{\mathsf{k}})$ then $Y_{\tau,n} = (I_{\mathsf{h}}\otimes\Psi[\tau]_n)Y(I_{\mathsf{h}}\otimes\Psi[\tau]_n)$, so let

$$X_{\tau,m,n} := (I_{\mathsf{h}}\otimes\Psi[\tau]_m\otimes\Psi[\tau]_n)X(I_{\mathsf{h}}\otimes\Psi[\tau]_m\otimes\Psi[\tau]_n) \tag{54}$$

for all $(m,n) \in \mathbb{Z}_+^{2,\uparrow}$. Having examined the case of multiplicity two, the general case is now clear.

Definition 5.1. For all $n \geq 1$, $X \in \mathcal{B}(\mathsf{h}\otimes\widehat{\mathsf{k}}^{\otimes n})$, $t \in \mathbb{R}_+$ and $\tau \in \mathrm{T}$, let

$$\Sigma^n_\tau(X)_t := \sum_{\mathbf{p}\in\mathbb{Z}_+^{n,\uparrow}} \mathbb{1}_{\tau_{p_n+1}\in[0,t]} D^*_\tau \widetilde{s}_{\mathbf{p}}(X_{\tau,\mathbf{p}})D_\tau, \tag{55}$$

where $\widetilde{s}_{\mathbf{p}} : \mathcal{B}(\mathsf{h}\otimes\widehat{\mathsf{k}}^{\otimes n}) \to \mathcal{B}(\widetilde{\Gamma})$ is the normal $*$-homomorphism such that

$$B\otimes C_1\otimes\cdots\otimes C_n \mapsto B\otimes I_{\Gamma_{p_1)}}\otimes C_1\otimes P^{\omega}_{[p_1+1,p_2)}\otimes\cdots\otimes C_n\otimes P^{\omega}_{[p_n+1}, \tag{56}$$

in which C_m acts on $\widehat{\mathsf{k}}_{(p_m)}$ for $m = 1,\ldots,n$ and $P^\omega : x \mapsto \langle x,\omega\rangle\omega$ acts on $\widehat{\mathsf{k}}_{(q)}$ for all $q \geq p_1$ such that $q \notin \{p_1,\ldots,p_n\}$, and

$$X_{\tau,\mathbf{p}} := (I_{\mathsf{h}}\otimes\Psi[\tau]_{p_1}\otimes\cdots\otimes\Psi[\tau]_{p_n})X(I_{\mathsf{h}}\otimes\Psi[\tau]_{p_1}\otimes\cdots\otimes\Psi[\tau]_{p_n}). \tag{57}$$

This is the discrete analogue of the vacuum-adapted n-fold quantum Wiener integral of X.

Theorem 5.1. *If $n \geq 1$, $X \in \mathcal{B}(\mathsf{h}\otimes\widehat{\mathsf{k}}^{\otimes n})$ and $t \in \mathbb{R}_+$ then*

$$\Sigma^n_\tau(X)_t = \Lambda^n_\Omega\big(X\otimes Q_\tau\mathbb{E}^\tau; P_\tau\big)^\tau_t \to \Lambda^n_\Omega\big(X\big)_t \tag{58}$$

strongly on $\widetilde{\mathcal{E}}$ as $|\tau| \to 0$.

Proof. If $\mathbf{p} \in \mathbb{Z}_+^{n,\uparrow}$ then, with the obvious extension of notation,

$$\begin{aligned}\langle u\varepsilon(f), D^*_\tau\widetilde{s}_{\mathbf{p}}(X_{\tau,\mathbf{p}})D_\tau v\varepsilon(g)\rangle = \\ \prod_{m=0}^{p_1-1}\langle\widehat{f_\tau(m)},\widehat{g_\tau(m)}\rangle\langle u\otimes\bigotimes_{k\in\mathbf{p}}\widehat{f_\tau(k)}, X_{\tau,\mathbf{p}}[v\otimes\bigotimes_{k\in\mathbf{p}}\widehat{g_\tau(k)}]\rangle\end{aligned} \tag{59}$$

for all u, $v \in \mathsf{h}$ and f, $g \in \mathsf{K}$. Furthermore, as

$$\Psi[\tau]_k\widehat{f_\tau(k)} = \frac{1}{\sqrt{\tau_{k+1}-\tau_k}}\int_{\tau_k}^{\tau_{k+1}}\widehat{f(t)}\,\mathrm{d}t, \tag{60}$$

it follows that

$$\Big\langle u \otimes \bigotimes_{k\in\mathbf{p}} \widehat{f_\tau(k)}, X_{\tau,\mathbf{p}}[v \otimes \bigotimes_{k\in\mathbf{p}} \widehat{g_\tau(k)}]\Big\rangle = \int_{[\tau_{\mathbf{p}},\tau_{\mathbf{p}+1}[} \big\langle u \otimes \widehat{f}^{\otimes n}(\mathbf{t}), X[v \otimes \widehat{P_\tau g}^{\otimes n}(\mathbf{t})]\big\rangle \,\mathrm{d}\mathbf{t}, \quad (61)$$

which gives the identity. That the limit is as claimed may be established by writing the difference $\Sigma^n_\tau(X) - \Lambda^n_\Omega(X)$ as

$$\Lambda^n_\Omega(X \otimes Q_\tau \mathbb{E}^\tau; P_\tau)^\tau - \Lambda^n_\Omega(X \otimes Q_\tau \mathbb{E}^\tau; P_\tau) + \Lambda^n_\Omega(X \otimes (Q_\tau \mathbb{E}^\tau - \mathbb{E}); P_\tau) + \Lambda^n_\Omega(X \otimes \mathbb{E}; P_\tau) - \Lambda^n_\Omega(X \otimes \mathbb{E}; I_{\mathsf{K}})$$

and employing Proposition 3.3, Theorem 3.2 and Proposition 3.2. □

6. Product formulae

Entia non sunt multiplicanda praeter necessitatem.
– William of Ockham.

Remark 6.1. Given $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{Z}_+^{m,\uparrow}$ and $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{Z}_+^{n,\uparrow}$ with $p_m < q_1$, let

$$\mathbf{p} \cup \mathbf{q} := (p_1, \ldots, p_m, q_1, \ldots, q_n) \in \mathbb{Z}_+^{m+n,\uparrow}. \quad (62)$$

If the normal $*$-homomorphism

$$\widetilde{\sigma}_{n,m} : \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n} \otimes \widehat{\mathsf{k}}^{\otimes m}) \to \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes m} \otimes \widehat{\mathsf{k}}^{\otimes n})$$

is that determined by the transposition $B \otimes C \otimes D \mapsto B \otimes D \otimes C$ then, letting

$$Y \rhd X := [\widetilde{\sigma}_{n,m}(Y \otimes I_{\widehat{\mathsf{k}}^{\otimes m}})](X \otimes (P^\omega)^{\otimes n}) \quad (63)$$

$$\text{and} \quad X \lhd Y := (X \otimes (P^\omega)^{\otimes n})[\widetilde{\sigma}_{n,m}(Y \otimes I_{\widehat{\mathsf{k}}^{\otimes m}})], \quad (64)$$

it is readily verified that

$$\widetilde{s}_{\mathbf{q}}(Y)\widetilde{s}_{\mathbf{p}}(X) = \widetilde{s}_{\mathbf{p}\cup\mathbf{q}}(Y \rhd X) \quad \text{and} \quad \widetilde{s}_{\mathbf{p}}(X)\widetilde{s}_{\mathbf{q}}(Y) = \widetilde{s}_{\mathbf{p}\cup\mathbf{q}}(X \lhd Y) \quad (65)$$

for all $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes m})$ and $Y \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$. Furthermore, for any $\tau \in \mathrm{T}$,

$$(Y \rhd X)_{\tau,\mathbf{p}\cup\mathbf{q}} = Y_{\tau,\mathbf{q}} \rhd X_{\tau,\mathbf{p}} \quad \text{and} \quad (X \lhd Y)_{\tau,\mathbf{p}\cup\mathbf{q}} = X_{\tau,\mathbf{p}} \lhd Y_{\tau,\mathbf{q}}, \quad (66)$$

since $\Psi[\tau]_p C \Psi[\tau]_p P^\omega = \Psi[\tau]_p C P^\omega \Psi[\tau]_p$ for all $p \in \mathbb{Z}_+$ and $C \in \mathcal{B}(\widehat{\mathsf{k}})$. The following Proposition is an immediate consequence of these observations.

Proposition 6.1 (Fubini). *If* $X \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes m})$ *and* $Y \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$ *then*

$$\Sigma_\tau^{m+n}(Y \rhd X)_t = \sum_{\mathbf{q} \in \mathbb{Z}_+^{n,\uparrow}} \mathbb{1}_{\tau_{q_n+1} \in [0,t]} D_\tau^* \widetilde{s}_{\mathbf{q}}(Y_{\tau,\mathbf{q}}) D_\tau \Sigma_\tau^m(X)_{\tau_{q_1}} \tag{67}$$

$$\text{and} \quad \Sigma_\tau^{m+n}(X \lhd Y)_t = \sum_{\mathbf{q} \in \mathbb{Z}_+^{n,\uparrow}} \mathbb{1}_{\tau_{q_n+1} \in [0,t]} \Sigma_\tau^m(X)_{\tau_{q_1}} D_\tau^* \widetilde{s}_{\mathbf{q}}(Y_{\tau,\mathbf{q}}) D_\tau \tag{68}$$

for all $\tau \in \mathrm{T}$ *and* $t \in \mathbb{R}_+$.

Theorem 6.1 (Quantum Itô product formula). *If* X, $Y \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}})$ *then*

$$\Lambda_\Omega^1(Y)\Lambda_\Omega^1(X) = \Lambda_\Omega^2(Y \rhd X) + \Lambda_\Omega^2(Y \lhd X) + \Lambda_\Omega^1(Y\Delta X), \tag{69}$$

where $\Delta \in \mathcal{B}(\mathsf{h} \otimes \widehat{\mathsf{k}})$ *denotes the orthogonal projection onto* $\mathsf{h} \otimes \mathsf{k}$.

Proof. Note first that if $\alpha_{\tau,n} := (\tau_{n+1} - \tau_n)^{1/2}$ for all $\tau \in \mathrm{T}$ and $n \in \mathbb{Z}_+$ then

$$(I_\mathsf{h} \otimes \Psi[\tau]_n)^2 = I_\mathsf{h} \otimes \Psi[\tau]_n^2 = \Delta + \alpha_{\tau,n}^2 \Delta^\perp, \tag{70}$$

whence

$$Y_{\tau,n} X_{\tau,n} = (X\Delta Y)_{\tau,n} + \alpha_{\tau,n}^2 (X\Delta^\perp Y)_{\tau,n}. \tag{71}$$

This working, the fact that $\widetilde{s}_m$ is a homomorphism for all $m \in \mathbb{Z}_+$ and the identities (67–68) with $m = n = 1$ imply that

$$\Sigma_\tau(Y)_t \Sigma_\tau(X)_t = \Sigma_\tau^2(Y \rhd X)_t + \Sigma_\tau^2\big(Y \lhd X\big)_t + \Sigma_\tau(Y\Delta X)_t + Z_t^\tau \tag{72}$$

for all $\tau \in \mathrm{T}$ and $t \in \mathbb{R}_+$, where

$$Z_t^\tau := \sum_{m=0}^\infty \mathbb{1}_{\tau_{m+1} \in [0,t]} \alpha_{\tau,m}^2 D_\tau^* \widetilde{s}_m\big((Y\Delta^\perp X)_{\tau,m}\big) D_\tau. \tag{73}$$

Working as in the proof of Theorem 4.1 (compare (49)) shows that

$$\begin{aligned} Z_t^\tau u\varepsilon(f) = \Lambda_\Omega(Y\Delta^\perp X \otimes W^\tau; I_\mathsf{K})_{\tau_n} u\varepsilon(f) \\ + \mathcal{I}_{\tau_n}\big(V[u \otimes (P_\tau f - f)(\cdot)] \otimes W_\cdot^\tau \varepsilon(f)\big) \end{aligned} \tag{74}$$

for all $t \in [\tau_n, \tau_{n+1}[$, $u \in \mathsf{h}$ and $f \in \mathsf{K}$, where $V := \Delta Y \Delta^\perp X \Delta$ and

$$W_t^\tau := Q_\tau \sum_{n=0}^\infty \mathbb{1}_{t \in [\tau_n, \tau_{n+1}[} \alpha_{\tau,n}^2 \mathbb{E}_{\tau_n} \qquad \forall\, t \in \mathbb{R}_+. \tag{75}$$

Now $W_t^\tau \to 0$ in norm as $|\tau| \to 0$, since $\|W_t^\tau\| \leq \sup_{n\geq 1} \alpha_{\tau,n}^2$ for all $t \in \mathbb{R}_+$, so $Z_{\tau,n} \to 0$ strongly on $\widetilde{\mathcal{E}}$, by Theorem 3.1 and Itô isometry. Combining this with Theorem 5.1, it follows that

$$\Sigma_\tau(Y)_t\Sigma_\tau(X)_t \to \Lambda_\Omega^2(Y \rhd X)_t + \Lambda_\Omega^2(Y \lhd X)_t + \Lambda_\Omega^1(Y\Delta X)_t \tag{76}$$

strongly on $\widetilde{\mathcal{E}}$ as $|\tau| \to 0$ and this gives the result. □

Remark 6.2. The quantum Itô formula (69) may be compared to that valid for the usual form of adaptedness [14, Exercise after Proposition 3.20].

7. Further development

Unbounded hopes were placed on each successive extension
– George Bernard Shaw, *Socialism: Principles and Outlook*, Shavian Tract No. 4, The Illusions of Socialism and Socialism: Principles and Outlook (1956).

This section contains little analysis, but sets out the basic situation once one moves beyond bounded integrands.

Definition 7.1. An *admissible triple* $(\mathsf{h}_0, \mathsf{k}_0, S)$ is a dense subspace $\mathsf{h}_0 \subseteq \mathsf{h}$, a dense subspace $\mathsf{k}_0 \subseteq \mathsf{k}$ and a subset $S \subseteq \mathsf{K}$ such that

(i) each $f \in S$ has compact support,
(ii) $f(t) \in \mathsf{k}_0$ for all $t \in \mathbb{R}_+$ and $f \in S$
and (iii) $\mathcal{E}_S := \mathrm{lin}\{\varepsilon(f) : f \in S\}$ is dense in $\mathcal{F}$.

Definition 7.2. If $X \in \mathcal{L}(\mathsf{h}_0 \odot \widehat{\mathsf{k}}_0; \mathsf{h} \otimes \widehat{\mathsf{k}})$, where h_0 is a subspace of h, k_0 is a subspace of k and $\widehat{\mathsf{k}}_0 := \mathbb{C} \oplus \mathsf{k}_0 \subseteq \widehat{\mathsf{k}}$, then

$$\widetilde{s}_n(X) := U_n^*(X \odot I_{\Gamma_n)} \odot P_{[n+1}^\omega)U_n \in \mathcal{L}\Big(\mathsf{h}_0 \odot \bigodot_{m=0}^{\infty} \widehat{\mathsf{k}}_0; \widetilde{\Gamma}\Big) \tag{77}$$

for all $n \in \mathbb{Z}_+$, where the unitary operator $U_n : \widetilde{\Gamma} \to \widetilde{\Gamma}$ is such that

$$u \otimes \bigotimes_{m=0}^{\infty} x_m \mapsto u \otimes x_n \otimes \bigotimes_{m=0}^{n-1} x_m \otimes \bigotimes_{m=n+1}^{\infty} x_m \tag{78}$$

and

$$\bigodot_{m=0}^{\infty} \widehat{\mathsf{k}}_0 := \mathrm{lin}\Big\{ \bigotimes_{m=0}^{\infty} x_m \Big| x_m \in \widehat{\mathsf{k}}_0 \ \forall\, m \geq 0,\ \exists\, l \in \mathbb{Z}_+ : x_l = x_{l+1} = \cdots = \omega \Big\}. \tag{79}$$

Proposition 7.1. *Let* $(\mathsf{h}_0, \mathsf{k}_0, S)$ *be admissible. If* $X \in \mathcal{L}(\mathsf{h}_0 \odot \widehat{\mathsf{k}}_0; \mathsf{h} \otimes \widehat{\mathsf{k}})$ *and* $t \in \mathbb{R}_+$ *are such that* $\int_0^t \|X[u \otimes \widehat{f(s)}]\|^2 \,\mathrm{d}s < \infty$ *for all* $u \in \mathsf{h}_0$ *and* $f \in S$ *then*

$$\Sigma_\tau(X)_t := \sum_{n=0}^{\infty} \mathbb{1}_{\tau_{n+1} \in [0,t]} D_\tau^* \widetilde{s}_n(X_{\tau,n}) D_\tau \to \Lambda_\Omega(X)_t \tag{80}$$

weakly on $\widetilde{\mathcal{E}}_S := \mathsf{h}_0 \odot \mathcal{E}_S$ *as* $|\tau| \to 0$*, where* $\Lambda_\Omega(X) := \Lambda_\Omega(X \odot \mathbb{E}; I_{\mathsf{K}})$.

Proof. Note that

$$\begin{aligned}\langle u\varepsilon(f), D_\tau^* \widetilde{s}_n(X) D_\tau v\varepsilon(g)\rangle \\ = \prod_{m=0}^{n-1} \langle \widehat{f_\tau(m)}, \widehat{g_\tau(m)}\rangle \int_{\tau_n}^{\tau_{n+1}} \langle u \otimes \widehat{P_\tau f(s)}, X[v \otimes \widehat{g(s)}]\rangle \,\mathrm{d}s\end{aligned} \tag{81}$$

for all $n \in \mathbb{Z}_+$, u, $v \in \mathsf{h}$ and f, $g \in S$, so if $t \in [\tau_n, \tau_{n+1}[$ then, as $|\tau| \to 0$,

$$\begin{aligned}\langle u\varepsilon(f), \Sigma_\tau(X)_t v\varepsilon(g)\rangle &= \int_0^{\tau_n} \langle u \otimes \widehat{P_\tau f(s)}, X[v \otimes \widehat{g(s)}]\rangle \langle \varepsilon(f), Q_\tau \mathbb{E}_s^\tau \varepsilon(g)\rangle \,\mathrm{d}s \\ &\to \int_0^t \langle u \otimes \widehat{f(s)}, X[v \otimes \widehat{g(s)}]\rangle \langle \varepsilon(f), \varepsilon(1_{[0,s[} g)\rangle \,\mathrm{d}s. \qquad \square\end{aligned}$$

Remark 7.1. Similarly, if $X \in \mathcal{L}(\mathsf{h}_0 \odot \widehat{\mathsf{k}}_0^{\odot n}; \mathsf{h} \otimes \widehat{\mathsf{k}}^{\otimes n})$ and $t \in \mathbb{R}_+$ are such that

$$\int_{\Delta_n(t)} \left\| X[u \otimes \widehat{f}^{\otimes n}(\mathbf{t})] \right\|^2 \mathrm{d}\mathbf{t} < \infty \qquad \forall\, u \in \mathsf{h}_0,\ f \in S \tag{82}$$

then $\Sigma_\tau^n(X)_t \to \Lambda_\Omega^n(X)_t$ weakly on $\widetilde{\mathcal{E}}_S$ as $|\tau| \to 0$; the proper definitions of $\Sigma_\tau^n(X)$ and $\Lambda_\Omega^n(X)$ should be clear from the above.

Acknowledgements

This is an expanded version of a talk given at the 27th Conference on Quantum Probability and its Applications, which was held at the University of Nottingham in July 2006; the hospitality of its organisers is acknowledged with pleasure. Thanks are extended to Professor Martin Lindsay, for a remark clarifying the proof of Theorem 3.1 and extensive for comments on a previous draft, which (hopefully) have led to a much improved presentation of this material; thanks are also extended to Dr Stephen Wills, for many helpful conversations on these ideas. The author is an Embark Postdoctoral Fellow at University College Cork, funded by the Irish Research Council for Science, Engineering and Technology.

Bibliography

1. L. Accardi & A. Bach, *Central limits of squeezing operators*, in: Quantum probability and applications IV (Rome, 1987), L. Accardi & W. von Waldenfels (eds.), Lecture Notes in Mathematics 1396, Springer, Berlin, 1989, 7–19.
2. S. Attal, *Approximating the Fock space with the toy Fock space*, Séminaire de Probabilités XXXVI, J. Azéma, M. Émery, M. Ledoux & M. Yor (eds.), Lecture Notes in Mathematics 1801, Springer, Berlin, 2003, 477–491.
3. S. Attal & Y. Pautrat, *From repeated to continuous quantum interactions*, Ann. Henri Poincaré 7 (2006), 59–104.
4. A.C.R. Belton, *Some self-adjoint quantum semimartingales*, Proc. London Math. Soc. (3) 92 (2006), 791–816.
5. A.C.R. Belton, *Random-walk approximation to vacuum cocycles*, `arXiv:math.OA/0702700`, version 2, 2007.
6. L. Bouten, R. van Handel & M.R. James, *A discrete invitation to quantum filtering and feedback control*, `arXiv:math.PR/0606118`, version 4, 2006.
7. T.A. Brun, *A simple model of quantum trajectories*, Amer. J. Phys. 70 (2002), 719–737.
8. U. Franz & A. Skalski, *Approximation of quantum Lévy processes by quantum random walks*, `arXiv:math.FA/0703339`, version 1, 2007.
9. J. Gough, *Holevo-ordering and the continuous-time limit for open Floquet dynamics*, Lett. Math. Phys. 67 (2004), 207–221.
10. J. Gough & A. Sobolev, *Stochastic Schrödinger equations as limit of discrete filtering*, Open Sys. Inf. Dyn. 11 (2004), 235–255.
11. R.V. Kadison & J.R. Ringrose, *Fundamentals of the theory of operator algebras. Volume II: Advanced theory*, Graduate Studies in Mathematics 16, American Mathematical Society, Providence, Rhode Island, 1997.
12. B. Kümmerer, *Quantum Markov processes and applications in physics*, in: Quantum independent increment processes II, M. Schürmann & U. Franz (eds.), Lecture Notes in Mathematics 1866, Springer, Berlin, 2006, 259–330.
13. M. Leitz-Martini, *Quantum stochastic calculus using infinitesimals*, doctoral thesis, Eberhard Karls Universität Tübingen, 2001. `http://w210.ub.uni-tuebingen.de/dbt/volltexte/2002/458/`
14. J.M. Lindsay, *Quantum stochastic analysis — an introduction*, in: Quantum independent increment processes I, M. Schürmann & U. Franz (eds.), Lecture Notes in Mathematics 1865, Springer, Berlin, 2005, 181–271.
15. J.M. Lindsay & K.R. Parthasarathy, *The passage from random walk to diffusion in quantum probability. II*, Sankhyā Ser. A 50 (1988), 151–170.
16. P.-A. Meyer, *Éléments de probabilités quantiques. I–V*, in: Séminaire de Probabilités XX, J. Azéma & M. Yor (eds.), Lecture Notes in Mathematics 1204, Springer, Berlin, 1986, 186–312.
17. P.-A. Meyer, *Éléments de probabilités quantiques. X. Approximation de l'oscillateur harmonique (d'après L. Accardi et A. Bach)*, in: Séminaire de Probabilités XXIII, J. Azéma, P.-A. Meyer & M. Yor (eds.), Lecture Notes in Mathematics 1372, Springer, Berlin, 1989, 175–182.
18. K.R. Parthasarathy, *The passage from random walk to diffusion in quantum probability*, J. Appl. Probab. 25A (1988), 151–166.

19. Y. Pautrat, *From Pauli matrices to quantum Itô formula*, Math. Phys. Anal. Geom. 8 (2005), 121–155.
20. L. Sahu, *Quantum random walks and their convergence*, `arXiv:math.OA/0505438`, version 1, 2005.
21. K.B. Sinha, *Quantum random walk revisited*, in: Quantum Probability, M. Bożejko, W. Młotkowski & J. Wysoczański (eds.), Banach Center Publications 73, Polish Academy of Sciences, Warsaw, 2006, 377–390.

REGULAR SOLUTIONS OF QUANTUM STOCHASTIC DIFFERENTIAL EQUATIONS

FRANCO FAGNOLA*

Department of Mathematics, Politecnico di Milano
20133 Milano, Italy
E-mail: franco.fagnola@polimi.it
www.mate.polimi.it/qp

We prove that the operators U_t solving a quantum stochastic differential equation $dU_t = F_\beta^\alpha U_t d\Lambda_\alpha^\beta(t)$ with F_β^α bounded operators in the initial space map vectors of the form $u \otimes \mathrm{e}(f)$ in the domain of any power of field operators $\Lambda_\nu^\mu(t)$. This proves that compositions of operators like $\Lambda_\nu^\mu(t)U_t$ make sense on the usual domain of Hudson-Parthasarathy quantum stochastic calculus. Applications to quantum continual measurements and quantum filtering are pointed out.

Keywords: Quantum stochastic differential equation; unbounded operators; regularity; output fields.

1. Introduction

Let U be a contractive solution of a right Hudson-Parthasarathy (H-P) equation

$$dU_t = \sum_{\alpha,\beta} F_\beta^\alpha U_t d\Lambda_\alpha^\beta(t), \qquad U_0 = \mathbb{1}. \tag{1}$$

Here F_β^α are bounded operators on the initial space h satisfying the appropriated conditions for U_t to be a contraction and Λ_α^β are the fundamental noises of Boson Fock quantum stochastic calculus. In the applications of quantum stochastic calculus to quantum measurement (see, for instance, Barchielli[1], Bouten, van Handel and James[4]) it is a quite natural question to ask whether we can define operator processes like $(U_t^* \Lambda_\alpha^\beta(t) U_t)_{t\geq 0}$ as operator processes on the tensor product of h with the Fock space with a domain including vectors of the form $u \otimes \mathrm{e}(f)$. Indeed, this is the usual fundamental domain where all the operator processes of Hudson-Parthasarathy quantum stochastic calculus must be defined.

Clearly, since the operators $\Lambda_\alpha^\beta(t)$ are unbounded, it is necessary to check that the contractions U_t solving (1) map vectors $u \otimes \mathrm{e}(f)$ in the domain of $\Lambda_\alpha^\beta(t)$. Following the terminology of classical PDE, this property will be called *regularity*. Indeed, a function solving a PDE is often called regular if it is differentiable, i.e. it belongs to the domain of some differential operator, or it satisfies some growth condition at infinity, i.e. it belongs to the domain of a multiplication operator by $|x|^k$, or both.

In this note we prove that solutions of (1) are regular and compositions of operators like $\Lambda_\nu^\mu(t)U_t$ make sense of the usual domain of Hudson-Parthasarathy quantum stochastic calculus. Our method is inspired by the a priori estimate technique introduced by Fagnola and Wills[6] to study H-P equations with unbounded F_β^α of vectors $U_t u \otimes \mathrm{e}(f)$ in the graph norm of a suitable operator.

2. Notation and preliminaries

We consider the usual framework of Hudson-Parthasarathy (see Hudson and Parthasarathy[7], Parthasarathy[9]) quantum stochastic calculus in Boson Fock spaces.

The *initial space* h is a complex separable Hilbert space, the *noise space* is the Boson Fock space $\mathcal{F}$ on $L^2(\mathbb{R}_+;\mathsf{k})$ where k is a complex separable Hilbert space. Let $(e_k)_{k\geq 1}$ be an orthonormal basis of k and let

$$\begin{aligned}\mathcal{M} = \big\{ & f \in L^2(\mathbb{R}_+;\mathsf{k}) \cap L^\infty_{\mathrm{loc}}(\mathbb{R}_+;\mathsf{k}) \mid \langle e_k, f(t)\rangle = 0 \\ & \qquad \text{identically in } t \text{ for all but a finite number of } k\text{'s} \big\} \\ \mathcal{E} = \lin\{ & \mathrm{e}(f) : f \in \mathcal{M}\}.\end{aligned}$$

where $\mathrm{e}(f)$ is the exponential vector associated to the test function f.

The quantum noises in $\mathcal{F}(L^2(\mathbb{R}_+;\mathsf{k}))$, $\{\Lambda_\beta^\alpha \mid \alpha,\beta \geq 0\}$ are defined by

$$\begin{array}{lcll} \Lambda_\beta^0(t) & = & A^+(1_{[0,t]} \otimes |e_\beta\rangle) = A_\beta^+(t) & \text{if } \beta > 0, \\ \Lambda_\beta^\alpha(t) & = & \Lambda(1_{[0,t]} \otimes |e_\beta\rangle\langle e_\alpha|) & \text{if } \alpha,\beta > 0, \\ \Lambda_0^\alpha(t) & = & A(1_{[0,t]} \otimes \langle e_\alpha|) = A^\alpha(t) & \text{if } \alpha > 0, \\ \Lambda_0^0(t) & = & t\,\mathbb{1}_{\mathcal{F}} & \end{array}$$

where A^+, A, Λ denote respectively the creation, annihilation and gauge

operators in $\mathcal{F}$ defined, for each $u \in \mathsf{h}$ and each $\mathrm{e}(f) \in \mathcal{F}$, by

$$\begin{aligned} A^+(1_{[0,t]} \otimes |e_m\rangle) u\mathrm{e}(f) &= \frac{d}{d\epsilon} u\mathrm{e}(f + \epsilon 1_{[0,t]} e_m)\Big|_{\epsilon=0} \\ \Lambda(1_{[0,t]} \otimes |e_m\rangle\langle e_\ell|) u\mathrm{e}(f) &= -i\frac{d}{d\epsilon} u\mathrm{e}(e^{i\epsilon 1_{[0,t]} \otimes |e_m\rangle\langle e_\ell|} f)\Big|_{\epsilon=0} \\ A(1_{[0,t]} \otimes \langle e_\ell|) u\mathrm{e}(f) &= \big\langle e_\ell 1_{[0,t]}, f\big\rangle\, u\mathrm{e}(f). \end{aligned}$$

Let $\mathcal{H}$ be the Hilbert space $\mathsf{h} \otimes \mathcal{F}$. Operators defined on h or $\mathcal{F}$ will be identified with their natural extensions to $\mathcal{H}$. Vectors $u \otimes \mathrm{e}(f)$ ($u \in \mathsf{h}$, $f \in \mathcal{M}$) will be denoted by $u\mathrm{e}(f)$ omitting the tensor product.

We suppose that the reader has some familiarity with quantum stochastic calculus and recall that the quantum Itô formula, as an algebraic rule for multiplying stochastic differentials, reads as

$$d\Lambda^\mu_\nu d\Lambda^\beta_\alpha = \hat{\delta}^\mu_\alpha d\Lambda^\beta_\nu \tag{2}$$

where $\hat{\delta}^\mu_\alpha = 1$ if $\mu = \alpha > 0$ and $\hat{\delta}^\mu_\alpha = 0$ otherwise. Summation over repeated will be understood with the agreement that Greek indices start from 0 and Roman indices start from 1.

Solutions $(U_t)_{t\geq 0}$ to (1) are adapted operator processes on $\mathcal{H}$ such that

$$U_t = \mathbb{1}_{\mathcal{H}} + \sum_{\alpha,\beta} \int_0^t F^\alpha_\beta U_s d\Lambda^\beta_\alpha(s)$$

where the integral is the usual Hudson-Parthasarathy quantum stochastic integral and $\mathbb{1}_{\mathcal{H}}$ denotes the identity operator on $\mathcal{H}$. In order to avoid indices and simplify computations with stochastic differential it is very useful to introduce a "matrix notation" inspired by a similar one introduced by Belavkin[2].

The matrix of bounded operators $[F^\alpha_\beta]_{\alpha,\beta\geq 0}$ defines an operator F on the Hilbert space $\mathsf{h} \otimes (\mathbb{C} \oplus \mathsf{k}) = \mathsf{h} \oplus (\mathsf{h} \otimes \mathsf{k})$. The operator on the same space determined by the matrix $\mathrm{diag}(0, \mathbb{1}_{\mathsf{h}}, \ldots, \mathbb{1}_{\mathsf{h}})$ of operators on h will be denoted by $\Delta(\mathbb{1}_{\mathsf{h}})$. The operator $\Delta(\mathbb{1}_h)$ is associated with the matrix $[\hat{\delta}^\alpha_\beta \mathbb{1}_{\mathsf{h}}]_{\alpha,\beta\geq 0}$. Quantum stochastic differentials $F^\alpha_\beta d\Lambda^\beta_\alpha(s)$ will be denoted $d\Lambda_F(s)$ and the quantum Itô formula (2), reads as

$$d\Lambda_G(s) d\Lambda_F(s) = d\Lambda_{G\Delta(\mathbb{1}_h)F}(s).$$

The condition necessary and sufficient condition for U_t to be a contraction is now

$$F + F^* + F^*\Delta(\mathbb{1}_{\mathsf{h}})F \leq 0. \tag{3}$$

Since all the operators F^{α}_{β} are bounded, this condition implies that the operator F is bounded (see Lindsay and Wills[8] Prop. 7.3).

The domain of an operator X will be denoted by $\mathrm{Dom}(X)$.

3. Regularity

We start by recalling some basic facts on domains of operators $\Lambda^{\beta}_{\alpha}(t)$ that have been introduced above as operators on the exponential domain $\mathcal{E}$.

Let $\mathcal{F}(K)$ be the boson Fock space over a complex separable Hilbert space K. Following Parthasarathy[9] Ch.II Sect.20, we recall the definition of number operator (or differential second quantization). Given a self-adjoint operator H on K, consider the one-parameter unitary group $(\mathrm{e}^{-isH})_{s\in\mathbb{R}}$ on K generated by $-iH$. The second quantization

$$\Gamma(\mathrm{e}^{-isH})\mathrm{e}(f) = \mathrm{e}(\mathrm{e}^{-isH}f)$$

yields a strongly continuous one-parameter unitary group $(\Gamma(\mathrm{e}^{-isH}))_{s\in\mathbb{R}}$ on $\mathcal{F}(K)$. The Stone generator $\Lambda(H)$ is self-adjoint by definition. Moreover, it can be shown (see Parthasarathy[9] Prop. 20.7) that the linear manifold $\mathcal{E}$ of exponential vectors is an essential domain for $\Lambda(H)$.

When $K = L^2(\mathbb{R}_+;\mathsf{k})$ and $H = 1_{[0,t]}\otimes|e_\ell\rangle\langle e_\ell|$ ($(e_\ell)_{\ell\geq 1}$ is an orthonormal basis of k) we find that $\Lambda(H)$ is the unique self-adjoint extension of the operators $\Lambda^{\ell}_{\ell}(t)$ defined in Sect. 2 on the domain $\mathcal{E}$. When $H = 1_{[0,t]}\otimes\mathbb{1}_{\mathsf{k}}$, we denote by N_t the generator of the unitary group $(\Gamma(\mathrm{e}^{-isH}))_{s\in\mathbb{R}}$ and call *number operator process* the family $(N_t)_{t\geq 0}$.

The operators $N_t, \Lambda^1_1(t), \ldots$ commute because the operators $1_{[0,t]}\otimes\mathbb{1}_{\mathsf{k}}, 1_{[0,t]}\otimes|e_1\rangle\langle e_1|, \ldots$ commute (Parthasarathy[9] Prop. 20.7).

The following facts are easily checked.

Lemma 3.1. *The operators N_t and $\Lambda^{\ell}_{\ell}(t)$ are positive. The domain $\mathrm{Dom}(\Lambda^{\ell}_{\ell}(t))$ is contained in the domain $\mathrm{Dom}(N_t)$ and*

$$N_t = \sum_{\ell\geq 1}\Lambda^{\ell}_{\ell}(t)$$

the series being strongly convergent on $\mathrm{Dom}(N_t)$ for all $t\geq 0$.

Proof. From the definition of N_t, for all $g, f\in\mathcal{M}$, we have

$$\begin{aligned}\langle \mathrm{e}(g), N_t\,\mathrm{e}(f)\rangle &= i\frac{d}{ds}\left\langle \mathrm{e}(g), \Gamma\left(\mathrm{e}^{-is1_{[0,t]}\otimes\mathbb{1}_{\mathsf{k}}}\right)\mathrm{e}(f)\right\rangle\Big|_{s=0}\\ &= i\frac{d}{ds}\mathrm{e}^{\langle g,\mathrm{e}^{-is1_{[0,t]}}f\rangle}\Big|_{s=0} = \langle g, 1_{[0,t]}f\rangle\mathrm{e}^{\langle g,f\rangle}.\end{aligned}$$

Therefore, for all $\xi = \sum_j z_j \mathrm{e}(f_j) \in \mathcal{E}$ ($z_j \in \mathbb{C}, f_j \in \mathcal{M}$), we have

$$\langle \xi, N_t \xi \rangle = \left\langle \sum_j z_j f_j \otimes \mathrm{e}(f_j), \left(1_{[0,t]} \otimes \mathbb{1}_{\mathsf{k}} \otimes \mathbb{1}_{\mathcal{F}}\right) \left(\sum_k z_k f_k \otimes \mathrm{e}(f_k) \right) \right\rangle \geq 0$$

because the operator $1_{[0,t]} \otimes \mathbb{1}_{\mathsf{k}} \otimes \mathbb{1}_{\mathcal{F}}$ on $L^2(\mathbb{R}_+; \mathsf{k}) \otimes \mathcal{F}$ is clearly positive. In the same way, replacing $1_{[0,t]} \otimes \mathbb{1}_{\mathsf{k}} \otimes \mathbb{1}_{\mathcal{F}}$ by $1_{[0,t]} \otimes |e_\ell\rangle\langle e_\ell| \otimes \mathbb{1}_{\mathcal{F}}$ we find

$$\langle \xi, \Lambda_\ell^\ell(t) \xi \rangle = \left\langle \xi', \left(1_{[0,t]} \otimes |e_\ell\rangle\langle e_\ell| \otimes \mathbb{1}_{\mathcal{F}}\right) \xi' \right\rangle \geq 0$$

where $\xi' = \sum_k z_k f_k \otimes \mathrm{e}(f_k)$. Therefore the identity holds on the exponential domain $\mathcal{E}$ because

$$\sum_{\ell \geq 1} 1_{[0,t]} \otimes |e_\ell\rangle\langle e_\ell| \otimes \mathbb{1}_{\mathcal{F}} = 1_{[0,t]} \otimes \mathbb{1}_{\mathsf{k}} \otimes \mathbb{1}_{\mathcal{F}}$$

in the strong operator topology on $L^2(\mathbb{R}_+; \mathsf{k}) \otimes \mathcal{F}$. Since $\mathcal{E}$ is an essential domain for the commuting self-adjoint operators N_t and $\Lambda_\ell^\ell(t)$, it also implies that $\Lambda_\ell^\ell(t) \leq N_t$ and then $(\Lambda_\ell^\ell(t))^2 \leq N_t^2$.

Let $\xi \in \mathrm{Dom}(N_t)$ and let $(\xi_n)_{n \geq 1}$ be a sequence in $\mathcal{E}$ such that

$$\lim_{n \to \infty} \xi_n = \xi, \qquad \lim_{n \to \infty} N_t \xi_n = N_t \xi. \tag{4}$$

The inequality $(\Lambda_\ell^\ell(t))^2 \leq N_t^2$ implies then

$$\left\| \Lambda_\ell^\ell(t)(\xi_n - \xi_m) \right\|^2 \leq \| N_t(\xi_n - \xi_m) \|^2 .$$

Therefore the sequence $(\Lambda_\ell^\ell(t)(\xi_n))_{n \geq 1}$ converges and $\xi \in \mathrm{Dom}(\Lambda_\ell^\ell(t))$ for all $\ell \geq 1$. Finally we can check the identity on $\xi \in \mathrm{Dom}(N_t)$ approximating ξ with $\xi_n \in \mathcal{E}$. □

We now study the domains of annihilation $A^\ell(t)$ and creation $A_\ell^+(t)$. Both these operators are closed because they are mutually adjoint on $\mathcal{E}$ and so they have a densely defined adjoint operator. We shall denote by the same symbol also their closures.

Lemma 3.2. *The domains* $\mathrm{Dom}(A^\ell(t))$ *and* $\mathrm{Dom}(A_\ell^+(t))$ *are contained in* $\mathrm{Dom}(N_t)$ *and, for all* $\xi \in \mathrm{Dom}(N_t)$*, we have*

$$\sum_{\ell \geq 1} \| A^\ell(t) \xi \|^2 = \| N_t^{1/2} \xi \|^2 .$$

Proof. We first check the above identity on vectors $\xi \in \mathcal{E}$ of the form $\sum_{k=1}^n z_k \mathrm{e}(f_k)$ with $n \geq 1$, $z_1, \ldots, z_n \in \mathbb{C}$, $f_1, \ldots, f_n \in \mathcal{M}$. In this case we

have

$$\begin{aligned}\sum_{\ell\geq 1}\|A^\ell(t)\xi\|^2 &= \sum_{\ell\geq 1}\sum_{j,k=1}^n \langle z_j f_j^\ell, 1_{[0,t]}\rangle\langle 1_{[0,t]}, z_k f_k^\ell\rangle\langle \mathrm{e}(f_j), \mathrm{e}(f_k)\rangle\\ &= \sum_{j,k=1}^n \langle \mathrm{e}(f_j), \mathrm{e}(f_k)\rangle \sum_{\ell\geq 1}\langle z_j f_j, 1_{[0,t]} e_\ell\rangle\langle 1_{[0,t]} e_\ell, z_k f_k\rangle\\ &= \sum_{j,k=1}^n \langle \mathrm{e}(f_j), \mathrm{e}(f_k)\rangle \int_0^t \langle z_j f_j(s), z_k f_k(s)\rangle_{\mathsf{k}} ds.\end{aligned}$$

Moreover we have also

$$\begin{aligned}\|N_t^{1/2}\xi\|^2 &= \sum_{j,k=1}^n \bar{z_j} z_k \sum_{\ell\geq 1}\langle \mathrm{e}(f_j), \Lambda_\ell^\ell(t)\mathrm{e}(f_k)\rangle\\ &= \sum_{j,k=1}^n \bar{z_j} z_k \sum_{\ell\geq 1}\int_0^t \bar{f_j^\ell}(s) f_k^\ell(s) ds\, \langle \mathrm{e}(f_j), \mathrm{e}(f_k)\rangle\\ &= \sum_{j,k=1}^n \langle \mathrm{e}(f_j), \mathrm{e}(f_k)\rangle \int_0^t \langle z_j f_j(s), z_k f_k(s)\rangle_{\mathsf{k}} ds\end{aligned}$$

and the claimed identity holds on the domain $\mathcal{E}$.

Now, for all $\xi \in \mathrm{Dom}(N_t)$, letting $(\xi_n)_{n\geq 1}$ be a sequence satisfying (4), we find that the sequence $(A^\ell(t)\xi_n)_{n\geq 1}$ is convergent. This shows that $\xi \in \mathrm{Dom}(A^\ell(t))$. Moreover, the identity holds on $\mathrm{Dom}(N_t)$. The canonical commutation relation $A^\ell(t)A_\ell^+(t) = A_\ell^+(t)A^\ell(t) + t\mathbb{1}$ implies then

$$\left\|A_\ell^+(t)\xi\right\|^2 = \left\|A^\ell(t)\xi\right\|^2 + t\|\xi\|^2 \leq \|N_t\xi\|^2 + t\|\xi\|^2.$$

Therefore we can prove that $\mathrm{Dom}(A_\ell^+(t))$ is contained in $\mathrm{Dom}(N_t))$ by the same sequence argument. □

The following formula, giving the stochastic differential of processes $(N_t^n)_{t\geq 0}$, is a special case of the quantum Itô formula in Belavkin's pseudo-Poisson form (see Belavkin[3]).

Lemma 3.3. *For all $n \geq 1$ we have*

$$dN_t^n = \sum_{\ell\geq 1}\left((N_t+1)^n - N_t^n\right) d\Lambda_\ell^\ell(t).$$

Proof. The above formula is clearly true for $n = 1$. Suppose it has been established for an integer n, then, by the quantum Itô formula, for all

$f, g \in \mathcal{M}$, we can write $\langle \mathrm{e}(g), N_t^{n+1}\mathrm{e}(f)\rangle = \langle N_t\mathrm{e}(g), N_t^n\mathrm{e}(f)\rangle$ in the form

$$\begin{aligned}
&\int_0^t \Big\{ \langle \mathrm{e}(g), N_s^n\mathrm{e}(f)\rangle + \langle N_s\mathrm{e}(g), ((N_s+1)^n - N_s^n)\mathrm{e}(f)\rangle \\
&\quad + \langle \mathrm{e}(g), ((N_s+1)^n - N_s^n)\mathrm{e}(f)\rangle \Big\} g_\ell(s) f^\ell(s) ds \\
= &\int_0^t \Big\{ \langle \mathrm{e}(g), (N_s^n + (N_s+1)((N_s+1)^n - N_s^n))\,\mathrm{e}(f)\rangle \Big\} g_\ell(s) f^\ell(s) ds.
\end{aligned}$$

This completes the proof. □

Fix a positive integer n. For all $\varepsilon > 0$ consider the process $(Z_t^\varepsilon)_{t\geq 0}$ given by

$$Z_t^\varepsilon = (N_t+1)^n \mathrm{e}^{-\varepsilon(N_t+1)}. \tag{5}$$

The operators Z_t^ε are bounded. Indeed, it can be easily shown that $\|Z_t^\varepsilon\| \leq n^n(\varepsilon e)^{-n}$.

Proposition 3.1. *The adapted process $(Z_t^\varepsilon)_{t\geq 0}$ satisfies*

$$dZ_t^\varepsilon = \sum_{\ell\geq 1} Y_t^\varepsilon Z_t^\varepsilon d\Lambda_\ell^\ell(t), \quad Z_0^\varepsilon = e^{-\varepsilon}\mathbb{1}.$$

where $(Y_t^\varepsilon)_{t\geq 0}$ is the bounded operator process

$$Y_t^\varepsilon = \big((N_t+2)(N_t+1)^{-1}\big)^n e^{-\varepsilon} - \mathbb{1}$$

satisfying

$$-\mathbb{1} \leq Y_t^\varepsilon \leq (2^n-1)\,\mathbb{1}.$$

Proof. The function

$$\varphi : \mathbb{R} \to \mathbb{R}, \qquad \varphi(x) = (x+1)^n \mathrm{e}^{-\varepsilon(x+1)}$$

is analytic. Therefore it suffices to apply Lemma 3.3. □

Proposition 3.2. *Suppose that the operators F_β^α are bounded and let U be a contractive solution of the right QSDE (1). For all $\varepsilon > 0$, $u \in \mathsf{h}$, $f \in \mathcal{M}$ and $t \geq 0$ we have*

$$\|Z_t^\varepsilon U_t u e(f)\|^2 \leq \exp\big((4^n-1)(1+2\|F\|)\|f\|^2\big)\,\|u\|^2 \tag{6}$$

Proof. Computing the stochastic differential of $Z_t^\varepsilon U_t$ by the Itô formula we have

$$\begin{aligned} dZ_t^\varepsilon U_t &= F_\beta^\alpha Z_t^\varepsilon U_t d\Lambda_\alpha^\beta + Y_t^\varepsilon Z_t^\varepsilon U_t d\Lambda_\ell^\ell + F_\beta^\alpha Y^\varepsilon Z_t^\varepsilon U_t d\Lambda_\ell^\ell d\Lambda_\alpha^\beta \\ &= F_\beta^\alpha Z_t^\varepsilon U_t d\Lambda_\alpha^\beta + \hat\delta_\beta^\alpha Y_t^\varepsilon Z_t^\varepsilon U_t d\Lambda_\alpha^\beta + \hat\delta_\alpha^\ell F_\beta^\alpha Y^\varepsilon Z_t^\varepsilon U_t d\Lambda_\ell^\beta \\ &= \left(F_\beta^\alpha + \left(\hat\delta_\beta^\alpha + \hat\delta_\alpha^\ell F_\beta^\ell\right) Y_t^\varepsilon\right) Z_t^\varepsilon U_t d\Lambda_\alpha^\beta. \end{aligned}$$

Another application of the quantum Itô formula yields

$$\begin{aligned} \langle Z_t^\varepsilon U_t u\mathrm{e}(f), Z_t^\varepsilon U_t u\mathrm{e}(f)\rangle &= \mathrm{e}^{-2\varepsilon\|f\|^2}\|u\|^2 \\ &+ \int_0^t \left\langle Z_s^\varepsilon U_s u\mathrm{e}(f), M_\beta^\alpha(s) Z_s^\varepsilon U_s u\mathrm{e}(f)\right\rangle f_\alpha(s) f^\beta(s) ds \end{aligned}$$

where M_β^α are the operator processes on h given by

$$M_\beta^\alpha(s) = F_\beta^\alpha + F_\alpha^{\beta *} + F_\alpha^{\ell *} F_\beta^\ell + \left(\hat\delta_\alpha^\ell + F_\alpha^\ell\right)^* \left(\hat\delta_\beta^\ell + F_\beta^\ell\right) \left((Y_s^\varepsilon)^2 + 2Y_s^\varepsilon\right).$$

The operators F_β^α determine an operator $F = [F_\beta^\alpha]$ on $\mathsf{h} \otimes (\mathbb{C} \oplus \mathsf{k})$. In the same way, for each $s \geq 0$, the $M_\beta^\alpha(s)$ determine an operator $M(s)$ on $\mathsf{h} \otimes (\mathbb{C} \oplus \mathsf{k}) \otimes \mathcal{F}$. Notice that, by Proposition 3.1 we have

$$-\mathbb{1} \leq \left((Y_s^\varepsilon)^2 + 2Y_s^\varepsilon\right) \leq (4^n - 1)\,\mathbb{1}.$$

Therefore we obtain the inequality

$$\begin{aligned} M(s) &\leq (F + F^* + F^*\Delta(\mathbb{1}_\mathsf{h})F) \otimes \mathbb{1}_\mathcal{F} \\ &+ (4^n - 1)\,(\Delta(\mathbb{1}_\mathsf{h}) + F)^*\Delta(\mathbb{1}_\mathsf{h})(\Delta(\mathbb{1}_\mathsf{h}) + F) \otimes \mathbb{1}_\mathcal{F} \end{aligned}$$

where $\Delta(\mathbb{1}_\mathsf{h})$ denotes the projection $\mathrm{diag}(0, \mathbb{1}_\mathsf{h}, \dots)$ on $\mathsf{h} \otimes (\mathbb{C} \oplus \mathsf{k})$. The contractivity condition (3) for U implies that the first term in the right-hand side is a negative operator. Moreover, it impels that the operator F is bounded (see Lindsay and Wills[8]) and, letting $\Delta(\mathbb{1}_\mathsf{h})^\perp = \mathbb{1}_{h\otimes(\mathbb{C}\oplus\mathsf{k})} - \Delta(\mathbb{1}_\mathsf{h})$, the orthogonal of the projection $\Delta(\mathbb{1}_\mathsf{h})$ on $h \otimes (\mathbb{C} \oplus \mathsf{k})$,

$$\begin{aligned} (\Delta(\mathbb{1}_\mathsf{h}) + F)^*\Delta(\mathbb{1}_\mathsf{h})(\Delta(\mathbb{1}_\mathsf{h}) + F) &= \Delta(\mathbb{1}_\mathsf{h}) + F^*\Delta(\mathbb{1}_\mathsf{h}) + \Delta(\mathbb{1}_\mathsf{h})F + F^*\Delta(\mathbb{1}_\mathsf{h})F \\ &\leq \Delta(\mathbb{1}_\mathsf{h}) - F^*\Delta(\mathbb{1}_\mathsf{h})^\perp - \Delta(\mathbb{1}_\mathsf{h})^\perp F \\ &\leq (1 + 2\|F\|)\,\mathbb{1}_{\mathsf{h}\otimes\mathsf{k}}. \end{aligned}$$

Therefore we have

$$M(s) \leq (4^n - 1)\,(1 + 2\|F\|)\mathbb{1}_\mathcal{H}.$$

Then our computation with the quantum Itô formula leads finally to the inequality

$$\|Z_t^\varepsilon U_t u\mathrm{e}(f)\|^2 \leq \mathrm{e}^{-2\varepsilon\|f\|^2}\|u\|^2 + (4^n - 1)(1 + 2\|F\|) \int_0^t \|Z_s^\varepsilon U_s u\mathrm{e}(f)\|^2\, |f(s)|^2 ds.$$

The conclusion follows then from the Gronwall's Lemma. □

The following Lemma is an immediate consequence of the spectral theorem

Lemma 3.4. *Let X be a positive self-adjoint operator on a Hilbert space K and let $X^{(\varepsilon)}$ be the bounded operator $(X+\mathbb{1})e^{-\varepsilon(X+\mathbb{1})}$ $(\varepsilon>0)$. A vector $u\in K$ belongs to the domain of X if and only if*

$$\sup_{\varepsilon>0}\left\|X^{(\varepsilon)}u\right\|^2<\infty.$$

We can now prove that the operators U_t solving the right equation (1) map all vectors $u\mathrm{e}(f)$ with $u\in\mathsf{h}$, $f\in\mathcal{M}$ in the domain of the operators $\Lambda_\alpha^\beta(t)$.

Theorem 3.1. *Let U be a contractive solution of the right equation (1). Suppose that the operators F_β^α are bounded. Then, for all $n\geq 1$, $\ell,m\geq 1$, $t,s\geq 0$, $u\in\mathsf{h}$ and $f\in\mathcal{M}$, the vector $U_tu\mathrm{e}(f)$ belongs to the domain of the operators N_t^n, $(\Lambda_\alpha^\beta(t))^n$ and*

$$A^\ell(t)U_t=\int_0^t\left(\hat\delta_\beta^\ell+F_\beta^\ell\right)U_sd\Lambda_0^\beta(s)+\int_0^tF_\beta^\alpha A^\ell(s)U_sd\Lambda_\alpha^\beta(s)\tag{7}$$

$$A_\ell^+(t)U_t=\int_0^tU_sdA_\ell^+(s)+\int_0^tF_\beta^\alpha A_\ell^+(s)U_sd\Lambda_\alpha^\beta(s)\tag{8}$$

$$\Lambda_m^\ell(t)U_t=\int_0^t\left(\hat\delta_\beta^\ell+F_\beta^\ell\right)U_sd\Lambda_m^\beta(s)+\int_0^tF_\beta^\alpha\Lambda_m^\ell(s)U_sd\Lambda_\alpha^\beta(s)\tag{9}$$

Proof. The right-hand side of the *a priori* estimate (6) is independent of ε. Therefore, $U_tu\mathrm{e}(f)$ belongs to the domain of the operators N_t^n by Lemma 3.4. Moreover, it belongs also to $\mathrm{Dom}(\Lambda_\alpha^\beta(t)^n)$ by Lemma 3.1 and Lemma 3.2.

Let $u,v\in\mathsf{h}$ and $f,g\in\mathcal{M}$. By the Itô formula we have

$$\begin{aligned}\left\langle A_\ell^\dagger(t)v\mathrm{e}(g),U_tu\mathrm{e}(f)\right\rangle&=\int_0^t\left\langle A_\ell^\dagger(s)v\mathrm{e}(g),F_\beta^\alpha U_su\mathrm{e}(f)\right\rangle g_\alpha(s)f^\beta(s)ds\\&+\int_0^t\left\langle v\mathrm{e}(g),U_su\mathrm{e}(f)\right\rangle f^\ell(s)ds\\&+\int_0^t\left\langle v\mathrm{e}(g),F_\beta^\ell U_su\mathrm{e}(f)\right\rangle f^\beta(s)ds.\end{aligned}$$

Now, by Proposition 3.2 and Lemma 3.4, $U_su\mathrm{e}(f)$ belongs to the domain of N_t. Therefore, by Lemma 3.2, $U_su\mathrm{e}(f)$ belongs to the domain of $A_\ell(s)$.

Therefore, since $(F^\alpha_\beta)^*$ commutes with $A^\dagger_\ell(s)$, we find

$$\begin{aligned}\left\langle A^\dagger_\ell(s)v\mathrm{e}(g), F^\alpha_\beta U_s u\mathrm{e}(f)\right\rangle &= \left\langle A^\dagger_\ell(s)(F^\alpha_\beta)^* v\mathrm{e}(g), U_s u\mathrm{e}(f)\right\rangle\\ &= \left\langle (F^\alpha_\beta)^* v\mathrm{e}(g), A_\ell(s)U_s u\mathrm{e}(f)\right\rangle\\ &= \left\langle v\mathrm{e}(g), F^\alpha_\beta A_\ell(s)U_s u\mathrm{e}(f)\right\rangle.\end{aligned}$$

Therefore we have

$$\begin{aligned}\langle v\mathrm{e}(g), A_\ell(t)U_t u\mathrm{e}(f)\rangle &= \int_0^t \left\langle v\mathrm{e}(g), F^\alpha_\beta A_\ell(s)U_s u\mathrm{e}(f)\right\rangle g_\alpha(s)f^\beta(s)ds\\ &\quad + \int_0^t \left\langle v\mathrm{e}(g), (\hat\delta^\ell_\beta + F^\ell_\beta)U_s u\mathrm{e}(f)\right\rangle f^\beta(s)ds.\end{aligned}$$

This proves (7). The proof of (8), (9) is similar; we omit it. □

Theorem 3.1 shows that, for any bounded operator $X = [X^\alpha_\beta]_{\alpha,\beta\geq 0}$ on $\mathsf{h}\otimes(\mathbb{C}\oplus\mathsf{k})$, the operator process $(U_t^*\Lambda_X(t)U_t)_{t\geq 0}$ is well defined on the algebraic tensor product of the initial space h with the linear manifold $\mathcal{E}$ of exponential vectors (with test function in $\mathcal{M}$. Computing the stochastic differential in the matrix notation we find immediately

$$dU_t^*\Lambda_X(t)U_t = U_t^* d\Lambda_{\left(1+\Delta(\mathbb{1}_\mathsf{h})F\right)^* X(1+F\Delta(\mathbb{1}_\mathsf{h}))-X\big)}(t)U_t.$$

This allows us to write the QSDE satisfied by the operator process $(U_t^*\Lambda_X(t)U_t)_{t\geq 0}$ by a simple matrix computation.

Acknowledgments

We would like to thank A. Barchielli for pointing out this problem and discussions on quantum measurements and filtering.

Bibliography

1. A. Barchielli, *Continual Measurements in Quantum Mechanics and Quantum Stochastic Calculus*. In: S. Attal, A. Joye, C.-A. Pillet (eds.) *Open Quantum Systems III*. LNM 1882 pp. 207–291. Springer-Verlag, Berlin Heidelberg 2006.
2. V. P. Belavkin, A new form and *-algebraic structure of quantum stochastic integrals in Fock space. *Rend. Sem. Mat. Fis. Milano* **LVIII** (1988), 117–193.
3. V. P. Belavkin, The quantum functional Itô formula has the pseudo-Poisson structure $df(X) = [f(X+B) - f(X)]^\alpha_\beta d\Lambda^\beta_\alpha$. *Quantum probability & related topics*, 81–85, QP-PQ, VIII, World Sci. Publ., River Edge, NJ, 1993.
4. L. Bouten, R. van Handel and M. James, An Introduction to Quantum Filtering. `arXiv:math.OC/0601741 v1 30 Jan 2006`.

5. F. Fagnola, Quantum Stochastic Differential Equations and Dilations of Completely Positive Semigroups. In: S. Attal, A. Joye, C.-A. Pillet (eds.) *Open Quantum Systems II*. LNM 1881 pp. 183–220. Springer-Verlag, Berlin Heidelberg 2006.
6. F. Fagnola and S. J. Wills, Solving quantum stochastic differential equations with unbounded coefficients. *J. Funct. Anal.* **198** (2003), 279–310.
7. R. L. Hudson and K. R. Parthasarathy, Quantum Itô's formula and stochastic evolutions, *Comm. Math. Phys.* **93** (1984), 301–323 .
8. J. M. Lindsay and S. J. Wills, Existence, positivity and contractivity for quantum stochastic flows with infinite dimensional noise. *Probab. Theory Related Fields* **116** (2000), no. 4, 505–543.
9. K. R. Parthasarathy, *An introduction to quantum stochastic calculus*, Monographs in Mathematics 85, (Birkhäuser Verlag, Basel, 1992).

FROM ALGEBRAIC TO ANALYTIC DOUBLE PRODUCT INTEGRALS

ROBIN L. HUDSON

School of Mathematics, University of Loughborough, Loughborough, Leicestershire LE11 3TU Great Britain

The algebraic theory of double product integrals and particularly its role in the quantisation of Lie bialgebras is described. When the underlying associative algebra is that of the Itô differentials of quantum stochastic calculus such product integrals are formally represented as operators which are infinite sums of iterated integrals in Fock space. In this paper we describe some of the analytic problems encountered in making such sums rigourously meaningful, as well as the expected properties of such analytic double product integrals.

Keywords: double product integral, quantum stochastic integral, iterated integral, quantum Yang-Baxter equation

1. Introduction

An algebraic theory of double product integrals consisting of formal power series and thereby avoiding convergence questions has been developed. The original motivation for this theory was to construct a solution of the quantum Yang-Baxter equation from the datum of a generator in the form of a solution of the classical Yang-Baxter equation , thereby effecting a general quantisation procedure for quasitriangular Lie bialgebras in the sense of[3,12]. It emerged[10] that in general the relevant double product integral was not generated directly by the corresponding cYBe solution but by a formal power series with vanishing zero-order coefficient of which the given solution was the first-order coefficient and whose higher-order coefficients are determined by a hierarchy of inhomogeneous linear equations previously considered in a related but distinct context in[2].

More recently the analytic double product integral generated by $dA^\dagger \otimes dA - dA \otimes dA^\dagger$,where $A^\dagger$ and A are the creation and annihilation processes of quantum stochastic calculus[13,14], was constructed in the form of a family of operators in Fock space satisfying quantum stochastic differential equa-

tions[7], as second quantisations of an explicitly evaluated double continuous product of infinitesimal rotations[8]. The existence of a general theory of quantum stochastic double product integrals in Fock space is impeded by the fact that one of the simplest such integrals, that generated by $zd\Lambda\otimes d\Lambda$ where Λ is the conservation process of quantum stochastic calculus[14], is divergent in the sense that its matrix elements between exponential vectors have zero radius of convergence as power series in z. Another disincentive to constuction of such a general theory is awareness that the algebraic double products of most interest, those that furnish solutions of the qYBe, have generators which are themselves formal power series as noted above and thus give rise, when converted to Fock space objects using quantum stochastic calculus by replacing the formal parameter by a complex variable, to analytic quantum stochastic differential equations of a novel type for which the usual arguments for convergence of the series obtained by the Picard iterative method of solution do not apply without modification.

In this paper we consider some aspects of the transition from the algebraic, or "formal" in the language of some analysts, to an analytic theory of double product integrals in Fock space In Section 2 we recall so-called Itô Hopf algebra over an underlying associative algebra which reduces to the shuffle product Hopf algebra[2] when the underlying algebra is trivial, all products vanishing. In Section 3 we review the algebraic theory of double product integrals and their characterisation by their behaviour under the coproduct and in Section 4 their use to construct solutions of the quantum Yang-Baxter equation which are double product integrals and to use them to quantise Lie bialgebras. In Section 5 we show how, when the underlying algebra is that of Itô differentials in quantum stochastic calculus, the Itô Hopf algebra possesses a family of representations in Fock space in which homogeneous product tensors are mapped to iterated stochastic integrals, and through which the coproduct is related to the continuous tensor product splittings of Fock space (Theorem 5). Finally in Section 6 we consider some analytic problems arising when double product integrals are transferred to processes in Fock space by this mechanism.

2. The Itô-Hopf algebra.

Let there be given a not necessarily unital associative algebra $\mathcal{L}$ over the complex field $\mathbb{C}$. The vector space

$$\mathcal{T}(\mathcal{L}) = \mathbb{C}\oplus\mathcal{L}\oplus(\mathcal{L}\otimes\mathcal{L})\oplus(\mathcal{L}\otimes\mathcal{L}\otimes\mathcal{L})\oplus\cdots\oplus(\otimes^n\mathcal{L})\oplus\cdots$$

of all tensors over $\mathcal{L}$ becomes a Hopf algebra when equipped with the following structure.

- The associative product is defined by $\alpha\beta = \gamma = (\gamma_0, \gamma_1, \gamma_2, \ldots)$, where the nth rank component γ_n of γ is
$$\gamma_n = \sum_{A\cup B=\{1,2,\ldots,n\}} \alpha^A_{|A|}\beta^B_{|B|},$$
where the sum is over the 3^n decompositions of $\{1, 2, \ldots, n\}$ into the union of two subsets, we use the place notation so that $\alpha^A_{|A|}$ indicates that the $|A|$th rank component of α is to be placed in those copies of $\mathcal{L}$ within $\bigotimes^n \mathcal{L}$ which are labelled by elements of A, $\beta^B_{|B|}$ is defined analogously and if $A \cap B \neq \emptyset$ double occupancies of copies of $\mathcal{L}$ are reduced by multiplication in $\mathcal{L}$.
- The unit element is $(1, 0, 0, \ldots)$.
- The coproduct Δ is defined by its action on homogeneous product tensors
$$\Delta(L_1\otimes L_2\otimes\cdots L_n) = \sum_{j=0}^{n}(L_1\otimes L_2\otimes\cdots L_j)\bigotimes(L_{j+1}\otimes L_{j+2}\otimes\cdots L_n). \tag{1}$$
- The counit ε is defined by
$$\varepsilon(\alpha_0, \alpha_1, \alpha_2, \ldots) = \alpha_0.$$
- The antipode S acts on homogeneous product tensors embedded in $\mathcal{T}(\mathcal{L})$ as

$S(L) = -L$;
$S(L_1 \otimes L_2 \otimes \cdots L_n) = (-1)^n(L_n \otimes L_{n-1} \otimes \cdots L_1) + \text{lower rank terms.}$

Note that when $\mathcal{L}$ is trivial, all products vanishing, the sum defining the product reduces to one over pairs of disjoint subsets whose union $\{1, 2, \ldots, n\}$, and the Hopf algebra is then the shuffle product Hopf algebra of[2]. The proof that it is indeed a Hopf algebra for general $\mathcal{L}$, in particular that the coproduct remains multiplicative for the modified product, is in[9].

The kernel $\mathcal{K}(\varepsilon)$ of the counit consists of all tensors with vanishing zero-rank component. The map
$$\varepsilon_1 : \mathcal{K}(\varepsilon) \ni (0, \alpha_1, \alpha_2, \ldots) \mapsto \alpha_1 \in \mathcal{L}$$
is a homomorphism of associative algebras from $\mathcal{K}(\varepsilon)$ to $\mathcal{L}$.

Let $\mathcal{U}(\mathcal{L})$ denote the universal enveloping algebra of the Lie algebra got by equipping $\mathcal{L}$ with the commutator product. Then the map

$$\phi : \mathcal{L} \ni L \mapsto (0, L, 0, 0, \ldots) \in \mathcal{T}(\mathcal{L}) \tag{2}$$

is a Lie algebra homomorphism from the Lie algebra $\mathcal{L}$ to $\mathcal{T}(\mathcal{L})$ when the latter is also equipped with the commutator Lie bracket. Denote its universal extension by Φ.

Theorem 2.1. Φ *is an isomorphism of Hopf algebras from* $\mathcal{U}(\mathcal{L})$ *onto the sub-Hopf algebra of* $\mathcal{T}(\mathcal{L})$ *consisting of symmetric tensors.*

Proof. That Φ is an isomorphism of unital associative algebras is proved in[11] in the context that $\mathcal{L}$ is the algebra of Itô differentials of quantum stochastic calculus, and the proof does not depend on this context. To prove that

$$\Delta\Phi = (\Phi \otimes \Phi)\Delta \tag{3}$$

where on the right hand side Δ is the coproduct of $\mathcal{U}(\mathcal{L})$, which is the universal extension of the Lie algebra homomorphism

$$\mathcal{L} \ni L \mapsto L \otimes 1_{\mathcal{U}(\mathcal{L})} + 1_{\mathcal{U}(\mathcal{L})} \otimes L \in \mathcal{U}(\mathcal{L}) \otimes \mathcal{U}(\mathcal{L}),$$

note first that (3) holds for actions on arbitrary $L \in \mathcal{L} \subset \mathcal{U}(\mathcal{L})$ by (2) and (1). That (3) holds in full generality now follows from the facts that $\mathcal{U}(\mathcal{L})$ is generated as an algebra by $\mathcal{L}$ and that both coproducts and the maps Φ and $\Phi \otimes \Phi$ are multiplicative. That Φ intertwines antipodes follows similarly from the facts that it does so by definition when applied to each $L \in \mathcal{L} \subset \mathcal{U}(\mathcal{L})$ (whose antipode in $\mathcal{U}(\mathcal{L})$ is $-L$), that Φ is multiplicative and that antipodes are antimultiplicative. □

We introduce the (right) *differential map* d from $\mathcal{T}(\mathcal{L})$ to $\mathcal{T}(\mathcal{L}) \otimes \mathcal{L}$ by linear extension of its action on homogeneous product tensors

$$d(1_{\mathcal{T}(\mathcal{L})}) = 0,\ d(L_1 \otimes L_2 \otimes \cdots L_n) = (L_1 \otimes L_2 \otimes \cdots L_{n-1}) \otimes L_n;$$

equivalently, for arbitrary T,

$$d(T) = (\mathrm{id}_{\mathcal{T}(\mathcal{L})} \otimes \varepsilon_1)(\Delta(T) - T \otimes 1_{\mathcal{T}(\mathcal{L})}).$$

Here the right hand side is well defined because $\Delta(T) - T \otimes 1_{\mathcal{T}(\mathcal{L})} \in \mathcal{T}(\mathcal{L}) \otimes \mathcal{K}(\varepsilon)$, as is clear by linear extension of the case when T is a homogeneous product tensor which follows from (1).

Note that d maps $\mathcal{U}(\mathcal{L})$ to $\mathcal{U}(\mathcal{L}) \otimes \mathcal{L}$.

3. Algebraic theory of double product integrals.

Let $\mathcal{A}$ be a unital associative algebra which we call the *system algebra* and let $k[h] \in h(\mathcal{A}\otimes\mathcal{L})[[h]]$ be a formal power series with vanishing zero-order term and coefficients in the tensor product algebra $\mathcal{A}\otimes\mathcal{L}$. We define the *forward* and *backward product integrals generated by* $k[h]$, using place notation, as the formal power series with coefficients in $\mathcal{A}\otimes\mathcal{T}(\mathcal{L}) = \bigoplus_{n=0}^{\infty}\mathcal{A}\otimes(\bigotimes^n\mathcal{L})$

$$\overrightarrow{\prod}(1+dk[h]) = \sum_{n=0}^{\infty} k[h]^{0,1}k[h]^{0,2}\dots k[h]^{0,n}$$
$$\overleftarrow{\prod}(1+dk[h]) = \sum_{n=0}^{\infty} k[h]^{0,n}k[h]^{0,n-1}\dots k[h]^{0,1}$$

where the zero-order term of the series is defined to be $1_{\mathcal{A}\otimes\mathcal{T}(\mathcal{L})}$ in each case. Equivalently they are the solutions $X[h]$ and $Y[h]$ of the algebraic stochastic differential equations

$$(\mathrm{id}_{\mathcal{A}}\otimes d)X[h] = X[h]^{1,2}k[h]^{1,3},\ (\mathrm{id}_{\mathcal{A}}\otimes\varepsilon)X[h] = 1_{\mathcal{A}}$$
$$(\mathrm{id}_{\mathcal{A}}\otimes d)Y[h] = k[h]^{1,3}Y[h]^{1,2},\ (\mathrm{id}_{\mathcal{A}}\otimes\varepsilon)Y[h] = 1_{\mathcal{A}}.$$

If $\mathcal{A}$ is non-unital we may still define the *decapitated* product integrals $\widehat{\overrightarrow{\prod}}(1+dk[h])$ and $\widehat{\overleftarrow{\prod}}(1+dk[h])$ either by omitting the zero-order term in the series expansions or as the solutions $\hat{X}[h]$ and $\hat{Y}[h]$ of

$$(\mathrm{id}_{\mathcal{A}}\otimes d)\hat{X}[h] = \left(\hat{X}[h]^{1,2} + 1^2{}_{\mathcal{T}(\mathcal{L})}\right)k[h]^{1,3},\ (\mathrm{id}_{\mathcal{A}}\otimes\varepsilon)\hat{X}[h] = 0_{\mathcal{A}}$$
$$(\mathrm{id}_{\mathcal{A}}\otimes d)\hat{Y}[h] = k[h]^{1,3}\left(\hat{Y}[h]^{1,2} + 1^2{}_{\mathcal{T}(\mathcal{L})}\right),\ (\mathrm{id}_{\mathcal{A}}\otimes\varepsilon)\hat{Y}[h] = 0_{\mathcal{A}}.$$

Note that these belong to $h(\mathcal{L}\otimes\mathcal{T}(\mathcal{L}))[[h]]$. Analogous definitions of the product integrals $\overrightarrow{\prod}(1+dk[h]), \overleftarrow{\prod}(1+dk[h]), \widehat{\overrightarrow{\prod}}(1+dk[h])$ and $\widehat{\overleftarrow{\prod}}(1+dk[h])$, now belonging to $(\mathcal{T}(\mathcal{L})\otimes\mathcal{A})[[h]] = (\bigoplus_{n=0}^{\infty}(\bigotimes^n\mathcal{L})\otimes\mathcal{A})[[h]]$ and to $h(\mathcal{T}(\mathcal{L})\otimes\mathcal{A})[[h]]$ respectively, hold for a generator $k[h]\in h(\mathcal{L}\otimes\mathcal{A})[[h]]$.

Now let $r[h]\in h(\mathcal{L}\otimes\mathcal{L})[[h]]$ be a formal power series with vanishing zero-order term and coefficients in $\mathcal{L}\otimes\mathcal{L}$. Taking the first copy of $\mathcal{L}$ in $\mathcal{L}\otimes\mathcal{L}$ as the system algebra we may form the decapitated product integral $\widehat{\overleftarrow{\prod}}(1+dr[h]) \in h(\mathcal{L}\otimes\mathcal{T}(\mathcal{L}))[[h]]$. Using this as a generator we may then form $\overrightarrow{\prod}(1+\widehat{\overleftarrow{\prod}}(1+dr[h]))$. Alternatively, taking the second copy of $\mathcal{L}$ as the system algebra we may form first $\widehat{\overrightarrow{\prod}}(1+dr[h])\in h(\mathcal{T}(\mathcal{L})\otimes\mathcal{L})[[h]]$ and then using this as a generator, $\overleftarrow{\prod}(1+\widehat{\overrightarrow{\prod}}(1+dr[h]))$. In fact[9] these two

constructions yield one and the same element of $(\mathcal{T}(\mathcal{L}) \otimes \mathcal{T}(\mathcal{L}))[[h]]$ which we define to be the forward-backward double product integral

$$\overrightarrow{\prod}^{\leftarrow} (1 + dr[h]) = \overrightarrow{\prod}(1 + \widehat{\overleftarrow{\prod}}(1 + dr[h])) = \overleftarrow{\prod}(1 + \widehat{\overrightarrow{\prod}}(1 + dr[h])).$$

The double product integrals $\overleftarrow{\prod}^{\rightarrow} (1 + dr[h])$, $\overrightarrow{\prod}^{\rightarrow} (1 + dr[h])$ and $\overleftarrow{\prod}^{\leftarrow} (1 + dr[h])$ are defined similarly by appropriate reversals of arrows.

Double product integrals can be characterised as elements $(\mathcal{T}(\mathcal{L}) \otimes \mathcal{T}(\mathcal{L}))[[h]]$ of by their behaviour under the actions of the coproduct[9]. For example the forward-backward integral is characterised as follows

Theorem 3.1. *Let $W[h]$ be a nonzero element of $(\mathcal{T}(\mathcal{L}) \otimes \mathcal{T}(\mathcal{L}))[[h]]$. Then the following are equivalent:*

- *There exists $r[h] \in h(\mathcal{L} \otimes \mathcal{L})[[h]]$ such that $W[h] = \overrightarrow{\prod}^{\leftarrow} (1 + dr[h])$.*
- $$\left(\Delta \otimes id_{\mathcal{T}(\mathcal{L})}\right) W[h] = W[h]^{1,3} W[h]^{2,3}$$
  $$\left(id_{\mathcal{T}(\mathcal{L})} \otimes \Delta\right) W[h] = W[h]^{1,3} W[h]^{1,2}.$$

Similar characterisations for otherwise directed double product integrals are found by reversing arrows in the first condition and reversing products in the second.

For $r[h] \in h(\mathcal{L} \otimes \mathcal{L})[[h]]$ the *quasi-inverse* is the unique element $r'[h]$ of $h(\mathcal{L} \otimes \mathcal{L})[[h]]$ such that

$$r[h] + r'[h] + r[h]r'[h] = r'[h] + r[h] + r'[h]r[h] = 0.$$

The following theorem is proved in[10].

Theorem 3.2. *The inverse of $\overrightarrow{\prod}^{\leftarrow} (1 + dr[h])$ is $\overleftarrow{\prod}^{\rightarrow} (1 + dr'[h])$ where $r'[h]$ is the quasi-inverse of $r[h]$.*

4. The quantum Yang-Baxter equation and quantisation of Lie bialgebras.

In[10] the following theorem is proved.

Theorem 4.1. *A necessary and sufficient condition that $R[h] = \overrightarrow{\prod}^{\leftarrow} (1 + dr[h])$ satisfy the quantum Yang-Baxter equation*

$$R[h]^{1,2} R[h]^{1,3} R[h]^{2,3} = R[h]^{2,3} R[h]^{1,3} R[h]^{1,2}$$

in $(\mathcal{T}(\mathcal{L})\otimes\mathcal{T}(\mathcal{L})\otimes\mathcal{T}(\mathcal{L})[[h]]$ *is that* $r[h]$ *satisfy*

$$r[h]^{1,2}r[h]^{1,3}+r[h]^{1,2}r[h]^{2,3}+r[h]^{1,3}r[h]^{2,3}+r[h]^{1,2}r[h]^{1,3}r[h]^{2,3}$$
$$=r[h]^{1,3}r[h]^{1,2}+r[h]^{2,3}r[h]^{1,2}+r[h]^{2,3}r[h]^{1,3}+r[h]^{2,3}r[h]^{1,3}r[h]^{1,2} \quad (4)$$

in $(\mathcal{L}\otimes\mathcal{L}\otimes\mathcal{L})[[h]]$.

Note that the condition (4) is equivalent to $1+dr[h]$ satisfying the quantum Yang-Baxter equation in the algebra $(\mathcal{L}'\otimes\mathcal{L}'\otimes\mathcal{L}')[[h]]$ where $\mathcal{L}'$ is got by adjoining a unit element 1 to $\mathcal{L}$.

Equating coefficients of powers of h in (4) we obtain, for the lowest order coefficient $r_1\in\mathcal{L}\otimes\mathcal{L}$,

$$r_1^{1,2}r_1^{1,3}+r_1^{1,2}r_1^{2,3}+r_1^{1,3}r_1^{2,3}=r_1^{1,3}r_1^{1,2}+r_1^{2,3}r_1^{1,2}+r_1^{2,3}r_1^{1,3},$$

that is, r_1 satisfies the classical Yang-Baxter equation

$$CYB(r_1)\equiv\left[r_1^{1,2},r_1^{1,3}\right]+\left[r_1^{1,2},r_1^{2,3}\right]+\left[r_1^{1,3},r_1^{2,3}\right]=0,$$

while for higher order coefficients $CYB(r_n)$,is equal to a polynomial expression in $r_1^{j.k},r_2^{j.k},\dots r_{n-1}^{j.k}$, $(j,k)=(1,2),(1,3),(2,3)$. It can be shown[2] that, for given r_1 satisfying $CYB(r_1)=0$, the resulting hierarchy of inhomogeneous linear equations for the successive higher order equations has solutions so that a corresponding solution $r[h]$ of (4) can be constructed. For example, if $dr_1=d\Lambda\otimes dA-dA\otimes d\Lambda$. where $d\Lambda$ and dA are the Itô differentials of the conservation and annihilation processes Λ and A of quantum stochastic calculus satisfying $(d\Lambda)^2=d\Lambda$, $dAd\Lambda=dA$, $d\Lambda dA=(dA)^2=0$, then it may be verified, firstly, that $CYB(dr_1)=0$, and, secondly, that the hierarchy is satisfied by taking each $dr_n=0$ for $n\geq 2$. Hence $\overrightarrow{\prod}^{\leftarrow}(1+h(d\Lambda\otimes dA-dA\otimes d\Lambda))$ satisfies the quantum Yang Baxter equation.

A solution of the quantum Yang-Baxter equation of form

$$R[h]=\overrightarrow{\prod}^{\leftarrow}(1+dr[h])$$

may be used to construct a deformation coproduct

$$\Delta[h](\alpha[h])=R[h]\Delta(\alpha[h])R[h]^{-1}$$

in particular for the algebra $\mathcal{A}[[h]]$of formal power series with coefficients in $\mathcal{T}(\mathcal{L})$ whose zero order coefficients lie in the symmetric subalgebra $\mathcal{S}(\mathcal{L})\subset\mathcal{T}(\mathcal{L})$. Indeed it is clear that the map $\Delta[h]$ defined by this equation inherits multiplicativity from Δ, while its coassociativity follows from that of Δ together with the quantum Yang-Baxter equation for $R[h]$ and

$R[h]^{-1}$. The resulting deformation Hopf algebra provides a simple method of quantisation of the Lie bialgebra got by equipping the Lie algebra $\mathcal{L}$ with the cobracket δ which is the infinitesimal of $\Delta[h]$, given by

$$\begin{aligned}\delta(L) &= \frac{1}{h}\left(\Delta[h](L) - \Delta_{\rm op}[h](L)\right) + o(h)\\ &= (L\otimes 1 + 1\otimes L)(r_1 - \tau_{(2.1)}r_1) - (r_1 - \tau_{(2.1)}r_1)(L\otimes 1 + 1\otimes L).\end{aligned}$$

Here $\Delta_{\rm op}[h]$ denotes the opposite coproduct, $\Delta_{\rm op}[h] = \tau_{(2.1)}\Delta[h]$ where (for an arbitrary vector space $\mathcal{V}$) $\tau_{(2.1)}$ is the flip linear transformation in $\mathcal{V}\otimes\mathcal{V}$ which exchanges the components of product tensors.

Thus the Lie bialgebra $\mathcal{L}$ is the quasitriangular Lie bialgebra generated by the solution r_1 of the classical Yang Baxter equation[3,12] and we have an effective general method for the quantisation of such Lie bialgebras.

5. Representation and corepresentation properties in Fock space.

In this section we take $\mathcal{L}$ to be the algebra $\mathcal{L} = \mathbb{C}\langle d\Lambda_\alpha^\beta; \alpha,\beta = 0,1,2,\ldots,N\rangle$ of Itô differentials of N-dimensional quantum stochastic calculus, with product defined by

$$d\Lambda_\alpha^\beta d\Lambda_\delta^\gamma = \hat\delta_\alpha^\gamma d\Lambda_\delta^\beta$$

where $\hat\delta_\alpha^\gamma = 1$ if $\alpha = \gamma \neq 0$ and $\hat\delta_\alpha^\gamma = 0$ otherwise. The corresponding processes consist of operators in the Fock space $\mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N))$ over the Hilbert space $L^2(\mathbb{R}_+;\mathbb{C}^N) = L^2(\mathbb{R}_+)\otimes\mathbb{C}^N$ acting on the exponential domain, with matrix elements between exponential vectors given by

$$\left\langle e(f), \Lambda_\alpha^\beta(t)e(g)\right\rangle = \int_0^t f^\beta(s)g_\alpha(s)\,ds\,\left\langle e(f),e(g)\right\rangle$$

where for $f = (f_1,f_2,\ldots,f_N)^T \in L^2(\mathbb{R}_+;\mathbb{C}^N)$, $f_0\equiv 1$ and $f^\alpha(s) = \bar f_\alpha(s)$.

The iterated stochastic integral

$$\int_{a<s_1<s_2<\cdots<s_n<b} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\,d\Lambda_{\alpha_2}^{\beta_2}(s_2)\ldots d\Lambda_{\alpha_n}^{\beta_n}(s_n)$$

between $a<b\in\mathbb{R}_+$ may be defined by its matrix elements between exponential vectors

$$\begin{aligned}&\left\langle e(f), \int_{a<s_1<s_2<\cdots<s_n<b} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\,d\Lambda_{\alpha_2}^{\beta_2}(s_2)\ldots d\Lambda_{\alpha_n}^{\beta_n}(s_n)e(g)\right\rangle\\ &= \int_{a<s_1<s_2<\cdots<s_n<b} f^{\beta_1}(s_1)g_{\alpha_1}(s_1)f^{\beta_2}(s_2)g_{\alpha_2}(s_2)\ldots f^{\beta_n}(s_n)g_{\alpha_n}(s_n)\left\langle e(f),e(g)\right\rangle.\end{aligned}$$

The action of its adjoint on exponential vectors is obtained by exchanging α and β indices while conserving the order of the terms in the integrator;

$$\left(\int d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2)\ldots\, d\Lambda_{\alpha_n}^{\beta_n}(s_n)\right)^\dagger$$
$$= \int d\Lambda_{\beta_1}^{\alpha_1}(s_1)\, d\Lambda_{\beta_2}^{\alpha_2}(s_2)\ldots\, d\Lambda_{\beta_n}^{\alpha_n}(s_n),$$

where the integrals are performed over the domain $a < s_1 < s_2 < \cdots < s_n < b$.

For $a < b \in \mathbb{R}_+$ and $X \in \mathcal{T}(\mathcal{L})$ we define an operator $J_a^b(X)$ on the exponential domain by linear extension of the prescription

$$J_a^b\left(d\Lambda_{\alpha_1}^{\beta_1} \otimes d\Lambda_{\alpha_2}^{\beta_2} \otimes \ldots \otimes\, d\Lambda_{\alpha_n}^{\beta_n}\right)$$
$$= \int_{a<s_1<s_2<\cdots<s_n<b} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2)\ldots\, d\Lambda_{\alpha_n}^{\beta_n}(s_n).$$

Then the map J_a^b is a weak representation of the algebra $\mathcal{T}(\mathcal{L})$ in the sense that[11].

Theorem 5.1. *For arbitrary $X, Y \in \mathcal{T}(\mathcal{L})$ and exponential vectors $e(f), e(g)$*

$$\left\langle \left(J_a^b(X)\right)^\dagger e(f), J_a^b(Y)e(g)\right\rangle = \left\langle e(f), J_a^b(XY)e(g)\right\rangle.$$

Note that each operator belonging to the range of J_a^b is biadapted to the interval $]a,b]$ in the sense that it takes the form of a double ampliation $\mathrm{id}_{\mathcal{F}(L^2(\mathbb{R}_+;]0,a];\mathbb{C}^N)} \otimes P_a^b \otimes \mathrm{id}_{\mathcal{F}(L^2(\mathbb{R}_+;]b,\infty];\mathbb{C}^N)}$ in the joint splitting at a and b.

$$\mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N))$$
$$= \mathcal{F}(L^2(\mathbb{R}_+;]0,a];\mathbb{C}^N) \otimes \mathcal{F}(L^2(\mathbb{R}_+;]a,b];\mathbb{C}^N) \otimes \mathcal{F}(L^2(\mathbb{R}_+;]b,\infty];\mathbb{C}^N) \quad (5)$$

where P_a^b is an operator on the exponential domain in $\mathcal{F}(L^2(\mathbb{R}_+;]a,b];\mathbb{C}^N)$.

The following theorem must be well-known to many quantum probabilists but a formal proof seems to be lacking in the literature, so one is given here. For $a < b < c \in \mathbb{R}_+$ we denote by $F_a^c(b)$ the splitting isomorphism from $\mathcal{F}(L^2(\mathbb{R}_+;]a,c];\mathbb{C}^N)$ to $\mathcal{F}(L^2(\mathbb{R}_+;]a,b];\mathbb{C}^N) \otimes \mathcal{F}(L^2(\mathbb{R}_+;]b,c];\mathbb{C}^N)$ under which each exponential vector $e(f)$ is mapped to the product vector $e(f|_{]a,b]}) \otimes e(f|_{]b,c]})$. We identify elements of the range of J_a^b with the operators of which they are the double ampliations corresponding to the double splitting (5).

Theorem 5.2. *For $a < b < c \in \mathbb{R}_+$*

$$F_a^c(b) \cdot J_a^c = (J_a^b \otimes J_b^c)\Delta. \tag{6}$$

Proof. Expressing the simplex

$$\Delta_a^c(n) = \{(s_1, s_2, \ldots, s_n) \in \mathbb{R}_+^n : a < s_1 < s_2 < \cdots < s_n < c\}$$

as the disjoint union of Cartesian products

$$\Delta_a^c(n) = \bigcup_{j=0}^{n} \left(\Delta_a^b(j) \times \Delta_b^c(n-j)\right),$$

where, by definition, the Cartesian product of a set with $\Delta_a^b(0)$ or with $\Delta_b^c(0)$ is itself, we have for arbitrary $n \in \mathbb{N}$ and $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, , \beta_2, \ldots, \beta_n \in \{0, 1, 2, \ldots, N\}$

$$\begin{aligned}
&F_a^c(b) J_a^c(d\Lambda_{\alpha_1}^{\beta_1} \otimes d\Lambda_{\alpha_2}^{\beta_2} \otimes \ldots \otimes d\Lambda_{\alpha_n}^{\beta_n}) \\
&= F_a^c(b) \int_{a<s_1<s_2<\cdots<s_n<c} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2) \ldots d\Lambda_{\alpha_n}^{\beta_n}(s_n) \\
&= F_a^c(b) \int_{(s_1,s_2,\ldots,s_n)\in\Delta_a^c(n)} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2) \ldots d\Lambda_{\alpha_n}^{\beta_n}(s_n) \\
&= F_a^c(b) \int_{(s_1,s_2,\ldots,s_n)\in\bigcup_{j=0}^{n}\left(\Delta_a^b(j)\times\Delta_b^c(n-j)\right)} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2) \ldots d\Lambda_{\alpha_n}^{\beta_n}(s_n) \\
&= \sum_{j=0}^{n} \int_{(s_1,s_2,\ldots,s_j)\in\Delta_a^b(j)} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2) \ldots d\Lambda_{\alpha_j}^{\beta_j}(s_j) \\
&\quad \otimes \int_{(s_{j+1},s_{j+2},\ldots,s_n)\in\Delta_b^c(n-j)} d\Lambda_{\alpha_{1+1}}^{\beta_{j+1}}(s_{j+1})\, d\Lambda_{\alpha_{j+2}}^{\beta_{j+2}}(s_{j+2}) \ldots d\Lambda_{\alpha n}^{\beta n}(s_n) \\
&= \sum_{j=0}^{n} \int_{a<s_1<s_2<\cdots<s_j<b} d\Lambda_{\alpha_1}^{\beta_1}(s_1)\, d\Lambda_{\alpha_2}^{\beta_2}(s_2) \ldots d\Lambda_{\alpha_j}^{\beta j}(s_j) \\
&\quad \otimes \int_{b<s_{j+1}<s_{j+2}<\cdots<s_n<c} d\Lambda_{\alpha_{1+1}}^{\beta_{j+1}}(s_{j+1})\, d\Lambda_{\alpha_{j+2}}^{\beta_{j+2}}(s_{j+2}) \ldots d\Lambda_{\alpha n}^{\beta n}(s_n) \\
&= \sum_{j=0}^{n} J_a^b(d\Lambda_{\alpha_1}^{\beta_1} \otimes d\Lambda_{\alpha_2}^{\beta_2} \otimes \ldots \otimes d\Lambda_{\alpha_j}^{\beta_j}) \otimes \cdots \cdot J_b^c(d\Lambda_{\alpha_{j+1}}^{\beta_{j+1}} \otimes d\Lambda_{\alpha_{j+2}}^{\beta_{j+2}} \otimes \ldots \otimes d\Lambda_{\alpha_n}^{\beta_n}) \\
&= (J_a^b \otimes J_b^c)\Delta(d\Lambda_{\alpha_1}^{\beta_1} \otimes d\Lambda_{\alpha_2}^{\beta_2} \otimes \ldots \otimes d\Lambda_{\alpha_n}^{\beta_n}).
\end{aligned}$$

The result now follows by linearity of both sides of (6). □

6. Double product integrals in Fock space.

Given intervals $a < b$ and $c < d$ in $\mathbb{R}_+$ and a double product such as $R[h] = \overrightarrow{\prod}^{\leftarrow} (1 + dr[h])$ where the coefficients of $dr[h]$ belong to the algebra $\mathcal{L}$ of Itô differentials we may try to construct a corresponding operator on the exponential domain in the double Fock space

$$\mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N)) \otimes \mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N)) = \mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N) \oplus L^2(\mathbb{R}_+;\mathbb{C}^N))$$

by applying the homomorphism $J_a^b \otimes J_c^d$ to the coefficients of the formal power series $R[h]$and replacing h by a complex variable z. Assuming convergence we may thus define a family of operators on the exponential domain

$$R_{a,c}^{b,d}(z) = \overrightarrow{\prod_{(s,t)\in]a,b]\times]c,d]}}^{\leftarrow} (1 + r(z;ds,dt)) = \left(J_a^b \otimes J_c^d\right) \overrightarrow{\prod}^{\leftarrow} (1 + dr[h])\,|_{h=z}\ .$$

In view of Theorem 3 and Theorem 6 we should expect these to satisfy

$$R_{a,c}^{b,d}(z) = R_{a,c}^{x,d}(z) R_{x,c}^{b,d}(z) = \ R_{a,y}^{b,d}(z) R_{a,c}^{b,y}(z),\ a < x < b, c < y < d, \qquad (7)$$

at least in the weak sense where the operator products are circumvented by moving the first operator to its adjoint in exponential matrix elements, that is, for arbitrary exponential vectors $e(f), e(g), e(p), e(q)$

$$\begin{aligned}
&\left\langle e(f) \otimes e(g), R_{a,c}^{b,d}(z) e(p) \otimes e(q) \right\rangle \\
&= \left\langle \left(R_{a,c}^{x,d}(z)\right)^\dagger e(f) \otimes e(g), R_{x,c}^{b,d}(z) e(p) \otimes e(q) \right\rangle \\
&= \left\langle \left(R_{a,y}^{b,d}(z)\right)^\dagger e(f) \otimes e(g), R_{a,c}^{b,y}(z) e(p) \otimes e(q) \right\rangle,
\end{aligned}$$

as well as

$$\left(\overrightarrow{\prod_{(s,t)\in]a,b]\times]c,d]}}^{\leftarrow} (1 + r(z;ds,dt)) \right)^\dagger = \overleftarrow{\prod_{(s,t)\in]a,b]\times]c,d]}}^{\rightarrow} (1 + r^\dagger(\bar{z};ds,dt))$$

$$\left(\overrightarrow{\prod_{(s,t)\in]a,b]\times]c,d]}}^{\leftarrow} (1 + r(z;ds,dt)) \right)^{-1} = \overleftarrow{\prod_{(s,t)\in]a,b]\times]c,d]}}^{\rightarrow} (1 + r'(z;ds,dt))$$

where $r^\dagger[h]$ is formed by appling the tensor product involution to the coefficients of $r[h]$ and $r'[h]$ is the quasiinverse of $r[h]$, the latter again at least

in the weak sense that

$$\begin{aligned}&\langle e(f)\otimes e(g), e(p)\otimes e(q)\rangle\\ &= \left\langle \prod_{(s,t)}^{\leftarrow\rightarrow}(1+r^{\dagger}(\bar{z};ds,dt))e(f)\otimes e(g), \prod_{(s,t)}^{\leftarrow\rightarrow}(1+r^{'}(z;ds,dt))e(p)\otimes e(q)\right\rangle\\ &= \left\langle \prod_{(s,t)}^{\rightarrow\leftarrow}(1+r'^{\dagger}(z;ds,dt))e(f)\otimes e(g), \prod_{(s,t)}^{\rightarrow\leftarrow}(1+r(z;ds,dt))e(p)\otimes e(q)\right\rangle,\end{aligned}$$

where $(s,t)\in]a,b]\times]c,d]$.

However an analytic theory of double product integrals in Fock space of this kind is obstructed by divergence problems. It is evident that if the formal power series $r[h]$ has many nonzero coefficients these may be of a novel and possibly intractable kind. But even if we make the simpifying assumption that only the first order coefficient r_1 is nonzero (which we do henceforth) divergence problems remain. For example the (commutative) double product $\prod(1+zd\Lambda\otimes d\Lambda)$,where Λ is the conservation process of one-dimensional quantum stochastic calculus is divergent for all nonzero values of z in the sense that for arbitrary f and g non-zero on $]a,b]$ and f',g' non zero on $]c,d]$, the exponential matrix element

$$\left\langle e(f)\otimes e(f'), \prod_{(s,t)\in]a,b]\times]c,d]}(1+zd\Lambda(s)\otimes d\Lambda(t))e(g)\otimes e(g')\right\rangle$$

has zero radius of convergence in z. On the other hand it is clear (since $(dA)^2=0$) that $\prod_{(s,t)\in]a,b]\times]c,d]}(1+zdA(s)\otimes dA(t))$ exists in this sense and is given by $\exp\left(a(\chi_a^b)\otimes a(\chi_c^d)\right)$ where $a(\chi_a^b)$ is the annihilation operator corresponding to the indicator function of the interval $]a,b]$. More interestingly the double product $\prod_{(s,t)\in]a,b]\times]c,d]}(1+z(dA^{\dagger}(s)\otimes dA(t)-dA(s)\otimes dA^{\dagger}(t))$ exists (for all real z) as a unitary operator satisfying (7)[7,8]. It is conjectured that $\prod_{(s,t)\in]a,b]\times]c,d]}(1+z(d\Lambda(s)\otimes dA(t)-d\Lambda(s)\otimes dA^{\dagger}(t))$ exists and continues to satisfy the quantum Yang-Baxter equation in $\mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N))\otimes\mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N))\otimes\mathcal{F}(L^2(\mathbb{R}_+;\mathbb{C}^N))$.

Bibliography

1. N Bourbaki, Algebra I, Addison-Wesley (1974), Chapter III Section 11.5, Lemma 1.
2. B Enriquez, Quantization of Lie bialgebras and shuffle algebras of Lie algebras, Selecta Mathematica (new series) **7** 321-407 (2001).
3. P Etingof and O Schiffman *Lectures on quantum groups*, International Press (1998).

4. T M W Eyre, *Quantum stochastic calculus and representations of Lie bialgebras*, Springer lecture Notes in Mathematics **1692 (**1998).
5. R L Hudson, Itô calculus and quantisation of Lie bialgebras, Annales de l'Institue H Poicaré-PR **41** (P-A Meyer memorial volume), 375-390 (2005).
6. R L Hudson, The method of double product integrals in quantisation of Lie bialgebras, pp303-319, in *Quantum Probability and Infinite Dimensional Analysis from Foundations to Applications*, Proceedings, Greifswald 2003, eds M Schürmann et al, World Scientific (2005).
7. R L Hudson, An analytic double product integral, to appear in Proceedings of 2005 QP workshop.
8. R L Hudson, A double dilation constructed from a double product of rotations, to appear in Markov Processes and Applications (J T Lewis mmemorial volume, 2007).
9. R L Hudson and S Pulmannová, Double product integrals and Enriquez quantisation of Lie bialgebras I: The quasitriangularity relations, Journal of Mathematical Physics **45**, 2090-2105 (2004).
10. R L Hudson and S Pulmannová, Double product integrals and Enriquez quantisation of Lie bialgebras II: The quantum Yang-Baxter equation, Letters in Mathematical Physics **72**, 211-224 (2005).
11. R L Hudson and S Pulmannová, Chaotic expension of elements of the universal enveloping algebra of a Lie algebra associated with a quantum stochastic calculus, Proceedings of the London Mathematical Society (3rd Series) **77**, 462-480 (1998).
12. S Majid, *Foundations of quantum group theory*, Cambridge University press (1995).
13. P-A Meyer, *Quantum probability for probabilists*, Springer lecture Notes in Mathematics **1538** (1993)
14. K R Parthasarathy, *An introduction to quantum stochastic calculus*, Birkhäuser (1992)

PRODUCT SYSTEMS; A SURVEY WITH COMMUTANTS IN VIEW

MICHAEL SKEIDE*

Dipartimento S.E.G.e S., Università degli Studi del Molise, Via de Sanctis, 86100 Campobasso, Italy, E-mail: skeide@math.tu-cottbus.de, Homepage: http://www.math.tu-cottbus.de/INSTITUT/lswas/_skeide.html

The theory of product systems both of Hilbert spaces (Arveson systems) and product systems of Hilbert modules has reached a status where it seems appropriate to rest a moment and to have a look at what is known so far and what are open problems. However, the attempt to give an approximately complete account in view pages is destined to fail already for Arveson systems since Tsirelson, Powers and Liebscher have discovered their powerful methods to construct large classes of examples. In this survey we concentrate on that part of the theory that works also for Hilbert modules. This does not only help to make a selection among the possible topics, but it also helps to shed some new light on the case of Arveson systems. Often, proofs that work for Hilbert modules also lead to simpler proofs in the case of Hilbert spaces. We put emphasis on those aspect that arise from recent results about *commutants* of von Neumann correspondences, which, in the case of Hilbert spaces, explain the relation between the Arveson system and the Bhat system associated with an E_0–semigroup on $\mathcal{B}(H)$.

*K*eywords: Quantum dynamics, quantum probability, Hilbert modules, product systems, E_0–semigroups. 2000 AMS-Subject classification: 46L53; 46L55; 46L08; 60J25; 81S25; 12H20.

1. Introduction

A *product system of Hilbert spaces* is a family $E^{\otimes} = \big(E_t\big)_{t\in\mathbb{R}_+}$ of Hilbert spaces that factor as

$$E_{s+t} \;\cong\; E_s \otimes E_t$$

by means of an associative bilinear multiplication $E_s \times E_t \ni (x_s, y_t) \mapsto x_s y_t \in E_{s+t}$. (Depending on the application, there are also technical con-

*MS is supported by research funds of the University of Molise and the Italian MIUR (PRIN 2005).

ditions about continuity or measurability of sections. We speak about this later on.) The definition of such product systems is due to Arveson.[Arv89a] It is motivated by Arveson's construction that associates with every E_0*–semigroup* (a semigroup of normal unital endomorphisms) of the $\mathcal{B}(H)$ of all adjointable operators on a Hilbert space H a product system $E^{A^{\otimes}}$. If H is infinite-dimensional and separable, then the product system determines the E_0–semigroup up to *cocycle conjugacy*. In a series of four articles[Arv89a,Arv90a,Arv89b,Arv90b] Arveson showed the *fundamental theorem*, namely, that every product system of Hilbert spaces is the one associated with a suitable E_0–semigroup. Thus, there is a one-to-one correspondence between product systems (up to isomorphism) and E_0–semigroups (up to cocycle conjugacy). In the sequel, we will speak about an ***Arveson system*** if we intend a product system of Hilbert spaces. In particular, we will speak about the Arveson system associated with an E_0–semigroup on $\mathcal{B}(H)$.

Meanwhile, product systems of Hilbert *bimodules* or *correspondences* made appearance in many contexts. Bhat and Skeide[BS00] constructed a product system of correspondences over a (unital) C^*–algebra $\mathcal{B}$ from a (unital) CP-semigroup on $\mathcal{B}$. (See also the discussion of Muhly and Solel[MS02] in Remark 6.6.) This construction overcomes constructions by Bhat[Bha96] and by Arveson[Arv96] who construct an Arveson system starting from a CP-semigroup on $\mathcal{B}(H)$ by, first, *dilating* in a unique way the CP-semigroup to a *minimal* E_0–semigroup and, then, constructing the Arveson system of that E_0–semigroup. The construction of,[BS00] instead, is direct and allows, then, to construct the minimal dilation in a transparent way. Only later, Skeide[Ske02,Ske05a,Ske04] associated in several ways with an E_0–semigroup a product system. Now the E_0–semigroup acts on the algebra $\mathcal{B}^a(E)$ of all adjointable maps on a Hilbert $\mathcal{B}$–module E. Although historically earlier, the approach to product systems from CP-semigroups (that is, *irreversible* quantum dynamics) has the disadvantage that not all product systems arise in that way. While one of the latest results (Skeide;[Ske07] still in preparation) asserts that *cum grano salis* every product system comes from an E_0–semigroup (that is, *reversible* quantum dynamics in a sense we specify later on). So the approach via E_0–semigroups allows a more coherent discussion. In this survey we will concentrate on this connection between product systems and E_0–semigroups, while we will have no space to discuss also the connections with CP-semigroup and their dilations; see Skeide.[Ske03b] Also basic classification of product systems must be sacrificed; see Skeide.[Ske03b,Ske06g]

The basic factorization property of the symmetric Fock space

$$\Gamma(H_1 \oplus H_2) \;=\; \Gamma(H_1) \otimes \Gamma(H_2)$$

(H_1 and H_2 some Hilbert spaces) has drawn attention since a long time. In the form

$$\Gamma(L^2([r,t],K)) \;=\; \Gamma(L^2([r,s],K)) \otimes \Gamma(L^2([s,t],K)), \quad r \le s \le t \qquad (*)$$

(K a Hilbert space) it made appearance in the work of Araki[Ara70] and Streater[Str69] on current representations of Lie algebras, in the work of Parthasarathy and Schmidt[PS72] about Lévy processes (culminating in Schürmann's work[Sch93] on quantum Lévy processes) and in quantum stochastic calculus on the symmetric Fock space initiated by Hudson and Parthasarathy.[HP84]

Let us put $E_t = \Gamma(L^2([0,t],K))$. Then, from the beginning, there are two possibilities to use $(*)$ in order to define an isomorphism $E_s \otimes E_t \cong E_{s+t}$, namely,

$$E_s \otimes E_t \;\cong\; \mathrm{s}_t E_s \otimes E_t \;\cong\; E_{s+t} \quad \text{and} \quad E_s \otimes E_t \;\cong\; E_s \otimes \mathrm{s}_s E_t \;\cong\; E_{s+t},$$

where $\mathrm{s}_t \colon \Gamma(L^2([0,s],K)) \to \Gamma(L^2([t,t+s],K))$ is the time shift. If we consider the *CCR-flow*, that is, the E_0–semigroup induced on $\mathcal{B}\big(\Gamma(L^2(\mathbb{R}_+,K))\big)$ by the time shift, then the associated Arveson system is E_t with the second choice of an isomorphism, that is, with the time shift acting on the right factor in $E_s \otimes E_t$. However, Bhat discovered a second possibility to associate an Arveson with an E_0–semigroup. In the case of the CCR-flow one obtains the same Hilbert spaces E_t but with the first choice of an isomorphism, that is, with the time shift acting on the left factor in $E_s \otimes E_t$. More generally, the ***Bhat system*** associated with any E_0–semigroup shows always to be *anti-isomorphic* to the associated Arveson system.

This ambivalence in the tensor product of Hilbert spaces, where we may switch the order of factors without changing (up to canonical isomorphism) the resulting space, is by far less innocent than it appears at the first sight. Nothing like this is true in the module case for the tensor product of correspondences over $\mathcal{B}$. (It is very well possible that in one order their tensor product is $\{0\}$, while in the other order it is not.) In fact, we will see that the construction of a product system of correspondences over $\mathcal{B}$ from an E_0–semigroup on $\mathcal{B}^a(E)$ for some Hilbert $\mathcal{B}$–module E corresponds to the construction of the Bhat system of an E_0–semigroup on $\mathcal{B}(H)$. Also the construction of product system following the ideas of Arveson is still possible. However, it yields a product system of correspondences over the commutant $\mathcal{B}'$ of $\mathcal{B}$ and works nicely only for von Neumann algebras $\mathcal{B}$. The relation

between these two product systems is that one is the *commutant* of the other. The commutant of a correspondence was introduced in Skeide.[Ske03a]

Let $\mathcal{B}$ be a C^*–algebra. Recall that a ***pre-Hilbert $\mathcal{B}$–module*** is a right $\mathcal{B}$–module E with a sesquilinear inner product $\langle\bullet,\bullet\rangle\colon E\times E\to\mathcal{B}$ satisfying $\langle x,x\rangle\geq 0$ for all $x\in E$ (positivity), $\langle x,yb\rangle=\langle x,y\rangle b$ for all $x,y\in E; b\in\mathcal{B}$ (right linearity), and $\langle x,x\rangle=0\implies x=0$ (definiteness). If definiteness is missing, then E is a ***semi-Hilbert $\mathcal{B}$–module***. (Properties like $\langle x,y\rangle^*=\langle y,x\rangle$ and $\langle xb,y\rangle=b^*\langle x,y\rangle$ are automatic.) The most basic property of the inner product in a semi-Hilbert $\mathcal{B}$–module is the following ***Cauchy-Schwartz inequality***

$$\langle x,y\rangle\langle y,x\rangle\ \leq\ \|\langle y,y\rangle\|\,\langle x,x\rangle.$$

By Cauchy-Schwartz inequality it is possible to quotient out length-zero elements. By Cauchy-Schwartz inequality $\|x\|:=\sqrt{\langle x,x\rangle}$ defines a norm on the pre-Hilbert module E. If E is complete in that norm, then E is a ***Hilbert $\mathcal{B}$–module***. By Cauchy-Schwartz inequality the operator norm turns the algebra of bounded adjointable operators $\mathcal{B}^a(E)$ on the pre-Hilbert module E into a pre–C^*–algebra. Recall that a map a on E is ***adjointable***, if it admits an adjoint a^* such that $\langle x,ay\rangle=\langle a^*x,y\rangle$ for all $x,y\in E$. Every adjointable map is closeable. Therefore, by the *closed graph theorem*, an adjointable map on a Hilbert module is bounded, automatically.

In order to speak about product systems we need the (internal) tensor product, and the tensor product is among bimodules or *correspondences*. If $\mathcal{A}$ is another C^*–algebra, then a ***correspondence*** from $\mathcal{A}$ to $\mathcal{B}$ (or a $\mathcal{A}$–$\mathcal{B}$–correspondence) is a Hilbert $\mathcal{B}$–module with a ***nondegenerate*** representation of $\mathcal{A}$ by adjointable operators. If $\mathcal{A}=\mathcal{B}$, then we speak also of a correspondence over $\mathcal{B}$ or of a $\mathcal{B}$–correspondence.[a] The (***internal***) ***tensor product*** of a correspondence E from $\mathcal{A}$ to $\mathcal{B}$ and a correspondence F from $\mathcal{B}$ to $\mathcal{C}$ is the unique correspondence $E\odot F$ from $\mathcal{A}$ to $\mathcal{C}$ that is generated by elementary tensors $x\odot y$ with inner product

$$\langle x\odot y,x'\odot y'\rangle\ =\ \big\langle y',\langle x,x'\rangle y'\big\rangle \tag{1}$$

and the obvious bimodule operation. Uniqueness is, in the sense of a universal property, up to canonical isomorphism. (In two realizations, simply identify the elementary tensors. For a construction take the vector

[a]The nondegeneracy condition is crucial in all what follows. For the right action of $\mathcal{B}$ on a Hilbert $\mathcal{B}$–module it is automatic. (Exercise: Why?) But, there are left actions that act degenerately. However, in that case we will never say $\mathcal{A}$–$\mathcal{B}$–module, but rather speak of a (possibly degenerate) representation of $\mathcal{A}$.

space tensor product $E \otimes F$, define a semiinner product by (1) and divide by the length-zero elements.) The tensor product applies also if E is just a Hilbert $\mathcal{B}$–module, as every Hilbert $\mathcal{B}$–module E may be viewed as a correspondence from $\mathcal{B}^a(E)$ to $\mathcal{B}$. This also shows that $E \odot F$ carries a canonical nondegenerate left action of $a \in \mathcal{B}^a(E)$ which we denote by $a \odot \mathsf{id}_F$ or, sometimes, simply by a, too. (Attention! The unital embedding $\mathcal{B}^a(E) \to \mathcal{B}^a(E) \odot \mathsf{id}_F \subset \mathcal{B}^a(E \odot F)$ need not be faithful.) By $\mathcal{B}^{a,bil}(F)$ we denote the space of those elements $a \in \mathcal{B}^a(F)$ that are ***bilinear***, that is, which fulfill $a(by) = b(ay)$ for all $b \in \mathcal{B}, y \in F$. There is an embedding $\mathcal{B}^{a,bil}(F) \to \mathsf{id}_E \odot \mathcal{B}^{a,bil}(F) \subset \mathcal{B}^a(E \odot F)$. If E is ***full***, that is, if the ***range ideal*** $\mathcal{B}_E := \overline{\mathsf{span}}\langle E, E\rangle$ in $\mathcal{B}$ is $\mathcal{B}$, then one may show that this embedding is an isomorphism onto the relative commutant of $\mathcal{B}^a(E) \odot \mathsf{id}_F$ in $\mathcal{B}^a(E \odot F)$.

If $(v, w) \mapsto v \cdot w$ is bilinear or sesquilinear operation, then VW is the set $\{v \cdot w \colon v \in V, w \in W\}$. We ***do not*** adopt the convention that $VW = \mathsf{span}\, VW$ or even $VW = \overline{\mathsf{span}}\, VW$.

2. The product system associated with an E_0–semigroup

Let $\mathbb{S}$ be one of the (additive) semigroups $\mathbb{R}_+$ or $\mathbb{N}_0$ (with identity 0). We will refer to $\mathbb{S} = \mathbb{R}_+$ also as the ***continuous time*** case and to $\mathbb{S} = \mathbb{N}_0$ as the ***discrete*** case. We are mainly interested in the continuous time case. In associating with an E_0–semigroup a product system, there is no difference between the discrete and the continuous time case. But knowing how to deal with the discrete case will play a crucial role in showing the converse statement in Sections 3 and 4. For the forward direction in this section, we will discuss first the Hilbert space case and then gradually pass to modules.

Let $\vartheta = \big(\vartheta_t\big)_{t\in\mathbb{S}}$ be an E_0–semigroup on the algebra $\mathcal{B}(H)$ of all adjointable operators on a Hilbert space H. Recall that an E_0***–semigroup*** ϑ on a unital $*$–algebra is a semigroup of unital endomorphisms. If the $*$–algebra is $\mathcal{B}(H)$, then we will require that these endomorphisms are normal, while for the time being we do note pose continuity conditions regarding time dependence of ϑ_t. We mentioned already, that there are essentially two ways to associate with ϑ a product system of Hilbert spaces (***Arveson system***, for short). The first one is Arveson's original construction from,[Arv89a] the second one is due to Bhat.[Bha96] However, only the second construction due to Bhat allows for a direct generalization to Hilbert modules. Arveson's construction works nicely only for von Neumann modules and results in a different product system, the *commutant system*. Even for Hilbert spaces the results of the constructions need not coincide; see

Footnote q. Moreover, he results are related to the original E_0–semigroup in different ways, namely, one (Arveson) by what we will call a *right dilation* and the other (Bhat) by what we will call a *left dilation*. Starting with this section we will discuss product systems and their relations with E_0–semigroups in terms that correspond rather to Bhat's construction. The generalization of Arveson's approach requires the *commutant* of a von Neumann correspondence. We will discuss these things starting from Section 5.

In[Bha96] Bhat chooses a unit vector $\xi \in H$ and defines the Hilbert subspaces

$$E_t^B \;:=\; \vartheta_t(\xi\xi^*)H \tag{2}$$

of H. (Once for all, for an element x in a space with an inner product, we define the map $x^*\colon y \mapsto \langle x, y\rangle$. Consequently, xy^* is the ***rank-one operator*** $z \mapsto x\langle y, z\rangle$.) It is easy to show that the bilinear maps

$$(x, y_t) \;\longmapsto\; xy_t \;:=\; \vartheta_t(x\xi^*)y_t \quad\text{and}\quad (x_s, y_t) \;\longmapsto\; x_s y_t \;:=\; \vartheta_t(x_s\xi^*)y_t \tag{3}$$

define isometries $v_t\colon H \otimes E_t^B \to H$ and $u_{s,t}\colon E_s^B \otimes E_t^B \to E_{s+t}^B$. (Exercise: Check that $u_{s,t}$ is into E_{s+t}.) Using a bounded approximate unit of finite-rank operators and normality, one may show that v_t is surjective. (We discuss this in a minute in the more general context; see Equation (4) and the exercise suggested there.) Now $u_{s,t}$ is just the restriction of v_t to the subspace $E_s^B \otimes E_t^B$ of $H \otimes E_t^B$ and v_t^* maps the subspace E_{s+t} of H into $E_s \otimes E_t$. (Exercise!) This shows that $u_{s,t}$ is onto E_{s+t}. We find

$$\begin{aligned}(xy_s)z_t \;&=\; \vartheta_t((xy_s)\xi^*)z_t \;=\; \vartheta_t(\vartheta_s(x\xi^*)y_s\xi^*)z_t\\ &=\; \vartheta_{s+t}(x\xi^*)\vartheta_t(y_s\xi^*)z_t \;=\; \vartheta_{s+t}(x\xi^*)(y_s z_t) \;=\; x(y_s z_t)\end{aligned}$$

and, by restriction, $(x_r y_s)z_t = x_r(y_s z_t)$. Therefore, the family $E^{B\otimes} = \big(E_t^B\big)_{t\in\mathbb{S}}$ is an (algebraic) Arveson system[b] and the v_t iterate associatively with that product system structure. We call $E^{B\otimes}$ the ***Bhat system associated with*** ϑ.[c]

[b] "Algebraic" refers to that we are not posing any continuity or measurability condition on $E^{B\otimes}$.

[c] Of course, the construction of $E^{B\otimes}$ depends on ξ. But we explain in Proposition 2.4 that all Arveson systems we obtain from different choices are isomorphic. Moreover, we will single out the result of yet another construction as ***the*** Bhat system of ϑ. (That construction has the advantage that it works with choosing a distinguished unit vector. But, even in the Hilbert space case, its simple proof cannot be understood without

In general, whenever for an Arveson system $E^\otimes$ we have a Hilbert space $L \neq \{0\}$ and a family $w^\otimes$ of unitaries $v_t\colon L \otimes E_t \to L$ that iterates associatively with the product system structure, we call the pair $(v^\otimes, L)$ a ***left dilation*** $v^\otimes$ of $E^\otimes$ to L.[d] In that case, by setting $\vartheta^v_t(a) := v_t(a \otimes \mathsf{id}_t)v_t^*$ we define an E_0–semigroup ϑ^v on $\mathcal{B}(L)$. (The semigroup property corresponds exactly to the associativity condition.) Moreover, it is easy to check that the Bhat system of ϑ^v is $E^\otimes$ by identifying $x_t \in E_t$ with $v_t(\xi \otimes x_t) \in \vartheta^v_t(\xi\xi^*)L$. (Exercise: Verify that this identification does not only preserve the spaces but also the product system structure.) In the case of the Bhat system $E^{B^\otimes}$ of an E_0–semigroup ϑ on $\mathcal{B}(H)$ and the left dilation v_t of $E^{B^\otimes}$ to H as constructed before, it follows from

$$\begin{aligned} v_t(a \otimes \mathsf{id}_t)v_t^*(xy_t) &= v_t(a \otimes \mathsf{id}_t)(x \otimes y_t) = v_t(ax \otimes y_t) \\ &= \vartheta_t(ax\xi^*)y_t = \vartheta_t(a)\vartheta_t(x\xi^*)y_t = \vartheta_t(a)(xy_t) \end{aligned}$$

that $\vartheta^v_t = \vartheta_t$. We summarize:

2.1 Proposition. *Let $E^\otimes$ be an (algebraic) Arveson system. The problem of finding an E_0–semigroup that has $E^\otimes$ as associated Bhat system is equivalent to the problem of finding a left dilation of $E^\otimes$.*

Now suppose that ϑ is an E_0–semigroup acting on $\mathcal{B}^a(E)$ where E is a Hilbert $\mathcal{B}$–module. In order to obtain a representation theory of $\mathcal{B}^a(E)$ on E in analogy with that of $\mathcal{B}(H)$, we need a condition that replaces normality. The crucial point is that a normal representation of $\mathcal{B}(H)$ is determined completely by what it does to the rank-one operators. In particular, if the representation is nondegenerate, then already the action of the rank-one operators alone has to be nondegenerate. (For a unital representation of $\mathcal{B}(H)$, this nondegeneracy condition is equivalent to normality!) We will require that all unital endomorphisms ϑ_t of $\mathcal{B}^a(E)$ are nondegenerate in that sense, that is, we require that for all $t \in \mathbb{S}$ the set $\vartheta_t(EE^*)E$ is total in E. It can be shown that this is equivalent to say that the unital representation ϑ_t is ***strict***; see, for instance,[MSS06]

To begin with, suppose that E has a ***unit vector*** ξ, that is, $\langle\xi,\xi\rangle = \mathbf{1} \in \mathcal{B}$. That means, in particular, that $\mathcal{B}$ is unital and that E is full. We

knowing Hilbert modules; see Remark 2.3.) If we want to emphasize the unit vector ξ, we will say the Bhat system of ϑ based on ξ.

[d] Note that by associativity and the requirement that $u_{0,0}$ is the canonical identification, it follows that also v_0 is the canonical identification. Indeed, suppose u is the unique unitary in $\mathcal{B}(L)$ such that $v_0(x\otimes 1) = ux$. Then $ux = v_0(x\otimes 1) = v_0((v_0(u^*x\otimes 1))\otimes 1) = v_0(u^*x \otimes u_{0,0}(1\otimes 1)) = v_0(u^*x\otimes 1) = uu^*x = x$, so that $u = \mathsf{id}_L$.

showed in Skeide[Ske02] that, in this case, the whole construction of a product system *à la* Bhat *cum grano salis* goes through, as before. As in (2), we define Hilbert $\mathcal{B}$–submodules $E_t := \vartheta_t(\xi\xi^*)E$ of E. The *grano salis* we had to add in[Ske02] is the definition of a left action of $\mathcal{B}$ on E_t that turns it into a correspondence over $\mathcal{B}$. This left action is

$$bx_t \;:=\; \vartheta_t(\xi b\xi^*)x_t.$$

(Exercise: Check that this defines a unital representation of $\mathcal{B}$ by operators on E_t.) Once more, the (balanced $\mathbb{C}$–bilinear) mappings in (3) define isometries $v_t\colon E\odot E_t$ and $u_{s,t}\colon E_s\odot E_t\to E_{s+t}$. (We invite the reader to check that these maps, indeed, preserve inner products.) To see that v_t is surjective, simply observe that the elements of the total subset $\vartheta_t(EE^*)E$ can be written as

$$\vartheta_t(xy^*)z \;=\; \vartheta_t(x\xi^*\xi y^*)z \;=\; \vartheta_t(x\xi^*)\vartheta_t(\xi y^*)z \;=\; v_t(x\odot\vartheta_t(\xi y^*)z) \qquad (4)$$

where, clearly, $\vartheta_t(\xi y^*)z\in E_t$. (Exercise: Go back to the Hilbert space case and give a formal proof of surjectivity under the apparently weaker assumption of normality, modifying the preceding argument suitably.) Of course, also here $\vartheta^v_t(a) := v_t(a\odot\mathsf{id}_t)v_t^*$ gives back $\vartheta_t(a)$. Surjectivity of $u_{s,t}$ can be checked as in the Hilbert space case. And by

$$\begin{aligned} bu_{s,t}(x_s\odot y_t) \;&=\; \vartheta_{s+t}(\xi b\xi^*)\vartheta_t(x_s\xi^*)y_t \\ &=\; \vartheta_t(\vartheta_s(\xi b\xi^*)x_s\xi^*)y_t \;=\; \vartheta_t(bx_s\xi^*)y_t \;=\; u_{s,t}(bx_s\odot y_t) \end{aligned}$$

we see that the unitaries $u_{s,t}$ are even bilinear.

We summarize: The family $E^\odot = \big(E_t\big)_{t\in\mathbb{S}}$ with the unitaries $u_{s,t}\in\mathcal{B}^{a,bil}(E_s\odot E_t, E_{s+t})$ is an (algebraic) ***product system*** of correspondences over $\mathcal{B}$. That means, the product $(x_s,y_t)\mapsto x_sy_t := u_{s,t}(x_s\odot y_t)$ is associative, $E_0=\mathcal{B}$ and $u_{t,0}$ and $u_{0,t}$ are the canonical identifications. Moreover, the product system is ***full*** in the sense that each E_t is full, and the pair $(v^\odot, E)$ with $v^\odot = \big(v_t\big)_{t\in\mathbb{S}}$ is a ***left dilation*** of $E^\odot$ to E. By this we mean that the unitaries $v_t\in\mathcal{B}^a(E\odot E_t, E)$ iterate associatively with the product system structure ***and*** that E is full. The E_0–semigroup $\vartheta^v = \big(\vartheta^v_t\big)_{t\in\mathbb{S}}$ is the ϑ we started with.

2.2 Remark. Note that if $(v^\odot, L)$ is a left dilation, then $\mathcal{B}_L\subset\mathcal{B}_{E_t}$ so that full L implies that every E_t is full. The condition that L be full replaces the condition $L\neq\{0\}$ of nontriviality in the Hilbert space case. In fact, the only Hilbert space that is not a full Hilbert $\mathbb{C}$–module is $\{0\}$. For nonfull $E^\odot$ the concept of left dilation is not defined.

The idea of left dilation is that, if a product system $E^\odot$ gives rise to an E_0–semigroup ϑ^v via a left dilation $(v^\odot, L)$, then the E_0–semigroup should determine that product system uniquely. By this we mean, if we have another product system with a left dilation to the same L such that the induced E_0–semigroups coincide, then the two product systems should be isomorphic. For full L this is a special case of Proposition 2.4 below. If we would weaken to not necessarily full L, then uniqueness fails as soon as L is not full. (An extrem example would be $L = \{0\}$ to which every product system could be "dilated".) If we have a pair $(v^\odot, L)$ that fulfills all conditions of a left dilation except fullness of L, then we speak of a left ***quasi dilation***. Also a quasi dilation defines an E_0–semigroup ϑ^v on $\mathcal{B}^a(L)$.

Every product system $E^\odot$ of correspondences over $\mathcal{B}$ with a quasi dilation $v^\odot$ to L has a subsystem $F^\odot$ of full correspondences

$$F_t := \bigcap_{t_1+\ldots+t_n=t} \overline{\mathsf{span}}\big(\mathcal{B}_L E_{t_n} \mathcal{B}_L \ldots \mathcal{B}_L E_{t_1} \mathcal{B}_L\big)$$

over $\mathcal{B}_L$. It is easy to check (exercise!) that the restriction of the quasi dilation of $E^\odot$ to that subsystem $F^\odot$ is, now, a left dilation of $F^\odot$ to the full Hilbert $\mathcal{B}_L$–module L inducing the same E_0–semigroup on $\mathcal{B}^a(L)$. By Proposition 2.4, which holds also for nonunital $\mathcal{B}_L$, such a product system is determined uniquely by ϑ^v.

We owe the reader to say a few words about the construction of ***the*** unique product system of an E_0–semigroup in the general case. (The reader who is satisfied considering only the full unital case, may skip this and pass to Proposition 2.4, immediately.) Again this is nothing but the theory of (strict) representations of $\mathcal{B}^a(E)$, now in its most general form. The theory of unital strict representations ϑ of $\mathcal{B}^a(E)$ on another Hilbert module F over a possibly different C^*–algebra $\mathcal{C}$ ***and*** the theory of arbitrary representations on a von Neumann module have been settled in Muhly, Skeide and Solel.[MSS06] In the strict and unital case there is a correspondence F_ϑ from $\mathcal{B}$ to $\mathcal{C}$ such that $F \cong E \odot F_\vartheta$ and $\vartheta(a)$ is just amplification $a \odot \mathsf{id}_{F_\vartheta}$. In the not necessarily strict case, the representation on a von Neumann module decomposes into a strict unital part, and a part that annihilates the ***algebra of finite-rank operators*** $\mathcal{F}(E) := \mathsf{span}\, EE^*$ and, therefore, also the C^****–algebra of compact operators*** $\mathcal{K}(E) = \overline{\mathcal{F}(E)}$.

We repeat briefly what the construction asserts in the case of an E_0–semigroup ϑ on $\mathcal{B}^a(E)$ as discussed in Skeide.[Ske04] To begin with, we do not assume that the Hilbert $\mathcal{B}$–module E is full. For every $t \in \mathbb{S}$ we turn E into a correspondence ${}_tE$ from $\mathcal{B}^a(E)$ to $\mathcal{B}$ by defining the left

action $ax = \vartheta_t(a)x$. By the nondegeneracy condition we posed on ϑ_t, already the action of $\mathcal{F}(E) \subset \mathcal{B}^a(E)$ alone on ${}_tE$ is nondegenerate. In other words, we may also view ${}_tE$ as correspondence from $\mathcal{K}(E)$ to $\mathcal{B}$. We turn $E^* = \{x^* \colon x \in E\}$ into a correspondence from $\mathcal{B}$ to $\mathcal{B}^a(E)$ by defining the inner product $\langle x^*, y^* \rangle := xy^*$ and the bimodule action $bx^*a := (a^*xb^*)^*$. As $\mathcal{B}^a(E)_{E^*} = \mathcal{K}(E)$ and $\mathcal{B}_E E^*$ is total in E^* we may view E^* also as a full correspondence from $\mathcal{B}_E$ to $\mathcal{K}(E)$.

It is easy to verify that $E \odot E^* \cong \mathcal{K}(E)$ via $x \odot y^* \mapsto xy^*$ and $E^* \odot E \cong \mathcal{B}_E$ via $x^* \odot y \mapsto \langle x, y \rangle$, as correspondences over $\mathcal{K}(E)$ and over $\mathcal{B}_E$, respectively.[e] If we define the correspondence $E_t := E^* \odot {}_tE$ over $\mathcal{B}_E$, then

$$E \odot E_t \;=\; E \odot (E^* \odot {}_tE) \;\cong\; (E \odot E^*) \odot {}_tE \;\cong\; \mathcal{K}(E) \odot {}_tE \;\cong\; {}_tE$$

via

$$v_t \colon x \odot (y^* \odot_t z) \;\longmapsto\; \vartheta_t(xy^*)z,$$

where we denote the elementary tensor of elements $y^* \in E^*$ and $z \in {}_tE$ as $y^* \odot_t z$. Note that this is an isomorphism of correspondences from $\mathcal{B}^a(E)$ to $\mathcal{B}$ so that the canonical action $a \odot \mathsf{id}_t$ of $a \in \mathcal{B}^a(E)$ on the left-hand side corresponds to the canonical action $\vartheta_t(a)$ of a on the right-hand side. It is readily verified (exercise!) that

$$(x^* \odot_s y) \odot (x'^* \odot_t y') \;\longmapsto\; x^* \odot_{s+t} (\vartheta_t(yx'^*)y')$$

defines an (obviously, bilinear) unitary $u_{s,t} \colon E_s \odot E_t \to E_{s+t}$ and that this product is associative. In other words, $E^\odot = \big(E_t\big)_{t\in\mathbb{S}}$ is a product system of $\mathcal{B}_E$–correspondences and $v^\odot = \big(v_t\big)_{t\in\mathbb{S}}$ is a left dilation of $E^\odot$ to E giving back ϑ as ϑ^v. If we want to have a concise construction that works for all E_0–semigroups, then we speak about this $E^\odot$ as ***the*** product system associated with ϑ.

[e] Effectively, as $\mathcal{K}(E)$–$\mathcal{B}_E$–correspondence, E is a Morita equivalence from $\mathcal{K}(E)$ to $\mathcal{B}_E$ and E^* its *inverse* under tensor product. In general, what we nowadays call a ***Morita equivalence*** from $\mathcal{A}$ to $\mathcal{B}$, is a full correspondence F from $\mathcal{A}$ to $\mathcal{B}$ for which the canonical homomorphism $\mathcal{A} \to \mathcal{B}^a(F)$ defines an isomorphism onto $\mathcal{K}(F)$. (Rieffel,[Rie74a] who introduced the concept, called F an *imprimitivity bimodule*.) With this isomorphism the $\mathcal{B}$–$\mathcal{K}(F)$–correspondence F^* can be viewed as $\mathcal{B}$–$\mathcal{A}$–correspondence. As $\mathcal{B}$ is canonically isomorphic to $\mathcal{K}(F^*)$, also F^* is a Morita equivalence. Almost all what follows, essentially noting that tensoring with a Morita equivalence may be undone by tensoring with its inverse, was already known to Rieffel. What we added to his *imprimitivity theorem* [Rie74a, Theorem 6.23], essentially the representation theory of $\mathcal{F}(E)$ on a Hilbert space, is the extension to $\mathcal{B}^a(E)$ and that the representation space may be a Hilbert module.

2.3 Remark. There is a price to be paid, for that this construction works for all E_0–semigroups. The members $E_t = E^* \odot_t E$ are abstract tensor products, while in every other construction, also Arveson's for Hilbert spaces, the E_t are subspaces of one fixed Banach space (of E in the construction *à la* Bhat with a unit vector and of $\mathcal{B}(H)$ in Arveson's construction; see Section 5).

Also, the proof is very elegant and simple. But, unlike the other proofs, even in the Hilbert space case it requires some basic knowledge of Hilbert modules. (H^* is a correspondence from $\mathbb{C}$ to $\mathcal{K}(H)$ and we have to calculate tensor products with this correspondence.) We recommend as an intriguing exercise to redo the theory of normal representation of $\mathcal{B}(H)$ along the preceding proof. See [Ske05a, Remark 2.2] and [MSS06, Example 1.5].

After the preceding discussion of the general nonunital and even nonfull case, we will now concentrate on full product systems. What happens if we have two left dilations $(v^{1^\odot}, L^1)$ and $(v^{2^\odot}, L^2)$ of a full product system $E^\odot$? In the case of Hilbert spaces Arveson's answer (in terms of left dilation) is, the two E_0–semigroups ϑ^{v_1} and ϑ^{v2} must be cocycle conjugate. However, this statement relies on the fact that Arveson's E_0–semigroups all act on a $\mathcal{B}(H)$ where H is infinite-dimensional and separable. In other words, the Hilbert spaces L_1 and L_2 have the same dimension and, therefore, are isomorphic. The general case of Hilbert modules, is an (interesting) open problem.[f] However, if the Hilbert modules L^1 and L^2 are isomorphic, then we have the same result as Arveson.[g] In Skeide[Ske02] we discussed the case with unit vectors. Here we state a slightly more general result directly in terms of left dilations.

2.4 Proposition. *Let L be a Hilbert $\mathcal{B}$–module. Then for two left dilations $(v^{i^\odot}, L)$ of two full product systems $E^{i^\odot}$ $(i = 1, 2)$ to L the product systems are* ***isomorphic*** *(that is, there is a family $u_t \colon E^1_t \to E^2_t$ of bilinear unitaries fulfilling $u_{s+t}(x_s y_t) = u_s(x_s)u_t(y_t)$ and $u_0 = \mathsf{id}_{\mathcal{B}}$), if and only if the*

[f] Left (quasi) dilations of the same product system admit direct sums. We expect that it is possible to develop a decomposition theory for left dilations in terms of smallest building blocks.

[g] We have even more: Suppose $\mathcal{B}^a(L^1)$ and $\mathcal{B}^a(L^2)$ are strictly isomorphic, so that there is a Morita equivalence M such that $L^2 = L^1 \odot M$; see Footnote e. Then two E_0–semigroups ϑ^{v^i} on $\mathcal{B}^a(L^i)$ $(i = 1, 2)$ are cocycle conjugate (in the obvious way), if and only if their product systems $E^{i^\odot}$ are ***Morita equivalent*** via the same Morita equivalence M, that is, there is an isomorphism between the product systems $E^{1^\odot}$ and $M \odot E^{2^\odot} \odot M^* := \big(M \odot E^2_t \odot M^*\big)_{t\in\mathbb{S}}$. For that, L^1 and L^2 need not even be modules over the same C^*–algebra. See Skeide[Ske04] for details.

*E_0–semigroups ϑ^{v^1} and ϑ^{v^2} are **cocycle conjugate** (that is, there is a family $\mathfrak{u}_t \in \mathcal{B}^a(L)$ of unitaries with $\mathfrak{u}_0 = \mathsf{id}_L$ such that $\vartheta^{v^1}_{s+t}(\mathfrak{u}_{s+t}) = \mathfrak{u}_s\vartheta^{v^1}_s(\mathfrak{u}_t)$ and $\vartheta^{v^2}_t(a) = \mathfrak{u}_t\vartheta^{v^1}_t(a)\mathfrak{u}_t^*$).*

Proof. (Sketch.) If there is a family u_t, then $\mathfrak{u}_t := v^2_t(\mathsf{id}_L \odot u_t)v^{1\,*}_t$ fulfills the desired properties. (Exercise!) For the backward direction recall that L is necessarily full. Therefore, $E^i_t = L^* \odot L \odot E^i_t$ as correspondences over $\mathcal{B}$. Moreover, $L \odot E^i_t$ is isomorphic to ${}_tL^i$ via v^i_t, if we define ${}_tL^i$ as L when viewed as correspondence from $\mathcal{B}^a(L)$ to $\mathcal{B}$ via ϑ^i_t. Clearly, $\mathfrak{u}_t\colon L \to L$ is an isomorphism of correspondences when viewed as mapping ${}_tL^1 \to {}_tL^2$. In other words, $u_t := \mathsf{id}_{L^*} \odot v^{2\,*}_t\mathfrak{u}_t v^1_t$ is an isomorphism $E^1_t = L^* \odot L \odot E^1_t \to L^* \odot L \odot E^2_t = E^2_t$. These u_t form an isomorphism of product systems. (Exercise!) ■

We summarize: Every E_0–semigroup ϑ on $\mathcal{B}^a(E)$ leads to a product system $E^\odot$. If E is full, then so is $E^\odot$ and $E^\odot$ is related to ϑ via a left dilation $v^\odot$ to E giving back ϑ as ϑ^v. Every other left dilation of $E^\odot$ to E leads to an E_0–semigroup cocycle conjugate to ϑ and two E_0–semigroups on $\mathcal{B}^a(E)$ have isomorphic product systems, if and only if they are cocycle conjugate.

If E is not full, then we may still associate with an E_0–semigroup ϑ on $\mathcal{B}^a(E)$ a product system $E^\odot$. This product system consists of full correspondences over $\mathcal{B}_E$. So if we simply restrict to $\mathcal{B}_E$, then we are in the full situation. There is no problem to consider E_t as a (no longer full) correspondence over $\mathcal{B}$. However, if we insist in having a product system of correspondences over $\mathcal{B}$, then we must replace $E_0 = \mathcal{B}_E$ with $\mathcal{B}$. This causes a sharp discontinuity at $t = 0$. Such a product system with $\mathcal{B}_{E_t} = \mathcal{B}_E \neq \mathcal{B}$ for $t > 0$ will never be continuous in the sense of Definition 4.2 below. Product systems where $\mathcal{B}_{E_t}$ increases in continuous way to $\mathcal{B}$ may have interesting left quasi dilations if there is a nontrivial subalgebra $\mathcal{C}$ of $\mathcal{B}$ such that $\mathcal{B}_{E_t} \supset \mathcal{C}$ for all t. But the investigation of quasi dilations, so far, has not yet been tackled.

Even if $\bigcap_{t\in\mathbb{S}} \mathcal{B}_{E_t} = \{0\}$ we obtain interesting structures, if we weaken, in the definition of left quasi dilation, unitarity of v_t to isometry.

2.5 Example. Put $E = \mathcal{B} = C_0(0,\infty)$, define the Hilbert submodules $E_t = C_0(t,\infty)$ of $\mathcal{B}$, and let $\mathfrak{s}_t$ denote the usual *right shift*. We turn E_t into a correspondence over $\mathcal{B}$ by defining the left action $b.x_t := \mathfrak{s}_t(b)x_t$. We leave it as an exercise to check that $v_t(x \odot y_t) = \mathfrak{s}_t(x)y_t$ defines an isometry $E\odot E_t \to E$, and that the restriction $u_{s,t}$ to $E_s\odot E_t$ defines a bilinear unitary

onto E_{s+t}. Clearly, the $u_{s,t}$ turn $E^{\odot} = \big(E_t\big)_{t\in\mathbb{S}}$ into a product system and the v_t iterate associatively with the product system structure. Note that v_t is not adjointable, so it is not possible to define an E–***semigroup*** (that is, a semigroup of not necessarily unital endomorphism) ϑ^v on $\mathcal{B}^a(E) = C_b(0,\infty)$, the multiplier algebra $M(\mathcal{B})$ of $\mathcal{B}$. (In fact, such a semigroup should have to send the identity of $\mathcal{B}^a(E)$ to the indicator function $\mathbb{1}_{[t,\infty)}$ which is not in $C_b(0,\infty)$.) But, s_t does define an E–semigroup on $\mathcal{K}(E) = \mathcal{B}$, and $E^{\odot}$ may be considered as the product system of that E–semigroup. The difficulty disappears for von Neumann modules; see Bhat and Lindsay[BL05] and Skeide[Ske04,Ske06e] for the obvious generalizations from E_0–semigroups to E–semigroups.

The question whether to every full product system $E^{\odot}$ there exists a left dilation and, therefore, an E_0–semigroup that has $E^{\odot}$ as associated product system is the subject of the following two sections.

3. Arveson systems and E_0–semigroups

One of the most important results about Arveson systems is that every Arveson system is the Arveson system associated with an E_0–semigroup; see Section 5 for the terminology we use here. Therefore we refer to this result as the *fundamental theorem* about Arveson systems. By Observation 3.1, below, this is equivalent to say that every Arveson system is the Bhat system associated with an E_0–semigroup as described in Section 2 or, by Proposition 2.1, to say that every Arveson system admits a left dilation.

Arveson showed the fundamental theorem in the last of the four articles.[Arv89a,Arv90a,Arv89b,Arv90b] After laying the basis in,[Arv89a] in[Arv90a] he introduced the *spectral C^*–algebra* of an Arveson system, that is, essentially, the C^*–algebra generated by the representing operators of a sufficiently faithful representation (see below) of the Arveson system. A universal property asserts that representations of the Arveson system are in one-to-one correspondence with representations of the spectral algebra. In[Arv89b] he analyzed that sort of representations, the singular representations, that do not lead to E_0–semigroups. (In the discrete case, a singular representation corresponds to the defining representation of a Cuntz algebra on the full Fock space.) In[Arv90b] he used the precise knowledge of the singular representations to construct an essential representation to every Arveson system, and this is equivalent to constructing an E_0–semigroup (Proposition 5.1). Independently of proving the fundamental theorem, the spectral algebra and the deep analysis of its representations in[Arv90a,Arv89b,Arv90b] is interesting

in its own right and has been subject to intense research, among others, by Zacharias and Hirshberg.[Zac00a,Zac00b,HZ03,Hir04,Hir05b,Hir05a] We shall not discuss the spectral algebra.

In the meantime, there is a proof due to Liebscher,[Lie03] which is similarly involved as Arveson's. (See also Footnote j.) Since Skeide[Ske06a] we have a proof of the fundamental theorem that fits into a few pages. Shortly after, Arveson[Arv06] presented a proof which frees the construction in[Ske06a] from a not actually difficult but quite tedious verification of the associativity condition. In Skeide[Ske06d] it is shown that the result of Arveson's construction[Arv06] is, indeed, unitarily equivalent to (a special case of) the construction in.[Ske06a] While Arveson's approach[Arv90a,Arv89b,Arv90b] via the spectral algebra definitely is not applicable to the case of Hilbert modules, the proof(s) in[Ske06a,Arv06,Ske06d] generalize in a (more or less) straightforward way to Hilbert modules; see[Ske06c] (E_0–semigroups) and[Ske06f] (essential representations) for Hilbert modules and[Ske07] (in preparation) for von Neumann modules.

In this section we discuss the case of Hilbert spaces. The versions we have, so far, for modules we discuss in Section 4. We should like to say that we will describe the construction of a left dilation, because it is this construction which generalizes to Hilbert modules. What Arveson constructs (be it in[Arv89b] or in[Arv06]) is an essential representation or, what is the same, a *right* dilation; see the end of this section. For Hilbert spaces these concepts may be easily translated into each other; see Observation 3.1. (In fact, in[Ske06d] we translated[Ske06a] into a right dilation in order to compare with.[Arv06] Here we will proceed the other way round.)

The hard problem is the continuous time case $\mathbb{S} = \mathbb{R}_+$ in absence of so-called *units*. A ***unit*** for an Arveson system $E^\otimes$ is a family $\xi^\otimes = \big(\xi_t\big)_{t\in\mathbb{S}}$ of elements $\xi_t \in E_t$ that fulfills

$$\xi_s\xi_t \;=\; \xi_{s+t} \tag{5}$$

and $\xi_0 = 1$. Arveson excludes the ***trivial*** case where $\xi_t = 0$ for all $t > 0$. We do not want to exclude it at all as a possibility. Nevertheless, in these notes we shall assume tacitly that a unit is nontrivial.[h] We say a unit $\xi^\otimes$ is ***unital***, if $\langle\xi_t, \xi_t\rangle = 1$ for all $t \in \mathbb{S}$.

If $E^\otimes$ has a unital unit $\xi^\otimes$, then already Arveson [Arv89a, appendix]

[h] If we speak about *continuous* units, as almost everywhere in the Hilbert modules case, then nontriviality is automatic. Instead, it is a well-known obstacle in semigroup theory that just measurability is not enough.

constructed a right dilation by an inductive limit. We discuss here the version for left dilations from[BS00,BBLS04] that will work also for Hilbert modules. For every $s, t \in \mathbb{S}$ we define an isometric embedding $E_t \to E_{s+t}$ by $x_t \mapsto \xi_s x_t$. The E_t together with these embeddings form an inductive system. We denote by L its inductive limit. The factorization $u_{s,t}\colon E_s \otimes E_t \to E_{s+t}$ under the limit $s \to \infty$ gives rise to a factorization $v_t\colon L \otimes E_t \to L$. Associativity of the product system structure $u_{s,t}$ guarantees that the v_t form a left dilation of $E^{\otimes}$. All the $\xi_t \in E_t$ in the inductive limit appear as the same vector ξ and $v_t(\xi \otimes \xi_t) = \xi$. We leave it as an exercise to show that the Bhat system associated with the E_0–semigroup ϑ^v via ξ is $E^{\otimes}$.

The problem is that, in the continuous time case, there are loads of product systems without units.[i] However, it is always possible to find a unit for a product system in the discrete case $\mathbb{S} = \mathbb{N}_0$. Simply take any unit vector $\xi_1 \in E_1$ and put $\xi_n = \xi_1^n$. Then $\xi^{\otimes} = \big(\xi_n\big)_{n\in\mathbb{N}_0}$ is a unital unit for $E^{\otimes}$. Existence of left dilations for discrete Arveson systems is the starting point of the construction in Skeide.[Ske06a]

So let $E^{\otimes} = \big(E_t\big)_{t\in\mathbb{R}_+}$ be an Arveson system. Suppose we have a left dilation $\breve{v}_n$ of the discrete subsystem $\big(E_n\big)_{n\in\mathbb{N}_0}$ of $E^{\otimes}$ to $\breve{L}$. (This can be the preceding inductive limit construction based on a unit vector $\xi_1 \in E_1$, but it need not.) We try now to "lift" this left dilation of the discrete subsystem to a left dilation of the whole system.[j] To that goal we consider the direct integrals $\int_a^b E_\alpha\, d\alpha$ $(0 \le a < b \le \infty)$. We put $L := \breve{L} \otimes \int_0^1 E_\alpha\, d\alpha$. For $t \in \mathbb{R}_+$ we define $n := [t]$, the unique integer such that $t - n \in [0, 1)$. Then

[i] Powers[Pow87] showed existence of nonspatial E_0–semigroups on $\mathcal{B}(H)$ by rather indirect means. And nonspatiality is equivalent to that the associated Arveson (or Bhat) system is unitless. The first constructive examples are due to Tsirelson.[Tsi00b] Bhat and Srinivasan[BS05] started a systematic investigation of Tsirelson's probabilistic constructions in a more functional analytic way, and discovered a large class of examples.

[j] This very similar to Riesz' proof of Stone's theorem; see Riesz and Sz.-Nagy.[RSN82] But, there are also similarities to Liebscher's proof in.[Lie03] However, the E_0–semigroup Liebscher constructs is *pure*, while ours is definitely nonpure.

the following identifications

$$\begin{aligned} L \otimes E_t &= \breve{L} \otimes \left(\int_0^1 E_\alpha \, d\alpha \right) \otimes E_t \cong \breve{L} \otimes \int_t^{1+t} E_\alpha \, d\alpha \\ &\cong \left(\breve{L} \otimes E_n \otimes \int_{t-n}^1 E_\alpha \, d\alpha \right) \oplus \left(\breve{L} \otimes E_{n+1} \otimes \int_0^{t-n} E_\alpha \, d\alpha \right) \\ &\cong \left(\breve{L} \otimes \int_{t-n}^1 E_\alpha \, d\alpha \right) \oplus \left(\breve{L} \otimes \int_0^{t-n} E_\alpha \, d\alpha \right) = L \end{aligned} \tag{6}$$

define a unitary $v_t \colon L \otimes E_t \to L$. In the step from the second line to the third one we have made use of the identifications $\breve{v}_n \colon \breve{L} \otimes E_n \to \breve{L}$ and $\breve{v}_{n+1} \colon \breve{L} \otimes E_{n+1} \to \breve{L}$ coming from the dilation of $\big(E_n\big)_{n\in\mathbb{N}_0}$. Existence of the dilation of the discrete subsystem means that $\breve{L}$ absorbs every tensor power of E_1. Just that how many factors E_1 have to be absorbed depends on whether $\alpha + t - n$ is bigger or smaller than 1.

The identifications in (6) suggest operations that act directly on sections $\big(\breve{x} \otimes y_\alpha\big)_{\alpha\in[0,1)}$ and give as result again a section in $\big(\breve{L} \otimes E_\alpha\big)_{\alpha\in[0,1)}$. It is a tedious but straightforward verification (one page in[Ske06a]) directly on sections that these identifications iterate associatively. But a word need to be said about the direct integrals, because these do no longer make sense without technical conditions.[k] The technical conditions on an Arveson system are such that $\big(E_t\big)_{t>0}$ is isomorphic to $(0,\infty) \times H_0$ as a Borel bundle, where H_0 is an infinite-dimensional separable Hilbert space such that $E_t \cong H_0$, and such that the product $(x_t, y_s) \mapsto x_t y_s$ is measurable. We will have to speak more about good choices of technical conditions in the module case. Here it is enough to know that the direct integrals gain a sense as $\int_a^b E_\alpha \, d\alpha = L^2([a,b), H_0)$. The fundamental theorem, in its precise formulation, involves the statement that under the technical conditions on an Arveson system, the constructed E_0–semigroup is strongly continuous in time. The self-contained proof in[Ske06a] is done by involving Observation 3.2. (More precisely, we construct in the same way also a right dilation of $E^\otimes$ to R, and show that the unitary group u_t on $L \otimes R$ defined in Observation 3.2 is weakly measurable, hence, by separability, strongly continuous.

[k]The direct integrals make sense immediately, if for α we choose the counting measure, that is, as direct sums. In this case, the E_0–semigroup we obtain from the left dilation has no chance to satisfy any reasonable continuity condition, because the left dilation involves a shift of sections. (Also the time shift on $\ell^2[0,1)$ is not weakly continuous.) Anyway: Every algebraic Arveson system admits a left dilation, though not always continuous.

And, this turns over to the E_0–semigroup ϑ.

Arveson[Arv06] was able to simplify considerably the proof of associativity in the special case, when the left dilation of the discrete subsystem $\big(E_n\big)_{n\in\mathbb{N}_0}$ is obtained as an inductive limit over the unit $\big(\xi_1^n\big)_{n\in\mathbb{N}_0}$ obtained from a unit vector $\xi_1 \in E_1$. Starting with that unit vector, Arveson defines the space of ***stable sections***, that is, of all locally square integrable sections $\big(y_\alpha\big)_{\alpha\in\mathbb{R}_+}$ that fulfill

$$\xi_1 y_\alpha \;=\; y_{\alpha+1} \tag{7}$$

for all sufficiently big α. It is not difficult to show that for two such sections y and z the expression $\int_a^{a+1}\langle y_\alpha, z_\alpha\rangle\, d\alpha$ is eventually constant, so that

$$\langle y, z\rangle \;:=\; \lim_{a\to\infty}\int_a^{a+1}\langle y_\alpha, z_\alpha\rangle\, d\alpha$$

defines a semiinner product on the space of stable sections. The kernel of this semiinner product consists of those stable sections which are eventually 0 almost everywhere. The quotient is a Hilbert space. In[Ske06d] we have shown that this Hilbert space is canonically isomorphic to L.[l] For a stable section y and an element $x_t \in E_t$ Arveson defines the stable section yx_t by setting

$$(yx_t)_\alpha \;=\; \begin{cases} y_{\alpha-t}x_t & t \geq \alpha, \\ 0 & \text{else.} \end{cases}$$

Then $y\otimes x_t \mapsto yx_t$ defines an isometry that, in the picture L, coincides with v_t (from which also surjectivity is immediate). The advantage of Arveson's approach[Arv06] is that associativity is immediate and that it gives an interpretation of the inductive limit in very concrete terms. The construction in[Ske06a] is slightly more general. It starts from the well-known observation that it is easy to obtain a left dilation for the discrete subsystem and the basic idea in (6) how to transform that dilation into a dilation of the whole system.

We close this section with some explanations about left and right dilations. Whenever for an Arveson system $E^\otimes$ we have a Hilbert space $R \neq \{0\}$ and a family $w^\otimes$ of unitaries $w_t\colon E_t\otimes R \to R$ that iterates associatively with the product system structure, we call the pair $(w^\otimes, R)$ a ***right dilation*** $w^\otimes$ of $E^\otimes$ to R.[m]

[l] Roughly speaking, in the construction of $\check{L}\otimes\int_0^1 E_\alpha\, d\alpha$ one has to interchange inductive limit and direct integral. In other words, one considers an inductive limit over $E_n \otimes \int_0^1 E_\alpha\, d\alpha = \int_n^{n+1} E_\alpha\, d\alpha$. This space corresponds to the subspace of stable sections that satisfy (7) for all $\alpha \geq n$.

[m] Also here it is automatic that w_0 is the canonical identification; cf. Footnote d.

A ***representation*** of an Arveson system $E^\otimes$ on a Hilbert space R is a family of maps $\eta_t\colon E_t \to \mathcal{B}(R)$ such that

$$\eta_t(x_t)^*\eta_t(y_t) \;=\; \langle x_t, y_t\rangle\,\mathsf{id}_R, \qquad \eta_t(x_t)\eta_s(y_s) \;=\; \eta_{t+s}(x_t y_s).$$

A representation is ***nondegenerate*** (or essential in[Arv89a]), if each η_t is nondegenerate, that is, if $\overline{\mathsf{span}}\,\eta_t(E_t)R = R$ for all $t \in \mathbb{S}$.

If we have a representation, then it is easy to check that $w_t\colon x_t \otimes y \mapsto \eta_t(x_t)y$ defines an isometry $w_t\colon E_t \otimes R \to R$. These isometries iterate associatively with the product system structure. If the representation is nondegenerate, then the w_t form a right dilation. Conversely, if we have a right dilation w_t, then $\eta_t(x_t)\colon y \mapsto w_t(x_t \otimes y)$ defines a nondegenerate representation. Therefore, it is the same to speak about a right dilation or a nondegenerate representation. In[Arv89a,Arv90a,Arv89b,Arv90b] Arveson showed his fundamental theorem by establishing existence of a representation for every Arveson system $E^\otimes$.

3.1 Observation. For an Arveson system $E^\otimes$ we define the ***opposite*** Arveson system $E'^\otimes$ as the same family of Hilbert spaces (with the same measurable structure) but opposite multiplication $(x_s, y_t) \mapsto y_t x_s$. We may transform a left dilation $(v^\otimes, L)$ of $E^\otimes$ into a right dilation $(w'^\otimes, L)$ of the opposite system and a right dilation $(w^\otimes, R)$ of $E^\otimes$ into a left dilation $(v'^\otimes, L)$ of the opposite system. We simply have to reverse in all tensor products the order of the factors, that is, we put $w'_t(x_t \otimes y) := v_t(y \otimes x_t)$ and $v'_t(x \otimes y_t) := w_t(y_t \otimes x)$.

For correspondences the operation$(x_s, y_t) \mapsto y_t x_s$ will rarely define an isometry. In general, there is no opposite system for product systems of correspondences. However, for product systems of von Neumann correspondences there is the commutant, and the commutant of an Arveson system coincides with its opposite system; see Example 6.4.

3.2 Observation. Suppose $(v^\otimes, L)$ and $(w^\otimes, R)$ are a left and a right dilation, respectively, of $E^\otimes$. Then $u_t := (v_t \otimes \mathsf{id}_R)(\mathsf{id}_L \otimes w_t^*)$ defines a unitary group in $\mathcal{B}(L \otimes R)$. Moreover, the automorphism group $\alpha_t := u_t \bullet u_t^*$ leaves invariant $\mathcal{B}(L) \otimes \mathsf{id}_R$ for $t \geq 0$ and $\mathsf{id}_L \otimes \mathcal{B}(R)$ for $t \leq 0$. Then the restriction α_t $(t \geq 0)$ to $\mathcal{B}(L) \otimes \mathsf{id}_R \cong \mathcal{B}(L)$ defines an E_0–semigroup ϑ on $\mathcal{B}(L)$ which has $E^\otimes$ as associated Bhat system, while the restriction α_t $(t \leq 0)$ to $\mathsf{id}_L \otimes \mathcal{B}(R) \cong \mathcal{B}(R)$ defines an E_0–semigroup θ on $\mathcal{B}(R)$ which has $E'^\otimes$ as associated Bhat system. Canceling from the last phrase all statements about product systems, this is exactly the situation when Powers and Robinson[PR89] say ϑ and θ are ***paired***.

Note that the E_0–semigroup θ on $\mathcal{B}(R)$ has $E^{\otimes}$ as associated Arveson system in the sense of[Arv89a] and ϑ has $E'^{\otimes}$ as associated Arveson system. We explain this in Section 5.

4. Continuous product systems and E_0–semigroups

Let $E^{\odot} = \big(E_t\big)_{t\in\mathbb{S}}$ be a product system of correspondences over a unital C^*–algebra $\mathcal{B}$. Like in the Hilbert space case, a ***unit*** for $E^{\odot}$ is a family $\xi^{\odot} = \big(\xi_t\big)_{t\in\mathbb{S}}$ of elements $\xi_t \in E_t$ fulfilling (5) and $\xi_0 = \mathbf{1}$. We ***do not*** define what a unit is, if $\mathcal{B}$ is nonunital! The unit is ***unital***, if $\langle \xi_t, \xi_t\rangle = \mathbf{1}$ for all $t \in \mathbb{S}$. The construction of an E_0–semigroup from a unital unit works as for Hilbert spaces: We define isometric embeddings $E_t \to E_{s+t}$ by $x_t \mapsto \xi_s x_t$. Then, the inductive limit L factors as $v_t\colon L \odot E_t \to L$ via a left dilation $v^{\odot}$.

4.1 Remark. For Hilbert spaces the isometric embedding could be defined as $x_t \mapsto x_t\xi_t$, leading to a right dilation. In fact, this is what Arveson did in the appendix of[Arv89a] and what we did in[Ske06d] in the discrete case when we compared Arveson's construction of a right dilation in[Arv06] with.[Ske06a] From the beginning, such a construction of a right dilation can not be done for Hilbert modules: $x_t \mapsto x_t\xi_t$ is, in general, not an isometry. More precisely, it is an isometry, if and only if the unital unit $\xi^{\odot}$ is ***central***, that is, if $b\xi_t = \xi_t b$ for all $t \in \mathbb{S}, b \in \mathcal{B}$. Product systems that admit a central unital unit are classified as ***spatial*** (Skeide[Ske06g]) and admit classification results parallel to those for spatial Arveson systems. Product systems of von Neumann modules that admit (continuous) units are spatial automatically (Barreto, Bhat, Liebscher and Skeide[BBLS04]), while for Hilbert modules existence of a continuous unit is not enough to guarantee spatiality. We see, the problem of constructing an E_0–semigroup is difficult only for nonspatial product systems.

If in a discrete product system $E^{\odot} = \big(E_n\big)_{n\in\mathbb{N}_0}$ the member E_1 has a unit vector, then we may construct a left dilation of that discrete product system. However, there are discrete product systems where no member except E_0 has a unit vector. It is one of the main results of Skeide[Ske04] to show that every *full* discrete product system of correspondences over a unital C^*–algebra admits a left dilation; see Footnote p.

Now let us discuss the continuous time case. Let $E^{\odot} = \big(E_t\big)_{t\in\mathbb{R}_+}$ be a full product system of correspondences over a unital C^*–algebra $\mathcal{B}$. By[Ske04] we may choose a left dilation $(\check{v}^{\odot}, \check{L})$ of the discrete subsystem $\big(E_t\big)_{t\in\mathbb{N}_0}$. We

would like to proceed as in Section 3, defining the direct integrals $\int_a^b E_\alpha \, d\alpha$ and $L = \breve{L} \odot \int_0^1 E_\alpha \, d\alpha$ so that (6) had a chance to define a left dilation. Without posing precise technical conditions, this works only for direct sums, leading to noncontinuous E_0–semigroups; cf. Footnote k.

To motivate the technical conditions we will pose, we start from what, in the end, we wish to have. Given a full product system $E^\odot = \big(E_t\big)_{t\in\mathbb{R}_+}$ of correspondences over a unital C^*–algebra, we wish to have a full Hilbert module L and a left dilation $v_t\colon L\odot E_t \to L$ such that the E_0–semigroup ϑ^v on $\mathcal{B}^a(L)$ is ***strongly continuous***, that is, for all $a \in \mathcal{B}^a(L)$ and all $x \in L$, the function $t \mapsto \vartheta^v_t(a)x$ is continuous. (As usual with semigroups, it is sufficient to require strong continuity around 0. Since the ϑ^v_t are bounded uniformly, it is also sufficient to check continuity for a and x from total subsets of $\mathcal{B}^a(L)$ and L, respectively.)

So suppose we have a strongly continuous E_0–semigroup ϑ on $\mathcal{B}^a(E)$ where E is a full Hilbert module over a unital C^*–algebra $\mathcal{B}$. What can we say about the associated product system? How is continuity of ϑ reflected by the *bundle structure* of the product system? As a first step, we have to fix a version of the product system and of the left dilation relating it to ϑ. We have to take into account that the structures we derive might depend on that choice. To have a start, let us suppose that E has a unit vector ξ and construct product system $E^\odot$ and left dilation v_t of that product system from the unit vector ξ *à la* Bhat. The essential observation is that in this approach all E_t are identified as submodules of E. Moreover, for every $x \in E$ the function $t \mapsto \vartheta_t(\xi\xi^*)x \in E_t \subset E$ is continuous. Of course, if $x = y_t \in E_t$, then the section $t \mapsto x_t := \vartheta_t(\xi\xi^*)x$ assumes the value $x_t = y_t$ at t. It is not difficult to check that, whenever we have two sections x, y of $E^\odot$ such that the functions $t \mapsto x_t \in E_t \subset E$ and $t \mapsto y_t \in E_t \subset E$ are continuous, then also the function $(s,t) \mapsto x_s y_t \in E_{s+t} \subset E$ is continuous; see Skeide.[Ske03b] This motivates the following definition from.[Ske03b,Ske06c]

4.2 Definition. Let $E^\odot = \big(E_t\big)_{t\in\mathbb{R}_+}$ be a product system of correspondences over a C^*–algebra $\mathcal{B}$ with a family $i = \big(i_t\big)_{t\in\mathbb{R}_+}$ of isometric embeddings $i_t\colon E_t \to \widehat{E}$ into a Hilbert $\mathcal{B}$–module $\widehat{E}$. Denote by

$$CS_i(E^\odot) \;=\; \Big\{x = \big(x_t\big)_{t\in\mathbb{R}_+} : x_t \in E_t, t \mapsto i_t x_t \text{ is continuous}\Big\}$$

the set of ***continuous sections*** of $E^\odot$ (with respect to i). We say $E^\odot$ is ***continuous*** (with respect to i), if the following conditions are satisfied.

(1) For every $y_t \in E_t$ we can find a continuous section $x \in CS_i(E^\odot)$ such that $x_t = y_t$.
(2) For every pair $x, y \in CS_i(E^\odot)$ of continuous sections the function

$$(s,t) \longmapsto i_{s+t}(x_s y_t)$$

is continuous.

We say two embeddings i and i' have the same ***continuous structure***, if $CS_i(E^\odot) = CS_{i'}(E^\odot)$.

4.3 Remark. The definition says, roughly speaking, that $E^\odot$ is a subbundle of the trivial Banach bundle $[0,\infty) \times \widehat{E}$. Note that this is even weaker than Arveson's requirement that the part $t > 0$ of an Arveson system be *Borel isomorphic* to the trivial bundle $(0,\infty) \times H_0$. Of course, this bundle is also a Banach bundle, and the condition just means that the Borel structure of an Arveson system is that induced from the continuous structure of a trivial Banach bundle. We allow even for subbundles.

The only difference between a *continuous* Arveson system (that is, an algebraic Arveson system that is continuous in the sense of Definition 4.2) and a *measurable* Arveson system (that is, an Arveson system in the sense of Arveson[Arv89a]) consists in whether multiplication is required continuous or just measurable. (Note that in[Ske06a] we used Arveson's condition, that is, just measurable multiplication.) If, for a general product system, we pose just measurability as assumption, then the construction we describe in the sequel should provide us with a (weakly) measurable E_0–semigroup. Only under separability assumptions this will be enough to show strong continuity. We did not yet put into practice a measurable version.

In Definition 4.2 we do not require that $E^\odot$ is full, nor that $\mathcal{B}$ is unital. However, if $\mathcal{B}$ is unital, then we have the following lemma from.[Ske06a]

4.4 Lemma. *If $\mathcal{B}$ is unital, then a continuous product system $E^\odot$ of correspondences over $\mathcal{B}$ contains a continuous section $\zeta \in CS_i(E^\odot)$ that consists entirely of unit vectors and fulfills $\zeta_0 = \mathbf{1}$. In particular, every E_t contains a unit vector (and, therefore, is full).*

The proof relies on the fact that the invertible elements of $\mathcal{B}$ form an open subset (so that there is a continuous section that consists of unit vectors at least for all sufficiently small t), and on the fact that the tensor product of unit vectors is again a unit vector (so that small pieces of that section can be used to compose a global continuous section of unit vectors).

We see that not only the continuous product system is full automatically[n], but that all of its members contain a unit vector. That is, we may not only use a unit vector in E_1 to construct a left dilation of the discrete subsystem, but we may even adapt Arveson's construction of an E_0–semigroup word by word as described in Section 3.

4.5 Remark. Note that, like in the construction of the dilation of the discrete subsystem, also for imitating Arveson's proof it is indispensable that we define the space of *stable* sections as in (7), by multiplying the unit vector from the left. Multiplying from the right (as in[Arv06]) will cause that the inner product $\int_a^{a+1}\langle y_\alpha, z_\alpha\rangle\, d\alpha$ is no longer eventually constant.

Arveson's proof in,[Arv06] once understood the idea, is easier to carry out than our proof in.[Ske06a] This is, why in[Ske06c] we followed Arveson's road to prove the *fundamental theorem* for continuous product systems of correspondences over a unital C^*–algebra. The result:

4.6 Theorem. *Every continuous product system of correspondences over a unital C^*–algebra is the continuous product system associated with a strictly continuous E_0–semigroup that acts on the algebra of all adjointable operators on a Hilbert module with a unit vector.*

Note that the technically most difficult part in the proof of that theorem is to show that the continuous structure induced by the constructed E_0–semigroup and the unit vector, is the same we started with.

4.7 Remark. So suppose, once more, we have a strongly continuous E_0–semigroup ϑ on $\mathcal{B}^a(E)$ where E is a full Hilbert module over a unital C^*–algebra $\mathcal{B}$. If E has a unit vector, then it is not difficult to show that the continuous structure induced on the product system $E^\odot$ associated with ϑ does not depend on the choice of that unit vector. As a sort of surprise, if the members of a product system (continuous or not) derived from an E_0–semigroup on $\mathcal{B}^a(E)$ have unit vectors, this does not mean that the full Hilbert module E need have a unit vector.[o] However, if a

[n] Note that this may fail, if $\mathcal{B}$ is nonunital. Indeed, the product system from Example 2.5 with the canonical embedding $C_0(t,\infty)\to C_0(0,\infty)$ is continuous. But, none of the E_t is full but E_0.

[o] As a trivial example, take the Hilbert M_2–modules $E=\mathbb{C}_2$ where we put $\mathbb{C}_n := (\mathbb{C}^n)^*$. The only E_0–semigroup ϑ on $\mathcal{B}^a(E)=\mathbb{C}$ is $\vartheta_t=\mathsf{id}_{\mathbb{C}}$, its product system simply $E_t=M_2$ with multiplication as identification $E_s\odot E_t=E_{s+t}$. Also the left dilation v_t is simply the canonical identification $E\odot M_2=E$. Nontrivial examples may be obtained by working in the present one via direct sum constructions.

full E does not have a unit vector, then a little lemma from[Ske04] asserts that a finite direct sum E^n of copies of E will have a unit vector. (The proof uses, again, that the invertibles form an open subset, and Cauchy-Schwartz inequality: If E is full, then $\mathbf{1}$ is approximated by finite sums $b = \sum_{i=1}^{n} \langle x_i, y_i \rangle$. If the approximation is sufficiently good, then b is invertible. b can be interpreted as inner product of elements in $X, Y \in E^n$. A simple application of Cauchy-Schwartz inequality shows that also $\langle X, X \rangle$ must be invertible so that $X\sqrt{\langle X, X\rangle^{-1}}$ is a unit vector in E^n.)[p] The strongly continuous E_0–semigroup ϑ on $\mathcal{B}^a(E)$ may be lifted to a strongly continuous E_0–semigroup on $\mathcal{B}^a(E^n) = M_n(\mathcal{B}^a(E))$ (acting pointwise with ϑ on the matrix elements), having the same product system $E^\odot$ as ϑ. Now $E^\odot$ can be induced from a unit vector E^n. Also here the continuous structure depends neither on how big n is, nor on which unit vector we choose.

It is an open problem how to define a continuous structure on $E^\odot$ without reference to a unit vector. A solution might be to give a definition of continuous product systems in terms of Banach bundles (as Hirshberg's[Hir04] for Borel bundles), that is, by giving explicitly a set of sections that are supposed to be continuous and that determine the structure of the bundle. The product system will, then, be considered as obtained via.[MSS06,Ske04] A candidate for the generating set of continuous sections would be the set

$$\Big\{ \big(x^* \odot_t y\big)_{t\in\mathbb{R}_+} \mid x, y \in E \Big\}.$$

It is unclear in how far a definition of continuous product system as Banach bundle (generated by a compatible set of continuous sections) is already sufficient to run through the proof of.[Ske06c] It might be necessary to find a further condition that substitutes the condition being a subbundle of a trivial bundle.

4.8 Remark. We mentioned in Remark 4.1 that Arveson's construction of

[p] This result is key in the proof of[Ske04] that every full discrete product system $E^\odot = \big(E_n\big)_{n\in\mathbb{N}_0}$ of correspondences over a unital C^*–algebra admits a left dilation. In fact, if E_1^n contains a unit vector, then intuitively also $M_{n\cdot\infty,\infty}(E_1) = M_\infty(E)$ should contain a unit vector. The problem is that $M_\infty(E_1)$, usually, is not big enough, but a suitable *strict completion* is. This problem does not appear in the version for von Neumann modules. The price to be paid is that the analogue of the lemma for von Neumann modules requires a direct sum $E_1^{\mathfrak{n}}$ with arbitrary cardinality $\mathfrak{n}$. Now $M_{\mathfrak{n}}(E_1)$ is nothing but $\mathcal{B}^{\mathfrak{n}} \odot E_1 \odot \mathcal{B}_{\mathfrak{n}}$. In other words, the tensor powers of $M_{\mathfrak{n}}(E_1)$ form a product system $\mathcal{B}^{\mathfrak{n}} \odot E^\odot \odot \mathcal{B}_{\mathfrak{n}}$ Morita equivalent to $E^\odot$ in the sense of Footnote g. Once a unit vector in $M_{\mathfrak{n}}(E_1)$ is established, we find a left dilation of $\mathcal{B}^{\mathfrak{n}} \odot E^\odot \odot \mathcal{B}_{\mathfrak{n}}$. And one of the major results of[Ske04] asserts that a product system admits a left dilation, if (and only if) it is Morita equivalent to another product system that admits a left dilation.

a right dilation (that is, of an essential representation) for the Hilbert space case, with tensoring a unit vector from the right, fails for modules. However, in Skeide[Ske06f] we pointed out that the construction, indeed, can be saved if we tensor something different from the right. This "something different" is a unit vector not for E_1 but for the member E'_1 of the commutant of the product system $E^{\odot}$. And an element of E_t can be tensored with en element from E'_s in a reasonable way. We explain the commutant of a product system in Section 6. But we do not have the space to explain any detail (in particular, how the C^*–setting of this section fits into the von Neumann–setting of Section 6) and refer the reader to.[Ske06f] However, we mention that the existence of a unit vector in E'_1 follows by Hirshberg's result[Hir05a] that every full discrete product system with faithful left action admits a right dilation. This result is *dual* to[Ske04] in the sense of commutant. And, as a matter of fact, the condition that the left actions of the product system be faithful is dual to fullness under commutant. For right dilations it is as indispensable as fullness is for left dilations.

5. The Arveson system of an E_0–semigroup on $\mathcal{B}(H)$

Preparing Section 6, in this section we review Arveson's construction of an Arveson system from an E_0–semigroup ϑ on $\mathcal{B}(H)$ and compare it with Bhat's. Arveson defines for every ϑ_t the intertwiner space

$$E_t^A \;:=\; \big\{x'_t \in \mathcal{B}(H)\colon \vartheta_t(a)x'_t = x'_t a \;(a \in \mathcal{B}(H))\big\}.$$

It is readily verified that $x_t'^* y'_t \in \mathcal{B}(H)' = \mathbb{C}\,\mathsf{id}_H$, so that $x_t'^* y'_t = \langle x'_t, y'_t\rangle\,\mathsf{id}_H$ defines a an inner product turning E_t^A into a Hilbert space. Similarly, it is an easy exercise to check that the bilinear mappings

$$(x'_t, y) \;\longmapsto\; x'_t y \qquad\qquad (x'_t, y'_s) \;\longmapsto\; x'_t y'_s$$

define isometries $w'_t\colon E_t^A \otimes H \to H$ and $u'_{t,s}\colon E_t^A \otimes E_s^A \to E_{t+s}^A$. Less obvious is that w'_t is surjective. (Here, normality of ϑ_t is essential; see Lemma 6.1 in the more general context.) But, from surjectivity of w'_t, w'_s and w'_{t+s} it is immediate (exercise!) that also $u'_{t,s}$ is surjective. From $(x'_t y'_s)z'_r = x'_t(y'_s z'_r)$ we see that the $u'_{t,s}$ iterate associatively on multiple tensor products. In other words, the family $E^{A^{\otimes}} = \big(E_t^A\big)_{t\in\mathbb{S}}$ forms an (algebraic) Arveson system. We call $E^{A^{\otimes}}$ the ***Arveson system associated with*** ϑ. But, we have more than just the Arveson system of ϑ. From $(x'_t y'_s)z = x'_t(y'_s z)$ it follows that the w'_t define a right dilation $(w'^{\otimes}, H)$ of $E^{A^{\otimes}}$.

In general, if $(w^{\otimes}, R)$ is a right dilation of $E^{\otimes}$, by setting $\vartheta_t^w(a) := w_t(\mathsf{id}_t \otimes a)w_t^*$ we define an E_0–semigroup ϑ^w on $\mathcal{B}(R)$. It is easy to check

that the Arveson system of ϑ^w is $E^{\otimes}$ by identifying $x_t \in E_t$ with the intertwiner

$$w_t(x_t \otimes \mathsf{id}_R)\colon z \longmapsto w_t(x_t \otimes z).$$

In the case of the Arveson system $E^{A^{\otimes}}$ of an E_0–semigroup ϑ on $\mathcal{B}(H)$ and the right dilation w'_t of $E^{A^{\otimes}}$ to H as constructed before, it follows from

$$\begin{aligned}(w'_t(\mathsf{id}_t \otimes a)w'^*_t)(x'_t y) &= w'_t(\mathsf{id}_t \otimes a)(x'_t \otimes y)\\ &= w'_t(x'_t \otimes ay) = x'_t ay = \vartheta_t(a)(x'_t y)\end{aligned}$$

that $\vartheta_t^{w'} = \vartheta_t$. We summarize:

5.1 Proposition. *Let $E^{\otimes}$ be an (algebraic) Arveson system. The problem of finding an E_0–semigroup that has $E^{\otimes}$ as associated Arveson system is equivalent to the problem of finding a right dilation of $E^{\otimes}$ or, equivalently, a nondegenerate representation.*

Let us return to the E_0–semigroup ϑ on $\mathcal{B}(H)$ and ask what the relation is between $E^{A^{\otimes}}$ and $E^{B^{\otimes}}$. Of course, the dimension of the multiplicity space of an endomorphism ϑ_t is unique, no matter whether we factor it out to the left (right dilation) or to the right (left dilation). Therefore, $E_t^A \cong E_t^B$ as Hilbert spaces. But what happens to the identifications $u'_{t,s}$ and $u_{s,t}$? Can we identify E_t^A with E_t^B in such a way that these identifications are preserved? At this point the reader will have noticed that in the discussion of the Arveson system of ϑ we discussed an identification of $E_t^A \otimes E_s^A$ with E_{t+s}^A, while for $E^{B^{\otimes}}$ we chose the opposite direction of times. From this, the reader might guess what the answer will be: $E^{A^{\otimes}}$ and $E^{B^{\otimes}}$ turn out to be anti-isomorphic, that is , $E^{A^{\otimes}}$ is the opposite product system of $E^{B^{\otimes}}$. The two need not be isomorphic.[q]

Let us see why $E^{A^{\otimes}}$ and $E^{B^{\otimes}}$ are anti-isomorphic. Let $x_t \in E_t^B$ and define $x'_t\colon h \mapsto \vartheta_t(h\xi^*)x_t$ $(= v_t(h \otimes x_t))$. It follows that

$$\vartheta_t(a)x'_t h = \vartheta_t(a)\vartheta_t(h\xi^*)x_t = \vartheta_t(ah\xi^*)x_t = x'_t ah,$$

[q] Tsirelson[Tsi00a] provided us with an explicit example of a product system that is not isomorphic to its opposite product system. However, taking into account the notion of *paired* E_0–semigroup from Powers and Robinson[PR89] (see Observation 3.2) together with their information that the relation defined by *pairing* fails to be transitive, one may deduce the same result from [Arv89a, Theorem 3.4]: Two E_0–semigroups on $\mathcal{B}(H)$ are paired, if and only if their product systems are anti-isomorphic.

so that $x'_t \in E_t^A$. Clearly, $u_t \colon x_t \mapsto x'_t$ is an isometry. Now let $x'_t \in E_t^A$ and define $x_t = x'_t\xi$. Then

$$\vartheta_t(\xi\xi^*)x_t \;=\; \vartheta_t(\xi\xi^*)x'_t\xi \;=\; x'_t(\xi\xi^*)\xi \;=\; x'_t\xi \;=\; x_t,$$

so that $x_t \in E_t^B$. Clearly, $u'_t \colon x'_t \mapsto x_t$ is an isometry. Moreover,

$$(u_tu'_tx'_t)h \;=\; (u_tx'_t\xi)h \;=\; \vartheta_t(h\xi^*)x'_t\xi \;=\; x'_t(h\xi^*)\xi \;=\; x'_th.$$

In other words, u_t and u'_t are a pair of inverse unitaries. Now let us see what u'_{t+s} does to a tensor product $u'_{t,s}(x'_t \otimes y'_s)$.

$$\begin{aligned} u'_{t+s}u'_{t,s}(x'_t \otimes y'_s) \;&=\; u'_{t+s}(x'_ty'_s) \;=\; (x'_ty'_s)\xi \;=\; x'_t(y'_s\xi)\\ &= x'_t(u'_sy'_s) \;=\; x'_t((u'_sy'_s)\xi^*)\xi \;=\; \vartheta_t((u'_sy'_s)\xi^*)x'_t\xi\\ &= \vartheta_t((u'_sy'_s)\xi^*)(u'_tx'_t) \;=\; u_{s,t}((u'_sy'_s) \otimes (u'_tx'_t)). \end{aligned}$$

In other words, $u'_{t+s}u'_{t,s} = u_{s,t}(u'_s \otimes u'_t)$, that is, the family u'_t establishes an anti-isomorphism $E^{A\otimes} \to E^{B\otimes}$.

6. E_0–Semigroups and product systems *à la* Arveson: Commutants of von Neumann correspondences

Of course, the construction in Section 2 of the product system associated with an E_0–semigroup on $\mathcal{B}^a(E)$ works also if E is a von Neumann module. (After all a von Neumann module is also a Hilbert module.) In presence of a unit vector it is even clear that the product system consists of von Neumann modules. (The ranges of projections on von Neumann modules are von Neumann modules.) The point is that in the assumptions on the E_0–semigroup it is sufficient that the endomorphisms of the von Neumann algebra $\mathcal{B}^a(E)$ be only normal, not necessarily strict. In Skeide[Ske05a] we provided a generalization of Bhat's approach (without unit vectors, not along the lines of[MSS06]) that works for every von Neumann module (and, of course, gives a product system isomorphic to that constructed along the lines of[MSS06]). As all modifications to be done are plain, we do not discuss them here.

The approach we want to discuss here, is the generalization of Arveson's approach, as discovered in Skeide[Ske03a] together with the commutant of von Neumann correspondences. To that goal we have to spend some time to review the necessary notions and facts about von Neumann modules, von Neumann correspondences and their commutants. The correspondence between a von Neumann algebra and its commutant is bijective. In order that this desirable property remains true for commutants of von Neumann

correspondences and does not degenerate to an equivalence, we have to choose our categories carefully. The correct category that allows to view the commutant as a bijective functor is the category of *concrete* von Neumann correspondences; Skeide.[Ske06b] In the sequel, we discuss only the case relevant to us, namely, correspondences over $\mathcal{B}$. (See also Remark 6.3.)

Before we can speak about concrete von Neumann correspondences, we have to speak about concrete von Neumann modules. Recall that a von Neumann algebra is a strongly closed $*$–algebra $\mathcal{B} \subset \mathcal{B}(G)$ of operators acting nondegenerately on a Hilbert space G. As usual, by $\mathcal{B}' \subset \mathcal{B}(G)$ we denote the commutant of $\mathcal{B}$. Similarly, a ***concrete von Neumann $\mathcal{B}$–module*** is a subset E of $\mathcal{B}(G,H)$, where H is another Hilbert space, such that

(1) E is a right $\mathcal{B}$–submodule of $\mathcal{B}(G,H)$, that is, $E\mathcal{B} \subset E$,
(2) E is a pre-Hilbert $\mathcal{B}$–module with inner product $\langle x,y\rangle = x^*y$, that is, $E^*E \subset \mathcal{B}$,
(3) E acts nondegenerately on G, that is, $\overline{\mathsf{span}}\, EG = H$, and
(4) E is strongly closed in $\mathcal{B}(G,H)$.

If we wish to underline the Hilbert space H, we will also write the pair (E,H) for the concrete von Neumann $\mathcal{B}$–module. One may show (see Skeide[Ske00,Ske05b]) that a subset E of $\mathcal{B}(G,H)$ fulfilling 1–3 (that is, E is a concrete pre-Hilbert $\mathcal{B}$–module) is a concrete von Neumann $\mathcal{B}$–module, if and only if E is self-dual[r], that is, if and only if E is a W^*–module over the von Neumann algebra $\mathcal{B} \subset \mathcal{B}(G)$ considered as a W^*–algebra. By $\mathfrak{cv}\mathfrak{N}_{\mathcal{B}}$ we denote the ***category of concrete von Neumann $\mathcal{B}$–modules*** with the adjointable maps $a \in \mathcal{B}^a(E_1,E_2)$ as morphisms. The definition of *concrete* von Neumann modules and their category is due to Skeide,[Ske06b] while the definition of von Neumann modules is due to Skeide;[Ske00] see Footnote s.

Identifying $xg \in H$ with $x \odot g \in E \odot G$, we see from 3 that H and $E \odot G$ are canonically isomorphic.[s] Giving E as a subset of $\mathcal{B}(G,H)$ from

[r] Recall that a Hilbert $\mathcal{B}$–module is ***self-dual***, if every bounded right linear map $E \to \mathcal{B}$ has the form $x \mapsto \langle y,x\rangle$ for a suitable $y \in E$.

[s] In fact, if E is a pre-Hilbert module over a pre-C^*–algebra $\mathcal{B} \subset \mathcal{B}(G)$, then one may construct the Hilbert space $E \odot G$ with an embedding $x \mapsto L_x \in \mathcal{B}(G, E \odot G)$ where we put $L_x g := x \odot g$, transforming E into a concrete pre-Hilbert $\mathcal{B}$–module $(E, E \odot G)$. For a von Neumann algebra $\mathcal{B} \subset \mathcal{B}(G)$ we defined in[Ske00] that E is a von Neumann $\mathcal{B}$–module, if its image in $\mathcal{B}(G, E \odot G)$ is strongly closed. Of course, in that way also a W^*–module over a W^*–algebra M may be turned into a von Neumann module after choosing a faithful normal unital representation of M on a Hilbert space G, thus, turning M into a von Neumann algebra.

the beginning, is crucial for that the commutant, later on, will be bijective. The fact that H is canonically isomorphic to the tensor product $E \odot G$ is, however, by far more inspiring from the algebraic point of view.

For instance, every adjointable operator $a \in \mathcal{B}^a(E_1, E_2)$ amplifies to an operator $a \odot \mathsf{id}_G \in \mathcal{B}(E_1 \odot G, E_2 \odot G)$. Consequently, a gives rise to and is determined uniquely by an operator in $\mathcal{B}(H_1, H_2)$ that acts as $x_1 g \mapsto (ax_1)g$. We shall denote this operator by the same symbol a and identify in that way $\mathcal{B}^a(E_1, E_2)$ as a subset of $\mathcal{B}(H_1, H_2)$. It is easy to show that $\mathcal{B}^a(E_1, E_2)$ is strongly closed in $\mathcal{B}(H_1, H_2)$. In particular, $\mathcal{B}^a(E) \subset \mathcal{B}(H)$ is a von Neumann algebra acting on H.

Those operators on the second factor G in $E \odot G$ that embed into $\mathcal{B}(E \odot G)$ are the $\mathcal{B}$–$\mathbb{C}$–linear operators on G. Of course, $\mathcal{B}^{bil}(G) = \mathcal{B}'$ is nothing but the commutant of $\mathcal{B}$. So, the (clearly, normal and nondegenerate) representation $b' \mapsto \mathsf{id}_E \odot b'$ of $\mathcal{B}'$ on $E \odot G$ gives rise to a normal nondegenerate representation ρ' of $\mathcal{B}'$ on H which acts as $\rho'(b')xg = xb'g$. We call ρ' the ***commutant lifting*** associated with E.

From the commutant lifting ρ' we obtain back E as the space

$$E \;=\; C_{\mathcal{B}'}(\mathcal{B}(G,H)) \;:=\; \big\{x \in \mathcal{B}(G,H)\colon \rho'(b')x = xb' \;(b' \in \mathcal{B}')\big\} \qquad (8)$$

of *intertwiners* for the natural actions of $\mathcal{B}'$. This was known already to Rieffel.[Rie74b] In[Ske05b] we proved it by simply calculating the double commutant of the linking von Neumann algebra in $\mathcal{B}(G \oplus H)$:

$$\begin{pmatrix} \mathcal{B} & E^* \\ E & \mathcal{B}^a(E) \end{pmatrix}'' \;=\; \left\{\begin{pmatrix} b' & 0 \\ 0 & \rho'(b') \end{pmatrix}\colon b' \in \mathcal{B}'\right\}' \;=\; \begin{pmatrix} \mathcal{B} & C_{\mathcal{B}'}(\mathcal{B}(H,G)) \\ C_{\mathcal{B}'}(\mathcal{B}(G,H)) & \rho'(\mathcal{B}')' \end{pmatrix}.$$

This proof also shows that the commutant $\rho'(\mathcal{B}')'$ of the range of ρ' in $\mathcal{B}(H)$ may be identified with the von Neumann algebra $\mathcal{B}^a(E) \subset \mathcal{B}(H)$. By doing the computation for $E = E_1 \oplus E_2$ one also shows that $\mathcal{B}^a(E_1, E_2)$ is just $\mathcal{B}^{bil}(H_1, H_2)$, the space of operators that *intertwine* the commutant liftings ρ'_2 and ρ'_1.) Conversely, if (ρ', H) is a normal nondegenerate representation of $\mathcal{B}'$ on the Hilbert space H, then $E := C_{\mathcal{B}'}(\mathcal{B}(G,H))$ as in (8) defines a concrete von Neumann $\mathcal{B}$–module in $\mathcal{B}(G,H)$, which gives back ρ' as commutant lifting. The only critical task, nondegeneracy in Condition 3, is settled by the following result.

6.1 Lemma (Muhly and Solel [MS02, Lemma 2.10]). *If ρ' is a nondegenerate normal representation of $\mathcal{B}'$ on a Hilbert space H, then the intertwiner space $C_{\mathcal{B}'}(\mathcal{B}(G,H))$ acts nondegenerately on G.*[t]

[t] Denote by $P \in \left(\begin{smallmatrix} \mathcal{B} & E^* \\ E & \mathcal{B}^a(E) \end{smallmatrix}\right)' = \left\{\left(\begin{smallmatrix} b' & 0 \\ 0 & \rho'(b') \end{smallmatrix}\right)\colon b' \in \mathcal{B}'\right\}$ the projection onto the invariant

We find that

$$(E,H) \longleftrightarrow (\rho',H) \qquad a\in\mathcal{B}^a(E_1,E_2) \longleftrightarrow a\in\mathcal{B}^{bil}(H_1,H_2) \tag{9}$$

establishes a bijective functor between the category $\mathfrak{cv}\mathfrak{N}_{\mathcal{B}}$ of concrete von Neumann $\mathcal{B}$–modules and the ***category*** ${}_{\mathcal{B}'}\mathfrak{cv}\mathfrak{N}$ ***of normal nondegenerate representations of*** $\mathcal{B}'$ with the intertwiners $\mathcal{B}^{bil}(H_1,H_2)$ as morphisms. (The preceding correspondence was established in Skeide[Ske03a] as an equivalence between the category von Neumann $\mathcal{B}$–modules and ${}_{\mathcal{B}'}\mathfrak{cv}\mathfrak{N}$. As a von Neumann $\mathcal{B}$–module E, first, must be turned into a concrete von Neumann $\mathcal{B}$–modules $(E,E\odot G)$, the correspondence is not bijective but only an equivalence. The precise formulation above, where the functor is, really, bijective and not only an equivalence, is due to.[Ske06b])

A ***concrete von Neumann correspondence*** over a von Neumann algebra $\mathcal{B}$ is a concrete von Neumann $\mathcal{B}$–module (E,H) with a left action of $\mathcal{B}$ such that $\rho\colon\mathcal{B}\to\mathcal{B}^a(E)\to\mathcal{B}(H)$ defines a normal (nondegenerate, of course) representation of $\mathcal{B}$ on H. We call ρ the ***Stinespring representation*** associated with E.

6.2 Remark. The GNS-correspondence of a (normal) CP-map T on $\mathcal{B}$ is the unique Hilbert (von Neumann) $\mathcal{B}$–correspondence E which has a vector $\xi\in E$ that generates E as a (von Neumann) correspondence and gives back T as $T(b)=\langle\xi,b\xi\rangle$; see Paschke.[Pas73] For this GNS-correspondence E, the representation ρ is, indeed, the Stinespring representation, while ρ' is (a restriction of) the representation constructed by Arveson[Arv69] in the section called "lifting commutants".

By ${}_{\mathcal{B}}\mathfrak{cv}\mathfrak{N}_{\mathcal{B}}$ we denote the ***category of concrete von Neumann correspondences from*** $\mathcal{B}$ ***to*** $\mathcal{B}$ with the bilinear adjointable maps $a\in\mathcal{B}^{a,bil}(E_1,E_2)$ as morphisms. (For adjointable maps, only left $\mathcal{B}$–linearity has to be checked.) We observe that $\rho(\mathcal{B})\subset\mathcal{B}^a(E)=\rho'(\mathcal{B}')'$, that is, ρ' and ρ have mutually commuting ranges. As this is very close to correspondences in the sense of Connes[Con80] (if $\mathcal{B}$ is in *standard form*, then $\mathcal{B}'\cong\mathcal{B}^{op}$), we introduce the ***category of concrete Connes correspondences*** ${}_{\mathcal{B}}\mathfrak{c}\mathfrak{C}_{\mathcal{B}}$ whose objects are triples (ρ',ρ,H) such that ρ' and ρ are a pair of normal nondegenerate representations of $\mathcal{B}'$ and of $\mathcal{B}$, respectively, on H with mutually commuting ranges, and with those maps in $\mathcal{B}(H_1,H_2)$ as morphisms

subspace $\overline{\mathsf{span}}\left(\begin{smallmatrix}\mathcal{B} & E^*\\ E & \mathcal{B}^a(E)\end{smallmatrix}\right)\left(\begin{smallmatrix}G\\ H\end{smallmatrix}\right)$. This subspace contains G so that $P=\left(\begin{smallmatrix}\mathbf{1} & 0\\ 0 & \rho'(\mathbf{1})\end{smallmatrix}\right)$. Since ρ' is nondegenerate, the statement follows.

that intertwine both actions that of $\mathcal{B}'$ and that of $\mathcal{B}$. Extending the correspondence between concrete von Neumann $\mathcal{B}$–modules and representations of $\mathcal{B}'$, we find a find bijective functor between the category of concrete von Neumann $\mathcal{B}$–correspondences (E, H) and the category of concrete Connes correspondences (ρ', ρ, H). In[Ske03a] we observed this as an equivalence for von Neumann correspondences, while the bijective version for concrete von Neumann correspondences is from.[Ske06b]

A last almost trivial observation (once again in[Ske03a] up to equivalence and in,[Ske06b] really, bijective) consists in noting that in the representation picture the roles of the representations ρ' and ρ are absolutely symmetric. That is, ${}_{\mathcal{B}}\mathfrak{c}\mathfrak{C}_{\mathcal{B}} \cong {}_{\mathcal{B}'}\mathfrak{c}\mathfrak{C}_{\mathcal{B}'}$. Therefore, if we switch $\mathcal{B}$ and $\mathcal{B}'$, that is, if we interprete ρ as commutant lifting of $\mathcal{B}$, the commutant of $\mathcal{B}'$, and ρ' as Stinespring representation of $\mathcal{B}'$, by

$$E' := C_{\mathcal{B}}(\mathcal{B}(G,H)) := \big\{x' \in \mathcal{B}(G,H) \colon \rho(b)x' = x'b \ (b \in \mathcal{B})\big\} \qquad (10)$$

we obtain a von Neumann $\mathcal{B}'$–module which is turned into a von Neumann $\mathcal{B}'$–correspondence by defining a left action via ρ'. We call E' the ***commutant*** of E. The commutant is a bijective functor from the category of concrete von Neumann $\mathcal{B}$–correspondences onto the category of concrete von Neumann $\mathcal{B}'$–correspondences (in each case with the bilinear adjointable maps as morphisms that are, really, the same algebra $\mathcal{B}^a(E) \cap \mathcal{B}^a(E') = \rho'(\mathcal{B}')' \cap \rho(\mathcal{B})'$ of operators in $\mathcal{B}(H)$). Obviously, $E'' := (E')' = E$.

6.3 Remark. Muhly and Solel[MS04] have discussed (independently) a version of the commutant for W^*–algebras, called *σ–dual*, where σ is a faithful representation of the underlying W^*–algebra, that must be chosen, and the σ–dual depends on σ (up to Morita equivalence of correspondences[MS05a]). An extension to correspondences from $\mathcal{A}$ to $\mathcal{B}$ was first done in the setting of σ–duals in.[MS05a] In[Ske06b] we discussed the version for von Neumann algebras and (concrete) von Neumann correspondences.

We remark that the functor $\mathfrak{cv}\mathfrak{N}_{\mathcal{B}} \leftrightarrow {}_{\mathcal{B}'}\mathfrak{cv}\mathfrak{N}$ in (9) fits canonically into the setting of the commutant functor as ${}_{\mathbb{C}}\mathfrak{cv}\mathfrak{N}_{\mathcal{B}} \longleftrightarrow {}_{\mathcal{B}'}\mathfrak{cv}\mathfrak{N}_{\mathbb{C}'}$, if we consider $\mathbb{C} = \mathbb{C}' \subset \mathcal{B}(\mathbb{C}) = \mathbb{C}$ as a von Neumann algebra.

The tensor product of Connes correspondences is tricky to describe in terms that do not explicitly involve the von Neumann correspondences to which they correspond. It requires that the von Neumann algebra is a W^*–algebra in standard form and parts from Tomita-Takesaki theory and the result depends manifestly on the choice of a normal semifinite weight;

see, for instance, Takesaki [Tak03, Section IX.3]. Also the tensor product of W^*–correspondences, although definitely less involved, still has the problem that the usual tensor product must be completed in a suitable σ–*topology*, and this topology is defined rather *ad hoc*.

The tensor product two of von Neumann correspondences E_1 and E_2 is easy to obtain (and unique up to unitary equivalence): Simply construct $E_1 \odot E_2 \odot G$ and determine the strong closure of $E_1 \underline{\odot} E_2$ in $\mathcal{B}(G, E_1 \odot E_2 \odot G)$ or, equivalently, determine the intertwiner space $C_{\mathcal{B}'}(\mathcal{B}(G, E_1 \odot E_2 \odot G))$, a purely algebraic problem, like determining the double commutant of a $*$–algebra of operators. Up to canonical isomorphism it is not important whether we construct first $E_1 \underline{\odot} E_2$ and then $(E_1 \underline{\odot} E_2) \odot G$ or first $E_2 \odot G$ and then $E_1 \odot (E_2 \odot G)$. If we have concrete von Neumann correspondences (E_1, H_1) and (E_2, H_2) it occurs to be more adapted to construct $E_1 \odot H_2$ as the space H_2, canonically isomorphic to $E_2 \odot G$, is given from the beginning. By slight abuse of notation we shall denote the concrete von Neumann correspondence obtained in that way by $E_1 \odot E_2 \subset \mathcal{B}(G, E_1 \odot H_2)$, using the same symbol $\odot$ as for the tensor product of C^*–correspondences. Anyway, no matter how we obtained $E_1 \odot E_2 \odot G$, as $(E_1 \underline{\odot} E_2) \odot G$, as $E_1 \odot (E_2 \odot G)$ or as $E_1 \odot H_2$, to fix an isomorphism from the concrete von Neumann correspondence $(E_1 \odot E_2, E_1 \odot E_2 \odot G)$ to a concrete von Neumann correspondence (F, K) simply means to fix a unitary $u \in \mathcal{B}(E_1 \odot E_2 \odot G, K)$ that intertwines both the commutant liftings of $\mathcal{B}'$ and the Stinespring representations of $\mathcal{B}$.

The notations established so far allow to state and prove that the commutant establishes a bijective functor between the category $\mathfrak{cv}\mathfrak{N}_{\mathcal{B}}^{\odot}$ of product systems of concrete von Neumann $\mathcal{B}$–correspondences and the category $\mathfrak{cv}\mathfrak{N}_{\mathcal{B}'}^{\odot}$ of product systems of concrete von Neumann $\mathcal{B}'$–correspondences; see Skeide[Ske03a,Ske06e] and Muhly and Solel.[MS05b] A morphism between two objects $E^{\odot}$ and $F^{\odot}$ in $\mathfrak{cv}\mathfrak{N}_{\mathcal{B}}^{\odot}$ is a family $a^{\odot} = \big(a_t\big)_{t\in\mathbb{S}}$ of maps $a_t \in \mathcal{B}^{a,bil}(E_t, F_t)$ that fulfills $a_s \odot a_t = a_{s+t}$ and $a_0 = \mathsf{id}_{\mathcal{B}}$.

We sketch this very briefly. Suppose $E^{\odot}$ is a product system of concrete von Neumann correspondences (E_t, H_t) over $\mathcal{B}$ and denote by (ρ'_t, ρ_t, H_t) the corresponding concrete Connes correspondence. The familiy $u_{s,t}\colon E_s \odot E_t \to E_{s+t}$ that determines the product system structure, in the picture of Hilbert spaces is captured by unitaries $u_{s,t} \in \mathcal{B}(E_s \odot H_t, H_{s+t})$ that intertwine both the actions of $\mathcal{B}$ and the actions of $\mathcal{B}'$ on these Hilbert spaces. The associativity condition reads $u_{r,s+t}(x_r \odot u_{s,t}(y_s \odot h_t)) = u_{r+s,t}(u_{r,s}(x_s \odot y_r) \odot h_t)$. The double meaning in this formula of $u_{r,s+t}$, $u_{s,t}$,

$u_{r+s,t}$ as operators between Hilbert spaces and of $u_{r,s}$ as operator between correspondences is not exactly satisfactory. We will circumvent this (purely formal) difficulty in[Ske07] by giving a different definition of the product system structure in terms of *representations of Hilbert modules*[u] where the $u_{s,t}$ will appear no longer as a defining object but as a derived one. Here we limit ourselves to explain how the product system structure of $E^{\odot}$ gives rise to a product system structure of $E'^{\odot}$ by a giving some (canonical) isomorphisms as we did in.[Ske03a]

Indeed, to define a (bilinear) unitary $E'_t \odot E'_s \to E'_{t+s}$ we have to establish a unitary $u'_{t,s}\colon E'_t \odot H_s \to H_{t+s}$ intertwining the relevant representations. We have

$$E_s \odot E_t \odot G \;\cong\; E_s \odot E'_t \odot G \;\cong\; E'_t \odot E_s \odot G \;\cong\; E'_t \odot E'_s \odot G,$$

where we used two times $E_t \odot G \cong H_t \cong E'_t \odot G$ and, in the middle, an isomorphism that, indeed, simply flips $x_s \odot y'_t \odot g$ to $y'_t \odot x_s \odot g$. (This flip is the only place where we have to compute something.) Attaching elements $b \in \mathcal{B}$ and $b' \in \mathcal{B}'$ in every part to the places where they act naturally, we see that the suggested isomorphism intertwines their actions. This chain of isomorphisms written down for an arbitrary tensor product of correspondences (even over different von Neumann algebras; see Remark 6.3) shows clearly that the commutant flips orders in tensor products:

$$(E \odot F)' \;\cong\; F' \odot E'; \tag{11}$$

see [Ske03a, Theorem 2.3] and [MS05a, Lemma 3.3]. For our scope here, defining $u'_{t,s}$, it is sufficient to look at the chain

$$E'_t \odot H_s \;=\; E'_t \odot \overline{\mathsf{span}}(E_s G) \;\cong\; E_s \odot \overline{\mathsf{span}}(E'_t G) \;=\; E_s \odot H_t \overset{u_{s,t}}{\cong} H_{s+t}. \tag{12}$$

This chain shows clearly how $u_{s,t}$ enters and that the flip of the elements x'_t ans y_s is the only thing where we are really doing something.[v] It is routine to show that the $u'_{t,s}$ defined in that way turn $E'^{\odot}$ into a product system, the ***commutant system*** of $E^{\odot}$.

6.4 Example. Suppose $E^{\otimes}$ is an Arveson system. We turn E_t in to the concrete von Neumann correspondence $(E_t, H_t = E_t)$ by identifying $x_t \in E_t$

[u] This is not to be confused with the term *covariant representation* of a correspondence in the work of Muhly and Solel starting with.[MS98]

[v] Behind this flip there is a sort of tensor product among a von Neumann $\mathcal{B}$–module and a von Neumann $\mathcal{B}'$–module, resulting into a von Neumann $(\mathcal{B} \cap \mathcal{B}')'$–module. In the picture of Connes correspondences this is closely related to Sauvageot.[Sau80,Sau83] We will discuss this in detail in.[Ske07]

with the map $\lambda \mapsto x_t\lambda$ in $\mathcal{B}(\mathbb{C}, E_t)$. Both the Stinespring representation of $\mathbb{C} \subset \mathcal{B}(\mathbb{C})$ and the commutant lifting of $\mathbb{C}' = \mathbb{C}$ are simply the map $\lambda \mapsto \lambda\,\mathsf{id}_t$. So $E'_t = C_{\mathbb{C}}(\mathcal{B}(\mathbb{C}, E_t)) = E_t$. But, the unitary $u'_{t,s}\colon E'_t \otimes E'_s \to E'_{t+s}$ suggested by (12) means: Take $x'_t \otimes y'_s$, express $y'_s \in H_s = E_s$ as $y_s 1 \in E_s\mathbb{C}$ with $y_s = y'_s$, flip the tensors to $y_s \otimes x'_t 1$ and interpret $x'_t 1 \in E'_t\mathbb{C}$ as element in $x_t \in H_t = E_t = E'_t$. In this special case, thanks to $E_t = E'_t$ (a formula, that in the general case has no sense), everything can be done at once by defining $u'_{t,s}(x'_t \otimes y'_t) := u_{s,t}(y'_s \otimes x'_t)$.

As a more elaborate example we discuss now the construction from[Ske03a] that, first, generalizes Arveson's construction of a product system from an E_0–semigroup and, then, shows that the commutant of that system is the product system associated with the E_0–semigroup. So let (E, H) be a strongly full von Neumann module over a von Neumann algebra $\mathcal{B} \subset \mathcal{B}(G)$. ***Strongly full*** means that $\langle E, E\rangle$ generates $\mathcal{B}$ as von Neumann algebra. It is easy to show (exercise!) that E is strongly full, if and only if the associated commutant lifting (ρ', H) is faithful. Let ϑ be an E_0–semigroup of (normal unital) endomorphisms ϑ_t of $\mathcal{B}^a(E) \subset \mathcal{B}(H)$. Like Arveson we define the intertwiner space[w]

$$E'_t \;:=\; \big\{x'_t \in \mathcal{B}(H)\colon \vartheta_t(a)x'_t = x'_t a \;(a \in \mathcal{B}^a(E))\big\}.$$

Since $x'^*_t y'_t \in \mathcal{B}^a(E)' = \rho'(\mathcal{B}')$ and ρ' is faithful, we may define an inner product

$$\langle x'_t, y'_t\rangle \;:=\; \rho'^{-1}(x'^*_t y'_t)$$

on E'_t with values in $\mathcal{B}'$. It is plain to verify that E'_t with this inner product and the bimodule operation $b'_1 x'_t b'_2 := \rho'(b'_1)x'_t\rho'(b'_2)$ is a von Neumann correspondence over $\mathcal{B}'$. Observe that E'_t acts nondegenerately on H by Lemma 6.1 and that it is the only space of intertwiners of ϑ_t and $\mathsf{id}_{\mathcal{B}^a(E)}$ with this property. It follows that

$$x'_t \odot y'_s \;\mapsto\; x'_t y'_s$$

defines an isomorphism from $E'_t \odot E'_s$ onto E'_{t+s} turning $E'^{\odot} = \big(E'_t\big)_{t\in\mathbb{S}}$ into a product system.

Like the Arveson system of an E_0–semigroup on $\mathcal{B}(H)$, the product system $E'^{\odot}$ associated à la Arveson with the E_0–semigroup ϑ on

[w] Note that the interwiners are calculated in $\mathcal{B}(H)$ of which $\mathcal{B}^a(E)$ is a von Neumann subalgebra. Taking the analogy with[Arv89a] too literally would mean to determine only the interwiners in $\mathcal{B}^a(E)$ what is not approximately as useful.

$\mathcal{B}^a(E) \subset \mathcal{B}(H)$ comes along with a ***faithful nondegenerate representation*** $\eta'_t\colon E'_t \to \mathcal{B}(H)$ on the Hilbert space H. By this we mean two things: Firstly, the η'_t are isometric in the sense that $\eta'_t(x'_t)^*\eta'_t(y'_t) = \eta'_0(\langle x'_t, y'_t\rangle)$, and the η'_t are multiplicative in the sense that $\eta'_t(x'_t)\eta'_s(y'_s) = \eta'_{t+s}(x'_t y'_s)$. From this it follows that for each t the pair (η'_t, η'_0) is an isometric covariant representation of E'_t in the sense of Muhly and Solel.[MS98] In particular, η'_0 is a representation of $\mathcal{B}'$. And, secondly, the η'_t are normal and faithful in the sense that their unique extension to a representation of the *von Neumann linking algebra* $\left(\begin{smallmatrix}\mathcal{B}' & E'^*_t\\ E'_t & \mathcal{B}^a(E'_t)\end{smallmatrix}\right)$ on $\mathcal{B}(H\oplus H) = M_2(\mathcal{B}(H))$ is normal and faithful. One may show that this is the case, if and only if η'_0 is normal and faithful. In our case, η'_t is simply the canonical embedding $E'_t \to \mathcal{B}(H)$. In particular, $\eta'_0 = \rho'$. Like for Arveson systems, speaking about a faithful nondegenerate representation is the same as speaking about a ***right dilation*** of $E'^{\odot}$ to R in the sense that R is Hilbert space with a faithful normal representation ρ' of $\mathcal{B}'$ and $w'_t\colon E'_t \odot R \to R$ is a family of unitaries in $\mathcal{B}^{bil}(E'_t \odot R, R)$ that iterates associatively with the product system structure. If we have such a right dilation then $\vartheta^{w'}_t(a) = w'_t(\mathsf{id}_{E'_t} \odot a)w'^*_t$ defines an E_0–semigroup on the concrete von Neumann $\mathcal{B}$–module (L, R) determined by (ρ', R), giving back $E'^{\odot}$ as product system of intertwiners. Note that faithfulness of ρ' implies that L is strongly full and that the left action of $\mathcal{B}'$ on all E'_t is faithful. We say the product system $E'^{\odot}$ is ***faithful***.

Returning to the E_0–semigroup ϑ on $\mathcal{B}^a(E)$ and its product system *à la* Arveson $E'^{\odot}$, let us turn E'_t into a concrete von Neumann correspondence $(H_t, \sigma_t, \sigma'_t)$ by defining the Hilbert space $H_t = E'_t \odot G$, the commutant lifting $\sigma_t(b) = \mathsf{id}_{E'_t} \odot b$ of $(\mathcal{B}')' = \mathcal{B}$, and the Stinespring representation $\sigma'_t(b') = b' \odot \mathsf{id}_G$ of $\mathcal{B}'$. Then its commutant

$$E_t \;:=\; C_{\mathcal{B}'}(\mathcal{B}(G, H_t)) \;=\; \big\{x_t \in \mathcal{B}(G, H_t)\colon \sigma'_t(b')x_t = x_t b' \;(b' \in \mathcal{B}')\big\}$$

is a concrete von Neumann correspondence over $\mathcal{B}$ with left action via σ_t. If we now apply (11) to $E'_t \odot H$ what we find is $H' \odot E_t = E \odot E_t$, because (H', H), the commutant of $(H, H) \in {}_{\mathcal{B}'}\mathfrak{cv}\mathfrak{N}_{\mathbb{C}}$, is nothing but (E, H); see Remark 6.3. The (bilinear!) isomorphisms $w'_t\colon E'_t \odot H \to H$ give, therefore, rise to isomorphisms $v_t := (w'_t)'\colon E \odot E_t \to E$. (Again, we do not show that associativity is respected.) This transition between right dilations of $E'^{\odot}$ and left dilations of $E^{\odot}$ does not depend on that we started from an E_0–semigroup on $\mathcal{B}^a(E)$. In fact, we have the complete analogue of Observation 3.1 including that $E^{\odot}$ is strongly full, if and only if $E'^{\odot}$ is faithful. We mention that this correspondence between left dilations of $E^{\odot}$ (that is, E_0–semigroup having $E^{\odot}$ as associated product system) and

right dilations of $E'^{\odot}$ (that is, nondegenerate representations of $E'^{\odot}$) is due to.[Ske04] It has been generalized to not necessarily unital endomorphism semigroups and not necessarily nondegenerate representations in [Ske06e, Theorem 3.6(3)]. Just that in[Ske04,Ske06e] we did not yet use the terminology of dilations of a product system.

If $E^{\odot}$ is both strongly full and faithful, then also Observation 3.2 remains true. In[Ske07] we will show that a *strongly continuous* product system $E^{\odot}$ of (concrete) von Neumann correspondences admits a *strongly continuous* left dilation, if it is strongly full, and a *strongly continuous* right dilation, if it is faithful.[x] In this case, the unitary group according to Observation 3.2 is strongly continuous and so are the two E_0–semigroups ϑ and θ. This means that the commutant of $E^{\odot}$ is also derived from a strongly continuous E_0–semigroup and, therefore, possesses a ***strongly continuous*** structure. We do not describe here the definition of *strongly continuous* product systems. Apart from missing space, at the time being we have more than one candidate for a definition, and all candidates work well. We did not yet find out which one we should consider the best one. Anyway, the results will allow to show the following theorem.

6.5 Theorem. *The commutant of a strongly continuous, strongly full and faithful product system is strongly continuous, strongly full and faithful, too.*

Muhly and Solel[MS05b] have a similar result for *measurable* product systems under separability assumptions. However, while our proof relies essentially on the product system structure (in that we have to construct a left and a right dilation and to use the semigroup structure encoded by them), their proof is rather a result on general measurable bundles of correspondences and works by a reduction to the analogue result for von Neumann algebras due to Effros.[Eff65] In general it is far from being obvious why a bundle of intertwiners between bundles of Banach modules should admit (strongly) continuous sections.

6.6 Remark. The list of dualities may be extended. For instance, the fact that (by using *quasi orthonormal bases* of von Neumann $\mathcal{B}$–modules, as suit-

[x] Unlike continuous product systems (unital $\mathcal{B}$) where existence of a unit vector in E_1 was automatic, it is an open problem whether strongly continuous product systems always have unit vectors. (The C^*–proof does no longer work, because the invertibles are not open in the strong topology.) So, the results from[Ske04] (existence of a left and right dilation of the discrete subsystem) and the basic idea of[Ske06a] (how to turn it into a dilation of the continuous time system) become indispensable.

able substitute for orthonormal bases of Hilbert spaces) every von Neumann $\mathcal{B}$–module is a complemented submodule of a *free* von Neumann $\mathcal{B}$–module, may be used to prove the *amplification-induction theorem* on the representations ρ' of $\mathcal{B}'$. In the presence of invariant vector states there is a duality between CP-maps from $\mathcal{A}$ to $\mathcal{B}$ and CP-maps from $\mathcal{B}'$ to $\mathcal{A}'$ (Albeverio and Hoegh-Krohn[AHK78]) that includes a duality between *tensor dilations* of a CP-maps on $\mathcal{B}$ and *extensions* from $\mathcal{B}'$ to $\mathcal{B}(G)$ of the dual of that CP-map; see Gohm and Skeide.[GS05] Applying the duality of CP-maps to the canonical embedding of a subalgebra $\mathcal{A} \subset \mathcal{B}$ into $\mathcal{B}$ (both in standard form) and translating back the dual map $\mathcal{B}' \to \mathcal{A}'$ into a map $\mathcal{B} \to \mathcal{A}$ via twofold *Tomita conjugation*, one obtains the *Accardi-Cecchini conditional expectation*[AC82] that coincides with the usual conditional expectation whenever the latter exists; see also Accardi and Longo.[Lon84,AL93]

We mention the duality between Rieffel's *Eilenberg-Watts theorem*[Rie74b] about functors between categories of representations of von Neumann algebras and Blecher's *Eilenberg-Watts theorem*[Ble97] about functors between categories of Hilbert modules. When the latter is restricted to von Neumann modules, the two *Eilenberg-Watts theorems* are dual to each other under the commutant; see.[Ske06b]

Last but surely not least there is the duality between the product system of a CP-semigroup in Bhat and Skeide[BS00] and the product system constructed from the same CP-semigroup by Muhly and Solel.[MS02] The latter is the the commutant of the former, a problem left open in[MS02] that lead to the notion of commutant of correspondences and product systems in.[Ske03a]

Bibliography

AC82. L. Accardi and C. Cecchini, *Conditional expectations in von Neumann algebras and a theorem of Takesaki*, J. Funct. Anal. **45** (1982), 245–273.

AHK78. S. Albeverio and R. Hoegh-Krohn, *Frobenius theory for positive maps of von Neumann algebras*, Commun. Math. Phys. **64** (1978), 83–94.

AL93. L. Accardi and R. Longo, *Martingale convergence of generalized conditional expectations*, J. Funct. Anal. **118** (1993), 119–130.

Ara70. H. Araki, *Factorizable representations of current algebra*, Publ. Res. Inst. Math. Sci. **5** (1970), 361–422.

Arv69. W. Arveson, *Subalgebras of C^*–algebras*, Acta Math. **123** (1969), 141–224.

Arv89a. ———, *Continuous analogues of Fock space*, Mem. Amer. Math. Soc., no. 409, American Mathematical Society, 1989.

Arv89b. ———, *Continuous analogues of Fock space III: Singular states*, J. Operator Theory **22** (1989), 165–205.

Arv90a. ———, *Continuous analogues of Fock space II: The spectral C^*–algebra*, J. Funct. Anal. **90** (1990), 138–205.

Arv90b. ———, *Continuous analogues of Fock space IV: essential states*, Acta Math. **164** (1990), 265–300.

Arv96. ———, *Dynamical invariants for noncommutative flows*, Operator algebras and quantum field theory (S. Doplicher, R. Longo, J.E. Roberts, and L. Zsido, eds.), International Press, 1996, pp. 476–514.

Arv06. ———, *On the existence of E_0–semigroups*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **9** (2006), 315–320.

BBLS04. S.D. Barreto, B.V.R. Bhat, V. Liebscher, and M. Skeide, *Type I product systems of Hilbert modules*, J. Funct. Anal. **212** (2004), 121–181, (Preprint, Cottbus 2001).

Bha96. B.V.R. Bhat, *An index theory for quantum dynamical semigroups*, Trans. Amer. Math. Soc. **348** (1996), 561–583.

BL05. B.V.R. Bhat and J.M. Lindsay, *Regular quantum stochastic cocycles have exponential product systems*, Quantum Probability and Infinite Dimensional Analysis — From Foundations to Applications (M. Schürmann and U. Franz, eds.), Quantum Probability and White Noise Analysis, no. XVIII, World Scientific, 2005, pp. 126–140.

Ble97. D.P. Blecher, *A new approach to Hilbert C^*–modules*, Math. Ann. **307** (1997), 253–290.

BS00. B.V.R. Bhat and M. Skeide, *Tensor product systems of Hilbert modules and dilations of completely positive semigroups*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **3** (2000), 519–575, (Rome, Volterra-Preprint 1999/0370).

BS05. B.V.R. Bhat and R. Srinivasan, *On product systems arising from sum systems*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **8** (2005), 1–31.

Con80. A. Connes, *Correspondences*, His hand-written notes, unpublished, 1980.

Eff65. E.G. Effros, *The Borel structure of von Neumann algebras on a separable Hilbert space*, Pac. J. Math. **15** (1965), 1153–1164.

GS05. R. Gohm and M. Skeide, *Constructing extensions of CP-maps via tensor dilations with the help of von Neumann modules*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **8** (2005), 291–305, (arXiv: math.OA/0311110).

Hir04. I. Hirshberg, *C^*–Algebras of Hilbert module product systems*, J. Reine Angew. Math. **570** (2004), 131–142.

Hir05a. ———, *Essential representations of C^*–correspondences*, Int. J. Math. **16** (2005), 765–775.

Hir05b. ———, *On the universal property of Pimsner-Toeplitz C^*–algebras and their continuous analogues*, J. Funct. Anal. **219** (2005), 21–33.

HP84. R.L. Hudson and K.R. Parthasarathy, *Quantum Ito's formula and stochastic evolutions*, Commun. Math. Phys. **93** (1984), 301–323.

HZ03. I. Hirshberg and J. Zacharias, *On the structure of spectral algebras and their generalizations*, Advances in quantum dynamics (G.L. Price,

B.M. Baker, P.E.T. Jorgensen, and P.S. Muhly, eds.), Contemporary Mathematics, no. 335, American Mathematical Society, 2003, pp. 149–162.

Lie03. V. Liebscher, *Random sets and invariants for (type II) continuous tensor product systems of Hilbert spaces*, Preprint, arXiv: math.PR/0306365, 2003, To appear in Mem. Amer. Math. Soc.

Lon84. R. Longo, *Solution of the factorial Stone-Weierstrass conjecture*, Invent. Math. **76** (1984), 145–155.

MS98. P.S. Muhly and B. Solel, *Tensor algebras over C^*–correspondences: representations, dilations, and C^*–envelopes*, J. Funct. Anal. **158** (1998), 389–457.

MS02. ———, *Quantum Markov processes (correspondences and dilations)*, Int. J. Math. **51** (2002), 863–906, (arXiv: math.OA/0203193).

MS04. ———, *Hardy algebras, W^*–correspondences and interpolation theory*, Math. Ann. **330** (2004), 353–415, (arXiv: math.OA/0308088).

MS05a. ———, *Duality of W^*-correspondences and applications*, Quantum Probability and Infinite Dimensional Analysis — From Foundations to Applications (M. Schürmann and U. Franz, eds.), Quantum Probability and White Noise Analysis, no. XVIII, World Scientific, 2005, pp. 396–414.

MS05b. ———, *Quantum Markov semigroups (product systems and subordination)*, Preprint, arXiv: math.OA/0510653. To appear in Int. J. Math., 2005.

MSS06. P.S. Muhly, M. Skeide, and B. Solel, *Representations of $\mathcal{B}^a(E)$*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **9** (2006), 47–66, (arXiv: math.OA/0410607).

Pas73. W.L. Paschke, *Inner product modules over B^*–algebras*, Trans. Amer. Math. Soc. **182** (1973), 443–468.

Pow87. R.T. Powers, *A non-spatial continuous semigroup of $*$–endomorphisms of $\mathcal{B}(\mathfrak{H})$*, Publ. Res. Inst. Math. Sci. **23** (1987), 1053–1069.

PR89. R.T. Powers and D. Robinson, *An index for continuous semigroups of $*$–endomorphisms of $\mathcal{B}(H)$*, J. Funct. Anal. **84** (1989), 85–96.

PS72. K.R. Parthasarathy and K. Schmidt, *Positive definite kernels, continuous tensor products, and central limit theorems of probability theory*, Lect. Notes Math., no. 272, Springer, 1972.

Rie74a. M.A. Rieffel, *Induced representations of C^*-algebras*, Adv. Math. **13** (1974), 176–257.

Rie74b. ———, *Morita equivalence for C^*–algebras and W^*–algebras*, J. Pure Appl. Algebra **5** (1974), 51–96.

RSN82. F. Riesz and B. Sz.-Nagy, *Vorlesungen über Funktionalanalysis*, Deutsch, 1982.

Sau80. J.-L. Sauvageot, *Produites tensoriels de Z–modules*, Preprint 23, Publ. Univ. P. & M. Curie, 1980.

Sau83. ———, *Sur le produit tensoriel relatif d'espace de Hilbert*, J. Operator Theory **9** (1983), 237–252.

Sch93. M. Schürmann, *White noise on bialgebras*, Lect. Notes Math., no. 1544,

Springer, 1993.

Ske00. M. Skeide, *Generalized matrix C^*–algebras and representations of Hilbert modules*, Mathematical Proceedings of the Royal Irish Academy **100A** (2000), 11–38, (Cottbus, Reihe Mathematik 1997/M-13).

Ske02. ———, *Dilations, product systems and weak dilations*, Math. Notes **71** (2002), 914–923.

Ske03a. ———, *Commutants of von Neumann modules, representations of $\mathcal{B}^a(E)$ and other topics related to product systems of Hilbert modules*, Advances in quantum dynamics (G.L. Price, B.M. Baker, P.E.T. Jorgensen, and P.S. Muhly, eds.), Contemporary Mathematics, no. 335, American Mathematical Society, 2003, (Preprint, Cottbus 2002, arXiv: math.OA/0308231), pp. 253–262.

Ske03b. ———, *Dilation theory and continuous tensor product systems of Hilbert modules*, Quantum Probability and Infinite Dimensional Analysis (W. Freudenberg, ed.), Quantum Probability and White Noise Analysis, no. XV, World Scientific, 2003, Preprint, Cottbus 2001, pp. 215–242.

Ske04. ———, *Unit vectors, Morita equivalence and endomorphisms*, Preprint, arXiv: math.OA/0412231, 2004.

Ske05a. ———, *Three ways to representations of $\mathcal{B}^a(E)$*, Quantum Probability and Infinite Dimensional Analysis — From Foundations to Applications (M. Schürmann and U. Franz, eds.), Quantum Probability and White Noise Analysis, no. XVIII, World Scientific, 2005, (arXiv: math.OA/0404557), pp. 504–517.

Ske05b. ———, *Von Neumann modules, intertwiners and self-duality*, J. Operator Theory **54** (2005), 119–124, (arXiv: math.OA/0308230).

Ske06a. ———, *A simple proof of the fundamental theorem about Arveson systems*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **9** (2006), 305–314, (arXiv: math.OA/0602014).

Ske06b. ———, *Commutants of von Neumann Correspondences and Duality of Eilenberg-Watts Theorems by Rieffel and by Blecher*, Quantum probability (M. Bozejko, W. Mlotkowski, and J. Wysoczanski, eds.), Banach Center Publications, vol. 73, Polish Academy of Sciences — Institute of Mathematics, 2006, (arXiv: math.OA/0502241), pp. 391–408.

Ske06c. ———, *E_0–semigroups for continuous product systems*, Preprint, arXiv: math.OA/0607132, 2006.

Ske06d. ———, *Existence of E_0–semigroups for Arveson systems: Making two proofs into one*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **9** (2006), 373–378, (arXiv: math.OA/0605480).

Ske06e. ———, *Isometric dilations of representations of product systems via commutants*, Preprint, arXiv: math.OA/0602459, 2006.

Ske06f. ———, *Nondegenerate representation of continuous product systems*, Preprint, arXiv: math.OA/0607362, 2006.

Ske06g. ———, *The index of (white) noises and their product systems*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **9** (2006), 617–655, (Rome, Volterra-Preprint 2001/0458, arXiv: math.OA/0601228).

Ske07. ———, *Dilations of product sytems and commutants of von Neumann modules*, in preparation, 2007.

Str69. R.F. Streater, *Current commutation relations, continuous tensor products and infinitely divisible group representations*, Local quantum theory (R. Jost, ed.), Academic Press, 1969.

Tak03. M. Takesaki, *Theory of operator algebras II*, Encyclopaedia of Matematical Sciences, no. 125 (number VI in the subseries Operator Algebras and Non-Commutative Geometry), Springer, 2003.

Tsi00a. B. Tsirelson, *From random sets to continuous tensor products: answers to three questions of W. Arveson*, Preprint, arXiv: math.FA/0001070, 2000.

Tsi00b. ———, *From slightly coloured noises to unitless product systems*, Preprint, arXiv: math.FA/0006165, 2000.

Zac00a. J. Zacharias, *Continuous tensor products and Arveson's spectral algebras*, Mem. Amer. Math. Soc., no. 680, American Mathematical Society, 2000.

Zac00b. ———, *Pure infiniteness of spectral algebras*, J. Funct. Anal. **178** (2000), 381–395.

CLIFFORD ALGEBRAS, RANDOM GRAPHS, AND QUANTUM RANDOM VARIABLES

RENÉ SCHOTT

IECN and LORIA, Université Henri Poincaré-Nancy I,
BP 239, 54506 Vandoeuvre-lès-Nancy, France
E-mail: schott@loria.fr

G. STACEY STAPLES

Department of Mathematics and Statistics,
Southern Illinois University Edwardsville,
Edwardsville, IL 62026-1653, USA
E-mail: sstaple@siue.edu

For fixed $n > 0$, the space of finite graphs on n vertices is canonically associated with an abelian, nilpotent-generated subalgebra of the Clifford algebra $\mathcal{C}\ell_{2n,2n}$, which is canonically isomorphic to the $2n$-particle fermion algebra. Using the generators of the subalgebra, an algebraic probability space of "Clifford adjacency matrices" associated with finite graphs is defined. Each Clifford adjacency matrix is a quantum random variable whose m^{th} moment corresponds to the number of m-cycles in the graph G. Each matrix admits a canonical "quantum decomposition" into a sum of three algebraic random variables: $a = a^{\Delta} + a^{\Upsilon} + a^{\Lambda}$, where a^{Δ} is classical while a^{Υ} and a^{Λ} are quantum. Moreover, within the Clifford algebra context the NP problem of cycle enumeration is reduced to matrix multiplication, requiring no more than n^4 Clifford (geometric) multiplications within the algebra.

Keywords: Clifford algebra, geometric algebra, quantum probability, fermions, graphs, Hamiltonian cycles, self-avoiding walks, random graphs, quantum computing

1. Introduction

Links between quantum probability and graph theory have been explored in a number of works. Hashimoto, Hora, and Obata[1] obtained limit theorems for increasing sequences of graphs (G_n) whose adjacency matrices admit a *quantum decomposition* $A_n = {A_n}^{+} + {A_n}^{-}$. Examples include Cayley graphs,

Johnson graphs, and distance-regular graphs.

Obata[2] uses this approach to focus on star graphs, which are obtained by gluing together the common origins of a finite number of copies of a given graph. The adjacency matrices of star graphs admit a quantum decomposition of the form $A_n = {A_n}^+ + {A_n}^- + {A_n}^\circ$. Star graphs are of particular interest because they are related to Boolean independence in quantum probability.

Homogeneous trees are also of interest in quantum probability. These are related to the free independence of Voiculescu[3].

Comb graphs, which provide models of Bose-Einstein condensation, are related to monotone independence discovered by Lu[4] and Muraki[5]. Accardi, Ben Ghorbal, and Obata[6] computed the vacuum spectral distribution of the comb graph by decomposing the adjacency matrix into a sum of monotone independent random variables.

Another work of interest is that of Franz Lehner[7], who investigated the relationships among non-crossing partitions, creation and annihilation operators, and the cycle cover polynomial of a graph. In that work, the cycle indicator polynomials of particular digraphs are used to understand the partitioned moments and cumulants occurring in Fock spaces associated to characters of the infinite symmetric group of Bożejko and Guţă[8].

In contrast to the works cited above, the philosophy of the current work is to begin with an arbitrary finite graph and then to construct an associated algebraic probability space in which the moments of random variables reveal information about the graph. The graph need possess no particular relationship to notions of independence or Fock spaces.

Let $G = (V, E)$ denote a finite graph with vertices V and edges $E \subset V \times V$. Letting $n = |V|$, the vertices of G can be represented by unit coordinate vectors in $\mathbb{R}^n$.

Recalling the inner product $\langle u, v\rangle = u^\dagger v$ in $\mathbb{R}^n$, and letting A denote the adjacency matrix of G, a well known result in graph theory states that $\langle x_0, A^k x_0\rangle$ corresponds to the number of closed k-walks based at vertex $x_0 \in V$.

We are interested in the related problem of recovering the *k-cycles* based at any vertex x_0. This can be done for any finite graph with the methods described here.

Clifford algebras and quantum logic gates have been discussed in works by W. Li[9] and Vlasov[10]. The role of Clifford algebras in quantum computing has been considered in works by Havel and Doran[11], Matzke[12], and others.

In the Clifford algebra context, i.e., in terms of the number of multi-

plications performed in the algebra, the computational complexity of enumerating the Hamiltonian cycles in a graph on n vertices is $O(n^4)$. In this context, some graph problems are moved from complexity class NP into class P.

2. Clifford Adjacency Matrices

Definition 2.1. Let ${\mathcal{C}\ell_n}^{\text{nil}}$ be the associative algebra generated by commuting nilpotents $\{e_{\{i\}}\}_{1\leq i\leq n}$ along with the unit scalar $e_\emptyset = 1 \in \mathbb{R}$. In particular, for $1 \leq i, j \leq n$, the generators of ${\mathcal{C}\ell_n}^{\text{nil}}$ satisfy

$$e_{\{i\}}e_{\{j\}} = e_{\{j\}}e_{\{i\}}, \text{ and} \tag{1}$$

$$e_{\{i\}}e_{\{i\}} = 0. \tag{2}$$

Definition 2.2. For $n > 0$ and nonnegative integers p, q satisfying $p + q = n$, the Clifford algebra $\mathcal{C}\ell_{p,q}$ is defined as the associative algebra generated by the unit scalar $\mathbf{e}_\emptyset = 1$ and the collection $\{\mathbf{e}_i\}_{1\leq i\leq n}$ subject to multiplication rules

$$\mathbf{e}_i\,\mathbf{e}_j + \mathbf{e}_j\,\mathbf{e}_i = \begin{cases} 0 & \text{if } i \neq j, \\ 2 & \text{if } i = j \leq p, \\ -2 & \text{if } i = j > p. \end{cases} \tag{3}$$

It can be shown that for $n > 0$, the Clifford algebra $\mathcal{C}\ell_{2n,2n}$ is canonically isomorphic to the $2n$-particle fermion algebra. Using this correspondence, the algebra ${\mathcal{C}\ell_n}^{\text{nil}}$ is generated within $\mathcal{C}\ell_{2n,2n}$ by $e_\emptyset = 1$ along with the set $\{e_{\{i\}}\}_{1\leq i\leq n}$, where $e_{\{i\}} = {f_i}^+{f_{n+i}}^+$. Here f_i^+ denotes the i^{th} fermion creation operator.

In terms of Pauli matrices, the generators of ${\mathcal{C}\ell_n}^{\text{nil}}$ can be written as

$$e_{\{i\}} = \sigma_0^{\otimes(i-1)} \otimes (\sigma_x + i\sigma_y) \otimes \sigma_0^{\otimes(n-i)}. \tag{4}$$

Remark 2.1. Writing the generators $\{e_{\{i\}}\}_{1\leq i\leq n}$ in terms of fermion annihilation operators is an acceptable alternative.

For each $n > 0$, the algebra ${\mathcal{C}\ell_n}^{\text{nil}}$ has dimension 2^n and is spanned by unit multi-vectors, or *blades*, of the form

$$e_{\underline{i}} = \prod_{\iota\in\underline{i}} e_{\{\iota\}}, \tag{5}$$

where $\underline{i} \in 2^{[n]}$ is a multi-index in the power set of $[n] = \{1, 2, \ldots, n\}$. In other words, blades are indexed by *subsets* of the n-set.

Let $u \in \mathcal{C}\ell_n{}^{\mathrm{nil}}$, and observe that u has the canonical expansion

$$u = \sum_{\underline{i} \in 2^{[n]}} u_{\underline{i}}\, e_{\underline{i}}, \tag{6}$$

where $u_{\underline{i}} \in \mathbb{R}$ for each $\underline{i}$.

Definition 2.3. Define the *dual* $u^\star$ of $u \in \mathcal{C}\ell_n{}^{\mathrm{nil}}$ by

$$u^\star = \sum_{\underline{i} \in 2^{[n]}} u_{\underline{i}}\, e_{[n]\setminus \underline{i}}. \tag{7}$$

Definition 2.4. The *scalar sum evaluation* of u is defined by

$$\varphi(u) = \sum_{\underline{i} \in 2^{[n]}} u_{\underline{i}}. \tag{8}$$

Remark 2.2. The scalar sum evaluation is equivalent to the *Clifford 1-norm* $\|u\|_1 = \sum_{\underline{i} \in 2^{[n]}} |u_{\underline{i}}|$ when all the coefficients in the expansion are non-negative.

Lemma 2.1. *Let $u, v \in \mathcal{C}\ell_n{}^{\mathrm{nil}}$, and let $\alpha \in \mathbb{R}$. Let $\left\langle \mathcal{C}\ell_n{}^{\mathrm{nil}} \right\rangle_+$ denote those elements of $\mathcal{C}\ell_n{}^{\mathrm{nil}}$ having strictly non-negative coefficients. Then, the scalar sum evaluation $\varphi : \mathcal{C}\ell_n{}^{\mathrm{nil}} \to \mathbb{R}$ satisfies*

$$\varphi(1) = 1\,, \ \textit{and} \tag{9}$$

$$\varphi(\alpha\, u + v) = \alpha\, \varphi(u) + \varphi(v). \tag{10}$$

Moreover,

$$u \in \left\langle \mathcal{C}\ell_n{}^{\mathrm{nil}} \right\rangle_+ \Rightarrow \varphi(u^\star\, u) \geq 0. \tag{11}$$

Proof. The result follows immediately from the definitions and is omitted.

Definition 2.5. Let $\{e_{\{i\}}\}_{1 \leq i \leq n}$ denote the orthonormal nilpotent generators of $\mathcal{C}\ell_n{}^{\mathrm{nil}}$. Associated with any finite graph $G = (V, E)$ on n vertices is a *Clifford adjacency matrix* a defined by

$$a_{ij} = \begin{cases} e_{\{j\}} & \text{if } (v_i, v_j) \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Proposition 2.1. *Let a denote the Clifford adjacency matrix of a graph G on n vertices. Let x_0 represent any fixed vertex in the graph. Let X_m denote the number of m-cycles based at x_0 in G. Then,*

$$\varphi\left(\langle x_0, a^m x_0 \rangle\right) = \vartheta X_m, \tag{13}$$

where

$$\vartheta = \begin{cases} 2 & \textit{if the graph is undirected,} \\ 1 & \textit{otherwise.} \end{cases} \tag{14}$$

Proof. Proof is by induction on m. When $m = 2$,

$$\left(a^2\right)_{ii} = (a \times a)_{ii} = \sum_{\ell=1}^{n} a_{i\ell} a_{\ell i}. \tag{15}$$

By construction of the Clifford adjacency matrix,

$$a_{i\ell} \equiv \text{s.a. 1-walk } v_i \to v_\ell, \text{ and} \tag{16}$$

$$a_{\ell i} \equiv \text{ s.a. 1-walk } v_\ell \to v_i, \tag{17}$$

where s.a. is short for self-avoiding. Hence the product of these terms corresponds to 2-cycles $v_i \to v_i$.

Now assuming the proposition holds for m and considering the case $m + 1$,

$$\left(a^{m+1}\right)_{ii} = (a^m \times a)_{ii} = \sum_{\ell=1}^{n} (a^m)_{i\ell}\, a_{\ell i}. \tag{18}$$

Considering a general term of the sum,

$$(a^m)_{i\ell} = \sum_{\text{s.a. } m\text{-walks } w_m : v_i \to v_\ell} w_m, \text{ and} \tag{19}$$

$$a_{\ell i} = \begin{cases} \text{s.a. 1-walk } w_1 : v_\ell \to v_i, & \text{if } (v_\ell, v_i) \in E \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

It should then be clear that terms of the product

$$(a^m)_{i\ell}\, a_{\ell i} \tag{21}$$

are nonzero if and only if they correspond to self-avoiding $(m + 1)$-walks $v_i \to v_\ell \to v_i$. Summing over all vertices v_ℓ gives the total number of $m + 1$-cycles based at v_i.

The correction factor $\vartheta = 2$ in the undirected case accounts for the two possible orientations of each cycle. An orientation correction is not needed when the graph is directed. □

3. Quantum Random Variables

Let $\mathcal{M}$ denote the *-algebra of $n \times n$ Clifford matrices with involution defined by $a^* = (a^\star)^\dagger$. Here, * denotes the dual defined by linear extension of $e_{\underline{i}}{}^\star = e_{[n]\setminus\underline{i}}$, and † denotes the matrix transpose.

Definition 3.1. The *norm* of $a \in \mathcal{M}$ is defined by

$$\|a\|^2 = \int \mathrm{tr}(a^* a)\, de_n \cdots de_1. \tag{22}$$

Remark 3.1. This can be regarded as the Frobenius norm of a Clifford matrix[13].

Let $\varphi(u)$ denote the scalar sum evaluation of $u \in \mathcal{C}\ell_n{}^{\mathrm{nil}}$, and define $\psi : \mathcal{M} \to \mathbb{R}$ by

$$\psi(a) = \frac{\varphi\left(\mathrm{tr}(a)\right)}{n}. \tag{23}$$

Direct computation shows that ψ is a linear mapping and that $\psi(1_{\mathcal{M}}) = 1$. The positivity requirement for states: $\psi(a^*\, a) \geq 0$, does not generally hold in the *-algebra, however.

The set of Clifford adjacency matrices $\mathcal{A} \subset \mathcal{M}$ together with matrix multiplication defined in terms of the geometric product constitutes a multiplicative semigroup with involution $*$. On this semigroup, the map ψ satisfies the positivity requirement, and $(\mathcal{A}, \psi)$ is a quantum probability space.

The proofs of the following propositions are virtually identical and follow naturally from Proposition 2.1.

Proposition 3.1 (Cycles in finite graphs). *Let a be the Clifford adjacency matrix of an n-vertex graph, G, and let X_m denote the number of m-cycles in G. Then,*

$$\psi(a^m) = \frac{\vartheta\, m}{n} X_m, \tag{24}$$

where ϑ is defined as in (14)*. In other words, a is a quantum random variable whose m^{th} moment in the state ψ corresponds to the number of m-cycles occurring in the graph associated with a.*

Proposition 3.2 (Time-homogeneous random walks on graphs). *Let M denote a stochastic matrix corresponding to an n-state Markov chain, and let τ denote the Clifford stochastic matrix defined by*

$$\tau_{ij} = M_{ij}\, e_{\{j\}}\,, \tag{25}$$

where $e_{\{j\}}$ is a nilpotent generator of the abelian algebra $\mathcal{C}\ell_n{}^{\mathrm{nil}}$.

Let the state ψ be defined as in (23)*, and let ϑ be defined as in* (14)*. Define the random variable w_m on the space of m-step walks by*

$$w_m = \begin{cases} 1 & \text{if the } m\text{-walk forms a cycle,} \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

Then,

$$\psi(\tau^m) = \frac{\vartheta\, m}{n} \mathbb{E}(w_m). \tag{27}$$

In other words, τ is a quantum random variable whose m^{th} moment corresponds to the expected number of m-walks forming cycles in the n-state Markov chain associated with matrix M.

Proposition 3.3 (Cycles in random graphs). *Consider a random directed graph $G = (V, E)$ on n vertices, corresponding to edge-existence matrix A, defined by*

$$A_{ij} = \mathbb{P}\left((v_i, v_j) \in E\right). \tag{28}$$

It is assumed that the probabilities are pairwise-independent. Let ξ denote the Clifford matrix defined by

$$\xi_{ij} = A_{ij}\, e_{\{j\}}, \tag{29}$$

where $e_{\{j\}}$ is a nilpotent generator of the abelian algebra ${\mathcal{C}\ell_n}^{\text{nil}}$.

Let the state ψ be defined as in (23)*, let ϑ be defined as in* (14)*, and define the random variable z_m as the number of m-cycles occurring in the graph. Then,*

$$\psi(\xi^m) = \frac{\vartheta\, m}{n} \mathbb{E}(z_m). \tag{30}$$

That is, ξ is a quantum random variable whose m^{th} moment corresponds to the expected number of m-cycles occurring in the graph.

4. Quantum Decomposition of Clifford Matrices

In the works of Hashimoto, Hora, and Obata[1] and Obata[2], fixing a vertex v_0 in a finite graph induces a stratification of all the vertices by associating each vertex with the length of the shortest path linking it with v_0. This stratification is then used to define a quantum decomposition of the graph's adjacency matrix.

The Clifford adjacency matrix of a graph also admits a quantum decomposition as the sum of three algebraic random variables: one classical,

and two quantum. The decomposition considered here differs from that of Hashimoto, et al.

The Clifford adjacency matrix a of any finite graph has the canonical quantum decomposition

$$a = a^{\Delta} + a^{\Lambda} + a^{\Upsilon}, \tag{31}$$

where

$$a^{\Delta}{}_{ij} = \delta_{ij} a_{ij}, \tag{32}$$

$$a^{\Upsilon}{}_{ij} = \theta_{ij} a_{ij}, \tag{33}$$

$$a^{\Lambda}{}_{ij} = (1 - \theta_{ij}) a_{ij}. \tag{34}$$

Here δ_{ij} is the Kronecker delta function, and θ_{ij} is the *ordering symbol* defined by

$$\theta_{ij} = \begin{cases} 1 & \text{if } i < j \\ 0 & \text{otherwise.} \end{cases} \tag{35}$$

The matrix a^{Δ} resides in the semigroup $\mathcal{D}$ of diagonal Clifford adjacency matrices with matrix multiplication, which is commutative. Hence, a^{Δ} is a classical algebraic random variable in the space $(\mathcal{D}, \psi)$.

The matrices a^{Λ} and a^{Υ} reside in the semigroups $\mathcal{L}$ and $\mathcal{U}$ of lower- and upper-triangular Clifford adjacency matrices with matrix multiplication, respectively, which are noncommutative. Hence, a^{Λ} and a^{Υ} are quantum algebraic random variables in the spaces $(\mathcal{L}, \psi)$ and $(\mathcal{U}, \psi)$, respectively.

5. Extending to Infinite Dimensions

Denote by $\left(\mathcal{C}\ell_n{}^{\text{nil}}\right)^n$ the n-fold Cartesian product of $\mathcal{C}\ell_n{}^{\text{nil}}$. Let $\{x_i\}_{1 \le i \le n}$ denote the collection of basis coordinate vectors of the form $x_i = (0, \ldots, 0, \underbrace{1}_{i^{\text{th}} \text{ pos.}}, 0, \ldots, 0)$ such that for $\vec{u} = (u_1, u_2, \ldots, u_n) \in \left(\mathcal{C}\ell_n{}^{\text{nil}}\right)^n$, one has

$$\langle \vec{u}, x_i \rangle = u_i. \tag{36}$$

The following finite-dimensional result follows from the previous results and definitions.

Proposition 5.1 (Generating function method). *Fix $n > 0$, and let G be a finite graph on n vertices with associated Clifford adjacency matrix a. Coordinate basis vectors $x_1, \ldots, x_n$ are also associated with the vertices*

of G by construction of a. Let X_k denote the number of k-cycles based at an arbitrary vertex x_0 in G. Then,

$$\vartheta X_k = \varphi\left(\frac{\partial^k}{\partial t^k}\left\langle x_0, \exp(ta)\, x_0\right\rangle\right)\Big|_{t=0}. \tag{37}$$

Define the infinite-dimensional nilpotent-generated algebra $\mathcal{C}\ell^{\mathrm{nil}}$ by

$$\mathcal{C}\ell^{\mathrm{nil}} = \bigoplus_{n=1}^{\infty} \mathcal{C}\ell_n{}^{\mathrm{nil}}. \tag{38}$$

Define the space $\mathcal{V}\left(\mathcal{C}\ell^{\mathrm{nil}}\right)$ by

$$\mathcal{V}\left(\mathcal{C}\ell^{\mathrm{nil}}\right) = \bigoplus_{n=1}^{\infty} \left(\mathcal{C}\ell_n{}^{\mathrm{nil}}\right)^n. \tag{39}$$

The algebra $\mathcal{M}_n$ of $n \times n$ matrices with entries in $\mathcal{C}\ell_n{}^{\mathrm{nil}}$ is a $*$-algebra of operators on $\left(\mathcal{C}\ell_n{}^{\mathrm{nil}}\right)^n$. Define the $*$-algebra $\mathcal{M}$ of operators on $\mathcal{V}\left(\mathcal{C}\ell^{\mathrm{nil}}\right)$ by

$$\mathcal{M} = \bigoplus_{i=1}^{\infty} \mathcal{M}_n. \tag{40}$$

Remark 5.1. Note that with inner product defined by

$$(a, b) = \int \mathrm{tr}(a^* b)\, de_{\{n\}}, \ldots, de_{\{1\}}, \tag{41}$$

$\mathcal{M}$ is a Hilbert space.

Denote by $\mathcal{A}_n$ the semigroup of Clifford adjacency matrices associated with finite graphs on n vertices under matrix multiplication. Denote by $\mathcal{A}$ the collection of operators in $\mathcal{M}$ that represent Clifford adjacency matrices. In other words,

$$\mathcal{A} = \{a \in \mathcal{M} : (\exists \xi \in \mathcal{A}_n)(\forall i \in \{1, \ldots, n\})\,[\langle x_i, a x_i\rangle = \langle x_i, \xi x_i\rangle]\}. \tag{42}$$

Observe that φ is a state on the semigroup $\mathcal{A}$.

Define a sequence of operators $a_1, a_2, \ldots \in \mathcal{A}$ such that for each n, a_n is an operator on $\left(\mathcal{C}\ell_n{}^{\mathrm{nil}}\right)^n$. The sequence $\{a_n\}$ will be said to converge to the operator a if for each $k \geq 0$ and any coordinate basis vector x_0, the following equation holds:

$$\lim_{n\to\infty} \varphi\left(\left\langle x_0, {a_n}^k x_0\right\rangle\right) = \varphi\left(\left\langle x_0, a^k x_0\right\rangle\right). \tag{43}$$

Denote this convergence by $a_n \xrightarrow{w} a$, and notice that this implies

$$\lim_{n\to\infty} \varphi\left(\left\langle x_0, \exp(a_n)\, x_0\right\rangle\right) = \varphi\left(\left\langle x_0, \exp(a)\, x_0\right\rangle\right). \tag{44}$$

Proposition 5.2 (Ascending chains). *Let $a_1, a_2, \ldots \in \mathcal{A}$ be a sequence of operators such that for each $n > 0$, a_n represents the Clifford adjacency matrix associated with a finite graph on n vertices. Let t be a real-valued parameter.*

Let $X_k(n)$ denote the number of k-cycles based at vertex x_0 in the n^{th} graph of the sequence. Then,

$$\vartheta X_k(n) = \varphi\left(\frac{\partial^k}{\partial t^k}\left\langle x_0, \exp\left(t a_n\right) x_0\right\rangle\right)\Big|_{t=0}. \tag{45}$$

Assuming $a_n \xrightarrow{w} a$ as $n \to \infty$ and fixing $k \geq 1$, one finds

$$\vartheta \lim_{n\to\infty} X_k(n) = \varphi\left(\frac{\partial^k}{\partial t^k}\left\langle x_0, \exp\left(ta\right) x_0\right\rangle\right)\Big|_{t=0}. \tag{46}$$

6. Applications

Letting X_k denote the number of k-cycles contained in a simple random graph on n vertices, the Clifford adjacency matrix method has been used to recover higher moments of ${X_k}$[14].

The Clifford adjacency matrix method has been used to give a graph-theoretic construction of iterated stochastic integrals[15]. In particular, stochastic integrals of $L^2 \otimes \mathcal{C}\ell_{p,q}$-valued processes are defined and recovered with this approach.

A graph-free Clifford-algebraic construction of iterated stochastic integrals of classical processes has also been developed using sequences within a growing chain of fermion algebras[16].

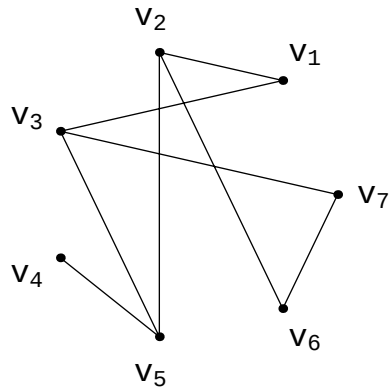
Clifford-algebraic generating functions for Stirling numbers of the second kind, Bell numbers, and Bessel numbers have been defined[17].

Example 6.1. In Figure 1, the number of 5-cycles contained in a randomly generated, 7-vertex graph are recovered by examining the trace of a^5.

6.1. *Complexity*

Within the ${\mathcal{C}\ell_n}^{\text{nil}}$ context, the problem of cycle enumeration in finite graphs is reduced to matrix multiplication. Hence, enumerating the k cycles in a graph on n vertices requires kn^3 *algebra multiplications*. Consequently, fixing k and allowing n to vary, enumerating the k-cycles in a finite graph on n vertices has complexity (in the ${\mathcal{C}\ell_n}^{\text{nil}}$-algebraic context) equal to $O(n^3)$.

Let $Clops(n)$ denote the number of algebra multiplications performed by an algorithm processing a data set of size n. An algorithm will be said



$$A= \begin{pmatrix} 0 & e_{\{2\}} & e_{\{3\}} & 0 & 0 & 0 & 0 \\ e_{\{1\}} & 0 & 0 & 0 & e_{\{5\}} & e_{\{6\}} & 0 \\ e_{\{1\}} & 0 & 0 & 0 & e_{\{5\}} & 0 & e_{\{7\}} \\ 0 & 0 & 0 & 0 & e_{\{5\}} & 0 & 0 \\ 0 & e_{\{2\}} & e_{\{3\}} & e_{\{4\}} & 0 & 0 & 0 \\ 0 & e_{\{2\}} & 0 & 0 & 0 & 0 & e_{\{7\}} \\ 0 & 0 & e_{\{3\}} & 0 & 0 & e_{\{6\}} & 0 \end{pmatrix}$$

```
In[49]:= Simplify[Tr[MatrixPower[A, 5]]] / 5 / 2

Out[49]= e{1,2,3,6,7} + e{2,3,5,6,7}
In[42]:= ScalarSum[%]

Out[42]= 2
```

Fig. 1. Counting the 5-cycles in a graph on 7 vertices.

to have $\mathcal{C}\ell$-complexity $O(f(n))$ if for every $k \in \mathbb{N}$, $\exists c \in \mathbb{R}$ such that

$$n \geq k \Rightarrow |Clops(n)| \leq c|f(n)|. \tag{47}$$

The following lemmas reflect the time complexity of computing powers of matrices.

Lemma 6.1. *Fixing k and allowing n to vary, enumerating the k-cycles in a finite graph on n vertices has $\mathcal{C}\ell$-complexity $O(n^3)$.*

Lemma 6.2. *Enumerating the Hamiltonian cycles in a finite graph on n vertices has $\mathcal{C}\ell$-complexity $O(n^4)$.*

These lemmas emphasize the potential power of a computer architecture

based on the ${\mathcal{C}\ell_n}^{\text{nil}}$ algebra. Given a device in which the product of any pair $u, v \in \mathcal{C}\ell^{\text{nil}}$ could be computed in polynomial time, the NP class problem of cycle enumeration would be reduced to class P.

Bibliography

1. Y. Hashimoto, A. Hora and N. Obata, Central limit theorems for large graphs: method of quantum decomposition, *J. Math. Phys.* **44**, 71 (2003).
2. N. Obata, Quantum probabilistic approach to spectral analysis of star graphs, *Interdisciplinary Information Sciences* **10**, 41 (2004).
3. D. Voiculescu, K. Dykema and A. Nica, *Free Random Variables, CRM Monograph Series* (American Mathematical Society, Providence, RI, 1992).
4. Y.-G. Lu, An interacting free Fock space and the arcsine law, *Probab. Math. Stat.* **17**, 149 (1997).
5. N. Muraki, Noncommutative Brownian motion in monotone Fock space, *Commun. Math. Phys.* **183**, 557 (1997).
6. L. Accardi, A. Ben Ghorbal and N. Obata, Monotone independence, comb graphs and Bose-Einstein condensation, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **7**, 419 (2004).
7. F. Lehner, Cumulants in noncommutative probability theory III. Creation and annihilation operators on Fock spaces, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **8**, 407 (2005).
8. M. Bożejko and M. Guţă, Functors of white noise associated to characters of the infinite symmetric group, *Comm. Math. Phys.* **229**, 209 (2002).
9. W. Li, Clifford algebra and quantum logic gates, *Advances in Applied Clifford Algebras* **14**, 225 (2004).
10. A. Vlasov, Clifford algebras and universal sets of quantum gates, *Phys. Rev. A* **63**, p. 054302 (2001).
11. T. Havel and C. Doran, Geometric algebra in quantum information processing, *S.J. Lomonaco and H.E. Brandt, eds. Quantum Computation and Information, Contemporary Mathematics* **305**, 81 (2002).
12. D. Matzke, Quantum computation using geometric algebra, Ph.D. dissertation in the Department of Electrical Engineering, University of Texas at Dallas, 2002.
13. G.S. Staples, Norms and generating functions in Clifford algebras, `http://www.siue.edu/~sstaple/index_files/froben20051126.pdf`, (2005).
14. R. Schott and G.S. Staples, Nilpotent adjacency matrices and random graphs, `http://www.siue.edu/~sstaple/index_files/RndGrphs.pdf`, (2006).
15. G.S. Staples, Graph-theoretic approach to stochastic integrals with Clifford algebras, *Journal of Theoretical Probability* (To appear).
16. G.S. Staples, Multiple stochastic integrals via sequences in geometric algebras, `http://www.siue.edu/~sstaple/index_files/Stochmsr3.pdf`, (2005).
17. R. Schott and G.S. Staples, Partitions and Clifford algebras, `http://www.loria.fr/~schott/bessel2march06.pdf`, (2006).

THE SET OF DENSITY OPERATORS MODELLED ON AN ORLICZ SPACE

R. F. STREATER

Department of Mathematics, King's College London, Strand, London WC2R 2LS, England
E-mail: raymond.streater@kcl.ac.uk

Suppose that the quantum Hilbert space of vectors is separable. We consider a subset of density operators which happen to have the property that their entropy is finite. We furnish the set with a norm coming from the Orlicz theory in which the Young function is defined on small quadratic forms, relative to the positive self-adjoint operator defined as minus the logarithm of the state. We argue that the collection of states is a Banach manifold, generalising to quantum theory the manifold introduced in classical estimation theory by Pistone and Sempi in 1995.

1. Introduction and Motivation

Suppose that $\mathcal{H}$ is an infinite-dimensional Hilbert space, and $\mathcal{B}(\mathcal{H})$ is the W^*-algebra of bounded operators on $\mathcal{H}$, and Σ_+ is the set of all normal states on $\mathcal{B}(\mathcal{H})$. In elementary equilibrium statistical physics, we assume that any physical system is of finite size, and its state $\rho \in \Sigma_+$ has finite entropy:

$$S(\rho) := -\mathrm{Tr}(\rho \log \rho) < \infty.$$

The set Σ spanned by Σ_+ possesses the trace-norm; however, this is not a good norm, since any neighbourhood of a density operator ρ contains a dense set of states of infinite entropy. Such a state cannot be near ρ in any physical sense: we expect any state arising in the dynamics to have entropy less than that of the equilibrium state to which it converges at large time, and this is finite for all finite systems.

We therefore seek a norm which will exclude states of infinite entropy from the space. Thus we seek a quantum generalization of the work of Pistone and Sempi[10], who developed a theory of non-parametric estimation in classical probability. We summarise this next.

2. The Work of Pistone and Sempi

2.1. *Young Functions*

A Young function is a convex map $\Phi : \mathbf{R} \to \mathbf{R}^+ \cup \{+\infty\}$ such that

(1) $\Phi(x) = \Phi(-x)$ for all $x \in \mathbf{R}$.
(2) $\Phi(0) = 0$.
(3) $\lim_{x\to\infty} \Phi(x) = +\infty$.

Given Φ, we define the *epigraph* of Φ to be the set $E(\Phi)$:

$$E(\Phi) := \{(x, y) : y \geq \Phi(x)\}.$$

Then $E(\Phi)$ is convex and

$E(\Phi)$ is closed if and only if Φ is lower semi-continuous.

More, if $E(\Phi)$ is not closed, then its closure is the epigraph of a Young function, which is the lower-continuous version of Φ.

Here are four examples of Young functions.

- $\Phi_1(x) = \cosh x - 1$.
- $\Phi_2(x) = e^{|x|} - |x| - 1$.
- $\Phi_3(x) = (1 + |x|)\log(1 + |x|) - |x|$.
- $\Phi^p(x) = |x|^p, 1 \leq p < \infty$.

The *Legendre-Fenchel transform*, $\Phi^*(y)$, of a Young function Φ is defined for $y \in \mathbf{R}$ by

$$\Phi^*(y) = \sup_x \{xy - \Phi(x)\}.$$

This is a Young function, which is always lower-semi-continuous. Examples: $\Phi_2 = \Phi_3^*$, and $\Phi^p = (\Phi^q)^*$ when $1/p + 1/q = 1$. If Φ is a lower-semi-continuous Young function, then $(\Phi^*)^* = \Phi$. Because of this, it is convenient (without loss of generality in the context in which we work) to limit ourselves to lower-semi-continuous Young functions. We say two Young functions Φ and Ψ are *equivalent* if there exist two positive numbers, say c, C, and $x_0 > 0$ such that if $x > x_0$, we have

$$\Phi(cx) \leq \Psi(x) \leq \Phi(Cx), \text{ if } x > x_0.$$

We then write $\Phi \equiv \Psi$. For example, $\Phi_1 \equiv \Phi_2$. Duality is an operation on equivalence classes; that is, $\Phi \equiv \Psi \Rightarrow \Phi^* \equiv \Psi^*$.

Finally, we describe the Δ_2-class of Young functions. We say that a Young function Φ lies in the Δ_2-class, or satisfies the Δ_2-condition, if and

only if there exists $\kappa > 0$ and $x_0 > 0$ such that

$$\Phi(2x) \leq \kappa\Phi(x) \qquad \text{for all } x \geq x_0.$$

For example, Φ^p and Φ_3 satisfy the Δ_2-condition, but Φ_1 and Φ_2 do not. Clearly, if two Young functions are equivalent, then either both or neither lies in the Δ_2-class.

2.2. *Orlicz Theory*

Given a space X, a σ-ring $\mathcal{B}$ of subsets, called the measurable sets of X, and a measure μ on $(X, \mathcal{B})$, then the *Orlicz class relative to the Young function* Φ is the set of functions $u : X \to \mathbf{R}$ obeying

$$u \in \mathcal{B} \qquad \text{and} \qquad \int_X \Phi(u(x))\mu dx < \infty.$$

This is a convex set; it is a vector space if and only if $\Phi \in \Delta_2$. If not, then the convex span of the Orlicz class is the *Orlicz space relative to* Φ; this can be used to define a Banach space when furnished with the so-called Luxemburg norm; this is the following: let $a > 0$ be chosen; then

$$\|u\|_L = \inf_{r>0} \left\{ r : \int_X \Phi\left(\frac{u(x)}{r}\right) < a \right\}$$

defines a norm on the space of continuous functions lying in the Orlicz space, and a semi-norm more generally on the Orlicz space. More, for different positive a, all these norms are equivalent. The space of functions of finite Luxemburg norm, modulo the set of zero norm, is a Banach space: it is complete. There is an analogue of Hölder's inequality:

$$\int_X |uv|\mu(dx) \leq 2\|u\|_L \, \|v\|_L^* .$$

Note the extra factor 2. It might be mentioned that Orlicz gave a different definition of the Orlicz norm; this turned out to be equivalent to that later used by Luxemburg, which is a generalisation to infinite dimensions of the work of Minkowski, who called the quantities entering his theory the *gauges*.

By way of example, by taking $X = \mathbf{R}$ and μ the Lebesgue measure, and $\mathcal{B}$ the Borel σ-ring, we get Zygmund space if we use Φ_1 or Φ_2; however, by using Φ_3 we get a different space, namely the space called $L \log L$, while Φ^p gives a norm equivalent to

$$\|u\|_p := \left(\int |u(x)|^p dx \right)^{1/p} .$$

One can show that equivalent Young functions leads to equivalent norms.

Now, the Zygmund space is not separable, while the spaces $L \log L$ and L^p are. This is because Φ_1 and Φ_2 do not obey the Δ_2 condition, while Φ_3 and Φ^p do. Indeed, the Banach space obtained from the Orlicz space relative to Φ in this way is separable if and only if Φ obeys Δ_2.

2.3. *Non-parametric Estimation Theory*

Suppose that we have two random variables, X and Y say, and we have a bivariate plot of data. Suppose that this tells us the experimental value y_1 of Y when $X = x_1$ is observed, the value y_2 of Y when $X = x_2$ is observed, and so on up to $X = x_k$. The elementary theory of statistics tells us how to find the linear regression of one random variable, Y, on the other, here X. The answer gives us the "best" linear function $Y = aX + b$; it is found by minimising (by varying a and b) the sum of the square errors:

$$\sum_{i=1}^{k} (ax_i + b - y_i)^2.$$

Good texts mention that if we had done the calculation the other way round, to find X as a linear function of Y, we would (in general) get a different answer from what we get by solving the previous one for X as a function of Y; so the new answer has a different name: the regression of X on Y. The reason is that choosing Y and its difference-squared puts a metric on the space of probability distributions on the line which depends on the choice made; choosing X and its difference-squared puts a different metric on the same space, the line. Fisher noticed this, and obtained a coordinate-free metric on the space, now called the Fisher metric. Fisher also generalised this to the case of n linearly independent variables.

Pistone and Sempi[10] wanted to generalise this to infinite-dimensions; this is not an easy problem. They start with a measure space $(X, \mathcal{B}, \mu)$, where μ is not necessarily a probability measure; its purpose is to specify the impossible events, that is, those $A \in \mathcal{B}$ for which $\mu(A) = 0$. The states are determined by a probability density $f(x)$ such that

(1) $f(x) > 0$ μ-almost everywhere.
(2) $\mathbf{E}_\mu[f] := \int_X f(x)\mu(dx) = 1$.

Alternatively, one may postulate that the state ν under consideration is equivalent to μ. Then the Radon-Nikodym theorem tells us that $f(x) = d\nu/d\mu$. So far, no smoothness conditions are postulated for f.

Let f_0 obey the two conditions (1) and (2). We seek a norm on the "states" f such that if $\|f - f_0\|$ is small enough, then the free energy of $f\mu$ is finite. Some observers of physics seem to be saying that the use of the terms "entropy" and "free energy" in so many fields is due to habit, and might lead to unnecessary limitations on the theory. However, the free energy, defined by the function

$$\psi_f(u) := \log\left\{\int_X f_0 e^u \mu(dx)\right\} \tag{1}$$

$$= \log\left\{\mathbf{E}_{f_0\mu}\left[e^u\right]\right\} \tag{2}$$

exists under the assumption that f is positive except for a set of measure zero, and is not related to an energy conserved in time. Thus, the use of states of the form

$$f = f_0 \exp\left\{u - \psi_{f_0}(u)\right\} \tag{3}$$

where ψ_{f_0} is taken to be finite, expresses merely that we are interested in some family of states in eq. (3) which have a non-zero probability density. To specify the family more clearly, Pistone and Sempi limit the choice of f to be such that the corresponding u obeys: there exists $\epsilon > 0$ such that

$$\psi_{\mu f_0}(\lambda u) = \log\left\{\mathbf{E}_{\mu f_0}\left[e^{-\lambda u}\right]\right\} \tag{4}$$

is finite for all real λ with $|\lambda| \leq \epsilon$. This implies that all moments, $\mathbf{E}_{\mu f_0}\left[u^k\right]$, $k = 1, 2, \ldots$, of the random variables u are finite, and that the moment generating function given by (4) is analytic in λ in the circle $|\lambda| \leq \epsilon$.

Such a u is said to belong to the *Cramér class* of random variables. Pistone and Sempi show that the Cramér class is complete when it is furnished with the Luxemburg norm got from the Young function Φ_1:

$$\|u\|_L := \inf_r \left\{r > 0 : \mathbf{E}_\mu\left[f_0 \cosh\frac{u}{r}\right] - 1 < 1\right\}. \tag{5}$$

They show that the map

$$u \mapsto \exp\left\{u - \psi_{f_0}(u)\right\} f_0 \tag{6}$$

maps the unit sphere of the Cramér class into the class of distributions equivalent to μ. The map is bijective if we limit the u so that $\mathbf{E}_\mu\left[f_0 u\right] = 0$. Such random variables are called *scores*. Since $f_0 > 0$ we may write $f_0 = e^{-h_0}$ and

$$f = \exp\left\{-h_0 - u - \psi_{f_0}(u)\right\}. \tag{7}$$

This has the analogy with equilibrium statistical physics, with the identifications:

h_0 = the "free" Hamiltonian; u = the perturbing potential; $\psi_{f_0}(u)$ = the free energy.

The Fisher metric is

$$\frac{\partial^2 \psi_f(u)}{\partial u_i \partial u_j}.$$

One can construct a Banach manifold, which is the classical information manifold, out of the possible states; the set consists of all the f_0, and the topology (in the neighbourhood of any f_0) is determined by the Luxemburg norm eq. (5). Pistone and Sempi were able to show the equivalence of the norms based at f_0 and f_0' for any point f lying in the overlap of two spheres, one around f_0 and the other around f_0'. Since Φ_1 does not satisfy the Δ_2-condition, this manifold is not separable.

The work of Pistone and Sempi is beginning to be used by other statisticians in non-parametric estimation theory. The question arises, how can we generalise to quantum probability?

3. The quantum information manifold

3.1. *The quantum Fisher metric*

In finite dimensions, quantum probability becomes matrix probability, and the first question is to define what is to be the analogue to the Fisher metric. This problem was formalised by Chentsov and Morozova[3]. In the classical case, a step in a linear stochastic process is determined by a *stochastic map* on the observables. This could be any positive linear map on the set of observables that maps the identity into itself. Then it can be shown that any metric on the space of observables that is reduced by any stochastic map, or left the same, is a multiple of the Fisher metric. The idea is that the stochastic map introduces some error into the calculation; thus whether or not an observed value of the random variable to be estimated is close to the true value should be less clear after the stochastic map than before. So the distance between any two random variables should be less with all stochastic maps, or be left the same. So, in the quantum case, we need the concept of a stochastic map. This is a completely positive linear map, leaving the identity operator unchanged. Then, the quantum generalisation of the Fisher metric should decrease, or stay the same, with any quantum stochastic map.

Unfortunately, there is more than one metric with this property. Chentsov went about classifying the possible ones, a job that was completed for finite dimensions by Petz[9] after Chentsov died. I shall use one

of these, the Bogoliubov-Kubo-Martin metric, which seems to me to be the best. It was advocated by Ingarden *et al.* in[5], and later by Petz and Toth[11]. At the point $\rho := e^{-H}$ of the manifold of states, the BKM-metric between two bounded operators A, B is

$$\langle A, B\rangle_{BKM} := \int_0^1 d\alpha \mathrm{Tr}\left(Ae^{-\alpha H}Be^{(1-\alpha)H}\right). \tag{8}$$

In[14] I show that this makes sense in infinite dimensions if $\rho \in \mathcal{M}$. The metric is not simply related to the norm needed for the topology; we need an Orlicz norm, rather than the Hilbert norm given by the metric.

3.2. *Quantum Orlicz spaces*

Let $\mathcal{H}$ be a separable Hilbert space, and denote by Σ the set of density operators on $\mathcal{H}$, and by Σ_+ the set of faithful density operators. We have suggested[15] that the quantum analogue of the set of density functions f used by Pistone and Sempi might be the set of faithful density operators ρ on $\mathcal{H}$ such that there exists $\beta_0 \in [0, 1)$, depending on ρ, such that ρ^β is of trace-class for all $\beta > \beta_0$. Thus, interesting states lie in the point set $\mathcal{M}$:

$$\mathcal{M} = \bigcup_{0<\beta<1} \mathcal{C}_\beta \cap \Sigma_+. \tag{9}$$

A state in the manifold of states thus lies in some Schatten class $\mathcal{C}_\beta$, in the unfashionable case $\beta < 1$. This is a complete, quasi-normed complex vector space, with quasi-norm given by $\left(\mathrm{Tr}\,\rho^\beta\right)^{1/\beta}$. The topologies defined by two different values of β are not equivalent, and we shall furnish $\mathcal{M}$ with a different topology, given by a quantum Orlicz norm.

We start with a point $\rho_0 \in \mathcal{M}$, and write $\rho_0 = e^{-H_0}$. We see that H_0 is a positive self-adjoint operator. Let X be a quadratic form on $\mathcal{H}$ that is form bounded relative to the quadratic form of H_0. This means that

(1) $\mathrm{Dom}\, X \supseteq \mathrm{Dom}\, H_0^{1/2}$.
(2) There exist real numbers $a \geq 0$ and b such that

$$|X(\phi, \phi)| \leq a\langle H_0^{1/2}\phi, H_0^{1/2}\phi\rangle + b\|\phi\|^2.$$

The parameter a is called a form bound of X relative to H_0, and the infimum of all form bounds a, for which b is anything finite, is a semi-norm:

$$\|X\|_K := \inf\{a : a \text{ is an } H_0 - \text{bound of } X\}. \tag{10}$$

It is well known that if $\|X\|_K < 1$ then there is a unique self-adjoint operator with form domain $\mathrm{Dom}\, H_0^{1/2}$ taking the values of the sum of the forms X and H_0; this operator is bounded below.

In classical information theory, we need to restrict the manifold to the scores. We showed[16] that the following generalised mean is finite, when X is a H-form-small quadratic form, and $\rho \in \mathcal{M}$:

$$\rho \cdot X := \mathrm{Tr}\left(\int_0^1 \rho^t X \rho^{(1-t)} dt\right). \tag{11}$$

We seek a Young function that is defined for such X, which define the convex set of quantum scores.

Work[8] by W. Kunze suggests using a classical Young function, Φ, and defining by operator calculus the function $\Phi(|\tilde{X}\|)$, where $\tilde{X}$ is the rearranged operator defined by a self-adjoint operator X. Al-Rashid and Zegarlinski[1] develop this work, as does Jenkova[6]. This idea does not work for forms. In the next section, I propose the Young function chosen in[14].

4. The Quantum Young Function

Let $\rho := e^{-H}$ be a density operator. For any quadratic form X, defined on $\mathrm{Dom}\, H^{1/2}$, and small-enough relative to the form of H, we define

$$\Phi(X) := \frac{1}{2}\mathrm{Tr}\left(e^{-(H+X)} + e^{-(H-X)}\right) - 1. \tag{12}$$

Then we claim

Theorem 4.1. $\Phi(X)$ obeys

(1) $\Phi(X)$ is finite for all forms X with small-enough H-form bound, say $\|X\|_K < c$.
(2) $X \mapsto \Phi(X)$ is convex on the set of H-small-enough forms.
(3) $\Phi(X) = \Phi(-X)$.
(4) $\phi(0) = 0$ and if $X \neq 0$, then $\Phi(X) > 0$ or $\Phi(X) = \infty$.

This is proved when X is an H-small operator in[14,15]. The map $X \mapsto \mathrm{Tr}\, e^{H+X}$ is convex when extended to H-form-bounded forms X of small-enough H-form bound. We now give a proof of this. H has discrete spectrum, say $\lambda_1 \leq \lambda_2 \leq \ldots$; since $\exp(-H)$ is of trace class, we see that $\lambda_n \to \infty$ as $n \to \infty$. Let $H = \sum_j \lambda_j P_j$ be the spectral resolution of H, and put

$$Q_n := \sum_{j=1}^{n} P_j, n = 1, 2, \ldots.$$

Each normalised eigen-vector, say $\psi_j, j = 1, 2, \ldots$, of H lies in $\mathrm{Dom}\, H^{1/2}$, so we have that $X_n := Q_n X Q_n$ is a bounded operator. Moreover, for any sufficiently small H-form bounded X, we have

$$\langle \psi_j, e^{(H+X)} \psi_j \rangle = \lim_{n\to\infty} \langle \psi_j, e^{-(H+X_n)} \psi_j \rangle. \tag{13}$$

Since exponentials are positive, the traces are absolutely convergent, and we have for any H-form small-enough X and Y, and any $\lambda \in (0, 1)$,

$$\begin{aligned}
\mathrm{Tr}\, e^{-(H+\lambda X+(1-\lambda)Y)} &= \sum_j \langle \psi_j, e^{-(H+\lambda X+(1-\lambda)Y)} \psi_j \rangle \\
&= \sum_{j=1}^{\infty} \langle \psi_j, e^{-\lambda(H+X)+(1-\lambda)(H+Y)} \psi_j \rangle \\
&= \sum_{j=1}^{\infty} \lim_{n\to\infty} \langle \psi_j, e^{-(\lambda(H+X_n)+(1-\lambda)(H+Y_n))} \\
&\leq \sum_{j=1}^{\infty} \lim_{n\to\infty} \left\{ \langle \psi_j, \left(\lambda e^{-(H+X_n)} + (1-\lambda) e^{-(H+Y_n)} \right) \psi_j \rangle \right\} \\
&\quad \text{by convexity with bounded perturbations} \\
&= \sum_{j=1}^{\infty} \left\{ \lambda \langle \psi_j, e^{-(H+X)} \psi_j \rangle + (1-\lambda) \langle \psi_j, e^{-(H+Y)} \psi_j \rangle \right\} \\
&= \lambda \mathrm{Tr}\, e^{-(H+X)} + (1-\lambda) \mathrm{Tr}\, e^{-(H+Y)}.
\end{aligned}$$

Thus, $\mathrm{Tr}\, e^{-(H+X)}$ is convex in X in a 'hood of $X = 0$, and so $\Phi(X)$ is too.

The closure of the epigraph of $\lambda \mapsto \Phi(\lambda X)$ defines a lower-semicontinuous quantum Young function, which we now again call Φ, and which we take to define the norm

$$\|X\| := \inf_r \left\{ r : \Phi\left(\frac{X}{r} \right) < 1 \right\}. \tag{14}$$

The function $\lambda \mapsto \Phi(\lambda X)$ is continuous in $\lambda \in (-\epsilon, \epsilon)$ provided that it is finite in the interval $(-\epsilon, \epsilon)$. For, we proved[15] that the value of $\mathrm{Tr}\, e^{-(H+\lambda X)}$ lies between $\mathrm{Tr}\, e^{-(1+\lambda a)H}$ and $\mathrm{Tr}\, e^{-(1-\lambda a)H}$, and so is continuous at $\lambda = 0$. For a general λ, we replace H by $H + \lambda' X + \lambda X$ and show that this is continuous at $\lambda = 0$. Another proof of these facts follows from the Golden-Symanzik-Thompson inequality, in the form:

Theorem 4.2. Let A, B be self-adjoint operators on the separable Hilbert space $\mathcal{H}$, such that $\mathrm{Dom}\, A^{1/2} \cap \mathrm{Dom}\, B^{1/2} := \mathcal{D}$ is dense in $\mathcal{H}$. Let $A \overset{\circ}{+} B$

be the form sum of A and B, defined on $\mathcal{D}$. Then

$$\operatorname{Tr} e^{A \overset{\circ}{+} B} \leq \operatorname{Tr} \left(e^A e^B \right). \tag{15}$$

Simon gives a sketch of the proof of this, in his book[13], Theorem (8.5). He used a result of Kato[7], and showed for any symmetric norm Φ, (that is, the norm obeys $\Phi(X) = \Phi(X^*)$) defined on the semigroups, that if

$$e^{A/2} e^B e^{A/2}$$

has finite Φ norm, then so has e^{A+B}; and then we have

$$\Phi \left(e^{A+B} \right) \leq \Phi \left(e^{A/2} e^B e^{A/2} \right).$$

As an alternative, we have the proof of Hiai[4].

To apply this form of the inequality, let $H > 0$ be self-adjoint on $\mathcal{H}$ be such that $e^{-H} \in \mathcal{M}_{\beta_1}$, and let X be an H-form-bounded quadratic form, and put $\beta(H \overset{\circ}{+} X) = A$ and $(1-\beta)H = B$. Then $\operatorname{Dom} A, B = \operatorname{Dom} H^{1/2}$ is dense in $\mathcal{H}$, and $H \overset{\circ}{+} \beta X = A \overset{\circ}{+} B$. Moreover, we have

$$\begin{aligned} \operatorname{Tr} e^{-(H \overset{\circ}{+} \beta X)} = \operatorname{Tr} e^{A \overset{\circ}{+} B} \leq \operatorname{Tr}(e^A e^B) = \operatorname{Tr}(e^{-\beta(H \overset{\circ}{+} X)} e^{-(1-\beta)H} \\ \leq \| e^{-\beta(H \overset{\circ}{+} X)} \|_\infty \ \| e^{-(1-\beta)H} \|_1 \end{aligned} \tag{16}$$

by Hölder's inequality. By assumption, $e^{-(1-\beta)H}$ is of trace-class for all $1 - \beta > \beta_1$, so the perturbed state is of trace-class is finite for all $0 < \beta < 1 - \beta_1$. Similarly, one shows that $e^{-\alpha(H \overset{\circ}{+} \beta X)}$ is of trace class for all α close enough to 1, and so it lies in $\mathcal{M}_\alpha$.

The first patch of the manifold around the point $\rho = e^{-H}$ is defined in a way following[10]; it consists of all forms X such that the norm $\|X\|_0$ is sufficiently small. The topology of this set is that determined by this Orlicz norm. At a point Y with $\rho_{X+Y} := e^{-(H+Y+\psi(Y))} \in \mathcal{M}$, we could use the norm $\| \cdot \|_{X+Y}$, and proceed as do Pistone and Sempi; to make sense, on points lying in the overlap of these two spheres, we need to prove that the topologies are equivalent. This was done in my paper[16].

We have then a Banach manifold modelled on a Banach space, which is flat under the (+1)-connection. It is not trivial, however, since there are no coordinates around $\rho = e^{-H}$ that cover the whole space. The whole manifold is convex, however[14].

We note that Φ does not satisfy the Δ_2-condition, and so leads to a non-separable Banach manifold. Pistone has advocated a smaller, separable manifold in the classical case. We would expect, however, that in

quantum mechanics, such a small manifold would not be enough even for non-relativistic estimation theory.

Bibliography

1. Al-Rashid, M. H. A., and Zegarlinski, B., Non-commutative Orlicz spaces associated with a state, (*to appear in Studia Mathematica.*)
2. Chentsov, N. N., *Statistical Decision Rules and Optimal Inference*, Translations of Mathematical Monographs, (Amer. Math. Soc., Providence, R. I., **58**, 1982.)
3. Chentsov, N. N., and Morozova, E. A., Markov maps in non-commutative probability theory and mathematical statistics, *Probability Theory and Mathematical Statistics*, Vol. II, Vilnius, 1985, 287-310, (VNU Sci Press, Utrecht, 1987.)
4. Hiai, F., Trace-norm convergence of exponential product formula, *Letters in Mathematical Physics*, **33**, 147-158, 1995.
5. Ingarden, R. S., Janysek, H., Kossakowski, A., and Kawaguchi, T., Information geometry of quantum mechanical systems, *Tensor*, **37**, 105-111, 1982.
6. Jenkova, A., Affine connections, duality and divergences for aa von Neumann algebra, arXiv/math-ph/0311004.
7. Kato, T., Trotter's product-formula for an arbitrary pair of self-adjoint contraction semi-groups, *Advances in Math* Supplement Stud., **3**, 185-195, (Academic Press, New York, 1978.)
8. Kunze, W., Non-commutative Orlicz spaces and generalized Arens algebras, *Math. Nachr.*, **147**, 123-, 1990.
9. Petz, D., Monotone metrics on matrix spaces, *Linear Algebra and Appl.*, **244**, 81-96, 1996.
10. Pistone, G., and Sempi, C., An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one, *Annals Stat.*, **23**, 1543-1561, 1995.
11. Petz, D., and Toth, G., The Bogoliubov inner product in quantum statistics, *Letters in Mathemaical Phys.*, **27**, 205-216, 1993.
12. Ruskai, M. B., *Commun. Math. Phys.*, **26**, 280-289, 1972.
13. Simon, B., *Trace Ideals and Their Applications*, *Mathematical Surveys and Monographs, Vol. 120*, (Amer. Math. Soc., 2005.)
14. Streater, R. F., The information manifold for relatively bounded potentials, in *Problems of the Modern Mathematical Physics*, Eds. Vladimirov, V. S., Slavnov, A. A., and Katanaev, M. O., *Proc. V. A. Steklov Institute Math.*, **228**, 205-223, 2000.
15. Streater, R. F., Quantum Orlicz spaces in information geometry, *Open Systems and Information Dynamics*, **11**, 350-375, 2004.
16. Streater, R. F., The Banach manifold of quantum states, *to appear.*

QUANTUM EXTENSIONS OF THE CLASSICAL DOMINATION PRINCIPLE

VERONICA UMANITÀ*

Dipartimento di Matematica "F. Brioschi", Politecnico di Milano, Milano, I-20133, Italy
E-mail: veronica.umanita@polimi.it
www.mate.polimi.it

We explore some natural reformulations of the Classical Domination Principle for Quantum Dynamical Semigroups on arbitrary von Neumann algebras, and we investigate if such extensions are true or not.

Keywords: QDS; Domination principle; Superharmonic element; Potential.

1. Introduction

Let us consider a measurable space $(E, \mathcal{E})$ with a σ-finite measure μ on it and $\mathcal{T}$ a positive submarkovian semigroup on $L^\infty(E, \mathcal{E}, \mu)$. The Domination Principle of classical Potential Theory (*Classical Domination Principle*, CDP) says that a superharmonic function f majorizing the potential $\mathcal{U}(g)$ of a function g on the set $\{g > 0\}$ is greater than $\mathcal{U}(g)$ everywhere (see[3]). The CDP has noteworthy applications to classical probability theory as, for instance, the maximum principle.

Here instead is a possible interpretation of the principle on a physical ground: if f represents the electric potential generated by a density charge h, that is f is the solution of the Poisson equation

$$\Delta f = -4\pi h \quad (h \geq 0),$$

then, in order for the potential $\mathcal{U}(h)$ generated by h to majorize the potential generated by another density charge g, it suffices that the inequality holds where g is strictly positive, i.e. where the charge "given by g" is actually present.

In this paper we study non-commutative extensions of the CDP to a Quantum Dynamical Semigroup (QDS) on an arbitrary von Neumann algebra $\mathcal{A}$. We show that the closest and most natural reformulation (Conjecture

3.1) of the CDP does not hold true in general (see example 3.1). Indeed, we prove that, in case $\mathcal{A}$ is non abelian, one can always find a submarkovian QDS and two positive elements that violate the non-commutative version of the CDP (see Theorem 3.2). On the other hand, we give two additional conditions under which our reformulation is true (see Theorem 3.3).

2. Preliminaries

We denote by $\mathcal{A}$ an arbitrary von Neumann algebra on a complex Hilbert space $\mathcal{H}$ and by $\mathbf{1}$ its unit. We recall that a *Quantum Dynamical Semigroup* (QDS) $\mathcal{T}$ on $\mathcal{A}$ is a weak* continuous semigroup of normal completely positive maps $\{\mathcal{T}_t\}_{t\geq 0}$ on $\mathcal{A}$; $\mathcal{T}$ is *submarkovian* (resp. *submarkovian*) if $\mathcal{T}_t(\mathbf{1}) \leq \mathbf{1}$ (resp. $\mathcal{T}_t(\mathbf{1}) = \mathbf{1}$) for all $t \geq 0$.
For every operator x on $\mathcal{H}$, we denote by $[x]$ the orthogonal projection onto the closure of $x(\mathcal{H})$ and we call it the *range projection of* x. It is a well-known fact that $x \in \mathcal{A}$ implies $[x] \in \mathcal{A}$.
We say that a positive element $x \in \mathcal{A}$ is *superharmonic* if $\mathcal{T}_t(x) \leq x$ for all $t \geq 0$ (see[4]).

Given a positive operator $x \in \mathcal{A}$, we set

$$\mathfrak{U}(x)[u] = \int_0^\infty \langle u, \mathcal{T}_s\,(x)\,u\rangle ds$$

and we call x *integrable* if $\mathfrak{U}(x)[u] < \infty$ for every $u \in \mathcal{H}$; in this case, since $\mathfrak{U}(x)$ is a symmetric, positive and closed quadratic form thanks to Theorem 3.13a and Lemma 3.14a of Ref. 5, it is represented by a bounded self-adjoint operator (see Theorems 2.1, 2.6 and 2.23 of[5]) $\mathcal{U}(x) \in \mathcal{A}$ which is called the *potential of* x.
We denote by $\mathcal{A}_{int}$ the set of all integrable elements.

If $\mathcal{T}$ is a submarkovian QDS on $\mathcal{A} = L^\infty(E, \mathcal{E}, \mu)$ ($(E, \mathcal{E})$ measurable space, μ a σ-finite measure on E), then the *Classical Domination Principle* (see Theorem XII.27 of[3]) states

Classical Domination Principle 1 (CDP). *If f is superharmonic and g is integrable, the inequality $f \geq \mathcal{U}(g)$ on $\{g > 0\}$ implies $f \geq \mathcal{U}(g)$ everywhere.*

3. Reformulation of the domination principle

As a first step, we have to give a meaning to the statement

$$\text{“ } x \geq \mathcal{U}(y) \text{ on } \{y > 0\} \text{ ”}, \tag{1}$$

$x, y \in \mathcal{A}$. A natural extension of (1) could be

$$pxp \geq p\mathcal{U}(y)p, \tag{2}$$

where p denotes the spectral projection of y associated with the interval $]0, +\infty[$, that is $p = [y]$: indeed, since for all $g \in L^\infty(E, \mathcal{E}, \mu)$ the spectral measure of g is given by $E^g(S) = M_{1_{g^{-1}(S)}}$, the inequality $pfp \geq p\mathcal{U}(g)p$ translates into $1_{\{g>0\}}f \geq 1_{\{g>0\}}\mathcal{U}(g)$, which exactly means that $f \geq \mathcal{U}(g)$ on $\{g > 0\}$.

Therefore, our reformulation of the domination principle is the following:

Conjecture 3.1. *If x, y are positive elements of $\mathcal{A}$, x superharmonic and y integrable, and $p = [y]$, then*

$$pxp \geq p\mathcal{U}(y)p \Longrightarrow x \geq \mathcal{U}(y).$$

Unfortunately, we show below that this conjecture is not true: indeed, condition (2) only concerns "diagonal" elements and so what it may and does happen in the following example is that $pxp = p\mathcal{U}(y)p$, and yet $p(x - \mathcal{U}(y))p^\perp$ is not 0.

Example 3.1. Let $\mathcal{A} = \mathcal{B}(\ell^2(\mathbb{N}, \mathbb{C}))$ and

$$\mathcal{L}(x) := S^* x S - x$$

for all $x \in \mathcal{A}$, where $Se_k = e_{k+1}$ and $(e_k)_{k\geq 0}$ is the canonical orthonormal basis of $\ell^2(\mathbb{C})$. Since $\mathcal{L}(\mathbf{1}) = 0$ and $\mathcal{L}$ has the Lindblad form, $\mathcal{L}$ generates a uniformly continuous QMS on $\mathcal{A}$.
Given $x \in \mathcal{A}$, let $p_{k,j}(t) := \langle e_k, \mathcal{T}_t(x)e_j\rangle$ for all $t \geq 0$, $k, j \in \mathbb{N}$; then, by direct computation, we have

$$\begin{aligned} p'_{k,j}(t) &= \langle e_k, \mathcal{L}(\mathcal{T}_t(x))e_j\rangle = (\mathcal{L}_*(|e_j\rangle\langle e_k|))(\mathcal{T}_t(x)) \\ &= -p_{k,j}(t) + p_{k+1,j+1}(t), \end{aligned}$$

so that

$$\langle e_k, \mathcal{T}_t(x)e_j\rangle = \sum_{l\geq 0} e^{-t}\frac{t^l}{l!}\langle e_{l+k}, xe_{j+l}\rangle$$

is the unique solution of the above equation with initial data $p_{k,l}(0) = \langle e_k, xe_j\rangle$. Therefore,

$$\langle e_k, \mathcal{U}(x)e_j\rangle = \sum_{l\geq 0}\langle e_{l+k}, xe_{j+l}\rangle = \sum_{l\geq 0}\langle S^k e_l, xS^j e_l\rangle = \sum_{l\geq 0}\langle e_l, S^{*k}xS^l e_j\rangle,$$

which implies

$$\mathcal{U}(x) = \sum_{k\geq 0} S^{*k} x S^k.$$

Put $y = |e_0\rangle\langle e_0|$ and $z = 2^{-1}\sum_{i,j=0}^{1} |e_i\rangle\langle e_j|$; since such elements are clearly integrable, in particular we can define $x = \mathcal{U}(z)$. Moreover, we can think of x, y and z as

$$y = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad z = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad x = \frac{1}{2}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix};$$

we have that x is superharmonic (it is a potential), $\mathcal{U}(y) = y$, $p = y$ and $pxp = y = p\mathcal{U}(y)p$, yet

$$x - \mathcal{U}(y) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \not\geq 0.$$

In addition, we will show that, if the algebra $\mathcal{A}$ is not abelian, we can always find a submarkovian QDS $\mathcal{T}$ and two positive elements x, y in $\mathcal{A}$, x superharmonic and y integrable, which do not satisfy Conjecture 3.1. To prove this, we introduce some preliminary results.

Lemma 3.1. *Let $a, b \in \mathcal{A}_+$ be such that $b \leq a$ and let $p = [b]$. Then, $|\langle pu, ap^\perp u\rangle|^2 \leq \langle p^\perp u, ap^\perp u\rangle\langle pu, (a-b)pu\rangle$ for all $u \in \mathcal{H}$.*

Proof. The inequality $b \leq a$ implies $\langle pv, (a-b)pv\rangle + 2\Re\langle pv, ap^\perp v\rangle + \langle p^\perp v, ap^\perp v\rangle \geq 0$ for all $v \in \mathcal{H}$, for $pbp = b$; taking $v = \lambda pu + p^\perp u$ with $u \in \mathcal{H}$, $\lambda = \mu e^{i\theta}$, $\mu \in \mathbb{R}$ and θ which satisfies

$$\Re\langle pu, ap^\perp u\rangle \sin\theta = \Im\langle pu, ap^\perp u\rangle \cos\theta, \tag{3}$$

we obtain $\mu^2\langle pu, (a-b)pu\rangle + 2\mu\Re(e^{i\theta}\langle pu, ap^\perp u\rangle) + \langle p^\perp u, ap^\perp u\rangle \geq 0$, which means $[\Re(e^{i\theta}\langle pu, ap^\perp u\rangle)]^2 \leq \langle p^\perp u, ap^\perp u\rangle\langle pu, (a-b)pu\rangle$. But

$$[\Re(e^{i\theta}\langle pu, ap^\perp u\rangle)]^2 = (\Re\langle pu, ap^\perp u\rangle\cos\theta + \Im\langle pu, ap^\perp u\rangle\sin\theta)^2 = |\langle pu, ap^\perp u\rangle|^2$$

by virtue of (3), and so the conclusion follows. □

Definition 3.1. We say that $\mathcal{A}$ satisfies the *weak domination property* (WDP) if for each $a, b \in \mathcal{A}_+$ the inequality $pap \geq pbp$ with $p = [b]$ implies $a \geq b$.

Theorem 3.1. *Any von Neumann algebra $\mathcal{A}$ which satisfies the (WDP) is abelian.*

Proof. Since any von Neumann algebra is generated by its projections, it is enough to show that $[p, z] = 0$ for all $z \in \mathcal{A}$ and projections $p \in \mathcal{A}$. Moreover, since $z = z^+ - z^-$ with z^+, z^- positive elements in $\mathcal{A}$ we can suppose $z \in \mathcal{A}_+$.
We put $a := p + pzp^\perp + p^\perp zp + p^\perp z^2 p^\perp$ and we claim that a is positive, or equivalently, that

$$\|pu\|^2 + 2\Re\langle pu, zp^\perp u\rangle + \|zp^\perp u\|^2 \geq 0$$

for all $u \in \mathcal{H}$. But $|\Re\langle pu, zp^\perp u\rangle| \leq \|pu\| \, \|zp^\perp u\|$ for all $u \in \mathcal{H}$ by Cauchy-Schwarz inequality, so

$$\|pu\|^2 + \|zp^\perp u\|^2 + 2\Re\langle pu, zp^\perp u\rangle \geq 2\|pu\| \, \|zp^\perp u\| +$$

$$+2\Re\langle pu, zp^\perp u\rangle \geq 2|\Re\langle pu, zp^\perp u\rangle| + 2\Re\langle pu, zp^\perp u\rangle \geq 0$$

for all $u \in \mathcal{H}$. Hence $a \geq 0$.
Since $[p] = p$ and $pap = p$, thanks to (WDP) we have $a \geq p$, so that $pzp^\perp = p^\perp zp = 0$ by Lemma 3.1. Therefore $z = pzp + p^\perp zp^\perp$, i.e. $[p, z] = 0$. □

In the end, we have the following

Theorem 3.2. *If $\mathcal{A}$ is non abelian, then there exist a submarkovian QDS $\mathcal{T}$, $x, y \in \mathcal{A}_+$, x superharmonic and y integrable, such that $pxp \geq p\mathcal{U}(y)p$ but $x \not\geq \mathcal{U}(y)$, where $p = [y]$.*

Proof. Since $\mathcal{A}$ is non abelian, by Corollary 3.1 there exist $x, y \in \mathcal{A}_+$ such that $pxp \geq pyp$ but $x \not\geq y$, $p = [y]$. Put $\mathcal{T}_t(z) = e^{-t}z$ for all $z \in \mathcal{A}$, $t \geq 0$; then $(\mathcal{T}_t)$ is clearly a submarkovian QDS on $\mathcal{A}$; moreover x is superharmonic and

$$\int_0^\infty \langle u, \mathcal{T}_t(y)u\rangle dt = \int_0^\infty e^{-t}\langle u, yu\rangle dt = \langle u, yu\rangle$$

for all $u \in \mathcal{H}$, so that y is integrable and $\mathcal{U}(y) = y$.
Therefore, $pxp \geq pyp = p\mathcal{U}(y)p$, yet $x \not\geq \mathcal{U}(y)$. □

Notice that the domination principle holds at least in a trivial case, i.e. when x and y belong to an abelian $\mathcal{T}$-invariant subalgebra of $\mathcal{A}$.
Instead, we found a weaker and quite natural condition under which the domination principle is fulfilled.

Theorem 3.3. *Let $\mathcal{T}$ be a submarkovian QDS on $\mathcal{A}$. If x is superharmonic, y integrable and $p = [y]$, then*

$$\left.\begin{array}{c} pxp \geq p\mathcal{U}(y)p \\ [x,p]=0 \text{ and } p \text{ superhamonic} \end{array}\right\} \Longrightarrow x \geq \mathcal{U}(y).$$

Proof. If $[x,p] = 0$, then $x = pxp + p^{\perp}xp^{\perp}$; moreover, since $p\mathcal{A}p$ is invariant (p is superharmonic), we have $\mathcal{U}(y) = p\mathcal{U}(y)p$, so that $x \geq \mathcal{U}(y)$ when $pxp \geq p\mathcal{U}(y)p$. □

Remark 3.1. The hypotesis "$[x,p] = 0$" is very natural to assume, for in the physical setting, we need observables to be comparable. On the other hand, "p superharmonic" is a quite strong condition: indeed, for instance, in the abelian case of electric potential, it means that the electric field generated by a density charge g supported by p stays confined in the region where g is actually present.
As a matter of fact, a more cogent hypothesis would be that x, $\mathcal{U}(y)$ and p (or x, $\mathcal{U}(y)$ and y) commute, but it is very difficult to find nontrivial examples satisfying such conditions.

4. Analysis of the proof of the abelian case

In this section we analyze the proof given by Dellacherie and Meyer in Ref. 3 for abelian algebras, and we try to understand what goes wrong with the non-commutative case.
Their proof is based upon the following two facts (proved in Theorems $IX.33$ and $IX.33$ of Ref. 2):

(a) If $\mathcal{T}$ is a submarkovian QDS on $L^{\infty}(E,\mathcal{E},\mu)$ and $h \in L^{\infty}(E,\mathcal{E},\mu)$ is a positive function, then there exists the smallest superharmonic function $R(h)$ which majorizes h.
(b) If h belongs to $\mathcal{A}_{int}$, the smallest superharmonic function which majorizes $\mathcal{U}(h) + ch$ on $\{h > 0\}$ is $\mathcal{U}(h) + ch$ itself, i.e.

$$R((\mathcal{U}(h) + ch)1_{\{h>0\}}) = \mathcal{U}(h) + ch$$

for every positive constant c.

In particular, if f and g are positive functions, f superharmonic and g integrable, with $f \geq \mathcal{U}(g)$ on $M := \{g > 0\}$, we clearly have that

$$f + \lambda^{-1}\|g\| \geq (\mathcal{U}(g) + \frac{g}{\lambda})1_M$$

for all $\lambda > 0$. Since $f + \lambda^{-1}\|g\|$ is also superharmonic, by a) and b) it follows that

$$f + \lambda^{-1}\|g\| \geq R((\mathcal{U}(g) + \frac{g}{\lambda})1_M) = \mathcal{U}(g) + \frac{g}{\lambda},$$

and therefore $f \geq \mathcal{U}(g)$ by letting $\lambda \to \infty$.

Remark 4.1. A simple consequence of b) is that $R(\mathcal{U}(h)1_{\{h>0\}}) = \mathcal{U}(h)$ for all $h \in \mathcal{A}_{int}$.
Indeed, given a superharmonic function f which majorizes $\mathcal{U}(h)1_{\{h>0\}}$, we have $f + c\|h\| \geq R((\mathcal{U}(h) + ch)1_{\{h>0\}}) = \mathcal{U}(h) + ch$ for every $c > 0$, so that $f \geq \mathcal{U}(h)$ by letting $c \to 0$.

Let now $\mathcal{A}$ be an arbitrary von Neumann algebra and $\mathcal{T}$ a submarkovian QDS on $\mathcal{A}$. We want to see if properties a) and b) can be generalized to this case.
We have the following:

Theorem 4.1. *If z is a positive element in $\mathcal{A}$, then there exists a minimal superharmonic element which majorizes z.*

Proof. Let $\mathcal{S} = \{\, y \in \mathcal{A}_+ : y \text{ is superharmonic, } z \leq y \,\}$: $\mathcal{S}$ is not empty, since it contains $\|z\|\mathbf{1}$. We define a partial order on $\mathcal{S}$ by letting $y_1 \preceq y_2$ when $y_1 \geq y_2$. Therefore, given a totally ordered subset T of $\mathcal{S}$, the net $\{y\}_{y \in T}$ is decreasing in $\mathcal{A}_+$ and therefore there exists $y_0 := \inf_T y$, which clearly majorizes z. Moreover, since $\mathcal{T}_t(y_0) \leq \mathcal{T}_t(y) \leq y$ for all $y \in T$, $t \geq 0$, we have also $\mathcal{T}_t(y_0) \leq y_0$, so that y_0 belongs to $\mathcal{S}$ and it is a majorant (w.r.t. $\preceq$) of T. Applying Zorn Lemma we find a maximal element of $\mathcal{S}$, which clearly is a minimal superharmonic element majorizing z. □

Unfortunately, such a minimal element will not be unique in general, as the following example shows.

Example 4.1. We consider $\mathcal{A}$ and $\mathcal{L}$ as in example 3.1. Then, a positive element $x = \sum_{i,j \geq 0} x_{i,j} \, |e_i, \rangle\langle e_j \,|$ is superharmonic if and only if

$$\sum_{i,j \geq 0} (x_{i+1,j+1} - x_{i,j}) \, |e_i\rangle\langle e_j \,| \leq 0. \tag{4}$$

We set

$$z = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and we want to show that the smallest superharmonic element which majorizes z does not exist.
A simple computation proves that

$$z_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is superharmonic and $z_0 \geq z$. Moreover, given $w = \sum_{i,j\geq 0} w_{i,j} \mid e_i\rangle\langle e_j \mid$ superharmonic with $z \leq w \leq z_0$, we have in particular

$$\begin{pmatrix} 1-w_{0,0} & -w_{0,1} \\ -w_{1,0} & 1-w_{1,1} \end{pmatrix} \geq 0 \quad \text{and} \quad \begin{pmatrix} w_{0,0} & w_{0,1} \\ w_{1,0} & w_{1,1}-1 \end{pmatrix} \geq 0,$$

so that $w_{0,0} \leq 1$, $w_{1,1} = 1$ and $w_{0,1} = w_{1,0} = 0$. But condition (4) implies $w_{1,1} \leq w_{0,0}$, and so $w_{0,0} = w_{1,1} = 1$. Therefore, by $w \leq z_0$ we get $w_{i,j} = 0$ for every $i, j > 1$, i.e. $w = z_0$. This proves that z_0 is minimal for the set of superharmonic elements which majorize z. But it is not the minimum of such a set, because

$$y = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

is superharmonic and satisfies $y \geq z$, yet

$$y - z_0 = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$$

is neither positive nor negative.

Therefore, in general, we can not give a meaning to $R(z)$ for all $z \geq 0$; in any case, given $y \in \mathcal{A}_{int}$, even if $R(p\mathcal{U}(y)p + cy)$ would exist for all $c \geq 0$, with $p = [y]$, property b) would not generally be satisfied, since the necessary condition $R(p\mathcal{U}(y)p) = \mathcal{U}(y)$ given in Remark 4.1 is not automatically true. Indeed, the following example shows that the inequality $\mathcal{U}(y) \geq p\mathcal{U}(y)p$ can fail.

Example 4.2. We consider $\mathcal{A}$ and $\mathcal{L}$ as in example 3.1 and we set

$$y = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then, $p = y/2$ and

$$\mathcal{U}(y) - p\mathcal{U}(y)p = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} - \frac{5}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 3 & -1 \\ -1 & -1 \end{pmatrix} \not\geq 0.$$

Notice that it is not enough to add the condition $\mathcal{U}(y) \geq p\mathcal{U}(y)p$ to obtain the domination principle: indeed, just in the example 3.1 we had $\mathcal{U}(y) = y = p\mathcal{U}(y)p$, $pxp \geq p\mathcal{U}(y)p$, but $x \not\geq \mathcal{U}(y)$.

5. An alternative reformulation

In this section we present a different interpretation of condition (1) in order to take into account the non-commutativity of the algebra. Unfortunately, we will see that, in general, the domination principle still does not hold with this new reformulation.

We may interpret (1) in the following way:

$$px + xp \geq p\mathcal{U}(y) + \mathcal{U}(y)p, \tag{5}$$

where $p = [y]$, x is superharmonic and y integrable.
Clearly (5) implies $pxp \geq p\mathcal{U}(y)p$, and the two conditions are equivalent in the abelian framework.

We can then translate the domination principle in this way:

Conjecture 5.1. *Let x, y be positive elements of $\mathcal{A}$, x superharmonic, y integrable and let $p = [y]$. Then*

$$px + xp \geq p\mathcal{U}(y)p \Longrightarrow x \geq \mathcal{U}(y), \tag{6}$$

or, equivalently,

$$\left.\begin{array}{r} pxp \geq p\mathcal{U}(y)p \\ pxp^{\perp} = p\mathcal{U}(y)p^{\perp} \end{array}\right\} \Longrightarrow\ p^{\perp}xp^{\perp} \geq p^{\perp}\mathcal{U}(y)p^{\perp}. \tag{7}$$

The equivalence of (6) and (7) is a simple consequence of the following

Proposition 5.1. *Let z, w be positive elements in $\mathcal{A}$. The following conditions are equivalent:*

(1) $pz + zp \geq pw + wp$;
(2) $pzp \geq pwp$ and $pzp^{\perp} = pwp^{\perp}$.

Proof. If we assume condition 2), then $pz + zp = 2pzp + pzp^{\perp} + p^{\perp}zp \geq 2pwp + pwp^{\perp} + p^{\perp}wp = pw + wp$.
Conversely, if condition 1) holds, we obtain

$$\langle u, (z - w)u\rangle + \Re(\lambda\langle u, (z - w)v\rangle) \geq 0$$

for all $u \in p(\mathcal{H})$, $v \in p^{\perp}(\mathcal{H})$ and $\lambda \in \mathbb{C}$.
Taking $\lambda \in \mathbb{R}$ we have $pzp \geq pwp$ and $\Re\langle u, zv\rangle = \Re\langle u, wv\rangle$, while with $\lambda = i\mu \in i\mathbb{R}$ it follows $\Im\langle u, zv\rangle = \Im\langle u, wv\rangle$. □

Notice that, if $p = [y] = [\mathcal{U}(y)]$, then (7) is trivially satisfied but, in general, this version of domination principle is neither true.

Example 5.1. Let $\mathcal{A}$ and $\mathcal{L}$ as in the example 3.1. We set

$$y = \frac{1}{2}\begin{pmatrix} 4 & 2i \\ -2i & 1 \end{pmatrix}, \quad z = \frac{1}{2}\begin{pmatrix} 7 & 4i \\ -4i & 1 \end{pmatrix}, \quad x \doteq \mathcal{U}(z) = \frac{1}{2}\begin{pmatrix} 8 & 4i \\ -4i & 1 \end{pmatrix}.$$

Then $p = (2/5)\, y$ and

$$\mathcal{U}(y) = \frac{1}{2}\begin{pmatrix} 5 & 2i \\ -2i & 1 \end{pmatrix}.$$

An easy computation show that $px + xp \geq p\mathcal{U}(y) + \mathcal{U}(y)p$, but

$$x - \mathcal{U}(y) = \frac{1}{2}\begin{pmatrix} 3 & 2i \\ -2i & 0 \end{pmatrix} \not\geq 0.$$

6. A weaker reformulation

If $\mathcal{A}$ is a $\mathcal{C}^*$-algebra, then we can write any self-adjoint element x as $x = x^+ - x^-$, where $x^+ = (|\, x \,| + x)/2$ and $x^- = (|\, x \,| - x)/2$ are positive. Clearly such a decomposition is not unique but, if $\mathcal{A}$ is abelian, then this is the minimal one, i.e. the equality $x = x_1 - x_2$ with x_1, x_2 positive implies $x_1 \geq x^+$ and $x_2 \geq x^-$.
However, for finite-dimensional non-abelian algebras, just the following inequalities

$$\lambda_j(x_1) \geq \lambda_j(x^+) \quad \text{and} \quad \lambda_j(x_2) \geq \lambda_j(x^-)$$

are fulfilled (see Lemma IX.4.1 of[1]), where $\lambda_j(y)$ denotes the j-th eigenvalue of y with respect to decreasing order.
This suggest the following weaker reformulation of the domination principle:

Conjecture 6.1. *Assume $\mathcal{A}$ is finite-dimensional. If x, y are positive elements of $\mathcal{A}$, x superharmonic and y integrable, and $p = [y]$, then*

$$pxp \geq p\mathcal{U}(y)p \Longrightarrow \lambda_j(x) \geq \lambda_j(\mathcal{U}(y)) \text{ for every } j. \tag{8}$$

We test Conjecture 6.1 on the usual example in order to understand if such a formulation may hold or not.

Example 6.1. We consider $\mathcal{A}$ and $\mathcal{L}$ as in Example 3.1. Let

$$y = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad z = \begin{pmatrix} 7 & 0 \\ 0 & (3 - 2\sqrt{5})/2 \end{pmatrix}$$

and

$$x = \mathcal{U}(z) = \frac{1}{2}\begin{pmatrix} 17 - 2\sqrt{5} & 0 \\ 0 & 3 - 2\sqrt{5} \end{pmatrix}.$$

An easy calculation shows that $p = (1/2)y$,

$$\mathcal{U}(y) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \lambda_1(\mathcal{U}(y)) = \frac{3+\sqrt{5}}{2} \quad \text{and} \quad \lambda_2(\mathcal{U}(y)) = \frac{3-\sqrt{5}}{2}.$$

Therefore, we have

$$pxp = \frac{10-2\sqrt{5}}{4}\, y \geq \frac{5}{4}\, y = p\mathcal{U}(y)p,$$

but

$$\lambda_2(x) = \frac{3-2\sqrt{5}}{2} \not\geq \lambda_2(\mathcal{U}(y)),$$

so that Conjecture 6.1 also is false.

Bibliography

1. R. Bhatia, *Matrix analysis* (Springer-Verlag, New York 1997).
2. C. Dellacherie and P. A. Meyer, *Probabilités et potentiel. Chapitres IX-XI*, 2nd edition (Hermann, Paris 1987).
3. C. Dellacherie and P. A. Meyer, *Probabilités et potentiel. Chapitres XII-XVI*, 2nd edition (Hermann, Paris 1987).
4. F. Fagnola and R. Rebolledo, Subharmonic projections for a quantum semigroup, *J. Math. Phys.* **43**, no. 2, 1074–1082 (2002).
5. T. Kato, *Perturbation theory for linear operator*, Corr. printing of the 2nd edition (Springer-Verlag, New York, 1980).

ANALYSIS IN OPERATOR SPACES

BOGUSŁAW ZEGARLIŃSKI

Department of Mathematics,
Imperial College London, UK,
&
Chaire d'Excellence Pierre de Fermat
ADERMIP & UPS Toulouse b.zegarlinski@imperial.ac.uk

Keywords: Dissipative dynamics, Noncommutative algebras, Hypercontractivity, Nonlinear dissipative semigroups, Noncommutative Orlicz Spaces, Entropic switch

1. Introduction

We briefly review a number of issues of analysis in operator spaces. We begin from introducing and recalling few results (and open problems) on ergodicity of dissipative dynamics in inductive limit C^* algebra. As a motivation we present later some recent achievements in commutative case saying that one can obtain stronger ergodicity results by analysis based on a notion of hypercontractivity of dissipative dynamics in suitable functional ($\mathbb{L}_p$ and Orlicz) spaces. The hypercontractivity property, besides helping in study the ergodicity, has an intricate relation to the entropy and the particle structure of a system which provides additional motivation to research in this area.

In recent years one observes a growing interest in studying nonlinear dissipative dynamics including large interacting classical and quantum systems. To show that some interesting progress in this domain is possible we present some recent results for finite and infinite spin systems. One interesting point for such dissipative semigroups is an interesting phase transition: a dissipation may be contractive for $\mathbb{L}_p$ spaces with large indices p, but not so for small ones.

While the noncommutative $\mathbb{L}_p$ spaces associated to a state were introduced long time ago and seem to be well established, until recently much less attention was given to the noncommutative counterpart of Orlicz spaces. We

briefly sketch recent results in this area motivated by potential applications to analysis of coercive inequalities or information geometry.
At the end we give a brief account on the topics of entropic switch which are related to interplay of entropy and noncommutativity (nonlinearity).

2. Dissipative Dynamics in Noncommutative Spaces *Some Results & Problems*

Basic Setup. The first element of our setup is the following.

- $\{\mathcal{A}, ||\cdot||, \{\mathcal{A}_X, X \subset\subset \mathbb{Z}^d\}\}$ - the inductive limit C^* algebra over a finite dimensional complex matrix algebra $\mathbf{M}$, with a norm $||\cdot||$

In such algebra one has a large family of partial traces which have all the features of perfect conditional expectations.

- $\mathbf{Tr}_X : \mathcal{A} \longrightarrow \mathcal{A}_{X\complement}$ partial traces

$$\mathbf{Tr}_X f = \int \nu_X(dU)\ U^* f U$$

where $\nu_X(dU)$ is the normalised Haar's measure on the group of all unitaries U in $\mathcal{A}_X$.

This expression involving "local" unitaries U may be very convenient when passing to a representation associated to some nontrivial (Gibbs) state ω which when restricted to a finite dimensional subalgebra has a density with respect to the normalised trace.
On the algebra it is natural to introduce the following completely positive and unit preserving maps.

- Generalised Conditional Expectations $E_X : \mathcal{A} \longrightarrow \mathcal{A}_{X\complement}$

$$E_X(f) \equiv \mathbf{Tr}_X\ \gamma_X^* f \gamma_X$$

for some $0 < \gamma_X \in \mathcal{A}$, $\mathbf{Tr}_X\ \gamma_X^* \gamma_X = \mathbf{1}$

In case one could think of a state as given by a limit of the following form

$$\omega(f) \equiv \lim \lim_{\Lambda \to \mathbb{Z}^d} Tr\,(\rho_\Lambda f)$$

with

$$\rho_\Lambda \equiv \frac{e^{-H_\Lambda}}{Tre^{-H_\Lambda}}$$

one could have the following example.

- Example: $E_X \equiv E_{X,\phi}$

$$\gamma_X = \lim_{\Lambda \to \mathbb{Z}^d} e^{-\frac{1}{2}H_\Lambda} \left(Tr_X e^{-H_\Lambda}\right)^{-\frac{1}{2}} \tag{1}$$

where $H_\Lambda \equiv \sum_{Y \cap \Lambda \neq 0} \phi_Y$, $\phi_X \in \mathcal{A}_X$.

$$||\phi|| \equiv \sum_{X \ni 0} |X| \cdot ||\phi_X|| < \infty$$

The limit in (1) exists in the algebra for Classical Potentials by definition satisfying

$$\forall Y, Z \quad [\phi_Y, \phi_Z] = 0.$$

In general one can show existence of γ_X by passing to the representation in von Neumann algebra,[32]. In this case the generalised conditional expectation is a bounded selfadjoint operator in the symmetric $\mathbb{L}_{2,\frac{1}{2}}(\omega)$ space with a scalar product

$$< f, g >_{\omega,s} \equiv \lim_{\Lambda \to \mathbb{Z}^d} Tr\ \rho_\Lambda^{\frac{1}{2}} f^* \rho_\Lambda^{\frac{1}{2}} g$$

For applications we describe below, one would need to have more detailed control of behaviour of approximation of γ_X by local elements of the algebra. In particular it would be interesting to study the following problem.
Question 1 : Is the limit stable w.r.t. small perturbations $\lambda\psi \equiv \{\lambda\psi_X\}_{X \subset\subset \mathbb{Z}^d}$, $[\psi_X, \psi_X] \neq 0$ for ∞-many X ? Given the generalised conditional expectations one may like to introduce the following:

- Dissipative Generators

$$\mathcal{L}f \equiv \sum_j (E_{X+j,\phi}(f) - f) + \varepsilon\delta_\psi(f).$$

where

$$\delta_\psi(f) \equiv \lim_{\Lambda \to \mathbb{Z}^d} \sum_{X \subset \Lambda} [\psi_X, f].$$

To ensure possible uniformity of description as well as the ergodicity it would be important to have well defined generator including possibly infinite sum. This is exactly the point where more information would be required about γ_{X+j}. In case if classical potential Φ is used to define the generalised conditional expectation, it is possible to prove that

$$P_t \equiv P_t^{(\phi,\varepsilon\psi)} \equiv e^{t\mathcal{L}}$$

is a well defined dissipative semigroup on $\mathcal{A}$. Since in the generator we can include a derivation corresponding to another potential Ψ which is not classical and noncommuting with Φ, we get qualitatively new situation not encountered in the theory of Glauber type dynamics for classical spin systems. Then in some nontrivial domain one has the following result (cf.[37] and references therein).

- Strong Ergodicity & Stability : "If for $\varepsilon = 0$

$$|||P_t f||| \leq e^{-mt} C |||f||| \tag{2}$$

then similar is true for $\varepsilon < \varepsilon_0$, for some ε_0."

Thus a limit $\lim_{t\to\infty} P_t f$ exists and defines a state on the algebra. It must depend in some way on both Φ and Ψ and one may conjecture that it is a Gibbs state for some well defined (nonclassical) potential.
Unfortunately the method used to show (2) is very restrictive and necessarily leaves open intriguing question: what happens in the situation when the Hamiltonian dynamic term corresponding to Ψ is large and potentially may dominate large time behaviour. In fact this method does not allow even to get the ergodicity results for "large" classical potentials on noncommutative algebra, (in which case the derivation term is absent), in a vicinity of "critical point" where the correlations in the equilibrium state are becoming of longer and longer range. Given the experience with classical systems where an essential progress was made in tackling this issue by employing coercive inequalities in suitable functional spaces, one would like to ask:
Question 2 : Can one improve the state of the matter by analysis in Noncommutative $\mathbb{L}_p$ spaces ? With this in mind, in the next section we briefly review recent progress in coercive inequalities and applications to classical systems which involve infinite noncommutative set of derivations.

3. Coercive Inequalities for Gibbs Measures associated to Systems of Noncommutative Fields.

When studying a commutative system described by an infinite number of random variables it may be useful to represent the underlying space as a product space $(\Omega, \Sigma) = (\mathbb{M}, \mathcal{B})^{\mathcal{R}}$, with $\mathbb{M}$ being a smooth (Riemannian) manifold and $\mathcal{R} \equiv \{1, .., n\}, \mathbb{N}, \mathbb{Z}^d, ..,$ *or a more general infinite graph* $\mathcal{G}$. In this situation an important role is played by a (Smooth) Local Specification, that is a family of probability kernels $\{E_\Lambda : \Lambda \subset \mathcal{G}, |\Lambda| \in \mathcal{R}\}$, indexed by finite subsets of $\mathcal{R}$, which satisfy all the properties of the regular

conditional expectations including the following compatibility condition [i] $\Lambda_1 \subset \Lambda_2 \implies E_{\Lambda_2} E_{\Lambda_1} f = E_{\Lambda_1} f$. In a variety of practical applications E_Λ can be introduced as follows

$$E_\Lambda(f)(\widetilde{\omega}) = \delta_{\widetilde{\omega}_{\mathcal{R}\setminus\Lambda}} \left(\frac{\int_{\mathbb{M}^\Lambda} e^{-U_\Lambda} f \, d\omega_\Lambda}{\int_{\mathbb{M}^\Lambda} e^{-U_\Lambda} \, d\omega_\Lambda} \right)$$

with a "natural" finite dimensional product measure $d\omega_\Lambda$ and some (additive) interaction energy functional U_Λ. With the help of local specification following Dobrushin - Lanford - Ruelle, one can give apriori meaning to a state of the infinite dimensional system as follows

- **Gibbs Measure**

$$\mu E_\Lambda(f) = \mu(f) \qquad (DLR)$$

which simply means that E_Λ are equal to the conditional expectations associated to μ. The description with the local states E_Λ, provided the corresponding interaction is sufficiently smooth, allows also an introduction of a family of symmetric Markov generators using [ii]

$$E_{\{j\}} |\nabla_j g|^2 \equiv -E_{\{j\}} g L_j g$$

Then the (DLR) condition allows one to define

$$\mu |\nabla g|_2^2 \equiv \sum_{j \in \mathcal{R}} \mu |\nabla_j g|^2 = -\mu(gLg)$$

Similar definitions are possible with the use of the fields in place of the gradient.
One of the first important entropy bounds is the following

(i) Log-Sobolev Inequality

$$\mu \mathbf{v^2} \log \left(\frac{\mathbf{v^2}}{\mu \mathbf{v^2}} \right) \leq \mathbf{c}\, \mu |\nabla \mathbf{v}|_{\mathbf{2}}^{\mathbf{2}} \qquad (\mathbf{LS}_2)$$

As it was shown by L.Gross[4], this coercive inequality is equivalent a strong contractivity property of the semigroup (which action allows to revert Hölder inequality)

(ii) $\mathbb{L}_p(\mu)$ - Hypercontractivity :

$$||e^{tL} f||_{\mathbb{L}_{p(t)}} \leq ||f||_{\mathbb{L}_2}, \quad with \quad p(t) = 1 + e^{t/c} > 2$$

For more detailed history of this subject as well as exposition of some applications in infinite dimensional analysis see[11] and references therein. While the above Log-Sobolev inequality is relevant in the situations where the Lipschitz random variables have at least the Gaussian tails, more recently it was possible to go beyond this restriction by studying generalised entropy bounds of the following form.

(iii) **F**-Sobolev Inequality ([33])

$$\mu \mathbf{v^2 F}\left(\frac{\mathbf{v^2}}{\mu \mathbf{v^2}}\right) \leq \mathbf{c_F}\, \mu|\nabla \mathbf{v}|_{\mathbf{2}}^{\mathbf{2}} \qquad (\mathbf{FS}_2)$$

with F concave, non-decreasing on $\mathbb{R}^+$, vanishing at 1, ($\sim \{\log_+(x)\}^\beta$, $\beta \in (0,1]$).

It was shown in[33] that such coercive inequality is equivalent to the following interesting hypercontractivity property in Orlicz spaces.

(iv) $\mathbb{L}_{\Phi_p}(\mu)$ - Hypercontractivity :

$$||e^{tL} f||_{\Phi_{p(t)}} \leq ||f||_{\mathbb{L}_2},\ \textit{where}\ \ \Phi_p(x) \equiv x^2 e^{p\cdot F(x^2)}$$

In infinite dimensions for Gibbs Measures the corresponding coercive inequality was shown in[35] (see also[12]). Finally we mention yet another form of coercive inequalities which involves Orlicz spaces.

(v) Orlicz-Sobolev Inequalities

$$||(f-\mu(f))^2||_\Phi \leq C_\Phi\ \mu|\nabla f|_2^2 \qquad (\mathbf{OS}_2)$$

3.1. *Log-Sobolev Inequality for Hörmander Type Markov Generators*

More recently a progress has been made in case of infinite dissipative systems described by the following generators

$$\mu|\mathbf{X}g|_2^2 = \sum_{j\in\mathcal{R}} \mu|X_j g|^2 = -\mu(g\mathcal{L}g)$$

where $|X_j g|^2 \equiv \sum_{\alpha=1,..,k} |X_j^\alpha(\omega_j)\frac{\partial}{\partial\omega_{j,\alpha}} g|^2$ with $1 < k \leq dim(\mathbb{M})$.
The system of fields $\{X_j\}_{j\in\mathcal{R}}$ is noncommutative and may be degenerate. In such the case to achieve dissipativity one may take advantage of noncommutativity to get the following condition is satisfied.

*L*ocal Hörmander Condition : $\exists N \in \mathbb{N}\ \forall \omega_j \in \mathbb{M}, j \in \mathcal{R}$,

$$\mathbf{Span}\left\{X_j^\alpha,\, ad_{X_j^{\alpha_1}}(X_j^\alpha), .., \, ad_{X_j^{\alpha_N}} .. ad_{X_j^{\alpha_1}}(X_j^\alpha)\right\} = \mathbf{T}_{\omega_j}\mathbb{M} \qquad (\mathbf{H}_N)$$

where $\alpha, \alpha_l = 1, .., k$, and $ad_Y(X) \equiv [X, Y]$, and $\mathbf{T}_{\omega_j}\mathbb{M}$ denotes the tangent space at ω_j.
The corresponding finite dimensional theory has a rich and interesting history including works of Kolmogorov, Hörmander, Malliavin, Kusuoka&Stroock, and many others; for a more detailed description we refer to[36].

Theorem 3.1.[12]
Let $\mathbb{M}$ *be a smooth, connected, boundaryless manifold* $dim(\mathbb{M}) \in [3, \infty)$. *Suppose Regular Conditional Expectations* $\{E_\Lambda : |\Lambda| \leq L_0 \in \mathbb{N}\}$ *satisfy uniform mixing condition*

$$|E_\Lambda(f_{\Lambda_1}, g_{\Lambda_2})| \leq Const\ e^{-M\ dist(\Lambda_1, \Lambda_2)}$$

Then there exists $c \in (0, \infty)$

$$\mu\left(v^2 \log \frac{v^2}{\mu v^2}\right) \leq c\, \mu |\mathbf{X} v|^2 \qquad (\mathbf{XLS_2})$$

Corollaries
The real benefit from proving this type of coercive inequality comes in the form of multitude of abstract consequences including the following.

- Exponential Bounds Assume (XLS_2).

$$\mu e^{tf} \leq e^{Ct^2 ||\ |Xf|\ ||_\infty^2 + t\mu f}$$

- Spectral Gap : There exists $m \in (0, \infty)$

$$\mu\left(e^{t\mathcal{L}} f - \mu f\right)^2 \leq e^{-2mt} \mu\left(f - \mu f\right)^2$$

- Exponential Decay of Entropy : There exists $\alpha \in (0, \infty)$

$$\mu\left(P_t f \log \frac{P_t f}{\mu f}\right) \leq e^{-\alpha t} \mu\left(f \log \frac{f}{\mu f}\right)$$

With an extra information on finite speed of propagation of information and extra hard work one arrives at the following strong estimate in the uniform norm.

- Strong Exponential Decay to Equilibrium : *There exists* $m \in (0,\infty)$

$$\|P_t f - \mu f\|_u \leq e^{-mt}\ C_{\Lambda_0}\ |||f|||$$

for any observable f *localised in a bounded subset* $\Lambda_0 \subset \mathbb{Z}^d$ *with finite Lipschitz seminorm.*

The key point here is that by analysis in functional spaces, if hypercontractivity can be proved one go far beyond traditional methods which require much stronger technical restrictions. For example for Glauber type dynamics in the ferromagnetic NN Ising model, hypercontractivity method works till the critical point,[34] !!!

4. Hypercontractivity in Noncommutative Case.

In this section we briefly review a (very slow) progress in Hypercontractivity in Noncommutative $\mathbb{L}_p(\omega)$ spaces in some type III cases considered in[29,30].
For interesting development in an area where one deals with certain representations of commutation relations and type II situations see[1–8] and references therein.

We consider infinite product state as follows

- $\omega \equiv \otimes_l\, \omega_{\Lambda_l}$, where $\omega_{\Lambda_l} \equiv \mathbf{Tr}_{\Lambda_l}(\rho_{\Lambda_l}\,\cdot\,)$
$0 < \rho_{\Lambda_l} \leqq ||\rho_{\Lambda_l}|| < \infty$, $\Lambda_l \cap \Lambda_k = \emptyset$ for $k \neq l$.

It is known that if density matrices ρ_{Λ_l} are nontrivial, the corresponding state leads to type III von Neumann algebra in GNS representation. A nice thing about the infinite product state is that one has explicit formulas for $\mathbb{L}_p(\omega)$ norms on a dense set of local elements.

- $\mathbb{L}_p(\omega, s)$ norms

$$||f||^p_{\mathbb{L}_p(\omega,s)} \equiv \mathbf{Tr}\left|\ \rho_\Lambda^{\frac{1-s}{p}} f \rho_\Lambda^{\frac{s}{p}}\ \right|^p,$$

for $f \in \mathcal{A}_\Lambda$ with $\rho_\Lambda \equiv \prod_{\Lambda_l \cap \Lambda \neq \emptyset} \rho_{\Lambda_l}$.

with the scalar product for $\mathbb{L}_2(\omega)$ space given as follows

- $\mathbb{L}_2(\omega, s)$ scalar product

$$< f\,,\, g >_{\mathbb{L}_2(\omega,s)} \equiv \mathbf{Tr}\left(\rho_\Lambda^{\frac{1-s}{2}} f^* \rho_\Lambda^{\frac{s}{2}} g\right)\ .$$

In this setup we can introduce a Markov generator symmetric in $\mathbb{L}_2(\omega, s)$, $\forall s \in [0,1]$,

$$\mathcal{L}f \equiv \sum_{l\in\mathcal{R}} (E_{\Lambda_l}(f) - f)$$

defined with the following Generalised Conditional Expectation

$$E_X(f) \equiv \mathbf{Tr}_X \left(\xi^*_{\Lambda_l} f \xi_{\Lambda_l}\right) \ ,$$

where for a bounded set $X \subset \mathcal{R}$,

$$\xi_X \equiv \rho^{\frac{1}{2}}_{\Lambda_l} \left(\mathbf{Tr}_X \rho_{\Lambda_l}\right)^{-\frac{1}{2}} .$$

For dissipative semigroup P_t corresponding to the above given generator the following is true.

- **Hypercontractivity**: The Markov semigroup $P_t \equiv e^{t\mathcal{L}}$ satisfies

$$||P_t f||_{\mathbb{L}_{p(t)}(\omega,s)} \leq ||f||_{\mathbb{L}_2(\omega,s)} \qquad (\mathbf{H})$$

  for any $s \in [0,1]$ with $p(t) \equiv 1 + e^{\alpha t}$, with some $\alpha > 0$.

By the theory developed in[28] this implies the following relative entropy bound.

- **Weak product property**: For $s = \frac{1}{2}$

$$QEnt_2(f) \leq c_{\Lambda_0} \langle f, -\sum_{i\in\mathbb{Z}^d} (E_i(f) - f)\rangle_{\mathbb{L}_2(\omega,\frac{1}{2})} \qquad (\mathbf{LS})$$

  where

$$QEnt_p(f) \equiv \lim_{\Lambda\to\mathcal{R}} \mathbf{Tr} \left|\rho^{1/2p} f \rho^{1/2p}\right|^p \left(\log\left|\rho^{1/2p} f \rho^{1/2p}\right| - 1/p \log\rho\right) .$$

In other words one concludes in our noncommutative framework that if the Log-Sobolev inequality is true for a state on a finite dimensional algebra it is also true for any corresponding product state. In commutative case the product property is known since mid 70's and provide a fundamental basis of the related theory.

Remark : Here, perhaps because of the technical reasons, we obtain a finite constant, but not the same as for the finite dimensional state one started with. Clearing this up would be an interesting task.

Hypercontractivity $\iff$? Particle Structure :

In the noncommutative case the key strategy to prove the Logarithmic Sobolev inequality goes via the following interesting route involving spectral characteristics of the underlying theory.

Suppose we have the following Particle Structure

- Invariant Subspaces $\mathbb{L}_2(\mu) = \mathcal{H}_0 \oplus_{n\in\mathbb{N}} \mathcal{H}_n$

$$\mathcal{H}_j \perp \mathcal{H}_k,\ k \neq j \quad , P_t(\mathcal{H}_n) \subset \mathcal{H}_n$$

- Spectrum

$$\forall n \in \mathbb{N} \quad \sigma(\mathcal{L}|\mathcal{H}_n) \subset (-\infty, -\varepsilon n), \quad \varepsilon > 0$$

- Gaussian Bounds: $\exists C > 0\ \forall n \in \mathbb{N} \quad \forall \quad f \in \mathcal{H}_n$

$$||f||_4 \leq C^n ||f||_2,$$

In this situation corresponding semigroup is hypercontractive (**H**) and so the corresponding Log-Sobolev inequality (**LS**) is true.
While this direct route to hypercontractivity can sometimes be employed for some (simple) models with interactions[29], in general the particle structure remains one of the bigest challenges of mathematical physics exploration.

5. Nonlinear Jump Processes in $Dim = \infty$

In this section we outline briefly a recent progress in study of nonlinear jump type dissipative dynamics for infinite interacting systems.
Given probability kernels $E_j \equiv E_{\Lambda_0+j}$ on $\Omega = \mathbb{M}^{\mathbb{Z}^d}$, $\mathbb{M}$ a finite set and $\Lambda_0 \subset \mathbb{Z}^d$ a box, associated to a Gibbs measure, one can consider the following nonlinear generalisation of Glauber type generators.

Nonlinear Generator

$$\mathfrak{L}(v) \equiv \sum_{i\in\mathbb{Z}^d} \mathfrak{L}_i(v) \equiv \sum_{i\in\mathbb{Z}^d} \big(\psi^{-1} \circ E_i \circ \psi(v) - v\big)$$

with some (suitably chosen) convex increasing positive function ψ. One shows that the corresponding semigroup is well defined and has all the properties of the Markov semigroup (except linearity). Namely we have the following result.

Theorem 5.1. *The Cauchy problem*

$$\frac{\partial}{\partial t} v(t) \quad = \mathfrak{L}(v(t))$$

$$v(t=0) = f$$

has a unique solution given by a nonlinear Markov semigroup $(\mathfrak{P}_t)_{t\in\mathbb{R}^+}$ *satisfying*

- *Preservation of Constants*
$$\mathfrak{P}_t(\alpha\mathbb{1}) = \alpha\mathbb{1}$$
- *Positivity*
$$\mathfrak{P}_t(f^2) \geq 0$$
- *Contractivity*
$$||\mathfrak{P}_t(f)||_\infty \leq ||f||_\infty$$
- ⋆ *Super-invariance Property* $\quad \mu f \ \leq \ \mu\mathfrak{P}_t f$

We remark that in some class of similar operators, including the ones defined as above with $\psi(x) \equiv e^x$, when one has a single term in the generator one can write the following explicit formula which we call a nonlinear Poisson semigroup

$$\mathfrak{P}_t(f) = e^{-t}f + \int_0^t ds \log E \exp\{e^{-s}f\}.$$

From this one can see that the following further interesting properties are true

- *"Orthogonality"*
$$\mathfrak{P}_t(f + \alpha\mathbb{1}) = \mathfrak{P}_t(f) + \alpha\mathbb{1}$$
- *Convexity*
$\forall q \in [0,1] \qquad \mathfrak{P}_t(qf + (1-q)g) \leq q\mathfrak{P}_t(f) + (1-q)\mathfrak{P}_t(g)$

By a nonlinear generalisation of the product formula these properties transfer through to the infinite dimensional case.

Ergodicity

Under suitable uniqueness condition for the probability kernels one can extend the classical strong ergodicity results as follows.

Theorem 5.2. *Suppose a "Uniform Uniqueness Condition" is satisfied for* E_j, $j \in \mathbb{Z}^d$.
Then $\exists\, M \in (0,\infty)$

$$|||\mathfrak{P}_t f||| \leq e^{-Mt}|||f|||$$

with usual triple bar (Lipschitz) seminorms.

This implies the existence of limiting function $\alpha(f) \equiv \lim_t \mathfrak{P}_t f$. While it is a nonlinear functional, thanks to convexity we can talk about tangent functionals which brings us back to more conventional realm of probabilities.

6. Noncommutative Orlicz Spaces Associated to a State

Suppose we are given a faithfull normal state

$$\omega(f) \equiv Tr(\rho f)$$

and therefore the following family of scalar products

$$< f, g >_{\omega,s} \equiv Tr\left(\rho^s f^* \rho^{1-s} g\right)$$

for each $s \in (0,1)$. In such the situation the theory of noncommutative L_p spaces is well established[14–19]. While the theory of noncommutative Orlicz spaces associated to a trace where introduced already some time ago[20], only recently the more general cases were studied[13,23]. Below we briefly list some results and possible directions.

$\mathbb{L}_{p,s}(\rho)$ **Norms**, $p \in [1,\infty)$, $s \in [0,1]$

$$||f||^p_{p,s,\rho} \equiv Tr\left| \rho^{\frac{s}{p}} f \rho^{\frac{(1-s)}{p}} \right|^p$$

Orlicz Functionals I

$$\Xi(f) \equiv \int ||f||^p_{p,s,\rho} \, \nu(dp)$$

with $\int \lambda^p \, \nu(dp) < \infty$, for any $\lambda \in (0,\infty)$. **Orlicz Functionals II**

$$\Lambda_{\Phi,\rho,s}(f) \equiv Tr\Phi\left(\left|\left(\Phi^{-1}(\rho)\right)^s f \left(\Phi^{-1}(\rho)\right)^{1-s}\right|\right)$$

- Universal Normalisation $\Lambda_{\Phi,\rho,s}(\mathbf{1}) = 1$
- Strict Convexity $\forall |\alpha| + |\beta| \leqq 1$

$$\Lambda_{\Phi,\rho,s}(\alpha f + \beta g) \leqq |\alpha| \Lambda_{\Phi,\rho,s}(f) + |\beta| \Lambda_{\Phi,\rho,s}(g)$$

- Duality and Young Inequality

$$< f, g >_{\omega,s} \leqq \Lambda_{\Phi,\rho,s}(f) + \Lambda_{\Psi,\rho,s}(f)$$

where $\Psi(x) = \sup_{y \in \mathbb{R}} (xy - \Phi(y))$

Orlicz Functionals III

$$\mathcal{O}_{A,\Phi,\rho,s}(f) \equiv \frac{1}{2}\Lambda_{\Phi,\rho,s}(f + A) + \frac{1}{2}\Lambda_{\Phi,\rho,s}(f - A) - 1$$

where $\Lambda_{\Phi,\rho,s}(A) = 1$.
Similar Properties as in I.

Bogoliubov-Kubo-Mori Scalar Product and Generalised Young Inequality.
In case of the following scalar product which plays an important role in nonequilibrium quantum statistical mechanics

$$< f, g >_{BKM} \equiv \int_0^1 ds\ Tr\left(\rho^s f^* \rho^{1-s} g\right)$$

it is natural to consider the following generalised Young inequality

$$< f, g >_{BKM} \leqq \int_0^1 ds\ \{\Lambda_{\Phi_s,\rho,s}(f) + \Lambda_{\Psi_s,\rho,s}(f)\ \}$$

where $\Lambda_{\Phi_s,\rho,s}$ is a family of Orlicz functional introduced above.

Streater's Functionals[23]

$$\frac{1}{2} Tr e^{log\rho + X} \ + \ \frac{1}{2} Tr e^{log\rho - X} - 1$$

Other Orlicz Functionals
One should expect that the following functionals are interesting.

- Functional with Renormalised Logarithm ([28])

$$N(f) \ \equiv \ \sup_{c\in\mathbb{R}} \Bigg\{ Tr \Big|\ \rho^{\frac{1}{4}}\ (f+c)\ \rho^{\frac{1}{4}}\ \Big|^2 \left(\log \Big|\ \rho^{\frac{1}{4}}\ (f+c)\ \rho^{\frac{1}{4}}\ \Big|^2 - \log \rho \right)$$

$$- \ ||f+c||^2_{2,1/2,\rho} \log ||f+c||^2_{2,1/2,\rho} \Bigg\}$$

$$= \ \sup_{c\in\mathbb{R}} \Bigg\{ Tr \Big|\ \rho^{\frac{1}{4}}\ (f+c)\ \rho^{\frac{1}{4}}\ \Big|^2 \left(\log \Big(\ \rho^{\frac{1}{2}}\ \Big|\alpha_{i\frac{1}{4}}(f) + c\Big|^2\ \rho^{\frac{1}{2}} \Big) - \log \rho \right)$$

$$- \ ||f+c||^2_{2,1/2,\rho} \log ||f+c||^2_{2,1/2,\rho} \Bigg\}$$

This kind of relative entropy functionals show up in the study of the flow through a bundle of noncommutative L_p spaces associated to the state, which is relevant for the hypercontractive semigroups in such the setup[28].

- Relative Entropies[26,27].

$$H(f) \equiv \sup_c Tr\left(\rho^{1/2}\eta(\Delta_{P_{f+c},\rho})\rho^{1/2}\right).$$

where $P_f \equiv \rho^{\frac{1}{2}} f \rho^{\frac{1}{2}}$ and $\Delta_{P,Q} \equiv L_P R_Q^{-1}$ with L_P and R_P being left and right multiplication operators, respectively, and where η is an operator convex function on $(0,\infty)$ such that $\eta(1) = 0$.
This convex and homogeneous functionals originate from the noncommutative information geometry ([27] and references therein).

7. Entropic Switch and Phase Transition in Orlicz Spaces

In this section we discuss briefly phenomena arising from an interplay of entropy and nonlinearity. As the entropy is an addive quantity, one can guess that the phenomena can also be sensitive to the size of the system. In particular this can be relevant to the class of Orlicz functionals which we discuss at the end in a commutative setup (having some other noncommutative flavour). We illustrate what can happen on the example of a simple nonlinear quantity, called a iota index, which is defined as follows.

Definition 7.1. (Index $\boldsymbol{\iota}$) For probability density ρ on $\mathbb{R}^n$ and an Orlicz function Φ, satisfying the doubling property $\exists C \in (0,\infty)\forall z \in \mathbb{R}\ \Phi(2z) \leq C\Phi(z)$, we define

$$\boldsymbol{\iota}_\rho(\Phi) = \int_0^\infty \int_{\mathbb{R}^n} \frac{1}{s}\Phi\left(s\Phi^{-1}(\rho(x))\right)\ d_n x\ e^{-s}\ ds$$

The index is well defined for all densities ρ. For monomials it is independent of the density ρ and is equal to the usual Euler gamma function [a]. Otherwise it has the following properties.

7.1. *Shannon Entropy versus Index $\boldsymbol{\iota}$*

(i) Index is sensitive to Entropy: $\forall c \in (0,\infty)$ let $\rho_c = \frac{1}{c}\rho(c^{-1/d}x)$

$$\mathrm{S}(\rho_c) = \mathrm{S}(\rho) + \log c.$$

Hence

[a] Note that ι is given by the group commutator; i.e. is also a measure of noncommutativity.

$$\mathrm{S}(\rho_c) \longrightarrow_{c\longrightarrow 0} -\infty \qquad \text{and} \qquad \mathrm{S}(\rho_c) \longrightarrow_{c\longrightarrow \infty} +\infty$$

and we have

$$S(\rho_\sigma) \to \pm\infty \implies \boldsymbol{\iota}(\rho_\sigma, \Phi) \to \boldsymbol{\Gamma}_{_{Euler}}(\kappa^{\pm})$$

where $\boldsymbol{\Gamma}_{_{Euler}}(\kappa^{\pm})$ denotes the Euler's Gamma function with $\kappa^{\pm}$ being Karamata-Matuszewska indices of Φ, that is the leading exponents describing behaviour of the Orlicz function near zero and infinity, respectively[b].

(ii) For a fixed normalised signal $\widetilde{\rho}$, the entropy of its mixture with a signal ρ_c is asymptotically determined by the stretched density. This is because for any $\alpha \in (0,1)$ one has

$$\alpha\mathrm{S}(\widetilde{\rho})+(1-\alpha)\mathrm{S}(\rho_c) \;\leqq\; \mathrm{S}(\alpha\widetilde{\rho}+(1-\alpha)\rho_c) \;\leqq \alpha\mathrm{S}(\widetilde{\rho})+(1-\alpha)\mathrm{S}(\rho_c) +\mathrm{S}(\alpha,1-\alpha),$$

where $S(\alpha, 1-\alpha)$ denotes the entropy of Bernoulli distribution ($p = \alpha, q = 1-\alpha$), and one has

$$\mathrm{S}(\alpha\widetilde{\rho}+(1-\alpha)\rho_c) \longrightarrow_{c\longrightarrow 0} -\infty \quad \text{and} \quad \mathrm{S}(\alpha\widetilde{\rho}+(1-\alpha)\rho_c) \longrightarrow_{c\longrightarrow \infty} +\infty$$

It is different in case of the index

$$\lim_{c\longrightarrow 0} \iota(\Phi, \alpha\widetilde{\rho} + (1-\alpha)\rho_c) \;=\; \lim_{c\longrightarrow 0} \iota(\Phi, \rho_c)$$

while

$$\lim_{c\longrightarrow \infty} \iota(\Phi, \alpha\widetilde{\rho} + (1-s)\rho_c) \;=\; \frac{1}{\alpha}\iota(\Phi, \alpha\widetilde{\rho}) \neq \lim_{c\longrightarrow \infty} \iota(\Phi, \rho_c)$$

where the index ι for non-normalised signal $\alpha\widetilde{\rho}$ is defined formally by the formula introduced before.

(iii) (*Collective Switching*) $\rho_N(\mathbf{x}) \equiv \prod_{j=1,..,N} \rho_j(x_j)$

$$\mp \inf_N \frac{1}{N} S(\rho_N) > 0 \implies \boldsymbol{\iota}(\rho_N, \Phi) \to \boldsymbol{\Gamma}_{_{Euler}}(\kappa^{\pm})$$

Many systems originate from a single excited cell (e.g. biological organisms or business companies), which divide into descendant cells or accommodate others (e.g. by buying other companies). During the growth of number of cells, if the average entropy does not remain balanced, the nonlinear control parameter ι is changing its value.

[b]One can see that Karamata-Matuszewska indices give rise to a natural traces on the group algebra Δ_2 (which have some other interesting applications in connection to **Topo**logic **No**nlinear **Ne**tworks)

(iv) (*Ideal Balance Principle*)

$$\frac{1}{N}S(\rho_N) \approx 0 \implies$$

$$\min\left\{\mathbf{\Gamma}_{_{Euler}}(\kappa^{\pm})\right\} \ < \ \boldsymbol{\iota}_{\rho_N}(\Phi) \ < \ \max\left\{\mathbf{\Gamma}_{_{Euler}}(\kappa^{\pm})\right\}$$

A system with the global nonlinear control can grow without being switched away, provided its entropy remains well balanced.

7.2. *Phase Transitions in Orlicz Spaces*

Motivated by our earlier examples, consider the following Orlicz functional on the algebra of functions.

$$\mathbb{I}_{\rho,\Phi}(f) \equiv \int_{\mathbb{R}^d} \Phi\left(\Phi^{-1}(\rho(x))f(x)\right)\ dx$$

with a probability density ρ and similarly for complementary function. Such functionals possess the following *Universal Normalisation Property*

$$\mathbb{I}_{\rho,\Phi}(\mathbf{1}) = 1$$

which the noncommutative L_p spaces associated to a state naturally do have. In case of (commutative) family of monomials we obtain the usual L_p type functionals.[c]

We mention first that for complementary Orlicz functions Φ and Ψ^{-1} and any $\rho \geqq 0$, one has the following inequality

$$\rho \leqq \Phi^{-1}(\rho)\Psi^{-1}(\rho) \leqq 2\rho$$

Therefore the following Young type inequality is true $\forall \lambda, \vartheta \in \mathbb{R}$.

$$|\lambda\vartheta|\rho \leqq \Phi^{-1}(\rho)|\lambda| \cdot \Psi^{-1}(\rho)|\vartheta| \leqq \Phi(\Phi^{-1}(\rho)\lambda) + \Psi(\Psi^{-1}(\rho)\vartheta)$$

This in turn implies the following *Young Inequality* for functions

$$\left|< f, g >_{L_2(\rho dx)}\right| \leq \mathbb{I}_{\rho,\Phi}(f) + \mathbb{I}_{\rho,\Psi}(g)$$

[c]In case when Φ are not monomials, similarly to the quantum L_p spaces, one can introduces the horizontal interpolation of functionals

$$I_{\Phi,s}(f) \equiv \int \Phi(f\Phi^{-1}(\rho^s))\rho^{1-s}dx$$

for any $s \in [0,1]$.

To study large dimensional behaviour consider a product type density $\rho_n \equiv \rho^{\otimes n}$. Then for $\Phi \in \Delta_2$ with $\kappa_\Phi \neq \mathrm{K}_\Phi$, we have corresponding functional

$$\mathbb{I}_{n,\Phi}(f) \equiv \int_{(\mathbb{R}^d)^n} \Phi\left(\Phi^{-1}(\rho_n(\mathbf{x}))f(\mathbf{x})\right)\, d\mathbf{x}$$

Given a smooth cylinder function dependent on given m coordinates $\mathbf{y}$, we have

$$\mathbb{I}_{n,\Phi}(f) \equiv \int_{(\mathbb{R}^d)^m} \left(\int_{(\mathbb{R}^d)^k} \Phi\left(\Phi^{-1}(\rho_m(\mathbf{y})\rho_n(\mathbf{x}))f(\mathbf{y})\right)\, d\mathbf{x} \right)\, d\mathbf{y}$$

$$= \int_{(\mathbb{R}^d)^m} \left(\int_{(\mathbb{R}^d)^k} \frac{\Phi\left(\Phi^{-1}(\rho_m(\mathbf{y})\rho_k(\mathbf{x}))f(\mathbf{y})\right)}{\Phi\left(\Phi^{-1}(\rho_m(\mathbf{y})\rho_k(\mathbf{x}))\right)}\, \nu_{\geqq k}\left(d\mathbf{x}\right) \right)\, \nu\left(d\mathbf{y}\right)$$

where $m + k = n$ while $\nu\left(d\mathbf{x}\right)$ and $\nu_{\geqq k}\left(d\mathbf{x}\right)$ is the infinite product of measures ρdx and related conditional expectation given variables $\mathbf{y}$. For given point $\mathbf{y}$, setting $\lambda \equiv \rho_m(\mathbf{y})$ and $z \equiv f(\mathbf{y})$, the integrand we need to study has the following form

$$\frac{\Phi\left(\Phi^{-1}(\lambda\rho_k(\mathbf{x}))z\right)}{\Phi\left(\Phi^{-1}(\lambda\rho_k(\mathbf{x}))\right)}$$

This allows us to conclude with the following claim.

Let ν be an infinite product of measures ρdx. Then the following is true for any cylinder function f.

(i) *If* $\mathrm{H}(\rho) > 0$, *then*

$$\lim_{n \longrightarrow \infty} \mathbb{I}_{n,\Phi}(f) = \int |f|^{\kappa_\Phi} d\nu$$

(ii) *If* $\mathrm{H}(\rho) < 0$, *then*

$$\lim_{n \longrightarrow \infty} \mathbb{I}_{n,\Phi}(f) = \int |f|^{\mathrm{K}_\Phi} d\nu$$

(iii) *If* $\mathrm{H}(\rho) = 0$, *then*

$$\lim_{n \longrightarrow \infty} \mathbb{I}_{n,\Phi}(f) = \frac{1}{2}\int |f|^{\kappa_\Phi} d\nu + \frac{1}{2}\int |f|^{\mathrm{K}_\Phi} d\nu$$

where in the above κ_Φ and K_Φ denote the exponent of Φ at 0 and ∞, respectively.

Consider now a family of Orlicz functions $\{\Phi_\theta \in \Delta_2\}_{\theta\in\Theta}$. We note that

$$|\lambda\vartheta|\rho \leqq \int \Phi_\theta(\Phi_\theta^{-1}(\rho)\lambda)\ \mathbf{m}(d\theta) + \int \Psi_\theta(\Psi_\theta^{-1}(\rho)\vartheta)\ \mathbf{m}(d\theta)$$

where the functional

$$\mathbb{I}_{\rho,\mathbf{m},\Theta}(f) \equiv \int \Phi_\theta(\Phi_\theta^{-1}(\rho(x))f(x))\ dx\ \mathbf{m}(d\theta)$$

is nonnegative and convex (and similarly for Ψ_θ family).
Performing the infinite dimensional limit we arrive at

$$\mathbb{I}_{\mu,\mathbf{m},\Theta}(f) \equiv \int |f|^{\mathrm{K}(\theta)}\ d\mu\ \mathbf{m}(d\theta)$$

where $\mathrm{K}(\theta) \equiv \mathrm{K}(\mathfrak{h}, \Phi_\theta)$ is the exponent dependent on the mean entropy $\mathfrak{h}$ of the infinite system and the exponents of Φ_θ. Changing the order of integration, which can be justified under general assumptions, we can represent our functional as follows

$$\mathbb{I}_{\mu,\mathbf{\Phi}}(f) \equiv \int \mathbf{\Phi}(f)\ d\mu$$

where $\mathbf{\Phi}$ can be represented as the following Mellin transform

$$\mathbf{\Phi}(f) \equiv \int X_p(f)\ \widetilde{\mathbf{m}}(dp)$$

with

$$\widetilde{\mathbf{m}}(dp) \equiv \Gamma(p+1)\mathbf{m}(d\mathrm{K}^{-1}(p))$$

Thus in this case we have a possibility to get a nontrivial Orlicz functional at infinity.

Bibliography

1. P. Biane, Free hypercontractivity. Commun. Math. Phys. 184 (1997) 457-474.
2. M. Bożejko, B. Kümmerer & R. Speicher, q-Gaussian processes: non-commutative and classical aspects. Commun. Math. Phys. 185 (1997) 129-154.
3. E. Carlen & E. Lieb, Optimal hypercontractivity for Fermi fields and related noncommutative integration inequalities. Commun. Math. Phys. 155 (1993) 26-46.
4. L. Gross, Hypercontractivity and logarithmic Sobolev inequalities for the Clifford- Dirichlet form. Duke Math. J. 42 (1975) 383-396.
5. T. Kemp, Strong hypercontractivity in non-commutative holomorphic spaces. Preprint.
6. I. Królak, Contractivity properties of the Ornsetein-Uhlenbeck semigroup for general commutation relations. To appear in Mathematische Zeitschrift.

7. J.M. Lindsay, Gaussian hypercontractivity revisited. J. Funct. Anal. 92 (1990) 313-324
8. J.M. Lindsay, P.A. Meyer, Fermionic hypercontractivity. In: Quantum probability VII, Singapore: Wold Scientific, 1992, pp. 211-220
9. P. Fougères & B. Zegarliński, Semi-Linear Problems in Infinite Dimensions, J. Funct. Analysis 228 (2005) 39-88.
10. S. Bobkov & B. Zegarliński, Entropy Bounds and Isoperimetry, Memoires AMS 2005
11. A. Guionnet & B. Zegarliński, Lectures on Logarithmic Sobolev Inequalities, Séminaire de Probabilités XXXVI, Lec. Notes in Math. 1801, Springer 2003.
12. P.Ługiewicz & B. Zegarliński, Coercive Inequalities for Hörmander Type Generators in Infinite Dimensions IC05.
13. M. H. A. Al-Rashed & B. Zegarliński, Noncommutative Orlicz Spaces Associated to a State, **I** to appear in Studia Mathematica; **II** IC2006 Preprint.
14. U. Haagerup , L_p-spaces associated with an arbitrary von Neumann algebra in *Algèbres d'opérateurs et leurs applications en physique mathématique*, Colloques internationaux du CNRS, No. 274, Marseille 20-24 juin 1977, 175-184. Éditions du CNRS, Paris 1979.
15. M. Hilsum, Les espaces L_p d'une algèbre de von Neumann définies par la dérivée spatiale, J. Func. Anal. 40 (1981), 151-169.
16. H. Kosaki, Application of the complex interpolation method to a von Neumann algebra (Non-commutative L_p-spaces), J. Func. Anal. 56 (1984), 29-78.
17. N. V. Trunov , On a noncommutative analogue of the space L_p. Izvestiya VUZ. Matematika. 23 (1979), 69-77. (Soviet Math. Transl.)
18. F. J. Yeadon, Non-commutative L_p-spaces, Math. Proc. Camb. Phil. Soc. 77 (1975), 91-102.
19. R. A. Kunze , L_p Fourier transforms on locally compact unimodular groups, Trans. Amer. Math. Soc. 89 (1958) 519-540.
20. W. Kunze , Noncommutative Orlicz Spaces and Generalized Arens Algebras, Mth. Nachr. 147(1990) 123-138.
21. W. A. Majewski, On completely positive nonlinear dynamical semigroups, J. Phys. A: Math. Gen. 23 (1990) L359
22. M. Czachor, M. Kuna, S.B. Leble & J. Naudts, Nonlinear von Neumann type equations, quant-ph/9904110.
23. R.F. Streater, Quantum Orlicz Spaces in Information Geometry, KCL 2004.
24. Yong Moon Park, Remarks on the Structure of Dirichlet Forms on Standard Forms of von Neumann Algebras, arXiv:math-ph/Prep 2004.
25. Yong Moon Park, Ergodic Property of Markovian Semigroups on Standard Forms of von Neumann Algebras, arXiv:math-ph/Prep 2005.
26. D. Petz, Quasi-entropies for finite quantum systems, Rep. Math. Phys. 23 (1986) 57–65.
27. A. Leśniewski & M.B. Ruskai, Monotone Riemannian metrics and relative entropy on noncommutative probability spaces, J. of Math. Phys. 40 (1999) 5702-5724
28. R. Olkiewicz & B. Zegarliński , Hypercontractivity in Noncommutative L_P Spaces, J. of Funct. Analysis 161(1999) 246-285

29. T. Bodineau & B. Zegarliński, Hypercontractivity via Spectral Theory, Infinite Dim. Analysis, Quantum Probab. and Rel. Tops. 3 (2000) 15-31.
30. W. Hebisch, R. Olkiewicz & B. Zegarliński, On upper bound for the quantum entropy, Linear Algebra and its Appl. 329 (2001) 89-96.
31. Lihu Xu , R. Olkiewicz & B. Zegarliński, Nonlinear Problems in Infinite Interacting Particle Systems, preprint IC 2006 & Noncommutative Work in Progress.
32. A.W. Majewski & B. Zegarliński, Quantum Stochastic Dynamics II, Rev. Math. Phys. 8, Nr5 (1996) 689-713.
33. F. Barthe, P. Cattiaux & C. Roberto. Interpolated inequalities between exponential and gaussian, orlicz hypercontractivity and isoperimetry. To appear in Revista Matemática Iberoamericana, 2005.
34. F. Martinelli, E. Olivieri, R.H. Schonmann, For 2-D lattice spin systems Weak Mixing Implies Strong Mixing, Comm. Math. Phys. 165, no. 1 (1994), 33-47.
35. C. Roberto & B. Zegarliński, Orlicz-Sobolev inequalities for sub-Gaussian measures and ergodicity of Markov semi-groups, J. Func. Anal. 2007
36. N. Varopoulos, L. Saloff-Coste, T. Coulhon Analysis and Geometry on Groups (Cambridge Tracts in Mathematics) Cambridge University Press 1992.
37. B. Zegarliński, Analysis of Classical and Quantum Interacting Particle Systems, pp 241 - 336 in *Quantum Interacting Particle Systems*, Eds L. Accardi and F. Fagnola, World Sci 2002
38. B. Zegarliński, Entropic Switch, *Essays for Science Studies* Vol. I., *http://matrixpress.co.uk*

PART B

Quantum Statistics, Filtering and Control

This page intentionally left blank

Quantum Filtering and Optimal Control

VIACHESLAV P. BELAVKIN and SIMON EDWARDS

School of Mathematical Sciences
University of Nottingham, UK
E-mail: vpb@maths.nottingham.ac.uk
URL: http://www.maths.nott.ac.uk/personal/vpb/

Quantum mechanical systems exhibit an inherently probabilistic nature upon measurement which excludes in principle the singular direct observability case. Quantum theory of time continuous measurements and quantum filtering developed by VPB on the basis of semi-Markov independent increment models for quantum noise and quantum nondemolition (QND) observability is generalized for demolition indirect measurements of quantum unstable systems satisfying the microcausality principle. The reduced quantum feedback-controlled dynamics is described both by linear semi-Markov and nonlinear conditionally-Markov stochastic master equations. Using this scheme for diffusive and counting measurement to describe the stochastic evolution of the open quantum system under the continuous indirect observation and working in parallel with classical indeterministic control theory, we show the conditionally-Markov Bellman equations for optimal feedback control of the *a posteriori* stochastic quantum states conditioned upon these measurements. The resulting Bellman equation for the diffusive observation is then applied to the explicitly solvable quantum linear-quadratic-Gaussian (LQG) problem which emphasizes many similarities with the corresponding classical control problem.

Keywords: Quantum Trajectories, Quantum Probability, Quantum Stochastics, Quantum Filtering, Quantum Control

1. Introduction

This paper introduces a new *quantum stochastic* framework of *semi-Markov processes* for study the problems of quantum observation and feedback control and outlines their solutions. In orthodox quantum mechanics, which treats only closed Hamiltonian quantum dynamics of *unobserved* microsystems, there is no observation problem since the measurement, even if it is considered for an interpretation, is always shifted to the final stage where is assumed to be performed instantaneously. However, there is no instantaneous measrement in the nature and it is meaningless to consider the closed

quantum dynamics of a system under feedback control without opening it for continuous observation. This requires a solution of the dynamical problem of quantum measurement by reconciling the deterministic continuous dynamics with quantum state diffusions and jumps under the continuous observation which was realized in.[4,6,7,9,11]

With technological advances now allowing the possibility of continuous monitoring and rapid manipulations of systems at the quantum level[2,47] there is an increasing awareness of the importance of quantum feedback control in applications of quantum information; such as the dynamical problems of quantum error corrections for quantum computations. The theory of quantum feedback control formally developed by Belavkin for these puposes in the 80's has recently been applied in many contexts including state preparation,[36,45,63] purification,[52,53] risk-sensitive control[54,55] and quantum error correction.[1,49] It has also been studied from the practical point of view of stability theory[64] and it gives along with[46] a comprehensive discussion of the comparison of classical and quantum control techniques.

The main ingredients of quantum control are essentially the same as in the classical case. One controls the system by coupling to an external control field which modifies the system in a desirable manner. The desired objectives of the control can be encoded into a *cost function* along with any other stipulations or restrictions on the controls such that the minimization of this cost indicates optimality of the control process. The most interesting case of quantum *closed-loop* or *feedback* control is based on the theory of quantum stochastic processes and their indirect continuous observations developed by Belavkin in a series of papers.[6,8,15,21,25] This laid down the foundation of quantum prediction and nolinear filtering theory[9,11,18,25] in analogy to the classical stochastic prediction theory which is based upon the nonlinear (Stratonovich) filtering equations. This work was continued at the beginning of 90's[17,19,20,25,26,28,38] in the direction of the foundations for a new quantum stochastic mechanics with trajectories of continual observations giving a microscopic theory of continuous reductions[24,31] and spontaneous localizations[30,39] for quantum diffusions,[13,16] quantum jumps[12,37] and other mixed stochastic quantum trajectories.[3,18,29]

In order to demonstrate the power of this new event enhanced quantum mechanics in a most effective way, it was applied right from the beginning[7,9,11] to solve the typical problems of quantum feedback control in parallel to the work on classical stochastic control with partial observations first introduced by Stratonovich.[59,61] The problem of optimal quantum feedback control was separated into quantum filtering which provides

optimal estimates of the quantum state variables (density operators), and a classical optimal control problem based on the output of the quantum filter. The classical noise, which is filtered out by passing from the prior to the posterior quantum states, comes from the irreducible disturbance to the quantum system during observation (a consequence of the interaction with measurement apparatus). Unlike the classical case this is an unavoidable feature of quantum measurement, since the state of an individual quantum system is not an observable but hidden.

The purpose of this paper is to build on the original work of the first author and present an accessible account of the theory of quantum continual measurements, quantum causality and predictions and optimal quantum feedback control in a natural generality of semi-unitary feedback controlled evolutions and semi-Markov quantum stochastic processes. Firstly we introduce the necessary concepts and mathematical tools from modern quantum theory including quantum probability, continuous causal (non-demolition) measurements, quantum stochastic calculus, and quantum filtering. Next the quantum Bellman equations for optimal feedback control with diffusive and counting measurement schemes are informally derived. The latter results were first stated in[11] without derivation and a consideration for the diffusive case was recently given in.[62] We conclude with an application of these results to the multi-dimensional quantum Linear-Quadratic-Gaussian (LQG) problem, and a discussion of its comparison with the corresponding classical results. However, we first start from a simple example of single mode quantum stochastic linear open system[6] which admits continuous obsrvation, filtering and the feedback control. It allows us to set up notations and clearly demonstrates not only the similarity but also the difference of classical and quantum feedback control theories which can be observed in the *microduality principle*, a more elaborated duality between quantum linear Gaussian filtering and classical linear optimal control.

1.1. *A controlled quantum particle model*

The quantum linear filtering and optimal quadratic control with additive quantum Gaussian noises was first studied and resolved by Belavkin in a series of quantum measurement and filtering problems[4,6,24] and based on these quantum feedback control papers.[5,9,11] It was noticed that, in order to realize continuously quantum trajectory measurements for a fundamentally unobservable noncommutative Heisenberg coordinate process $\mathfrak{q}(t)$, one should open the quantum system by coupling it with a classical stochastic observable process. The simplest example of a single quantum

Gaussian oscillator matched with a pair of transmission lines, reducible in the sense of[6] to a complex one-dimensional channel, was taken as the first quantum feedback model in the preprint.[5] However, more similar to the classical case irreducible quantum linear models require at least two real dimensions instead of a single complex one. In order to demonstrate this similarity we may now use the multidimensional quantum linear model derived in the last part of this paper for application on higher dimensional systems which may not have the reduced complex representation. The continuously observed and controlled quantum particle in one dimension is the simplest such example.

In the traditional approach to quantum mechanics the canonical variables q, p of a classical object (particle) phase state $s_{\rm o} = (q, p)$ are replaced by selfadjoint operators $\mathtt{s}_{\rm o} = (\mathtt{q}, \mathtt{p})$ of position $\mathtt{s}_q = \mathtt{q}$ and momentum $\mathtt{s}_p = \mathtt{p}$ satisfying the canonical commutation relation (CCR)

$$[\mathtt{q}, \mathtt{p}] := \mathtt{qp} - \mathtt{pq} = \mathrm{i}\hbar \mathtt{I}^{\rm o}. \tag{1}$$

Here $\mathtt{I}^{\rm o}$ is the identity operator on a Hilbert space $H_{\rm o}$ of the CCR representation (1) and $\hbar$ is Planck's constant, which for our purpose could be any constant $\hbar \neq 0$. The Hamiltonian $\mathtt{p}^2/2\mu$ of free mass $\mu \geq 0$ is perturbed as $\mathtt{H}^u = \mathtt{p}^2/2\mu + \gamma u \mathtt{q}$ by a control $u(t) \in \mathbb{R}$ through the potential $\phi(t, \mathtt{q}) = \gamma u(t)\, \mathtt{q}$.

The particle is assumed to be coupled not only to the forward channel of the controlling force $f = -\gamma u$, but also to an observation channel locally represented on increasing histories Fock spaces $F_0^t = \Gamma(\mathcal{E}_0^t)$ of the observation intervals $(0, t]$. This opens the free particle Heisenberg dynamics $\dot{\mathtt{q}} = \mathtt{p}/\mu$, $\dot{\mathtt{p}} = -\gamma u \mathtt{I}$ perturbing the Schrödinger's unitary evolution operators $\mathrm{U}_{\rm o}^t$ on $H_{\rm o}$ to *casual interaction* unitaries $\mathrm{V}_t : \mathcal{H} \to H_t$ mapping adaptively the Hilbert product components $F_{[0}^t = H_{\rm o} \otimes F_0^t$ of $\mathcal{H} = H_{\rm o} \otimes F_0$ onto the components $F_0^{t]} = F_0^t \otimes H_{\rm o}$ of $H_t = F_0^{t]} \otimes F_t$, where $F_t = \Gamma(\mathcal{E}_t)$ is the Fock space of future after t. The perturbed Heisenberg dynamics $\hat{\mathrm{s}}_{\rm o}(t) = \mathrm{V}_t^\dagger \left(\mathrm{s}_{\rm o}^{t]} \otimes \mathrm{I}_t\right) \mathrm{V}_t$, where $\mathrm{s}_{\rm o}^{t]} = \mathrm{I}_0^t \otimes \mathtt{s}_{\rm o}$, is described by the adapted operators $\hat{\mathrm{s}}_{\rm o}(t) = \hat{\mathrm{s}}_{\rm o}^{t]} \otimes \mathrm{I}_t$ on $\mathcal{H} = F_{[0}^t \otimes F_t$ satisfying the *canonical pair of linear Langevin equations*

$$\mathrm{d}\hat{\mathrm{q}}^t + \lambda \hat{\mathrm{q}}(t)\, \mathrm{d}t = \frac{1}{\mu} \hat{\mathrm{p}}(t)\, \mathrm{d}t + \mathrm{dv}_q^t, \ \hat{\mathrm{q}}(0) = \mathtt{q} \otimes \mathrm{I}_0 \equiv \mathtt{q}, \tag{2}$$

$$\mathrm{d}\hat{\mathrm{p}}^t + \lambda \hat{\mathrm{p}}(t)\, \mathrm{d}t = \mathrm{du}_p^t - \gamma \widehat{u}(t)\, \mathrm{d}t, \quad \hat{\mathrm{p}}(0) = \mathtt{p} \otimes \mathrm{I}_0 \equiv \mathtt{p}, \tag{3}$$

as a case of (101). Here $\mathrm{d}\mathtt{x}^t$ denotes the forward increment of any process $\mathtt{x}(t)$, $\mathtt{u}_p = \mathtt{I}^{\rm o} \otimes \mathrm{u}_p$, $\mathtt{v}_p = \mathtt{I}^{\rm o} \otimes \mathrm{v}_q$ are quantum noises adaptively represented

by independent increment processes $\mathrm{u}_p(t)$, $\mathrm{v}_p(t)$ on F_0^t and $\widehat{u}(t)$ is given by a feedback control low which will be specified later, $\widehat{u}(t) = u(t)\,\mathtt{I}$ if no feedback, with the identity operator $\mathtt{I} = \mathrm{I}_{[0}^t \otimes \mathrm{I}_t$ on $\mathcal{H}$, where I_t and $\mathrm{I}_{[0}^t$ are the identities respectively on F_t and $F_{[0}^t$. The parameter $\lambda = \frac{1}{2}(\beta\varepsilon + \gamma\varphi)$ is defined by the real coefficients of linear combinations $\mathrm{u}_p = \beta\mathrm{u}_e + \gamma\mathrm{u}_f$, $\mathrm{v}_q = -\varepsilon\mathrm{v}_e - \varphi\mathrm{v}_f$ given by two independent pairs $(\mathrm{u}_j, \mathrm{v}_j)$ of *canonical noncommuting Wiener noises*, where the suffix $j = e, f$ stands for measurement error and perturbation force respectively in the observation and control coherent channels.

The pairs $(\mathrm{u}_j, \mathrm{v}_j)$ can be realized on vacuum quantum channel states by the Hermitian parts $\mathrm{u}_j = \hbar\,\mathsf{Im}\left[\mathrm{A}_j^+\right] = \mathrm{u}_j^\dagger$, $\mathrm{v}_j = 2\,\mathsf{Re}\left[\mathrm{A}_j^+\right] = \mathrm{v}_j^\dagger$ of quantum creation processes $\mathrm{A}_j^+ = \frac{1}{2}\mathrm{v}_j + \frac{\mathrm{i}}{\hbar}\mathrm{u}_j$ in the Fock Hilbert space F_0 at any $t > 0$. The forward increments $\mathrm{d}\mathrm{A}_j^+(t)$ are independent of the processes $\mathrm{A}_j^+(t)$ such that the quanta created on each interval $[0, t)$ due to forward and backward actions onto the quantum particle from the control and observation channels are spontaneously emitted independently of the quanta created on $[t, t + \mathrm{d}t)$. Instead of trying to explain the physical origin of the noncommutativity of these actions, we simply note that the Wiener noise combinations u_p, v_q must satisfy the following CCR relations

$$\mathrm{v}_q(s)\,\mathrm{u}_p(r) - \mathrm{u}_p(r)\,\mathrm{v}_q(s) = (r \wedge s)\,\mathrm{i}\hbar\lambda\mathrm{I} \tag{4}$$

as necessary and sufficient condition for preservation $[\hat{\mathrm{q}}(t), \hat{\mathrm{p}}(t)] = \mathrm{i}\hbar\mathtt{I}$ of the CCR (1) by the system (2), (3) due to mutual commutativity of the independent pairs (q, p) and $(\mathrm{v}_q, \mathrm{u}_p)$ on $\mathcal{H}$.

By integrating (2) with the solution

$$\hat{\mathrm{p}}(t) = \mathrm{e}^{-\lambda t}\mathtt{p} + \int_0^t \mathrm{e}^{(s-t)\lambda}\left(\mathrm{d}\mathrm{u}_p - \gamma\widehat{u}(s)\,\mathrm{d}s\right)$$

of (3) it can be easily found from (1) that

$$[\hat{\mathrm{p}}(r), \hat{\mathrm{q}}(s)] = \frac{\mathrm{i}\hbar}{\mu}\,|r - s|\,\mathrm{e}^{-\lambda|r-s|}\mathtt{I} \neq 0.$$

Therefore the family $\{\hat{\mathrm{s}}_q(t)\}$ of selfadjoint quantum trajectory operators is incompatible, and therefore cannot be represented as a classical stochastic process and directly observed.

1.2. *Quantum trajectory measurement model*

In order to resolve the fundamental quantum causality problem to enable statistical predictions and filtering of unobservable quantum stochastic trajectories, Belavkin noticed in[5,6] that the quantum output processes in open

quantum systems can be directly observed by continuous measurements as in the classical case if the additive error noises satisfy an error-perturbation CCR relation with the quantum Langevin forces in the quantum stochastic equations like (2), (3). For example, in order to monitor and predict continuously in time the present and future quantum particle trajectory $\hat{\mathrm{s}}_q(t)$ in the channel $\dot{\mathtt{y}} = \beta\hat{\mathrm{s}}_q + \dot{\mathrm{v}}$ with the additive standard white noise $\dot{v}(t)$ represented in $\mathcal{H}$ by a generalized operator field $\dot{\mathrm{v}}(t)$, it should satisfy the CCR relation $\left[\mathrm{u}_p(r), \dot{\mathrm{v}}(s)\right] = \mathrm{i}\hbar\beta\mathtt{I}$ if $r > s$, otherwise $\left[\mathrm{u}_p(r), \dot{\mathrm{v}}(s)\right] = 0$, as it was shown in.[11,38]

This causality follows from the *Belavkin's nondemolition principle* requiring the past output operators $\mathtt{y}(s)$, $s \le t$ to be compatible (i.e. commuting) with the present and any future quantum trajectory operators $\mathtt{x}(t)$:

$$[\mathtt{x}(t), \mathtt{y}(s)] = 0 \quad \forall s, t \ge s \tag{5}$$

It easy to show that this condition is satisfied for both canonical operators $\hat{\mathrm{q}}, \hat{\mathrm{p}}$ in place of $\mathtt{x}$ and the output process $\mathtt{y} = \mathtt{y}_e$ evolving the measurement error process $\mathtt{v}_e = \mathtt{I}^{\mathrm{o}} \otimes \mathrm{v}_e$ in the observation channel with $\dot{\mathrm{v}}_e(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{v}_e(t)$. Indeed, representing the classical Wiener process v_e, realizing the commutative family $(\mathrm{v}_e(r) : r \in (0,t])$ after interaction with the particle by the family of operators $\mathtt{y}_e(t) = \mathrm{V}_t^{\dagger}\mathrm{v}_e(t)\,\mathrm{V}_t$ commuting also with $\hat{\mathrm{s}}_{\mathrm{o}}(t)$, one can obtain the linear estimation channel equation

$$\mathrm{d}\mathtt{y}_e = \beta\hat{\mathrm{s}}_q(t)\,\mathrm{d}t + \mathtt{I}^{\mathrm{o}} \otimes \mathrm{dv}_e \tag{6}$$

for $\mathtt{y}_e(t)$ as the particular case of (104). Here the input process appears as measurement error noise with commutative independent increments dv_e, representing the standard Wiener process such that $(\mathrm{dv}_e)^2 = \mathrm{d}t$, but non-commuting with the perturbative force u_p, since, as it follows from (107), they have imaginary quantum Itô product

$$\mathrm{dv}_e\mathrm{du}_p = \frac{\beta\hbar}{2\mathrm{i}}\mathrm{d}t\mathrm{I},\ \mathrm{dv}_e\mathrm{dv}_q = -\varepsilon\mathrm{d}t\mathrm{I}. \tag{7}$$

Therefore, the measurement error noise v_e and perturbation noise u_p satisfy the *error-perturbation CCR* (105):

$$[\mathrm{u}_p(r), \mathrm{v}_e(s)] = (r \wedge s)\,\mathrm{i}\hbar\beta\mathrm{I}. \tag{8}$$

This is necessary condition for quantum causality (or Belavkin's *quantum nondemolition condition*), written in the form (5), is also sufficient for statistical predictability of *quantum hidden (in the future) trajectories*

$(\hat{\mathrm{s}}_{\mathrm{o}}(r) : r \geq t)$ with respect to the *classical observed (in the past) trajectories* $(\mathrm{y}_e(r) : r \leq t)$ for each t. From this the Heisenberg *error-perturbation uncertainty principle* follows in the precise form of the Belavkin inequality[6,24]

$$(\mathrm{du}_p)^2 \geq \left(\frac{\beta\hbar}{2}\right)^2 \mathrm{d}t\mathrm{I} \ \text{ if } \ (\mathrm{dv}_e)^2 = \mathrm{d}t\mathrm{I}, \tag{9}$$

in terms of the standardized intensities of perturbation u_p and error v_e defined respectively in (3) and (6).

Thus, we have the two-dimensional case

$$\boldsymbol{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ \boldsymbol{\lambda} = \begin{pmatrix} \frac{\beta}{2} & \frac{\mathrm{i}\varepsilon}{\hbar} \\ \frac{\gamma}{2} & \frac{\mathrm{i}\varphi}{\hbar} \end{pmatrix} \tag{10}$$

of the general quantum linear open system considered in the last Section, with $\boldsymbol{\lambda}$ as the direct sum $\boldsymbol{\lambda}_e \oplus \boldsymbol{\lambda}_f$ of two rows $\boldsymbol{\lambda}_e, \boldsymbol{\lambda}_f$ corresponding to $\boldsymbol{b}_e = (\beta, 0)$, $\boldsymbol{e} = (0, \varepsilon)$, $\boldsymbol{b} = (\gamma, 0)$, $\boldsymbol{e}_f = (0, \varphi)$. From this one can find the matrices (108) and (120) which are turned to be diagonal,

$$\boldsymbol{g} = \begin{pmatrix} \zeta_q & 0 \\ 0 & \zeta_p \end{pmatrix}, \ \ \boldsymbol{h} = \begin{pmatrix} \eta_q & 0 \\ 0 & \eta_p \end{pmatrix},$$

having eigenvalues $\zeta_q = \varphi^2$, $\zeta_p = (\hbar/2)^2 \left(\beta^2 + \gamma^2\right) = \eta_q$, $\eta_p = \varepsilon^2$ if they are taken to satisfy the microduality principle as defined in the last Section.

2. Quantum Stochastics and Langevin Equations

Time-continuous indirect observations of quantum open objects are described by representing a classical measured process $v(t)$ as a *preoutput noise* with respect to a *reference product measure*, $\mathsf{Q} = \mathsf{Q}^t \otimes \mathsf{Q}_t$ for any t, on a measurable space $(\Omega, \mathfrak{A})$ idetified with the space $(\Upsilon, \mathfrak{B})$ of random trajectories $v(t) = v(t, \omega)$. This means that Q is infinitely divisible, $\mathsf{Q}^{t+s} = \mathsf{Q}^t \otimes \mathsf{Q}_t^s$ for any $s > 0$, and we may assume that $v(t)$ are the previsible realizations of a càdlàg process $t \mapsto x^t(\omega)$ with independent increments $\Delta^s x^t = x(t^s) - x(t)$, where $t^s = t + s > t$, and $x(t) = \lim_{r \nearrow t} x^r \equiv x^{t-}$ denotes the left limit as the previsible version $x^{t-} = v(t)$ of x^t. Such measure can be causally induced from a reference product state $\mathrm{R} = \mathrm{R}^t \otimes \mathrm{R}_t$ on an infinitely divisible W*-algebra $\mathcal{A} \supseteq \mathcal{M}$ assuming the Hilbert space divisibility $G^t = G^t \otimes G_t^s$ for the infinite divisibility $\mathrm{R}^t = \mathrm{R}^t \otimes \mathrm{R}_t^s$ at each t. One can take G unitary equivalent to $H = L^2_{\mathfrak{B}}(\Upsilon, \mathsf{Q})$ with $\mathrm{R} = |g\rangle\langle g|$ as one-dimensional projector given by $g(v) \simeq 1$ and represent each real (or

complex) process $x(t)$ by commuting operators $\mathrm{x}(t) \simeq \breve{x}(t)$ of multiplication $\breve{x}(t)\, h = x(t)\, h$ densely defined at least if $\|x(t)\|_{\mathsf{Q}}^2 = \mathbb{E}_{\mathsf{Q}}\left[|x(t)|^2\right] < \infty$ almost for all t.

Thus, each product measure $\mathsf{Q} = \mathsf{Q}|\mathfrak{B}$ can be induced from a *pure* quantum state $\mathrm{P} = |1\rangle\langle 1|$ on the operator algebra $\mathcal{A} = \mathcal{B}(G)$, given by

$$\mathsf{Q}^t(\Delta) = \langle 1 \mid \mathrm{E}(\Delta)\, 1\rangle \equiv \langle \mathrm{P}, e_{\mathrm{x}}(\Delta)\rangle\,, \quad \Delta \in \mathfrak{B}_t^s. \tag{11}$$

Here $\mathrm{E}(\Delta) = e_{\mathrm{x}}(\Delta)$ denotes the joint spectral measure $e_{\mathrm{x}}(\Delta) = 1_{x^{-1}(\Delta)}(\mathrm{x})$ of all commuting selfadjoint operators $\mathrm{x} = \{\mathrm{x}(t)\}$ acting in an infinitely divisible Hilbert product space G as the adapted representation

$$\mathrm{x}(t) = \int_{\Upsilon} \upsilon(t)\, e_{\mathrm{x}}(\mathrm{d}\upsilon) \simeq \jmath_{\mathrm{x}}^t(x)$$

of a classical process $x(t)$. Here $\jmath_{\mathrm{x}}^t(x) = \jmath_{\mathrm{x}}(x)\,|\,G^t$ is defined by the classical-quantum correspondence $\mathrm{x}(t) = \jmath_{\mathrm{x}}(x) \simeq \breve{x}(t)$ of commuting operators $\mathrm{x}_r(t) = \mathrm{x}(t) - \mathrm{x}(r)$ for all $r \le t$ with their realization $x_r(t) = x(t) - x(r)$ by the real-valued random variables $x(t,\omega) = \upsilon(t)$ as spectral values of $\mathrm{x}(t)$.

In this section we also consider quantum open dynamics for a controlled object with indirect observation, extending the orthodox framework of the closed quantum dynamics on a simple noncommutative W*-algebra $\mathbb{A} = \mathcal{B}(H_{\mathrm{o}})$ to a more general framework of *quantum stochastics* (QS) on an arbitrary W*-algebra, icluding the classical commutative case $\mathbb{A} \simeq L^\infty$. (For the definitions and the basic facts related to W*-algebras see the Appendix A.) It is described by an *interaction dynamics* $\{\mathrm{U}_t\}$ usually defined by unitary Hilbert space operators U_t, called *scatterings*, having the inverse $\mathrm{U}(t) = \mathrm{U}_t^\dagger$. The operators $\mathrm{U}(t)$ are assumed to be adapted on $H_t = G^{t]} \otimes F_t$, embedding at any time $t > 0$ the Hilbert product spaces $G^{t]} = G^t \otimes H_{\mathrm{o}}$ of the quantum object H_{o} and increasing measurement history Hilbert spaces $G^t \simeq L^2$ into an initial Hilbert space $F_{[0} = H_{\mathrm{o}} \otimes F_0$ of the object with quantum noise Hilbert space $F_0 = F_0^t \otimes F_t$ usually represented as the Fock space $\Gamma(\mathcal{E}_0)$ over $\mathcal{E}_0 = L^2\left(\mathbb{R}_0^+ \to \mathfrak{k}\right)$. If $\mathcal{B}_t \subseteq \mathcal{B}(H_t)$ are proper W*-subalgebras $\mathcal{B}_t = \mathcal{A}^t \bar{\otimes} \mathcal{N}_{[t}$ with increasing history W*-algebras $\mathcal{A}^t \subseteq \mathcal{B}(G^t)$, the reversible Schrödinger interaction dynamics $\mathrm{U}_t : F_{[0} \to H_t$ may induce an irreversible Heisenberg dynamics $\mathcal{B}_t \to \mathcal{B}_0$ even if the object plus noise W*-algebra $\mathcal{N}_{[t} = \mathbb{A}^{\mathrm{o}} \bar{\otimes} \mathcal{N}_t$ is simple, $\mathcal{N}_{[t} = \mathcal{B}\left(F_{[t}\right)$. Thus, the dynamics of even indirect, nondemolition measurements,[28] induced by the reversible scatterings on semisimple von Neumann algebras $\mathcal{B}_t = \mathcal{M}^t \bar{\otimes} \mathcal{N}_{[t}$ with increasing Abelian W*-algebras $\mathcal{M}^t \subset \mathcal{B}(G^t)$ and decreasing noncommutative $\mathcal{N}_{[t}$, becomes irreversible, since $\mathrm{U}_t^\dagger \mathcal{B}_t \mathrm{U}_t \subset \mathcal{B}_0 = \mathcal{N}_{[0}$ for all $r < t$

if $\mathcal{N}_{[t} = \mathcal{B}\left(F_{[t}\right)$. We will not require the unitarity but only *semiunitarity* $\mathrm{U}_t\mathrm{U}_t^\dagger = \mathrm{I}_t$, i.e. isometricity of $\mathrm{U}_t^\dagger = \mathrm{U}\left(t\right)$, allowing for strict embeddings $G^t \subset F_0^t$. By such relaxing the unitarity $\mathrm{U}_t^\dagger = \mathrm{U}_t^{-1}$, applicable only for the reversible dynamics of stable particles without observation, we will be able to describe also the demolition indirect measurement processes in unstable quantum systems by semiunitary (*coisometric*) dynamical scatterings U_t.

2.1. *Classical stochastic calculus in Fock space*

An appropriate candidate for the infinitely divisible Hilbert space suitable to accommodate a classical independent increment vector-process $v_\bullet = (v_1, \ldots, v_d)$ on any interval I_0^t of $\mathbb{R}_0^+ = \{r > 0\}$ is the Guichardet-Fock space $F_0 = \Gamma\left(\mathcal{E}_0\right)$ over the Hilbert space $\mathcal{E}_0$ of L^2-functions $\xi_0 : \mathbb{R}_0^+ \mapsto \mathfrak{k}$. (The definitions and notation of Fock space, taken from,[25] are summarized in the Appendix B.) In this case G is adaptively vacuum-embedded into F_0 by considering only the square-integrable vector-functions $\xi_0^\bullet \in \mathcal{E}_0$ as columns $\xi_0^\bullet = \left(\xi_0^j\right)^{j\in J_\bullet}$ with values in a subspace $\mathfrak{k}^\bullet = \mathbb{C}^d$ of $\mathfrak{k} \equiv \mathfrak{k}^\circ$ given in an orthonormal basis indexed by $J \supseteq \{1, \ldots, d\} \equiv J_d$.

The Fock space $F_0 = \Gamma\left(\mathcal{E}_0\right)$ is an exponential of the Hilbert space $\mathcal{E}_0 = L^2\left(\mathbb{R}_0^+ \to \mathfrak{k}\right)$ in the sense that it is generated by the *exponential vectors* $\left\{\xi_0^\otimes | \xi_0 \in \mathcal{E}_0\right\}$ such that

$$\left\|\xi_0^\otimes\right\|_{F_0}^2 = \exp\left[\left\|\xi_0\right\|_{\mathcal{E}}^2\right], \text{ where } \left\|\xi_0\right\|_{\mathcal{E}}^2 = \int_0^t \left\|\xi\left(r\right)\right\|_{\mathfrak{k}}^2 \mathrm{d}r.$$

The exponential vectors are infinitely divisible due to $\xi_0^{t+s} \simeq \xi_0^t \oplus \xi_t^s$, with $\left(\xi_0^t\right)^\otimes \in F_0^t$ defining the *coherent state vectors*

$$f_{\xi_0^t} = \exp\left\{-\frac{1}{2}\left\|\xi_0^t\right\|_{\mathcal{E}}^2\right\}\xi_0^{t\otimes}, \ \left\|\xi_0^t\right\|_{\mathcal{E}}^2 = \int_0^t \left\|\xi\left(r\right)\right\|_{\mathfrak{k}}^2 \mathrm{d}r. \tag{12}$$

by $\xi_0^t = \xi | I_0^t \in \mathcal{E}_0^t$. The coherent states $\mathrm{P}_\xi^t = |f_\xi^t\rangle\langle f_\xi^t|$ induce the *product states* on increasing subalgebras $\mathcal{A}^t \subseteq \mathcal{B}\left(G^t\right)$ factor-embedded into $\mathcal{B}\left(F_0^t\right)$ such that it satisfies the divisibility condition $\mathcal{A}^{t+s} \simeq \mathcal{A}^t \bar{\otimes} \mathcal{A}_t^s$ for all $t, s > 0$ with $\mathcal{A}_t^s$ represented in $\mathcal{B}\left(F_t^s\right)$. In particular, they can induce any infinitely divisible classical state with compatible probability measures $\mathsf{Q}^t = \epsilon_\xi^t \circ e_{\mathrm{x}}$ on $\mathfrak{B}^t$ defining classical expectations $\mathbb{E}_{\mathsf{Q}}^t = \epsilon_\xi^t \circ \jmath_{\mathrm{x}}$ on $\mathcal{M}^t = L_{\mathfrak{B}}^\infty\left(\Upsilon^t, \mathsf{Q}^t\right)$ as restrictions of the coherent expectations $\epsilon_\xi^t\left[\mathrm{B}\right] = \left\langle \mathrm{P}_\xi^t, \mathrm{B}\right\rangle$ to the local Abelian algebras $\mathcal{M}^t = \jmath_{\mathrm{x}}\left(\mathcal{M}^t\right) \equiv \mathbb{A}_{\mathrm{x}}^t$.

There are two basic classical stationary processes with additive independent increments and zero expectation which can be induced from the

coherent state P_{ξ}^{t} in Fock space with complex $\xi^{j}=\mathrm{e}^{\mathrm{i}\theta_j}\left|\xi_j\right|\equiv\xi_j^{*}$ indexed by a subset $J_\bullet$ of an orthonormal basis set J of $\mathfrak{k}$: The standard Wiener vector-valued process $w_\bullet=(w_j)$ and the scaled Poisson compound process $m_\bullet=(m_j)$ of $m_j=\left(n_j-\mathbb{E}_{\mathsf{Q}}\left[n_j\right]\right)/\left|\xi_i\right|$ compensated by the expectations

$$\mathbb{E}_{\mathsf{Q}}\left[n_j\left(t\right)\right]=\int_0^t\left|\xi_j\left(r\right)\right|^2\mathrm{d}r=\epsilon_{\xi}^{t}\left[\mathrm{n}_j\left(t\right)\right].$$

Here n_j are the usual (noncompensated) Poisson processes of the intensity $\left|\xi_j\right|^2$ which can be realized as independent number processes n_j in their eigenrepresentation with respect to a coherent vector-state in F_0^t up to any finite time t. The forward differential increments $\mathrm{d}v_j^t$ of these $v_j\left(t\right)$ satisfy quite distinct Itô multiplication tables

$$\mathrm{d}w_i\mathrm{d}w_k=\delta_{ik}\mathrm{d}t,\ \mathrm{d}m_i\mathrm{d}m_k=\delta_{ik}\left(\mathrm{d}t+\frac{1}{\left|\xi_i\right|}\mathrm{d}m_i\right),\tag{13}$$

with the Wiener nilpotent table as the central limit of the Poisson table at $\left|\xi_i\right|\longrightarrow\infty$. These tables can be easily obtained from an operator calculus in Fock subspace G over $\mathcal{E}_\bullet\simeq\mathbb{C}^d\otimes L^2\left(\mathbb{R}_0^+\right)$ by representing $\mathrm{d}w_j$ and $\mathrm{d}m_j$ as

$$\mathrm{d}\mathrm{w}_j=2\,\mathsf{Re}\left[\mathrm{e}^{\mathrm{i}\theta_j}\mathrm{d}\mathrm{A}_j^{+}-\left|\xi_j\right|\mathrm{d}\mathrm{A}_{-}^{+}\right],\ \ \mathrm{d}\mathrm{m}_j=\frac{1}{\left|\xi_j\right|}\mathrm{d}\mathrm{A}_j^{j}-\left|\xi_j\right|\mathrm{d}\mathrm{A}_{-}^{+}.\tag{14}$$

Here $2\,\mathsf{Re}\left[\mathrm{A}_j^{+}\right]=\mathrm{A}_j^{+}+\mathrm{A}_{-}^{j}$ is doubled Hermitian part of the creation (or annihilation A_{-}^{j}) operators $\mathrm{A}_j^{+}=\mathrm{A}_{-}^{j\dagger}$, $\mathrm{A}_{-}^{+}\left(t\right)$ denotes the preservation process $\mathrm{A}_{-}^{+}\left(t\right)=t\mathrm{I}$ and $\mathrm{A}_j^{j}\left(t\right)=\mathrm{n}_j\left(t\right)$ denotes the number processes in its eigenrepresentation by the diagonal elements of the exchange matrix-process $\mathrm{A}_\bullet^\bullet=\left(\mathrm{A}_i^k\right)_{i\in J_\bullet}^{k\in J_\bullet}$. The infinitesimal increments

$$\mathrm{d}\mathrm{A}_\iota^\kappa\left(t\right)=\mathrm{A}_\iota^\kappa\left(t+\mathrm{d}t\right)-\mathrm{A}_\iota^\kappa\left(t\right)$$

of these canonical quantum processes satisfy the Hudson-Parthasarathy (HP) table[51]

$$\begin{aligned}\mathrm{d}\mathrm{A}_{-}^\bullet\left(t\right)\mathrm{d}\mathrm{A}_\bullet^\bullet\left(t\right)&=\mathrm{d}\mathrm{A}_{-}^\bullet\left(t\right),\quad \mathrm{d}\mathrm{A}_{-}^\bullet\left(t\right)\mathrm{d}\mathrm{A}_\bullet^{+}\left(t\right)=\mathrm{d}t\mathrm{I},\\ \mathrm{d}\mathrm{A}_\bullet^\bullet\left(t\right)\mathrm{d}\mathrm{A}_\bullet^\bullet\left(t\right)&=\mathrm{d}\mathrm{A}_\bullet^\bullet\left(t\right),\quad \mathrm{d}\mathrm{A}_\bullet^\bullet\left(t\right)\mathrm{d}\mathrm{A}_\bullet^{+}\left(t\right)=\mathrm{d}\mathrm{A}_\bullet^{+}\left(t\right)\end{aligned}$$

with all other increment multiplications vanishing, which can be written all in one pseudo-Poisson multiplication table

$$\mathrm{d}\mathrm{A}_\iota^\nu\left(t\right)\mathrm{d}\mathrm{A}_\mu^\kappa\left(t\right)=\delta_\mu^\nu\mathrm{d}\mathrm{A}_\iota^\kappa\left(t\right)\quad\forall\iota,\mu\in\left\{-,J\right\},\nu,\kappa\in\left\{J,+\right\}\tag{15}$$

as it was found by Belavkin in.[10,25] Note that the matrix $\left\{\mathrm{A}_\iota^\kappa\left(t\right)\right\}_{\iota=-,\bullet,+}^{\kappa=-,\bullet,+}$, with $\mathrm{A}_i^{-}=0=\mathrm{A}_{+}^k$ for all $i,k\in J_\bullet$ and with the creations arranged

in the row $\mathrm{A}_\bullet^+ = \left(\mathrm{A}_i^+\right)_{i\in J}$ and annihilations arranged in the column $\mathrm{A}_-^\bullet(t) = \left(\mathrm{A}_-^j(t)\right)^{j\in J_\bullet}$, is pseudo-Hermitian $\mathbf{A}^\star = \mathbf{A}$ under the special involution[10,25] $\left(\mathrm{A}_{-\iota}^\kappa\right)^\star = \left(\mathrm{A}_{-\kappa}^{\iota\dagger}\right)$ as $\mathrm{A}_-^+ = \mathrm{A}_-^{+\dagger}$ $\mathrm{A}_\bullet^\bullet = \mathrm{A}_\bullet^{\bullet\dagger}$. Thus, the Belavkin involution is pseudo-Hermitian conjugation

$$\mathbf{A}^\star = \left\{\begin{matrix} 0 & 0 & 1 \\ 0 & \mathrm{I} & 0 \\ 1 & 0 & 0 \end{matrix}\right\} \left\{\begin{matrix} 0 & \mathrm{A}_-^{\bullet\dagger} & \mathrm{A}_-^{+\dagger} \\ 0 & \mathrm{A}_\bullet^{\bullet\dagger} & \mathrm{A}_\bullet^{+\dagger} \\ 0 & 0 & 0 \end{matrix}\right\} \left\{\begin{matrix} 0 & 0 & 1 \\ 0 & \mathrm{I} & 0 \\ 1 & 0 & 0 \end{matrix}\right\} \equiv \mathbf{J}\mathbf{A}^\dagger\mathbf{J}$$

defined by the Hermitian conjugation $\dagger$ in F_0 and the reflection $-(-,\bullet,+) = (+,\bullet,-)$ leaving the usual indices $\bullet = i, k \in J_\bullet$ unchanged.

The forward differentials dA_μ^ν, giving a linear basis of the vacuum Itô $\star$-algebra, serve as noncommutative integrators for the QS integrals

$$\mathrm{V}_r(t) = \int_r^t \mathbf{K}\cdot\mathrm{d}\mathbf{A} := \int_r^t \mathrm{K}_\nu^\mu(s)\,\mathrm{dA}_\mu^\nu(s) \equiv \mathfrak{i}_r^t(\mathbf{K}) \tag{16}$$

defined by four integrands $\mathbf{K} = \mathrm{K}_\nu^\mu$ with $\mathrm{K}_\nu^\mu\mathrm{dA}_\mu^\nu = \sum_{\mu,\nu}\mathrm{K}_\nu^\mu\mathrm{dA}_\mu^\nu$ as operator-valued functions integrable in the QS sense defined in[25] where the following proposition was proved.

Proposition 2.1. *The QS integrals (16) satisfy the $\star$-property in the form $\mathfrak{i}_r^t(\mathbf{K})^\dagger = \mathfrak{i}_r^t(\mathbf{K}^\star)$ under the pseudo-Hermitian involution $\left(\mathrm{K}_{-\nu}^\mu\right)^\star = \left(\mathrm{K}_{-\mu}^{\nu\dagger}\right)$ and the noncommutative Itô product rule*

$$\mathrm{d}(\mathrm{V}\mathrm{V}^\dagger) = \mathrm{d}(\mathrm{V})\,\mathrm{V}^\dagger + \mathrm{V}\mathrm{d}\left(\mathrm{V}^\dagger\right) + \mathrm{d}(\mathrm{V})\,\mathrm{d}(\mathrm{V}^\dagger), \tag{17}$$

where $\mathrm{dV} = \mathrm{K}_\nu^\mu\mathrm{dA}_\mu^\nu$ (the usual summation convention is assumed) and the QS Itô correction is given for the adapted *QS integrands $\mathrm{K}_\nu^\mu(t)$ in terms of the matrix product (138) as*

$$\mathrm{d}\mathfrak{i}_r^t(\mathbf{K})\,\mathrm{d}\mathfrak{i}_r^t(\mathbf{K}^\star) = \mathrm{d}\mathfrak{i}_r^t(\mathbf{K}\mathbf{K}^\star). \tag{18}$$

As an application of the quantum Itô algebra matrix representation (138) one can easily check that $\mathbf{d}(c)\,\mathbf{d}(c)^\star = \mathbf{d}(c\varepsilon c^*) + c^\dagger c\mathbf{d}_-^+$ for the quadruples

$$\mathbf{d}(c^*) = \begin{bmatrix} d_\bullet^- & d_+^- \\ d_\bullet^\bullet & d_+^\bullet \end{bmatrix}(c^*) = \mathbf{d}(c)^\star,\ \mathbf{d}_-^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

given by $\mathbf{d}(c) = c^j\mathbf{d}_j$ with the scalar-valued

$$d_k^i(c) = \varepsilon\delta_k^i c^i, \quad \mathrm{e}^{-\mathrm{i}\theta_k} d_+^k(c) = (1-\lambda)\,c^k = \mathrm{e}^{\mathrm{i}\theta_k} d_k^-(c) \tag{19}$$

and $d_+^-(c) = \sum_j \varepsilon^{-1}\lambda(\lambda-2)c^j$ as linear functions of the $\mathbb{C}$-valued test functions $c^j(t)$ arranged into the column $c = \left(c^j\right)$. This gives for the forward increments $\mathrm{dx}^\varepsilon(c) = c^j(t)\,\mathrm{dx}_j^\varepsilon$ of the QS integrals $\mathrm{x}_j^\varepsilon(c) = \int_0^t \mathbf{d}(c)\cdot\mathrm{d}\mathbf{A}$ that the product $\mathrm{dx}_t^\varepsilon(c)\,\mathrm{dx}_t^\varepsilon(c^*)$ is defined by the standard multiplication table

$$\mathrm{dx}_i^\varepsilon \mathrm{dx}_k^\varepsilon = \delta_{ik}\left(\mathrm{d}t + \varepsilon_k \mathrm{dx}_k^\varepsilon\right) = \mathrm{dx}_k^\varepsilon \mathrm{dx}_i^\varepsilon \tag{20}$$

representing the commutative Itô tables (13) for $\lambda_j = \varepsilon\left|\xi_j\right| = 0$ and $\lambda_j = \varepsilon\left|\xi_j\right| = 1$. Moreover, one can also easily find that $\mathrm{dx}_j^\varepsilon \mathrm{dx}_j^{\varepsilon'} \neq \mathrm{dx}_j^{\varepsilon'} \mathrm{dx}_j^\varepsilon$ if $\varepsilon \neq \varepsilon'$, in particular, for $\lambda_i = 0$ and $\lambda_k = 1$ it follows straightforward from

$$\mathrm{dm}_k \mathrm{dw}_i = \delta_{ki}\frac{1}{\xi_i}\mathrm{dA}_i^+ \ , \quad \mathrm{dw}_i \mathrm{dm}_k = \delta_{ik}\frac{1}{\xi_k^*}\mathrm{dA}_-^k .$$

Note that this cannot be realized in the classical category of the commutative processes unless the type ε is assumed to be a function of j (and possibly of time t), $\varepsilon_j(t) = \varepsilon_j'(t)$, such that the Wiener and the Poisson processes are indexed by disjoint subsets respectively J_w and J_m as it is always in the classical case: $\mathrm{d}w_i^t \mathrm{d}m_k^t = 0 = \mathrm{d}m_k^t \mathrm{d}w_i^t$. The joint probability measure for the stationary processes can also be induced from the coherent states as $\mathsf{Q}^t(\Delta) = \left\langle \mathrm{P}_\xi^t, e_\mathrm{x}(\Delta)\right\rangle$ on each $\mathfrak{B}^t$ with constant nonzero amplitudes $\left|\xi_j(r)\right| = \varepsilon_j^{-1}\lambda_j$ for all $r \in (0,t]$. Note that the described representation is eigen representation for the processes m_j but not for w_j, and that the amplitudes may only be locally integrable in the sense $\int_0^t \left|\xi_j(r)\right|^2 \mathrm{d}r < \infty$ such that there might be no limiting coherent state P_ξ which would globally induce the probability distribution Q.

As we prove below, any such Q can be also globally induced from the vacuum state $\mathrm{Q} = \mathrm{P}_0$,

$$\mathsf{Q}(\Delta) := \left\langle \delta^\emptyset \mid e_\mathrm{v}(\Delta)\,\delta^\emptyset \right\rangle \equiv \left\langle \mathrm{Q}, e_\mathrm{v}(\Delta)\right\rangle, \tag{21}$$

given by the *vacuum vector* $\delta_\emptyset = f_0$ of F_0 as the only stationary coherent state vector corresponding to the only translationary invariant vector-function $\xi(t) = 0$ square-integrable on $\mathbb{R}_0^+$. This can be done even in the nonstationary case by replacing each x_j^ε with locally unitary equivalent v_j^ε given by

$$\mathrm{dv}_j^\varepsilon = \varepsilon_j \mathrm{dA}_j^j + \mathrm{dA}_-^j + \mathrm{dA}_j^+ \equiv \mathbf{b}_j^\varepsilon \cdot \mathrm{d}\mathbf{A}, \tag{22}$$

where the canonical quadruple $\mathbf{b}_j^\varepsilon$ can be obtained form the quadruples $\mathbf{d}_j^\varepsilon$ for $\mathrm{dx}_j^\varepsilon = \mathbf{d}_j^\varepsilon(t)\cdot\mathrm{d}\mathbf{A}$ by a linear transformation

$$\mathbf{d}_j^\varepsilon = \begin{bmatrix} d_j^{\bullet\dagger} & d_j \\ \delta_j^\bullet \varepsilon_j \delta_j^{\bullet\dagger} & d_j^\bullet \end{bmatrix} \mapsto \mathbf{b}_j^\varepsilon = \begin{bmatrix} \delta_j^{\bullet\dagger} & 0 \\ \delta_j^\bullet \varepsilon_j \delta_j^{\bullet\dagger} & \delta_j^\bullet \end{bmatrix}. \tag{23}$$

with $d_j^\bullet = \delta_j^\bullet (1-\lambda_j)\,\mathrm{e}^{\mathrm{i}\theta_j}$ and $d_j = (\lambda_j - 2)\,|\xi_j|$ corresponding to (19). This equivalence is defined at any $t \geq r + \mathrm{d}r$ by the unitary Weyl operator $\mathrm{U}_\xi^t = \mathrm{W}_t(\mathbf{a}_\xi)$ as the multiplicative QS integral $\mathrm{W}_t(\mathbf{a}) = \exp\{\mathrm{i}_0^t(\ln \mathbf{g})\}$ resolving the the exponential QS differential equation

$$\mathrm{dW}_t(\mathbf{a}) = \mathrm{W}_t(\mathbf{a})\,\mathbf{a}(t)\cdot \mathrm{d}\mathbf{A}(t),\ \ \mathrm{W}_0(\mathbf{a}) = \mathrm{I} \tag{24}$$

It is given in terms of the Weyl generator $\mathbf{a} = \mathbf{1} - \mathbf{g}$ as the quadruple $[a_\kappa^\iota]_{\kappa=+,\bullet}^{\iota=-,\bullet}$ determined by the Weyl germ $\mathring{\mathrm{W}}_{t^+} = \mathbf{g}(t)\,\mathrm{W}_t$ given by the pseudo-unitary matrix $\mathbf{g} = \{g_\kappa^\iota\}_{\kappa=-,\bullet,+}^{\iota=-,\bullet,+}$, $\mathbf{g}^{-1} = \mathbf{g}^\star$, with the complex entries $g_\kappa^\iota(t)$ defined by the column $\xi^\bullet(t) = \xi_\bullet^\dagger(t)$ as $g_+^i = \xi^i$, $g_k^- = -|\xi_k|$, $g_k^i = \mathrm{e}^{\mathrm{i}\theta_k}\delta_k^i$, $g_+^- = -\frac{1}{2}\|\xi\|_{\mathfrak{k}}^2$ in

$$\mathbf{g}_\xi = \begin{bmatrix} -|\xi_\bullet| & -\frac{1}{2}\|\xi_\bullet\|_{\mathfrak{k}}^2 \\ \mathrm{e}^{\mathrm{i}\theta_\bullet} & \xi^\bullet \end{bmatrix} \equiv [g_\kappa^\iota]_{\kappa=\bullet,+}^{\iota=-,\bullet} \tag{25}$$

and $g_-^- = 1 = g_+^+$, $g_k^+ = 0 = g_-^i$ for all $i,k \in J_\bullet$.

Theorem 2.1. *Let $\mathfrak{b}$ be an Itô $\star$-algebra of the quadruples $\mathbf{b} = (b_\kappa^\iota)_{\kappa=\bullet,+}^{\iota=-,\bullet}$ with pointwise matrix multiplication of locally bounded entries $b_\kappa^\iota(\cdot) \in L_{\mathrm{loc}}^\infty(\mathbb{R}_+)$, and $\mathcal{B}^t$ be the W*-algebra generated by the Weyl exponentials $\mathrm{W}_t(\mathbf{b})$ satisfying to (24) with $b_+^-(r) = 0$ and $b_k^i(t) = (\exp\{\mathbf{l}(t)\})_k^i - \delta_k^i$ given by the matrix exponent of $\mathbf{l}(t) = [l_\kappa^\iota(t)] \in \mathfrak{b}(t)$. Then*

$$\mathrm{W}_t(\mathbf{b}) = \mathrm{e}^{\int_0^t b_+^i(r)\mathrm{dA}_i^+(r)}\mathrm{e}^{\int_0^t l_k^i(r)\mathrm{dA}_i^k(r)}\mathrm{e}^{\int_0^t b_k^-(r)\mathrm{dA}_-^k(r)} \equiv \exp\left\{\mathrm{i}_0^t(\mathbf{l})\right\} \tag{26}$$

such that $\epsilon_0[\mathrm{W}_t(\mathbf{b})] := \langle f_0|\mathrm{W}_t(\mathbf{b})\,f_0\rangle = 1 = \langle f_\xi|\mathrm{W}_t(\mathbf{d})\,f_\xi\rangle \equiv \epsilon_\xi^t[\mathrm{W}_t(\mathbf{d})\,f]$, where $\mathrm{W}_t(\mathbf{d}) = \mathrm{U}_\xi^t\mathrm{W}_t(\mathbf{b})\,\mathrm{U}_\xi^{t\dagger}$ is given as $\exp\{\mathrm{i}_0^t(\ln[\mathbf{1}+\mathbf{d}])\}$ by $\mathbf{d} = \mathbf{g}_\xi\mathbf{b}\mathbf{g}_\xi^\star$ with

$$\mathrm{e}^{-\mathrm{i}\theta_j}c_+^j = b_+^j - c_k^j\xi^k,\ c_j^-\mathrm{e}^{\mathrm{i}\theta_j} = b_j^- - \xi_i c_j^i,\quad \mathrm{e}^{-\mathrm{i}\theta_j}c_k^j\mathrm{e}^{\mathrm{i}\theta_k} = b_k^j \tag{27}$$

and $c_+^- + \xi_j c_+^j + c_k^-\xi^k + \xi_j c_k^j\xi^k = b_+^-$. (The summation rule is applied with $\xi^j = \xi_j^*$.) This defines the statistical equivalence

$$\epsilon_0^t\left[\mathrm{W}_t\left(\mathbf{g}_\xi^\star\mathbf{d}\mathbf{g}_\xi\right)\right] = \mathrm{e}^{\int_0^t b_+^-(r)\mathrm{d}r} = \epsilon_\xi^t[\mathrm{W}_t(\mathbf{d})] \tag{28}$$

of the locally coherent-induced state on the W-algebras $\mathbb{A}_{\mathrm{x}}^t$ generated by $\mathrm{x}(t) = \mathrm{i}_0^t(\mathbf{d})$ with $\mathbf{d} \in \mathfrak{d} := \mathbf{g}_\xi\mathfrak{b}\mathbf{g}_\xi^\star$ and the vacuum state on the local W*-algebras $\mathbb{C}_{\mathrm{v}}^t$ generated by $\mathrm{v}(t) = \mathrm{i}_0^t(\mathbf{b})$ with $\mathbf{b} = \mathbf{g}_\xi^\star\mathbf{d}\mathbf{g}_\xi$.*

Proof. The existence and uniqueness of the solution $\mathrm{U}_\xi^t = \mathrm{W}_t(\mathbf{a}_\xi)$ to the equation (24) follows[21,58] from the local boundedness of $|a_k^i(r)| = \delta_k^i$, local L^2-integrability of the complex functions $|a_i^-(r)| = |\xi_i(r)| = a_+^i(r)$ and

local L^1-integrability of $a_+^-(r) = \frac{1}{2}\|\xi(r)\|_{\mathfrak{k}}^2$. The unitarity $\mathrm{U}_\xi^\dagger = \mathrm{U}_\xi^{-1}$ follows[21] from the directly verifiable pseudo-unitarity $\mathbf{g}_\xi^\star \mathbf{g}_\xi = \mathbf{1} = \mathbf{g}_\xi \mathbf{g}_\xi^\star$ of the triangular matrix-function $\mathbf{g}_\xi(t) = \mathbf{1} + \mathbf{a}_\xi(t)$, can be checked by applying the quantum Itô rule (17) to $\mathrm{W}^\dagger\mathrm{W} = \mathrm{I}$ with (18) giving for $\mathbf{K} = \mathrm{W} \otimes \mathbf{a}_\xi$, $\mathbf{K}^\star = \mathrm{W}^\dagger \otimes \mathbf{a}_\xi^\star$

$$\begin{aligned}\mathrm{d}\left(\mathrm{W}^\dagger\mathrm{W}\right) &= \left[\mathrm{W}^\dagger\mathrm{W} \otimes (\mathbf{a}^\star + \mathbf{a} + \mathbf{a}^\star\mathbf{a})\right] \cdot \mathrm{d}\mathbf{A} \\ &= \left[\mathrm{W}^\dagger\mathrm{W} \otimes (\mathbf{g}^\star\mathbf{g} - \mathbf{1}) \cdot \mathrm{d}\mathbf{A}\right] = 0.\end{aligned}$$

This all can also be verified by the direct substitution of the exponential formula (26) with $b_\kappa^\iota = a_\kappa^\iota - \delta_-^\iota \delta_\kappa^+ a_+^-$ into $\mathrm{W}_t(\mathbf{a}) = \mathrm{W}_t(\mathbf{b}) \exp\left\{\int_0^t a_+^-(r)\,\mathrm{d}r\right\}$, which follows from the functional QS Itô formula.[21,27] Applying this normal ordered exponential to the vacuum vector $\delta^\emptyset = f_0$ we have

$$\mathrm{U}_\xi^t f_0 = \exp\left\{\int_0^t \left[a_+^-(r)\,\mathrm{d}r + a_+^i(r)\,\mathrm{dA}_i^+\right]\right\} f_0 = f_\xi,$$

as $a_+^\bullet(r) = \xi^\bullet(r) \equiv \xi(r)$, $a_+^-(r) = -\frac{1}{2}\|\xi(r)\|_{\mathfrak{k}}^2$, which can also be obtained from $\mathrm{dA}_-^k f_0 = 0$ and $\mathrm{dA}_i^k(t) f_0 = 0$ in (24):

$$\mathrm{dU}_\xi^t \delta^\emptyset = \left(\xi^i(t)\,\mathrm{dA}_i^+(t) - \frac{1}{2}\|\xi(t)\|_{\mathfrak{k}}^2\right)\mathrm{d}t\mathrm{U}_\xi^t\delta^\emptyset.$$

Due to the commutativity of the disjoint increments dA_μ^ν the solution $\mathrm{U}_\xi^s(t) = \mathrm{W}_{t^s}(t)$ of (24) with $\mathrm{W}_t(t) = \mathrm{I}$ satisfies at any $t^s = t + s > r$ the hemigroup property $\mathrm{U}_\xi^s(t)\,\mathrm{U}_\xi^t = \mathrm{U}_\xi^{t+s} \simeq \mathrm{U}_\xi^t \otimes \mathrm{U}_{t,\xi}^s$. This proves the local unitary equivalence $\mathrm{W}_t(\mathbf{b}) \simeq \mathrm{U}_\xi^{t+s}\mathrm{W}_t(\mathbf{b})\,\mathrm{U}_\xi^{t+s\dagger} = \mathrm{W}_t(\mathbf{d})$ for any $s > 0$, where $\mathrm{W}_t(\mathbf{d}) = \mathrm{U}_\xi^t\mathrm{W}_t(\mathbf{b})\,\mathrm{U}_\xi^{t\dagger} = \mathrm{W}_t(\mathbf{gbg}^\star)$ as it follows from the Itô rule

$$\mathrm{dW}_t(\mathbf{d}) = \mathrm{W}_t(\mathbf{d})\,(\mathbf{aba}^\star + \mathbf{ab} + \mathbf{ba}^\star + \mathbf{b}) \cdot \mathrm{d}\mathbf{A}.$$

Thus, $\mathrm{W}_t(\mathbf{b})\,f_0 = \mathrm{U}_\xi^\dagger \mathrm{W}_t(\mathbf{d})\,f_\xi$, and therefore we have the equivalence (28) as $\langle f_0|\mathrm{dW}_t(\mathbf{b})\,f_0\rangle = b_+^-(t)\,\mathrm{d}t$, or directly from the exponential formula for $\mathrm{W}(\mathbf{b})$ with arbitrary $b_+^- \in L_{\mathrm{loc}}^\infty(\mathbb{R}_+)$. □

Corollary 2.1. *The equivalence (28) imply the correspondence*

$$\epsilon_0[e_{\mathrm{v}_\bullet}(\Delta)] = \mathsf{Q}(\Delta) = \epsilon_\xi^t[e_{\mathrm{x}_\bullet}(\Delta)] \quad \forall \Delta \in \mathfrak{B}^t, t > 0 \tag{29}$$

if the vector process $\mathrm{v}_\bullet = (\mathrm{v}_j)$ *is defined by* $\mathrm{dv}_j = \mathbf{b}_j \cdot \mathrm{d}\mathbf{A}$ *with* $\mathbf{b}_j = \mathbf{g}_\xi^\star \mathbf{d}_j \mathbf{g}_\xi$ *given by commuting pseudo-Hermitian integrands* $\mathbf{d}_j = \mathbf{d}_j^\star$ *defining the increments* $\mathrm{dx}_j = \mathbf{d}_j \cdot \mathrm{d}\mathbf{A}$ *of* $\mathrm{x}_\bullet = (\mathrm{x}_j)$. *In particular, the coherent-induced probability measure* $\mathsf{Q}_\varepsilon^t = \epsilon_\xi^t \circ e_{\mathbf{x}^\varepsilon}|\mathfrak{B}^t$, *given by the affine mixtures*

$$\mathrm{dx}_j^\varepsilon = \lambda_j^\varepsilon(t)\,\mathrm{dm}_j + \left(1 - \lambda_j^\varepsilon(t)\right)\mathrm{dw}_j \tag{30}$$

of (14) with $\lambda_j^\varepsilon = \varepsilon_j \left|\xi_j\right|$*, coincides on each* $\mathfrak{B}^t$ *with the vacuum-induced probability measure* $\mathsf{Q}_\varepsilon = \epsilon_0 \circ e_{\mathrm{v}_\bullet^\varepsilon}$ *for (22) representing the classical independent increments of* $v_\bullet^\varepsilon = \left(v_j^\varepsilon\right)$ *defined by zero expectations and the Itô table*

$$\mathrm{d}v_i^\varepsilon \mathrm{d}v_k^\varepsilon = \delta_{ik}\left(\mathrm{d}t + \varepsilon_k \mathrm{d}v_k^\varepsilon\right) = \mathrm{d}v_k^\varepsilon \mathrm{d}v_i^\varepsilon. \tag{31}$$

Indeed, the statistical equivalence (29) can also be extended to any classical stochastic vector-process $v_\bullet = (v_j)$ with the same expectation and Itô table as the QS vector process $\mathrm{v}_\bullet = (\mathrm{v}_j)$. In particular, one can easily see that $\epsilon_\xi\left[\mathrm{V}_c^t\right] = 1 = \mathbb{E}_\mathsf{Q}\left[V_c^t\right]$ for the solutions of quantum and classical Itô equations

$$\mathrm{dV}_c^t = \mathrm{V}_c^t c^j\left(t\right)\mathrm{dv}_j^\varepsilon, \;\; \mathrm{d}V_c^t = V_c^t c^j\left(t\right)\mathrm{d}v_j^\varepsilon \tag{32}$$

given respectively by the normal ordered exponentials $\mathrm{V}_c^t = \mathrm{W}_t\left(\mathbf{b}_c\right)$ with $\mathbf{b}_c = \left[b_\kappa^t\right] = c^j \mathbf{b}_j$ and the classical stochastic exponentials

$$V_c^t = \exp\left\{\frac{1}{\varepsilon_j}\int_0^t \left[\ln\left(1+\varepsilon_j c^j\left(r\right)\right)\left(\mathrm{d}v_j^\varepsilon + \frac{1}{\varepsilon_j}\mathrm{d}r\right) - c^j\left(r\right)\mathrm{d}r\right]\right\}$$

of $v_c\left(t\right) = \int_0^t c^i\left(r\right)\mathrm{d}v_i^r$, where $\frac{1}{\varepsilon}\ln\left(1+\varepsilon c\right) = c$ if $\varepsilon = 0$. Since

$$\mathrm{dV}_c^t \mathrm{dV}_c^{t\dagger} = \mathrm{V}_c^t c^j\left(t\right)\left(\varepsilon_j \mathrm{dv}_j^\varepsilon + \mathrm{d}t\right)\bar{c}^j\left(t\right)\mathrm{V}_c^{t\dagger}$$

and similar for $\left|\mathrm{d}V_c^t\right|^2$, the classical and quantum stochastic exponentials obey the same multiplication rule

$$\mathrm{V}_c^t \mathrm{V}_c^{t\dagger} = \mathrm{V}_{c\dot{+}\bar{c}}^t \mathrm{e}^{\int_0^t \|c(r)\|_{\mathfrak{k}}^2 \mathrm{d}r}, \;\; \left|V_c^t\right|^2 = V_{c\dot{+}\bar{c}}^t \mathrm{e}^{\int_0^t \|c(r)\|_{\mathfrak{k}}^2 \mathrm{d}r},$$

where $c^j \dot{+} \bar{c}^j = c^j + \varepsilon_j\left|c^j\left(t\right)\right|^2 + \bar{c}^j$. This proves the statistical equivalence

$$\epsilon_\xi^t\left[\mathrm{V}_c^t \mathrm{V}_c^{t\dagger}\right] = \mathrm{e}^{\int_0^t \|c(r)\|_{\mathfrak{k}}^2 \mathrm{d}r} = \mathbb{E}_\mathsf{Q}\left[\left|V_c^t\right|^2\right]$$

of the classical – quantum correspondence for the processes $v_\bullet^\varepsilon$ and $\mathrm{x}_\bullet^\varepsilon$, and similar for the coherent states on the Weyl operators $\mathrm{X}_c^t = \mathrm{W}_t\left(\mathbf{d}\right)$ given by the tests $\mathbf{d}$ in (19) as Wick exponentials of $\mathrm{x}_c\left(t\right) = \int_0^t c^i\left(r\right)\mathrm{dx}_i^r$.

Note that not only translationary stationary but any infinitely divisible normal state on a W*-product subalgebra $\mathcal{B} \subseteq \mathcal{B}\left(F_0\right)$ can be globally induced from the vacuum state on $\mathcal{B}\left(F_0\right)$ by choosing F_0 in a canonical way[22,23] with in general time-dependent Hilbert space $\mathfrak{k}\left(t\right)$.

2.2. *Quantum stochastic differential equations*

Let $\mathcal{M}_0^{t]} = \mathbb{A}_{\mathrm{x}}^t$ denote the W*-algebra equivalent to $\mathcal{M}_0^t \bar{\otimes} \mathbb{A}^{\mathrm{o}}$ on $G_t^{s]} = G_t^s \otimes H_{\mathrm{o}}$, where $\mathcal{M}_0^t = \mathbb{C}_{\mathrm{x}}^t$ be the Fock space representation of a commutative W*-algebra $\mathbb{A}_{\mathrm{x}}^t \subseteq \mathcal{B}(G^t)$ generated as above by an observed process $\mathrm{x}(r), r \leq t$, and $\mathrm{U}_t(t^s) : G_t^{s]} \to F_{[t}^s$ with $t^s = t + s > t$ be linear isometries from $G_t^{s]}$ onto $F_{[t}^s = H_{\mathrm{o}} \otimes F_t^s$ forming a hemigroup in the sense

$$\mathrm{U}_r(t)\,\mathrm{U}_t(t^s) = \mathrm{U}_r(t^s) \quad \forall r < t < t^s,$$

when $\mathrm{U}_r(t)$ are extended adaptively in both sides to the isometries $H_t \to H_r$ of the Hilbert spaces $H_t = G^t \otimes F_{[t}$. The adjoint operators $\mathrm{U}_t(r) := \mathrm{U}_r(t)^\dagger$ with $r < t$ describe a Schrödinger picture interaction evolution $\psi_t = \mathrm{U}_t(r)\,\psi_r$ with $\mathrm{U}_t(t) = \mathrm{I}_t$. Such adapted evolutions correspond to quantum open Markovian systems and in the stationary case are defined as $\mathrm{U}_t(t^s) = \theta_s(\mathrm{U}_{t_s}(t))$ at $t_s = t - s$ by a one-parameter cocycle $\mathrm{U}(s) = \mathrm{U}_0(s)$ with respect to the right shift semigroup $(\theta_s)_{s>0}$ of W*-endomorphisms $\theta_s : \mathcal{B}_t \to \mathcal{B}_{t^s}$ on $\mathcal{B}_t = \mathbb{C}_{\mathrm{x}}^t \bar{\otimes} \mathcal{N}_t$. The shift describes the free irreversible evolution $\mathbb{C}_{\mathrm{x}}^{t-s} \bar{\otimes} \mathcal{N}_t \to \mathbb{C}_{\mathrm{x}}^t \bar{\otimes} \mathcal{N}_{t+s}$ of the semiquantum bath described by the preoutput W*-algebras $\mathcal{M}_0^t = \mathbb{C}_{\mathrm{x}}^t$ generated by the observable process x on the Fock space F_0^t up to time t and the input quantum noise W*-algebras $\mathcal{N}_t = \mathcal{B}(F_t)$ by injective shifting of each subalgebra $\mathcal{M}_{t-s}^{s]} \simeq \mathcal{M}_{t-s}^s \bar{\otimes} \mathbb{A}^{\mathrm{o}}$ into $\mathcal{N}_{[t}^s \simeq \mathbb{A}^{\mathrm{o}} \bar{\otimes} \mathcal{N}_t^s$.

Hudson and Parthasarathy[51,58] derived a QS *forward equation*

$$\mathrm{dU}(t) = \mathrm{U}(t)\left(\mathrm{K}_+^- \mathrm{d}t + \mathrm{K}^- \mathrm{dA}_-^\circ + \mathrm{K}_+ \mathrm{dA}_\circ^+ + \mathrm{K}\mathrm{dA}_\circ^\circ\right) \tag{33}$$

defining a stationary QS-continuous evolution $\mathrm{U}_r(t)$ for the case $G^t = F_0^t$ by iterating the corresponding QS integral equation

$$\mathrm{U}(t) = \mathrm{U}(r) + \mathfrak{i}_r^t(\mathrm{U}\mathbf{K}) \text{ with } \mathrm{U}(r) = \mathrm{I}. \tag{34}$$

It is written in terms the QS integrators $\mathrm{A}_\mu^\nu(t)$ of four adapted operator-valued coefficients K_+^-, $\mathrm{K}_\circ^- = \mathrm{K}^-$, $\mathrm{K}_+^\circ = \mathrm{K}_+$ and $\mathrm{K}_\circ^\circ = \mathrm{K}$ which are usually taken from the quantum object algebra, $\mathrm{K}_\nu^\mu \simeq \mathtt{K}_\nu^\mu \in \mathbb{A}^{\mathrm{o}}$. They gave necessary and sufficient algebraic conditions for unitarity of this solution in terms of such constant bounded operators $\mathtt{K}_\nu^\mu$, which can be easily written in terms of the *scatterings* $\mathtt{S}_\nu^\mu = \delta_\nu^\mu \mathtt{I}^{\mathrm{o}} + \mathtt{K}_\nu^\mu$ as

$$\mathtt{S}^\dagger = \mathtt{S}^{-1},\ \mathtt{S}^- = -\mathtt{S}_+^\dagger \mathtt{S},\ 2\,\mathsf{Re}\left(\mathtt{S}_+^-\right) = -\mathtt{S}_+^\dagger \mathtt{S}_+. \tag{35}$$

($\mathsf{Re}(\mathtt{S})$ denotes the Hermitian part $\left(\mathtt{S} + \mathtt{S}^\dagger\right)/2$ of an operator $\mathtt{S}$).

It is important for the control to have also the time-dependent $\mathtt{K}_\nu^\mu(t) \in \mathbb{A}^{\mathrm{o}}$, and even $\mathbb{C}_{\mathrm{x}}^t$-adapted operator-valued coefficients $\mathtt{K}_\nu^\mu(t, \mathrm{x}) \in \mathbb{A}_{\mathrm{x}}^t$ for feedback control, where $\mathbb{A}_{\mathrm{x}}^t \simeq \mathbb{C}_{\mathrm{x}}^t \bar{\otimes} \mathbb{A}^{\mathrm{o}}$ to be specified by the preoutput algebras

$\mathbb{C}_{\mathrm{x}}^t \subseteq \mathcal{B}(G^t)$. To this end the HP equation ( ) was generalized by Belavkin in[10,21,25] for any increasing family $(\mathcal{A}^t)$ of W*-algebras $\mathcal{A}^t \subseteq \mathcal{B}(G^t)$ containing $\mathbb{C}_{\mathrm{x}}^t$. It is usually generated by an independent increment subprocess up to time t of the pseudo-Poisson quantum noise $\mathbf{A}(t)$ represented on F_0^t by the quadruple $\left[\mathrm{A}_\mu^\nu\right]_{\mu=-,\circ}^{\nu=\circ,+}$ where A_μ° denotes a column of A_μ^j for each $\mu = -, m$ and $\mathrm{A}_\circ^\nu$ denotes a row of A_j^ν for each $\nu = n, +$ with j, m, n say, in $J = \{0, 1, \ldots, d\}$. (Or quadratic block-matrix $\left\{\mathrm{A}_\mu^\nu\right\}$ with $\mathrm{A}_\mu^\nu = 0$ if $\mu = +$ or $\nu = -$.) It was shown that the HP unitarity relations (35) for an adapted unitary evolution U can simply be expressed as $\mathbf{S}^\star = \mathbf{S}^{-1}$ in terms of the *right QS logarithmic germ*

$$\mathbf{S}(t) = \mathrm{U}_r(t)^{-1}\, \mathring{\mathrm{U}}_r\left(t^+\right) \equiv \mathring{\mathrm{U}}_t\left(t^+\right) \quad \forall r < t$$

as the block-matrix $\mathbf{S} = \mathbf{I}_t + \{\mathrm{K}_\nu^\mu\}_{\nu=-,\circ,+}^{\mu=-,\circ,+}$ with the elements $\mathrm{S}_+^-(t) \in \mathcal{A}^{t]}$, $\mathrm{S}^-(t) \in \mathcal{A}^{t]} \bar{\otimes} \mathfrak{k}$, $\mathrm{S}_+(t) \in \mathcal{A}^{t]} \bar{\otimes} \mathfrak{k}^\dagger$ and $\mathrm{S}(t) \in \mathcal{A}^{t]} \bar{\otimes} \mathcal{B}(\mathfrak{k})$ and equal to zero if $\mu > \nu$ in terms of the order $- < \circ < +$. Here $\mathbf{I}_t = \{\delta_\kappa^\mu \mathrm{I}_t\}$ with the identity I_t on H_t and $\mathbf{S}^\star$ denotes the pseudo-Hermitian conjugation (Belavkin involution) defined at each t by $\left(\mathrm{S}_{-\nu}^\mu\right)^\star = \left(\mathrm{S}_{-\mu}^{\nu\dagger}\right)$ in terms of the reflection $-\mp = \pm$ on the index set $\{-, J, +\}$ leaving invariant the values $\mu, \nu \in J$, see more on this and for the definition of the QS germ in the Appendix and.[21,34] Moreover, it was proved in[21] that $\mathrm{U}_t(t^s)$ is isometric iff the scattering germ $\mathbf{S}(t) = \mathring{\mathrm{U}}_t(t^+)$ is pseudo-isometric, $\mathbf{S}(t)^\star \mathbf{S}(t) = \mathbf{I}_t$ for all $t \leq t^s$, and this remains true, with $\mathbf{I}_t = \{\delta_\kappa^\iota \mathrm{I}_t\}_{\kappa=-,\bullet,+}^{\iota=-,\bullet,+}$ indexed by a subset $J_\bullet \subseteq J$ of $\iota = i$, $\kappa = k$, say, in $J_\bullet = \{1, \ldots, d\}$ for the rectangular $\mathbf{S} = \{\mathrm{S}_\kappa^\mu\}$ with $\kappa \in \{+, J_\bullet, -\}$ and $\mu \in \{-, J, +\}$. This condition, which can be written in the form (35) for $\mathrm{S}^- = \mathrm{K}_\bullet$, $\mathrm{S}_+^- = \mathrm{K}_+^-$ with isometricity $\mathrm{S}_\bullet(t)^\dagger \mathrm{S}_\bullet(t) = \mathrm{I}_t$ replacing the unitarity $\mathrm{S}^\dagger = \mathrm{S}^{-1}$ by semi-unitarity of $\mathrm{S}^\dagger$, defined as the adjoint to an isometry $\mathrm{S} = \mathrm{S}_\bullet(t)$ of $\mathfrak{k}^\bullet \otimes H_\mathrm{o}$ into $\mathfrak{k} \otimes H_\mathrm{o}$, is also necessary (and sufficient under a natural QS-integrability[21]) in the case of unbounded $\mathrm{K}_\kappa^\mu(t)$, densely defined operators in H_t.

Let us now consider an adapted QS semimartingale $\mathtt{X}(t)$ on $H_{[0}$, by which we simply mean an operator-valued process $\mathtt{X}(t) = \mathtt{X}(r) + \mathfrak{i}_r^t(\mathbf{D})$ that has a QS derivative $\mathbf{D}_\mathtt{X}^t = \mathbf{D}(t)$ given at all $t > r$ by a locally QS-integrable operator-functions $\mathtt{D}_\nu^\mu(t)$ (with zero values for $\mu = +$ and for $\nu = -$). It was proved in[25] that if $\mathtt{X}(t)$ is intertwined by the isometries $\mathrm{U}_r(t)$ on H_t with $\check{\mathtt{X}}(t)$,

$$\mathtt{X}(t)\, \mathrm{U}_r(t) = \mathrm{U}_r(t)\, \check{\mathtt{X}}(t) \quad \forall t > r,$$

then $\check{\mathtt{X}}(t) = \check{\mathtt{X}}(r) + \mathfrak{i}_r^t(\mathbf{C}) = \mathrm{U}_r(t)^\dagger\, \mathtt{X}(t)\, \mathrm{U}_r(t)$. Moreover, the QS derivative $\mathbf{D}_\mathtt{X}^t = [\mathrm{C}_\kappa^\iota]_{\kappa=+,\bullet}^{\iota=-,\bullet} \equiv \mathbf{C}(t)$ is uniquely defined by intertwining $\mathbf{J}_\mathtt{X}^t \mathbf{G}_r(t) =$

$\mathbf{G}_r(t)\,\mathbf{J}^t_{\mathtt{\breve{X}}}$ the *right QS germ*

$$\mathbf{J}^t_{\mathtt{X}} = \mathring{\mathtt{X}}\left(t^+\right) = \mathbf{X}(t) + \mathbf{D}(t) \equiv \mathbf{J}(t)$$

with the germ $\mathbf{J}_{\breve{\mathtt{X}}} = \breve{\mathbf{X}} + \mathbf{D}_{\breve{\mathtt{X}}}$ of $\breve{\mathtt{X}}(t)$ by the pseudo-isometry $\mathbf{G}_r(t) = \mathrm{U}_r(t)\,\mathbf{S}(t)$ as the right QS germ $\mathbf{J}^t_{\mathrm{U}_r} = \mathring{\mathrm{U}}_r(t^+)$ of $\mathrm{U}_r(t)$. Note that in the adapted case $\mathtt{X}(t) \simeq \mathtt{X}^{t]} \otimes \mathrm{I}_t$ this can be shown[25] by applying the HP Itô formula[50] corresponding to $\mathbf{X} = \{\mathtt{X}\delta^\mu_\nu\} \equiv \mathtt{X}\mathbf{I}$:

$$\mathbf{D}_{\breve{\mathtt{X}}} = \mathbf{G}_r{}^\star\mathbf{J}\mathbf{G}_r - \breve{\mathtt{X}}\mathbf{I} = \mathbf{G}_r{}^\star\mathbf{D}\mathbf{G}_r + \mathbf{G}_r{}^\star\mathtt{X}\mathbf{G}_r - \breve{\mathtt{X}}\mathbf{I},$$

where $\mathbf{I}(t) = \{\delta^t_\kappa \mathrm{I}_t\} = \mathbf{J}^t_{\mathrm{I}}$. However, it is more convenient to use the Belavkin rule[21] that *The QS germ of any product (or a function) of (not necessary commuting) processes is the product (or the function) of their germs*. It holds even in the nonadapted case,[21] and it simply implies the HP Itô formula in terms of the QS derivatives $\mathbf{D}^t_{\breve{\mathtt{X}}} = \mathbf{J}^t_{\breve{\mathtt{X}}} - \breve{\mathbf{X}}(t)$ with $\breve{\mathbf{X}}(t) = \breve{\mathtt{X}}(t)\,\mathbf{I}_t$ for the adapted $\breve{\mathtt{X}}(t)$. If

$$\mathbf{J}^t_{\mathtt{X}_r} = \mathbf{G}_r(t)\,\mathbf{J}^t_{\breve{\mathtt{X}}}\mathbf{G}_r(t)^\star \equiv \boldsymbol{\gamma}_r\left(t, \mathbf{J}^t_{\breve{\mathtt{X}}}\right), \tag{36}$$

then $\mathtt{X}_r(t)$ can be uniquely defined by this rule as the forward solution

$$\mathtt{X}_r(t) = \mathrm{U}_r(t)\,\breve{\mathtt{X}}(t)\,\mathrm{U}_r(t)^\dagger \equiv \alpha_r\left(t, \breve{\mathtt{X}}(t)\right),$$

corresponding to the boundary condition $\mathtt{X}_r(r) = \breve{\mathtt{X}}(r) \in \mathcal{A}^{r]}$ for the *QS Langevin equation*

$$\mathrm{d}\mathtt{X}(t) = \left(\boldsymbol{\gamma}_r\left(t, \mathbf{J}^t_{\breve{\mathtt{X}}}\right) - \mathtt{X}(t)\,\mathbf{1}\right)\cdot \mathrm{d}\mathbf{A}(t). \tag{37}$$

Obviously this is the case when the germ transformation (36) satisfies the unitality condition $\boldsymbol{\gamma}(\mathbf{I}) = \mathbf{I}$, i.e. if $\mathbf{G}$ is pseudounitary, otherwise $\mathbf{P}_r(t) = \boldsymbol{\gamma}_r(t, \mathbf{I}) = \mathbf{P}_r(t)^\star$ is a multiplicative hemigroup of pseudo-Hemitian projectors $\mathbf{P}_r^2 = \mathbf{P}_r$. This simply follows for $\mathbf{J} = \mathbf{I}$ from the $\star$-multiplicativity of the linear map $\boldsymbol{\gamma}_r(t) = \mathring{\alpha}_r(t^+)$ such that

$$\boldsymbol{\gamma}_r(\mathbf{J}^\star\mathbf{J}) = \mathbf{G}_r\mathbf{J}^\star\mathbf{J}\mathbf{G}_r^\star = \mathbf{G}_r\mathbf{J}^\star\mathbf{G}_r^\star\mathbf{G}_r\mathbf{J}\mathbf{G}_r^\star = \boldsymbol{\gamma}_r(\mathbf{J})^\star\,\boldsymbol{\gamma}_r(\mathbf{J}) \tag{38}$$

for any pseudo-isometric $\mathbf{G}$. These projectors define the QS germ $\mathring{\mathtt{F}}_0(t^+) = \mathbf{P}_0(t)$ of decreasing orthoprojectors $\mathtt{F}_0(t) = \mathrm{U}_0(t)\,\mathrm{U}_0(t)^\dagger \geq \mathtt{F}_0(t^s)$, representing on $H_0 = F_{[0}$ the quantum *property of survival* (the nondemolition property) for an open unstable quantum particle if it has not been completely observed up to time t.

We shall say that the hemigroup U *consistently* describes a *quantum open dynamical system* on a fixed object W*-algebra $\mathbb{A}^{\mathrm{o}} \subseteq \mathcal{B}(H_{\mathrm{o}})$ with

respect to a W*-product system $(\mathcal{A}^t)$ of W*-algebras on (G^t) if it induces in the Heisenberg picture

$$\alpha_r(t,\ddot{\mathrm{X}}) = \mathrm{U}_r(t)\,\ddot{\mathrm{X}}\mathrm{U}_r(t)^\dagger, \quad \ddot{\mathrm{X}} \in \mathcal{A}^{t]} \tag{39}$$

a forward-adapted input-output transformation $\alpha_t(t^s) : \mathcal{A}^{t+s]} \to \mathcal{A}^{t]}\bar{\otimes}\mathcal{N}_t^s$ satisfying the *consistency condition*

$$\alpha_t(t^s,\mathrm{XYZ}) = \mathrm{X}\alpha_t(t^s,\mathrm{Y})\,\mathrm{Z} \;\; \forall \mathrm{X} \in \mathcal{A}^t, \mathrm{Y} \in \mathcal{A}_t^{s]}, \mathrm{Z} \in \mathcal{N}_{t^s} \tag{40}$$

where $\mathcal{A}_t^{s]} \simeq \mathcal{A}_t^s\bar{\otimes}\mathbb{A}^{\mathrm{o}}$. Note that the last condition is met by assuming that $\mathrm{U}_t(t^s)$ is adapted and $[\mathrm{U}_t(t^s),\mathcal{A}^t] = 0$ for all $t < t^s$. Obviously the two-sided adapted HP evolution corresponding to the *Markovian case* $\alpha_t\left(t^s,\mathcal{A}_t^{s]}\right) \subseteq \mathcal{N}_{[t}^s$ trivially satisfies the condition (40) for $\mathbb{A}^{\mathrm{o}} = \mathcal{B}(H_{\mathrm{o}})$ and any $\mathcal{A}^t \subseteq \mathcal{B}(G^t)$, but in general it requires much less.

Proposition 2.2. *The dynamical consistency condition (40) is equivalent to the* semi-Markovianity $\alpha_t\left(t^s,\mathcal{A}_t^{s]}\right) \subseteq \mathcal{C}^{t]}\bar{\otimes}\mathcal{N}_t^s$, *where* $\mathcal{C}^{t]} \simeq \mathcal{C}^t\bar{\otimes}\mathbb{A}^{\mathrm{o}}$ *is given by an Abelian W*-algebra* $\mathcal{C}^t$ *from the center* $C(\mathcal{A}) = \mathcal{A}\cap\mathcal{A}'$ *of* $\mathcal{A}^t$.

Proof. Since $\left[\mathcal{A}_t^{s]},\mathcal{A}^t\right] = 0$, the concistency (40) is equivalent to the commutatativity $\left[\alpha_t\left(t^s,\mathcal{A}_t^{s]}\right),\mathcal{A}^t\right] = 0$ for the adapted $\alpha_t\left(t^s,\mathcal{A}_t^{s]}\right) \simeq \mathbb{B}_t^{s]}\bar{\otimes}\mathrm{I}_{t^s}$. Thus,

$$\alpha_t\left(t^s,\mathcal{A}_t^{s]}\right) \subseteq \left(\mathcal{A}^{t]}\bar{\otimes}\mathcal{N}_t^s\right)\backslash\mathcal{A}^t = \left(\mathcal{A}^{t]}\backslash\mathcal{A}^t\right)\bar{\otimes}\mathcal{N}_t^s = \left(\mathcal{A}^t\backslash\mathcal{A}^t\right)\bar{\otimes}\mathcal{N}_{t]}^s,$$

where the relative commutant $\mathcal{A}^t\backslash\mathcal{A}^t$ by definition is the center $C(\mathcal{A}^t)$. □

Thus, the consistent quantum open dynamics $\alpha_t(t^s)$ can be controlled not only by the current states on the quantum object and noise algebra $\mathcal{N}_{[t}^s$, but also by the classical history states on an increasing commutative W*-subalgebras $\mathcal{C}^t \subseteq C(\mathcal{A}^t)$.

We shall assume that the *stochastic control* algebras $(\mathcal{C}^t)$ form a W*-product system $\mathcal{C}^{t+s} \simeq \mathcal{C}^t\bar{\otimes}\mathcal{C}_t^s$ being generated, $\mathcal{C}^t = \mathbb{C}_{\mathrm{u}}^t$, by a joint representation $\{\mathrm{u}(r) : r \le t\}$ of one or several classical controlling processes $u(r,\omega)$ up to t, and shall call them *feedback control* algebras with respect to $(\mathbb{C}_{\mathrm{x}}^t)$ if they are adapted in the sense $\mathbb{C}_{\mathrm{u}}^t \subseteq \mathbb{C}_{\mathrm{x}}^t$ for all t to the preoutput algebras $(\mathbb{C}_{\mathrm{x}}^t)$. By reducing the total preoutput algebras $\mathcal{A}^t$ to the relative commutants $\mathcal{A}^t\backslash\mathbb{C}_{\mathrm{x}}^t$, we can have the admissable feedback control algebras $\mathbb{C}_{\mathrm{u}}^t$ coinciding with $\mathcal{M}^t = \mathbb{C}_{\mathrm{x}}^t$ if no other constraints, apart from causality,

is *a priori* given. Due to the consistency condition (40) the Heisenberg dynamics $\alpha_r(t) = \alpha^{\mathrm{u}}_r(t)$ controlled by the states on $\mathbb{C}^r_{\mathrm{u}}$ are decomposable for any $\mathrm{Y} \in \mathcal{A}^t$ with respect to the central subalgebra $\mathbb{C}^r_{\mathrm{u}}$ as

$$\alpha^{\mathrm{u}}_r(t, \mathrm{Y}_t) = \int_{\Upsilon^r} \alpha^{\upsilon}_r(t, \mathrm{Y}^{\upsilon}_t)\, e_{\mathrm{u}}(\mathrm{d}\upsilon) = \alpha^{\mathrm{u}}_r(t, \mathrm{Y}^{\mathrm{u}}_t) \tag{41}$$

for any $\mathrm{Y}_t \in \mathcal{A}^{t]}$ decomposed as $\mathrm{Y}_t = \int_{\Upsilon^r} \mathrm{Y}^{\upsilon}_t e_{\mathrm{u}}(\mathrm{d}\upsilon)$ such that $\alpha^{\upsilon}_t(t^s, \mathrm{Y}^{\upsilon}_{t^s}) \in \mathcal{N}^s_{[t}$ a.s. for all $\upsilon \in \Upsilon^t$ if $\mathrm{Y}_{t^s} \in \mathbb{C}^t_{\mathrm{u}} \bar{\otimes} \mathcal{A}^{s]}_t$.

In fact, since physically realizable control of quantum dynamics can be performed through the input quantum channels by controlling the input noise states on the algebra $\mathcal{N}$ we might need only the QS evolutions controlled by the input stochastic coherent state vectors (12). They can be defined by an adapted process $\varsigma_j(t, \omega) = \zeta^u_{r,j}(t, \upsilon_r)$ of complex stochastic amplitudes indexed by $j \in J$, where $u = \{u(t), t \le r\}$ and $\upsilon_r \in \Upsilon_r$. Let $\mathrm{U}^{\mathrm{u}}_t(t^s) = \mathrm{U}^{s\dagger}_{\widehat{\zeta}^{\mathrm{u}}_t} \mathrm{U}_t(t^s)$ denote a coherent controlled QS isometric evolution given by QS Weyl transform $\mathrm{U}^{t-r}_{\widehat{\zeta}^{\mathrm{u}}_r} = \mathrm{W}_t\left(r, \mathbf{a}_{\widehat{\zeta}^{\mathrm{u}}_r}\right)$ with the column $\widehat{\zeta}^{\mathrm{u}}_r(t) = \varsigma_\circ(t, \mathrm{u}, \hat{\mathrm{v}}_r)^\dagger$ of

$$\widehat{\zeta}^{\mathrm{u}}_{r,j} = \int_{\Upsilon_r} \zeta^{\mathrm{u}}_{r,j}(\upsilon)\, e_{\hat{\mathrm{v}}}(\mathrm{d}\upsilon) \equiv \zeta^{\mathrm{u}}_{r,j}(t, \hat{\mathrm{v}}_r),$$

where $\hat{\mathrm{v}}_r(t) = \mathtt{y}_r(t)$ denotes the output representation of the commutative input processes $\breve{\mathrm{v}}_r(t) = \mathrm{x}_r(t)$ such that $\mathtt{y}_r(t)\mathrm{U}^{\mathrm{u}}_r(t) = \mathrm{U}^{\mathrm{u}}_r(t)\mathrm{x}_r(t)$. Then $\mathrm{U}_t(t^s)$ intertwines $\hat{\mathrm{X}}\mathrm{U}^{\mathrm{u}}_t = \mathrm{U}^{\mathrm{u}}_t \breve{\mathtt{X}}$ the representations $\hat{\mathrm{X}}(t^s) = \mathrm{U}^{s\dagger}_{\widehat{\zeta}^{\mathrm{u}}_t} \mathtt{X}(t^s)\, \mathrm{U}^{s}_{\widehat{\zeta}^{\mathrm{u}}_t}$ of $\mathtt{X}(t^s)$ on $F_{[0}$ intertwined by the (uncontrolled) evolution $\mathrm{U}_t(t^s)$ with $\breve{\mathtt{X}}(t^s)$ on H_{t^s}. Here the Weyl transform $\mathrm{U}^{t-r}_{\widehat{\zeta}^{\mathrm{u}}_r}(r)$ is defined as a solution of the QS equation (24) at $t > r$ with $\mathrm{W}_r(r) = \mathrm{I}$ by the generators $\mathbf{a}_{\widehat{\zeta}^{\mathrm{u}}_r}(t) = \mathbf{g}_{\widehat{\zeta}^{\mathrm{u}}_r(t)} - \mathbf{1}$ with QS vector-amplitudes $\widehat{\zeta}^{\mathrm{u}}_r = \left(\widehat{\zeta}^{\mathrm{u},j}_r : j \in J\right)$ such that the coherent controlled Heisenberg transform

$$\alpha^{\mathrm{u}}_r(t, \breve{\mathtt{X}}) = \mathrm{U}^{\mathrm{u}}_r(t)\, \breve{\mathtt{X}} \mathrm{U}^{\mathrm{u}}_r(t)^\dagger = \hat{\mathrm{X}}_r(t)\, \hat{\mathrm{F}}_r(t)$$

is given by the QS Weyl transform $\hat{\mathrm{F}}_r(t)$ of the orthoprojector $\mathtt{F}_r(t)$ for uncontrolled $\alpha_t(t^s, \breve{\mathtt{X}}) = \mathtt{X}_t(t^s)\, \mathtt{F}_r(t)$.

Similarly, one can always represent the observable and controlling real-valued processes $x(t)$ and $u(t)$ in Fock space with respect as

$$\mathrm{y}(t) = \mathrm{U}^{t\dagger}_{\xi^{\mathrm{v}}_0} \mathrm{u}(t)\, \mathrm{U}^{t}_{\xi^{\mathrm{v}}_0}, \quad \mathrm{v}(t) = \mathrm{U}^{t\dagger}_{\xi^{\mathrm{v}}_0} \mathrm{x}(t)\, \mathrm{U}^{t}_{\xi^{\mathrm{v}}_0}$$

to the vacuum state-vector $\delta_\emptyset \in \Gamma(\mathcal{E}_\bullet)$ if they were locally given in the Fock spaces $G^t = \Gamma(\mathcal{E}^t_\bullet)$ with the reference measure induced by the coherent vector-states $\mathrm{U}^t_{\xi^{\mathrm{v}}_0}\delta_\emptyset$. Here $\xi^{\mathrm{v}}_0 = \left\{\xi^{\mathrm{v}}_\bullet(t)^\dagger \,|t \ge 0\right\}$ are given by the

row-amplitudes $\xi^x_\bullet(t) \in \mathfrak{k}_\bullet$ which may causally depend on x represented in $\Gamma(\mathcal{E}_0^\bullet)$. The dynamics $\alpha^{\rm u}_r(t)$ changes covariantly under such transformation to

$$\beta_r(t,{\rm X}) = {\rm V}_r(t)\,{\rm X}_t{\rm V}_r(t)^\dagger \equiv \mathtt{X}_r(t)\,, \quad {\rm X}_t \in \mathbb{A}^t_{\rm v} \tag{42}$$

on the locally unitary equivalent algebras $\mathbb{A}^s_{\rm v} = {\rm U}^{t\dagger}_{\xi^{\rm v}_0}\mathbb{A}^s_{\rm x}{\rm U}^t_{\xi^{\rm v}_0}$ by perturbing the controlled isometric evolution ${\rm U}_t(t^s)$ to ${\rm V}_t(t^s) = {\rm U}_t(t^s)\,{\rm U}^s_{\xi^{\rm v}_t}(t)$. Here the QS Weyl transform ${\rm U}^{t-r}_{\xi^{\rm v}_r}(r) = {\rm W}_t\left(r, \mathbf{a}_{\xi^{\rm v}_r}\right)$ is defined as a unitary operator on the Fock space $G = \Gamma(\mathcal{E}_0^\bullet)$ by the complex amplitudes $\xi^x_j(t)$, $j \in J_\bullet$ which may causally depend on the preoutput stochastic processes $x(\upsilon,t) = \upsilon(t)$.

Theorem 2.2. *Let $\varsigma_\circ = (\varsigma_j)_{j\in J}$ be stochastic row-amplitude with $\mathfrak{B}^{t-r}_r$-measurable sections $\upsilon \mapsto \zeta^u_{r,j}(t,\upsilon) = \varsigma_j(t,u,\upsilon)$ as spectral values of $\widehat{\zeta}^u_{r,j}$ a.s. in $L^2_{\rm loc}(\mathbb{R}^+_0)$ for each $\omega = (u,\upsilon)$, and let $\xi^{\rm v}_\bullet(t)$ be row-vector amplitude with spectral values ξ^υ_j a.s. in $L^2_{\rm loc}(\mathbb{R}^+_0)$, $j \in J_\bullet$. Assume that $\breve{\mathtt{X}}(t) = {\rm U}^t_{\xi^{\rm v}_0}{\rm X}(t)\,{\rm U}^{t\dagger}_{\xi^{\rm v}_0}$ is the Weyl transform of an $\mathbb{A}^t_{\rm v}$-adapted semi-martingale ${\rm X}(t)$, and let $\beta^-_r({\rm X}) = {\rm U}_r{\rm X}{\rm U}^\dagger_r$ denote the uncontrolled QS dynamics $\mathbb{A}^t_{\rm v} \to \mathbb{A}^r_{\rm v}\bar{\otimes}\mathcal{N}_r$ given by the germ $\mathbf{S}(t) = \mathring{\rm U}_t(t^+)$ at each $t \geq r$ given by a quadruple $[{\rm S}^\nu_\kappa]^{\nu=-,\circ}_{\kappa=\bullet,+}$ of QS-integrable $\mathbb{A}^{\rm o}$-valued functions ${\rm S}^\mu_\kappa(t) \simeq {\rm I}_t\bar{\otimes}{\rm S}^\mu_\kappa(t)$. Then the coherent controlled QS dynamics (42) is defined by the isometries ${\rm V}_t(t^s) = {\rm U}^{s\dagger}_{\widehat{\zeta}^{\rm u}_t}{\rm U}_t(t^s)\,{\rm U}^s_{\xi^{\rm v}_t}$ with the logarithmic QS germ $\mathring{\rm V}_t(t^+) = \mathbf{I}(t) + \mathbf{L}(t) \equiv \mathbf{T}(t)$ given by the injection $\mathbf{I}(t) = \{\delta^\mu_\kappa {\rm I}_t\}^{\mu=-,\circ,+}_{\kappa=-,\bullet,+}$ and a quadruple $\mathbf{T}(t,{\rm v}) = \mathbf{S}^{\rm v}(t)\,\mathbf{g}_{\xi^{\rm v}(t)}$ of controlled generators*

$$\mathbf{T} = \begin{bmatrix} ({\rm K}^{\rm v}_\bullet - \xi^{\rm v}_\bullet)\,{\rm e}^{{\rm i}\theta^{\rm v}_\bullet} & {\rm K}^{\rm v}_+ + {\rm K}^{\rm v}_\bullet\xi^{{\rm v}\dagger}_\bullet - \frac{1}{2}\left\|\xi^{\rm v}_\bullet\right\|^2_{\mathfrak{k}_\bullet}{\rm I} \\ {\rm S}^{\rm v}_\bullet{\rm e}^{{\rm i}\theta^{\rm v}_\bullet} & {\rm S}^{\rm v}_+ + {\rm S}^{\rm v}_\bullet\xi^{{\rm v}\dagger}_\bullet \end{bmatrix} \equiv \begin{bmatrix} {\rm L}_\bullet & {\rm L}_+ \\ {\rm T}_\bullet & {\rm T}_+ \end{bmatrix}, \tag{43}$$

$$\mathbf{S}^{\rm v} = \begin{bmatrix} {\rm K}_\bullet + \zeta^{\rm v}_\circ{\rm S}_\bullet & {\rm K}_+ + \zeta^{\rm v}_\circ{\rm S}_+ - \frac{1}{2}\left\|\zeta^{\rm v}_\circ\right\|^2_{\mathfrak{k}}{\rm I} \\ {\rm e}^{-{\rm i}\theta^{\rm v}_\circ}{\rm S}_\bullet & {\rm e}^{-{\rm i}\theta^{\rm v}_\circ}({\rm S}_+ - \zeta^{\rm v}{\rm I}) \end{bmatrix} \equiv \begin{bmatrix} {\rm K}^{\rm v}_\bullet & {\rm K}^{\rm v}_+ \\ {\rm S}^{\rm v}_\bullet & {\rm S}^{\rm v}_+ \end{bmatrix}. \tag{44}$$

Proof. Indeed, according to quantum Itô formula the germ of the product ${\rm V}_t(t^s) = {\rm U}^{s\dagger}_{\widehat{\zeta}^{\rm u}_t}{\rm U}_t(t^s)\,{\rm U}^s_{\xi^{\rm v}_t} = {\rm U}^{\rm u}_t(t^s)\,{\rm U}^s_{\xi^{\rm v}_t}$ is the product of the germs,

$$\begin{aligned} {\rm V}_r(t)\,\mathbf{T}(t) &= \mathbf{g}^\star_{\zeta^{\rm u}_r(t,{\rm y}_r)}{\rm U}^{t-r\dagger}_{\widehat{\zeta}^{\rm u}_r}{\rm U}_r(t)\,\mathbf{S}(t)\,{\rm U}^{t-r}_{\xi^{\rm v}_r}\mathbf{g}_{\xi^{\rm v}_r(t)} \\ &= {\rm U}^{\rm u}_r(t)\,\mathbf{g}^\star_{\zeta^{\rm u}_r(t,{\rm x}_r)}{\rm U}^{t-r}_{\xi^{\rm v}_r}\mathbf{S}(t)\,\mathbf{g}_{\xi^{\rm v}_r(t)} \\ &= {\rm U}^{\rm u}_r(t)\,{\rm U}^{t-r}_{\xi^{\rm v}_r}\mathbf{g}^\star_{\zeta^{\rm u}_r(t,{\rm v}_t)}\mathbf{S}(t)\,\mathbf{g}_{\xi^{\rm v}_r(t)}, \end{aligned}$$

where we rearranged the order by intertwining the representations ${\rm x}_r$ with ${\rm y}_r$ by ${\rm U}^{\rm u}_r(t)$ and ${\rm v}_r$ with ${\rm x}_t$ by ${\rm U}^{t-r}_{\xi^{\rm v}_r}$ and using the commutativity of ${\rm u} =$

$\{\mathrm{u}(t)\,|t\leq r\}$ with $\mathrm{U}_{\xi^{\mathrm{v}}_r}^{t-r}$. Therefore, by taking $r=t$ we obtain $\mathbf{T}(t)=\mathbf{g}^{\star}_{\zeta^{\mathrm{u}}(t)}\mathbf{S}(t)\,\mathbf{g}_{\xi^{\mathrm{v}}_r(t)}$ since by assumption of causality $\zeta^u_t(t,\upsilon)$ does not depend on $\upsilon_t\in\Upsilon_t$.

Multiplying the germ-matrices $\mathbf{S}^{\mathrm{y}}(t)=\mathbf{g}^{\star}_{\zeta^{\mathrm{u}}(t)}\mathbf{S}(t)$ and $\mathbf{g}_{\xi^{\mathrm{v}}(t)}$ corresponding to the germ quadruples

$$\mathbf{S}^{\mathrm{y}}(t)=\begin{bmatrix}\mathrm{K}^{\mathrm{y}}_{\bullet}&\mathrm{K}^{\mathrm{y}}_{+}\\ \mathrm{S}^{\mathrm{y}}_{\bullet}&\mathrm{S}^{\mathrm{y}}_{+}\end{bmatrix},\ \mathbf{g}_{\xi^{\mathrm{v}}(t)}=\begin{bmatrix}-|\xi^{\mathrm{v}}_{\bullet}(t)|-\|\xi^{\mathrm{v}}_{\bullet}(t)\|^2/2 & \\ \mathrm{e}^{\mathrm{i}\theta_{\bullet}(t,\mathrm{v})} & \xi^{\mathrm{v}}_{\bullet}(t)^{\dagger}\end{bmatrix}$$

we immediately obtain the result (43) and (44). □

3. Stochastic Master and Filtering Equations

Let us fix an initial product state $\mathtt{R}_{[0}=\varrho\otimes\mathrm{R}_0$ on $\mathbb{B}_{[0}\simeq\mathbb{A}^{\mathrm{o}}\bar{\otimes}\mathcal{N}_0$, with a reference state R_0 on $\mathcal{N}_0$ induced by normal, say, coherent states $\mathrm{P}_{\zeta^t_0}$ on each $\mathcal{N}^t_0$ given by $\zeta_0|I^t_0\equiv\zeta^t_0\in L^2(I^t_0\to\mathfrak{k})$. We denote by $\mathtt{R}^{\star}_{[0}$ the corresponding expectation $\epsilon_{[0}[\mathtt{X}]=\langle\mathtt{R}_{[0},\mathtt{X}\rangle$ on $\mathbb{B}_{[0}$, and by $\mathrm{R}^{\star}_t$ the conditional expectation on $\mathcal{B}_t\simeq\mathcal{A}^{t]}\bar{\otimes}\mathcal{N}_t$ as locally normal projection onto $\mathcal{A}^{t]}$, say, coherent-induced by

$$\epsilon^{t]}_{\mathcal{A}}[\mathrm{Z}]=\left(\mathrm{I}^{t]}\otimes f_{\zeta_t}\right)^{\dagger}\mathrm{Z}\left(\mathrm{I}^{t]}\otimes f_{\zeta_t}\right)\equiv\langle\mathrm{R}_t,\mathrm{Z}\rangle^{t]}_t\quad\forall\mathrm{Z}\in\mathcal{B}_t.$$

The adapted isometric dynamics $\alpha(t,\check{\mathtt{X}}(t))=\mathrm{U}(t)\,\check{\mathtt{X}}(t)\,\mathrm{U}(t)^{\dagger}$ is homomorphic, mapping $\mathcal{A}^{t]}\simeq\mathcal{A}^t\bar{\otimes}\mathbb{A}^{\mathrm{o}}$ into $\mathbb{B}^t_{[0}$ with preservation of all algebraic relations except, maybe, the identity which is mapped into the survival projector $\mathtt{F}(t)=\mathrm{U}(t)\,\check{\mathtt{X}}(t)\,\mathrm{U}(t)^{\dagger}$ in the case of noninvertable isometry $\mathrm{U}(t)$. It induces entangled object-output states described by the positive operators $\check{\mathtt{R}}^t_{[0}=\alpha^t_{\star}\left(\mathtt{R}_{[0}\right)\in\mathcal{A}^{t]}_{\star}$ affiliated to $\mathcal{A}^{t]}\simeq\mathcal{A}^t\bar{\otimes}\mathbb{A}^{\mathrm{o}}$ as the densities of the expectation functionals $\mathtt{R}^{\star}_{[0}\circ\alpha(t)$ induced by $\epsilon=\mathtt{R}^{\star}_{[0}$ on $\mathcal{A}^{t]}$:

$$\left\langle\alpha^t_{\star}\left(\mathtt{R}_{[0}\right),\check{\mathtt{X}}(t)\right\rangle_{t]}=\left\langle\mathtt{R}_{[0},\alpha\left(t,\check{\mathtt{X}}(t)\right)\right\rangle_{[0}\,,\quad\forall\check{\mathtt{X}}(t)\in\mathcal{A}^{t]}.\tag{45}$$

If $\mathrm{U}(t)^{\dagger}\,\mathtt{R}_{[0}\mathrm{U}(t)\in\mathcal{A}^{t]}$, obviously $\check{\mathtt{R}}^t_{[0}=\mathtt{U}^t\mathtt{R}_{[0}\mathtt{U}^{t\dagger}$ with respect to the induced trace pairing, where $\mathtt{U}^t=\mathrm{U}(t)^{\dagger}$. However, the completely positive (CP) map $\alpha^t_{\star}:\mathtt{R}_{[0}\mapsto\check{\mathtt{R}}^t_{[0}$, which we call *semimorphic* as preadjoint to homomorphic $\alpha(t)$, is not multiplicative in the case of noninvertible $\mathrm{U}(t)$. Note that the coisometric semimorphic dynamics $\alpha^t_{\star}$ is reversible only if $\alpha\left(t,\mathcal{A}^{t]}\right)=\mathbb{B}^t_{[0}$ which cannot be the case for $\mathcal{A}^{t]}=\mathbb{A}^t_{\mathrm{x}}\simeq\mathbb{C}^t_{\mathrm{x}}\bar{\otimes}\mathbb{A}^{\mathrm{o}}$ if $\mathbb{B}^t_{[0}=\mathcal{B}\left(F^t_{[0}\right)$ even for the unitary dynamics $\mathrm{U}(t)\,\mathtt{U}^t=\mathtt{I}$.

In the general case the induced states $\check{\mathtt{R}}^t_{[0}$ are controlled by the semimorphic dynamics which is well-defined by $\alpha(t)$-marginalizations (45) with

respect to the standard pairings $\left\langle \mathcal{A}_\star^{t]}, \mathcal{A}^{t]} \right\rangle_{t]}$ as preadjoint completely positive (CP) coisometries $\alpha_\star^t = \alpha(t)_\star$. As the maps into the growing spaces $\mathcal{A}_\star^{t]} = \alpha_\star^t\left(\mathbb{B}_{[0\star}^t\right)$ it cannot satisfy a detrministic but a QS forward equation as the predual to the Langevin equation (37). The main purpose of this Section is to derive this stochastic equation for a controlled QS dynamics $\alpha(t)$ with the vacuum state $\mathrm{R}_0 = \mathrm{P}_0$ on $\mathcal{N}_0$, in which case it can be defined as preadjoint to the deterministic backward differential equation for the vaccum-expected completely positive contractive dynamics

$$\mu(t, \check{\mathrm{X}}(t)) = \epsilon_{\mathcal{A}}^{0]}[\alpha(t, \check{\mathrm{X}}(t))] \equiv \mathrm{P}_0^\star \circ \alpha(t, \check{\mathrm{X}}(t)), \ \forall \check{\mathrm{X}}(t) \in \mathcal{A}^{t]}.$$

The general form of the forward QS differential equations for completely positive dynamics continuously entangling the object state $\varrho_t \vdash \mathbb{A}^{\mathrm{o}}$ with the output state P^t induced by $\check{\mathrm{R}}_{[0}^t$ on $\mathcal{A}^t$ was discovered by Belavkin in.[32,33] It is presented in the Appendix B in the form (140) as a QS generalization of the Lindblad equation. Here we shall give the derivation of this equation for the commutative algebras $\mathcal{A}^t = \mathbb{C}_{\mathrm{x}}^t$ generated by the preoutput noise $\{\mathrm{x}_\bullet(r) \,|r \leq t\}$ representing a classical independent increment process $v_\bullet(t) = x_\bullet(v,t)$ on the vacuum of Fock space $G^t = \Gamma(\mathcal{E}_\bullet^t)$ up to the time t.

In the case classical case of Abelian object algebra $\mathbb{A}^{\mathrm{o}}$ this generalized Lindblad equation is equivalent to the well-known Zakai linear stochastic equation describing the stochastic state evolutions driven by $x_\bullet(t)$ for the probability density function $\varpi_x^t = \mu_{\star x}^t(\varrho)$ of the object plus output state given by a probability measure with respect to the reference measure $\lambda^{\mathrm{o}} \otimes \mathsf{Q}$, given an initial density $\varrho \in \mathbb{A}_\star^{\mathrm{o}}$ with respect to a Radon (say, Lebesgues) measure λ^{o} on $\mathbb{A}^{\mathrm{o}}$. In the semiquantum case, when $\mathbb{A}^{\mathrm{o}} = \mathcal{B}(H_{\mathrm{o}})$, the stochastic process ϖ_x^t realizing the states $\varpi_{\mathrm{x}}^t = \check{\mathrm{R}}_{[0}^t \vdash \mathbb{A}_{\mathrm{x}}^t$ satisfies a linear stochastic master equation first derived by Belavkin in a stochastic Schrödinger form separately for the diffusive and counting processes x respectively in[12,13] and then in the unified form in.[18] As in the commutative case the solutions ϖ_x^t to the linear Belavkin equation with initially normalized $\varpi_x^0 = \varrho$ do not remain normalized as the elements $\varpi_x^t \in \mathbb{A}_\star^{\mathrm{o}}$ for each particular trajectory x, but ϖ_{x}^t as an elements of the preadjoint space $\mathcal{A}_\star^{t]}$ to $\mathcal{A}^{t]} = \mathbb{A}_{\mathrm{x}}^t$ is normalized to the decreasing probability of survival such that the stochastic process $p_x^t = \langle \varpi_x^t, \mathtt{I}^{\mathrm{o}} \rangle$ is a positive $\mathbb{C}_{\mathrm{x}}^t$-submartingale defining the Girsanov transformation as an input to output transformation of the reference probability measure Q to $\mathsf{P}^t(\mathrm{d}v) = p_x^t(v)\,\mathsf{Q}(\mathrm{d}v)$. The renormalized states $\varrho_x(t) = \varpi_x^t / p_x^t$ are the conditional states on $\mathbb{A}^{\mathrm{o}}$ which

are described by a nonlinear filtering equation first derived in the classical case of commutative $\mathbb{A}^{\mathrm{o}}$ by Stratonovich[60] and for the simple algebra $\mathbb{A}^{\mathrm{o}} = \mathcal{B}(H_{\mathrm{o}})$ by Belavkin in.[11,25]

3.1. *Quantum Master equations for continuous observation*

In this section we consider the reduced stochastically controlled quantum hemigroup dynamics which is described on the general W*-algebra $\mathbb{A}^{\mathrm{o}}$ by the vacuum expected controlled hemigroups $\mu_r^{\mathrm{u}}(t) = \mathrm{Q}_r^{\star} \circ \alpha_r^{\mathrm{u}}(t)$. Say, $\alpha_r^{\mathrm{u}}(t)$ is the Weyl-perturbed QS isometric evolution $\alpha_r(t)$ coherently controlled by the input sates $\mathrm{R}_r = \mathrm{P}_{\widehat{\zeta}_r^{\mathrm{u}}}$ such that $\mathrm{Q}_r^{\star} \circ \alpha_r^{\mathrm{u}}(t) = \mathrm{P}_{\widehat{\zeta}_r^{\mathrm{u}}}^{\star} \circ \alpha_r(t)$,where $\mathrm{u} = \{\mathrm{u}(t) | t \leq r\} \equiv \mathrm{u}^r$ is given by some classical stochastic processes $u(t, \upsilon)$ adaptively represented in the preoutput algebras $\mathcal{M}^t = \mathbb{C}_{\mathrm{x}}^t \mathtt{I}$ by the causality condition $\mathrm{u}(t) \vdash \mathbb{C}_{\mathrm{x}}^t$. Our purpose is to derive the linear dissipative equations for the general QS controlled isometric dynamics $\alpha_r^{\mathrm{u}}(t) : \mathcal{A}^{t]} \to \mathcal{A}^{r]} \otimes \mathcal{N}_r$ on W*-algebras $\mathcal{A}^{t]} \simeq \mathcal{A}^t \otimes \mathbb{A}^{\mathrm{o}}$ with the increasing $\mathcal{A}^t$ containing at the center the Abelian algebras $\mathbb{C}_{\mathrm{x}}^t$. Following the original derivation by Belavkin[11,18] we shall unify the two most important cases (14): The case of diffusive measurements with coherent induced probability corresponding to the Wiener preoutput process $\mathrm{x}_\bullet = \mathrm{w}_\bullet$ and the case of counting measurements corresponding to the Poisson preoutput process $\mathrm{x}_\bullet = \mathrm{m}_\bullet$.

The quantum state dynamics $\mu_\star^{\mathrm{u},s}(t) : \breve{\mathrm{R}}_{[0}^t \mapsto \breve{\mathrm{R}}_{[0}^{t+s}$ is defined in

$$\left\langle \mu_{\star t}^{\mathrm{u},s}\left(t, \breve{\mathrm{R}}_{[0}^t\right), \mathrm{Y}\right\rangle_{t+s]} := \left\langle \breve{\mathrm{R}}_{[0}^t, \mu_t^{\mathrm{u}}(t^s, \mathrm{Y})\right\rangle_{t]} \quad \forall \mathrm{Y} \in \mathcal{A}^{t+s]}$$

by the preadjoint maps $\mu_\star^{\mathrm{u},s}(t) = \mu_t^{\mathrm{u}}(t^s)_\star$. The maps $\mu_r^{\mathrm{u}}(t)_\star : \mathcal{A}_\star^{r]} \to \mathcal{A}_\star^{t]}$ are obviously completely positive (CP) and contractive as preadjoint to the compositions $\epsilon_{\mathcal{A}}^{r]} \circ \alpha_r^{\mathrm{u}}(t)$ of the isometries $\alpha_r^{\mathrm{u}}(t)$ and the CP projections $\epsilon_{\mathcal{A}}^{r]} = \mathrm{Q}_r^{\star}$, and they are conservative, preserving the normalization of the states $\breve{\mathrm{R}}_{[0}^t$ if $\alpha_r^{\mathrm{u}}(t, \mathrm{I}_t) = \mathrm{I}_r$ for all t. Due to the consistency condition (40) these compositions satisfy the *modularity* property

$$\mu_r^{\mathrm{u}}(t, \mathrm{XY}) = \mathrm{X}\mu_r^{\mathrm{u}}(t, \mathrm{Y}) \quad \forall \mathrm{X} \in \mathcal{A}^r, \mathrm{Y} \in \mathcal{A}^{t]} \tag{46}$$

implying the commutativity of $\mu_r^{\mathrm{u}}(t, \mathrm{Y})$ with respect to the central Abelian subalgebras $\mathbb{C}_{\mathrm{u}}^r \subseteq \mathbb{C}_{\mathrm{x}}^r$ in $\mathcal{A}^r$ generated by the feedback controls $\mathrm{u}(t)$. This is in the correspondence with (41) for $\alpha_r = \alpha_r^{\mathrm{u}}$ as the orthogonal decomposition

$$\mu_\star^{\mathrm{u},s}\left(t, \breve{\mathrm{R}}_{[0}^t\right) = \int_{\Upsilon^t} \mu_\star^{u,s}\left(t, \mathrm{P}_\upsilon^t\right) \mathrm{E}(\mathrm{d}\upsilon)\, \mu_\star^{\mathrm{u},s}\left(t, \mathrm{P}_{\mathrm{x}}^t\right) \tag{47}$$

of the $\mathcal{A}^t$-modular $\mu^s_\star(t) = \mu^{\mathrm{u},s}_\star(t)$ for any $\check{\mathtt{R}}^t_{[0} = \mathrm{P}^t_{\mathrm{x}} \in \mathcal{A}^{t]}_\star$ consistent with the central decompositions $\mathcal{A}^{t]} = \int_{\Upsilon^t} \mathcal{A}^\upsilon_t \mathrm{E}(\mathrm{d}\upsilon)$ of $\mathcal{A}^{t]}$ relatively to $\mathbb{C}^t_{\mathrm{x}} \subseteq C(\mathcal{A}^t)$ decomposed into $\mathbb{C}^t_\upsilon = \mathbb{C}\mathtt{I}^{\mathrm{o}}$.

Thus, the expected semi-Markov quantum dynamics is described by a forward hemigroup $\mu^{\mathrm{u},s}_\star(t)$ of normal contractive CP maps $\mathcal{A}^{t]}_\star \to \mathcal{A}^{t+s]}_\star$ satisfying the decomposability condition (47). If all $\mu^{\mathrm{u}}_t(t)$ are unital, it is Markov for the increasing family $\mathcal{A}^{t]}_\star$, however it is not Markov on the object state space $\mathbb{A}^{\mathrm{o}}_\star$ unless the CP maps $\tau_r(t) = \mu^{\mathrm{u}}_r(t)\,|\mathbb{A}^{\mathrm{o}}$ leave the algebra $\mathbb{A}^{\mathrm{o}}$ invariant, as it is in the case of uncontrolled dynamics $\mu_r(t) = \mathrm{Q}^\star_r \circ \alpha_r(t)$ described by the generator $\kappa(\mathtt{X}) = (\mathbf{K}\mathtt{X}\mathbf{K}^\star)^-_+ \in \mathbb{A}^{\mathrm{o}}$ on a dense domain of $\mathtt{X} \in \mathbb{A}^{\mathrm{o}}$ as

$$\kappa(\mathtt{X}) = \mathtt{K}_+\mathtt{X} + \sum_{j\in J_\bullet} \mathtt{K}_j\mathtt{X}\mathtt{K}^\dagger_j + \mathtt{X}\mathtt{K}^\dagger_+.$$

In this case the preadjoint maps $\tau^s_\star(t) = \tau_t(t^s)_\star$ can be found on $\mathbb{A}^{\mathrm{o}}_\star$ for any $t > r$ as the resolving maps $\varrho_{t+s} = \tau^s_\star(t, \varrho_t)$ for the coherent controlled *Lindblad master equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\varrho_t + \mathtt{K}_-\varrho_t + \varrho_t\mathtt{K}^\dagger_- = \sum_{j\in J_\bullet} \mathtt{K}^j\varrho_t\mathtt{K}^{j\dagger}, \quad \varrho_0 = \varrho \in \mathbb{A}^{\mathrm{o}}_\star. \tag{48}$$

Here the $\mathbb{A}^{\mathrm{o}}$-valued operators $\mathtt{K}_- = -\mathtt{K}^\sharp_+$ and $\mathtt{K}^j = \mathtt{K}^\sharp_j$ are defined by left adjoints $\left(\mathtt{K}^\sharp\varrho \,|\, \mathtt{X}\right) = (\varrho \,|\, \mathtt{K}\mathtt{X})$ with respect to a Hermitian pairing $(\mathtt{X} \,|\, \mathtt{B}) = \left\langle \mathtt{B}^\dagger, \mathtt{X}\right\rangle$ on $\mathbb{A}^{\mathrm{o}}$. (In the case of the trace pairing $(\varrho \,|\, \mathtt{X}) = \mathrm{tr}\left[\varrho^\dagger\mathtt{X}\right]$ the left adjoint operator $\mathtt{K}^\sharp$ is simply the Hilbert space adjoints, $\mathtt{K}^\sharp = \mathtt{K}^\dagger$, see the Appendix A).

In order to find the stochastic generators defining the stochastic master equation for the semi-Markov dynamics $\mu^{\mathrm{u}}_r(t)_\star$ we have to find first the adjoint nonstochastic equation for $\mu^{\mathrm{u}}_r(t)$ not on the object algebra $\mathbb{A}^{\mathrm{o}}$ but on the increasing algebras $\mathcal{A}^{t]} \simeq \mathcal{A}^t \bar{\otimes} \mathbb{A}^{\mathrm{o}}$. The dynamical maps $\mu^{\mathrm{u}}_t(t^s) = \mathrm{Q}^\star_t \circ \alpha^{\mathrm{u}}_t(t^s)$ describe the expected interaction evolution $\mu^{\mathrm{u}}_r(t, \mathrm{Y}_t)$ of any adapted *QS-semimartingale* Y_t by the general *master equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^{\mathrm{u}}_r(t, \mathrm{Y}_t) = \mu^{\mathrm{u}}_r\left(t, \kappa^{\mathrm{u}}\left(t, \mathbf{J}^t_{\mathrm{Y}}\right)\right), \quad \mathrm{Y}_t = \mathcal{A}^{t]}, \tag{49}$$

where the generator $\kappa^{\mathrm{u}}(\mathbf{J}) = (\mathbf{K}^{\mathrm{u}}\mathbf{J}\mathbf{K}^{\mathrm{u}\star})^-_+$ on the germ $\mathbf{J}(t) = \mathrm{Y}_t + \mathbf{D}^t_{\mathrm{Y}} \equiv \mathbf{J}^t_{\mathrm{Y}}$ is given by $\mathbf{K}^{\mathrm{u}}(t) = \mathbf{S}^{\mathrm{u}}(t) - \mathbf{I}(t)$ with $\mathbf{S}^{\mathrm{u}}(t) = \mathring{\mathrm{U}}^{\mathrm{u}}_t(t^+)$. This simply follows from

$$\mathrm{Q}^\star_t\left[\mathrm{dA}^\nu_\mu\right] = \left(\mathrm{I}^{t]} \otimes \delta^\dagger_\emptyset\right)\left(\mathrm{I}^{t]} \otimes \mathrm{dA}^\nu_\mu\right)\left(\mathrm{I}^{t]} \otimes \delta_\emptyset\right) = \delta^\nu_+\delta^-_\mu\mathrm{I}^{t]}\mathrm{d}t$$

by applying the vacuum conditional expectation $\epsilon_{\mathcal{A}}^{t]} = \mathrm{Q}_t^\star$ to (37). Evaluation of the matrix element $(\mathbf{K}^{\mathrm{u}}\mathbf{J}\mathbf{K}^{\mathrm{u}\star})_+^-$ yields a conditionally-CP map

$$\kappa^{\mathrm{u}}(\mathbf{J}) = \mathrm{J}_+^- + \mathrm{K}_+^{\mathrm{u}}\mathrm{Y} + \mathrm{Y}\mathrm{K}_+^{\mathrm{u}\dagger} + \mathrm{K}_\bullet^{\mathrm{u}}\mathrm{J}_+^\bullet + \mathrm{K}_\bullet^{\mathrm{u}}\mathrm{J}_\bullet^\bullet\mathrm{K}_\bullet^{\mathrm{u}\dagger} + \mathrm{J}_\bullet\mathrm{K}_\bullet^{\mathrm{u}\dagger} \tag{50}$$

for each t. In particular, in the case $\mathrm{J}_\nu^\mu(t) = \mathrm{Y}_r\delta_\nu^\mu$ corresponding to a constant $\mathrm{Y}_t = \mathrm{Y}_r \in \mathcal{A}_0^{r]}$ for all $t \geq r$, this reads as the Lindblad controlled generator[48,57] which is normally written in the decomposed form as

$$\kappa^{\mathrm{u}}(\mathrm{Y}) = \frac{\mathrm{i}}{\hbar}\left[\mathrm{H}^{\mathrm{u}}(t), \mathrm{Y}\right] + \Lambda_{\mathrm{u}}^\star(\mathrm{Y}) + \delta^{\mathrm{u}}(\mathrm{Y}). \tag{51}$$

Here $\mathrm{H}^{\mathrm{u}} = \hbar\,\mathsf{Im}\left[\mathrm{K}_+^{\mathrm{u}}\right]$ is the controlled Hamiltonian, the quantum diffusion part, given by

$$\Lambda_{\mathrm{u}}^\star(\mathrm{Y}) = \frac{1}{2}\left(\mathrm{S}_+^{\mathrm{u}\dagger}\left[\mathrm{Y}, \mathrm{S}_+^{\mathrm{u}}\right] + [\mathrm{S}_+^{\mathrm{u}\dagger}, \mathrm{Y}]\mathrm{S}_+^{\mathrm{u}}\right), \tag{52}$$

is uncontrolled, and $\delta^{\mathrm{u}}(\mathrm{Y}) = \mathrm{K}_\bullet^{\mathrm{u}}\mathrm{Y}\mathrm{K}_\bullet^{\mathrm{u}\dagger} - \mathrm{S}_+^{\mathrm{u}\dagger}\mathrm{Y}\mathrm{S}_+^{\mathrm{u}}$ is the quantum jump part controlled by $\mathrm{K}_\bullet^{\mathrm{u}}$ which can be written using isometricity condition $\mathrm{K}_\bullet^{\mathrm{u}} + \mathrm{S}_+^{\mathrm{u}\dagger}\mathrm{S}_\bullet^{\mathrm{u}} = 0$ as

$$\delta^{\mathrm{u}}(t, \mathrm{Y}) = \mathrm{S}_+^{\mathrm{u}}(t)^\dagger\left(\mathrm{S}_\bullet^{\mathrm{u}}(t)\,\mathrm{Y}_t\mathrm{S}_\bullet^{\mathrm{u}}(t)^\dagger - \mathrm{Y}_t\right)\mathrm{S}_+^{\mathrm{u}}(t). \tag{53}$$

We note that κ^{u} is decomposable,

$$\kappa^{\mathrm{u}}(t, \mathbf{J}_t) = \int_{\Upsilon^t}\kappa^u(t, \mathbf{J}_t^\upsilon)\,e_{\mathrm{u}}(\mathrm{d}\upsilon) \equiv \kappa^{\mathrm{u}}(t, \mathbf{J}_t^{\mathrm{u}})$$

due to the modularity (46), which is obviously satisfied if the uncontrolled generator $\kappa(\mathtt{Y}) = (\mathbf{K}\mathtt{Y}\mathbf{K}^\star)_+^-$ is given by the coefficients $\mathrm{K}_\iota^- \simeq \mathtt{K}_\iota^-$ as the $\mathbb{A}^{\mathrm{o}}$-valued functions $\mathtt{S}_\iota^-(t)$, $\iota = +, j \in J_\bullet$ on Υ^t. Thus, the contractive conditionally CP maps $\kappa^{\mathrm{u}}(t)$ defines a semi-Markov evolution on the increasing family of W*-algebras $\mathcal{A}^{t]}$ which is Markov in the conservative case corresponding to $\kappa^{\mathrm{u}}(t, \mathrm{I}) = \mathrm{O}$. It is usually not Markov with respect to $\mathbb{A}^{\mathrm{o}}$, but it is conditionally Markov on $\mathbb{A}^{\mathrm{o}}$ with respect to the central subalgebras $\mathbb{C}_{\mathrm{u}}^t \subseteq C(\mathcal{A}^t)$ if conditioned by the controlling stochastic process u generating these $\mathbb{C}_{\mathrm{u}}^t$. Note that in the case of the trivial center $C(\mathcal{A}^t)$ there is no feedback, $\mathbb{C}_{\mathrm{u}} = \mathbb{C} = \mathbb{C}_{\mathrm{x}}$, the controlled dynamics is autonomous on $\mathbb{A}^{\mathrm{o}}$ as open controlled by a (deterministic) process $u(t)$, corresponding to the sub-Markovian (or Markovian if unital) case $\alpha_t^u(t^s, \mathbb{A}^{\mathrm{o}}) \subset \mathcal{N}_{[t}^s$ given by $\mathbf{K}^u(t) \vdash \mathbb{A}^{\mathrm{o}}\mathbf{I}(t)$.

From now on we will consider quantum semi-Markov dynamics under time continuous measurements of the output processes $\mathtt{y}_j^\varepsilon(t) = \hat{\mathrm{v}}_j^\varepsilon(t)$ representing the commuting processes (22) on the vacuum after the controlled

QS interaction with the object by isometric evolutions $\mathrm{V}_0(t) = \mathrm{U}_0(t)\,\mathrm{U}^t_{\xi^{\mathrm{v}}_0}$. They are intertwined by the isometric evolution $\mathrm{U}_0(t)$ with $\mathrm{x}^\varepsilon_j(t) = \bar{\mathtt{v}}^\varepsilon_j(t)$, where $\mathtt{v}^\varepsilon_j(t) = \mathrm{U}^t_{\hat{\zeta}_0}\,\hat{\mathrm{v}}^\varepsilon_j(t)\,\mathrm{U}^{t\dagger}_{\hat{\zeta}_0}$, and with $\mathrm{v}^\varepsilon_j(t)$ by $\mathrm{V}_0(t)$:

$$\mathtt{y}^\varepsilon_j(t)\,\mathrm{V}_0(t) = \mathrm{V}_0(t)\,\mathrm{v}^\varepsilon_j(t) = \hat{\mathrm{v}}^\varepsilon_j(t)\,\mathrm{V}_0(t)\,, \tag{54}$$

uniquely determining $\mathtt{y}^\varepsilon_j(t)$ in the unitary case $\mathrm{V}_0\mathrm{V}_0^\dagger = \mathtt{I}_0$ as

$$\mathtt{y}^\varepsilon_j(t) = \beta_0\left(t^s, \mathrm{v}^\varepsilon_j(t)\right) = \alpha^{\mathrm{u}}_0\left(t^s, \mathrm{x}^\varepsilon_j(t)\right) \tag{55}$$

for any $s \geq 0$.

Applying the quantum Itô formula to the products $\hat{\mathrm{v}}^\varepsilon_j\mathrm{V}_0 = \mathrm{V}_0\mathrm{v}^\varepsilon_j$ defined by the QS derivatives $\mathbf{b}^\varepsilon_j$ and $\mathrm{V}_0\mathbf{L} = \mathrm{V}_0(\mathbf{T}-\mathbf{I})$ respectively in (23) and in (43), one can find that the output processes $\mathtt{y}^\varepsilon_j(t)$ can be defined as the QS integrals $\mathfrak{i}^t_0\left(\hat{\mathbf{D}}^\varepsilon_j\right) = \int_0^t \mathrm{d}\mathtt{y}^\varepsilon_j$ of

$$\mathrm{d}\mathtt{y}^\varepsilon_j = \left(\hat{\mathrm{L}}_j + \varepsilon_j\hat{\mathrm{L}}_j\hat{\mathrm{L}}^\dagger_j + \hat{\mathrm{L}}^\dagger_j\right)\mathrm{d}t + 2\,\mathsf{Re}\left(\mathrm{d}\mathtt{B}^+_j\hat{\mathrm{C}}^\dagger_j\right) + \varepsilon_j\mathrm{d}\mathtt{n}_j$$

Here $\mathtt{n}_j = \hat{\mathrm{T}}^{\circ\dagger}_j\left(\mathtt{I}^{\mathrm{o}}\otimes\mathrm{A}^\circ_\circ\right)\hat{\mathrm{T}}^\circ_j \equiv \mathtt{B}^j_j$, $\mathtt{B}^+_j = \left(\mathtt{I}^{\mathrm{o}}\otimes\mathrm{A}^+_\circ\right)\hat{\mathrm{T}}^\circ_j$ and $\mathtt{X} = \hat{\mathrm{L}}_j$, $\hat{\mathrm{C}}_j, \hat{\mathrm{T}}^\circ_j$ are the representations intertwined with $\mathrm{X} = \mathrm{L}_j$, $\mathrm{C}_j = \mathrm{I}_t + \varepsilon\mathrm{L}_j$ and $\mathrm{T}^\circ_j = \mathrm{I}^\circ_j + \mathrm{L}^\circ_j$, which are uniquely defined in the unitary case by (42).

Indeed, the equation (54) can be written on the differential level as $\hat{\mathbf{D}}^\varepsilon_j\mathbf{H}_0 = \mathbf{H}_0\mathbf{b}^\varepsilon_j$, where $\mathbf{H}_0 = \mathrm{V}_0\mathbf{T}$ is the germ of V_0 and $\hat{\mathbf{D}}^\varepsilon_j$ are the QS derivatives defining $\mathrm{d}\mathtt{y}^\varepsilon_j = \hat{\mathbf{D}}^\varepsilon_j\cdot\mathrm{d}\mathbf{A}$. It is equivalent to $\mathbf{D}^\varepsilon_j\mathbf{T} = \mathbf{T}\mathbf{b}^\varepsilon_j$ for

$$\mathbf{D}^\varepsilon_j = \begin{bmatrix} \mathrm{T}^{\circ\dagger}_j\left(\mathrm{I}_t + \varepsilon_j\mathrm{L}_j\right) & \mathrm{L}_j + \varepsilon\mathrm{L}_j\mathrm{L}^\dagger_j + \mathrm{L}^\dagger_j \\ \mathrm{T}^{\circ\dagger}_j\varepsilon_j\mathrm{T}^\circ_j & \left(\mathrm{I}_t + \varepsilon\mathrm{L}_j\right)^\dagger\mathrm{T}^\circ_j \end{bmatrix} = \mathbf{T}\begin{bmatrix} \delta^\bullet_j{}^\dagger & 0 \\ \delta^\bullet_j\varepsilon_j\delta^\bullet_j{}^\dagger & \delta^\bullet_j \end{bmatrix}\mathbf{T}^\star$$

having the elements D^μ_ν intertwined $\mathrm{V}_0\mathrm{D}^\mu_\nu = \hat{\mathrm{D}}^\mu_\nu\mathrm{V}_0$ with the elements $\hat{\mathrm{D}}^\mu_\nu$ of $\hat{\mathbf{D}}^\varepsilon_j$. This can be verified straightforward for $\mathbf{D}^\varepsilon_j = \mathbf{T}\mathbf{b}^\varepsilon_j\mathbf{T}^\star$ by taking into account that $\mathbf{T}^\star\mathbf{T} = \mathbf{I}$.

Thus, the output processes corresponding to the standard Wiener processes $\mathrm{w}_j = \mathrm{U}_{\xi^{\mathrm{v}}_\bullet}\mathrm{v}^0_j\mathrm{U}^\dagger_{\xi^{\mathrm{v}}_\bullet}$, given for $j \in \boldsymbol{J}_\bullet$ in (14) for $\varepsilon_j = 0$, realize after the interaction a simultaneous indirect measurement of the evolved generalized coordinates $2\,\mathsf{Re}\left[\mathrm{L}^\dagger_j\right]$:

$$\mathrm{d}\mathtt{y}^0_j = 2\,\mathsf{Re}\left(\hat{\mathrm{L}}^\dagger_j(t)\,\mathrm{d}t + \mathrm{d}\mathtt{B}^+_j\right) = \left(\hat{\mathrm{L}}_j + \hat{\mathrm{L}}^\dagger_j\right)(t)\,\mathrm{d}t + \mathrm{d}\mathtt{v}_j, \tag{56}$$

where $\mathtt{v}_j = 2\,\mathsf{Re}\,\mathtt{B}^+_j$, $\hat{\mathrm{L}}_j = \widehat{\mathrm{K}}_j\mathrm{e}^{\mathrm{i}\widehat{\theta}_j} - \left|\widehat{\xi}_j\right|$ with $\widehat{\mathrm{K}}_j\mathrm{V}_0 = \mathrm{V}_0\mathrm{K}^{\mathrm{y}}_j$ and similar for $\widehat{\xi}_j(t) = \xi^{\mathrm{y}}_j(t)$ intertwined with $\xi^{\mathrm{v}}_j(t)$ by $\mathrm{V}_0(t)$. Similarly, the output processes corresponding to the compensated Poisson processes $\mathrm{m}_j = \mathrm{U}_{\xi^{\mathrm{v}}_\bullet}\mathrm{v}^1_j\mathrm{U}^\dagger_{\xi^{\mathrm{v}}_\bullet}$

scaled by $\left|\xi_j\right| = 1/\varepsilon_j$ for $j \in \boldsymbol{J}_\bullet$ in (14) realize after the interaction a simultaneous indirect measurement of the evolved $\varepsilon_j^{-1}\left(\mathrm{C}_j^\varepsilon \mathrm{C}_j^{\varepsilon\dagger} - \mathrm{I}\right)$:

$$\mathrm{d}\mathtt{y}_j^\varepsilon = \varepsilon_j^{-1}\left(\hat{\mathrm{C}}_j^\varepsilon \hat{\mathrm{C}}_j^{\varepsilon\dagger} - \mathtt{I}\right)\mathrm{d}t + 2\,\mathtt{Re}\left(\mathrm{d}\mathtt{B}_j^+ \hat{\mathrm{C}}_j^{\varepsilon\dagger}\right) + \varepsilon_j \mathrm{d}\mathtt{n}_j, \tag{57}$$

where $\hat{\mathrm{C}}_j^\varepsilon = \mathtt{I} + \varepsilon_j \hat{\mathrm{L}}_j$ is given by the row $\varepsilon_\bullet^{-1} = \left(\left|\xi_j\right|\right)_{j\in J_\bullet}$ of $\xi_j \in L^2_{\mathrm{loc}}$ at each t.

In order to obtain the form of the coherently controlled generators on the algebras $\mathcal{A}^{t]} = \mathbb{A}_{\mathrm{v}}^t$ from the generator of the uncontrolled master equation (48) for these particular cases of observation represented with respect to the vacuum by noise $\mathtt{v}_\bullet^\varepsilon$, let us find the transformation of the generator (50) from the algebras $\mathcal{A}^{t]} = \mathbb{A}_{\mathrm{x}}^t$ given in the reference of the locally coherent states by the amplitudes $\xi_\bullet^{\mathrm{v}}$.

Theorem 3.1. *Let* $\mathrm{R}_t = \mathrm{P}_{\widehat{\zeta}_t^{\mathrm{u}}}$ *be the initial state on* $\mathcal{N}_t$ *locally induced by the coherent vector* $f^s_{\widehat{\zeta}_t^{\mathrm{u}}}$ *given by the stochastic row-amplitudes* $\zeta_j^{\mathrm{u}^t}(t)$ *controlled by an adapted process* $u(t)$ *on the reference probability space* $(\Upsilon^t, \mathfrak{B}^t, \mathsf{Q}^t)$ *with the probability measure* Q *induced by the coherent states* $f^t_{\xi_0^{\mathrm{x}}}$ *given by row-amplitudes* $\xi_\bullet^x$. *Then there exists a conditional expectation* $\hat{\epsilon}_{\mathcal{A}}^{t]} = \mathrm{P}^{s\star}_{\widehat{\zeta}_t^{\mathrm{u}}}$ *as normal projection onto* $\mathcal{A}^{t]}$ *such that the compositions* $\hat{\epsilon}_{\mathcal{A}}^{t]} \circ \beta_t^-(t^s)$ *form a normal hemigroup of unital CP maps coinciding with the modular with respect to the central W*-algebras* $\mathbb{A}_{\mathrm{y}}^t = \mathrm{U}^{t\dagger}_{\xi_0^{\mathrm{v}}} \mathbb{A}_{\mathrm{u}}^t \mathrm{U}^t_{\xi_0^{\mathrm{v}}}$ *maps* $\phi_t(t^s) = \mathsf{Q}_t^\star \circ \beta_t(t^s)$ *onto* $\mathcal{A}^{t]}$. *It satisfies the forward evolution equation similar to (49), given in terms of the germ* $\mathbf{J}_t = \mathring{\mathrm{X}}_{t^+}$ *by the controlled generator*

$$\lambda(\mathbf{J}) = \mathrm{L}_\bullet \mathrm{J}_+^\bullet + \mathrm{L}_\bullet\left(\mathrm{J}_\bullet^\bullet - \mathrm{X}\delta_\bullet^\bullet\right)\mathrm{L}_\bullet^\dagger + \mathrm{J}_\bullet^- \mathrm{L}_\bullet^\dagger + \lambda(\mathrm{X}), \tag{58}$$

in the case $\mathrm{J}_+^- = 0$ *corresponding to a vacuum martingale* $\mathrm{X}_t \in \mathcal{A}^{t]}$). *Here*

$$\mathrm{L}_j(t, \mathrm{v}) = \left(\mathrm{K}_j^{\mathrm{y}} - \xi_j^{\mathrm{v}}\mathrm{I}\right)\mathrm{e}^{\mathrm{i}\theta_\bullet^{\mathrm{v}}}, \quad j \in J_\bullet \tag{59}$$

$$\mathrm{L}_+(t, \mathrm{v}) = \mathrm{K}_+^{\mathrm{y}} + \mathrm{K}_\bullet^{\mathrm{y}} \xi_\bullet^{\mathrm{v}\dagger} - \frac{1}{2}\left\|\xi_\bullet^{\mathrm{v}}\right\|^2_{\mathfrak{k}_\bullet}\mathrm{I} \tag{60}$$

determines the Lindbladian $\lambda(t, \mathrm{X}) = \kappa^{\mathrm{y}}(t, \mathrm{X})$,

$$\lambda(t, \mathrm{X}) = \sum_{j\in J_\bullet} \mathrm{L}_j(t)\,\mathrm{X}\mathrm{L}_j(t)^\dagger + \mathrm{L}_+(t)\,\mathrm{X} + \mathrm{X}\mathrm{L}_+(t)^\dagger, \tag{61}$$

coinciding with (51) determined by

$$\mathrm{K}_j^{\mathrm{y}} = \mathrm{K}_j + \xi_\circ^{\mathrm{y}}\mathrm{S}_j, \qquad j \in J_\bullet \tag{62}$$

$$\mathrm{K}_+^{\mathrm{y}} = \mathrm{K}_+ + \xi_\circ^{\mathrm{y}}\mathrm{S}_+ - \left\|\xi_\circ^{\mathrm{y}}\right\|^2_{\mathfrak{k}}\mathrm{I}/2. \tag{63}$$

Proof. Indeed, the conditional expectation $\mathrm{P}^{s\star}_{\widehat{\zeta}^{\mathrm{u}}_t}$ exists due to the unitary equivalence $f^{s}_{\widehat{\zeta}^{\mathrm{u}}_t} = \mathrm{U}^{t}_{\widehat{\zeta}^{\mathrm{u}}_r}\delta_{\emptyset}$ under the Weyl transformation $\mathrm{U}^{s}_{\widehat{\zeta}^{\mathrm{u}}_t}$ with the covariant transformation of the dynamics given by $\mathrm{V}^{-}_{r}(t) = \mathrm{U}^{t}_{\widehat{\zeta}^{\mathrm{u}}_r}\mathrm{V}_r(t)$. Applying the vacuum conditional expectation to $\beta_t(t^s)$ which has the germ $\mathring{\beta}_t(t^+, \mathrm{X}_t) = \mathbf{T}(t)\,\mathring{\mathrm{X}}_{t^+}\mathbf{T}(t)^{\star}$ given by $\mathbf{L}(t) = \mathbf{T}(t) - \mathbf{I}(t)$ in (43), we immediately obtain (58) with (61), where $\mathrm{K}^{\mathrm{y}}_j(t)$, $\mathrm{K}^{\mathrm{y}}_+(t)$ are defined by the germ $\mathbf{R}(t) = \mathring{\mathrm{V}}^{-}_{t}(t^+)$ as the first row in (44). □

Note that the Lindblad generator λ for the coherently controlled (62), (63), decomposed as (51) can also be decomposed in the same way in terms of L_ι instead of $\mathrm{K}^{\mathrm{y}}_\iota$ and T_ι instead of $\mathrm{S}^{\mathrm{y}}_\iota$, $\iota = \bullet, +$, with $\mathrm{H} = \hbar\,\mathsf{Im}\,[\mathrm{L}_+]$, Λ defined by T_+ instead of $\mathrm{S}^{\mathrm{y}}_+$ and

$$\delta(\mathrm{X}) = \mathrm{T}^{\dagger}_{+}\left(\mathrm{T}_{\bullet}\mathrm{X}\mathrm{T}^{\dagger}_{\bullet} - \mathrm{X}\right)\mathrm{T}_{+}. \tag{64}$$

3.2. *Stochastic Master and filtering equations*

The quantum dynamics with observation in general is not autonomous on the object state space $\mathbb{A}^{\mathrm{o}}_{\star}$ but can be described on the increasing family of the preadjoint spaces $\mathcal{M}^{t]}_{\star} = \mathcal{M}^{t}_{\star}\bar{\otimes}\mathbb{A}^{\mathrm{o}}_{\star}$ to $\mathcal{M}^{t]} = \mathbb{A}^{t}_{\mathrm{x}}$ by stochastic maps $\vartheta^{t}_{\mathrm{x}} : \mathbb{A}^{\mathrm{o}}_{\mathrm{x}\star} \to \mathbb{A}^{t}_{\mathrm{x}\star}$ as preadjoints to the restrictions of $\mu(t)$ to $\mathcal{M}^{t} \subseteq \mathcal{A}^{t]}$. These stochastic maps will be defined as resolving for a *linear stochastic master equation* determining the states $\varpi^{t}_{\mathrm{x}} \in \mathcal{M}^{t}_{\star}$ on $\mathbb{A}^{t}_{\mathrm{x}}$ obtained from $\tilde{\mathrm{R}}^{t}_{[0} = \mathrm{P}^{t}_{\mathrm{x}} \in \mathcal{A}^{t]}_{\star}$ by the marginalizations $\iota^{t}_{\star}(\mathrm{P}^{t}_{\mathrm{x}})$ as preadjoint to the injections $\iota^{t} : \mathbb{A}^{t}_{\mathrm{x}} \subseteq \mathcal{A}^{t]}$.

Let us first derive the stochastic equation for the density operator ϖ^{t}_{v} describing the states on the central subalgebras $\mathbb{C}^{t}_{\mathrm{v}} \subseteq \mathcal{B}^{t]}$ generated by the canonical Fock space representation $\left\{\mathrm{v}^{\varepsilon}_{j}(r)\,|r \leq t\right\}$ of the commuting processes $\mathrm{v}^{\varepsilon}_{j}(t)$ with respect to the vacuum-induced reference state. It will determine ϖ^{t}_{v} from the initial $\varrho = \varpi^{0}_{\mathrm{v}} \in \mathbb{A}^{\mathrm{o}}_{\star}$ by the map $\vartheta^{t}_{\mathrm{v}} = \vartheta^{t}_{\mathrm{v}}(0)$ having the adjoint

$$\vartheta^{t}_{\mathrm{v}}(0)^{\star} = \phi_0(t)\,|\mathbb{A}^{t}_{\mathrm{v}} \equiv \vartheta^{\star}_{\mathrm{v}}(t)$$

on $\mathbb{A}^{t}_{\mathrm{v}} \simeq \mathbb{C}^{t}_{\mathrm{v}}\bar{\otimes}\mathbb{A}^{\mathrm{o}}$. To this end we apply the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\vartheta^{\star}_{\mathrm{v}}(t, \mathrm{Z}_t) = \vartheta^{\star}_{\mathrm{v}}(t, \lambda(t, \mathbf{J}_t))\,,\quad \mathrm{Z}_0 \simeq \mathtt{X} \in \mathbb{A}^{\mathrm{o}}$$

to the vacuum martingales $\mathrm{Z}_t = \mathrm{Z}_0\mathrm{V}^{t}_{c} \equiv \mathtt{X}^{\mathrm{y}}_{t}$ generating the algebras $\mathbb{A}^{t}_{\mathrm{v}}$, where V^{t}_{c} is Wick exponent of $x^{\varepsilon}_{c}(t)$ defined by the equation $\mathrm{d}\mathrm{V}^{t}_{c} = \mathrm{V}^{t}_{c}c^{j}\mathrm{d}\mathrm{v}^{\varepsilon}_{j}$

as in (32). Taking into account that the germ $\mathbf{J}_t = \mathrm{Z}_{t^+}$ in this case is given by

$$\mathrm{J}^i_k = \left(\delta^i_k + c^j\delta^i_j\varepsilon_j\delta^j_k\right)\mathrm{Z}_t, \quad \mathrm{J}^-_j = c^j\mathrm{Z}_t = \mathrm{J}^j_+$$

and $\mathrm{J}^-_+ = 0$, we obtain from (58)

$$\lambda(\mathbf{J}) = c^j\left(\mathrm{L}_j\mathrm{Z} + \varepsilon_j\mathrm{L}_j\mathrm{Z}\mathrm{L}^\dagger_j + \mathrm{Z}\mathrm{L}^\dagger_j\right) + \lambda(\mathrm{Z})$$

where $\lambda(t)$ is in (61). Note that, if $\varepsilon_j \neq 0$, the term in the parenthesis can be written simply as $\varepsilon_j^{-1}\left(\mathrm{C}^\varepsilon_j\mathrm{Z}\mathrm{C}^{\varepsilon\dagger}_j - \mathrm{Z}\right)$. Thus,

$$\mathrm{d}\left\langle\varpi^t_{\mathrm{v}}, \mathrm{Z}_t\right\rangle = \left\langle\varrho, \mathrm{d}\vartheta^\star_{\mathrm{v}}(t, \mathrm{Z}_t)\right\rangle = \left\langle\varpi^t_{\mathrm{v}}, \lambda(t, \mathbf{J})\right\rangle \mathrm{d}t.$$

This suggests that ϖ^t_{v} satisfies the linear stochastic master equation

$$\mathrm{d}\varpi^t_{\mathrm{v}} = \lambda_\star\left(\varpi^t_{\mathrm{v}}\right)\mathrm{d}t + \left(\mathrm{L}^j\varpi^t_{\mathrm{v}} + \varepsilon\mathrm{L}^j\varpi^t_{\mathrm{v}}\mathrm{L}^{j\dagger} + \varpi^t_{\mathrm{v}}\mathrm{L}^{j\dagger}\right)\mathrm{d}\mathrm{v}^\varepsilon_j, \tag{65}$$

as a particular semiquantum case of (139) for $\mathrm{R} = \varpi^t_{\mathrm{v}}$, where ε may depend on j and t. Here $\mathrm{L}^j = \mathrm{L}^\sharp_j$ and $\lambda_\star$ is preadjoint to $\lambda(t) = \kappa^{\mathrm{y}}(t)$ written as

$$\lambda_\star\left(t, \varpi^t_{\mathrm{v}}\right) = \sum_{j\in J_\bullet}\mathrm{L}^j(t)\,\varpi^t_{\mathrm{v}}\mathrm{L}^j(t)^\dagger - \mathrm{L}_-(t)\,\varpi^t_{\mathrm{v}} - \varpi^t_{\mathrm{v}}\mathrm{L}_-(t)^\dagger, \tag{66}$$

in terms of $\mathrm{L}_- = -\mathrm{L}^\sharp_+$. ($\mathrm{L}^\sharp$ denotes the left adjoint operator with respect to a standard pairing coinciding with the adjoint $\mathrm{L}^\dagger$ with respect to the trace pairing.) It can be easily verified by applying the classical Itô formula

$$\begin{aligned}\mathrm{d}\left\langle\varpi^t_{\mathrm{v}}, \mathrm{Z}_t\right\rangle &= \left\langle\mathrm{d}\varpi^t_{\mathrm{v}}, \mathrm{Z}_t\right\rangle + \left\langle\mathrm{d}\varpi^t_{\mathrm{v}}, \mathrm{d}\mathrm{Z}_t\right\rangle + \left\langle\varpi^t_{\mathrm{v}}, \mathrm{d}\mathrm{Z}_t\right\rangle \\ &= \left\langle\lambda_\star\left(\varpi^t_{\mathrm{v}}\right), \mathrm{Z}_t\right\rangle\mathrm{d}t + \left\langle\left(\mathrm{L}^j\varpi^t_{\mathrm{v}} + \varepsilon\mathrm{L}^j\varpi^t_{\mathrm{v}}\mathrm{L}^{j\dagger} + \varpi^t_{\mathrm{v}}\mathrm{L}^{j\dagger}\right), \mathrm{Z}_t\right\rangle.\end{aligned}$$

The stochastic master equation (65), derived on the simple algebra $\mathbb{A}^{\mathrm{o}} = \mathcal{B}(H_{\mathrm{o}})$ by Belavkin in,[11,18] can also be written with respect to x^ε_j by replacing v and y respectively for x and u, or as the classical stochastic equation in the spectral representation $x^\varepsilon_j = v^\varepsilon_j$ of x^ε_j (or equivalently, of v^ε_j) by simple substitution $x = v$ and $u = y$ replacing respectively x and u. Thus, the two most important cases $\varepsilon_j = 0$ and $\varepsilon_j = 1/\left|\xi_j\right|$ of the Belavkin master equation (65) can be written in terms of the corresponding Wiener processes $\mathrm{x}_j = \mathrm{w}_j$ and the compensated Poisson processes $\mathrm{x}_j = \mathrm{m}_j$ as

$$\mathrm{d}\varpi^t_{\mathrm{w}} = \lambda^{\mathrm{u}}_\star\left(\varpi^t_{\mathrm{w}}\right)\mathrm{d}t + \left(\mathrm{L}^j\varpi^t_{\mathrm{w}} + \varpi^t_{\mathrm{w}}\mathrm{L}^{j\dagger}\right)\mathrm{d}\mathrm{w}_j, \tag{67}$$

$$\mathrm{d}\varpi^t_{\mathrm{m}} = \lambda^{\mathrm{u}}_\star\left(\varpi^t_{\mathrm{m}}\right)\mathrm{d}t + \left(\mathrm{C}^j\varpi^t_{\mathrm{m}}\mathrm{C}^{j\dagger} + \varpi^t_{\mathrm{m}}\right)\mathrm{d}\mathrm{m}_j. \tag{68}$$

The solutions ϖ_{v}^t of these equations, considered as a classical stochastic $\mathbb{A}_\star^\circ$-valued process ϖ_υ^t on the reference probability space $(\Upsilon, \mathfrak{B}, \mathsf{Q})$ by resolving the corresponding classical stochastic equation

$$\mathrm{d}\varpi_\upsilon^t = \kappa_\star^u\left(\varpi_\upsilon^t\right)\mathrm{d}t + \left(\mathtt{L}^j\varpi_\upsilon^t + \varepsilon_j\mathtt{L}^j\varpi_\upsilon^t\mathtt{L}^{j\dagger} + \varpi_\upsilon^t\mathtt{L}^{j\dagger}\right)(t,\upsilon)\,\mathrm{d}v_j^\varepsilon,$$

for $\varpi_\upsilon^0 = \varrho \in \mathbb{A}_\star^\circ$, do not preserve the normalization $\langle\varrho, \mathtt{I}\rangle$ $(= \mathrm{tr}\{\varrho\}$, say) in $\mathbb{A}_\star^\circ$. However, the stochastic dynamics, described by forward hemigroup of the CP maps $\vartheta_{\upsilon_t}^{u,s}(t) : \varpi_\upsilon^t \mapsto \varpi_\upsilon^{t+s}$ resolving this equation, is contractive in the mean,

$$\left\langle\varpi_x^t, \mathtt{I}\right\rangle_{\mathsf{Q}} := \int\left\langle\varpi_\upsilon^t, \mathtt{I}\right\rangle\mathsf{Q}(\mathrm{d}\upsilon) \leq \left\langle\varpi_x^t, \mathtt{I}\right\rangle_{\mathsf{Q}} \quad \forall r < t$$

in $L^1_{\mathsf{Q}}(\Upsilon \to \mathbb{A}_\star^\circ)$ since by definition of the representation v it coincides with

$$\left\langle\varpi_{\mathrm{x}}^t, \mathrm{I}\right\rangle_{\mathbb{A}_{\mathrm{x}}^t} = \left\langle\mathtt{R}_{[0}, \mathtt{F}(t)\right\rangle_{[0} \leq \left\langle\varpi_{\mathrm{x}}^r, \mathrm{I}\right\rangle,$$

where $\mathtt{F}(t) = \alpha(t, \mathrm{I}^t)$ is a decreasing orthoprojector of quantum survival. It preserves the mean normalization of ϖ_υ^t only in the conservative case of nondemolition observation corresponding to $\mathtt{F}(t) = \mathtt{I}_{[0}$. The positive function $p_\upsilon^t = \langle\varpi_\upsilon^t, \mathtt{I}\rangle$ gives the density of the *output probability*

$$\mathsf{P}^t(\Delta) = \left\langle\mathtt{R}_{[0}, \mathtt{I}(t,\Delta)\right\rangle_{[0} = \left\langle\varpi_{\mathrm{x}}^t, \mathrm{E}(\Delta)\right\rangle = \int_\Delta p_\upsilon^t\mathsf{Q}(\mathrm{d}\upsilon)$$

of the events $\mathtt{I}(t,\Delta) = \alpha(t, \mathrm{E}(\Delta))$, where $\mathrm{E}(\Delta) = e_{\mathrm{x}}(\Delta)\,\mathrm{I}^{\mathrm{o}}$ on $\mathfrak{B}^t$, normalized for each t to the probability of nondemolition $\mathsf{P}^t(\Upsilon^t) = \left\langle\mathtt{R}_{[0}, \mathtt{F}(t)\right\rangle_{[0}$ of quantum object observed up to time t. It obviously satisfies the stochastic equation

$$\mathrm{d}p_\upsilon^t = \left\langle\varpi_\upsilon^t, \delta(t,\upsilon,\mathtt{I})\right\rangle\mathrm{d}t + \sum_j\left\langle\varpi_\upsilon^t, \mathtt{L}_j + \mathtt{L}_j\varepsilon_j\mathtt{L}_j^\dagger + \mathtt{L}_j^\dagger\right\rangle(t,\upsilon)\,\mathrm{d}v_j^\varepsilon$$

with $\delta(\mathtt{I}) = \mathtt{T}_+^\dagger\left(\mathtt{T}_\bullet\mathtt{T}_\bullet^\dagger - \mathtt{I}_\circ^\circ\right)\mathtt{T}_+ \leq 0$ coinciding for the coherent control with the *demping operator*

$$\delta^u(t,\mathtt{I}) = -\mathtt{S}_+(t)^\dagger\,\mathtt{E}_\circ^\circ(t)\,\mathtt{S}_+(t) - \zeta_\circ^u(t)\,\mathtt{E}_\circ^\circ(t)\,\zeta_\circ^u(t)^\dagger, \tag{69}$$

where $\mathtt{E}_\circ^\circ = \mathtt{I}_\circ^\circ - \mathtt{S}_\bullet\mathtt{S}_\bullet^\dagger$ is the orthoprojector of quantum object demolition (or exit) event. This implies that p_υ^t is submartingale, $\mathbb{E}_{\mathsf{Q}}[p_\upsilon^t|\mathfrak{B}^r] \leq p_\upsilon^t$, since

$$\mathbb{E}_{\mathsf{Q}}\left[\mathrm{d}p_\upsilon^t|\mathfrak{B}^t\right] = \left\langle\varpi_\upsilon^t, \delta^u(t,\mathtt{I})\right\rangle\mathrm{d}t \leq 0$$

for every $\varpi_\upsilon^t \geq 0$ due to $\mathtt{S}_\bullet^u\mathtt{S}_\bullet^{u\dagger} \leq \mathtt{I}_\circ^\circ = \delta_\circ^\circ\mathtt{I}$. (It is martingale in the conservative case zero demping $\delta^u(\mathtt{I}) = 0$.)

The *posterior states* on future subalgebras $\mathcal{B}_t^{s]}$ conditioned by a spectral value $\upsilon \in \Upsilon^t$ of the central subalgebras $\mathbb{C}_{\mathrm{v}}^t \subseteq C\left(\mathcal{B}^{t]}\right)$ of $\mathcal{B}^{t]} = \mathbb{C}_{\mathrm{v}}^t \bar{\otimes} \mathbb{A}^{\mathrm{o}}$ as a result of the nondemolition measurement up to time t are described by the posterior expectations

$$\mathbb{E}_t\left[\beta_t\left(t^s, \mathrm{Z}_t^s\right) | \mathfrak{B}^t\right](\upsilon) = \frac{1}{p_\upsilon^t}\left\langle \varpi_{\upsilon,\mathrm{v}}^{t+s}(t), \mathrm{Z}_t^s\right\rangle \equiv \epsilon_{\mathbb{C}_{\mathrm{v}}}^{t]}\left[\mathrm{Z}_t^s\right](\upsilon) \tag{70}$$

They are given in terms of $\mathcal{B}_{t\star}^{s]}$-valued conditional density

$$\varpi_{\upsilon,\mathrm{v}}^{t,s} = \int_{\Upsilon_t} \varpi_{\upsilon,x}^{t,s} e_{\mathrm{v}}(\mathrm{d}x) = \frac{1}{p_\upsilon^t}\varpi_{\upsilon,\mathrm{v}}^{t+s}(t), \ \upsilon \in \Upsilon^t$$

a.s. defined with respect to the reference expectation

$$\left\langle \varpi_{\upsilon,\mathrm{v}}^{t,s}, \mathrm{Z}_t^s\right\rangle = \int_{\Upsilon_t}\left\langle \varpi_{\upsilon,x}^{t,s}, \mathtt{Y}_t^s(x)\right\rangle_{\mathbb{A}^{\mathrm{o}}} \mathsf{Q}(\mathrm{d}x) \equiv \left\langle \varpi_{\upsilon,x}^{t,s}, \mathtt{Y}_t^s(x)\right\rangle_{\mathbb{A}^{\mathrm{o}}}^{\mathsf{Q}}.$$

where $x = \upsilon_t \in \Upsilon_t$ is the future of $\upsilon \in \Upsilon$ corresponding to the split $\upsilon = (\upsilon, x)$. Since $\varpi_{\upsilon,x}^{t+s} = \vartheta_x^s(t, \varpi_\upsilon^t)$, the posterior state on $\mathcal{M}_t^{s]}$ is determined as $\varpi_{\upsilon,\mathrm{v}}^{t,s} = \vartheta_{\mathrm{v}}^s(t, \varrho_\upsilon^t)$ by the object posterior state

$$\varrho_\upsilon^t = \varpi_{\upsilon,\mathrm{v}}^{t,0} = \frac{1}{p_\upsilon^t}\vartheta_\upsilon^t(\varrho) \equiv \varphi_\upsilon^t(\varrho) \tag{71}$$

which is a.s. in $\mathbb{A}_\star^{\mathrm{o}}$ not depending on any future $\mathrm{v} \in \mathcal{B}_t$. Thus, in order to predict the future quantum states $\varpi_{\upsilon,\mathrm{v}}^{t,s}$, it is sufficient to find the *object posterior states* ϱ_υ^t by a nonlinear filtering map φ_υ^t on the initial $\varrho_\upsilon^0 = \varrho$ and then to apply the linear stochastic dynamics $\vartheta_{\mathrm{v}}^s(t)$. Our goal is to derive a stochastic differential equation

$$\mathrm{d}\varrho_\upsilon^t = \widetilde{\lambda_\star}\left(t, \varrho_\upsilon^t\right)\mathrm{d}t + \sum_{j \in J_\bullet} \widetilde{\varkappa}_\varepsilon^j\left(t, \varrho_\upsilon^t\right)\mathrm{d}\widetilde{v}_j^\varepsilon \tag{72}$$

for ϱ_υ^t from (65), and for this purpose we may write the posterior density in any appropriate operator representation, say in the Fock representation $\varrho_{\mathrm{v}}^t = \varpi_{\mathrm{v}}^t / p_{\mathrm{v}}^t$, an eigen or input representation $\varrho_{\mathrm{x}}^t = \varpi_{\mathrm{x}}^t / p_{\mathrm{x}}^t$, or in the output representation $\varrho_{\mathrm{v}}^t = \varpi_{\mathrm{y}}^t / p_{\mathrm{y}}^t$.

The nonlinear filtering equation (72) specified by the coefficients $\widetilde{\varkappa}_\varepsilon^j$ for $\varepsilon_j = 0$ and $\varepsilon_j = 1$ was derived by Belavkin directly by finding the equation for the conditional expectations $\left\langle \varrho_{\mathrm{y}}^t, \mathtt{X}\right\rangle$ on $\mathtt{X} \in \mathbb{A}^{\mathrm{o}}$ with respect to the output processes y_j^ε in.[11,25] It was also obtained in[18] from the linear equation (65) in Fock space by applying the Itô formula

$$\mathrm{d}\left(p_{\mathrm{v}}^t \varrho_{\mathrm{v}}^t\right) = \varrho_{\mathrm{v}}^t \mathrm{d}p_{\mathrm{v}}^t + p_{\mathrm{v}}^t \mathrm{d}\varrho_{\mathrm{v}}^t + \mathrm{d}p_{\mathrm{v}}^t \mathrm{d}\varrho_{\mathrm{v}}^t$$

to $\varpi_{\mathrm{v}}^t = p_{\mathrm{v}}^t \varrho_{\mathrm{v}}^t$. Here we formulate the result in the mixed input x-representation with arbitrary ε_j.

Theorem 3.2. *Let the QS density $\varpi_{\mathrm{x}}^t \in \mathcal{M}_\star^{t]}$ satisfy the universal linear QS equation*

$$\mathrm{d}\varpi_{\mathrm{x}}^t = \lambda_\star\left(\varpi_{\mathrm{x}}^t\right)\mathrm{d}t + \sum_{j\in J_\bullet} \varkappa_\varepsilon^j\left(\varpi_{\mathrm{x}}^t\right)\mathrm{d}\mathrm{x}_j^\varepsilon, \tag{73}$$

where $\varkappa_\varepsilon^j\left(\varpi_{\mathrm{x}}^t\right) = \mathrm{L}^j\varpi_{\mathrm{x}}^t + \varpi_{\mathrm{x}}^t\mathrm{L}^{j\dagger} + \varepsilon\mathrm{L}^j\varpi_{\mathrm{x}}^t\mathrm{L}^{j\dagger}$, $\mathrm{L}^j = \mathrm{L}_j^\sharp$ $(= \mathrm{L}_j^\dagger)$ and $\lambda_\star$ is given in (66). Then the posterior density $\varrho_{\mathrm{x}}^t = \varpi_{\mathrm{x}}^t / \langle \varpi_{\mathrm{x}}^t, \mathtt{I}\rangle$ satisfies the universal nonlinear filtering equation

$$\mathrm{d}\varrho_{\mathrm{x}}^t = \widetilde{\lambda_\star}\left(\varrho_{\mathrm{x}}^t\right)\,\mathrm{d}t + \sum_{j\in J_\bullet} \widetilde{\varkappa}_\varepsilon^j\left(\varrho_{\mathrm{x}}^t\right)\,\mathrm{d}\widetilde{\mathrm{x}}_j^\varepsilon \tag{74}$$

defined by the drift coefficient

$$\widetilde{\lambda_\star}\left(\varrho_{\mathrm{x}}^t\right) = \lambda_\star\left(\varrho_{\mathrm{x}}^t\right) - \left\langle \varrho_{\mathrm{x}}^t, \mathrm{T}_+^\dagger\left(\mathrm{T}_\bullet\mathrm{T}_\bullet^\dagger - \mathrm{I}_\circ^\circ\right)\mathrm{T}_+\right\rangle \varrho_{\mathrm{x}}^t \tag{75}$$

the innovation processes $\widetilde{\mathrm{x}}_j^\varepsilon$ given by the Itô differentials

$$\mathrm{d}\widetilde{\mathrm{x}}_j^\varepsilon = \left[\mathrm{d}\mathrm{x}_j^\varepsilon(t) - \left\langle \varrho_{\mathrm{x}}^t, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle \mathrm{d}t\right],$$

where $\varkappa_\varepsilon^{j\star}(\mathrm{I}) = \mathrm{L}_j + \mathrm{L}_j^\dagger + \varepsilon_j\mathrm{L}_j\mathrm{L}_j^\dagger$, with the fluctuating coefficients

$$\widetilde{\varkappa}_\varepsilon^j\left(\varrho_{\mathrm{x}}^t\right) = \frac{\varkappa_\varepsilon^j\left(\varrho_{\mathrm{x}}^t\right) - \left\langle \varrho_{\mathrm{x}}^t, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle \varrho_{\mathrm{x}}}{1 + \varepsilon_j\left\langle \varrho_{\mathrm{x}}^t, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle}, \tag{76}$$

Proof. From the definition (71) of ϱ_υ^t and $\vartheta_{\upsilon,x}^{t+s}(r) = \vartheta_x^s(t) \circ \vartheta_\upsilon^t$ for $\vartheta_x^s = \vartheta_x^{u,s}$ we have

$$\varphi_{\upsilon,x}^{u,s}\left(t, \varrho_\upsilon^t\right) = \frac{1}{p_{\upsilon,x}^{t+s}}\vartheta_x^{u,s}\left(t, \varpi_\upsilon^t\right) = \frac{p_\upsilon^t}{p_{\upsilon,x}^{t+s}}\vartheta_x^{u,s}\left(t, \varrho_\upsilon^t\right).$$

This can be written in the form of the quasilinear transformation

$$\varphi_{\upsilon,x}^{u,s}\left(t, \varrho_\upsilon^t\right) = \frac{1}{\langle \varrho_\upsilon^t, \mathtt{M}_t^u(t^s)\rangle}\vartheta_x^{u,s}\left(t, \varrho_\upsilon^t\right)$$

in terms of the positive normalizing operator-function $\mathtt{M}_t(t^s) = \mu_t^u\left(t^s, \mathrm{I}_t^{s]}\right)$ defining

$$\frac{p_\upsilon^{t+s}}{p_\upsilon^t} = \frac{1}{p_\upsilon^t}\left\langle \varpi_\upsilon^t, \mu_t^u\left(t^s, \mathrm{I}_t^{s]}\right)\right\rangle = \left\langle \varrho_\upsilon^t, \mathtt{M}_t^u(t^s)\right\rangle.$$

Since $\mu_r^{\mathrm{u}}(t)$ is the identity map on $\mathbb{A}_{\mathrm{x}}^t$ at $r = t$, its preadjoint $\vartheta_{\mathrm{x}}^0(r)$ is also identity, and therefore $\varphi_{\upsilon,x}^{u,0}(t, \varrho_\upsilon^t) = \varrho_\upsilon^t$ as $\mathtt{M}_t^{\mathrm{u}}(t) = \mu_t^{\mathrm{u}}(t, \mathrm{I}_t) = \mathrm{I}_t$. Thus,

one should expect that ϱ_{x}^t has a stochastic differential $\mathrm{d}\varrho_{\mathrm{x}}^t = \varrho_{\mathrm{x,v}}^{t+\mathrm{d}t} - \varrho_{\mathrm{x}}$ of the form (74) with respect to the *innovation processes*

$$\widetilde{\mathrm{x}}_j^\varepsilon(t) = \mathrm{x}_j^\varepsilon(t) - \epsilon_{\mathcal{M}}^{t]}\left[\mathrm{x}_j^\varepsilon(t)\right]$$

such that $\epsilon_{\mathcal{M}}^{t]}\left[\widetilde{\varkappa}_j^{\mathrm{u}}(t,\varrho_t^{\mathrm{x}})\,\mathrm{d}\widetilde{\mathrm{x}}_j\right] = 0$, with a drift coefficient $\widetilde{\lambda}_\star(t,\varrho_{\mathrm{v}}^t)$ defining the conditional expectation

$$\epsilon_{\mathcal{M}}^{t]}\left[\mathrm{d}\varrho_{\mathrm{x}}^t\right](\upsilon) = \mathbb{E}\left[\mathrm{d}\varrho_x^t|\mathfrak{B}^t\right](\upsilon) = \widetilde{\lambda}_\star\left(t,\varrho_{\mathrm{x}}^t\right)\mathrm{d}t.$$

Using the Itô multiplication $\left(\mathrm{dx}_j^\varepsilon\right)^2 = \mathrm{d}t\mathrm{I}_t + \varepsilon_j\mathrm{dx}_j^\varepsilon$, $\mathrm{d}t\mathrm{dx}_j^\varepsilon = 0$, $(\mathrm{d}t)^2 = 0$ to obtain

$$\mathrm{d}p_{\mathrm{x}}^t\mathrm{d}\varrho_{\mathrm{x}}^t = \left\langle\varpi_{\mathrm{x}}^t, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle\widetilde{\varkappa}_\varepsilon^j\left(\varrho_{\mathrm{x}}^t\right)\left(\mathrm{d}t + \varepsilon_j\mathrm{dx}_j^\varepsilon\right)$$

for $\mathrm{d}p_{\mathrm{x}}^t = \left\langle\varpi_{\mathrm{x}}^t, \delta(\mathrm{I})\right\rangle\mathrm{d}t + \left\langle\varpi_{\mathrm{x}}^t, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle\mathrm{dx}_j^\varepsilon$ and (73), one can easily obtain

$$\left\langle\varpi_{\mathrm{x}}^t, \delta(\mathrm{I})\right\rangle\varrho_{\mathrm{x}}^t + p_{\mathrm{x}}^t\widetilde{\lambda^\star}\left(\varrho_{\mathrm{x}}^t\right) = \lambda^\star\left(\varpi_{\mathrm{x}}^t\right)$$

by equating the coefficients in $\mathrm{d}(p\varrho) = \mathrm{d}\varpi$ against $\mathrm{d}t$, and against $\mathrm{dx}_j^\varepsilon$,

$$\left\langle\varpi, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle\varrho + p\widetilde{\varkappa}_\varepsilon^j(\varrho) + \left\langle\varpi, \varkappa_\varepsilon^{j\star}(\mathrm{I})\right\rangle\widetilde{\varkappa}_\varepsilon^j(\varrho)\,\varepsilon = \varkappa_\varepsilon^j(\varpi).$$

This gives the nonlinear expressions (75) and (76) for the coefficients $\widetilde{\lambda^\star}$ and $\widetilde{\varkappa}_\varepsilon^j$ of the stochastic equation (73) with the *innovations* processes

$$\widetilde{\mathrm{x}}_j^\varepsilon(t) = \mathrm{x}_j^\varepsilon(t) - \int_0^t\left\langle\varrho_{\mathrm{x}}^t, \mathrm{L}^j + \mathrm{L}^{j\dagger} + \varepsilon\mathrm{L}^{j\dagger}\mathrm{L}^j\right\rangle\mathrm{d}s. \tag{77}$$

□

Note that the innovation processes (77) are *martingales* in the classical sense $\mathbb{E}\left[\widetilde{v}_j^\varepsilon(s)\,|\varrho,\upsilon_0^t\right] = \widetilde{v}_j^\varepsilon(t)$ for any $s \geq t$, satisfying the same multiplication table $\mathrm{d}t\mathrm{d}\widetilde{v}_j^\varepsilon = 0 = \mathrm{d}\widetilde{v}_j^\varepsilon\mathrm{d}t$ and $\left(\mathrm{d}\widetilde{v}_j^\varepsilon\right)^2 = \mathrm{d}t + \varepsilon_j\mathrm{d}\widetilde{v}_j^\varepsilon$ as the output processes $v_j^\varepsilon(t)$. In the case of Markovianity of the stochastic $\mathbb{A}_\star^{\mathrm{o}}$-valued process $\upsilon \mapsto \varrho_\upsilon^t$ this simply follows from $\mathbb{E}\left[\mathrm{d}\varrho_x^t|\varrho,\upsilon_\bullet^t\right] = \mathbb{E}\left[\mathrm{d}\varrho_x^t|\varrho\right] = \mathrm{d}\varrho^t$ for a fixed $\varrho_x^t = \varrho$, and therefore

$$0 = \mathbb{E}\left[\mathrm{d}\varrho_x^t|\varrho\right] - \widetilde{\lambda_\star}(\varrho)\,\mathrm{d}t = \widetilde{\varkappa}_\varepsilon^j(\varrho)\,\mathbb{E}\left[\mathrm{d}\widetilde{v}_j^\varepsilon|\varrho,\upsilon_0^t\right].$$

However, except the diffusive case $\varepsilon = 0$, the counting innovation processes $\widetilde{v}_j^\varepsilon$ are not independent increment processes with respect to the output probability measure P, as is $\widetilde{w}_\bullet = \widetilde{v}_\bullet^0$. Thus, the Girsanov transformation $\widetilde{m}_j$ of the standard Poisson (respectively Q) processes m_j are not Poisson, unlike the Girsanov transformation $\widetilde{w}_\bullet$ of the Wiener (respectively Q) processes w_j, which remain Gaussian under the transformed measure P.

In the case $\varepsilon_j = 0$ corresponding to (67) the Belavkin nonlinear filtering equation (74) has the form

$$\mathrm{d}\varrho_{\mathrm{x}}^t - \widetilde{\lambda_\star}\left(\varrho_{\mathrm{x}}^t\right)\,\mathrm{d}t = \widetilde{\varkappa}_0^j\left(\varrho_{\mathrm{x}}^t\right)\,\left(\mathrm{dw}_j^t - \left\langle \varrho_{\mathrm{x}}^t, \mathrm{L}^j + \mathrm{L}^{j\dagger}\right\rangle \mathrm{d}t\right) \tag{78}$$

where $\widetilde{\lambda_\star}(\varrho)$ is the controlled generator (66) and

$$\widetilde{\varkappa}_0^j(\varrho) = \varrho \mathrm{L}^{j\dagger} + \mathrm{L}^j \varrho - \left\langle \varrho, \mathrm{L}_j + \mathrm{L}_j^\dagger \right\rangle \varrho \equiv \varrho \widetilde{\mathrm{L}}^{j\dagger} + \widetilde{\mathrm{L}}^j \varrho \tag{79}$$

is the nonlinear *diffusion coefficient*. In the case $\varepsilon_j = 1$ corresponding to (68) the Belavkin nonlinear filtering equation has the form

$$\mathrm{d}\varrho_{\mathrm{x}}^t - \widetilde{\lambda_\star}\left(\varrho_{\mathrm{x}}^t\right)\,\mathrm{d}t = \widetilde{\varkappa}_1^j\left(\varrho_{\mathrm{x}}^t\right)\,\left(\mathrm{dm}_j^t - \left\langle \varrho_{\mathrm{x}}^t, \mathrm{C}_j \mathrm{C}_j^\dagger - I\right\rangle \mathrm{d}t\right), \tag{80}$$

where $\mathrm{C}^j = \mathrm{I} + \mathrm{L}^j = \mathrm{C}_j^\sharp$ defining the nonlinear *jump coefficient*

$$\widetilde{\varkappa}_1^j(\varrho) = \widetilde{\varsigma}^j(\varrho) - \varrho,\ \ \widetilde{\varsigma}^j(\varrho) = \frac{\mathrm{C}^j \varrho \mathrm{C}^{j\dagger}}{\langle \varrho, \mathrm{C}_j \mathrm{C}_j^\dagger\rangle} \equiv \widetilde{\mathrm{C}}^j \varrho \widetilde{\mathrm{C}}^{j\dagger}. \tag{81}$$

4. Optimal Quantum semi-Markov Feedback Control

We now couple the system to a stochastic control force (row-vector) $\boldsymbol{u} = (u_j)_{j\in J_f}$ via *forward* feedback channels indexed by $J_f \subseteq J$ which is usually (but not necessary) assumed to be disjoint with the subset $J_e \subseteq J_\bullet$ of *estimation* channels. The vector feedback control process $\boldsymbol{u}(t,\omega)$ will causally depend on the previsible stochastic output process $v_e = (v_j)_{j\in J_e} \equiv \upsilon$ with the components $v_j(t,\omega)$ represented in the Fock space by the preoutput processes $\mathrm{v}_j(t)$. The force perturbs interaction quantum dynamics $\left\{\mathrm{V}_t^-(r)\,|r<t\right\}$ described by the forward cocycle of coisometries $\mathrm{V}_t^-(r) = \mathrm{V}_r^-(t)^\dagger$ such that $\mathrm{V}_r^-(t^s)$ is replaced by a QS controlled hemigroup $\mathrm{V}_r(t^s) = \mathrm{V}^s(t)^\dagger$ of isometric solutions to HP equation (33) controlled pseudo-isometric germ $\mathbf{T} = \mathbf{I} + \mathbf{L}$ instead of $\mathbf{S}(t) = \mathring{\mathrm{U}}_t(t^+)$ for a given feedback *control law* $\boldsymbol{u}(t) = \boldsymbol{\ell}(t, v_e)$.

In principle, the control law could determine any QS dynamics by an arbitrary choice of any operator-valued QS-integrable function $\boldsymbol{\ell}(t, v_e) = \mathbf{L}(t, v_e)$ satisfying the condition $\mathbf{T}^\star\mathbf{T} = \mathbf{I}$. If $\mathsf{dim}\,(\mathfrak{k} \otimes H_{\mathrm{o}}) = \infty$, this would require infinite dimensional space for the possible values of vector $\boldsymbol{u}(t)$. However, in practice the control law $\mathbf{L}(t, v_e)$ as an adapted function of the output process v_e, normally satisfies, apart of the pseudoisometricity constraint $\mathbf{L} + \mathbf{L}^\star\mathbf{L} + \mathbf{L}^\star = \mathbf{0}$, many other constraints determining the matrix elements $\mathtt{L}_\kappa^\mu(t, v_e) \vdash \mathbb{A}_{\mathrm{v}}^t$. Our main interest will be restricted to the coherent control strategies given by the particular form (43) of $\mathbf{L}(t, v_e)$ with v_e in

place of $\mathrm{v} = \mathrm{v}_e$ and a fixed uncontrolled germ-scattering $\mathbf{S}(t)$ determining (44) (Which can also deterministically depend on t.)

Coherent control of thus described open semi-Markov quantum dynamics, using field channels indexed by $J_f \subseteq J$, is realized by controlling quantum noise in these channels by a real-valued vector-process $\boldsymbol{u}(t) \in \mathbb{R}^{d_f}$ via their coherent states. One can start with uncontrolled dynamics described by the object Hamiltonian $\mathrm{H}^{\circ} = \hbar\,\mathsf{Im}(\mathrm{S}_+)$, Lindbladian operators $\mathrm{K}_\bullet = -\mathrm{S}_+^\dagger \mathrm{S}_\bullet$ defining QS isometric evolution by (33) by the arbitrary vector-operator S_+ and isometric scattering matrix $\mathrm{S}_\bullet$. Then apply the coherent conditional expectation $\mathrm{P}^\star_{\widehat{\zeta}_t}$ to the corresponding uncontrolled QS flow $\alpha_t(t^s)$, $t^s = t + s \geq t$, with a QS amplitude $\widehat{\zeta}^u_t(t^s) = \frac{\mathrm{i}}{\hbar}\widehat{u}^{\circ}(t^s)$, where u° is the column $u^{\circ} = \left(u^j\right)^{j\in J} = \boldsymbol{u}^\dagger$ of real-valued $u^j = u_j$, $j \in J_f$ and $u^j = 0$, $j \notin J_f$ embedded into $\mathfrak{k}$ and represented at each t^s as a vector-function of the commuting output operators $\widehat{v}_{j,t}(t') = \hat{\mathrm{v}}_{j,t}(t')$, $t' \in (t, t^s]$ and possibly of $u = \{v_e(r) \,|r \leq t\}$. This effectively results in a coherent controlled backward hemigroup $\{\mu^u_r(t) \,|r \leq t\}$ of normal contractive CP maps $\mu^u_r(t) : \mathbb{A}^t_{\mathrm{x}} \to \mathbb{A}^t_{\mathrm{x}}$ as vacuum expected controlled homomorphic dynamics $\mu^u_r(t) = \mathrm{Q}^\star \circ \alpha^u_r(t)$. The resulting semi-Markov master dynamics with respect to the increasing algebras $\mathbb{C}^t_{\mathrm{x}}$ of the eventum histories may not be Markov on the quantum object algebra $\mathbb{A}^{\circ}$, but at least it is conditionally Markov $\mathbb{A}^{\circ}$. It can be found by resolving the controlled master equation (49) given by change of the object Hamiltonian $\mathtt{H}^{\circ}$ to H^u and the Lindblad operators $\mathrm{S}^-_\bullet$ to K^u_j according to (62) and (63):

$$\mathrm{H}^u = \mathrm{H}^{\circ} - \sum_{j\in J_f} u^j 2\,\mathsf{Re}\left(\mathrm{S}^j_+\right),\ \mathrm{K}^u_\bullet = -\mathrm{S}^{u\dagger}_+ \mathrm{S}^u_\bullet, \tag{82}$$

where $\mathrm{i}s^{\circ}\mathrm{S}^u_+ = \mathrm{S}_+ - \frac{\mathrm{i}}{\hbar}u^{\circ}\mathrm{I}$ and $\mathrm{i}s^{\circ}\mathrm{S}^u_\bullet = \mathrm{S}_\bullet$ corresponding to $\mathrm{e}^{\mathrm{i}\theta_j} = \mathrm{i}s^j$, $s^j = \mathrm{sign}u^j$ for $j \in J_f$ and no change for the other S^j_+ and $\mathrm{S}^j_\bullet$ with $j \notin J_f$. The corresponding semi-Markov dynamics, based on the observation of the output stochastic process $\left(\mathrm{y}^\varepsilon_j(t) : j \in J_e\right)$ given at the preoutput by the QS integrals $\mathrm{x}^\varepsilon_j(t) = \mathfrak{i}^t_0\left(\mathbf{d}^\varepsilon_j\right)$, is described by the nonlinear Belavkin quantum filtering equation (74) which we write now in the feedback controlled form by simple substitution of the inputs $\mathrm{x}_e = (\mathrm{x}_j)_{j\in J_e}$ and $\mathbf{u} = \boldsymbol{\ell}(\mathrm{x}_e)$ for the outputs $\mathrm{y}_e = \widehat{v}_e$ and $\widehat{\boldsymbol{u}} = \ell(\mathrm{y}_e)$ as

$$\mathrm{d}\varrho^t_{\mathrm{y}} = \widetilde{\kappa}^{\widehat{u}}_\star\left(\varrho^t_{\mathrm{y}}\right)\mathrm{d}t + \sum_{j\in J_e} \widetilde{\varkappa}^j_\varepsilon\left(\varrho^t_{\mathrm{y}}\right)\mathrm{d}\widetilde{\mathrm{y}}^\varepsilon_j. \tag{83}$$

Here $\widetilde{\kappa}^{\widehat{u}}_\star\left(\varrho^t_{\mathrm{y}}\right) = \kappa^{\widehat{u}}_\star\left(\varrho^t_{\mathrm{y}}\right) + \left\langle \varrho^t_{\mathrm{y}}, \mathrm{S}^{\widehat{u}\dagger}_+ \mathrm{E}^{\circ}_{\circ}\mathrm{S}^{\widehat{u}}_+ \right\rangle \varrho^t_{\mathrm{y}}$, the orthoprojector $\mathrm{E}^{\circ}_{\circ} = \mathrm{I}^{\circ}_{\circ} -$

$\mathrm{S}_\bullet^{\mathtt{y}}\mathrm{S}_\bullet^{\mathtt{y}\dagger}$ describes the demolition (exit) event, and

$$\kappa_\star^u\left(t,\varrho_\upsilon^t\right)=\sum_{j\in J_\bullet}\mathrm{K}^j\left(t\right)\varrho_\upsilon^t\mathrm{K}^j\left(t\right)^\dagger-\mathrm{K}_-^u\left(t\right)\varrho_\upsilon^t-\varrho_\upsilon^t\mathrm{K}_-^u\left(t\right)^\dagger,\qquad(84)$$

is the controlled Lindblad generator preadjoint to (51), given by $\mathrm{K}^j=\left(\mathrm{K}_j-\frac{\mathrm{i}}{\hbar}\boldsymbol{u}\mathrm{S}_j\right)^\sharp$ and $\mathrm{K}_-^u\left(t\right)=-\mathrm{K}_+^u\left(t\right)^\sharp$. (Where in the tracial case $\sharp=\dagger$.)

We are going to consider the quantum feedback coherent control problem in which for simplicity we assume that control channels J_f are disjoint to the set $J_e=J_{\mathrm{w}}\cup J_{\mathrm{m}}$ of estimation channels. In this case the controlling amplitude $\zeta\left(t\right)\in\mathfrak{k}_f$ is orthogonal to the subspace $\mathfrak{k}_e=\mathfrak{k}_{\mathrm{w}}\oplus\mathfrak{k}_{\mathrm{m}}$ of observation channels, so the output equations (56), (57) are not affected by the coherent control; this will simplify the optimal feedback control problem which we solve by applying dynamical programming to coherent controlled quantum states. The controlled posterior density operator $\varrho_\upsilon^t=\varphi_\upsilon^{t-r}\left(r,\varrho_r^\upsilon\right)$ can then be obtained from the relevant uncontrolled filtering equation by replacing only the drift map $\widetilde{\kappa_\star}$ for $\widetilde{\kappa}_\star^{\widehat{u}}$ as in (83) without changing the fluctuating maps (76). So, we have a controlled time-dependent nonlinear filtering dynamics ϱ_υ^t satisfying the Belavkin equation in the form, say (78) or (80), in which the fluctuating part under the above coherent control assumption is independent of $\boldsymbol{u}\left(t\right)$.

4.1. *Dynamical programming of quantum states*

In the search for optimal control inputs, it is desirable to allow the control to be determined in terms of measurement results on the system, particularly in the quantum setting where classical-quantum interfaces introduce an inevitable stochastic nature. A feedback strategy $\boldsymbol{\ell}=\{\boldsymbol{\ell}\left(t\right)\}$ consists of measurable adapted maps $\boldsymbol{\ell}\left(t\right):\Upsilon^t\to\boldsymbol{U}$ which give for each $0\leq t<T$ a control law $\boldsymbol{u}\left(t,\omega\right)=\boldsymbol{\ell}(t,\upsilon)$ as a causal function of only the current and previous output variables $\upsilon\left(t\right)=\upsilon_e\left(r\right)$: $r\in(0,t]$. Thus, the control law $\boldsymbol{\ell}\left(t\right)$ is a $\mathfrak{B}^t$-adapted random vector variable $\boldsymbol{\ell}\left(t,\upsilon\right)$ in an admissible domain $\boldsymbol{U}\subseteq\mathbb{R}^{d_f}$ on the probability space $(\Upsilon,\mathfrak{B},\mathsf{P})$ of the output measurement results, and it can be represented in the operator form as a vector-function

$$\widehat{\boldsymbol{\ell}}\left(t\right)=\int_{\Upsilon^t}\boldsymbol{\ell}(t,\upsilon)\mathtt{I}\left(\mathrm{d}\upsilon\right)\equiv\boldsymbol{\ell}(t,\widehat{\upsilon}_e)$$

of the commuting output variables $\widehat{\upsilon}_e\left(r\right)=\mathtt{y}_e\left(r\right)\equiv\hat{\mathrm{v}}_e\left(r\right)$, $r\leq t$, having the joint spectral measure $\mathtt{I}\left(\Delta\right)=1_\Delta\left(\mathtt{y}_e\right)$ on $(\Upsilon,\mathfrak{B})$.

Following the original formulation[5,7,36] of quantum optimal control theory, we assume that the quality of a control process on a quantum object

over a finite period $[r,T)$ starting at $t=r$ in a product state $\mathrm{R}_r=\varpi_{\mathrm{v}}^r\otimes\mathrm{Q}_r$ on the algebra $\mathcal{A}_r\simeq\mathbb{A}_{\mathrm{v}}^r\bar{\otimes}\mathcal{N}_r$ is judged by the integral expectation

$$\mathsf{J}_r(\mathrm{R}_r,\boldsymbol{\ell}_r)=\int_r^T\left\langle\mathrm{R}_r,\hat{\mathtt{C}}\left(t,\widehat{\boldsymbol{\ell}}\,(t)\right)\right\rangle\mathrm{d}t+\langle\mathrm{R}_r,\hat{\mathtt{S}}\,(T)\rangle\,. \tag{85}$$

Here the operator-valued cost function $\hat{\mathtt{C}}\,(t)$ of the law $\widehat{\boldsymbol{\ell}}\,(t)=\boldsymbol{\ell}\,(t,\widehat{v}_e)$ is defined by a measurable operator-valued positive cost function $\boldsymbol{u}\to\mathtt{C}\,(t,\boldsymbol{u})\in\mathbb{A}^{\mathrm{o}}$ on $\boldsymbol{U}$ and $\hat{\mathtt{S}}\,(T)$ is the terminal, or bequest cost, defined by a positive self-adjoint operator $\mathtt{S}\in\mathbb{A}^{\mathrm{o}}$ evolved as

$$\hat{\mathtt{C}}\left(t,\widehat{\boldsymbol{\ell}}\right):=\alpha_r^{\mathrm{u}}\left(t,\mathtt{C}(t,\boldsymbol{\ell}\,(t))\right),\ \hat{\mathtt{S}}\,(T):=\alpha_r^{\mathrm{u}}\left(T,\mathtt{S}\right) \tag{86}$$

(Or by unbounded positive operators $\mathtt{C}(t,\boldsymbol{u})$, $\mathtt{S}$ affiliated to $\mathbb{A}^{\mathrm{o}}$.) An alternative problem of *risk-sensitive* control has also been studied by James[54,55] where the cost is exponentiated to enforce higher penalties for undesirable behavior such that it replaces by multiplicativity the additivity of the integral (85).

When a control law is in place, the output operators $\mathtt{v}_j\,(t)\,\mathrm{U}\,(t)=\mathrm{U}\,(t)\,\mathrm{x}_j\,(t)$, changed to $\mathtt{y}_j\,(t)\,\mathrm{V}\,(t)=\mathrm{V}\,(t)\,\mathrm{v}_j\,(t)$, once again form a commuting family compatible with the future Heisenberg operators $\hat{\mathtt{X}}\,(t)=\beta\,(t,\mathtt{X}\,(t))$ as it was pointed out in.[11,14] This implies the existence (70) of the posterior expectations

$$\epsilon_{\mathbb{C}_{\mathrm{v}}}^{t]}\left[\mathtt{Y}_{\mathrm{v},\mathrm{x}}^s\right](\upsilon)=\left\langle\vartheta_{\mathrm{x}}^s\left(\varrho_\upsilon^t\right),\mathtt{Y}_{\upsilon,\mathrm{x}}^s\right\rangle\equiv\mathbb{E}_t^\upsilon\left[\left\langle\varrho_{\upsilon,x}^{t+s},\mathtt{Y}_{\upsilon,x}^s\right\rangle\right] \tag{87}$$

of $\mathtt{Y}_{\mathrm{v},\mathrm{x}}^s\in\mathbb{C}_{\mathrm{v}}^t\bar{\otimes}\mathcal{M}_r^s$ with respect to the classical realization $v_e\,(r)\,,r\leq t$ of the output operators $\mathtt{y}_{t]}$ for $t>r$. It is given by the posterior state $\varrho_\upsilon^t=\varphi_\upsilon^t\,(r,\varrho)$ as a solution of the general Belavkin equation (74) which now has dependence on the chosen control law $\boldsymbol{u}\,(t)=\boldsymbol{\ell}\,(t,v_e)$ through the generator (84), and it does not really depend on $\mathtt{y}_{r]}$ in the Markovian case considered in.[7] The existence of this conditional expectation permits the following theorem which lies at the heart of quantum feedback control based on the continuous observation. It allows to describe the expected cost-to-go $\mathsf{J}_r(\mathrm{R}_r,\boldsymbol{\ell}_r)$ by a classical expression for a truncated admissible strategy $\boldsymbol{\ell}_r\in\boldsymbol{L}_r$ when starting in an arbitrary state $\varrho_\upsilon^r=\varrho$ at time r.

Theorem 4.1. *The expectation (85) of the operator valued costs (86) when a feedback control law* $\mathbf{u}\,(t,\omega)=\boldsymbol{\ell}(t,\upsilon,v_r)$ *on* $\omega=(\upsilon,v_r)$, *given* $\upsilon=\{v\,(r)\,|r\leq t\}$, *is in operation can be written as a classical expectation* $\mathbb{E}_{\mathsf{P}}\left[\mathsf{J}_r^{v_e}(\varrho_{v_e}^r,\boldsymbol{\ell}_r)\right]$ *with respect to the output probability measure* $\mathsf{P}\,(\mathrm{d}\omega)=$

$p^t_\omega \mathsf{Q}\,(\mathrm{d}\omega)$ *of the conditional expectation*

$$\mathsf{J}^\upsilon_r(\varrho^r_\upsilon, \boldsymbol{\ell}_r) = \int_{\Upsilon_r} \mathsf{J}^\omega_r(\varrho_\omega, \boldsymbol{\ell}_r\,(\omega))\mathsf{P}\,(\mathrm{d}\omega|\upsilon) := \mathbb{E}^\upsilon_r[\mathsf{J}^{\upsilon,x}_r(\varrho_{\upsilon,x}, \boldsymbol{\ell}_r)], \tag{88}$$

where $x\,(\omega) = \upsilon_r \equiv \upsilon_r\,(\omega)$ *and* J^ω_r *and of the random cost-to-go function*

$$\mathsf{J}^\omega_r(\varrho_\omega, \boldsymbol{\ell}_r) = \int_r^T \langle \varrho^t_\omega, \mathtt{C}(\boldsymbol{\ell}(t,\omega)\rangle \mathrm{d}t + \langle \varrho^T_\omega, \mathtt{S}\rangle. \tag{89}$$

Here $\varrho^t_\omega = \varphi^{t-r}_\omega\,(r, \varrho)$ *is the solution to the controlled filtering equation corresponding to the chosen measurement processes* $\mathrm{x}_r = \{\mathrm{x}_r\,(t)\,|t > r\}$*, classically represented as* $\upsilon_r = \{\upsilon_r\,(t)\,|t > r\}$ *for the feedback strategy* $\boldsymbol{\ell}_r$ *with the initial condition* $\varrho^r_\omega = \varrho$ *for all* $\omega = (\upsilon, \upsilon_r)$*.*

Proof. Using the existence and state invariance of the conditional expectation and the classical isomorphism proved in the first Section, it is a straight forward application of the formula (87) to $\mathtt{X}\,(t) = \mathtt{C}\,(\boldsymbol{\ell}\,(t))$ and $\mathtt{X}\,(T) = \mathtt{S}$ in (85). We can write the expected cost as

$$\begin{aligned}
\mathsf{J}_r(\varrho, \boldsymbol{\ell}_r) &= \int_r^T \left\langle \varpi^r_{\mathrm{x}}, \mathrm{Q}^\star_r \left[\widehat{\mathtt{C}}\left(\widehat{\boldsymbol{\ell}}\,(t)\,, t\right)\right]\right\rangle \mathrm{d}t + \langle \varpi^r_{\mathrm{x}}, \mathrm{Q}^\star_r\,[\mathtt{S}]\rangle \\
&= \int_r^T \left\langle \vartheta^{t-r}_{\mathrm{v}}\,(\varpi^r_{\mathrm{x}})\,, \mathtt{C}\,(\mathbf{y}\,(t))\right\rangle \mathrm{d}t + \left\langle \vartheta^{T-r}_{\mathrm{v}}\,(\varpi^r_{\mathrm{x}})\,, \mathtt{S}\right\rangle \\
&= \int_r^T \int_{\Upsilon^t} \left(\left\langle \varpi^t_\upsilon, \mathtt{C}\,(\boldsymbol{\ell}\,(t,\upsilon))\right\rangle \mathrm{d}t + \left\langle \varpi^t_\upsilon, \mathtt{S}\right\rangle\right) \mathsf{Q}\,(\mathrm{d}\upsilon) \\
&= \int_r^T \int_{\Upsilon^t} \left(\left\langle \varrho^t_\upsilon, \mathtt{C}\,(\boldsymbol{\ell}\,(t,\upsilon))\right\rangle \mathrm{d}t + \left\langle \varrho^t_\upsilon, \mathtt{S}\right\rangle\right) p^t_\upsilon \mathsf{Q}\,(\mathrm{d}\upsilon) \\
&= \mathbb{E}_{\mathsf{P}}\left[\int_r^T \left\langle \varrho^t_\upsilon, \mathtt{C}\,(\boldsymbol{\ell}\,(t))\right\rangle \mathrm{d}t + \left\langle \varrho^t_\upsilon, \mathtt{S}\right\rangle\right].
\end{aligned}$$

□

We denote the subspace of admissible operator-valued feedback control segments $\boldsymbol{\ell}^s_r$ on the interval $[r, r+s)$ by $\boldsymbol{L}^s_r\,(\Upsilon)$. Note that no measurement results are available initially, so the initial control is deterministic, $\boldsymbol{\ell}\,(0, \upsilon) = \boldsymbol{u}\,(0)$, and also no controls are applied at the termination time T. It is too restrictive to consider only continuous sample paths $(\boldsymbol{\ell}(t, \upsilon) : t > 0)$, since, for example, a counting process $\{m\,(t)\}$ certainly does not have continuous sample paths. Instead we give the following definition of an admissible strategy.

Definition 4.1. An feedback control strategy $\boldsymbol{\ell}_r\,(\upsilon_r) = \{\boldsymbol{\ell}_r\,(t, \upsilon_r)\,|t > r\}$ determining for $\upsilon_r \in \Upsilon_r$ the adapted control laws $\boldsymbol{u}\,(t, \omega) = \boldsymbol{\ell}(t, \upsilon, \upsilon_r)$ is

called admissible if it forms previsible sample paths $\{\boldsymbol{u}(t,\omega)\,|t>r\}$ as left continuous, having the right limits, realizing the values in a given constraint sets $\boldsymbol{U}(t)$. Moreover, an admissible strategy $\boldsymbol{\ell}_r^o$ shall be called *optimal* if it realizes the infimum

$$\mathsf{S}(t,\varrho)=\inf_{\boldsymbol{\ell}_t\in\boldsymbol{L}_t}\mathbb{E}_{\mathsf{P}}\left[\mathsf{J}_t^{v_e}(\varrho_{v_e}^t,\boldsymbol{\ell}_t)|\varrho\right]=\mathsf{J}_r(\varrho,\boldsymbol{\ell}_t^o) \tag{90}$$

over the space $\boldsymbol{L}_r$ of admissible feedback control strategies, where $\mathsf{J}_r(\varrho,\boldsymbol{\ell}_r)$ is the expected cost

$$\mathbb{E}_{\mathsf{P}}\left[\mathsf{J}_r^{v_e}(\varrho_{v_e}^r,\boldsymbol{\ell}_r)|\varrho\right]=\int_{\Upsilon^t}\mathsf{J}_r^{\upsilon}(\varrho,\boldsymbol{\ell}_r(\upsilon))\mathsf{P}_{\varrho}(\mathrm{d}\upsilon)$$

for the control process determined by the feedback strategy $\boldsymbol{\ell}_r$ under the condition $\varrho_v^t=\varrho$.

Thus, we can treat the quantum feedback control problem as a classical stochastic control problem on the space of quantum states as it was first done by Belavkin in.[7,11,14] Here, instead of repeating the derivation of the corresponding Bellman equation for this special case, we will use the following result. It is essentially an application of the corresponding classical stochastic control separation lemma to the case of quantum state-valued stochastic controlled processes, and the detailed proof specified for this case can be found in.[42]

Theorem 4.2. *Suppose that $\mathsf{S}(t,\varrho)$ is a functional continuously differentiable in t, has continuous Frèchet derivatives of all order with respect to ϱ and satisfies*

$$\inf_{\boldsymbol{u}\in\boldsymbol{U}(t)}\left\{\langle\varrho_v^t,\mathtt{C}(\boldsymbol{u})\rangle+\mathbb{E}_{\mathsf{P}}\left[\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{S}(t,\varrho_x^t)|\mathfrak{B}^t\right](\upsilon)\right\}=0. \tag{91}$$

for all $0<t<T$ and $\mathsf{S}(T,\varrho)=\langle\varrho,\mathsf{S}\rangle$ for all $\varrho\in\mathfrak{s}(\mathbb{A}^{\mathrm{o}})$. Suppose also that $\boldsymbol{\ell}_r^o$ is the strategy built from the control laws attaining these minima fro $t>r$ within admissible control values, then $\mathsf{S}(r,\varrho)$ is the function which minimizes (88) under the condition $\varrho_v^r=\varrho$ and $\boldsymbol{\ell}_r^o$ is the optimal strategy for the control problem on (r,T).[a]

Sketch proof. Let $\{\boldsymbol{\ell}(t)\}$, $\{\varrho_v^t\}$ be any control and state trajectories resulting from an admissible strategy $\boldsymbol{\ell}_r$ and the state $\varrho_v^r=\varrho$, then from (91),

[a]Additional technical assumptions and mathematical rigour are required to formalise the proof of this theorem when dealing with unbounded operators which is beyond the scope of this paper. See recommended texts e.g.[41,56] for a formal classical treatment.

we have the inequality

$$\langle \varrho_{\upsilon}^{t}, \mathtt{C}(\boldsymbol{\ell}(t,\upsilon))\rangle + \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{S}(t,\varrho_{x}^{t})|\mathfrak{B}^{t}\right](\upsilon) \geq 0$$

which we integrate over $[r,T]$ and take the expectation to obtain due to convexity

$$\mathbb{E}\left[\int_{r}^{T}\langle \varrho_{\mathrm{x}}^{t}, \mathtt{C}(\boldsymbol{\ell}(t))\rangle \mathrm{d}t + \mathsf{S}(T,\varrho_{\mathrm{x}}^{T}) - \mathsf{S}(r,\varrho)\right] \geq 0$$

Since we have $\varrho_{\mathrm{x}}^{r} = \varrho$, we can rearrange and use the terminal condition to obtain

$$\mathsf{S}(r,\varrho) \leq \mathbb{E}\left[\int_{r}^{T}\langle \varrho_{\mathrm{x}}^{t}, \mathtt{C}(\boldsymbol{\ell}(t))\rangle \mathrm{d}t + \langle \varrho_{\mathrm{x}}^{T}, \mathtt{S}\rangle\right] = \mathsf{J}_{r}(\varrho,\boldsymbol{\ell}).$$

with equality when $\boldsymbol{\ell}_{r} = \boldsymbol{\ell}_{r}^{o}$ and so the lower bound $\mathsf{S}(r,\varrho)$ is attained, proving optimality of $\boldsymbol{\ell}_{r}^{o}$. □

A choice of controlled filtering equation is required to determine the stochastic trajectories ϱ_{x}^{t} along which to differentiate candidate solutions. The next two sections are concerned with the examples of feedback control with respect to continuous measurements of the diffusive process $\left(\mathtt{y}_{j}^{0} : j \in J_{\mathrm{w}}\right)$ and the counting process $\left(\mathtt{y}_{j}^{1} : j \in J_{\mathrm{m}}\right)$ respectively.

4.2. *Quantum state Bellman equations*

First let us introduce notations of differential calculus on the quantum state space $\mathfrak{s}(\mathbb{A}^{\mathrm{o}}) \subset \mathbb{A}_{\star}^{\mathrm{o}}$. Let $\mathsf{F} = \mathsf{F}[\cdot]$ be a (nonlinear) scalar function $\varrho \mapsto \mathsf{F}[\varrho]$ on $\mathfrak{s}(\mathbb{A}^{\mathrm{o}})$, then we say it admits a (Frèchet) derivative if there exists an $\mathbb{A}^{\mathrm{o}}$-valued function $\nabla_{\varrho}\mathsf{F}[\cdot]$ on $\mathbb{A}_{\star}^{\mathrm{o}}$ such that

$$\lim_{h\to 0}\frac{1}{h}\left\{\mathsf{F}[\cdot + h\varphi] - \mathsf{F}[\cdot]\right\} = \langle \varphi, \nabla_{\varrho}\mathsf{F}[\cdot]\rangle = (\varphi, \nabla_{\varrho}\mathsf{F}[\cdot]) \tag{92}$$

for each $\varphi = \varphi^{\dagger} \in \mathbb{A}_{\star}^{\mathrm{o}}$. In the same spirit, a Hessian $\nabla_{\varrho}^{\otimes 2} \equiv \nabla_{\varrho} \otimes \nabla_{\varrho}$ can be defined as a mapping from the functions F onto the $\mathbb{A}_{\mathrm{sym}}^{\mathrm{o}\bar{\otimes}2}$-valued functions, via

$$\begin{aligned}\lim_{h,h'\to 0}\frac{1}{hh'}&\left\{\mathsf{F}[\cdot + h\varphi + h'\varphi'] - \mathsf{F}[\cdot + h\varphi] - \mathsf{F}[\cdot + h'\varphi'] + \mathsf{F}[\cdot]\right\}\\ &= \langle \varphi\otimes\varphi', \nabla_{\varrho}\otimes\nabla_{\varrho}\mathsf{F}[\cdot]\rangle,\end{aligned} \tag{93}$$

and we say that the function F is twice continuously differentiable whenever $\nabla_{\varrho}^{\otimes 2}\mathsf{F}[\cdot]$ exists and is continuous in the uniform topology of $\mathbb{A}_{\star}^{\mathrm{o}}$.

4.2.1. *Diffusive Bellman equation*

We have the Itô rule $\mathrm{d}v_i^0\mathrm{d}v_k^0=\delta_{ik}\mathrm{Id}t$ with $i,k\in J_e=J_{\mathrm{w}}$ for the increments of diffusive processes $v_j^0=w_j$ which have the expectations

$$\mathbb{E}_t^u[\mathrm{d}v_j^0\,(t)\,|\mathfrak{B}^t]\,(\upsilon)=\langle\varrho_\upsilon^t,2\,\mathsf{Re}\,(\mathtt{L}_j)\,(\upsilon)\rangle\mathrm{d}t,$$

so using the classical Itô formula for the diffusive process v_j^0 we can show for any $\varrho=\varrho_\upsilon^t$

$$\begin{aligned}\mathbb{E}_t^u[\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{S}(t,\varrho)|\mathfrak{B}^t]=&\frac{\partial}{\partial t}\mathsf{S}(t,\varrho)+\langle\widetilde{\kappa}_\star^u\,(\varrho)\,,\nabla_\varrho\mathsf{S}(t,\varrho)\rangle\\&+\frac{1}{2}\sum_{j\in J_{\mathrm{w}}}\langle\widetilde{\varkappa}_0^j(\varrho)\otimes\widetilde{\varkappa}_0^j(\varrho),\nabla_\varrho^{\otimes2}\mathsf{S}(t,\varrho)\rangle\end{aligned}\tag{94}$$

where $\nabla_\varrho\mathsf{S}\in\mathbb{A}^{\mathrm{o}}$ denotes Frèchet derivative with respect to $\varrho\in\mathbb{A}_\star^{\mathrm{o}}$, and $\nabla_\varrho^{\otimes2}\mathsf{S}\in\mathbb{A}^{\mathrm{o}}\bar{\otimes}\mathbb{A}^{\mathrm{o}}$ denotes Hessian applied to S. Observing $\frac{\partial}{\partial t}\mathsf{S}(t,\varrho)$ is independent of $\boldsymbol{u}$ leads to the following corollary.

Corollary 4.1. *Suppose there exists a function $\mathsf{S}(t,\varrho)$ which is continuously differentiable in t, has continuous first and second order Frèchet derivatives with respect to ϱ and satisfies the following Bellman equation*

$$\begin{aligned}-\frac{\partial}{\partial t}\mathsf{S}(t,\varrho)=&\inf_{\boldsymbol{u}\in\boldsymbol{U}}\Big\{\,\langle\varrho,\mathtt{C}(\boldsymbol{u})\rangle+\langle\widetilde{\kappa}_\star^u\,(\varrho)\,,[\nabla_\varrho\mathsf{S}(t,\varrho)]\rangle\,\Big\}\\&+\frac{1}{2}\sum_j\langle\widetilde{\varkappa}_0^j(\varrho)\otimes\widetilde{\varkappa}_0^j(\varrho),\nabla_\varrho^{\otimes2}\mathsf{S}(t,\varrho)\rangle\end{aligned}\tag{95}$$

for all $t>0$, $\varrho\in\mathfrak{s}\,(\mathbb{A}^{\mathrm{o}})$ with the terminal condition $\mathsf{S}(T,\varrho)=\langle\varrho,\mathtt{S}\rangle$. Then the strategy $\boldsymbol{\ell}^o\,(t,\upsilon_{\mathrm{w}})=\boldsymbol{u}^o\,(t,\varrho_\upsilon^t)$ built from the control laws

$$\boldsymbol{u}^o\,(t,\varrho)=\arg\inf_{\boldsymbol{u}\in\boldsymbol{U}}\Big\{\,\langle\varrho,\mathtt{C}(\boldsymbol{u})\rangle+\langle\widetilde{\kappa}_\star^u,\nabla_\varrho\mathsf{S}(t,\varrho)\rangle\,\Big\}\tag{96}$$

for $0\leq t<T$ is optimal for the feedback control problem based on diffusive output measurements.

Note that the last line in (95) defines the Laplace operator

$$\triangle_0\mathsf{S}\,(\varrho)=\sum_{j\in J_{\mathrm{w}}}\Big\langle\widetilde{\varkappa}_0^j\,(\varrho)\otimes\widetilde{\varkappa}_0^j\,(\varrho)\,,\nabla_\varrho^{\otimes2}\mathsf{S}(t,\varrho)\Big\rangle$$

in the quantum state 'coordinates' $\varrho\in\mathbb{A}_\star^{\mathrm{o}}$ as the sufficient coordinates from the preadjoint space of the algebra $\mathbb{A}^{\mathrm{o}}$.

4.2.2. *Counting Bellman equation*

We have the Itô rules $\mathrm{d}v_i^1\mathrm{d}v_k^1=\delta_i^j\delta_k^j\mathrm{d}v_j^1$ for the increments of the counting processes $v_j^1=m_j$, $j\in J_e=J_{\mathrm{m}}$ which have the expectations

$$\mathbb{E}_t^u[\mathrm{d}v_j^1|\mathfrak{B}^t]\,(v)=\langle\varrho_v^t,\mathtt{M}_j\,(t)-\mathtt{I}\rangle\mathrm{d}t,$$

where $\mathtt{M}_j=(\mathtt{I}+\mathtt{L}_j)\,(\mathtt{I}+\mathtt{L}_j)^\dagger$, so using Itô formula for the counting processes v_j^t we can show for any $\varrho=\varrho_v^t$

$$\begin{aligned}\mathbb{E}_t^u[\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{S}(t,\varrho)|\mathfrak{B}^t]=&\frac{\partial}{\partial t}\mathsf{S}(t,\varrho)+\langle\widetilde{\kappa}_\star^u\,(\varrho)\,,\nabla_\varrho\mathsf{S}(t,\varrho)\rangle\\&+\frac{1}{2}\triangle_1\mathsf{S}\,(t,\varrho)\,.\end{aligned}\tag{97}$$

It is written similar to the diffusive case using the Feller Laplassian

$$\triangle_1\mathsf{S}\,(t,\varrho)=2\sum_{j\in J_{\mathrm{m}}}\langle\varrho,\mathtt{M}_j\,(t)\rangle\left(\mathsf{S}\left(t,\widetilde{\varsigma}^j\,(\varrho)\right)-\left(1+\left\langle\widetilde{\varkappa}_1^j\,(\varrho)\,,\nabla_\varrho\right\rangle\right)\mathsf{S}\,(t,\varrho)\right),$$

where $\widetilde{\varkappa}_1^j\,(\varrho)=\widetilde{\varsigma}^j\,(\varrho)-\varrho$. It has formal Taylor expansion

$$\triangle_1^j\mathsf{S}\,(t,\varrho)=2\sum_{n=2}^{\infty}\frac{1}{n!}\,\langle\varrho,\mathtt{M}_j\,(t)\rangle\left\langle\widetilde{\varkappa}_1^j\,(\varrho)^{\otimes n}\,,\nabla_\varrho^{\otimes n}\mathsf{S}\,(t,\varrho)\right\rangle$$

in terms of the higher order Frèchet derivatives $\nabla_\varrho^{\otimes n}$ starting from the Hessian $\nabla_\varrho^{\otimes 2}=\nabla_\varrho\otimes\nabla_\varrho$ determining the usual Laplace operator for the diffusive approximation of this counting measurement case.

Let us introduce the Pontryagin's 'Hamiltonian' in the 'coordinates' $\mathtt{Q}\,(\varrho)=\varrho_0-\varrho$ and 'momenta' $\mathtt{P}\in\mathbb{A}^{\mathrm{o}}$ as the Legendre-Fenchel transform

$$\mathsf{H}\,(\mathtt{Q},\mathtt{P})=\sup_{\boldsymbol{u}\in\boldsymbol{U}}\left\{\langle\widetilde{\kappa}_\star^u\,[\mathtt{Q}]\,,\mathtt{P}\rangle-\mathsf{L}\,(\mathtt{Q},\boldsymbol{u})\right\}$$

of the 'Lagrangian' $\mathsf{L}\,(\mathtt{Q},\boldsymbol{u})=\langle\varrho_0-\mathtt{Q},\mathtt{C}\,(\boldsymbol{u})\rangle$, where $\varrho_0=\mathtt{Q}-\varrho$ is any stationary element $\varrho_0=\varrho_0^\dagger$ in $\mathbb{A}_\star^{\mathrm{o}}$ such that $\widetilde{\kappa}_\star^u\,(\varrho_0)=0$ for any $\boldsymbol{u}\in\boldsymbol{U}$ (e.g. $\varrho_0=0$). Then one can write Bellman equation defining minimal expected cost-to-go (90) as the action function in the compact Jacobi-Feller form for $\varepsilon=1$, unified with the diffusive case corresponding to $\varepsilon=0$.

Corollary 4.2. *Suppose there exists a function $\mathsf{S}(t,\varrho)$ which is continuously differentiable in t, has continuous first order Frèchet derivatives with respect to ϱ and satisfies the following Bellman-Jacobi-Feller equation*

$$-\frac{\partial}{\partial t}\mathsf{S}(t,\varrho)+\mathsf{H}\,(\mathtt{Q}\,(\varrho)\,,\nabla_\varrho\mathsf{S}(t,\varrho))=\frac{1}{2}\,\triangle_\varepsilon^j\,\mathsf{S}\,(t,\varrho)$$

for all $t > 0$, $\varrho \in \mathbb{A}^{\circ}_{\star}$ *with the terminal condition* $\mathsf{S}(T,\varrho) = \langle \varrho, \mathtt{S}\rangle$. *Then the strategy* $\boldsymbol{\ell}^{o}(t, v_e) = \boldsymbol{u}^{\circ}(t, \varrho^{t}_{v})$ *built from the control laws*

$$\boldsymbol{u}^{\circ}(t,\varrho) = \arg\inf_{\boldsymbol{u}\in\boldsymbol{U}} \left\{ \langle \varrho, \mathtt{C}(\boldsymbol{u})\rangle + \langle \widetilde{\kappa}^{u}_{\star}(\varrho), \nabla_{\varrho}\mathsf{S}(t,\varrho)\rangle\rangle \right\} \tag{98}$$

is optimal for the feedback control problem based on counting output measurements.

Thus we have shown that without loss in optimality, one can reformulate the unobservable quantum feedback control problem into a feedback problem based on indirect QND measurements with feedback of the controlled conditional density matrix. However, the corresponding Hamilton-Jacobi-Bellman equation resulting from the minimization rarely has a regular solution $\mathsf{S}(t,\varrho)$ from which to construct the optimal feedback laws. We now study a specific quantum filtering and feedback case where such a control solution can be explicitly found, which is familiar as the only such example in the classical case.

5. Linear Quantum Stochastic Dynamical System

Let $\mathtt{s}_{\circ} = (\mathtt{s}_1, \ldots, \mathtt{s}_m)$ be the canonical self-adjoint operators $\mathtt{s}_i = \mathtt{s}_i^{\dagger}$, $i = 1, ..., m$ on the quantum object Hilbert space H° satisfying the canonical commutation relations (CCRs)

$$[\mathtt{s}_i, \mathtt{s}_k] := \mathtt{s}_i\mathtt{s}_k - \mathtt{s}_i\mathtt{s}_i = \mathrm{i}\hbar j_{ik}\mathtt{I}^{\circ}, \tag{99}$$

and $\mathrm{s}_{\circ}(t) = \mathrm{I}^{t} \otimes \mathtt{s}_{\circ} \otimes \mathrm{I}_t$ be their representation on $H_t = G^t \otimes H_o \otimes F_t$. Here $j_{ik} = -j_{ki} = j^{*}_{ik}$ are the real elements of antisymmetric matrix $\boldsymbol{j} = (j_{ik})$ defining the CCR for the operator row-vector $\mathrm{s}_{\circ} = (\mathrm{s}_j)$ in the matrix form as $[\mathrm{s}^{\intercal}_{\circ}, \mathrm{s}_{\circ}] = \mathrm{i}\hbar\boldsymbol{j}\mathrm{I}$. Usually $\boldsymbol{j}$ is the standard symplectic matrix $\boldsymbol{j}^{\intercal} = \boldsymbol{j}^{-1}$, but we may assume that $\boldsymbol{j}$ is nonstandard or even degenerate in order to include also the commuting (classical) random variables into $\mathrm{s}_{\circ} = (\mathrm{s}_1, \ldots, \mathrm{s}_m)$ as it is in the case of odd m. It is worth remarking at this point that the noncommuting operators s_i are secondary commuting (in the sense of commutativity with all the commutants (99)), and therefore they must be unbounded generators of the object W*-algebra $\mathbb{A}^{\circ}$.

We couple the open quantum system to $d_e = |J_e|$ estimation (side) channels with linear combinations $\mathrm{L}_j = \sum_i \lambda_{ji}\mathrm{s}_i$ indexed by a subset $J_e \subseteq \{1, \ldots, d\}$, given by a complex-valued $d_e \times m$ matrix $\boldsymbol{\lambda}_e = (\lambda_{ij})$ with $\lambda_{ij} = 0$ for $i \notin J_e$. At the output the channel, we perform a measurement of the preoutput process of operator row-vectors

$$\mathrm{v}_e(t) = 2\,\mathsf{Re}\left[\mathrm{A}^{-}_e(t)\right] = (\mathrm{v}_1, \ldots, \mathrm{v}_d)(t)$$

given by the output operators $\mathtt{y}_j(t) = \hat{\mathrm{v}}_j$ with $j \in J_e$, where $2\,\mathsf{Re}\left[\mathrm{A}_j^+\right] = \mathrm{A}_j^+ + \mathrm{A}_-^j$ is the canonical vacuum representation of the standard independent Wiener processes $x_j = w_j$ and $\hat{\mathrm{v}}_j(t) = \beta(t, \mathrm{v}_j(t))$ is defined by a controlled QS dynamics $\beta(t, \mathrm{Z}) = \mathrm{V}(t)\,\mathrm{Z}\,\mathrm{V}(t)^\dagger$ to be specified. The system is also coupled to $d_f = |J_f|$ feedback (input) channels by the operators $\mathrm{K}^j = \mathrm{L}_j^\dagger$ as linear combinations of (s_j), given by the row $\mathbf{L}_f = \mathrm{s}_\circ \boldsymbol{\lambda}_f^\intercal$ of operators L_j, $j \in J_f$ for a complex-valued $d_f \times m$ matrix $\boldsymbol{\lambda}_f = (\lambda_{ij})$ with $\lambda_{ij} = 0$ for $i \notin J_f$, and we apply a control strategy $\boldsymbol{\ell}(v_e)$ as a row $\boldsymbol{u} = (u_j)_{j \in J_f}$ representing an adapted stochastic processes $u_j(t) = \ell_j\left(t, v_e^{t]}\right)$ with real-values $u_j(t, \upsilon)$. Both matrices $\boldsymbol{\lambda}_e$, $\boldsymbol{\lambda}_f$ may depend on t, but they are always orthogonal such that $\boldsymbol{\lambda}^\dagger\boldsymbol{\lambda} = \boldsymbol{\lambda}_e^\dagger\boldsymbol{\lambda}_e + \boldsymbol{\lambda}_f^\dagger\boldsymbol{\lambda}_f$, where $\boldsymbol{\lambda} = \boldsymbol{\lambda}_e + \boldsymbol{\lambda}_f$ and $\boldsymbol{\lambda}^\dagger$ is transposed to complex conjugated matrix $\boldsymbol{\lambda}^* = \left(\lambda_{ij}^*\right)$.

Let the free dynamics of the quantum system be described by a quadratic Hamiltonian $\mathrm{H}^\circ = \frac{1}{2}\mathrm{s}_\circ\boldsymbol{\omega}\mathrm{s}_\circ^\intercal$ for a real symmetric $m \times m$ matrix $\boldsymbol{\omega}$. We now separate the controlling amplitude $\mathbf{u}(t)\,\boldsymbol{b}_f \mathrm{d}t$ in the additive superposition with the control noise $\hbar\,\mathsf{Im}\left(\mathrm{dA}_\circ^+\right)\boldsymbol{b}_f$ coming from the feedback channel given by the matrix $\boldsymbol{b}_f = 2\,\mathrm{Re}\,\boldsymbol{\lambda}_f$ on the row $\mathbf{u} = (u_j \mathrm{I})_{j \in J_f}$, so that the Hamiltonian in (51) is modelled by

$$\mathrm{H}^u(t) = \frac{1}{2}\mathrm{s}_\circ\boldsymbol{\omega}\mathrm{s}_\circ^\intercal + \mathbf{u}(t)\,\boldsymbol{b}_f\mathrm{s}_\circ^\intercal. \tag{100}$$

Using CCR's (99) and assuming that the QS dynamics is purely diffusive and nondemolition (no scattering, $\mathrm{S}_k^i = \mathrm{I}\delta_k^i$ in (33 with $J_\bullet = J$), we can easily evaluate the controlled Lindblad generator $\kappa^u(t, \mathrm{s}_i)$ on the generating operators $\mathrm{s}_i = \mathrm{s}_i^\dagger$ for controlled CP dynamics $\mu_0^u(t, \mathrm{X})$ on the object algebra $\mathbb{A}^\circ$, arising from time dependent quadratic Hamiltonian (100). Substituting $\mathrm{Y} = \mathrm{s}_j$ into the decomposed generator (51) with linear L^j in s_j and zero jump part, $\rho_\nu^\mu(\mathrm{Y}) = 0$, we obtain $\kappa^u(t, \mathbf{s}_i)$ as linear transformation of $\mathrm{s}_\circ = (\mathrm{s}_i)$ written in vector form as

$$\kappa^u(t, \mathrm{s}_\circ^\intercal) = \boldsymbol{j}(\boldsymbol{\omega} + \hbar\,\mathrm{Im}\left(\boldsymbol{\lambda}^\dagger\boldsymbol{\lambda}\right))\mathrm{s}_\circ^\intercal - \boldsymbol{c}_f\mathbf{u}(t)^\intercal$$

where $\boldsymbol{c}_f^\intercal = \boldsymbol{b}_f\boldsymbol{j}$. From this, we can obtain the coherent controlled quantum Langevin equation (37) for $\breve{\mathrm{X}} = \mathrm{s}_i$ with $\mathbf{J}_{\breve{\mathrm{X}}}^t = \mathrm{s}_i\mathbf{1}$, given by the QS germ $\boldsymbol{\gamma}(t) = \mathring{\beta}(t^+)$ for the Heisenberg controlled evolution of the row-operator $\hat{\mathrm{s}}_\circ(t) = \beta(t, \mathrm{s}_\circ)$ in the linear form

$$\mathrm{d}\hat{\mathrm{s}}_\circ^t + \left(\hat{\mathrm{s}}_\circ(t)\,\mathbf{a}^\intercal + \widehat{\boldsymbol{u}}(t)\,\boldsymbol{c}_f^\intercal\right)\mathrm{d}t = \mathrm{d}\mathtt{w}_\circ^t \tag{101}$$

derived in.[24] Here $\mathbf{a}^\intercal = (\hbar\,\mathrm{Im}(\boldsymbol{\lambda}^\intercal\boldsymbol{\lambda}^*) + \boldsymbol{\omega})\boldsymbol{j}$ for $\boldsymbol{\lambda} = \boldsymbol{\lambda}_e + \boldsymbol{\lambda}_f$, the feedback control is given by the row $\widehat{\boldsymbol{u}} = (\widehat{u}_j)_{j \in J_f}$ of Heisenberg operators $\widehat{u}_j(t) =$

$\beta(t, \mathrm{y}_j(t))$ defined by a control law $\mathbf{y}(t) = (\mathrm{y}_j)_{j\in J_f} = \boldsymbol{\ell}(t, \mathrm{v}_e)$, with the total quantum noise $\mathtt{w}_\mathrm{o} = \mathtt{I} \otimes \mathrm{w}_\mathrm{o}$ given by

$$\mathrm{w}_\mathrm{o} := 2\hbar\, \mathsf{Im}\left[\mathrm{A}^+_\bullet \boldsymbol{\lambda}\right] \boldsymbol{j} = \mathrm{u}_\mathrm{o} + \mathrm{v}_\mathrm{o}, \tag{102}$$

where $\mathrm{u}_\mathrm{o} = \mathrm{u}_\bullet \boldsymbol{c}^\intercal$ is total Langevin force and $\mathrm{v}_\mathrm{o} = \mathrm{v}_\bullet \boldsymbol{f}^\intercal$ is total Wiener noise given by left action of matrices

$$\boldsymbol{c}^\intercal = 2\,\mathrm{Re}\,\boldsymbol{\lambda}\boldsymbol{j}, \quad \boldsymbol{f}^\intercal = \hbar\,\mathrm{Im}\,\boldsymbol{\lambda}\boldsymbol{j}, \tag{103}$$

on rows $\mathrm{u}_\bullet = \hbar\,\mathsf{Im}\left(\mathrm{A}^+_1, \ldots, \mathrm{A}^+_d\right)$, $\mathrm{v}_\bullet = 2\,\mathsf{Re}\left(\mathrm{A}^+_1, \ldots, \mathrm{A}^+_d\right)$ of all quantum Langevin forces and all conjugate Wiener noises respectively, coming both from the estimation and feedback channels.

5.1. *Quantum linear continuous observation*

The output observable process is given by the row $\mathtt{y}_e = (\mathtt{y}_j)_{j\in J_e}$ of observables $\mathtt{y}_j(t) = \beta(t, \mathrm{v}_j) \equiv \hat{\mathrm{v}}_j(t)$ in estimation channel. It satisfies the linear in $\hat{\mathrm{s}}_\mathrm{o}(t)$ equation (56):

$$\mathrm{d}\mathtt{y}^t_e = \hat{\mathrm{s}}_\mathrm{o}(t)\, \boldsymbol{b}^\intercal_e \mathrm{d}t + \mathrm{d}\mathrm{v}^t_e \tag{104}$$

where $\boldsymbol{b}_e = 2\,\mathrm{Re}\,\boldsymbol{\lambda}_e$ and the quantum measurement noise is given by the row $\mathrm{v}_e = (\mathrm{v}_j)_{j\in J_e}$ of $\mathrm{v}_j = \mathtt{I}^\mathrm{o} \otimes \mathrm{v}_j$ having Gaussian independent increments on each measurement channel with zero mean and standard variance given by $d_e \times d_e$ identity matrix $\mathbf{1}_e$. Considered alone, this noise represents the standard d_e-dimensional classical Wiener process which we measure in the field after interaction with the quantum object by the output isomorphic transformation $\mathtt{y}_e(t) = \widehat{\mathrm{v}}_e(t)$ of $\mathrm{v}_e = 2\,\mathsf{Re}\,\mathrm{A}^+_e$. However, it does not commute with the quantum Langevin force u_o as

$$\left[\mathrm{u}^\intercal_\mathrm{o}(r), \mathrm{v}_e(s)\right] = (r \wedge s)\, \mathrm{i}\hbar \boldsymbol{c}_e \mathrm{I}, \tag{105}$$

due to the noncommutativity with $\mathrm{v}_e = (\mathrm{v}_j)_{j\in J_e}$ and commutativity with $\mathrm{u}_f = \hbar\,\mathsf{Im}\left[\mathrm{A}^+_f\right]$, resulting form the independence of $\mathrm{A}^+_f = \left(\mathrm{A}^+_j\right)_{j\in J_f}$. The fundamental CCR (105), defined by nonzero (if $\boldsymbol{j} \neq 0$) matrix $\boldsymbol{c}^\intercal_e = \boldsymbol{b}_e \boldsymbol{j}$, was first derived in a complex form in[5,6] and in even more general infinite dimensional setting in.[24] It expresses the Heisenberg error-perturbation uncertainty principle in the standard form

$$\mathrm{d}\mathrm{u}^\intercal_\mathrm{o} \mathrm{d}\mathrm{u}_\mathrm{o} \geq \frac{\hbar^2}{4} \boldsymbol{c}_e \boldsymbol{c}^\intercal_e \mathrm{I} \mathrm{d}t \;\text{ if }\; \mathrm{d}\mathrm{v}^\intercal_e \mathrm{d}\mathrm{v}_e = \mathbf{I}_e \mathrm{d}t \tag{106}$$

($\mathbf{I}_e = \mathbf{1}_e \mathrm{I}$), which was derived by Belavkin in[6,24] as necessary and sufficient condition for nondemolition causality of the observable past $(\mathtt{y}_e(r) : r \leq t)$ and quantum future described by $(\hat{\mathrm{s}}_\mathrm{o}(r) : r \geq t)$.

The error-perturbation Heisenberg principle implies the dependence of v_e and $\mathrm{w}_\mathrm{o} = \mathrm{u}_\mathrm{o} + \mathrm{v}_\mathrm{o}$ which may result in a nonzero covariance matrix $\boldsymbol{f}_e = \hbar \boldsymbol{j} \operatorname{Im}\left(\boldsymbol{\lambda}_e^\dagger\right)$ describing the real part of quantum Itô table

$$\mathrm{dw}_\mathrm{o}^\intercal \mathrm{dv}_e = \left(\boldsymbol{f}_e + \frac{\mathrm{i}\hbar}{2}\boldsymbol{c}_e\right) \mathrm{Id}t \tag{107}$$

as the sum of $\mathrm{du}_\mathrm{o}^\intercal \mathrm{dv}_e = \frac{\mathrm{i}\hbar}{2}\boldsymbol{c}_e \mathrm{d}t$ and $\mathrm{dv}_\mathrm{o}^\intercal \mathrm{dv}_e = \boldsymbol{f}_e \mathrm{d}t$. Note that although each component $\mathrm{w}_j = \mathrm{w}_j^\dagger$ of the row $\mathrm{w}_\mathrm{o} = (\mathrm{w}_1, \ldots, \mathrm{w}_m)$ for vector quantum noise (102) having the independent increments can also be realized as a classical Wiener process, these components mutually do not commute, having complex multiplication table $\mathrm{dw}_\mathrm{o}^\intercal \mathrm{dw}_\mathrm{o} = \hbar^2 \boldsymbol{\lambda}^\dagger \boldsymbol{\lambda} \mathrm{d}t$ with imaginary part defining the commutation relations

$$\left[\mathrm{w}_\mathrm{o}^\intercal(r), \mathrm{w}_\mathrm{o}(s)\right] = (r \wedge s)\, 2\mathrm{i}\hbar^2 \boldsymbol{j}^\intercal \operatorname{Im}\left(\boldsymbol{\lambda}^\dagger \boldsymbol{\lambda}\right) \boldsymbol{j} \mathrm{I}.$$

The symmetrized multiplication $\mathsf{Re}\left[\mathrm{dw}_\mathrm{o}^\intercal \mathrm{dw}_\mathrm{o}\right] = \hbar^2 \operatorname{Re}\left[\boldsymbol{\lambda}^\dagger \boldsymbol{\lambda}\right] \mathrm{d}t$ results in the symmetric covariance

$$\operatorname{Re}\left\langle \mathrm{w}_\mathrm{o}^\intercal(r)\, \mathrm{w}_\mathrm{o}(s)\right\rangle = (r \wedge s)\, \hbar^2 \boldsymbol{j}^\intercal \operatorname{Re}\left(\boldsymbol{\lambda}^\dagger \boldsymbol{\lambda}\right) \boldsymbol{j}^\intercal$$

defined by $\boldsymbol{\lambda}^\dagger \boldsymbol{\lambda} = \boldsymbol{\lambda}_e^\dagger \boldsymbol{\lambda}_e + \boldsymbol{\lambda}_f^\dagger \boldsymbol{\lambda}_f$. It can be parametrized as $\boldsymbol{f}_e^\intercal \boldsymbol{f}_e + \boldsymbol{g}$ with positive matrix

$$\boldsymbol{g} = \frac{\hbar^2}{4}\boldsymbol{c}_e \boldsymbol{c}_e^\intercal + \hbar^2 \boldsymbol{j}^\intercal \operatorname{Re}\left(\boldsymbol{\lambda}_f^\dagger \boldsymbol{\lambda}_f\right) \boldsymbol{j} \tag{108}$$

implying the error-perturbation uncertainty relation

$$\mathsf{Re}\left[\mathrm{dw}_\mathrm{o}^\intercal \mathrm{dw}_\mathrm{o}\right] \geq \left(\boldsymbol{f}_e^\intercal \boldsymbol{f}_e + \frac{\hbar^2}{4}\boldsymbol{c}_e \boldsymbol{c}_e^\intercal\right) \mathrm{Id}t$$

in terms of the total perturbative noise (102) in the linear Langevin equation (101) with respect to the standard normalized error noise in estimation channel (104).

5.2. *Quantum stochastic Gaussian filtering*

Let us denote by $r_\mathrm{o} = (r_i)$ the initial mean vector $\langle \mathtt{s}_\mathrm{o} \rangle$ of the canonical operator-row $\mathtt{s}_\mathrm{o} = (\mathtt{s}_i)$ by the component wise expectations $r_i = \langle \varrho, \mathtt{s}_i \rangle$ and by $\boldsymbol{\Xi}$ the symmetric matrix

$$\boldsymbol{\Xi} := \left(\frac{1}{2}\langle \varrho, \mathtt{s}_i \mathtt{s}_k + \mathtt{s}_k \mathtt{s}_i \rangle - r_i r_k\right)$$

of real covariances $\Xi_{ik} = \mathrm{Re}\langle \varrho, \mathtt{s}_i \mathtt{s}_k \rangle - r_i r_k$. Note that a quantum Gaussian state, like in the classical case, is completely determined by the row-vector r_o and the matrix $\boldsymbol{\Xi} = (\Xi_{ik})$, however, given an antisymmetric matrix $\boldsymbol{j}$, the matrix $\boldsymbol{\Xi}$ should not be just positive-definite, it must satisfy the Heisenberg uncertainty relation $2\boldsymbol{\Xi} \geq \pm \mathrm{i}\hbar \boldsymbol{j}$, preventing zero covariances in the case of nondegenerate $\boldsymbol{j}$.

As it was shown by Belavkin in[11,40] and in infinite dimensions in,[24] the filtering equation (78) preserves the Gaussianity of the posterior state,[24] so the posterior mean $\widehat{s}_\mathrm{o}(t) = \langle \varrho_\mathrm{x}^t, \mathtt{s}_\mathrm{o} \rangle$ of $\mathtt{s}_\mathrm{o} = \mathrm{I}^t \otimes \mathtt{s}_\mathrm{o}$ and the posterior matrix $\boldsymbol{\Sigma}(t)$ of symmetric error covariances given by the real part of

$$\left\langle \varrho_\mathrm{x}^t, \left(\mathtt{s}_i - \widehat{s}_i^t\right)\left(\mathtt{s}_k - \widehat{s}_k^t\right)\right\rangle = \langle \varrho_\mathrm{x}^t, \mathtt{s}_i \mathtt{s}_k \rangle - \widehat{s}_i^t \widehat{s}_k^t$$

form a set of sufficient coordinates for the quantum LQG system which agree with the initial mean and covariance for $\varrho_\mathrm{x}^0 = \varrho$. Applying (78) to the first and second symmetrized moments of $\mathtt{s}_\mathrm{o}$, provides posterior expectations of these sufficient coordinates for diffusive non-demolition measurement of the output operators y_e. These can be found as solutions to Belavkin's Kalman filter equation[6,24] written in vector form as

$$\mathrm{d}\widehat{s}_\mathrm{o}^t + (\widehat{s}_\mathrm{o}(t)\,\mathbf{a}^\intercal + \widehat{\boldsymbol{u}}(t)\,\boldsymbol{c}_f^\intercal)\mathrm{d}t = \mathrm{d}\widetilde{\mathtt{y}}_e^t \boldsymbol{k}^\intercal(t) \tag{109}$$

with initial condition $\widehat{s}_\mathrm{o}^0 = r_\mathrm{o}$ for the posterior mean,

$$\boldsymbol{k}(t) = \boldsymbol{\Sigma}(t)\boldsymbol{b}_e^\intercal + \boldsymbol{f}_e, \ \mathrm{d}\widetilde{\mathtt{y}}_e^t = \mathrm{d}\mathtt{y}_e^t - \widehat{s}_\mathrm{o}(t)\,\boldsymbol{b}_e^\intercal \mathrm{d}t \tag{110}$$

and for the symmetric error covariance we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\Sigma} = \boldsymbol{g} - \boldsymbol{\Sigma}\mathbf{a}_e^\intercal - \mathbf{a}_e\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{b}_e^\intercal \boldsymbol{b}_e \boldsymbol{\Sigma}, \tag{111}$$

$$\mathbf{a}_e = \mathbf{a} + \boldsymbol{f}_e \boldsymbol{b}_e, \quad \boldsymbol{\Sigma}(0) = \boldsymbol{\Xi}. \tag{112}$$

6. Optimal Quantum LQG Feedback Control

We aim to control an *output* QS linear evolution $\mathbf{z}_f(t) = \beta(t, \mathtt{y}_f)$ of a d_f-dimensional linear combination $\mathtt{y}_f = \mathtt{s}_\mathrm{o} \boldsymbol{e}_f^\intercal$, where $\boldsymbol{e}_f$ is a real, in general time dependent $d_f \times m$ matrix, represented in the Heisenberg picture as a row $\mathbf{z}_f = \hat{\mathtt{s}}_\mathrm{o} \boldsymbol{e}_f^\intercal = (\mathbf{z}_j)_{j \in J_f}$ by forcing each classical input process $u_j(t)$ to follow the quantum output $\mathbf{z}_j(t)$ causally controlled by the feedback force $\{\boldsymbol{u}(r) \,|r \leq t\}$ whilst constraining, for energy considerations, a positive quadratic functional of phase space operators $\hat{\mathtt{s}}_\mathrm{o}(t) = \beta(t, \mathtt{s}_\mathrm{o})$. Thus, our control objectives and restraints can be described by the general quadratic operator-valued risk (85) in the canonical form

$$\mathtt{C}(\boldsymbol{u}) = (\mathbf{u} - \mathtt{y}_f)(\mathbf{u} - \mathtt{y}_f)^\intercal + \mathtt{s}_\mathrm{o} \boldsymbol{h} \mathtt{s}_\mathrm{o}^\intercal \tag{113}$$

and $\mathsf{S} = \mathsf{s}_{\rm o}\boldsymbol{\Omega}\mathsf{s}_{\rm o}^{\intercal}$ for positive real symmetric $m \times m$ matrices $\boldsymbol{\Omega}$,$\boldsymbol{h}$.

Since the posterior expectation $\widehat{s}_{\rm o}(t)$ and symmetric covariance $\boldsymbol{\Sigma}(t)$ completely determine the quantum Gaussian posterior state $\varrho_{\rm x}^t$, they form a set of sufficient coordinates, $r_{\rm o} = \widehat{s}_{\rm o}(t)$ and $\boldsymbol{\Xi} = \boldsymbol{\Sigma}(t)$. So we may consider the variations of $\mathsf{S}(t,\varrho)$ only with respect to r and $\boldsymbol{\Xi}$ by partial derivatives of $\mathsf{S}(t, r_{\rm o}, \boldsymbol{\Xi})$

$$\begin{aligned}&\langle \delta\varrho, \nabla_\varrho \mathsf{S}(t,\varrho)\rangle && {\rm d}r_{\rm o}\partial_{\rm o}^{\intercal}\mathsf{S}(t,r_{\rm o},\boldsymbol{\Xi}) + \big(\partial^{\rm oo}\mathsf{S}(t,r_{\rm o},\boldsymbol{\Xi}), {\rm d}\boldsymbol{\Xi}\big)\\ &+\frac{1}{2}\langle \delta\varrho\otimes\delta\varrho, \nabla_\varrho^{\otimes 2}\mathsf{S}(t,\varrho)\rangle &= &+\frac{1}{2}\big(\partial_{\rm o}^{\intercal}\partial_{\rm o}\mathsf{S}(t,r_{\rm o},\boldsymbol{\Xi}), {\rm d}r_{\rm o}^{\intercal}{\rm d}r_{\rm o}\big).\end{aligned}$$

We use the notation $(\cdot,\cdot)$ to denote the matrix trace inner product $\big(\boldsymbol{d},\boldsymbol{f}\big) = {\rm Tr}[\boldsymbol{d}^{\intercal}\boldsymbol{f}]$ on the vector space of matrix configurations for the multi-dimensional system. This gives the directional derivatives along ${\rm d}r_j$ and ${\rm d}\Xi_{ik}$ as linear functions of the column $\partial_{\rm o}^{\intercal}\mathsf{S}$ of partial derivatives $\partial^i\mathsf{S} = \partial\mathsf{S}/\partial r_i$ and the matrices

$$\partial^{\rm oo}\mathsf{S} = \left(\frac{\partial}{\partial\Xi_{ik}}\mathsf{S}\right), \qquad \partial_{\rm o}^{\intercal}\partial_{\rm o}\mathsf{S} = \left(\frac{\partial}{\partial r_i}\frac{\partial}{\partial r_k}\mathsf{S}\right)$$

which are evaluated at $(\widehat{s}_{\rm o}(t), \boldsymbol{\Sigma}(t))$.

Inserting this parametrization into the Bellman equation (95) and minimizing gives the optimal control strategy

$$\widehat{\boldsymbol{\ell}}(t) = \frac{1}{2}\partial^{\rm o}\mathsf{S}\left(t, \widehat{s}_{\rm o}(t), \boldsymbol{\Sigma}(t)\right)\boldsymbol{c} + \widehat{s}_{\rm o}(t)\,\boldsymbol{e}_f^{\intercal} \tag{114}$$

where $\mathsf{S}(t, r_{\rm o}, \boldsymbol{\Xi})$ at $r_{\rm o} = \widehat{s}_{\rm o}(t)$, $\boldsymbol{\Xi} = \boldsymbol{\Sigma}(t)$ now satisfies the non-linear partial differential equation

$$\begin{aligned}-\tfrac{\partial}{\partial t}\mathsf{S}(t,r_{\rm o},\boldsymbol{\Xi}) = {}& \tfrac{1}{2}(r_{\rm o}\mathbf{a}^{\intercal}\partial_{\rm o}^{\intercal}\mathsf{S} + \partial_{\rm o}\mathsf{S}\mathbf{a}r_{\rm o}^{\intercal}) + r_{\rm o}\boldsymbol{h}r_{\rm o}^{\intercal}\\ &+\langle\partial^{\rm oo}\mathsf{S}, \boldsymbol{g} - \mathbf{a}\boldsymbol{\Xi} + \boldsymbol{\Xi}\mathbf{a}^{\intercal}\rangle + (\boldsymbol{h},\boldsymbol{\Xi})\\ &-(\tfrac{1}{2}\partial_{\rm o}\mathsf{S}\boldsymbol{c} + r_{\rm o}\boldsymbol{e}_f^{\intercal})(\tfrac{1}{2}\partial_{\rm o}\mathsf{S}\boldsymbol{c} + r_{\rm o}\boldsymbol{e}_f^{\intercal})^{\intercal}\\ &+\tfrac{1}{2}\left(\partial_{\rm o}^{\intercal}\partial_{\rm o}\mathsf{S} - \partial^{\rm oo}\mathsf{S}, (\boldsymbol{\Xi}\boldsymbol{b}^{\intercal} + \boldsymbol{f}_e)(\boldsymbol{\Xi}\boldsymbol{b}^{\intercal} + \boldsymbol{f}_e)^{\intercal}\right)\end{aligned} \tag{115}$$

which is the Hamilton-Jacobi-Bellman equation for this example.

6.1. *The microduality of quantum LQG problem*

It is well known from classical control theory that LQG control has a minimum cost-to-go which is quadratic in the state, so we try the candidate solution

$$\mathsf{S}(t, r_{\rm o}, \boldsymbol{\Xi}) = r_{\rm o}\boldsymbol{\Omega}(t)r_{\rm o}^{\intercal} + (\boldsymbol{\Omega}(t), \boldsymbol{\Xi}) + \mathsf{s}(t)$$

in the HJB equation (115). This separates the HJB equation into a set of coupled ordinary differential equations and gives the optimal feedback control strategy

$$\widehat{\boldsymbol{u}}\,(t) = \widehat{s}_{\rm o}\,(t)\,\boldsymbol{l}^{\intercal}(t),\ \boldsymbol{l}^{\intercal}(t) = \boldsymbol{\Omega}(t)\boldsymbol{c}_f + \boldsymbol{e}_f^{\intercal} \tag{116}$$

which is linear in the solution $\widehat{s}_{\rm o}$ to the filtering equation at time t where $\boldsymbol{\Omega}(t)$ satisfies the matrix Riccati equation

$$\begin{aligned} -\frac{\rm d}{{\rm d}t}\boldsymbol{\Omega} &= \boldsymbol{h} - \boldsymbol{\Omega}\mathbf{a}_f - \mathbf{a}_f^{\intercal}\boldsymbol{\Omega} - \boldsymbol{\Omega}\boldsymbol{c}_f\boldsymbol{c}_f^{\intercal}\boldsymbol{\Omega} \\ \mathbf{a}_f &= \mathbf{a} + \boldsymbol{c}_f\boldsymbol{e}_f,\ \ \boldsymbol{\Omega}(T) = \boldsymbol{\Omega} \end{aligned} \tag{117}$$

and the function $\mathsf{s}(t)$ with $\mathsf{s}(T) = 0$ is given by

$$-\frac{\rm d}{{\rm d}t}\mathsf{s}(t) = \Big((\boldsymbol{\Omega}(t)\boldsymbol{c}_f + \boldsymbol{e}_f^{\intercal})(\boldsymbol{\Omega}(t)\boldsymbol{c}_f + \boldsymbol{e}_f^{\intercal})^{\intercal}, \boldsymbol{\Sigma}(t)\Big) + (\boldsymbol{\Omega}(t), \boldsymbol{g})\,.$$

Integrating this we obtain the total minimal cost for the control experiment

$$\begin{aligned} \mathsf{S}(0, r_{\rm o}, \boldsymbol{\Xi}) = r_{\rm o}\boldsymbol{\Omega}_0 r_{\rm o}^{\intercal} + {\rm Tr}[\boldsymbol{\Omega}_0\boldsymbol{\Sigma}_0] + \int_0^T {\rm Tr}[\boldsymbol{\Omega}(t)\boldsymbol{g}]{\rm d}t \\ + \int_0^T {\rm Tr}[(\boldsymbol{\Omega}(t)\boldsymbol{c}_f + \boldsymbol{e}_f^{\intercal})^{\intercal}\boldsymbol{\Sigma}(t)(\boldsymbol{\Omega}(t)\boldsymbol{c}_f + \boldsymbol{e}_f^{\intercal})]{\rm d}t \end{aligned} \tag{118}$$

where $\Sigma_0 = \Xi$ and $\boldsymbol{\Omega}_0$ is the solution to (117) at time $t = 0$.

The equations (110)-(111) and (116)-(117) demonstrate the intrinsic duality between optimal quantum linear filtering and optimal classical linear control, which we call *microduality*. To make this duality more transparent let us introduce real matrices $\boldsymbol{e}$ and $\boldsymbol{f}$ defining the matrices $\boldsymbol{f}_e$ and $\boldsymbol{e}_f$ in (103) and (113) by $\boldsymbol{e}\boldsymbol{j} = \boldsymbol{f}_e^{\intercal}$ and $\boldsymbol{e}_f\boldsymbol{j} = \boldsymbol{f}^{\intercal}$, similar to $\boldsymbol{b}\boldsymbol{j} = \boldsymbol{c}_e^{\intercal}$ and $\boldsymbol{b}_f\boldsymbol{j} = \boldsymbol{c}^{\intercal}$ in terms of the estimation and feedback channel matrices $\boldsymbol{b} = \boldsymbol{b}_e$ and $\boldsymbol{c} = \boldsymbol{c}_f$. Then the microduality can be summarized in the table

Filt	$\mathbf{a}\boldsymbol{j}$	$\boldsymbol{b}\boldsymbol{j}$	$\boldsymbol{e}\boldsymbol{j}$	$\boldsymbol{k}\,(T-t)$	$\boldsymbol{g}\,(T-t)$	$\boldsymbol{\Sigma}\,(T-t)$
Con	$\boldsymbol{j}\mathbf{a}^{\intercal}$	$\boldsymbol{c}^{\intercal}$	$\boldsymbol{f}^{\intercal}$	$\boldsymbol{j}\boldsymbol{l}^{\intercal}\,(t)$	$\boldsymbol{j}\boldsymbol{h}\,(t)\,\boldsymbol{j}^{\intercal}$	$\boldsymbol{j}\boldsymbol{\Omega}\,(t)\,\boldsymbol{j}^{\intercal}$

(119)

in which the duality notations are made in filtering-control alphabetical order ($\boldsymbol{b}$,$\boldsymbol{c}$), ($\boldsymbol{e}$,$\boldsymbol{f}$), ($\boldsymbol{g}$,$\boldsymbol{h}$) and ($\boldsymbol{k}$,$\boldsymbol{l}$) and if the matrices $\mathbf{a}$, $\boldsymbol{b}$,$\boldsymbol{e}$ depend on t, they should be also taken at $t^{\intercal} = T - t$ for the duality with $\mathbf{a}^{\intercal}, \boldsymbol{c}^{\intercal}, \boldsymbol{f}^{\intercal}$ evaluated at t. This allows us to formulate and solve the dual classical control problem given the solution to quantum filtering problem with dual parameters. The microduality can be understood when we examine the nature of each of the methods used. Both methods involve the minimization of a quadratic function for linear, Gaussian systems, (i.e. the least squares error

for filtering and the quadratic cost for control). The time reversal in the dual picture is explained by the interchange of the input (feedback) and the output (estimation) channels together with the linear canonical transformation given by the symplectic matrix $\boldsymbol{j}$. Note that $\boldsymbol{e} = \hbar \operatorname{Im} \boldsymbol{\lambda}_e$, and $\boldsymbol{g}$ must be Hermitian-positive, satisfying the relation (108) due to Heisenberg uncertainty relation corresponding to the error-perturbation CCR's $[\mathrm{d}\boldsymbol{u}_{\mathrm{o}}^{\intercal}, \mathrm{dv}_e] = \mathrm{i}\hbar \boldsymbol{c}_e \mathrm{d}t\mathtt{I}$ in the Itô multiplication table (107). In order to complete this filtering-control microduality we may also set $\boldsymbol{f}^{\intercal} = \hbar \operatorname{Im} \boldsymbol{\lambda}_f \boldsymbol{j}$ to have the relation between $\boldsymbol{e}$ and $\boldsymbol{f}$ similar to the duality of $\boldsymbol{b} = 2 \operatorname{Re} \boldsymbol{\lambda}_e$ and $\boldsymbol{c}^{\intercal} = 2 \operatorname{Re} \boldsymbol{\lambda}_f^* \boldsymbol{j}$, and also assume that matrix $\boldsymbol{h}$ is also Hermitian-positive, satisfying

$$\boldsymbol{h} = (\hbar/2)^2 \boldsymbol{b}_f^{\intercal} \boldsymbol{b}_f + \hbar^2 \operatorname{Re} \left(\boldsymbol{\lambda}_e^{\dagger} \boldsymbol{\lambda}_e \right), \tag{120}$$

where $\boldsymbol{b}_f = 2 \operatorname{Im} \boldsymbol{\lambda}_f$. Note that although the condition (120) is not a requirement in this classical-quantum setting, in which only positivity of the combination $\boldsymbol{h} + \boldsymbol{e}_f^{\intercal} \boldsymbol{e}_f$ suffices.

6.2. *Example: Quantum particle feedback control*

We can now apply the obtained results to demonstrate optimal quantum filtering and optimal feedback control and their microduality on the model example considered in the introduction. The optimal estimates of the quantum particle position and momentum based on the nondemolition observation of free controlled quantum particle via the continuous measurement of $\mathtt{y}_e$ when a control strategy $\widehat{\boldsymbol{u}}(t) = \ell\left(t, \mathtt{y}_{t]}\right)$ is in place are given by quantum Kalman filter (109) in the form of a pair of the linear stochastic equations

$$\mathrm{d}\widehat{q}^t + \widehat{q}(t)\, \lambda \mathrm{d}t = \frac{1}{\mu} \widehat{p}(t)\, \mathrm{d}t + (\beta \sigma_q(t) - \varepsilon)\, \mathrm{d}\widetilde{\mathtt{y}}_e^t \tag{121}$$

$$\mathrm{d}\widehat{p}^t + \widehat{p}(t)\, \lambda \mathrm{d}t = \gamma \widehat{\boldsymbol{u}}(t)\, \mathrm{d}t + \beta \sigma_{qp}(t)\, \mathrm{d}\widetilde{\mathtt{y}}_e^t. \tag{122}$$

with $\widehat{q}(0) = r_q \mathtt{I}$, $\widehat{p}(t) = r_p \mathtt{I}$ (They were originally derived in[11,38] in the absence of a control channel $\gamma = 0$.) Here the estimation innovation process $\tilde{\mathtt{y}}_e$ describes the gain of information due to measurement of the output process $\mathtt{y}_e$ given by $\mathrm{d}\widetilde{\mathtt{y}}_e^t = \mathrm{d}\mathtt{y}_e^t - \beta \widehat{q}(t)\, \mathrm{d}t$, and the error covariances satisfy the Riccati equations

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\sigma_q &= \zeta_q + 2\left(\tfrac{1}{\mu}\sigma_{qp} + \sigma_q \delta\right) - (\beta \sigma_q)^2 \\ \tfrac{\mathrm{d}}{\mathrm{d}t}\sigma_{qp} &= \tfrac{1}{\mu}\sigma_p - (\lambda - \delta)\, \sigma_{qp} - \beta^2 \sigma_q \sigma_{qp} \\ \tfrac{\mathrm{d}}{\mathrm{d}t}\sigma_p &= \zeta_p - 2\lambda \sigma_p - (\beta \sigma_p)^2, \end{aligned} \tag{123}$$

where we denote $\delta = \frac{1}{2}(\beta\varepsilon - \varphi\gamma)$, with initial conditions

$$\sigma_q(0) = \sigma_q, \quad \sigma_{qp}(0) = \sigma_{qp}, \quad \sigma_p(0) = \sigma_p.$$

The Riccati equations for the error covariance in the filtered free particle dynamics have an exact solution[38] with profound implications for the ultimate quantum limit satisfying the Heisenberg uncertainty relations for the accuracy of optimal quantum state estimation via the continuous indirect quantum particle coordinate measurement.

The dual optimal control problem can be found by identifying the corresponding dual matrices from the table (119) which give the quantum control parameters of the quadratic cost operators

$$\begin{aligned} \mathtt{C}(u) &= (u\mathtt{I} - \mathtt{y})^2 + \eta_q \mathtt{q}^2 + \eta_p \mathtt{p}^2, \\ \mathtt{S} &= \omega_q \mathtt{q}^2 + \omega_{qp}(\mathtt{pq} + \mathtt{qp}) + \omega_p \mathtt{p}^2 \end{aligned} \tag{124}$$

corresponding to the dual output process given by $\mathtt{y} = \varphi\mathtt{p}$. For the linear Gaussian system (116) gives the optimal control strategy

$$\widehat{u}(t) = (\omega_{pq}(t)\,\widehat{p}(t) + \omega_p(t)\,\widehat{q}(t))\gamma \tag{125}$$

where the coefficients $\omega(t)$ are the solutions to the Riccati equations

$$\begin{aligned} -\tfrac{\mathrm{d}}{\mathrm{d}t}\omega_q(t) &= \eta_q - 2\lambda\sigma_q - (\gamma\sigma_q)^2 \\ -\tfrac{\mathrm{d}}{\mathrm{d}t}\omega_{qp}(t) &= \tfrac{1}{\mu}\omega_q - (\lambda + \delta)\,\omega_{qp} - \gamma^2\omega_p\omega_{qp} \\ -\tfrac{\mathrm{d}}{\mathrm{d}t}\omega_p(t) &= \eta_p + 2\left(\tfrac{1}{\mu}\omega_{qp} - \omega_p\delta\right) - (\gamma\omega_p)^2 \end{aligned} \tag{126}$$

with the terminal conditions

$$\omega_p(T) = \omega_p, \quad \omega_{qp}(T) = \omega_{qp}, \quad \omega_q(T) = \omega_q.$$

Note that as seen in this example the microduality table (119) requires not only an identification of the dual matrices by transposition and time reversal, but also symplectic interchange the quantum phase variables $(\mathtt{q}, \mathtt{p}) \leftrightarrow (\mathtt{p}, \mathtt{q})$. This is because the matrix of coefficients $\mathbf{a}$ is non-symmetric and nilpotent, so it is dual to its transpose only when we interchange the variables in the dual picture. Thus the optimal coefficients $\{\omega_p, \omega_{qp}, \omega_q\}(t)$ in the quadratic cost-to-go correspond to the minimal error covariances $\{\sigma_q, \sigma_{qp}, \sigma_p\}(T-t)$ in the dual picture. This demonstrates the linear microduality in the following specified form of the table (119)

Filtering $\mathtt{q}$	$\lambda - \mu^{-1}$	β	ε	$\boldsymbol{j}^{\intercal}\boldsymbol{k}$	$\boldsymbol{g}\boldsymbol{j}$	$\boldsymbol{\Sigma}\boldsymbol{j}$
Control $\mathtt{p}$	$\lambda - \mu^{-1}$	γ	φ	$\boldsymbol{l}^{\intercal}$	$\boldsymbol{j}\boldsymbol{h}$	$\boldsymbol{j}\boldsymbol{\Omega}$

showing the complete symmetry under the time reversal and exchange of (q, p), in which the coordinate observation is seen as completely dual to the feedback of momentum.

The minimal total cost for this linear quantum feedback experiment on the interval $[0, T]$ can be obtained by substitution of these solutions into (118):

$$\begin{aligned}\mathsf{S} &= \omega_q(\sigma_q(0) + r_q^2) + 2\omega_{qp}(\sigma_{qp}(0) + r_q r_p)\\ &\quad +\omega_p(0)(\sigma_p(0) + r_p^2) + \int_0^T (\hbar^2\omega_p(t) + \omega_{pq}^2(t)\sigma_q(t))\mathrm{d}t\\ &\quad + \int_0^T (\omega_p^2(t)\sigma_p(t) + 2\omega_{qp}(t)\omega_p(t)\sigma_{pq}(t))\mathrm{d}t.\end{aligned} \tag{127}$$

7. Discussion

The Bellman equations for quantum systems having separate diffusive and counting measurement schemes have been derived in continuous time under a general setup. This presents original derivations of the results stated in[11] (a derivation of the diffusive case has also appeared in[62]) and allows us to reformulate the optimal control problem for a fundamentally unobservable quantum system into a classical control problem on the Banach space of observable filtered states.

The multi-dimensional quantum LQG problem demonstrates the first application of the general quantum Bellman equation from which one can obtain the special cases of the Gaussian quantum oscillator[36] and quantum free particle.[38] Note that the fundamental difference between this example and the corresponding well studied classical case is in the observability of the system. The quantum noises introduced act only to account for the quantum backaction due to the incompatibility of quantum events. No further restrictions on the observability or additional classical noise are introduced. As such, the corresponding classical problem (when $\hbar \to 0$) admits direct observations of commuting operators $\hat{\mathrm{s}}_{\mathrm{o}}(t) = s_{\mathrm{o}}(t)$ for a deterministic classical system and has an optimal direct feedback strategy $\boldsymbol{u}(t) = s_{\mathrm{o}}(t)\boldsymbol{l}(t)$ and the minimal cost $\mathsf{S}(0, s_{\mathrm{o}}) = s_{\mathrm{o}}\boldsymbol{\Omega}_0 s_{\mathrm{o}}^{\intercal}$. Also this example clearly demonstrates the micro duality between quantum linear filtering and classical feedback control as a more elaborated duality involving also the linear symplectic transformation $\boldsymbol{j}$.

During the publication of this paper, there have appeared a number of recent works on quantum filtering and feedback control by Bouten et al.[42–44] (See also in this volume.)

8. Appendix

8.1. *A. Some definitions and facts on W*-algebras*

(1) A complex Banach algebra A with involution $a \mapsto a^*$ such that $\|a^*a\| = \|a\|^2$ is called C*-algebra, and W*-algebra if it is dual to a linear subspace $L \subseteq \mathbb{A}^\star$ (called preadjoint of $\mathbb{A} = L^\star$ if it is closed, denoted as $L = \mathbb{A}_\star$). They all can be realized as operator algebras $\mathcal{A} = \{\mathrm{A} : \mathcal{H} \to \mathcal{H}\}$ on a complex Hilbert space $\mathcal{H}$ with involution given by the Hermitian conjugation $\mathrm{A} \mapsto \mathrm{A}^\dagger$, and an operator W*-algebra is called von Neumann algebra if its unit is the identity operator I in $\mathcal{H}$. The simplest example of W*-algebra is the von Neumann algebra $\mathcal{B}(\mathcal{H})$ of all bounded operators acting in a complex Hilbert space $\mathcal{H}$. A von Neumann algebra $\mathcal{A}$ is called semisimple if $\mathcal{H}$ has an orthogonal decomposition into invariant subspaces $\mathcal{H}_i$ in which $\mathcal{A}$ is $\mathcal{A}(\mathcal{H}_i)$. Note that the commutant

$$\mathcal{A}' = \{\mathrm{B} \in \mathcal{B}(\mathcal{H}) \,|\, [\mathcal{A}, \mathrm{B}] = 0\},$$

where $[\mathcal{A}, \mathrm{B}] = \{\mathrm{AB} - \mathrm{BA} | \mathrm{A} \in \mathcal{A}\}$, is a von Neumann algebra which is semisimple iff it is commutant of an Abelian algebra $\mathcal{A}$.

(2) Let $\{\mathrm{r}_i\}$ (or $\{\mathcal{A}_i\}$) be a family of self-adjoint operators (operator algebras $\mathcal{A}_i$, $i \in I$) acting in $\mathcal{H}$, e.g. orthoprojectors $\mathrm{P}_i^2 = \mathrm{P}_i = \mathrm{P}_i^\dagger$. The W*-algebra generated by this family is defined as the smallest weakly closed self-adjoint sub-algebra $\mathcal{A} \subseteq \mathcal{A}(\mathcal{H})$ containing these operators (the algebras $\mathcal{A}_i$, $\mathcal{A} = \vee\mathcal{A}_i$), or their spectral projectors $\mathrm{P}_\Delta \in \mathcal{A}$ if r_i are unbounded in $\mathcal{H}$. (In the latter case $\mathrm{r}_i \vdash \mathcal{A}$, i.e. affiliated with $\mathcal{A}$.) Such $\mathcal{A}$ has an identity I, and if I is identified with identity operator, it consists of all bounded operators that commute with the bounded commutant $\mathcal{B} = \cap_i \{\mathrm{r}_i\}'$(or with $\mathcal{B} = \cap_i \mathcal{A}_i'$), i.e., it is the second commutant $\mathcal{A} = \{\mathrm{r}_i | i \in I\}'' = \mathcal{B}'$ of the family $\{\mathrm{r}_i\}$. The latter can be taken as the definition of the von Neumann algebra $\mathbb{C}_\mathrm{r}$ generated by the family $\{\mathrm{r}_i\}$.

(3) A (normal) state on a von Neumann algebra $\mathcal{A}$ is defined as a linear ultraweakly continuous functional $\epsilon : \mathcal{A} \to \mathbb{C}$ of expectations $\epsilon[\mathrm{A}]$ satisfying the positivity and normalization conditions $\epsilon[\mathrm{A}] \geq 0$, $\forall \mathrm{A} \geqq 0$, $\epsilon[\mathrm{I}] = 1$. ($\mathrm{A} \geqq 0$ signifies the Hermit positivity $\langle\psi \mid \mathrm{A}\psi\rangle \geqq 0$ $\forall\psi \in \mathcal{H}$.) It is usually given by a Hermitian-positive trace-one operator R as $\epsilon[\mathrm{A}] := \mathrm{tr}[\mathrm{RA}] \equiv \langle \mathrm{R}, \mathrm{A}\rangle_1$, and is called vector state if $\epsilon[\mathrm{A}] = \langle\psi \mid \mathrm{A}\psi\rangle$ corresponding to $\mathrm{R} = |\psi\rangle\langle\psi| \equiv \mathrm{P}_\psi$. The linear span of all *normal* states is isometric with the preadjoint space $\mathcal{A}_\star$ which can be identified with the space of *density operators* $\mathrm{R} \in \mathcal{A}$ (or affiliated to $\mathcal{A}$: $\mathrm{R} \vdash \mathcal{A}$) with re-

spect to a standard *pairing* $\langle \mathcal{A}_\star, \mathcal{A} \rangle$ extending a real symmetric densely defined form

$$\langle \mathrm{X}, \mathrm{Z} \rangle = \langle \mathrm{X}, \mathrm{Z} \rangle \,, \ \langle \mathrm{X}, \mathrm{Z} \rangle^* = \left\langle \mathrm{X}^\dagger, \mathrm{Z}^\dagger \right\rangle \ \ \forall X, Z \in \mathcal{D} \tag{128}$$

($\mathcal{A}_\star$ is isometric to the space $\mathcal{B}_\star(\mathcal{H})$ of all trace class operators R with respect to the standard trace $\mathrm{tr} = \mathrm{Tr}_{\mathcal{H}}$ only in the simple case $\mathcal{A} = \mathcal{B}(\mathcal{H})$.)

(4) If $\mathcal{A} \subseteq \mathcal{A}(\mathcal{H})$, the density R of a state $\epsilon[\mathrm{X}] = \langle \mathrm{R}, \mathrm{A} \rangle$ may not be a trace class operator, but it is uniquely defined by a Hermitian-positive mass one element, $\langle \mathrm{R}, \mathrm{I} \rangle = 1$, affiliated to $\mathcal{A}$. For semifinite algebras one can always choose a tracial pairing such that $\langle \mathrm{AX}, \mathrm{Z} \rangle = \langle \mathrm{X}, \mathrm{ZA} \rangle$ for all $\mathrm{A} \in \mathcal{A}$. For the general W*-algebra $\mathcal{A}$ there might be no such tracial pairing, but there is always a standard modular pairing,

$$\left\langle \mathrm{X}, \mathrm{ZA}^\dagger \right\rangle = \left\langle \mathrm{A}^\sharp \mathrm{X}, \mathrm{Z} \right\rangle \ \ \forall \mathrm{A} \in \mathcal{A}, \mathrm{X}, \mathrm{Z} \in \mathcal{D}. \tag{129}$$

The antilinear map $\mathrm{A} \mapsto \mathrm{A}^\sharp$ into the left multipliers of $\mathcal{A}_\star$ is densely defined as an involution $\mathrm{A}^{\sharp\sharp} = \mathrm{A}$ in the algebra $\mathcal{A}$. The *modular superoperator* $\delta(\mathrm{A}) = \mathrm{A}^{\sharp\flat}$ is given by the adjoint (right) involution $\mathrm{A}^\flat = \mathrm{A}^{\dagger\sharp\dagger}$ with respect to the Hermitian form $(\mathrm{X}|\mathrm{Z}) = \left\langle \mathrm{X}^\dagger, \mathrm{Z} \right\rangle$. Usually the pairing $\langle \mathrm{R}, \mathrm{A} \rangle$ is determined on a dense domain in $\mathcal{A}$ by a reference state or by a normal faithful semifinite weight $\mu[\mathrm{A}] = \langle \mathrm{I}, \mathrm{A} \rangle$ such that the density R is related to μ by

$$\langle \mathrm{R}, \mathrm{A} \rangle := \mu \left[\mathrm{A} \delta^{\frac{1}{2}}(\mathrm{R}) \right] \equiv \epsilon[\mathrm{A}] \,,$$

where $\delta^{\frac{1}{2}}(\mathrm{R}) = \mathrm{R}^{\sharp\dagger}$. In this case $(\mathrm{Y}|\mathrm{Y}) \geq 0$ for any $\mathrm{Y} \in \mathcal{D}$, but in general the form $(\mathrm{Y}|\mathrm{Y})$ may be defined by an indefinite metric on $\mathcal{D}$. If the algebra $\mathcal{A}$ is semifinite (i.e. there exists a faithful normal semi-finite trace $\mathrm{A} \mapsto \mathrm{trA} \equiv \mu[\mathrm{A}]$ on $\mathcal{A}$), then the normal states can be described by positive selfadjoint unit-trace operators $\mathrm{R} \in \mathcal{A}$ (or $\mathrm{R} \vdash \mathcal{A}$), by means of the tracial pairing $\langle \mathrm{R}, \mathrm{A} \rangle = \mu[\mathrm{AR}]$.

(5) Let $\mathcal{A}$, $\mathcal{B}$ be von Neumann algebras in respective Hilbert spaces $\mathcal{H}_0$ and $\mathcal{H}_1$, and let $\phi : \mathcal{B} \to \mathcal{A}$ be a linear map that transforms the operators $\mathrm{B} \in \mathcal{B}$ into operators $\mathrm{A} \in \mathcal{A}$ (called sometimes superoperator, or operation). The preadjoint map $\Phi = \phi_\star := \phi^\star | \mathcal{A}_\star$ in terms of $(\phi^\star(\mathrm{R}) \mid \mathrm{B}) = (\mathrm{R} \mid \phi(\mathrm{B}))$ to ultraweakly continuous superoperator is called *transfer map* if ϕ is completely positive (CP) in the sense

$$\sum_{i,k=1}^{\infty} \left\langle \psi_i \mid \phi\left(\mathrm{B}_i^\dagger \mathrm{B}_k \right) \psi_k \right\rangle \geqq 0, \qquad \forall \mathrm{B}_j, \psi_j \tag{130}$$

$(i = 1, \ldots, d_e < \infty)$, and unity-preserving: $\phi(\mathrm{I}_1) = \mathrm{I}_0$ (or $\phi(\mathrm{I}_1) \leq \mathrm{I}_0$). The CP condition is obviously satisfied if ϕ is normal homomorphism (or W*-representation) $\jmath : \mathcal{A} \to \mathcal{A}$, which is defined by the additional multiplicativity property $\jmath\left(\mathrm{B}^\dagger \mathrm{B}\right) = \jmath(\mathrm{B})^\dagger \jmath(\mathrm{B})$. The composition $\Phi[\mathrm{R}](\mathrm{B}) = \langle \mathrm{R}, \phi(\mathrm{B}) \rangle_\mu$ of a transfer map Φ with any state $\mathrm{R} \in \mathfrak{s}(\mathcal{A})$ is a state $\mathrm{S} \in \mathfrak{s}(\mathcal{B})$ described by the preadjoint action of the superoperator $\phi = \Phi^\star$ on R:

$$\langle \mathrm{R}, \phi(\mathrm{B}) \rangle = \langle \phi_\star(\mathrm{R}), \mathrm{B} \rangle, \forall \mathrm{B} \in \mathcal{B}, \mathrm{R} \in \mathcal{A}_\star.$$

A transfer map Φ is called spatial if

$$\Phi^\star(\mathrm{B}) = \mathrm{F}\mathrm{B}\mathrm{F}^\dagger, \quad \text{or} \quad \Phi(\mathrm{R}) = \mathrm{F}_\star \mathrm{R} \mathrm{F}_\star^\dagger, \tag{131}$$

where F is a linear coisometric operator $\mathcal{H}_1 \to \mathcal{H}_0$, $\mathrm{F}\mathrm{F}^\dagger = \mathrm{I}_0$ (or $\mathrm{F}\mathrm{I}_1\mathrm{F}^\dagger \leq \mathrm{I}_0$) called the *propagator* and $\mathrm{F}_\star = \mathrm{F}^\sharp$. Every transfer map is in the closed convex hull of spatial transfer maps, but there might be no extreme point in this hull.

8.2. *B. Quantum stochastic calculus in Fock space*

(1) Guichardet Fock space $F_t^s = \Gamma(\mathcal{E}_t^s)$ over the space $\mathcal{E}_t^s = L^2(I_t^s)$ of square integrable complex functions $\xi(r)$ on the interval $I_t^s = (t, t+s]$ is built as the Hilbert sum $\oplus_{n=0}^\infty \mathcal{E}_n$ of the spaces $\mathcal{E}_n = L^2(\Gamma_n)$ of square integrable functions $g_n : \tau \in \Gamma_n \mapsto \mathbb{C}$ of *finite chains* $\tau = \{t_1 < \ldots < t_n\} \subset I_t^s$ for all $n = 0, 1, \ldots$. Here $g_0(\emptyset) = c$ is a constant corresponding to only one – empty chain $\tau_0 = \emptyset$ of Γ_0, with $c = 1$ for the vacuum vector state $g_n = \delta_0^n$. It is $L^2(\Gamma(I_t^s))$, where $\Gamma(I_t^s)$ is disjoint union $\sum_{n=0}^\infty \Gamma_n(I_t^s)$ of the n-simplices $\Gamma_n(I_t^s)$. The integration on Γ is assumed over the Lebesgues sum $\mathrm{d}\tau = \sum \mathrm{d}\tau_n$ of measures $\mathrm{d}\tau_n = \mathrm{d}t_1 \ldots \mathrm{d}t_n$ on the simplices Γ_n of chains $\tau_n : |\tau_n| = n$ with the only atom $\mathrm{d}\tau_0 = 1$ at $\Gamma_0 = \{\emptyset\}$ such that

$$\sum_{n=0}^\infty \int_{\Gamma_n} \left\| \xi^{\otimes n}(\tau_n) \right\|^2 \mathrm{d}\tau_n = \exp\left[\int_0^\infty \|\xi(t)\|^2 \mathrm{d}t \right]$$

for the exponential vector-functions $\xi^{\otimes n}(\tau_n) = \xi(t_1) \cdots \xi(t_n)$ given by a single-point function $\xi \in \mathcal{E}$. It is isomorphic to both usual Fock spaces $\oplus_{n=0}^\infty \mathcal{E}_n^\pm$ of symmetric and antisymmetric functions $g_n(r_1, \ldots, r_n)$ extending $g(\tau_n)$ on $(I_t^s)^n$ with respect to the measure $(n!)^{-1} \mathrm{d}r_1 \cdots \mathrm{d}r_n$. Fock space is infinitely divisible in the multiplicative sense $F_{t-r}^{r+s} \simeq F_{t-r}^r \otimes F_t^s$ for any $r, s > 0$, which is a reflection of the additive divisibility $\mathcal{E}_{t-r}^{r+s} \simeq \mathcal{E}_{t-r}^r \oplus \mathcal{E}_t^s$ of $L^2\left(I_{t-r}^{r+s}\right)$. The generalization to the

multiple Fock-Guichardet case over the space $\mathcal{E}_t^s = L^2(I_t^s \to \mathfrak{k})$ of vector-valued functions $\xi(t)$ in a Hilbert space $\mathfrak{k}$ (say, of columns $\xi = (\xi^j : j \in J)$) is straight forward and can be found in.[10,25] All properties remain the same, and the only difference is that F_0^t is not L^2-space of scalar-valued functions g on Γ but is Hilbert integral $\int_{\tau\in\Gamma(I_t^s)}^{\oplus} \mathcal{K}(\tau)\,\mathrm{d}\tau$ of $\mathcal{K}(\tau) = \mathfrak{k}^{\otimes|\tau|} \equiv \mathcal{K}_n$, which is the Hilbert sum of spaces $L^2(\Gamma_n \to \mathcal{K}_n) = \mathfrak{k}^{\otimes|\tau|} \otimes L^2(\Gamma_n)$ of square integrable tensor-valued functions $f : \Gamma_n(I_t^s) \to \mathfrak{k}^{\otimes n}$.

(2) Four basic *integrators* A_-^+, A_-°, $\mathrm{A}_\circ^+$ and $\mathrm{A}_\circ^\circ$ of the universal quantum stochastic (QS) calculus[10,25] are operator-valued measures $\mathrm{A}_\mu^\nu(I)$ of preservation, annihilation, creation and exchange respectively. They define the basic QS integrals of the total QS integral as sum-integral

$$\mathfrak{i}_0^t(\mathbf{K}) = \sum_{\mu=-,\circ;\nu=\circ,+} \int_0^t \mathrm{K}_\nu^\mu(r)\,\mathrm{A}_\mu^\nu(\mathrm{d}r) \tag{132}$$

of four basic integrands K_+^-, $\mathrm{K}_\circ^-$, K_+° and $\mathrm{K}_\circ^\circ$ as measurable operator-valued functions $\mathrm{K}_\nu^\mu(r)$ in F by the following explicit formulas:

$$\left[\mathfrak{i}_0^t(\mathrm{K})_-^+\right] g(\tau) = \int_0^t [\mathrm{K}(r)\,g](\tau)\,\mathrm{d}r, \tag{133}$$

$$\left[\mathfrak{i}_0^t(\mathrm{K})_-^\circ\right] g(\tau) = \int_0^t [\mathrm{K}(r)\,\mathring{g}(r)](\tau)\,\mathrm{d}r, \tag{134}$$

$$\left[\mathfrak{i}_0^t(\mathrm{K})_\circ^+\, g\right](\tau) = \sum_{r\in\tau(t)} [\mathrm{K}(r)\,g](\tau\backslash r), \tag{135}$$

$$\left[\mathfrak{i}_0^t(\mathrm{K})_\circ^\circ\right] g(\tau) = \sum_{r\in\tau(t)} [\mathrm{K}(r)\,\mathring{g}(r)](\tau\backslash r). \tag{136}$$

Here $[\mathring{g}(r)](\tau) = g(r \sqcup \tau)$, where $r \sqcup \tau$ is union of disjoint $\tau \in \Gamma$ and $r \notin \tau$, $\tau\backslash r$ is difference of τ and a singleton $r \equiv \{r\} \subseteq \tau$ and $\tau(t) = \tau \cap [0,t)$. The functions $\mathrm{K}(r)$ should be L^p-integrable in a uniform operator topology,[10,25] with $p = 2/|s(\nu) - s(\mu)| = 1, 2, \infty$ defined by the signature $s(\mp) = \mp 1$ and $s(\circ) = 0$ of the indices μ, ν corresponding to $\mathrm{K} = \mathrm{K}_\nu^\mu$. Note that these definitions do not assume adaptedness of the integrands $\mathrm{K}_\nu^\mu(r)$ with respect to the algebras $\mathcal{A}_0^t = \mathcal{A}(F_0^t)$ as they generalize Hitsuda-Skorochod extended stochastic integral. The multiple version of this explicit QS-integration is straight forward if one indicates by index value $\circ$ a block indexed by $j \in J$, with values, say, in $\{0, 1, \ldots, d\}$, for all of which we set the signature $s(j) = 0$, and it can be found also in[10,25] as an nonadapted version of the Parthasarathy calculus[58] based on the coherent vectors.

(3) Itô rule (17) for QS integrals $\mathrm{V}(t) = \mathrm{V}(0) + \mathfrak{i}_0^t(\mathbf{K})$ with adapted four-integrand $\mathbf{K} = [\mathrm{K}_\nu^\mu]$ is based on the noncommutative Itô table (18) defined in terms of the matrix product

$$\mathbf{KL}^\star = \left\{\begin{matrix} 0 & \mathrm{K}_\circ^- & \mathrm{K}_+^- \\ 0 & \mathrm{K}_\circ^\circ & \mathrm{K}_+^\circ \\ 0 & 0 & 0 \end{matrix}\right\} \left\{\begin{matrix} 0 & \mathrm{L}_+^{\circ\dagger} & \mathrm{L}_+^{-\dagger} \\ 0 & \mathrm{L}_\circ^{\circ\dagger} & \mathrm{L}_\circ^{-\dagger} \\ 0 & 0 & 0 \end{matrix}\right\} \tag{137}$$

with involution $\mathrm{L}_{-\nu}^{\star\mu} = \mathrm{L}_{-\mu}^{\nu\dagger}$ in the standard matrix representation $\mathbf{K}(t) = \{\mathrm{K}_\nu^\mu(t)\}_{\nu=-,\circ,+}^{\mu=-,\circ,+}$ of the integrands forming a *noncommutative Itô ⋆-algebra*,[10,25] where we set $\mathrm{K}_-^\mu = 0 = \mathrm{K}_\nu^+$. This rule was derived in[51] for simple bounded integrands, and therefore can not be rigorously applied for multiple integration of QS equations except the special isometric case. In the general form presented here the QS Itô formula was proved for unbounded integrands in[21] where it was also extended to nonadapted integrands. The *functional noncommutative Itô formula* was also obtained in the pseudo-Poisson form as

$$\mathrm{d}f(\mathrm{V}(t)) = \left(f\left(\mathring{\mathrm{V}}\left(t^+\right)\right) - f(\mathrm{V}(t))\,\mathbf{1}\right)_\nu^\mu \mathrm{A}_\mu^\nu(\mathrm{d}t)\,. \tag{138}$$

Here $\mathbf{1}$ is unit matrix $\{\delta_\nu^\mu\}$ indexed by $\mu,\nu \in \{-,J,+\}$ with $\mathrm{V}\otimes\mathbf{1}$ denoted for any operator V simply as $\mathrm{V}\mathbf{1} \equiv \mathbf{V}$, and $\mathring{\mathrm{V}}(t^+) = \mathbf{J}_\mathrm{V}^t$ is the right QS *germ*[34] of the QS integral $\mathrm{V}(t)$ which is given as the matrix sum $\mathbf{J}_\mathrm{V}^t = \{\mathrm{V}\delta_\nu^\mu + \mathrm{K}_\nu^\mu\}_{\nu=-,\circ,+}^{\mu=-,\circ,+} \equiv \mathbf{V} + \mathbf{K}$ at each t by the quadruple $\mathbf{K}(t) = [\mathrm{K}_\nu^\mu(t)]$ of the QS derivatives $\mathrm{K}_\nu^\mu(t)$. (The summation convention over $\mu,\nu = -,\circ,+$ is applied.) Using this formula the HP differential conditions (35) for the QS unitarity of a QS interaction evolution $\mathrm{U}(t)$ were obtained as pseudo-unitarity condition in terms of the germ $\mathring{\mathrm{U}}(t^+) = \mathrm{U}(t)\,\mathbf{S}(t)$. Also the QS differential conditions of complete positivity, contractivity and projectivity were found in[34,35] respectively as conditions for complete pseudo-positivity, pseudo-contractivity and pseudo-projectivity of the corresponding QS germs $\mathring{\mathrm{V}}(t^+)$.

(4) Using quantum Itô formula, the general QS *evolution equation* was obtained in[32,33] for QS density operators $\mathrm{R}_t \vdash \mathcal{A}^{t]}$, defining normal states $\langle \mathrm{R}_t, \mathrm{B}\rangle$ evolved by a QS completely positive cocycle $\phi_\star^s(t) : \mathrm{R}_t \mapsto \mathrm{R}_{t+s}$ on an increasing family $\left(\mathcal{A}^{t]}\right)$ of W*-algebras $\mathcal{A}_0^t \bar{\otimes} \mathbb{A}^\circ$ on the Hilbert space $F_0^t \otimes H_\circ$. It has the generalized form of the linear QS master equation

$$\mathrm{dR} = (\mathrm{C}_\nu^\iota \mathrm{R}\mathrm{C}_\kappa^{\star\nu} - \mathrm{R}\delta_\kappa^\iota)\,\mathrm{dA}_\iota^\kappa,\ \mathrm{R}^0 = \varrho \otimes I_0 \tag{139}$$

Here $\mathbf{C} = \mathbf{I} + \mathbf{L} = \{\mathrm{C}^{\mu}_{\nu}\}$ is an adapted triangular bock-operator with $\mathrm{C}^{-}_{-} = \mathrm{I} = \mathrm{C}^{+}_{+}$ defining the germ $\mathring{\mathrm{V}}_{t^+} = \mathrm{V}_t\mathbf{C}\,(t)$ for the QS evolution equation $\mathrm{dV} = \mathrm{VL}^{\mu}_{\nu}\mathrm{dA}^{\nu}_{\mu}$ with $\mathrm{L}^{\mu}_{\nu} = 0$ if $\mu = +$ or $\nu = -$ for the adapted V_t (which is not assumed to be isometric). In the case of *martingale normalization condition* $\mathrm{S}^{-}_{\iota}\mathrm{S}^{\star\iota}_{+} = \mathrm{O}$ in terms of left adjoint operators $\mathrm{S}^{\mu}_{-\nu} = \mathrm{C}^{\nu\sharp}_{-\mu}$ with respect to a standard Hermitian pairing $(\mathbb{A}^{\mathrm{o}}_{\star}|\mathbb{A}^{\mathrm{o}})$ this equation describes QS object-output *entangling* process. It evolves an initial state $\mathrm{R}^0 = \varrho$ onto a QS state R_t satisfying normalization $\langle(\mathrm{R}_t \mid \mathrm{I})\,, \mathrm{I}_0\rangle_{\emptyset} = 1 = (\varrho \mid \mathrm{I})$ with respect to the modular pairing $\langle \mathrm{R}, \mathrm{B}\rangle_{\emptyset} = \left\langle \mathrm{R}^*\delta^{\emptyset} \mid \mathrm{B}\delta^{\emptyset}\right\rangle$ on an input algebra $\mathcal{A}_0$ with the faithful vacuum vector $\delta^{\emptyset} = 0^{\otimes}$. This is the most general QS equation preserving complete positivity and normalization in the martingale sense. Denoting $\mathrm{K}^{-}_{\iota} = -\mathrm{C}^{-}_{\iota} \equiv \mathrm{K}_{\iota}$ such that $\mathrm{C}^{\star\iota}_{+} = -\mathrm{K}^{\dagger}_{\iota}$, it can be written[32,35] as

$$\mathrm{dR} + 2\,\mathsf{Re}\left[\mathrm{K}_{\iota}\mathrm{R}\mathrm{dA}^{\iota}_{-}\right] = \left(\sum_{j} \mathrm{C}^{j}_{\kappa}\mathrm{R}\mathrm{C}^{j\dagger}_{\iota} - \mathrm{R}\delta^{\iota}_{\kappa}\right)\mathrm{dA}^{\kappa}_{\iota}. \tag{140}$$

More explicitly this generalized Belavkin equation can be written in terms of $\mathrm{K} = \mathrm{K}_{+}$, $\mathrm{L}^{i}_{+} = \mathrm{C}^{i}_{+} \equiv \mathrm{L}^{i}$ such that $\mathrm{C}^{j\dagger}_{+} = \mathrm{L}^{j\dagger}$ as

$$\begin{aligned}
&\mathrm{dR}_t + \left(2\,\mathsf{Re}\left[\mathrm{KR}_t\right] - \sum_{j}\mathrm{L}^{j}\mathrm{R}\mathrm{L}^{j\dagger}\right)\mathrm{d}t\\
&= \sum_{k} 2\,\mathsf{Re}\left[\left(\sum_{j}\mathrm{C}^{j}_{k}\mathrm{R}_t\mathrm{L}^{j\dagger} - \mathrm{K}_k\mathrm{R}_t\right)\mathrm{dA}^{k}_{-}\right]\\
&+ \sum_{ik}\left(\sum_{j}\mathrm{C}^{j}_{k}\mathrm{R}_t\mathrm{C}^{j\dagger}_{i} - \mathrm{R}_t\delta^{i}_{k}\right)\mathrm{dA}^{k}_{i}.
\end{aligned}$$

The martingale normalization condition can be written as

$$\mathrm{K} + \mathrm{K}^{\dagger} + \sum_{i}\mathrm{L}_i\mathrm{L}^{\dagger}_{i} = 0$$

in terms of left adjoint K, L_i to $-\mathrm{L}, \mathrm{L}^i$ such that $\mathrm{L} = -\mathrm{K}^{\sharp}, \mathrm{L}^i = \mathrm{L}^{\sharp}_{i}$ ($\mathrm{K} = -\mathrm{L}^{\dagger}_{+}, \mathrm{L}^i = \mathrm{L}^{\dagger}_{i}$ in the case of trace pairing) for any number of i's, and arbitrary K_i, C^{i}_{k}, $i, k = 1, \ldots, d$. This is a QS generalization of Lindblad equation[57] which is given by the generator (48) corresponding to the deterministic case $d = 0$.

1. C. Ahn, A. Doherty, and A. Landahl, "Continuous quantum error correction via quantum feedback control," *Phys. Rev. A*, vol. 65, p. 042301, 2002.
2. M. Armen, J. Au, J. Stockton, A. Doherty, and H. Mabuchi, "Adaptive homodyne measurement of optical phase," *Phys. Rev. Lett.*, vol. 89, p. 133602, 2002.
3. A. Barchielli and V. P. Belavkin, "Measurements continuous in time and posteriori states in quantum mechanics," *J Phys A Math Gen*, vol. 24, pp. 1495–1514, 1991.
4. V. P. Belavkin, "Optimal quantum filtration of markovian signals," *Problems of Control and Information Theory*, vol. 7, no. 5, pp. 345–360, 1978.
5. ——, "Optimal measurement and control in quantum dynamical systems," Instytut Fizyki, Copernicus University, Torun', Tech. Rep. 411, 1979.
6. ——, "Optimal filtering of markov signals with quantum white noise," *Radio Eng Electron Physics*, vol. 25, pp. 1445–1453, 1980.
7. ——, "On the theory of controlling observable quantum systems," *Automatica and Remote Control*, vol. 44, no. 2, pp. 178–188, 1983.
8. ——, "Reconstruction theorem for quantum stochastic processes," *Theoret Math Phys*, vol. 3, pp. 409–431, 1985.
9. ——, "Non-demolition measurement and control in quantum dynamical systems," in *Proc of C I S M Seminar on Information Complexity and Control in Quantum Physics, Udine 1985.*, A. Blaquiere, Ed. Wien–New York: Springer–Verlag, 1987, pp. 311–329.
10. ——, "A new form and $*$-algebraic structure of quantum stochastic integrals in fock space," in *Rendiconti del Seminario Matematico e Fisico di Milano LVIII*, 1988, pp. 177–193.
11. ——, "Non-demolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes," in *Proc of Bellmann Continuum Workshop 'Modelling and Control of Systems', Sophia–Antipolis 1988*, ser. Lecture notes in Control and Inform Sciences, A.Blaquiere, Ed., vol. 121. Berlin–Heidelberg–New York–London–Paris–Tokyo: Springer–Verlag, 1988, pp. 245–265.
12. ——, "A continuous counting observation and posterior quantum dynamics," *J Phys A Math Gen*, vol. 22, pp. L 1109–L 1114, 1989.
13. ——, "A new wave equation for a continuous non-demolition measurement," *Phys Letters A*, vol. 140, pp. 355–358, 1989.
14. ——, "Non-demolition stochastic calculus in fock space and nonlinear filtering and control in quantum systems," in *Proc of Fourteenth Winter School in Theor Phys, Karpacz 1988*, ser. Stochastic Methods in Mathematics and Physics. Singapore: World Scientific, 1989, pp. 310–324.
15. ——, "Stochastic calculus of input-output quantum processes and nondemolition filtering," in *Reviews on Newest Achievements in Science and Technology*, ser. Modern Problems of Mathematics. Moscow: VINITI, 1989, vol. 36, pp. 29–67.
16. ——, "A posterior Schrödinger equation for continuous non-demolition measurement," *J of Math Phys*, vol. 31, no. 12, pp. 2930–2934, 1990.
17. ——, "A posterior stochastic equations for quantum brownian motion," in

Proc of the 1989 Conference on Stochastic Methods in Experimental Sciences. Singapore: World Scientific, 1990, pp. 26–42.
18. ——, "Stochastic equations of quantum filtering," in *Prob. Theory and Math. Statistics*, B. et al., Ed. VSP/Mokslas, 1990, pp. 91–109.
19. ——, "A stochastic posterior Schrödinger equation for counting non-demolition measurement," *Letters in Math Phys*, vol. 20, pp. 85–89, 1990.
20. ——, "Continuous non-demolition observation, quantum filtering and optimal estimation," in *Quantum Aspects of Optical Communication, Proceedings, Paris 1990*, ser. Lecture notes in Physics, vol. 379. Berlin: Springer, 1991, pp. 151–163.
21. ——, "A quantum nonadapted Ito formula and stochastic analysis in Fock scale," *J of Funct Analysis*, vol. 102, no. 2, pp. 414–447, 1991.
22. ——, "Chaotic states and stochastic integrations in quantum systems," *Usp. Math Nauk (Russian Math Surveys)*, vol. 47, pp. 47–106, 1992.
23. ——, "Kernel representations of *-semigroups associated with infinitely divisible states," in *Quantum Probability and Related Topics*. World Scientific, 1992, vol. 7, pp. 31–50.
24. ——, "Quantum continual measurements and a posteriori collapse on CCR," *Comm. Math. Phys.*, vol. 146, pp. 611–635, 1992.
25. ——, "Quantum stochastic calculus and quantum nonlinear filtering," *Journal of Multivariate Analysis*, vol. 42, no. 2, pp. 171–201, 1992.
26. ——, "Quantum diffustion, measurement and filtering." *Probability Theory and its Application*, vol. 38, no. 4, pp. 742–757, 1993.
27. ——, "The unified ito formula has pseudo-poisson structure," *Journal of Mathematical Physics*, vol. 34, no. 4, pp. 1508–1518, 1993.
28. ——, "Nondemolition principle of quantum measurement theory," *Foundation of Physics*, vol. 24, no. 5, pp. 685–714, 1994.
29. ——, "A stochastic model of quantum observation," in *Proc of Symposium on the Foundations of Modern Physics*, P. B. et al, Ed. World Scientific, 1994, pp. 38–54.
30. ——, "A dynamical theory of quantum continuous measurement and spontaneous localization," *Russian Journal of Mathematical Physics*, vol. 3, no. 1, pp. 3–24, 1995.
31. ——, "The interplay of classical and quantum stochastics: Diffusion, measurement and filtering," in *Chaos – The Interplay Between Stochastic and Deterministic Behaviour*, ser. Lecture Notes in Physics. Springer, 1995, pp. 21–41.
32. ——, "On stochastic generators of completely positive cocycles," *Russ Journ of Math Phys*, vol. 3, no. 4, pp. 523–528, 1995.
33. ——, "On the general form of quantum stochastic evolution equation," in *Stochastic Analysis and Applications*, I. M. D. at al, Ed. World Scientific, 1996, pp. 91–106.
34. ——, "Positive definite germs of quantum stochastic processes," *C. R. Acad. Sci. Paris*, vol. 322, no. 1, pp. 385–390, 1996.
35. ——, "Quantum stochastic positive evolutions: Characterization, construction, dilation," *Commun. Math. Phys.*, vol. 184, pp. 533–566, 1997.

36. ——, "Measurement, filtering and control in quantum open dynamical systems," *Rep on Math Phys*, vol. 43, no. 3, pp. 405–425, 1999.
37. V. P. Belavkin and P. Staszewski, "A continuous observation of photon emission," *Reports in Mathematical Physics*, vol. 29, pp. 213–225, 1990.
38. ——, "Nondemolition observation of a free quantum particle," *Phys Rev A*, vol. 45, no. 3, pp. 1347–1357, 1992.
39. V. Belavkin and O. Melsheimer, "A stochastic Hamiltonian approach for quantum jumps, spontaneous localizations, and continuous trajectories," *Quantum Semiclass. Opt.*, vol. 8, pp. 167–187, 1996.
40. V. Belavkin and P. Staszewski, "A quantum particle undergoing continuous observation," *Physics Letters A*, vol. 140, pp. 359–362, 1989.
41. A. Bensoussan, *Stochastic Control of Partially Observable Systems.* Cambridge University Press, 1992.
42. L. M. Bouten and R. van Handel, "On the separation principle of quantum control, arXiv:math-ph/0511021," 2005.
43. ——, "Quantum filtering: a reference probability approach, arXiv:math-ph/0508006," 2005.
44. L. M. Bouten, R. van Handel, and M. James, "An introduction to quantum filtering, arXiv:math.OC/0601741," 2006.
45. A. C. Doherty, K. Jacobs, and G. Jungman, "Information, disturbance and Hamiltonian quantum feedback control," *Phys. Rev. A*, vol. 63, p. 062306, 2001.
46. A. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. Tan, "Quantum feedback and classical control theory," *Phys. Rev. A*, vol. 62, p. 012105, 2000.
47. J. Geremia, J. Stockton, A. Doherty, and H.Mabuchi, "Quantum Kalman filtering and the Heisenberg limit in atomic magnetometry," *Phys. Rev. Lett.*, vol. 91, p. 250801, 2003.
48. V. Gorini, A. Kossakowski, and E. Sudarshan, "Completely positive dynamical semigroups of N-level systems," *J. Math Phys.*, vol. 17, no. 5, pp. 821–825, 1976.
49. M. Gregoratti and R. F. Werner, "On quantum error-correction by classical feedback in discrete time," *J. Math. Phys.*, vol. 45, pp. 2600–2612, 2004.
50. R. L. Hudson and K. R. Parthasarathy, "Quantum Itôs formula and stochastic evolutions," *Communications in Mathematical Physics*, vol. 93, no. 3, pp. 301–323, 1984.
51. R. Hudson and K. Parthasarathy, "Quantum Itô's formula and stochastic evolutions," *Commun. Math. Phys.*, vol. 93, pp. 301–323, 1984.
52. K. Jacobs, "How to project qubits faster using quantum feedback," *Phys. Rev. A.*, vol. 67, p. 030301(R), 2003.
53. ——, "Optimal feedback control for the rapid preparation of a single qubit," in *Proceedings of the SPIE*, vol. 5468, 2004, pp. 355–364.
54. M. R. James, "Risk sensitive optimal control of quantum systems," *Phys. Rev. A.*, vol. 69, p. 032108, 2004.
55. ——, "A quantum Langevin formulation of risk sensitive optimal contol," *J. Opt. B: Quantum Semiclass. Opt.*, vol. 7, pp. S198–S207, 2005, special issue on quantum control.

56. P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, NJ, 1986.
57. G. Lindblad, "On the generators of quantum dynamical semigroups," *Commun. Math. Phys.*, vol. 48, pp. 119–130, 1976.
58. K. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*. Birkhäuser, Basel, 1992.
59. R. L. Stratonovich, "On the theory of optimal control. sufficient coordinates." *Automation and Remote Control.*, vol. 23, no. 7, pp. 910–917, 1962.
60. ——, *Conditional Markovian Processes and their Application to the Theory of Optimal Control*. Moscow: MGU, 1966.
61. R. Stratonovich, *Conditional Markov processes and their application to the theory of optimal control*. American Elsevier Publishing Company, Inc, New-York, 1968.
62. J. G. V. P. Belavkin and O. Smolianov, "Hamilton-jacobi-bellman equation for quantum feedback control," *J. Opt. B: Quantum and Semiclass. Opt.*, vol. 7, pp. S237– S244, 2005.
63. L. B. V. P. Belavkin and S. Edwards, "Bellman equations for optimal feedback control of qubit states," *J. of Physics B: Atomic, Molecular and Optical Physics*, vol. 38, pp. 151 – 160, 2005.
64. R. van Handel, J. K. Stockton, and H. Mabuchi, "Feedback control of quantum state reduction," *IEEE Trans. Automat. Control*, vol. 50, no. 6, pp. 768–780, 2005.

ON THE SEPARATION PRINCIPLE IN QUANTUM CONTROL

LUC BOUTEN* and RAMON VAN HANDEL†

Physical Measurement and Control 266-33, California Institute of Technology, Pasadena, CA 91125, USA
**E-mail: bouten@its.caltech.edu*
†E-mail: ramon@its.caltech.edu

It is well known that quantum continuous observations and nonlinear filtering can be developed within the framework of the quantum stochastic calculus of Hudson-Parthasarathy. The addition of real-time feedback control has been discussed by many authors, but the foundations of the theory still appear to be relatively undeveloped. Here we introduce the notion of a controlled quantum flow, where feedback is taken into account by allowing the coefficients of the quantum stochastic differential equation to be adapted processes in the observation algebra. We then prove a separation theorem for quantum control: the admissible control that minimizes a given cost function is a memoryless function of the filter, provided that the associated Bellman equation has a sufficiently regular solution. Along the way we obtain results on existence and uniqueness of the solutions of controlled quantum filtering equations and on the innovations problem in the quantum setting.

Keywords: Quantum stochastic control, quantum filtering, controlled quantum flow, separation theorem.

1. Introduction

Quantum feedback control is a branch of stochastic control theory that takes into account the inherent uncertainty in quantum systems. Though quantum stochastic control was first investigated in the 1980s in the pioneering papers of Belavkin[1,2] it is only recently that this has become a feasible technology, as demonstrated by recent laboratory experiments in quantum optics[3,4]. On the other hand, modern computing and sensing technology are rapidly reaching a level of miniaturization and sensitivity at which inherent quantum uncertainties can no longer be neglected. The development of control theoretic machinery for the design of devices that are robust in presence of quantum uncertainty could thus have important implications

for a future generation of precision technology.

Though not as mature as their classical counterparts, mathematical tools for quantum stochastic analysis have been extensively developed over the last two decades following the introduction of quantum stochastic calculus by Hudson and Parthasarathy[5]. Quantum stochastic differential equations are known to provide accurate Markov models of realistic quantum systems, particularly the atomic-optical systems used in quantum optics, and continuous-time optical measurements are also accurately described within this framework. Furthermore, nonlinear filtering theory for quantum systems has been extensively developed[6–8] and provides a suitable notion of conditioning for quantum systems. Nonetheless the theory of quantum control is still very much in its infancy, and despite the large body of literature on classical stochastic control only a few rigorous results are available in the quantum case. Our goal here is to make a first step in this direction by proving a quantum version of a simple but important theorem in classical control theory: a separation theorem for optimal stochastic controls.

To set the stage for the remainder of the article, let us demonstrate the idea with a simple (but important) example. We consider an atom in interaction with the vacuum electromagnetic field; the atom can emit photons into the field, and we allow ourselves to control the strength of a fixed atomic Hamiltonian. The system dynamics is given by the quantum stochastic differential equation (QSDE)

$$dj_t(X) = u(t)\, j_t(i[H,X])\, dt + j_t(\mathcal{L}[X])\, dt + j_t([X,L])\, dA_t^* + j_t([L^*,X])\, dA_t$$

where $j_t(X)$ denotes the atomic observable X at time t, and for now $u(t)$ is a deterministic control (i.e. an open loop control). If we perform homodyne detection in the field, we observe the stochastic process Y_t given by

$$dY_t = j_t(L + L^*)\, dt + dA_t^* + dA_t.$$

We now have a system-observation pair as in classical stochastic control. Inspired by results in classical control theory, we begin by finding a recursive equation for $\pi_t(X) = \mathbb{P}(j_t(X)|\mathcal{Y}_t)$, the conditional expectation of the atomic observable X at time t, given the observations Y_s up to time t. One obtains the nonlinear filter

$$\begin{aligned} d\pi_t(X) = {} & u(t)\, \pi_t(i[H,X])\, dt + \pi_t(\mathcal{L}[X])\, dt \\ & + (\pi_t(XL + L^*X) - \pi_t(L + L^*)\pi_t(X))\, (dY_t - \pi_t(L + L^*)\, dt). \end{aligned}$$

A crucial property of the conditional expectation is that the expectation of $\pi_t(X)$ equals the expectation of $j_t(X)$, i.e. $\mathbb{P}(\pi_t(X)) = \mathbb{P}(j_t(X))$. Suppose

we pose as our control goal the preparation of an atomic state with particular properties, e.g. we wish to find a control $u(t)$ such that after a long time $\mathbb{P}(j_t(X)) = \mathbb{P}_f(X)$ for some target state $\mathbb{P}_f$. Then it is sufficient to design a control that obeys this property for the filter $\pi_t(X)$. The advantage of using the filter is that $\pi_t(X)$ is only a function of the observations, and hence is always accessible to us, unlike $j_t(X)$ which is not directly observable. This approach was taken e.g. in[9,10].

We immediately run into technical problems, however, as we have assumed in the derivation of the filtering equation that $u(t)$ is a deterministic function whereas state preparation generally requires the use of a feedback control. One can of course simply replace the deterministic function $u(t)$ in the filter with some (feedback) function of the observation history, which is the approach generally taken in the literature, but does this new equation actually correspond to the associated controlled quantum system? The first issue that we resolve in this paper is to show that if $u(t)$ is replaced by some function of the observation history in the equation for $j_t(X)$, then the associated filtering equation corresponds precisely (as expected) to the open loop filtering equation obtained without feedback where the control $u(t)$ is replaced by the same function of the observations.

The separation of the feedback control strategy into a filtering step and a control step, as suggested above, is a desirable situation, as the filter can be calculated recursively and hence the control strategy is not difficult to implement. It is not obvious, however, that such a separation is always possible. Rather than taking state preparation as the control objective, consider the optimal control problem in which the control goal is to find a control strategy that minimizes a suitably chosen cost function. This is a common choice in control theory, and in general the cost function can even be expressed in terms of the nonlinear filter. However, it is not at all obvious that the *optimal* control at time t only depends on $\pi_t(X)$; in principle, the control could depend on the entire past history of the filter or even on some aspect of the observation process that is not captured by the filtering equation! The implementation of such a control would be awkward, as it would require the controller to have sufficient memory to store the entire observations history and enough resources to calculate an appropriate functional thereof.

Optimal control problems are often approached through the method of dynamic programming, which provides a candidate control strategy in separated form. We will show that if we can find a separated control strategy that satisfies the dynamic programming equations, then this strategy is in-

deed optimal even with respect to all non-separated strategies. Thus the fortunate conclusion is that even in the case of optimal controls we generally do not need to worry about non-separated control strategies. This establishes a foundation for the *separation principle* of quantum control, by which we mean that as a rule of thumb the design of quantum feedback controls can be reduced to a separate filtering step and a control step.

On the technical side, our treatment of quantum filtering proceeds by means of the reference probability approach[8] which is inspired by the approach of Zakai[11] in classical nonlinear filtering. Our treatment of the separation theorem is directly inspired by the classic papers of Wonham[12] and Segall[13]. A fully technical account of the results in this article will be presented in[14], and we apologize to the reader for the liberally sprinkled references to that paper. Here we will mostly neglect domain issues and similar technicalities, while we focus our attention on demonstrating the results announced above.

This paper is organized as follows. In section 2 we briefly recall some of the basic ideas of quantum probability theory, and we develop the reference probability approach to quantum filtering without feedback. In section 3 we introduce the notion of a controlled quantum flow and show that for such models the controlled filter takes the expected form. In section 4 we convert the filtering equations into classical stochastic differential equations and study their sample path properties; as a corollary, we obtain some results on the innovations problem. Finally, in section 5 we introduce the optimal control problem and prove a separation theorem.

2. Quantum probability and filtering

The purpose of this section is to briefly remind the reader of the basic ideas underlying quantum probability and filtering. For a more thorough introduction we refer to[8] and the references therein.

A quantum probability space $(\mathcal{A}, \mathbb{P})$ consists of a von Neumann algebra $\mathcal{A}$, defined on some underlying Hilbert space $\mathscr{H}$, and a normal state $\mathbb{P}$. If $\mathcal{A}$ is commutative then this definition is essentially identical to the usual definition in classical probability theory: indeed, the spectral theorem then guarantees that for some measure space (Ω, Σ, μ) there exists a *-isomorphism $\iota : \mathcal{A} \to L^\infty(\Omega, \Sigma, \mu)$ such that $\mathbb{P}(A) = \mathbb{E}_{\mathbf{P}}(\iota(A))$ for all $A \in \mathcal{A}$, where $\mathbb{E}_{\mathbf{P}}$ denotes the expectation w.r.t. the probability measure $\mathbf{P} \ll \mu$.[a] Thus any self-adjoint element of $\mathcal{A}$ represents a bounded random

[a]We will also denote the *-isomorphism as $\iota : (\mathcal{A}, \mathbb{P}) \to L^\infty(\Omega, \Sigma, \mu, \mathbf{P})$; this means that

variable (observable). In quantum models $\mathcal{A}$ is noncommutative, but in each realization we are only allowed to measure a commuting set of observables which generate a commutative von Neumann algebra $\mathcal{C}$. Hence if we fix a set of (commuting) observations to be performed in every realization of the experiment, then many computations reduce to classical probability theory.

To define conditional expectations in quantum probability, we simply "pull back" the associated notion from classical probability theory:

Definition 2.1 (Conditional expectation). *Let $(\mathcal{A}, \mathbb{P})$ be a quantum probability space and $\mathcal{C} \subset \mathcal{A}$ be a commutative von Neumann algebra. Define the commutant $\mathcal{C}' = \{A \in \mathcal{A} : AC = CA \ \forall C \in \mathcal{C}\}$. The map $\mathbb{P}(\cdot|\mathcal{C}) : \mathcal{C}' \to \mathcal{C}$ is called (a version of) the conditional expectation onto $\mathcal{C}$ if $\mathbb{P}(\mathbb{P}(A|\mathcal{C})C) = \mathbb{P}(AC)$ for all $A \in \mathcal{C}'$, $C \in \mathcal{C}$.*

Let us clarify the statement that this is nothing more than a classical conditional expectation. Let $A \in \mathcal{C}'$ be self-adjoint; then $\mathcal{C}_A = \mathrm{vN}(A, \mathcal{C})$, the von Neumann algebra generated by A and $\mathcal{C}$, is a commutative algebra, and the spectral theorem gives a *-isomorphism ι_A to some $L^\infty(\Omega_A, \Sigma_A, \mu_A, \mathbf{P}_A)$. But then we can simply calculate the classical conditional expectation and pull it back to the algebra: $\mathbb{P}(A|\mathcal{C}) = \iota_A^{-1}(\mathbb{E}_{\mathbf{P}_A}(\iota_A(A)|\sigma(\iota_A(\mathcal{C}))))$. This is in fact identical to Definition 2.1, and can be extended to any A by writing is as $B+iC$ where B, C are self-adjoint.

The definition we have given is less general than the usual definition[15]. In particular, we only allow conditioning onto a commutative algebra $\mathcal{C}$ from its commutant $\mathcal{C}'$. For statistical inference purposes (filtering) this is in fact sufficient: we wish to condition on the set of measurements made in a single realization of an experiment, hence they must commute; and it only makes sense to condition observables that are compatible with the observations already made, otherwise the conditional statistics would not be detectable by any experiment. Definition 2.1 has the additional advantage that existence, uniqueness, and all the basic properties can be proved by elementary means[8].

We list some of the most important properties of the conditional expectation: it exists, is a.s. unique (the difference between any two versions is zero with unit probability), and satisfies the least-squares property $|||A - \mathbb{P}(A|\mathcal{C})||| \le |||A - C|||$, $|||X|||^2 = \mathbb{P}(X^*X)$ for all $C \in \mathcal{C}$. Moreover we have (up to a.s. equivalence) linearity, positivity, invariance of the state $\mathbb{P}(\mathbb{P}(A|\mathcal{C})) = \mathbb{P}(A)$, the module property $\mathbb{P}(AB|\mathcal{C}) = B\,\mathbb{P}(A|\mathcal{C})$ for $B \in \mathcal{C}$,

the null sets are quotiented w.r.t. the measure μ, whereas $\mathbf{P} \ll \mu$ is the image of $\mathbb{P}$.

the tower property $\mathbb{P}(\mathbb{P}(A|\mathcal{B})|\mathcal{C}) = \mathbb{P}(A|\mathcal{C})$ if $\mathcal{C} \subset \mathcal{B}$, etc. Finally, we have the following Bayes-type formula:

Lemma 2.1 (Bayes formula[8]). *Let $(\mathcal{A}, \mathbb{P})$ be a quantum probability space and $\mathcal{C} \subset \mathcal{A}$ be a commutative von Neumann algebra. Let $V \in \mathcal{C}'$ be such that $V^*V > 0$, $\mathbb{P}(V^*V) = 1$. Then we can define a new state on $\mathcal{C}'$ by $\mathbb{Q}(A) = \mathbb{P}(V^*AV)$ and $\mathbb{Q}(X|\mathcal{C}) = \mathbb{P}(V^*XV|\mathcal{C}) \,/\, \mathbb{P}(V^*V|\mathcal{C})$ for $X \in \mathcal{C}'$.*

Up to this point we have only dealt with bounded operators. In the framework of quantum stochastic calculus, however, we unavoidably have to deal with unbounded operators that are affiliated to, not elements of, the various algebras mentioned above. As announced in the introduction, we largely forgo this issue here and we claim that all the results above (and below) can be extended to a sufficiently large class of unbounded operators. We refer to[14] for a complete treatment.

Let us now introduce a class of quantum models that we will consider in this paper. The model consists of an initial system, defined on a finite-dimensional Hilbert space $\mathscr{H}_0$, in interaction with an external (e.g. electromagnetic) field that lives on the usual Boson Fock space $\Gamma = \Gamma_s(L^2([0,T]))$. We will always work on a finite time horizon $[0,T]$ and we have restricted ourselves for simplicity to a single channel in the field (the case of multiple channels presents no significant complications; it can be treated in the same manner on a case by case basis[8]. A completely general theory can also be set up, e.g.[6], but the required notations seem unnecessarily complicated.) Throughout we place ourselves on the quantum probability space $(\mathcal{A}, \mathbb{P})$ where $\mathcal{A} = \mathcal{B} \otimes \mathcal{W}$, $\mathcal{B} = B(\mathscr{H}_0)$, $\mathcal{W} = B(\Gamma)$, and $\mathbb{P} = \rho \otimes \phi$ for some state ρ on $\mathcal{B}$ and the vacuum state ϕ on $\mathcal{W}$. We will use the standard notation $\Gamma_{t]} = \Gamma_s(L^2([0,t]))$, $\mathcal{W}_{t]} = B(\Gamma_{t]})$, etc. For $f \in L^2([0,T])$ we denote by $e(f) \in \Gamma$ the corresponding exponential vector, by $\Phi = e(0)$ the vacuum vector, and by A_t, A_t^* and Λ_t the fundamental noises. The reader is referred to[5,16–18] for background on quantum stochastic calculus.

For the time being we will not consider feedback control—we extend to this case in section 3. Without feedback, the interaction between the initial system and the field is given by the unitary solution of the Hudson-Parthasarathy QSDE

$$U_t = I + \int_0^t L_s U_s \, dA_s^* - \int_0^t L_s^* S_s U_s \, dA_s + \int_0^t (S_s - I) U_s \, d\Lambda_s - \int_0^t (iH_s + \tfrac{1}{2} L_s^* L_s) U_s \, ds.$$

Here L_t, S_t and H_t are bounded processes of operators in $\mathcal{B}$, S_t is unitary and H_t is self-adjoint. Without external controls the processes will usually be chosen to be time-independent, or we can imagine that e.g. the Hamiltonian $H_t = u(t)H$ is modulated by some *deterministic (open loop)* scalar control $u(t)$. In addition we specify an output noise that will be measured in the field; it takes the general form

$$Z_t = \int_0^t \lambda_s \, d\Lambda_s + \int_0^t \alpha_s \, dA_s^* + \int_0^t \alpha_s^* \, dA_s$$

where $\lambda : [0,T] \to \mathbb{R}$ and $\alpha : [0,T] \to \mathbb{C}$ are bounded scalar functions. Together U_t and Z_t provide a full description of a filtering problem: any initial system observable $X \in \mathcal{B}$ is given at time t by the flow $j_t(X) = U_t^* X U_t$, whereas the observation process that appears on our detector is given by $Y_t = U_t^* Z_t U_t$. Using the quantum Itô rules, we obtain the explicit expressions

$$\begin{aligned} dj_t(X) = {} & j_t(i[H_t, X] + L_t^* X L_t - \tfrac{1}{2}\{L_t^* L_t X + X L_t^* L_t\}) \, dt \\ & + j_t(S_t^*[X, L_t]) \, dA_t^* + j_t([L_t^*, X] S_t) \, dA_t + j_t(S_t^* X S_t - X) \, d\Lambda_t, \end{aligned} \quad (1)$$

$$\begin{aligned} dY_t = {} & \lambda_t \, d\Lambda_t + j_t \left(S_t^*(\alpha_t + \lambda_t L_t) \right) \, dA_t^* + j_t \left((\alpha_t^* + \lambda_t L_t^*) S_t \right) \, dA_t \\ & + j_t \left(\lambda_t L_t^* L_t + \alpha_t^* L_t + \alpha_t L_t^* \right) \, dt. \end{aligned} \quad (2)$$

This "system-theoretic" description in terms of the system-observations pair (1) and (2) is closest in spirit to the usual description of filtering and control problems in classical control theory. We will not explicitly use this representation, however.

We now turn to the filtering problem, i.e. the problem of finding an explicit representation for the conditional state $\pi_t(X) = \mathbb{P}(j_t(X)|\mathcal{Y}_t)$, $X \in \mathcal{B}$, where $\mathcal{Y}_t = \mathrm{vN}(Y_s : 0 \le s \le t)$ is the von Neumann algebra generated by the observations up to time t. Before we can go down this road we must prove that $\pi_t(X)$ is in fact well defined according to Definition 2.1. This is guaranteed by the following proposition.

Proposition 2.1 (Nondemolition property). *The observation process Y_t satisfies the self-nondemolition condition, i.e. $\mathcal{Y}_t$ is commutative for all $t \in [0,T]$, and is nondemolition with respect to the flow, i.e. $j_t(X) \in \mathcal{Y}_t'$ for all $X \in \mathcal{B}$ and $t \in [0,T]$.*

Proof. Let $\mathcal{Z}_t = \mathrm{vN}(Z_s : 0 \le s \le t)$. We begin by showing that $\mathcal{Z}_t$ is a

commutative algebra for all $t \in [0,T]$. To this end, define

$$Z(c,d) = \int_0^T c_s \, d\Lambda_s + \int_0^T d_s \, dA_s^* + \int_0^T d_s^* \, dA_s$$

so that $Z_t = Z(\lambda\chi_{[0,t]}, \alpha\chi_{[0,t]})$. Using the quantum Itô rules, we obtain

$$[Z(C,D), Z(c,d)] =$$
$$\int_0^T (C_s d_s - c_s D_s)\, dA_s^* + \int_0^T (D_s^* c_s - d_s^* C_s)\, dA_s + \int_0^T (D_s^* d_s - d_s^* D_s)\, dt.$$

But then we obtain $[Z_t, Z_{t'}] = 0$ for all $t, t' \in [0,T]$ by setting

$$C_s = \lambda_s \chi_{[0,t]}(s), \quad c_s = \lambda_s \chi_{[0,t']}(s), \quad D_s = \alpha_s \chi_{[0,t]}(s), \quad d_s = \alpha_s \chi_{[0,t']}(s).$$

We conclude that the process Z_t generates a commutative algebra.

Next, we claim that $U_s^* Z_s U_s = U_t^* Z_s U_t$ for all $s \leq t \in [0,T]$. To see this, let $E \in \mathcal{Z}_s$ be an arbitrary projection operator in the range of the spectral measure of Z_s. Using the quantum Itô formula, we obtain

$$j_t(E) = j_s(E) + \int_s^t j_\sigma(i[H_\sigma, E] + L_\sigma^* E L_\sigma - \tfrac{1}{2}\{L_\sigma^* L_\sigma E + E L_\sigma^* L_\sigma\})\, d\sigma$$
$$+ \int_s^t j_\sigma(S_\sigma^*[E, L_\sigma])\, dA_\sigma^* + \int_s^t j_\sigma([L_\sigma^*, E] S_\sigma)\, dA_\sigma + \int_s^t j_\sigma(S_\sigma^* E S_\sigma - E)\, d\Lambda_\sigma$$

where $j_t(X) = U_t^* X U_t$. But by construction E commutes with all $H_\sigma, L_\sigma, S_\sigma$, hence we obtain $j_t(E) = j_s(E)$. As this holds for all spectral projections E of Z_s, the assertion follows. We conclude that $Y_t = U_T^* Z_t U_T$ for all $t \in [0,T]$. But then $\mathcal{Y}_t = U_T^* \mathcal{Z}_t U_T$, and as $\mathcal{Z}_t$ is commutative $\mathcal{Y}_t$ must be as well.

It remains to prove the nondemolition condition. To this end, note that $j_t(X) \in U_t^* \mathcal{B} U_t$ and $\mathcal{Y}_t = U_t^* \mathcal{Z}_t U_t$. But $\mathcal{Z}_t$ consists of elements in $\mathcal{W}$, which clearly commute with every element in $\mathcal{B}$. The result follows directly. □

To obtain an explicit representation for the filtering equation we are inspired by the classical reference probability method of Zakai[11]. The idea of Zakai is to introduce a change of measure such that under the new (reference) measure the observation process is independent of the system observables, which significantly simplifies the calculation of conditional expectations. The Bayes formula relates the conditional expectations under the reference measure and the true measure. We will follow a similar route and make use of the quantum Bayes formula of Lemma 2.1. The main difficulty is the choice of an appropriate change of measure operator V.

In the classical reference probability method the change of measure is obtained from Girsanov's theorem. Unfortunately, there is no satisfactory noncommutative analog of the Girsanov theorem; even though Girsanov-like expressions can be obtained, they do not give rise to a change of state that lies in the commutant of the observations as required by Lemma 2.1. A different naive choice would be something like $\mathbb{R}(X) = \mathbb{P}(U_T X U_T^*)$, so that Y_t under $\mathbb{R}$ has the same statistics as the martingale Z_t under $\mathbb{P}$, but once again U_T does not commute with the observations. However, the latter idea (in a slightly modified form) can be "fixed" to work: starting from a change of state that is the solution of a QSDE, we can modify the QSDE somewhat so that the resulting solution still defines the same state but has the desired properties. This trick appears to have originated in a paper by Holevo[19]. We state it here in the following form.

Lemma 2.2. *Let $C_t, D_t, F_t, G_t, \tilde{C}_t, \tilde{F}_t$ be bounded processes, and let*

$$\begin{aligned} dV_t &= \{C_t\,d\Lambda_t + D_t\,dA_t^* + F_t\,dA_t + G_t\,dt\}V_t, \\ d\tilde{V}_t &= \{\tilde{C}_t\,d\Lambda_t + D_t\,dA_t^* + \tilde{F}_t\,dA_t + G_t\,dt\}\tilde{V}_t, \quad V_0 = \tilde{V}_0. \end{aligned}$$

Then $\mathbb{P}(V_t^* X V_t) = \mathbb{P}(\tilde{V}_t^* X \tilde{V}_t)$ for all $X \in \mathcal{B} \otimes \mathcal{W}$.

Proof. As any state ρ on $\mathcal{B}$ is a convex combination of vector states, it is sufficient to prove the Lemma for any vector state $\rho(B) = \langle v, Bv\rangle$, $v \in \mathscr{H}_0$. Hence it suffices to prove that $\langle V_t\,v\otimes\Phi, XV_t\,v\otimes\Phi\rangle = \langle \tilde{V}_t\,v\otimes\Phi, X\tilde{V}_t\,v\otimes\Phi\rangle$ for any $X \in \mathcal{B}\otimes\mathcal{W}$ and $v \in \mathscr{H}_0$, as $\mathbb{P} = \rho\otimes\phi$. But clearly this would be implied by $V_t\,v\otimes\Phi = \tilde{V}_t\,v\otimes\Phi\ \forall v \in \mathscr{H}_0$. Let us prove that this is in fact the case.

As all the coefficients of the QSDE for $V_t, \tilde{V}_t$ are bounded processes both V_t and $\tilde{V}_t$ have unique solutions. Consider the quantity

$$\|(V_t - \tilde{V}_t)\,v\otimes\Phi\|^2 = \langle (V_t - \tilde{V}_t)\,v\otimes\Phi, (V_t - \tilde{V}_t)\,v\otimes\Phi\rangle.$$

Using the quantum Itô rule we obtain (see e.g.[17])

$$\begin{aligned} \|(V_t - \tilde{V}_t)\,v\otimes\Phi\|^2 &= \int_0^t \langle (V_s - \tilde{V}_s)\,v\otimes\Phi, (G_s + G_s^*)(V_s - \tilde{V}_s)\,v\otimes\Phi\rangle\,ds \\ &\quad + \int_0^t \langle D_s(V_s - \tilde{V}_s)\,v\otimes\Phi, D_s(V_s - \tilde{V}_s)\,v\otimes\Phi\rangle\,ds. \end{aligned}$$

Note that the last integrand can be expressed as

$$\|D_s(V_s - \tilde{V}_s)\,v\otimes\Phi\|^2 \le \left[\sup_{t\in[0,T]} \|D_t\|^2\right] \|(V_s - \tilde{V}_s)\,v\otimes\Phi\|^2.$$

To deal with the first integrand, note that $G_s + G_s^*$ are self-adjoint bounded operators. Denote by G_s^+ the positive part of $G_s + G_s^*$, and by K_s^+ the square root of G_s^+ (i.e. $G_s^+ = K_s^{+*} K_s^+$). Then

$$\langle (V_s - \tilde{V}_s)\, v \otimes \Phi, (G_s + G_s^*)(V_s - \tilde{V}_s)\, v \otimes \Phi \rangle \leq \| K_s^+ (V_s - \tilde{V}_s)\, v \otimes \Phi \|^2.$$

But as $G_s + G_s^*$ is a bounded process, so is K_s^+ and we have

$$\| K_s^+ (V_s - \tilde{V}_s)\, v \otimes \Phi \|^2 \leq \left[\sup_{t \in [0,T]} \| K_t^+ \|^2 \right] \| (V_s - \tilde{V}_s)\, v \otimes \Phi \|^2.$$

Thus we obtain

$$\| (V_t - \tilde{V}_t)\, v \otimes \Phi \|^2 \leq C \int_0^t \| (V_s - \tilde{V}_s)\, v \otimes \Phi \|^2 \, ds$$

were by boundedness of D_t and K_t^+

$$C = \sup_{t \in [0,T]} \| K_t^+ \|^2 + \sup_{t \in [0,T]} \| D_t \|^2 < \infty.$$

But then by Gronwall's lemma $\| (V_t - \tilde{V}_t)\, v \otimes \Phi \| = 0$, and the Lemma is proved. □

We are now in the position to prove the main filtering theorem: a quantum version of the Kallianpur-Striebel formula. We consider separately two versions of the theorem for diffusive and counting observations, respectively (but the former also covers some cases with counting observations if $\lambda_t \neq 0$.)

Theorem 2.1 (Kallianpur-Striebel formula, diffusive case). *Suppose that $|\alpha_t| > 0$ for all $t \in [0,T]$. Let V_t be the solution of the QSDE*

$$V_t = I + \int_0^t \alpha_s^{-1} L_s V_s \, (\lambda_s \, d\Lambda_s + \alpha_s \, dA_s^* + \alpha_s^* \, dA_s) - \int_0^t (iH_s + \tfrac{1}{2} L_s^* L_s) V_s \, ds.$$

Then V_t is affiliated to $\mathcal{Z}_t'$ for every $t \in [0,T]$, and we have the representation $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$ with $\sigma_t(X) = U_t^* \mathbb{P}(V_t^* X V_t | \mathcal{Z}_t) U_t$, $X \in \mathcal{B}$.

Theorem 2.2 (Kallianpur-Striebel formula, counting case). *Suppose that $|\lambda_t| > 0$ for all $t \in [0,T]$. Let E_t be the solution of the QSDE*

$$E_t = I + \int_0^t \lambda_s^{-1} E_s \, \{ (1 - \alpha_s) dA_s^* - (1 - \alpha_s^*) dA_s \} - \frac{1}{2} \int_0^t \lambda_s^{-2} |1 - \alpha_s|^2 E_s \, ds$$

and let $C_t = E_t^* Z_t E_t$, $\mathcal{C}_t = \mathrm{vN}(C_s : 0 \leq s \leq t) = E_t^* \mathcal{Z}_t E_t$, so that

$$C_t = A_t^* + A_t + \int_0^t \lambda_s \, d\Lambda_s + \int_0^t \lambda_s^{-1} (1 - |\alpha_s|^2) \, ds.$$

Define V_t as the solution of the QSDE

$$V_t = I + \int_0^t (L_s - \lambda_s^{-1}(1-\alpha_s))V_s\,(\lambda_s\,d\Lambda_s + dA_s^* + dA_s) \\ - \int_0^t (iH_s + \tfrac{1}{2}L_s^*L_s + \tfrac{1}{2}\lambda_s^{-2}|1-\alpha_s|^2 - \lambda_s^{-1}(1-\alpha_s^*)L_s)V_s\,ds.$$

Then V_t is affiliated to $\mathcal{C}_t'$ for every $t \in [0,T]$, and we have the representation $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$ with $\sigma_t(X) = U_t^*E_t\mathbb{P}(V_t^*XV_t|\mathcal{C}_t)E_t^*U_t$, $X \in \mathcal{B}$.

Proof of Theorem 2.1. We begin by using a general transformation property of the quantum conditional expectation. Let U be a unitary operator and define a new state $\mathbb{Q}(X) = \mathbb{P}(U^*XU)$. Then $\mathbb{P}(U^*XU|U^*\mathcal{C}U) = U^*\mathbb{Q}(X|\mathcal{C})U$ provided that $\mathcal{C}$ is commutative and $X \in \mathcal{C}'$. The statement is easily verified by direct application of Definition 2.1. The reason we wish to perform such a transformation is that in order to apply Lemma 2.2, it will be more convenient if we condition with respect to $\mathcal{Z}_t$ rather than $\mathcal{Y}_t$. From this point on, we fix a $t \in [0,T]$ and define the new state $\mathbb{Q}(X) = \mathbb{P}(U_t^*XU_t)$. Evidently $\pi_t(X) = U_t^*\mathbb{Q}(X|\mathcal{Z}_t)U_t$.

We would now like to apply Lemma 2.1 to $\mathbb{Q}(X|\mathcal{Z}_t)$. Note that by Lemma 2.2 we obtain $\mathbb{P}(V_t^*XV_t) = \mathbb{P}(U_t^*XU_t) = \mathbb{Q}(X)$. Moreover V_t is affiliated to $\mathcal{B}\otimes\mathcal{Z}_t \subset \mathcal{Z}_t'$, as it is defined by a QSDE which is integrated against Z_t and has coefficients in $\mathcal{B}$ (this statement can be rigorously verified by an approximation argument.) Hence by Lemma 2.1 we obtain $\mathbb{Q}(X|\mathcal{Z}_t) = \mathbb{P}(V_t^*XV_t|\mathcal{Z}_t)/\mathbb{P}(V_t^*V_t|\mathcal{Z}_t)$. The result follows immediately. □

Proof of Theorem 2.2. The proof proceeds along the same lines as the proof of Theorem 2.1, except that in order to apply Lemma 2.2 we must make sure that the coefficient in front of dA_t^* in the output noise does not vanish. To this end we perform an additional rotation by E_t; the expression for C_t is obtained by direct application of the quantum Itô rules. If we introduce the transformed state $\mathbb{Q}(X) = \mathbb{P}(U_t^*E_tXE_t^*U_t)$, we obtain $\pi_t(X) = U_t^*E_t\mathbb{Q}(X|\mathcal{C}_t)E_t^*U_t$ for $X \in \mathcal{B}$ (we have used the fact that any such X commutes with E_t.) It remains to notice that $\mathbb{P}(V_t^*XV_t) = \mathbb{Q}(X)$ by Lemma 2.2 and by application of the quantum Itô rules to $E_t^*U_t$. The statement of the Theorem follows from Lemma 2.1. □

Now that we have the quantum Kallianpur-Striebel formulas, it is not difficult to obtain a recursive representation of the filtering equations. Let us demonstrate the procedure with Theorem 2.1. Using the quantum Itô

rules, we can write

$$V_t^* X V_t = X + \int_0^t V_s^*(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})V_s\, ds + \int_0^t V_s^*(\alpha_s^{-1} X L_s + \alpha_s^{-1*} L_s^* X + (\alpha_s^* \alpha_s)^{-1}\lambda_s L_s^* X L_s)V_s\, dZ_t$$

for every $X \in \mathcal{B}$ (and we use the shortened notation $dZ_t = \lambda_s\, d\Lambda_s + \alpha_s\, dA_s^* + \alpha_s^*\, dA_s$.) Taking the conditional expectation of both sides, we obtain

$$\mathbb{P}(V_t^* X V_t - X|\mathcal{Z}_t) = \int_0^t \mathbb{P}(V_s^*(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})V_s|\mathcal{Z}_s)\, ds + \int_0^t \mathbb{P}(V_s^*(\alpha_s^{-1} X L_s + \alpha_s^{-1*} L_s^* X + (\alpha_s^* \alpha_s)^{-1}\lambda_s L_s^* X L_s)V_s|\mathcal{Z}_s)\, dZ_t,$$

where the fact that we can pull the conditional expectation into the integrals can once again be verified by an approximation argument. It remains to rotate the expression by U_t; we obtain directly

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds + \int_0^t \sigma_s(\alpha_s^{-1} X L_s + \alpha_s^{-1*} L_s^* X + (\alpha_s^* \alpha_s)^{-1}\lambda_s L_s^* X L_s)\, dY_t$$

where dY_t is given by (2). This is the noncommutative counterpart of the unnormalized filtering equation of classical nonlinear filtering theory. Applying a similar procedure to Theorem 2.2 yields an unnormalized filtering equation for the counting case.

3. Controlled quantum flows and controlled filtering

In the previous section we considered a quantum system where the coefficients H_t, L_t, S_t were deterministic functions in $\mathcal{B}$; for example, we considered the case where L, S were constant in time and where the Hamiltonian $H_t = u(t)H$ was modulated by a deterministic control. This gives rise to a filtering equation, e.g. the unnormalized filtering equation that propagates $\sigma_t(\cdot)$, which also depends deterministically on the control $u(t)$. In a feedback control scenario, however, we would like to adapt the controls in real time based on the observations that have been accumulated; i.e., we want to make $u(t)$ a function of the observations Y_s up to time t. We now introduce the notion of a controlled QSDE, and show that this gives rise to a controlled filtering equation of the same form as in the previous section.

Definition 3.1 (Controlled quantum flow). *Given*

(1) an output noise Z_t *of the form*

$$Z_t = \int_0^t \Xi_s \, d\Lambda_s + \int_0^t \Upsilon_s \, dA_s^* + \int_0^t \Upsilon_s^* \, dA_s$$

such that Ξ_t, Υ_t *are adapted and affiliated to* $\mathcal{Z}_t = \mathrm{vN}(Z_s : 0 \le s \le t)$ *for every* $t \in [0,T]$, Ξ_t *is self-adjoint, and* $\mathcal{Z}_t$ *is a commutative algebra;*

(2) a controlled Hudson-Parthasarathy equation

$$U_t = I + \int_0^t L_s U_s \, dA_s^* - \int_0^t L_s^* S_s U_s \, dA_s + \int_0^t (S_s - I) U_s \, d\Lambda_s - \int_0^t (iH_s + \tfrac{1}{2} L_s^* L_s) U_s \, ds$$

where H_t, L_t, S_t *are affiliated to* $\mathcal{B} \otimes \mathcal{Z}_t$ *for every* $t \in [0,T]$,

the pair (j_t, Y_t), *where* $j_t(X) = U_t^* X U_t$ $(X \in \mathcal{B} \otimes \mathcal{Z}_t)$ *and* $Y_t = U_t^* Z_t U_t$, *is called a* controlled quantum flow j_t *with* observation process Y_t.

To use this definition we must impose sufficient regularity conditions on the various processes involved so that these are indeed well defined. From this point onward we assume that $\Xi_t, \Upsilon_t, H_t, L_t, S_t$ are bounded measurable processes, i.e.

$$\sup_{t \in [0,T]} \|\Xi_t\| < \infty, \qquad t \mapsto \Xi_t \psi \text{ is measurable } \forall\, \psi \in \mathscr{H}_0 \otimes \Gamma,$$

and similarly for the other processes. Under such assumptions (essentially bounded control requirements) we can show[14] that Z_t is well defined and that U_t has a unique unitary solution (see also[17,20,21] for related results.)

Before we discuss filtering in the context of Definition 3.1, let us clarify the significance of this definition. First, note that we can use the quantum Itô rules to obtain the system-observations pair

$$dj_t(X) = j_t(i[H_t, X] + L_t^* X L_t - \tfrac{1}{2}\{L_t^* L_t X + X L_t^* L_t\})\, dt + j_t(S_t^*[X, L_t])\, dA_t^* + j_t([L_t^*, X] S_t)\, dA_t + j_t(S_t^* X S_t - X)\, d\Lambda_t, \quad (3)$$

$$dY_t = \Xi_t \, d\Lambda_t + j_t\left(S_t^*(\Upsilon_t + \Xi_t L_t)\right) dA_t^* + j_t\left((\Upsilon_t^* + \Xi_t L_t^*) S_t\right) dA_t + j_t\left(\Xi_t L_t^* L_t + \Upsilon_t^* L_t + \Upsilon_t L_t^*\right) dt, \quad (4)$$

which is simply the controlled counterpart of (1), (2). The essential thing to notice is that though the quantities H_t, L_t etc. that appear in the equation for the flow U_t are affiliated to the output noise $\mathcal{Z}_t$, the quantities that

appear in the system-observations model are in fact of the form $j_t(H_t)$, etc., which are affiliated to the observations $\mathcal{Y}_t$. Our model is extremely general and allows for any of the coefficients of the QSDE, and even the measurement performed in the field, to be adapted in real time based on the observed process. To illustrate the various types of control that are typically used, we give the following examples.

Example 3.1 (Hamiltonian feedback). *Consider the controlled flow*

$$dU_t = (L\,dA_t^* - L^*\,dA_t - \tfrac{1}{2}L^*L\,dt - iu_t(Z_{s\le t})H\,dt)\,U_t, \qquad Z_t = A_t + A_t^*.$$

That is, we have chosen $S_t = 0$, fixed $L, H \in \mathcal{B}$, $H = H^*$, and $u_t(Z_{s\le t})$ is a bounded (real) scalar function of the output noise up to time t. This gives the system-observation pair

$$\begin{aligned} dj_t(X) &= j_t([X,L])\,dA_t^* + j_t([L^*,X])\,dA_t \\ &\quad + j_t(L^*XL - \tfrac{1}{2}(L^*LX + XL^*L))\,dt + u_t(Y_{s\le t})\,j_t(i[H,X])\,dt \end{aligned}$$

and $dY_t = dA_t + dA_t^* + j_t(L + L^*)\,dt$, where we have pulled the control outside j_t. This scenario corresponds to a fixed system-probe interaction, measurement and system Hamiltonian, where we allow ourselves to feed back some function of the observation history to modulate the strength of the Hamiltonian; see e.g.[9].

Example 3.2 (Coherent feedback). *The controlled quantum flow*

$$\begin{aligned} dU_t = ((L + u_t(Z_{s\le t})I)\,dA_t^* - (L^* + u_t(Z_{s\le t})^*I)\,dA_t \\ - \tfrac{1}{2}(L^*L + u_t(Z_{s\le t})^*u_t(Z_{s\le t})I + 2u_t(Z_{s\le t})L^*)\,dt)\,U_t, \end{aligned}$$

where $Z_t = A_t + A_t^*$, describes an initial system driven by a field in a coherent state, where we modulate the coherent state amplitude through the bounded (complex) control $u_t(Z_{s\le t})$[22]. As in the previous example, the control becomes a function of the observations when we transform to the system-theoretic description.

Example 3.3 (Adaptive measurement). *In this scenario we choose an uncontrolled flow $dU_t = (L\,dA_t^* - L^*\,dA_t - \frac{1}{2}L^*L\,dt - iH\,dt)\,U_t$, but the measurement in the probe field is adapted in real time by*

$$dZ_t = e^{-iu_t(Z_{s\le t})}dA_t^* + e^{iu_t(Z_{s\le t})}dA_t$$

where $u_t(Z_{s\le t})$ is a real scalar control function. This gives rise to

$$dY_t = dZ_t + \left[j_t(L)\,e^{iu_t(Y_{s\le t})} + j_t(L^*)\,e^{-iu_t(Y_{s\le t})}\right]\,dt.$$

Evidently the control determines which of the system observables $L\,e^{iu} + L^*\,e^{-iu}$ is detected in the probe. The possibility to adapt the measurement in real time is useful for the detection of quantities that are not described by a system observable, such as the phase of an optical pulse. See e.g.[3,23].

We now wish to solve the filtering problem for a controlled quantum flow. It turns out that when we use the reference probability method, very little changes in the procedure outlined above. In particular, the proofs of Proposition 2.1 and Lemma 2.2 extend readily to the controlled case, and it is straightforward to extend the proofs of Theorems 2.1 and 2.2 to prove the following statements.

Theorem 3.1 (Kallianpur-Striebel formula, diffusive case). *Suppose that Υ_t has a bounded inverse for all $t \in [0,T]$. Let V_t be the solution of the QSDE*

$$V_t = I + \int_0^t (\Xi_s\,d\Lambda_s + \Upsilon_s\,dA_s^* + \Upsilon_s^*\,dA_s)\,\Upsilon_s^{-1} L_s V_s - \int_0^t (iH_s + \tfrac{1}{2}L_s^*L_s)V_s\,ds.$$

Then V_t is affiliated to $\mathcal{Z}_t'$ for every $t \in [0,T]$, and we have the representation $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$ with $\sigma_t(X) = U_t^*\mathbb{P}(V_t^*XV_t|\mathcal{Z}_t)U_t$, $X \in \mathcal{B}\otimes\mathcal{Z}_t$.

Theorem 3.2 (Kallianpur-Striebel formula, counting case). *Suppose that Ξ_t has a bounded inverse for all $t \in [0,T]$. Let E_t be the solution of the QSDE ($|X|^2 = X^*X$)*

$$E_t = I + \int_0^t \Xi_s^{-1}E_s\,\{(1-\Upsilon_s)dA_s^* - (1-\Upsilon_s^*)dA_s\} - \frac{1}{2}\int_0^t \Xi_s^{-2}|1-\Upsilon_s|^2E_s\,ds.$$

Let $C_t = E_t^*Z_tE_t$, $\mathcal{C}_t = \mathrm{vN}(C_s : 0 \le s \le t) = E_t^*\mathcal{Z}_tE_t$, so that

$$C_t = A_t^* + A_t + \int_0^t E_s^*\Xi_sE_s\,d\Lambda_s + \int_0^t E_s^*\Xi_s^{-1}(1-\Upsilon_s^*\Upsilon_s)E_s\,ds.$$

Define V_t as the solution of the QSDE

$$\begin{aligned} V_t = I &+ \int_0^t (E_s^*\Xi_sE_s\,d\Lambda_s + dA_s^* + dA_s)\,E_s^*(L_s - \Xi_s^{-1}(1-\Upsilon_s))E_sV_s \\ &- \int_0^t E_s^*(iH_s + \tfrac{1}{2}L_s^*L_s + \tfrac{1}{2}\Xi_s^{-2}|1-\Upsilon_s|^2 - \Xi_s^{-1}(1-\Upsilon_s^*)L_s)E_sV_s\,ds. \end{aligned}$$

Then V_t is affiliated to $\mathcal{C}_t'$ for every $t \in [0,T]$, and we have $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$ with $\sigma_t(X) = U_t^*E_t\mathbb{P}(V_t^*E_t^*XE_tV_t|\mathcal{C}_t)E_t^*U_t$, $X \in \mathcal{B}\otimes\mathcal{Z}_t$.

We now obtain controlled filtering equations for $\sigma_t(\cdot)$ in recursive form. The following statements are readily verified using the quantum Itô rules.

Corollary 3.1 (Unnormalized filtering equation, diffusive case). *Suppose that Υ_t has a bounded inverse for all $t \in [0,T]$. Then $\sigma_t(\cdot)$ of Theorem 3.1 satisfies*

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s,X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\,ds + \int_0^t \sigma_s(\Upsilon_s^{-1} X L_s + \Upsilon_s^{-1*} L_s^* X + (\Upsilon_s^* \Upsilon_s)^{-1} \Xi_s\, L_s^* X L_s)\,dY_s$$

for $X \in \mathcal{B}$, where dY_t is given by (4) and $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$.

Corollary 3.2 (Unnormalized filtering equation, counting case). *Suppose that Ξ_t has a bounded inverse for all $t \in [0,T]$. Then $\sigma_t(\cdot)$ of Theorem 3.2 satisfies*

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s,X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\,ds + \int_0^t \sigma_s(\Xi_s\, L_s^* X L_s + \Upsilon_s^* X L_s + \Upsilon_s L_s^* X - \Xi_s^{-1}(1 - \Upsilon_s^* \Upsilon_s) X) \times (dY_s - j_s(\Xi_s^{-1}(1 - \Upsilon_s^* \Upsilon_s))\,ds)$$

for $X \in \mathcal{B}$, where dY_t is given by (4) and $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$.

What is the relation between the open loop filters of the previous section and the closed loop filters associated to a controlled quantum flow? Consider for example the case of Hamiltonian control where $Z_t = A_t + A_t^*$, $S_t = 0$, $L_t = L \in \mathcal{B}$ is constant and $H_t = u(t)H$ with $H \in \mathcal{B}$. In open loop $u(t)$ is a deterministic function and we saw in the previous section that this gives rise to the unnormalized filtering equation

$$d\sigma_t(X) = u(t)\,\sigma_t(i[H,X])\,dt + \sigma_t(L^* X L - \tfrac{1}{2}\{L^* L X + X L^* L\})\,dt + \sigma_t(XL + L^* X)\,dY_t.$$

In closed loop the model is given by the controlled quantum flow of Example 3.1, and by Corollary 3.1 we obtain the controlled unnormalized equation

$$d\sigma_t(X) = u(Y_{s\le t})\,\sigma_t(i[H,X])\,dt + \sigma_t(L^* X L - \tfrac{1}{2}\{L^* L X + X L^* L\})\,dt + \sigma_t(XL + L^* X)\,dY_t.$$

Evidently we obtain the same filter for the closed loop controlled quantum flow as we would obtain by calculating the open loop filter and then substituting a feedback control for the deterministic function $u(t)$. This is in

fact a general property of controlled filtering equations, as can be seen directly from the statement of Corollaries 3.1 and 3.2. Though this property is usually assumed to hold true in the literature, we see here that it follows from the definition of a controlled quantum flow.

4. Sample path properties and the innovations problem

In the remainder of this paper we will use explicitly the properties of quantum filtering equations in recursive form, as given e.g. in Corollaries 3.1 and 3.2. In the next section we will show that under suitable regularity conditions, the quantum optimal control problem is solved by a feedback control policy that at time t is only a function of the normalized solution of the unnormalized filtering equation at that time. In order for this to be sensible, we have to show that the controlled unnormalized filtering equation, and in particular its normalized form, has a unique strong solution. The purpose of this section is to investigate these and related properties of the solutions of recursive quantum filters.

The approach we will take is to convert the entire problem into one of classical stochastic analysis. Note that all the quantities that appear in the unnormalized filtering equations of Corollaries 3.1 and 3.2 are adapted and affiliated to the commutative algebra $\mathcal{Y}_t$. Thus, we may use the spectral theorem to map the filter onto a classical stochastic differential equation driven by the (classical) observations. This will allow us to manipulate the filter by using the Itô change of variables formula for jump-diffusions and puts at our disposal the full machinery of classical stochastic differential equations driven by semimartingales.

We begin by proving the following Proposition. This property will be crucial in the proof of the separation theorem; at this point, however, we are mostly interested in the fact that as a consequence, the observation process Y_t is a semimartingale.

Proposition 4.1 (Innovations martingale). *Recall that* $dZ_t = \Xi_t\, d\Lambda_t + \Upsilon_t\, dA_t^* + \Upsilon_t^*\, dA_t$ *as in Definition 3.1. Define the innovations process*

$$\overline{Z}_t = U_t^* Z_t U_t - \int_0^t \pi_s(\Xi_s L_s^* L_s + \Upsilon_s^* L_s + \Upsilon_s L_s^*)\, ds.$$

Then $\overline{Z}_t$ *is a* $\mathcal{Y}_t$*-martingale, i.e.* $\mathbb{P}(\overline{Z}_t|\mathcal{Y}_s) = \overline{Z}_s$ *for all* $s \le t \in [0,T]$.

Proof. We need to prove that $\mathbb{P}(\overline{Z}_t - \overline{Z}_s|\mathcal{Y}_s) = 0$ for all $s \le t \in [0,T]$, or equivalently $\mathbb{P}((\overline{Z}_t - \overline{Z}_s)K) = 0$ for all $s \le t \in [0,T]$ and $K \in \mathcal{Y}_s$. The

latter can be written as

$$\mathbb{P}(U_t^* Z_t U_t K) - \mathbb{P}(U_s^* Z_s U_s K) = \int_s^t \mathbb{P}(\pi_s(\Xi_s L_s^* L_s + \Upsilon_s^* L_s + \Upsilon_s L_s^*)K)\, ds$$

for all $s \le t \in [0,T]$ and $K \in \mathcal{Y}_s$. Since $\mathcal{Y}_s = U_t^* \mathcal{Z}_s U_t$ for all $s \le t \in [0,T]$, it is sufficient to show that

$$\mathbb{P}(U_t^* Z_t C U_t) - \mathbb{P}(U_s^* Z_s C U_s) = \int_s^t \mathbb{P}(U_t^*(\Xi_s L_s^* L_s C + \Upsilon_s^* L_s C + \Upsilon_s L_s^* C) U_t)\, ds$$

for all $s \le t \in [0,T]$ and $K \in \mathcal{Z}_s$. But this follows directly from (4). □

The innovations process is the starting point for martingale-based approaches to (quantum) filtering[6,7]. The idea there is to obtain a particular martingale which is represented as a stochastic integral with respect to the innovations. The method is complicated, however, by what is known as the innovations problem: it is not clear a priori whether the observations and the innovations generate the same (σ-)algebras[24], which is a prerequisite for the martingale representation theorem. The problem is resolved using a method by Fujisaki-Kallianpur-Kunita, where the Girsanov theorem is used to prove a special martingale representation theorem with respect to the innovations[24]. In contrast, the reference probability method is completely independent from the innovations problem. Though we will not need this fact to prove the separation theorem, we will see at the end of this section that the equivalence of the observations and innovations σ-algebras follows as a corollary from the existence and uniqueness theorems.

To return to the task at hand, it is evident from Proposition 4.1 that the observation process Y_t can be written as the sum of the innovations process, which is a martingale, and a process of finite variation, both of which are affiliated to the algebra generated by the observations. Hence if we map Y_t to its classical counterpart through the spectral theorem, we obtain a classical semimartingale.

Remark 4.1. For convenience, we will abuse our notation somewhat and denote by $Y_t, \overline{Z}_t, \pi_t(X), \sigma_t(X)$ both the corresponding quantum processes and the associated classical processes obtained through the spectral theorem. By $\tilde{\Xi}_t, \tilde{\Upsilon}_t$ we denote the classical processes obtained by applying the spectral theorem to $j_t(\Xi_t)$ and $j_t(\Upsilon_t)$, whereas $\tilde{L}_t, \tilde{H}_t$ are $(\dim \mathscr{H}_0 \times \dim \mathscr{H}_0)$-matrix valued processes obtained by applying the spectral theorem to each matrix element of $j_t(L_t)$ and $j_t(H_t)$. Hence $\tilde{\Xi}_t, \tilde{\Upsilon}_t$ are Y_t-adapted bounded scalar processes, whereas $j_t(L_t), j_t(H_t)$ are Y_t-adapted bounded matrix-valued processes.

We now map the unnormalized filtering equation of Corollary 3.1 to a classical stochastic differential equation. This gives

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s,X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\,ds$$
$$+ \int_0^t \sigma_{s-}(\Upsilon_{s-}^{-1} X L_{s-} + \Upsilon_{s-}^{-1*} L_{s-}^* X + (\Upsilon_{s-}^* \Upsilon_{s-})^{-1} \Xi_{s-}\, L_{s-}^* X L_{s-})\,dY_s$$

and similarly, we obtain the classical equivalent of Corollary 3.2

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s,X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\,ds$$
$$+ \int_0^t \sigma_{s-}(\Xi_{s-}\, L_{s-}^* X L_{s-} + \Upsilon_{s-}^* X L_{s-} + \Upsilon_{s-} L_{s-}^* X - \Xi_{s-}^{-1}(1 - \Upsilon_{s-}^* \Upsilon_{s-})X)\times$$
$$(dY_s - \tilde{\Xi}_s^{-1}(1 - \tilde{\Upsilon}_s^* \tilde{\Upsilon}_s)\,ds)$$

where now the stochastic integrals are classical Itô integrals with respect to the semimartingale Y_t[25] (the fact that we can map a quantum Itô integral with respect to fundamental processes to a classical Itô integral with respect to a semimartingale is verified by approximation.) As we are now dealing with stochastic processes on the level of sample paths, we have to choose a modification such that the processes are well defined—this is an issue that does not occur on the level of QSDEs. We will make the standard choice[25] that all our (semi)martingales are càdlàg, and include explicitly the left limits σ_{s-} etc. to enforce causality.

We are now ready to apply classical stochastic analysis to our quantum filtering equations. We begin by normalizing the equations using the classical Itô formula.

Proposition 4.2 (Nonlinear filtering equation, diffusion case). *Suppose $\Xi_t = 0$ and Υ_t has a bounded inverse for all $t \in [0,T]$. Then*

$$\pi_t(X) = \mathbb{P}(X) + \int_0^t \pi_s(i[H_s,X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\,ds$$
$$+ \int_0^t \big\{\pi_s(\Upsilon_s^{-1} X L_s + \Upsilon_s^{-1*} L_s^* X) - \pi_s(\Upsilon_s^{-1} L_s + \Upsilon_s^{-1*} L_s^*)\,\pi_s(X)\big\}\,d\overline{Z}_s$$

for $X \in \mathcal{B}$, where $\overline{Z}_t$ is the innovations process of Proposition 4.1. Furthermore, there is a $(\dim \mathscr{H}_0 \times \dim \mathscr{H}_0)$-matrix process ρ_t such that $\pi_t(X) = \mathrm{Tr}[X\rho_t]$ for all $X \in \mathcal{B}$, which satisfies the classical Itô stochastic

differential equation (w.r.t. the semimartingale observations Y_t)

$$\rho_t = \rho_0 + \int_0^t \left\{ -i[\tilde{H}_s, \rho_s] + \tilde{L}_s \rho_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \rho_s + \rho_s \tilde{L}_s^* \tilde{L}_s) \right\} ds$$
$$+ \int_0^t \left\{ \tilde{\Upsilon}_s^{-1} \tilde{L}_s \rho_s + \tilde{\Upsilon}_s^{-1*} \rho_s \tilde{L}_s^* - \mathrm{Tr}[(\tilde{\Upsilon}_s^{-1} \tilde{L}_s + \tilde{\Upsilon}_s^{-1*} \tilde{L}_s^*) \rho_s] \, \rho_s \right\} d\overline{Z}_s.$$

Proposition 4.3 (Nonlinear filtering equation, jump case). *Suppose that Ξ_t has a bounded inverse for all $t \in [0, T]$. Then*

$$\pi_t(X) = \mathbb{P}(X) + \int_0^t \pi_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\}) \, ds$$
$$+ \int_0^t \left\{ \frac{\pi_{s-}((\Upsilon_{s-} + \Xi_{s-} L_{s-})^* X (\Upsilon_{s-} + \Xi_{s-} L_{s-}))}{\pi_{s-}((\Upsilon_{s-} + \Xi_{s-} L_{s-})^* (\Upsilon_{s-} + \Xi_{s-} L_{s-}))} - \pi_{s-}(X) \right\} \tilde{\Xi}_{s-}^{-1} \, d\overline{Z}_s$$

for $X \in \mathcal{B}$, where $\overline{Z}_t$ is the innovations process of Proposition 4.1. Furthermore, there is a $(\dim \mathscr{H}_0 \times \dim \mathscr{H}_0)$-matrix process ρ_t such that $\pi_t(X) = \mathrm{Tr}[X \rho_t]$ for all $X \in \mathcal{B}$, which satisfies the classical Itô stochastic differential equation (w.r.t. the semimartingale observations Y_t)

$$\rho_t = \rho_0 + \int_0^t \left\{ -i[\tilde{H}_s, \rho_s] + \tilde{L}_s \rho_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \rho_s + \rho_s \tilde{L}_s^* \tilde{L}_s) \right\} ds$$
$$+ \int_0^t \left\{ \frac{(\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \rho_{s-} (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^*}{\mathrm{Tr}[(\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \rho_{s-} (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^*]} - \rho_{s-} \right\} \tilde{\Xi}_{s-}^{-1} \, d\overline{Z}_s.$$

Proof of Propositions 4.2 and 4.3. In pure diffusion case (Proposition 4.2) the observation process Y_t is a continuous semimartingale, and hence the normalization is easily verified by applying the Itô change of variables formula to the Kallianpur-Striebel formula $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$. In both cases, the conversion to the matrix form follows from the fact that by construction $\pi_t(X)$ is a random linear functional on $\mathcal{B}$, and hence there is a unique random matrix ρ_t such that $\pi_t(X) = \mathrm{Tr}[X \rho_t]$ for all $X \in \mathcal{B}$. The expressions given are easily seen to satisfy this requirement.

It remains to normalize $\sigma_t(X)$ of Proposition 4.3. Here we use the Itô change of variables formula for Stieltjes integrals, but the manipulations are somewhat more cumbersome. We begin by noting that

$$\tilde{Y}_t = \int_0^t \tilde{\Xi}_{s-}^{-1} (dY_s + \tilde{\Xi}_s^{-1} \tilde{\Upsilon}_s^* \tilde{\Upsilon}_s \, ds)$$

is a pure jump process with jumps of unit magnitude. To see this, we

calculate the quadratic variation

$$[\tilde{Y},\tilde{Y}]_t = \tilde{Y}_t^2 - 2\int_0^t \tilde{Y}_{s-}\, d\tilde{Y}_s = \tilde{Y}_t$$

by applying the quantum Itô rules to (4) and using the spectral theorem. $[\tilde{Y},\tilde{Y}]_t$ is by construction an increasing process, so $\tilde{Y}_t$ is also increasing and hence of finite variation. But any adapted, càdlàg finite variation process is a quadratic pure jump semimartingale[25], meaning that

$$\tilde{Y}_t = [\tilde{Y},\tilde{Y}]_t = \sum_{0<s\le t} (\tilde{Y}_s - \tilde{Y}_{s-})^2.$$

As $\tilde{Y}_t$ is an increasing pure jump process and $\tilde{Y}_s - \tilde{Y}_{s-} = (\tilde{Y}_s - \tilde{Y}_{s-})^2$, we conclude that $\tilde{Y}_t$ is a pure jump process with unit magnitude jumps. We can now rewrite the unnormalized filtering equation for the pure jump case in terms of $\tilde{Y}_t$. This gives

$$\begin{aligned}\sigma_t(X) = \mathbb{P}(X) &+ \int_0^t \sigma_s(i[H_s,X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ &+ \int_0^t \sigma_{s-}((\Upsilon_{s-} + \Xi_{s-}L_{s-})^* X (\Upsilon_{s-} + \Xi_{s-}L_{s-}) - X)\,(d\tilde{Y}_s - \tilde{\Xi}_s^{-2}\, ds)\end{aligned}$$

where the integral over $\tilde{Y}_t$ is a simple Stieltjes integral. To normalize $\sigma_t(X)$, we first use the change of variables formula for finite variation processes[25]

$$\begin{aligned}f(V_t) - f(V_0) = &\int_0^t f'(V_{s-})\, dV_s \\ &+ \sum_{0<s\le t} \{f(V_s) - f(V_{s-}) - f'(V_{s-})\,(V_s - V_{s-})\}\end{aligned}$$

to calculate $\sigma_t(I)^{-1}$. We obtain directly that

$$\sigma_s(I) - \sigma_{s-}(I) = (\sigma_{s-}(B_{s-}^* B_{s-}) - \sigma_{s-}(I))\,(\tilde{Y}_s - \tilde{Y}_{s-})$$

where $B_s = \Upsilon_s + \Xi_s L_s$. Similarly we have

$$\sigma_s(I)^{-1} - \sigma_{s-}(I)^{-1} = (\sigma_{s-}(B_{s-}^* B_{s-})^{-1} - \sigma_{s-}(I)^{-1})\,(\tilde{Y}_s - \tilde{Y}_{s-}).$$

We thus express the correction term in the change of variables formula as

$$\begin{aligned}\sum_{0<s\le t} \{\sigma_s(I)^{-1} - \sigma_{s-}(I)^{-1} + \sigma_{s-}(I)^{-2}\,(\sigma_s(I) - \sigma_{s-}(I))\} = \\ \int_0^t (\pi_{s-}(B_{s-}^* B_{s-}) + \pi_{s-}(B_{s-}^* B_{s-})^{-1} - 2)\,\sigma_{s-}(I)^{-1}\, d\tilde{Y}_s\end{aligned}$$

and we obtain

$$\sigma_t(I)^{-1} = 1 + \int_0^t \frac{\pi_s(B_s^* B_s) - 1}{\sigma_s(I)}\, \tilde{\Xi}_s^{-2}\, ds + \int_0^t \frac{\pi_{s-}(B_{s-}^* B_{s-})^{-1} - 1}{\sigma_{s-}(I)}\, d\tilde{Y}_s.$$

Finally, applying the integration by parts formula for Stieltjes integrals to the Kallianpur-Striebel formula $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$ gives

$$\begin{aligned}\pi_t(X) = \mathbb{P}(X) &+ \int_0^t \pi_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ &+ \int_0^t \left\{ \frac{\pi_{s-}(B_{s-}^* X B_{s-})}{\pi_{s-}(B_{s-}^* B_{s-})} - \pi_{s-}(X) \right\} (d\tilde{Y}_s - \tilde{\Xi}_s^{-2}\, \pi_s(B_s^* B_s)\, ds).\end{aligned}$$

and it remains to notice that $\tilde{\Xi}_{s-}^{-1}\, d\overline{Z}_s = d\tilde{Y}_s - \tilde{\Xi}_s^{-2}\, \pi_s(B_s^* B_s)\, ds$. □

Now that we have obtained the filtering equations in their sample path form, let us study the question of existence and uniqueness of solutions.

Proposition 4.4 (Existence and uniqueness). *Let $\tilde{\Xi}_s$, $\tilde{\Upsilon}_s$ and the matrix elements of $\tilde{L}_s$ and $\tilde{H}_s$ have càdlàg sample paths. Then both the unnormalized and nonlinear filtering equations have a unique strong solution with respect to Y_t. Moreover, matrix ρ_t is a.s. positive with unit trace, i.e. a density matrix, for every t.*

Proof. As $\sigma_t(X)$ is linear by construction, we can find a matrix process τ_t such that $\sigma_t(X) = \mathrm{Tr}[X\tau_t]$. As we did for ρ_t, we obtain directly that τ_t must satisfy

$$\begin{aligned}\tau_t = \rho_0 &+ \int_0^t \left\{ -i[\tilde{H}_s, \tau_s] + \tilde{L}_s \tau_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \tau_s + \tau_s \tilde{L}_s^* \tilde{L}_s) \right\} ds \\ &+ \int_0^t \left\{ \tilde{\Upsilon}_s^{-1} \tilde{L}_s \tau_s + \tilde{\Upsilon}_s^{-1*} \tau_s \tilde{L}_s^* \right\} dY_s,\end{aligned}$$

in the pure diffusion case, and in the pure jump case τ_t must satisfy

$$\begin{aligned}\tau_t = \rho_0 &+ \int_0^t \left\{ -i[\tilde{H}_s, \tau_s] + \tilde{L}_s \tau_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \tau_s + \tau_s \tilde{L}_s^* \tilde{L}_s) \right\} ds \\ &+ \int_0^t \left\{ (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \tau_{s-} (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^* - \tau_{s-} \right\} \times \\ &\qquad \tilde{\Xi}_{s-}^{-1} (dY_s - \tilde{\Xi}_s^{-1} (1 - \tilde{\Upsilon}_s^* \tilde{\Upsilon}_s)\, ds).\end{aligned}$$

Both these equations are finite-dimensional linear stochastic differential equations with càdlàg coefficients, for which the existence of a unique strong

solution is a standard result[25]. Thus the unique solution τ_t of these equations must indeed satisfy $\sigma_t(X) = \mathrm{Tr}[X\tau_t]$ (this need not be true if the solution were not unique—then it could be the case that only one of the solutions coincides with $\sigma_t(\cdot)$.)

To demonstrate the existence of a solution to the equations for ρ_t, note that by construction $\sigma_t(X)$ is a positive map (we will always assume that ρ_0 is chosen so that $\sigma_0(X) = \pi_0(X) = \mathrm{Tr}[X\rho_0]$ is a state, i.e. ρ_0 is a positive matrix with unit trace.) Hence by uniqueness τ_t must be a positive matrix for all t. Moreover, as the linear solution map $\tau_0 \mapsto \tau_t$ is a.s. invertible[25] τ_t is for all t a.s. not the zero matrix. This means that for all t the process $\rho_t = \tau_t/\mathrm{Tr}[\tau_t]$ is well-defined and satisfies the nonlinear filtering equations for ρ_t which we obtained previously. Hence we have explicitly constructed a solution to the equations for ρ_t, and moreover ρ_t is a.s. a positive, unit trace matrix for every t. Finally, from the Kallianpur-Striebel formula it is evident that $\pi_t(X) = \mathrm{Tr}[X\rho_t]$.

It remains to prove that there are no other solutions ρ_t that satisfy the nonlinear filtering equations, i.e. that the solution ρ_t constructed above is unique. To this end, suppose that there is a different solution $\bar\rho_t$ with $\bar\rho_0 = \rho_0$ that also satisfies the nonlinear filtering equation (with respect to dY_t). Define $\bar\Pi_t$ as the unique strong solution of the linear equation

$$\bar\Pi_t = 1 + \int_0^t \mathrm{Tr}[(\tilde\Upsilon_s^{-1}\tilde L_s + \tilde\Upsilon_s^{-1*}\tilde L_s^*)\bar\rho_s]\bar\Pi_s\, dY_s$$

in the pure diffusion case, and as the unique strong solution of

$$\bar\Pi_t = 1 + \int_0^t \Big\{\mathrm{Tr}[(\tilde\Upsilon_{s-} + \tilde\Xi_{s-}\tilde L_{s-})^*(\tilde\Upsilon_{s-} + \tilde\Xi_{s-}\tilde L_{s-})\bar\rho_{s-}] - 1\Big\}\bar\Pi_{s-}\times \tilde\Xi_{s-}^{-1}(dY_s - \tilde\Xi_s^{-1}(1 - \tilde\Upsilon_s^*\tilde\Upsilon_s)\, ds)$$

in the pure jump case. Using the Itô rules, one can verify that $\bar\Pi_t\bar\rho_t$ satisfies the same equation as τ_t. But τ_t is uniquely defined; hence we conclude that $\rho_t = (\bar\Pi_t/\mathrm{Tr}[\tau_t])\bar\rho_t$. By taking the trace of the nonlinear filtering equations it is easily verified that $\mathrm{Tr}[\rho_t] = \mathrm{Tr}[\bar\rho_t] = 1$ a.s. for all t, and hence a.s. $\rho_t = \bar\rho_t$. The proof is complete. □

To conclude the section, we will now prove that the innovations problem can be solved for the class of systems that we have considered under some mild conditions on the controls. It is likely that the innovations problem can be solved under more general conditions; however, as we will not need this result in the following, we restrict ourselves to the following case for sake of demonstration.

Proposition 4.5 (Innovations problem). *Let $\tilde{\Upsilon}_s$ and the matrix elements of $\tilde{L}_s$ and $\tilde{H}_s$ be càdlàg semimartingales that are adapted to the filtration generated by the innovations process. The the observations Y_t and the innovations $\overline{Z}_t$ generate the same σ-algebras.*

Proof. First, note that we can restrict ourselves to the case of diffusive observations. In the case of counting observations the result is trivial: as $Y_t - \overline{Z}_t$ is a continuous process, Y_t can be completely recovered as the discontinuous part of $\overline{Z}_t$. Moreover, $\overline{Z}_t$ is Y_t-measurable by construction. Hence Y_t and $\overline{Z}_t$ generate the same σ-algebras.

Things are not so simple in the diffusive case. Denote by Σ_t^Y the σ-algebra generated by Y_t up to time t, and similarly by $\Sigma_t^{\overline{Z}}$ the σ-algebra generated by $\overline{Z}_t$ up to time t. The inclusion $\Sigma_t^{\overline{Z}} \subset \Sigma_t^Y$ holds true by construction, so we are burdened with proving the opposite inclusion $\Sigma_t^Y \subset \Sigma_t^{\overline{Z}}$. This is essentially an issue of "causal invertibility": given only the stochastic process $\overline{Z}_t$, can we find a map that recovers the process Y_t in a causal manner? Clearly this would be the case[24] if the nonlinear filtering equation has a unique strong solution with respect to $\overline{Z}_t$; as we have already established that it has a unique strong solution with respect to Y_t the two solutions must then coincide, after which we can recover Y_t from the formula $dY_t = d\overline{Z}_t + \mathrm{Tr}[(\tilde{\Upsilon}_t^* \tilde{L}_t + \tilde{\Upsilon}_t \tilde{L}_t^*)\rho_t]\,dt$. Our approach will be precisely to demonstrate the uniqueness of ρ_t with respect to $\overline{Z}_t$.

To this end, consider the diffusive nonlinear filtering equation for ρ_t given in Proposition 4.2, where we now consider it to be driven directly by the martingale $\overline{Z}_t$ rather than by the observations Y_t. Now introduce the following quantities: let X_t be a vector which contains as entries all the matrix elements of ρ_t, $\tilde{L}_t$, $\tilde{H}_t$, and the process $\tilde{\Upsilon}_t$, and let K_t be a vector that contains as entries all the matrix elements of $\tilde{L}_t$, $\tilde{H}_t$, the process $\tilde{\Upsilon}_t$, and $\overline{Z}_t$. Then K_t is a vector of semimartingales and we can rewrite the nonlinear filtering equation in the form

$$X_t = X_0 + \int_0^t f(X_{s-})\,dK_s$$

for a suitably chosen matrix function f. By inspection, we see that $f(X)$ is polynomial in the elements of X and hence f is a locally Lipschitz function. Thus there is a unique solution of the nonlinear filtering equation with respect to K_s up to an accessible explosion time ζ[25]. By uniqueness, the nonlinear filtering solution ρ_t with respect to $\overline{Z}_t$ must coincide with the solution with respect to Y_t up to the explosion time ζ. This also implies that for all $t < \zeta$, the solution with respect to $\overline{Z}_t$ must be a positive unit

trace matrix. But the set of all such matrices is compact, and hence the accessibility of ζ is violated unless $\zeta = \infty$ a.s. We conclude that the unique solution ρ_t with respect to $\overline{Z}_t$ exists for all time. □

5. A separation theorem

We are now finally ready to consider the control problem for quantum diffusions; i.e., how do we choose the processes L_t, H_t, etc. in the controlled quantum flow to achieve a certain control goal? As we will be comparing different control strategies, we begin by introducing some notation which allow us to keep them apart.

Definition 5.1 (Control strategy). *A strategy $\mu = (\Xi, \Upsilon, S, L, H)$ is a collection of processes $\Xi_t, \Upsilon_t, S_t, L_t, H_t$ defined on $[0,T]$ that satisfy the conditions of Definition 3.1. Given μ, we denote by $\mu_t = (\Xi_t, \Upsilon_t, S_t, L_t, H_t)$ the controls at time t, by $\mu_{t]} = (\Xi_{[0,t]}, \Upsilon_{[0,t]}, S_{[0,t]}, L_{[0,t]}, H_{[0,t]})$ the strategy on the interval $[0,t]$, and similarly by $\mu_{[t}$ the strategy on the interval $[t,T]$.*

Each control strategy μ defines a different controlled quantum flow. To avoid confusion, we will label the various quantities that are derived from the controlled flow by the associated control strategy. For example, $j_t^\mu(X)$ and Y_t^μ are the flow and observations process obtained under the control strategy μ, etc.

Our next task is to specify the control goal. To this end, we introduce a cost function which quantifies how successful a certain control strategy is deemed to be. The best control strategy is the one that minimizes the cost.

Definition 5.2 (Cost function). *Let C^μ be a process of positive operators, possibly dependent on the control strategy μ, such that C_t^μ is affiliated to $\mathrm{vN}(\mu_t, \mathcal{B})$, the algebra generated by the initial system and the control strategy at time t, for each $t \in [0,T]$. Let $C_T \in \mathcal{B}$. The total cost is defined by the functional*

$$J[\mu] = \int_0^T j_t^\mu(C_t^\mu)\,dt + j_T^\mu(C_T).$$

C_t^μ and C_T are called the running and terminal cost operators, respectively.

Ultimately our goal will be to find, if possible, an optimal control μ^* that minimizes the expected total cost $\mathbb{P}(J[\mu])$. Let us begin by converting the latter into a more useful form. Using the tower property of conditional

expectations, we have

$$\mathbb{P}(J[\mu]) = \mathbb{P}\left(\int_0^T \mathbb{P}(j_t^\mu(C_t^\mu)|\mathcal{Y}_t^\mu)\,dt + \mathbb{P}(j_T^\mu(C_T)|\mathcal{Y}_t^\mu)\right)$$

where $\mathcal{Y}_t^\mu$ is the algebra generated by Y_s^μ up to time t (note that the conditional expectations are well defined as $j_t^\mu(C_t^\mu)$ is affiliated to $j_t^\mu(\mathcal{B}) \otimes \mathcal{Y}_t^\mu$.) But then

$$\mathbb{P}(J[\mu]) = \mathbb{P}\left(\int_0^T \pi_t^\mu(C_t^\mu)\,dt + \pi_T^\mu(C_T)\right),$$

and we see that the expected cost can be calculated from the associated filtering equation only. As the filter lives entirely in the commutative algebra $\mathcal{Y}_T^\mu$ we can now proceed, as in the previous section, by converting the problem into a classical stochastic problem. To this end, we use the spectral theorem to map the commutative quantum probability space $(\mathcal{Y}_T^\mu, \mathbb{P})$ to the classical space $L^\infty(\Omega^\mu, \Sigma_T^\mu, \nu^\mu, \mathbf{P}^\mu)$ (note that different control strategies may not give rise to commuting observations, and hence the classical space depends on μ.) Thus we obtain the classical expression

$$\mathbb{P}(J[\mu]) = \mathbb{E}_{\mathbf{P}^\mu}\left(\int_0^T \pi_t^\mu(C_t^\mu)\,dt + \pi_T^\mu(C_T)\right).$$

We will use the same notations as in the previous section for the classical stochastic processes associated to L_t, H_t, etc. In addition, we denote by Σ_t^μ the σ-algebra generated by the observations Y_s^μ up to time t.

We have now defined the cost as a classical functional for an arbitrary control strategy. In practice not every control policy is allowed, however. First, note that in practical control scenarios only a limited number of controls are physically available; e.g. in the example of Hamiltonian feedback $H_t = u(t)H$ we can only modulate the strength $u(t)$ of a fixed Hamiltonian H, and we certainly cannot independently control every matrix element of S_t, L_t, etc. Moreover we have expressed the (classical) cost in terms of the filter state $\pi_t^\mu(\cdot)$; hence we should impose sufficient regularity conditions on the controls so that we can unambiguously obtain the filtered estimate from the observations using the nonlinear filtering equation of the previous section. To this end we introduce an admissible subspace of control strategies that are realizable in the control scenario of interest, and we require the solution of the optimal control problem to be an admissible control.

Definition 5.3 (Admissible controls). *Define the admissible range* $\mathscr{B}_t \subset \mathbb{R} \times \mathbb{C} \times \mathcal{B} \times \mathcal{B} \times \mathcal{B}$ *for every* $t \in [0, T]$. *A control strategy* μ *is*

admissible if $\mu_t \in \mathscr{B}_t$ *a.s. (in the sense that* $(\tilde{\Xi}_t, \tilde{\Upsilon}_t, \tilde{S}_t, \tilde{L}_t, \tilde{H}_t) \in \mathscr{B}_t$ *a.s.) for every* $t \in [0,T]$ *and* μ *has càdlàg sample paths. The set of all admissible control strategies is denoted by* $\mathscr{C}$*.*

Remark 5.1. Note that in order to satisfy Definition 3.1, $\mathscr{B}_t$ should be chosen to be a bounded set such that the only admissible S are unitary matrices and the only admissible H are self-adjoint matrices.

The optimal control problem is to find, if possible, an admissible control strategy μ^* that minimizes the expected total cost, i.e. to find a $\mu^* \in \mathscr{C}$ such that

$$\mathbb{P}(J[\mu^*]) = \min_{\mu \in \mathscr{C}} \mathbb{P}(J[\mu]).$$

In principle μ_t^* could depend on the entire history of observations up to time t. This would be awkward, as it would mean that the controller should have enough memory to record the entire observation history and enough resources to calculate a (possibly extremely complicated) functional thereof. However, as the cost functional only depends on the filter, one could hope that μ_t^* would only depend on ρ_t^μ, the solution of the nonlinear filtering equation at time t. This is a much more desirable situation as ρ_t^μ can be calculated recursively: hence we would not need to remember the previous observations, and the feedback at time t could be calculated simply by applying a measurable function to ρ_t^μ. A control strategy that separates into a filtering problem and a control map is called a separated strategy.

Definition 5.4 (Separated controls). *An admissible control strategy* $\mu \in \mathscr{C}$ *is said to be separated if there exists for every* $t \in [0,T]$ *a measurable map* $u_t^\mu : \mathcal{S}(\mathcal{B}) \to \mathscr{B}_t$ *such that* $\mu_t = u_t^\mu(\rho_t^\mu)$*, where* ρ_t^μ *is the matrix solution of the nonlinear filtering equation at time* t *and* $\mathcal{S}(\mathcal{B})$ *is the set of positive matrices in* $\mathcal{B}$ *with unit trace. The set of all separated admissible strategies is denoted by* $\mathscr{C}^0$*.*

The main technique for solving optimal control problems in discrete time is dynamic programming, a recursive algorithm that operates backwards in time to construct an optimal control strategy. The infinitesimal form of the dynamic programming recursion, Bellman's functional equation, provides a candidate optimal control strategy in separated form. The goal of this section is to prove that if we can find a separated strategy $\mu \in \mathscr{C}^0$ that satisfies Bellman's equation, then this strategy is indeed optimal with respect to *all* control strategies in $\mathscr{C}$, i.e. even those that are not separated. This result is known as a separation theorem. In classical stochastic control this

result was established for linear systems in a classic paper by Wonham[12] and for finite-state Markov processes by Segall[13]. The proof of the quantum separation theorem below proceeds along the same lines.

We begin by introducing the expected cost-to-go, i.e. the cost incurred on an interval $[t,T]$ conditioned on the observations up to time t.

Definition 5.5 (Expected cost-to-go). *Given an admissible control strategy μ, the expected cost-to-go at time t is defined as the random variable*

$$W[\mu](t) = \mathbb{E}_{\mathbf{P}^\mu}\left(\int_t^T \pi_s^\mu(C_s^\mu)\,ds + \pi_T^\mu(C_T)\,\middle|\,\Sigma_t^\mu\right).$$

The basic idea behind dynamic programming is as follows. Regardless of what conditional state we have arrived at at time t, an optimal control strategy should be such that the expected cost incurred over the remainder of the control time is minimized; in essence, an optimal control should minimize the expected cost-to-go. This is Bellman's principle of optimality. To find a control that satisfies this requirement, one could proceed in discrete time by starting at the final time T, and then performing a recursion backwards in time such that at each time step the control is chosen to minimize the expected cost-to-go. We will not detail this procedure here; see e.g.[26]. Taking the limit as the time step goes to zero gives the infinitesimal form of this recursion, i.e. Bellman's functional equation

$$-\frac{\partial V}{\partial t}(t,\theta) = \min_{u\in\mathscr{B}_t}\Big\{\mathscr{L}(u)V(t,\theta) + \mathrm{Tr}[\theta\,\tilde{C}_t^u]\Big\}, \qquad t\in[0,T],\ \theta\in\mathcal{S}(\mathcal{B})$$

subject to the terminal conditon $V(T,\theta) = \mathrm{Tr}[\theta\,C_T]$ (recall that C_t^μ is affiliated to $\mathrm{vN}(\mu_t,\mathcal{B})$, and hence $\tilde{C}_t^\mu = \tilde{C}_t^{\mu_t}$ can be considered a $\mathcal{B}$-valued measurable function of μ_t.) The value function $V(t,\theta)$ essentially represents the expected cost-to-go conditioned on the event that the solution ρ_t^μ of the nonlinear filtering equation takes the value θ at time t. If the minimum in Bellman's equation can be evaluated for all t and θ, then it defines a separated control strategy μ by

$$u_t^\mu(\theta) = \operatorname*{argmin}_{u\in\mathscr{B}_t}\Big\{\mathscr{L}(u)V(t,\theta) + \mathrm{Tr}[\theta\,\tilde{C}_t^u]\Big\}, \qquad t\in[0,T],\ \theta\in\mathcal{S}(\mathcal{B}).$$

Here $\mathscr{L}(u)$ denotes the infinitesimal generator of the matrix nonlinear filtering equation given the control u; i.e., it is the operator that satisfies for

any admissible control strategy μ the Itô change of variables formula

$$f(t,\rho_t^\mu) = f(s,\rho_s^\mu) + \int_s^t \left\{ \frac{\partial f}{\partial \sigma}(\sigma,\rho_\sigma^\mu) + [\mathscr{L}(\mu_\sigma) f](\sigma,\rho_\sigma^\mu) \right\} d\sigma + \int_s^t G_{\sigma-}^{\mu,f}(\rho_{\sigma-}^\mu)\, d\overline{Z}_\sigma$$

where f is any sufficiently differentiable function (C^2 in the diffusive case, C^1 in the pure jump case.) The expression for $\mathscr{L}(u)$ and $G_\sigma^{\mu,f}$ is standard[25] and can be obtained directly from Propositions 4.2 and 4.3.

Our brief discussion of dynamic programming is intended purely as a motivation for what follows. Even if we had given a rigorous description of the procedure, the solution of Bellman's equation can only give a candidate control strategy and one must still show that this control strategy is indeed optimal. Thus, rather than deriving Bellman's equation, we will now take it as our starting point and show that if we can find a separated control that solves it, then this control is optimal with respect to all admissible controls (i.e. there does not exist an admissible control strategy that achieves a lower expected total cost.)

Theorem 5.1 (Separation theorem). *Suppose there exists a separated admissible control strategy $\mu \in \mathscr{C}^0$ and a function $V : [0,T] \times \mathcal{S}(\mathcal{B}) \to \mathbb{R}$ such that*

(1) The function V is C^1 in the first variable and C^2 in the second variable (diffusive case), or C^1 in both variables (pure jump case).

(2) For all $t \in [0,T]$ and $\theta \in \mathcal{S}(\mathcal{B})$, the function V satisfies

$$\frac{\partial V}{\partial t}(t,\theta) + \mathscr{L}(u_t^\mu(\theta)) V(t,\theta) + \mathrm{Tr}\left[\theta\, \tilde{C}_t^{u_t^\mu(\theta)}\right] = 0.$$

(3) For all $t \in [0,T]$, $u \in \mathscr{B}_t$ and $\theta \in \mathcal{S}(\mathcal{B})$, the function V satisfies

$$\frac{\partial V}{\partial t}(t,\theta) + \mathscr{L}(u) V(t,\theta) + \mathrm{Tr}[\theta\, \tilde{C}_t^u] \geq 0.$$

(4) For all $\theta \in \mathcal{S}(\mathcal{B})$ the function V satisfies the terminal condition

$$V(T,\theta) = \mathrm{Tr}[\theta\, C_T].$$

Then the separated strategy μ is optimal in the class of all admissible strategies $\mathscr{C}$, i.e. $\mathbb{P}(J[\mu]) = \min_{\mu' \in \mathscr{C}} \mathbb{P}(J[\mu'])$.

Proof. We begin by showing that for the candidate optimal control μ, the value function $V(t,\rho_t^\mu)$ evaluated at the solution of the nonlinear filtering equation at time t equals the expected cost-to-go $W[\mu](t)$. To this end, we

substitute condition (2) with $\theta = \rho_t^\mu$ and the terminal condition (4) into Definition 5.5. This gives

$$W[\mu](t) = \mathbb{E}_{\mathbf{P}^\mu}\left(V(T,\rho_T^\mu) - \int_t^T \left\{ \frac{\partial V}{\partial s}(s,\rho_s^\mu) + [\mathscr{L}(u_s^\mu(\rho_s^\mu))V](s,\rho_s^\mu) \right\} ds \,\middle|\, \Sigma_t^\mu \right).$$

The purpose of condition (1) is to ensure that we can apply the Itô change of variables formula to $V(t,\rho_t^\mu)$. This gives

$$W[\mu](t) = \mathbb{E}_{\mathbf{P}^\mu}\left(V(t,\rho_t^\mu) + \int_t^T G_{s-}^{\mu,V}(\rho_{s-}^\mu)\, d\overline{Z}_s \,\middle|\, \Sigma_t^\mu \right).$$

But by a fundamental property of stochastic integrals[25], the stochastic integral of the bounded process $G_{t-}^{\mu,V}(\rho_{t-}^\mu)$ against the square-integrable martingale $\overline{Z}_t$ is itself a martingale. Hence the conditional expectation of the second term vanishes, and as ρ_t^μ is Σ_t^μ-measurable we obtain immediately

$$W[\mu](t) = V(t,\rho_t^\mu).$$

In particular, as $\rho_0^\mu = \rho$, we find that $V(0,\rho)$ equals the expected total cost

$$V(0,\rho) = \mathbb{P}(J[\mu]). \tag{5}$$

To show that μ is optimal, let $\psi \in \mathscr{C}$ be an arbitrary admissible control strategy. We follow essentially the same argument as before, but now in the opposite direction. Using the Itô change of variable formula, we find

$$V(t,\rho_t^\psi) = \mathbb{E}_{\mathbf{P}^\psi}\left(V(T,\rho_T^\psi) - \int_t^T \left\{ \frac{\partial V}{\partial s}(s,\rho_s^\psi) + [\mathscr{L}(\psi_s)V](s,\rho_s^\psi) \right\} ds \,\middle|\, \Sigma_t^\psi \right).$$

Using condition (4) and the inequality (3) with $\theta = \rho_t^\psi$ and $u = \psi_t$,

$$V(t,\rho_t^\psi) \le \mathbb{E}_{\mathbf{P}^\psi}\left(\pi_T^\psi(C_T) + \int_t^T \pi_s^\psi(C_s^\psi)\, ds \,\middle|\, \Sigma_t^\psi \right) = W[\psi](t).$$

But $\rho_0^\psi = \rho_0^\mu = \rho$, so together with Equation (5) we obtain the inequality

$$\mathbb{P}(J[\mu]) = V(0,\rho) \le \mathbb{P}(J[\psi])$$

for any $\psi \in \mathscr{C}$, which is the desired result. □

Remark 5.2. In a detailed study of the optimal control problem, separation theorems are only a first step. What we have not addressed here are

conditions under which Bellman's equation can in fact be shown have a solution, nor have we discussed conditions on the feedback function u_t^μ of a separated control strategy that guarantee that the closed-loop system is well-defined (i.e., that it gives rise to càdlàg controls.) These issues have yet to be addressed in the quantum case.

The simplicitly of the separation argument makes such an approach particularly powerful. The argument is ideally suited for quantum optimal control, as we only need to compare the solution V of a deterministic PDE to the solution of the controlled quantum filter separately for every control strategy. Hence we do not need to worry about the fact that different strategies give rise to different, mutually incompatible observation algebras—the corresponding filters are never compared directly. The argument is readily extended, without significant modifications, to a wide variety of optimal control problems: on a finite and infinite time horizon, with a free terminal time or with time-average costs (see e.g.[27]).

Another option is to rewrite the cost directly in terms of the unnormalized filtering equation, for example in the diffusive case:

$$\begin{aligned}\mathbb{P}(J[\mu]) &= \mathbb{P}\left(\int_0^T U_t^{\mu *} C_t^\mu U_t^\mu \, dt + U_T^{\mu *} C_T U_T^\mu\right) \\ &= \mathbb{P}\left(\int_0^T V_t^{\mu *} C_t^\mu V_t^\mu \, dt + V_T^{\mu *} C_T V_T^\mu\right) \\ &= \mathbb{R}^\mu\left(\int_0^T \sigma_t^\mu(C_t^\mu)\, dt + \sigma_T^\mu(C_T)\right),\end{aligned}$$

where $\mathbb{R}^\mu(X) = \mathbb{P}(U_T^\mu X U_T^{\mu *})$. Note that under $\mathbb{R}^\mu$, the observation process Y_t^μ is a Wiener process. Hence we can now perform dynamic programming, and find a separation theorem, directly in terms of the unnormalized filtering equation. This is particularly useful, for example, in treating the risk-sensitive control problem[28].

Finally, a class of interesting quantum control problems can be formulated using the theory of quantum stopping times[29]; this gives rise to optimal stopping problems and impulse control problems in the quantum context. Such control problems are explored in[30] using a similar argument to the one used above.

Acknowledgments

This work was supported by the NSF under contract PHY-0456720, and by the ARO under contract DAAD19-03-1-0073.

Bibliography

1. V. P. Belavkin, *Autom. Remote Control* **44**, 178 (1983).
2. V. P. Belavkin, Nondemolition stochastic calculus in Fock space and nonlinear filtering and control in quantum systems, in *Proceedings XXIV Karpacz winter school*, eds. R. Guelerak and W. Karwowski (World Scientific, 1988).
3. M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty and H. Mabuchi, *Phys. Rev. Lett.* **89**, p. 133602 (2002).
4. J. M. Geremia, J. K. Stockton and H. Mabuchi, *Science* **304**, 270 (2004).
5. R. L. Hudson and K. R. Parthasarathy, *Comm. Math. Phys.* **93**, 301 (1984).
6. V. P. Belavkin, *J. Multivar. Anal.* **42**, 171 (1992).
7. L. Bouten, M. I. Guţă and H. Maassen, *J. Phys. A* **37**, 3189 (2004).
8. L. Bouten and R. Van Handel, Quantum filtering: a reference probability approach (2005), preprint, arXiv:math-ph/0508006.
9. R. Van Handel, J. K. Stockton and H. Mabuchi, *IEEE Trans. Automat. Control* **50**, 768 (2005).
10. M. Mirrahimi and R. Van Handel, Stabilizing feedback controls for quantum systems (2005), *SIAM J. Control Optim.*, at press, arXiv:quant-ph/0510066.
11. M. Zakai, *Z. Wahrsch. th. verw. Geb.* **11**, 230 (1969).
12. W. M. Wonham, *SIAM J. Control* **6**, 312 (1968).
13. A. Segall, *IEEE Trans. Automat. Control* **AC-22**, 179 (1977).
14. L. Bouten and R. Van Handel, Controlled quantum stochastic processes (2007), in preparation.
15. M. Takesaki, *J. Funct. Anal.* **9**, 306 (1971).
16. P. Biane, Calcul stochastique noncommutatif, in *Lectures on Probability Theory, Saint-Flour, 1993*, ed. P. Bernard, Lecture Notes in Mathematics, Vol. 1608 (Springer, 1995).
17. P.-A. Meyer, *Quantum Probability for Probabilists* (Springer, 1993).
18. K. R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus* (Birkhäuser, 1992).
19. A. S. Holevo, *J. Soviet Math.* **56**, 2609 (1991), Translation of Itogi Nauki i Tekhniki, ser. sovr. prob. mat. 36, 3–28, 1990.
20. A. S. Holevo, *Proc. Roy. Soc. Edinburgh Sect. A* **126**, 375 (1996).
21. D. Goswami, J. M. Lindsay and S. J. Wills, *Math. Ann.* **319**, 647 (2001).
22. L. Bouten, Filtering and control in quantum optics (2004), Ph.D. thesis.
23. H. M. Wiseman, *Phys. Rev. Lett.* **75**, 4587 (1995).
24. R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes I: General Theory*, second edn. (Springer, 2001).
25. P. E. Protter, *Stochastic Integration and Differential Equations*, second edn. (Springer, 2004).
26. H. Kushner, *Introduction to stochastic control* (Holt Rinehart Winston, 1971).

27. B. Øksendal and A. Sulem, *Applied stochastic control of jump diffusions* (Springer, 2005).
28. M. R. James, *J. Opt. B Quantum Semiclass. Opt.* **7**, S198 (2005).
29. K. R. Parthasarathy and K. B. Sinha, *Probab. Th. Rel. Fields* **75**, 317 (1987).
30. R. Van Handel, Stopping times in quantum control (2006), in preparation.

CONCILIATION OF BAYES AND POINTWISE QUANTUM STATE ESTIMATION

RICHARD D. GILL

Mathematical Institute, Leiden University,
and EURANDOM*, Eindhoven, Netherlands*
E-mail: gill@math.leidenuniv.nl
www.math.leidenuniv.nl/∼gill

We derive an asymptotic lower bound on the Bayes risk when N identical quantum systems whose state depends on a vector of unknown parameters are jointly measured in an arbitrary way and the parameters of interest estimated on the basis of the resulting data. The bound is an integrated version of a quantum Cramér-Rao bound due to Holevo[13] and it thereby links the fixed N exact Bayesian optimality usually pursued in the physics literature with the pointwise asymptotic optimality favoured in classical mathematical statistics. By heuristic arguments the bound can be expected to be sharp. This does turn out to be the case in various important examples, where it can be used to prove asymptotic optimality of interesting and useful measurement-and-estimation schemes. On the way we obtain a new family of "dual Holevo bounds" of independent interest.

The paper is dedicated to Slava Belavkin in recognition of his pioneering work on quantum Cramér-Rao bounds, on the occasion of his 60th birthday. A more complete version will appear in *Annals of Statistics*.

Keywords: Statistical quantum information bounds, van Trees inequality, Cramér-Rao bounds

1. Introduction

The aim of this paper is to derive asymptotic information bounds for "quantum i.i.d. models" in quantum statistics. That is to say, one has N copies of a quantum system each in the same state depending on an unknown vector of parameters θ, and one wishes to estimate θ, or more generally a vector function of the parameters $\psi(\theta)$, by making some measurement on the N systems together. This yields data whose distribution depends on θ and on the choice of the measurement. Given the measurement, we therefore have a classical parametric statistical model, though not necessarily an i.i.d. model, since we are allowed to bring the N systems together before

measuring the resulting joint system as one quantum object. In that case the resulting data need not consist of (a function of) N i.i.d. observations, and a key quantum feature is that we can generally extract more information about θ using such "collective" or "joint" measurements than when we measure the systems separately. What is the best we can do as $N \to \infty$, when we are allowed to optimize both over the measurement and over the ensuing data-processing?

A heuristic, statistically motivated, approach to deriving methods with good properties for large N is to choose the measurement to optimize the Fisher information in the data, leaving it to the statistician to process the data efficiently, using for instance maximum likelihood or related methods, including Bayesian. This heuristic principle has already been shown to work in a number of special cases in quantum statistics. Since the measurement maximizing the Fisher information typically depends on the unknown parameter value this often has to be implemented in a two-step approach, first using a small fraction of the N systems to get a first approximation to the true parameter, and then optimizing on the remaining systems using this rough guess.

The approach favoured by many physicists is to choose a prior distribution and loss function on grounds of symmetry and physical interpretation, and then to *exactly* optimize the Bayes risk over all measurements and estimators, for any given N. This approach succeeds in producing attractive methods on those rare occasions when a felicitous combination of all the mathematical ingredients leads to a simple and analytically tractable solution. Now it has been observed in a number of problems that the two approaches result in asymptotically equivalent estimators, though the measurement schemes can be strikingly different. Heuristically, this can be understood to follow from the fact that, in the physicists' approach, for large N the prior distribution should become increasingly irrelevant and the Bayes optimal estimator close to the maximum likelihood estimator. Moreover, we expect those estimators to be asymptotically normal with variances corresponding to inverse Fisher information.

Here we link the two approaches by deriving a sharp asymptotic lower bound on the Bayes risk of the physicists' approach, in terms of the optimal Fisher information of the statisticians' approach. This enables us to conclude the asymptotic optimality of some heuristically motivated measurement-and-estimation schemes by showing that they attain the asymptotic bound. Sometimes one can find in this way asymptotically optimal solutions which are much easier to implement than the exactly optimal

solution of the physicists' approach. On the other hand, it also shows (if only heuristically) that the physicists' approach, when successful, leads to procedures which are *asymptotically* optimal for other prior distributions than those used in the computation, also for loss functions only locally equivalent to their loss function of choice, and also asymptotically optimal in a pointwise rather than a Bayesian sense.

We derive our main result by combining an existing quantum Cramér-Rao bound[13] with the van Trees inequality, a Bayesian Cramér-Rao bound from classical statistics[6,15]. The former can be interpreted as a bound on the Fisher information in an arbitrary measurement on a quantum system, the latter is a bound on the Bayes risk (for a quadratic loss function) in terms of the Fisher information in the data. This means that our result and its proof can be understood without any familiarity with quantum statistics. Of course, to appreciate the applications of the result, some further appreciation of "what is a quantum statistical model" is needed. The paper contains a brief summary of this; for more information the reader is referred to the literature[3,5]. For an overview of the "state of the art" in quantum asymptotic statistics see Hayashi's paper[11] which reprints papers of many authors together with introductions by the editor.

Let us develop enough notation to state the main result of the paper and compare it with the comparable result from classical statistics. Starting on familar ground with the latter, suppose we want to estimate a function $\psi(\theta)$ of a parameter θ, both represented by real column vectors of possibly different dimension, based on N i.i.d. observations from a distribution with Fisher information matrix $I(\theta)$. Let π be a prior density on the parameter space and let $\widetilde{G}(\theta)$ be a symmetric positive-definite matrix defining a quadratic loss function $l(\widehat{\psi}^{(N)}, \theta) = (\widehat{\psi}^{(N)} - \psi(\theta))^\top \widetilde{G}(\theta)(\widehat{\psi}^{(N)} - \psi(\theta))$. (Later we will use $G(\theta)$, without the tilde, in the special case when ψ is θ itself). Define the mean square error matrix $V^{(N)}(\theta) = \mathbb{E}_\theta(\widehat{\psi}^{(N)} - \psi(\theta))(\widehat{\psi}^{(N)} - \psi(\theta))^\top$ so that the risk can be written $R^{(N)}(\theta) = \operatorname{trace} \widetilde{G}(\theta) V^{(N)}(\theta)$. The Bayes risk is $R^{(N)}(\pi) = \mathbb{E}_\pi \operatorname{trace} \widetilde{G} V^{(N)}$. Here, $\mathbb{E}_\theta$ denotes expectation over the data for given θ, $\mathbb{E}_\pi$ denotes averaging over θ with respect to the prior π. The estimator $\widehat{\psi}^{(N)}$ is completely arbitrary. We assume the prior density to be smooth, compactly supported and zero on the smooth boundary of its support. Furthermore a certain quantity roughly interpreted as "information in the prior" must be finite. Then it is very easy to show[6], using the van Trees inequality, that under

minimal smoothness conditions on the statistical model,

$$\liminf_{N\to\infty} NR^{(N)}(\pi) \;\geq\; \mathbb{E}_\pi \operatorname{trace} GI^{-1} \tag{1}$$

where $G = \psi' \widetilde{G} \psi'^{\top}$ and ψ' is the matrix of partial derivatives of elements of ψ with respect to those of θ.

Now in quantum statistics the data depends on the choice of measurement and the measurement should be tuned to the loss function. Given a measurement $M^{(N)}$ on N copies of the quantum system, denote by $\overline{I}_M^{(N)}$ the average Fisher information (i.e., Fisher information divided by N) in the data. Holevo's[13] quantum Cramér-Rao bound, as extended by Hayashi and Matsumoto[12] to the quantum i.i.d. model, can be expressed as saying that, for all θ, G, N and $M^{(N)}$,

$$\operatorname{trace} G(\theta)(\overline{I}_M^{(N)}(\theta))^{-1} \;\geq\; \mathcal{C}_G(\theta) \tag{2}$$

for a certain quantity $\mathcal{C}_G(\theta)$, which depends on the specification of the quantum statistical model (state of one copy, derivatives of the state with respect to parameters, and loss function G) *at the point* θ only, i.e., on local or pointwise model features (see (7) below). According to as yet unpublished work of M. Hayashi the bound is asymptotically sharp. The idea behind his work is that locally, the quantum i.i.d. model is well approximated by a quantum Gaussian location model, a quantum statistical problem for which the Holevo bound is sharp[13].

We aim to prove that under minimal smoothness conditions on the quantum statistical model, and conditions on the prior similar to those needed in the classical case, but under essentially no conditions on the estimator-and-measurement sequence,

$$\liminf_{N\to\infty} NR^{(N)}(\pi) \;\geq\; \mathbb{E}_\pi \mathcal{C}_G \tag{3}$$

where, as before, $G = \psi' \widetilde{G} \psi'^{\top}$. The main result (3) is exactly the bound one would hope for, from heuristic statistical principles, and one may also expect it to be sharp, for the reasons mentioned above. In specific models of interest, the right hand side is often easy to calculate. Various specific measurement-and-estimator sequences, motivated by a variety of approaches, can also be shown in interesting examples to achieve the bound. The restrictions on the prior can often be relaxed by approximating the prior of interest, as we will show in our examples.

It was also shown by Gill and Levit[6], how—in the classical statistical context—one can replace a fixed prior π by a sequence of priors indexed by N, concentrating more and more on a fixed parameter value θ_0, at rate

$1/\sqrt{N}$. Following their approach would, in the quantum context, lead to the pointwise asymptotic lower bounds

$$\liminf_{N\to\infty} NR^{(N)}(\theta) \;\geq\; \mathcal{C}_G(\theta) \tag{4}$$

for each θ, for *regular* estimators, and to local asymptotic minimax bounds

$$\lim_{M\to\infty} \liminf_{N\to\infty} \sup_{\|\theta-\theta_0\|\leq N^{-1/2}M} NR^{(N)}(\theta) \;\geq\; \mathcal{C}_G(\theta_0) \tag{5}$$

for *all* estimators, but we do not further develop that theory here. In classical statistics the theory of Local Asymptotic Normality is the way to unify, generalise, and understand this kind of result. We do not yet have a theory of "Q-LAN" though there are indications that it may be possible to build such a theory. The results we obtain here using more elementary tools do give further support to the distant aim of building a Q-LAN theory.

The basic tools used in this paper have now all been mentioned, but as we shall see, the proof is not a routine application of the van Trees inequality. The missing ingredient will be provided by the following new *dual* bound to (2): for all θ, K, N and $M^{(N)}$,

$$\operatorname{trace} K(\theta)\overline{I}_M^{(N)}(\theta) \;\leq\; \mathcal{C}^K(\theta) \tag{6}$$

where $\mathcal{C}^K(\theta)$ actually equals $\mathcal{C}_G(\theta)$ for a certain G defined in terms of K (as explained in Theorem 3.2 below). This is an *upper* bound on Fisher information, in contrast to (2) which is a lower bound on inverse Fisher information. The new inequality (6) follows from the convexity of the sets of information matrices and of inverse information matrices for arbitrary measurements on a quantum system, and these convexity properties have a simple statistical explanation. Such dual bounds have cropped up incidentally in quantum statistics, for instance in Gill and Massar[7], but this is the first time a connection is established.

The argument for (6), and given that, for (3), is based on some general structural features of quantum statistics, and hence it is not necessary to be familiar with the technical details of the set-up. In the next section we will summarize the i.i.d. model in quantum statistics, focussing on the key facts which will be used in the proof of the dual Holevo bound (6) and of our main result, the asymptotic lower bound (3). These proofs are given in a subsequent section, where no further "quantum" arguments will be used. In a final section we will give three applications, leading to new results on some much studied quantum statistical estimation problems.

2. Quantum statistics: the i.i.d. parametric case.

The basic objects in quantum statistics are *states* and *measurements*, defined in terms of certain operators on a complex Hilbert space. To avoid technical complications we restrict attention to the finite-dimensional case, already rich in structure and applications, when operators are represented by ordinary (complex) matrices.

States and measurement The state of a d-dimensional system is represented by a $d \times d$ matrix ρ, called the *density matrix* of the state, having the following properties: $\rho^* = \rho$ (self-adjoint or Hermitian), $\rho \geq \mathbf{0}$ (non-negative), $\mathrm{trace}(\rho) = 1$ (normalized). "Non-negative" actually implies "self-adjoint" but it does no harm to emphasize both properties. $\mathbf{0}$ denotes the zero matrix; $\mathbf{1}$ will denote the identity matrix.

Example: when $d = 2$, every density matrix can be written in the form $\rho = \frac{1}{2}(\mathbf{1} + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3)$ where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the three Pauli matrices and where $\theta_1^2 + \theta_2^2 + \theta_3^2 \leq 1$. □

"Quantum statistics" concerns the situation when the state of the system $\rho(\theta)$ depends on a (column) vector θ of p unknown (real) parameters.

Example: a completely unknown two-dimensional quantum state depends on a vector of three real parameters, $\theta = (\theta_1, \theta_2, \theta_3)^\top$, known to lie in the unit ball. Various interesting submodels can be described geometrically: e.g., the equatorial plane; the surface of the ball; a straight line through the origin. More generally, a completely unknown d-dimensional state depends on $p = d^2 - 1$ real parameters. □

Example: in the previous example the two-parameter case obtained by demanding that $\theta_1^2 + \theta_2^2 + \theta_3^2 = 1$ is called the case of a two-dimensional pure state. In general, a state is called pure if $\rho^2 = \rho$ or equivalently ρ has rank one. A completely unknown pure d-dimensional state depends on $p = 2(d - 1)$ real parameters. □

A measurement on a quantum system is characterized by the outcome space, which is just a measurable space $(\mathcal{X}, \mathcal{B})$, and a *positive operator valued measure* (POVM) M on this space. This means that for each $B \in \mathcal{B}$ there corresponds a $d \times d$ non-negative self-adjoint matrix $M(B)$, together having the usual properties of an ordinary (real) measure (sigma-additive),

with moreover $M(\mathcal{X}) = \mathbf{1}$. The probability distribution of the outcome of doing measurement M on state $\rho(\theta)$ is given by the Born law, or trace rule: $\Pr(\text{outcome} \in B) = \text{trace}(\rho(\theta)M(B))$. It can be seen that this is indeed a bona-fide probability distribution on the sample space $(\mathcal{X}, \mathcal{B})$. Moreover it has a density with respect to the finite real measure $\text{trace}(M(B))$.

Example: the most simple measurement is defined by choosing an orthonormal basis of $\mathbb{C}^d$, say $\psi_1,\ldots,\psi_d$, taking the outcome space to be the discrete space $\mathcal{X} = \{1,\ldots,d\}$, and defining $M(\{x\}) = \psi_x\psi_x^*$ for $x \in \mathcal{X}$; or in physicists' notation, $M(\{x\}) = |\psi_x\rangle\langle\psi_x|$. One computes that $\Pr(\text{outcome} = x) = \psi_x^*\rho(\theta)\psi_x = \langle\psi_x|\rho|\psi_x\rangle$. If the state is pure then $\rho = \phi\phi^* = |\phi\rangle\langle\phi|$ for some $\phi = \phi(\theta) \in \mathbb{C}^d$ of length 1 and depending on the parameter θ. One finds that $\Pr(\text{outcome} = x) = |\psi_x^*\phi|^2 = |\langle\psi_x|\phi\rangle|^2$. □

So far we have discussed state and measurement for a single quantum system. This encompasses also the case of N copies of the system, via a tensor product construction, which we will now summarize. The joint state of N identical copies of a single system having state $\rho(\theta)$ is $\rho(\theta)^{\otimes N}$, a density matrix on a space of dimension d^N. A joint or collective measurement on these systems is specified by a POVM on this large tensor product Hilbert space. An important point is that joint measurements give many more possibilities than measuring the separate systems independently, or even measuring the separate systems adaptively.

Fact 2.1. State plus measurement determines probability distribution of data.

Quantum Cramér-Rao bound. Our main input is going to be Holevo's quantum Cramér-Rao bound[13], with its extension to the i.i.d. case due to Hayashi and Matsumoto[12].

Precisely because of quantum phenomena, different measurements, incompatible with one another, are appropriate when we are interested in different components of our parameter, or more generally, in different loss functions. The bound concerns estimation of θ itself rather than a function thereof, and depends on a quadratic loss function defined by a symmetric real non-negative matrix $G(\theta)$ which may depend on the actual parameter value θ. For a given estimator $\widehat{\theta}^{(N)}$ computed from the outcome of some measurement $M^{(N)}$ on N copies of our system, define its mean square error matrix $V^{(N)}(\theta) = \mathbb{E}_\theta(\widehat{\theta}^{(N)} - \theta)(\widehat{\theta}^{(N)} - \theta)^\top$. The risk function when using the quadratic loss determined by G is $R^{(N)}(\theta) = \mathbb{E}_\theta(\widehat{\theta}^{(N)} - \theta)^\top G(\theta)(\widehat{\theta}^{(N)} - \theta) = \text{trace}(G(\theta)V^{(N)}(\theta))$.

One may expect the risk of good measurements-and-estimators to decrease like N^{-1} as $N \to \infty$. The quantum Cramér-Rao bound confirms that this is the best rate to hope for: it states that for unbiased estimators of a p-dimensional parameter θ, based on arbitrary joint measurements on N copies,

$$NR^{(N)}(\theta) \;\geq\; \mathcal{C}_G(\theta) \;=\; \inf_{\vec{X}, V : V \geq Z(\vec{X})} \operatorname{trace}(G(\theta)V) \tag{7}$$

where $\vec{X} = (X_1, \ldots, X_p)$, the X_i are $d \times d$ self-adjoint matrices satisfying $\partial/\partial\theta_i \operatorname{trace}(\rho(\theta)X_j) = \delta_{ij}$; Z is the $p \times p$ self-adjoint matrix with elements $\operatorname{trace}(\rho(\theta)X_iX_j)$; and V is a real symmetric matrix. It is possible to solve the optimization over V for given $\vec{X}$ leading to the formula

$$\mathcal{C}_G(\theta) \;=\; \inf_{\vec{X}} \operatorname{trace}\bigl(\Re(G^{1/2}Z(\vec{X})G^{1/2}) + \operatorname{abs}\Im(G^{1/2}Z(\vec{X})G^{1/2})\bigr) \tag{8}$$

where $G = G(\theta)$. The absolute value of a matrix is found by diagonalising it and taking absolute values of the eigenvalues. We'll assume that the bound is finite, i.e., there exists $\vec{X}$ satisfying the constraints. A sufficient condition for this is that the Helstrom quantum information matrix H introduced in (17) below is nonsingular.

For specific interesting models, it often turns out not difficult to compute the bound $\mathcal{C}_G(\theta)$. Note, it is a bound which depends only on the density matrix of one system ($N = 1$) and its derivative with the respect to the parameter, and on the loss function, both at the given point θ. It can be found by solving a finite-dimensional optimization problem.

We will not be concerned with the specific form of the bound. What we are going to need, are just two key properties.

Firstly: the bound is local, and applies to the larger class of *locally unbiased estimators*. This means to say that *at the given point* θ, $\mathbb{E}_\theta \widehat{\theta}^{(N)} = \theta$, and at this point also $\partial/\partial\theta_i \, \mathbb{E}_\theta \widehat{\theta}_j^{(N)} = \delta_{ij}$. Now, it is well known that the "estimator" $\theta_0 + I(\theta_0)^{-1}S(\theta_0)$, where $I(\theta)$ is Fisher information and $S(\theta)$ is score function, is locally unbiased at $\theta = \theta_0$ *and achieves the Cramér-Rao bound there*. Thus the Cramér-Rao bound for *locally* unbiased estimators is sharp. Consequently, we can rewrite the bound (7) in the form (2) announced above, where $\overline{I}_M^{(N)}(\theta)$ is the *average* (divided by N) Fisher information in the outcome of an arbitrary measurement $M = M^{(N)}$ on N copies and the right hand side is defined in (7) or (8).

Fact 2.2. We have a family of computable lower bounds on the inverse average Fisher information matrix for an arbitrary measurement on N copies, given by (2) and (7) or (8),

Secondly, for given θ, define the following two sets of positive-definite symmetric real matrices, in one-to-one correspondence with one another through the mapping "matrix inverse". The matrices G occurring in the definition are also taken to be positive-definite symmetric real.

$$\mathcal{V} = \{V : \text{trace}(GV) \geq \mathcal{C}_G \ \forall \ G\}, \tag{9}$$

$$\mathcal{I} = \{I : \text{trace}(GI^{-1}) \geq \mathcal{C}_G \ \forall \ G\}. \tag{10}$$

In the appendix to this paper, we give an algebraic proof that that the set $\mathcal{I}$ is convex (for $\mathcal{V}$, convexity is obvious), and that the inequalities defining $\mathcal{V}$ define supporting hyperplanes to that convex set, i.e., all the inequalities are achievable in $\mathcal{V}$, or equivalently $\mathcal{C}_G = \inf_{V \in \mathcal{V}} \text{trace}(GV)$.

In fact, these properties have a statistical explanation, connected to the fact that the quantum statistical problem of collective measurements on N identical quantum systems approaches a quantum Gaussian problem as $N \to \infty$, see Guţă and Kahn[8]. It can be shown[10] (see also Hayashi, personal communication; Guţă, 2005, unpublished manuscript) that $\mathcal{V}$ *consists of all covariance matrices of locally unbiased estimators achievable (by suitable choice of measurement) on a certain p-parameter quantum Gaussian statistical model. The inequalities defining* $\mathcal{V}$ *are the Holevo bounds for that model, and each of those bounds is attainable.* Thus, for each G, there exists a $V \in \mathcal{V}$ achieving equality in $\text{trace}(GV) \geq \mathcal{C}_G$. It follows from this that $\mathcal{I}$ consists of all non-singular information matrices together with any non-singular matrix smaller than some information matrix, achievable by choice of measurement on the same quantum Gaussian model. Consider the set of information matrices attainable by some measurement together with all smaller matrices; and consider the set of variance matrices of locally unbiased estimators based on arbitrary measurements. Note that adding zero mean noise to a locally unbiased estimator preserves its local unbiasedness, so adding larger matrices to this set does not change it. The set of information matrices is convex: choosing measurement 1 with probability p and measurement 2 with probability q (and remembering your choice) gives a measurement whose Fisher information is the convex combination of the informations of measurements 1 and 2. Augmenting the set with all matrices smaller than something in the set, preserves convexity. (The set of variances of locally unbiased estimators is convex, by a similar randomization argument). Putting this together, we obtain

Fact 2.3. For given θ, both $\mathcal{V}$ and $\mathcal{I}$ defined in (9) and (10) are convex, and all the inequalities defining these sets are achieved by points in the sets.

See the appendix for a direct algebraic proof.

3. An asymptotic Bayesian information bound

We will now introduce the van Trees inequality, a Bayesian Cramér-Rao bound, and combine it with the Holevo bound (2) via derivation of a dual bound following from the convexity of the sets (7) and (8). We return to the problem of estimating the (real, column) vector function $\psi(\theta)$ of the (real, column) vector parameter θ of a state $\rho(\theta)$ based on collective measurements of N identical copies. The dimensions of ψ and of θ need not be the same. The sample size N is largely suppressed from the notation. Let V be the mean square error matrix of an arbitrary estimator $\widehat{\psi}$, thus $V(\theta) = \mathbb{E}_\theta(\widehat{\psi} - \psi(\theta))(\widehat{\psi} - \psi(\theta))^\top$. Often, but not necessarily, we'll have $\widehat{\psi} = \psi(\widehat{\theta})$ for some estimator of θ. Suppose we have a quadratic loss function $(\widehat{\psi} - \psi(\theta))^\top \widetilde{G}(\theta)(\widehat{\psi} - \psi(\theta))$ where $\widetilde{G}$ is a positive-definite matrix function of θ, then the Bayes risk with respect to a given prior π can be written $R(\pi) = \mathbb{E}_\pi \mathrm{trace}\, \widetilde{G}V$. We are going to prove the following theorem:

Theorem 3.1. *Suppose $\rho(\theta) : \theta \in \Theta \subseteq \mathbb{R}^p$ is a smooth quantum statistical model and suppose π is a smooth prior density on a compact subset $\Theta_0 \subseteq \Theta$, such that Θ_0 has a piecewise smooth boundary, on which π is zero. Suppose moreover the quantity $\mathcal{J}(\pi)$ defined in (15) below, is finite. Then*

$$\liminf_{N\to\infty} NR^{(N)}(\pi) \;\geq\; \mathbb{E}_\pi \mathcal{C}_{G_0} \tag{11}$$

where $G_0 = \psi' \widetilde{G} \psi'^\top$ (and assumed to be positive-definite), ψ' is the matrix of partial derivatives of elements of ψ with respect to those of θ, and $\mathcal{C}_{G_0}$ is defined by (7) or (8).

"Once continuously differentiable" is enough smoothness. Smoothness of the quantum statistical model implies smoothness of the classical statistical model following from applying an arbitrary measurement to N copies of the quantum state. Slightly weaker but more elaborate smoothness conditions on the statistical model and prior are spelled out by Gill and Levit[6]. The restriction that G_0 be non-singular can probably be avoided by a more detailed analysis.

Let $\overline{I}_M$ denote the average Fisher information matrix for θ based on a given collective measurement on the N copies. Then the van Trees inequality states that for all matrix functions C of θ, of size $\dim(\psi) \times \dim(\theta)$,

$$N\mathbb{E}_\pi \mathrm{trace}\, \widetilde{G}V \;\geq\; \frac{(\mathbb{E}_\pi \mathrm{trace}\, C\psi'^\top)^2}{\mathbb{E}_\pi \mathrm{trace}\, \widetilde{G}^{-1} C \overline{I}_M C^\top + \frac{1}{N}\mathbb{E}_\pi \frac{(C\pi)'^\top \widetilde{G}^{-1} (C\pi)'}{\pi^2}} \tag{12}$$

where the primes in ψ' and in $(C\pi)'$ both denote differentiation, but in the first case converting the vector ψ into the matrix of partial derivatives of elements of ψ with respect to elements of θ, of size $\dim(\psi) \times \dim(\theta)$, in the second case converting the matrix $C\pi$ into the column vector, of the same length as ψ, with row elements $\sum_j (\partial/\partial\theta_j)(C\pi)_{ij}$. To get an optimal bound we need to choose $C(\theta)$ cleverly.

First though, note that the Fisher information appears in the denominator of the van Trees bound. This is a nuisance since we have a Holevo's *lower* bound (2) to the *inverse* Fisher information. We would like to have an *upper* bound on the information itself, say of the form (6), together with a recipe for computing $\mathcal{C}^K$.

All this can be obtained from the convexity of the sets $\mathcal{I}$ and $\mathcal{V}$ defined in (10) and (9) and the non-redundancy of the inequalities appearing in their definitions. Suppose V_0 is a boundary point of $\mathcal{V}$. Define $I_0 = V_0^{-1}$. Thus I_0 (though not necessarily an attainable average information matrix $\overline{I}_M^{(N)}$) satisfies the Holevo bound for each positive-definite G, and attains equality in one of them, say with $G = G_0$. In the language of convex sets, and "in the V-picture", $\operatorname{trace} G_0 V = \mathcal{C}_{G_0}$ is a supporting hyperplane to $\mathcal{V}$ at $V = V_0$.

Under the mapping "matrix-inverse" the hyperplane $\operatorname{trace} G_0 V = \mathcal{C}_{G_0}$ in the V-picture maps to the smooth surface $\operatorname{trace} G_0 I^{-1} = \mathcal{C}_{G_0}$ touching the set $\mathcal{I}$ at I_0 in the I-picture. Since $\mathcal{I}$ is convex, the tangent plane to the smooth surface at $I = I_0$ must be a supporting hyperplane to $\mathcal{I}$ at this point. The matrix derivative of the operation of matrix inversion can be written $\mathrm{d}A^{-1}/\mathrm{d}x = -A^{-1}(\mathrm{d}A/\mathrm{d}x)A^{-1}$. This tells us that the equation of the tangent plane is $\operatorname{trace} G_0 I_0^{-1} I I_0^{-1} = \operatorname{trace} G_0 I_0^{-1} = \mathcal{C}_{G_0}$. Since this is simultaneously a supporting hyperplane to $\mathcal{I}$ we deduce that for all $I \in \mathcal{I}$, $\operatorname{trace} G_0 I_0^{-1} I I_0^{-1} \leq \mathcal{C}_{G_0}$. Defining $K_0 = I_0^{-1} G_0 I_0^{-1}$ and $\mathcal{C}^{K_0} = \mathcal{C}_{G_0}$ we rewrite this inequality as $\operatorname{trace} K_0 I \leq \mathcal{C}^{K_0}$.

A similar story can be told when we start in the I-picture with a supporting hyperplane (at $I = I_0$) to $\mathcal{I}$ of the form $\operatorname{trace} K_0 I = \mathcal{C}^{K_0}$ for some symmetric positive-definite K_0. It maps to the smooth surface $\operatorname{trace} K_0 V^{-1} = \mathcal{C}^{K_0}$, with tangent plane $\operatorname{trace} K_0 V_0^{-1} I V_0^{-1} = \mathcal{C}^{K_0}$ at $V = V_0 = I_0^{-1}$. By strict convexity of the function "matrix inverse", the tangent plane touches the smooth surface only at the point V_0. Moreover, the smooth surface lies above the tangent plane, but below $\mathcal{V}$. This makes V_0 the unique minimizer of $\operatorname{trace} K_0 V_0^{-1} I V_0^{-1}$ in $\mathcal{V}$.

It would be useful to extend these computations to allow singular I, G and K. Anyway, we summarize what we have so far in a theorem.

Theorem 3.2. *Dual to the Holevo family of lower bounds on average inverse information,* $\operatorname{trace} G\overline{I}_M^{-1} \geq \mathcal{C}_G$ *for each positive-definite* G*, we have a family of upper bounds on information,*

$$\operatorname{trace} K\overline{I}_M \leq \mathcal{C}^K \quad \text{for each} \quad K. \tag{13}$$

If $I_0 \in \mathcal{I}$ *satisfies* $\operatorname{trace} G_0 I_0^{-1} = \mathcal{C}_{G_0}$ *then with* $K_0 = I_0^{-1} G_0 I_0^{-1}$, $\mathcal{C}^{K_0} = \mathcal{C}_{G_0}$*. Conversely if* $I_0 \in \mathcal{I}$ *satisfies* $\operatorname{trace} K_0 I_0 = \mathcal{C}^{K_0}$ *then with* $G_0 = I_0 K_0 I_0$, $\mathcal{C}_{G_0} = \mathcal{C}^{K_0}$*. Moreover, none of the bounds is redundant, in the sense that for all positive-definite* G *and* K, $\mathcal{C}_G = \inf_{V \in \mathcal{V}} \operatorname{trace}(GV)$ *and* $\mathcal{C}^K = \sup_{I \in \mathcal{I}} \operatorname{trace}(KI)$*. The minimizer in the first equation is unique.*

Now we are ready to apply the van Trees inequality. First we make a guess for what the left hand side of (12) should look like, at its best. Suppose we use an estimator $\widehat{\psi} = \psi(\widehat{\theta})$ where $\widehat{\theta}$ makes optimal use of the information in the measurement M. Denote now by I_M the asymptotic normalized Fisher information of a sequence of measurements. Then we expect that the asymptotic normalized covariance matrix V of $\widehat{\psi}$ is equal to $\psi' I_M^{-1} \psi'^{\top}$ and therefore the asymptotic normalized Bayes risk should be $\mathbb{E}_\pi \operatorname{trace} \widetilde{G} \psi' I_M^{-1} \psi'^{\top} = \mathbb{E}_\pi \operatorname{trace} \psi'^{\top} \widetilde{G} \psi' I_M^{-1}$. This is bounded below by the integrated Holevo bound $\mathbb{E}_\pi \mathcal{C}_{G_0}$ with $G_0 = \psi'^{\top} \widetilde{G} \psi'$. Let $I_0 \in \mathcal{I}$ satisfy $\operatorname{trace} G_0 I_0^{-1} = \mathcal{C}_{G_0}$; its existence and uniqueness are given by Theorem 3.2. (Heuristically we expect that I_0 is asymptotically attainable). By the same Theorem, with $K_0 = I_0^{-1} G_0 I_0^{-1}$, $\mathcal{C}^{K_0} = \mathcal{C}_{G_0} = \operatorname{trace} G_0 I_0^{-1} = \operatorname{trace} \psi'^{\top} \widetilde{G} \psi' I_0^{-1}$.

Though these calculations are informal, they lead us to try the matrix function $C = \widetilde{G} \psi' I_0^{-1}$. Define $V_0 = I_0^{-1}$. With this choice, in the numerator of the van Trees inequality, we find the square of $\operatorname{trace} C\psi'^{\top} = \operatorname{trace} \widetilde{G} \psi' I_0^{-1} \psi'^{\top} = \operatorname{trace} G_0 V_0 = \mathcal{C}_{G_0}$. In the main term of the denominator, we find $\operatorname{trace} \widetilde{G}^{-1} \widetilde{G} \psi' I_0^{-1} \overline{I}_M I_0^{-1} \psi'^{\top} \widetilde{G} = \operatorname{trace} I_0^{-1} G_0 I_0^{-1} \overline{I}_M = \operatorname{trace} K_0 \overline{I}_M \leq \mathcal{C}^{K_0} = \mathcal{C}_{G_0}$ by the dual Holevo bound (13). This makes the numerator of the van Trees bound equal to the square of this part of the denominator, and using the inequality $a^2/(a+b) \geq a - b$ we find

$$N \mathbb{E}_\pi \operatorname{trace} GV \;\geq\; \mathbb{E}_\pi \mathcal{C}_{G_0} - \frac{1}{N} \mathcal{J}(\pi) \tag{14}$$

where

$$\mathcal{J}(\pi) \;=\; \mathbb{E}_\pi \frac{(C\pi)'^{\top} \widetilde{G}^{-1} (C\pi)'}{\pi^2} \tag{15}$$

with $C = \widetilde{G} \psi' V_0$ and V_0 uniquely achieving in $\mathcal{V}$ the bound $\operatorname{trace} G_0 V \geq \mathcal{C}_{G_0}$, where $G_0 = \psi'^{\top} \widetilde{G} \psi'$. Finally, provided $\mathcal{J}(\pi)$ is finite (which depends

on the prior distribution and on properties of the model), we obtain the asymptotic lower bound

$$\liminf_{N\to\infty} N\mathbb{E}_\pi \mathrm{trace}\,\widetilde{G}V \;\geq\; \mathbb{E}_\pi \mathcal{C}_{G_0}. \tag{16}$$

4. Examples

In the three examples discussed here, the loss function is derived from a very popular (among the physicists) figure-of-merit in state estimation called *fidelity*. Suppose we wish to estimate a state $\rho = \rho(\theta)$ by $\widehat{\rho} = \rho(\widehat{\theta})$. Fidelity measures the closeness of the two states, being maximally equal to 1 when the estimate and truth coincide. It is defined as $\mathrm{Fid}(\widehat{\rho}, \rho) = \left(\mathrm{trace}(\sqrt{\rho^{\frac{1}{2}}\widehat{\rho}\rho^{\frac{1}{2}}})\right)^2$ (some authors would call this *squared* fidelity). When both states are pure, thus $\rho = |\phi\rangle\langle\phi|$ and $\widehat{\rho} = |\widehat{\phi}\rangle\langle\widehat{\phi}|$ where ϕ and $\widehat{\phi}$ are unit vectors in $\mathbb{C}^d$, then $\mathrm{Fid}(\widehat{\phi}, \phi) = |\langle\widehat{\phi}|\phi\rangle|^2$. There is an important characterization of fidelity due to Fuchs[4] which both explains its meaning and leads to many important properties. Suppose M is a measurement on the quantum system. Denote by $M(\rho)$ the probability distribution of the outcome of the measurement M when applied to a state ρ. For two probability distributions P, $\widehat{P}$ on the same sample space, let p and $\widehat{p}$ be their densities with respect to a dominating measure μ and define the fidelity between these probability measures as $\mathrm{Fid}(\widehat{P}, P) = \left(\int \widehat{p}^{\frac{1}{2}} p^{\frac{1}{2}} \mathrm{d}\mu\right)^2$. In usual statistical language, this is the *squared Hellinger affinity* between the two probability measures. It turns out that $\mathrm{Fid}(\widehat{\rho}, \rho) = \inf_M \mathrm{Fid}(M(\widehat{\rho}), M(\rho))$, thus two states have small fidelity when there is a measurement which distinguishes them well, in the sense that the Hellinger affinity between the outcome distributions is small, or in other words, the L_2 distance between the root densities of the data under the two models is large.

Now suppose states are smoothly parametrized by a vector parameter θ. Consider the fidelity between two states with close-by parameter values θ and $\widehat{\theta}$, and suppose they are measured with the same measurement M. From the relation $\int p^{\frac{1}{2}}\widehat{p}^{\frac{1}{2}}\mathrm{d}\mu = 1 - \frac{1}{2}\|\widehat{p}^{\frac{1}{2}} - p^{\frac{1}{2}}\|^2$ and by a Taylor expansion to second order one finds $1 - \mathrm{Fid}(\widehat{P}, P) \approx \frac{1}{4}(\widehat{\theta} - \theta)^\top I_M(\theta)(\widehat{\theta} - \theta)$ where $I_M(\theta)$ is the Fisher information in the outcome of the measurement M on the state $\rho(\theta)$. We will define the *Helstrom* quantum information matrix $H(\theta)$ by the analogous relation

$$1 - \mathrm{Fid}(\widehat{\rho}, \rho) \approx \frac{1}{4}(\widehat{\theta} - \theta)^\top H(\theta)(\widehat{\theta} - \theta). \tag{17}$$

It turns out that $H(\theta)$ is the smallest "information matrix" such that $I_M(\theta) \leq H(\theta)$ for all measurements M.

Taking as loss function $l(\widehat{\theta},\theta) = 1 - \mathrm{Fid}(\rho(\widehat{\theta}),\rho(\theta))$ we would expect (by a quadratic approximation to the loss) that $\mathbb{E}_\pi \mathcal{C}_{\frac{1}{4}H}$ is a sharp asymptotic lower bound on N times the Bayes risk. We will prove this result for a number of special cases, in which by a fortuitous circumstance, the fidelity-loss function is *exactly* quadratic in a (sometimes rather strange) function of the parameter. The first two examples concern a two-dimensional quantum system and are treated in depth by Bagan and co-workers[1]; below we just outline some important features of the application. In the second of those two examples our asymptotic lower bound is an essential part of a proof of asymptotic optimality of a certain measurement-and-estimation scheme.

The third example concerns an unknown pure state of arbitrary dimension. Here we are present a short and geometric proof of a surprising but little known result of[9] which shows that an extraordinarily simple measurement scheme leads to an asymptotically optimal estimator (providing the data is processed efficiently). The analysis also links the previously unconnected Holevo and Gill-Massar bounds[7,13].

4.1. *Completely unknown spin half (d=2, p=3)*

Recall that a completely unknown 2-dimensional quantum state can be written $\rho(\theta) = \frac{1}{2}(\mathbf{1} + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3)$, where θ lies in the unit ball in $\mathbb{R}^3$. It turns out that $\mathrm{Fid}(\widehat{\rho},\rho) = \frac{1}{2}(1 + \widehat{\theta}\cdot\theta + (1-\|\widehat{\theta}\|^2)^{\frac{1}{2}}(1-\|\theta\|^2)^{\frac{1}{2}})$. Define $\psi(\theta)$ to be the four-dimensional vector obtained by adjoining $(1-\|\theta\|^2)^{\frac{1}{2}}$ to θ_1, θ_2, θ_3. Note that this vector has constant length 1. It follows that $1 - \mathrm{Fid}(\widehat{\rho},\rho) = \frac{1}{4}\|\widehat{\psi} - \psi\|^2$. This is a quadratic loss-function for estimation of $\psi(\theta)$ with $\widetilde{G} = \mathbf{1}$, the 4×4 identity matrix. By Taylor expansion of both sides, we find that $\frac{1}{4}H = \psi'^{\top}\widetilde{G}\psi' = G$ and conclude from Theorem 1 that N times $1-$ mean fidelity is indeed asymptotically lower bounded by $\mathbb{E}_\pi \mathcal{C}_{\frac{1}{4}H}$.

In Bagan and coworker's paper[1] the exactly optimal measurement-and-estimation scheme is derived and analysed in the case of a rotationally invariant prior distribution over the unit ball. The optimal *measurement* turns out not to depend on the (arbitrary) radial part of the prior distribution, and separates into two parts, one used for estimating the direction $\theta/\|\theta\|$, the other part for estimating the length $\|\theta\|$. The Bayes optimal estimator of the length of θ naturally depends on the prior. Because of these simplifications it is feasible to compute the asymptotic value of N times the (optimal) Bayes mean fidelity, and this value is $(3 + 2\mathbb{E}_\pi\|\theta\|)/4$.

The Helstrom quantum information matrix H and the Holevo lower

bound $\mathcal{C}_{\frac{1}{4}H}$ are also computed. It turns out that $\mathcal{C}_{\frac{1}{4}H}(\theta) = (3 + 2\|\theta\|)/4$. Our asymptotic lower bound is not only correct but also, as expected, sharp.

The van Trees approach does put some non-trivial conditions on the prior density π. The most restrictive conditions are that the density is zero at the boundary of its support and that the quantity (15) be finite. Within the unit ball everything is smooth, but there are some singularities at the boundary of the ball. So our main theorem does not apply directly to many priors of interest. However there is an easy approximation argument to extend its scope, as follows.

Suppose we start with a prior π supported by the whole unit ball which does not satisfy the conditions. For any $\epsilon > 0$ construct $\widetilde{\pi} = \widetilde{\pi}_\epsilon$ which is smaller than $(1 + \epsilon)\pi$ everywhere, and 0 for $\|\theta\| \geq 1 - \delta$ for some $\delta > 0$. If the original prior π is smooth enough we can arrange that $\widetilde{\pi}$ satisfies the conditions of the van Trees inequality, and makes (15) finite. N times the Bayes risk for $\widetilde{\pi}$ cannot exceed $1 + \epsilon$ times that for π, and the same must also be true for their limits. Finally, $\mathbb{E}_{\widetilde{\pi}_\epsilon} \mathcal{C}_{\frac{1}{4}H} \to \mathrm{E}_\pi \mathcal{C}_{\frac{1}{4}H}$ as $\epsilon \to 0$.

Some last remarks on this example: first of all, it is known that *only* collective measurements can asymptotically achieve this bound. Separate measurements on separate systems lead to strictly worse estimators. In fact, by the same methods one can obtain the sharp asymptotic lower bound $9/4$ (independent of the prior)[2], when one allows the measurement on the nth system to depend on the data obtained from the earlier ones. Instead of the Holevo bound itself, we use here the bound Gill–Massar bound[7], which is actually has the form of a dual Holevo bound. (We give some more remarks on this at the end of the discussion of the third example). Secondly, our result gives strong heuristic support to the claim[1] that the measurement-and-estimation scheme for a specific prior and specific loss function is also pointwise optimal in a minimax sense, or among regular estimators, for loss functions which are locally equivalent to fidelity-loss; and also asymptotically optimal in the Bayes sense for other priors and locally equivalent loss functions. In general, if the physicists' approach is successful in the sense of generating a measurement-and-estimation scheme which can be analytically studied and experimentally implemented, then this scheme will have (for large N) good properties independent of the prior and only dependent on local properties of the loss.

4.2. *Spin half: equatorial plane (d=2, p=2)*

Bagan et al.[1] also considered the case where it is known that $\theta_3 = 0$, thus we now have a two-dimensional parameter. The prior is again taken to be

rotationally symmetric. The exactly Bayes optimal measurement turns out (at least, for some N and for some priors) to depend on the radial part of the prior. Analysis of the exactly optimal measurement-and-estimation procedure is not feasible since we do not know if this phenomenon persists for all N. However there is a natural measurement, which is exactly optimal for some N and some priors, which one might conjecture to be asymptotically optimal for all priors. This sub-optimal measurement, combined with the Bayes optimal estimator given the measurement, can be analysed and it turns out that N times $1-$ mean fidelity converges to $1/2$ as $N \to \infty$, independently of the prior. Again, the Helstrom quantum information matrix H and the Holevo lower bound $\mathcal{C}_{\frac{1}{4}H}$ are computed. It turns out that $\mathcal{C}_{\frac{1}{4}H}(\theta) = 1/2$. This time we can use our asymptotic lower bound to prove that the natural sub-optimal measurement-and-estimator is in fact asymptotically optimal for this problem.

For a p-parameter model the best one could every hope for is that for large N there are measurements with $\overline{I}_M$ approaching the Helstrom upper bound H. Using this bound in the van Trees inequality gives the asymptotic lower bound on N times $1-$ mean fidelity of $p/4$. The example here is a special case where this is attainable. Such a model is called *quasi-classical*.

If one restricts attention to separate measurements on separate systems the sharp asymptotic lower bound is 1, twice as large[2].

4.3. *Completely unknown d dimensional pure state*

In this example we make use of the dual Holevo bound and symmetry arguments to show that in this example, the original Holevo bound for a natural choice of G (corresponding to fidelity-loss) is attained by an extremely large class of measurements, including one of the most basic measurements around, known as "standard tomography".

For a pure state $\rho = |\phi\rangle\langle\phi|$, fidelity can be written $|\langle\widehat{\phi}|\phi\rangle|^2$ where $|\phi\rangle \in \mathbb{C}^d$ is a vector of unit length. The state-vector can be multiplied by e^{ia} for an arbitrary real phase a without changing the density matrix. The constraint of unit length and the arbitrariness of the phase means that one can parametrize the density matrix ρ corresponding to $|\phi\rangle$ by $2(d-1)$ real parameters which we take to be our underlying vector parameter θ (we have d real parts and d imaginary parts of the elements of $|\phi\rangle$, but one constraint and one parameter which can be fixed arbitrarily).

For a pure state, $\rho^2 = \rho$ so $\mathrm{trace}(\rho^2) = 1$. Another way to write the fidelity in this case is as $\mathrm{trace}(\widehat{\rho}\rho) = \sum_{ij}(\Re(\widehat{\rho}_{ij})\Re(\rho_{ij}) + \Im(\widehat{\rho}_{ij})\Im(\rho_{ij}))$. So if we take $\psi(\theta)$ to be the vector of length $2d^2$ and of length 1 containing the

real and the imaginary parts of elements of ρ we see that $1 - \mathrm{Fid}(\widehat{\rho}, \rho) = \frac{1}{2}\|\widehat{\psi} - \psi\|^2$. It follows that $1-$ fidelity is a quadratic loss function in $\psi(\theta)$ with again $\widetilde{G} = \mathbf{1}$.

Define again the Helstrom quantum information matrix $H(\theta)$ for θ by $1 - \mathrm{Fid}(\widehat{\rho}, \rho) \approx \frac{1}{4}(\widehat{\theta} - \theta)^\top I_M(\theta)(\widehat{\theta} - \theta)$. Just as in the previous two examples we expect the asymptotic lower bound $\mathbb{E}_\pi \mathcal{C}_{\frac{1}{4}H}$ to hold for N times Bayes mean fidelity-loss, where $G = \frac{1}{4}H = \psi'^\top \widetilde{G} \psi'$.

Some striking facts are known about estimation of a pure state. First of all, from[14], we know that the Holevo bound is attainable, for all G, already at $N = 1$. Secondly, from[7] we have the following inequality

$$\mathrm{trace} H^{-1} \overline{I}_M \leq d - 1 \tag{18}$$

with *equality* (in the case that the state is completely unknown) for all *exhaustive* measurements $M^{(N)}$ on N copies of the state. Exhaustivity means, for a measurement with discrete outcome space, that $M^{(N)}(\{x\})$ is a rank one matrix for each outcome x. The meaning of exhaustivity in general is by the same property for the density $m(x)$ of the matrix-valued measure $M^{(N)}$ with respect to a real dominating measure, e.g., $\mathrm{trace}(M^{(N)}(\cdot))$. This tells us that (18) is one of the "dual Holevo inequalities". We can associate it with an original Holevo inequality once we know an information matrix of a measurement attaining the bound. We will show that there is an information matrix of the form $\overline{I}_M = cH$ attaining the bound. Since the number of parameters (and dimension of H) is $2(d-1)$ it follows by imposing equality in (18) that $c = \frac{1}{2}$. The corresponding Holevo inequality must be $\mathrm{trace} \frac{1}{2} H H^{-1} \frac{1}{2} H \overline{I}_M^{-1} \geq d - 1$ which tells us that $\mathcal{C}_{\frac{1}{4}H} = d - 1$.

The proof uses an invariance property of the model. For any unitary matrix U (i.e., $UU^* = U^*U = \mathbf{1}$) we can convert the pure state ρ into a new pure state $U\rho U^*$. The unitary matrices form a group under multiplication. Consequently the group can be thought to act on the parameter θ used to describe the pure state. Clearly the fidelity between two states (or the fidelity between their two parameters) is invariant when the same unitary acts on both states. This group action possesses the "homogenous two point property": for any two pairs of states such that the fidelities between the members of each pair are the same, there is a unitary transforming the first pair into the second pair.

We illustrate this in the case $d = 2$ where (first example, section 2), the pure states can be represented by the surface of the unit ball in $\mathbb{R}^3$. It turns out that the action of the unitaries on the density matrices translates into the action of the group of orthogonal rotations on the unit sphere.

Two points at equal distance on the sphere can be transformed by some rotation into any other two points at the same distance from one another; a constant distance between points on the sphere corresponds to a constant fidelity between the underlying states.

In general, the pure states of dimension d can be identified with the Riemannian manifold CP^{d-1} whose natural Riemannian metric corresponds locally to fidelity (locally, $1-$ fidelity is squared Riemannian distance) and whose isometries correspond to the unitaries. This space posseses the homogenous two point property, as we argued above. It is easy to show that the *only* Riemannian metrics invariant under isometries on such a space are proportional to one another. Hence the quadratic forms generating those metrics with respect to a particular parametrization must also be proportional to one another.

Consider a measurement whose outcome is actually an estimate of the state, and suppose that this measurement is *covariant* under the unitaries. This means that transforming the state by a unitary, doing the measurement on the transformed state, and transforming the estimate back by the inverse of the same unitary, is the same (has the same POVM) as the original measurement. The information matrix for such a measurement is generated from the squared Hellinger affinity between the distributions of the measurement outcomes under two nearby states, just as the Helstrom information matrix is generated from the fidelity between the states. If the measurement is covariant then the Riemannian metric defined by the information matrix of the measurement outcome must be invariant under unitary transformations of the states. Hence: *the information matrix of any covariant measurement is proportional to the Helstrom information matrix.*

Exhaustive covariant measurements certainly do exist. A particularly simple one is that, for each of the N copies of the quantum system, we independently and uniformly choose a basis of $\mathbb{C}^d$ and perform the simple measurement (given in an example in Section 2) corresponding to that basis.

The first conclusion of all this is: any exhaustive covariant measurement has information matrix $\overline{I}_M^{(N)}$ equal to one half the Helstrom information matrix. All such measurements attain the Holevo bound $\mathrm{trace}\frac{1}{4}H(\overline{I}_M^{(N)})^{-1} \geq d-1$. In particular, this holds for the i.i.d. measurement based on repeatedly choosing a uniformly distributed random basis of $\mathbb{C}^d$.

The second conclusion is that an asymptotic lower bound on N times $1-$ mean fidelity is $d-1$. Now the exactly Bayes optimal measurement-and-estimation strategy is known to achieve this bound. The measurement involved is a mathematically elegant collective measurement on the N copies

together, but hard to realise in the laboratory. Our results show that one can expect to asymptotically attain the bound by decent information processing (maximum likelihood? optimal Bayes with uniform prior and fidelity loss?) following an arbitrary *exhaustive covariant measurement*, of which the most simple to implement is the standard tomography measurement consisting of an independent random choice of measurement basis for each separate system.

In Gill and Massar's paper[7] the same bound as (18) was shown to hold for separable (and in particular, for adaptive sequential) measurements also in the mixed state case. Moreover in the case $d = 2$, any information matrix satisfying the bound is attainable already at $N = 1$. This is used in[2] to obtain sharp asymptotic bounds to mean fidelity for separable measurements on mixed qubits.

Acknowledgments

Thanks to Peter Jupp for advice on Riemannian manifolds, Manuel Ballester, Madalin Guţă and Jonas Kahn for their help in understanding the Holevo bound, and to Emili Bagan, Ramon Muñez-Tapia and Alex Monras whose work led to the questions studied here.

I acknowledge financial support from the European Community project RESQ, contract no IST-2001-37559.

Appendix A. Proof of convexity

The first step is to show that

$$\mathcal{V} \;=\; \mathrm{clos}\{V : V \geq Z(\vec{X}) \text{ for some } \vec{X}\} \tag{A.1}$$

where, as before, $\vec{X} = (X_1, \ldots, X_p)$, the X_i are $d \times d$ self-adjoint matrices satisfying $\partial/\partial\theta_i \,\mathrm{trace}(\rho(\theta)X_j) = \delta_{ij}$; Z is the $p \times p$ self-adjoint matrix with elements $\mathrm{trace}(\rho(\theta)X_iX_j)$; and V is a real symmetric matrix.

An easy computation shows that $Z(p\vec{X} + (1-p)\vec{Y}) \leq pZ(\vec{X}) + (1-p)Z(\vec{Y})$ (check that the second derivative w.r.t. p of $\langle\psi|Z(p\vec{X}+(1-p)\vec{Y})|\psi\rangle$ is non-negative, for any complex vector ψ.) This makes $\{V : V \geq Z(\vec{X}) \text{ for some } \vec{X}\}$, where V is self-adjoint, a convex set. Restricting to the real matrices in this set preserves convexity, as does taking the closure of the set. By convexity, the definition (7) tells us that the equations $\mathrm{trace}(GV) = \mathcal{C}_G$ define supporting hyperplanes to the set defined on the right hand side of (A.1). Since a closed convex set is the intersection of the closed halfspaces defined by its supporting hyperplanes, it follows that $\mathcal{V}$

as defined by (9) can also be specified as (A.1), and that all the Holevo bounds $\text{trace}(GV) \geq \mathcal{C}_G$ are attained in $\mathcal{V}$.

The convexity of $\mathcal{I}$, the set of inverses of elements of $\mathcal{V}$, is a lot more subtle. In the following argument I will suppose that the state $\rho(\theta)$ is strictly positive. The proof is easily adapted to the case of a model for a pure state. (More generally we need the notion of D-invariant model and the $\mathcal{L}^2$ spaces defined by a quantum state[12,13]).

We can consider our model with p parameters and a strictly positive density matrix as a submodel of the model of a completely unknown mixed state, which has $d^2 - 1$ parameters. Denote the parameter vector of the full model by ϕ. The submodel is parametrized by θ, a subvector of ϕ. I'll use the terminology interest parameter, nuisance parameter for the two subvectors of ϕ corresponding to submodel parameters and auxiliary parameters. Subscripts 1, 2 will be also used when we partition matrices or vectors according to these two parts. By the strict positivity of ρ we are working at a point in the interior of the full model (this is one of the reasons why the argument needs to be adapted for a pure-state model). Since $\text{trace}\,\rho = 1$, the partial derivatives of ρ with respect to the components of θ in submodel and ϕ in fullmodel are traceless (i.e., have trace zero). It is easy to see from this that we may restrict the elements X of $\vec{X}$, entering into the Holevo bounds for the submodel, and elements Y of $\vec{Y}$, entering into the Holevo bounds for the full model, to be such that $\text{trace}\,\rho Y = 0$. Such Y form a $d^2 - 1$ dimensional real Hilbert space $\mathcal{L}_0^2(\rho)$ under the innerproduct $\langle X, Y \rangle_\rho = \Re\,\text{trace}\,\rho XY$.

Let ρ_i' denote the partial derivative of ρ with respect to θ_i at the fixed parameter value under consideration. For the submodel, define the symmetric logarithmic derivatives $\lambda_i \in \mathcal{L}_0^2(\rho)$ by $\langle \lambda_i, X \rangle_\rho = \text{trace}\,\rho_i' X$ for all $X \in \mathcal{L}_0^2(\rho)$. The constraints $\text{trace}\,\rho_i' X_j = \delta_{ij}$ translate into constraints $\langle \lambda_i, X_j \rangle_\rho = \delta_{ij}$ for all $i, j \leq p$. In the full model, I'll use the notation $\vec{\mu}$ for the vector of symmetric logarithmic derivatives, and $\vec{Y}$ for a candidate vector of Y_i, each of length $d^2 - 1$. Of course, $\vec{\lambda}$ is a subvector of $\vec{\mu}$. In the full model, the constraints on $\vec{Y}$ translate into $\langle \mu_i, Y_j \rangle_\rho = \delta_{ij}$ for all $i, j \leq d^2 - 1$. The μ_i form a basis of $\mathcal{L}_0^2(\rho)$ of linearly independent vectors.

Now in the full model, the constraints on the Y_i make them uniquely defined. Thus for the full model, the set $\mathcal{V}_{\text{full}}$ is the set of all $(d^2-1) \times (d^2-1)$ real matrices W exceeding the fixed self-adjoint matrix $Z_{\text{full}} = Z(\vec{Y})$. Unfortunately, Z_{full} is singular. But we may describe $\mathcal{I}_{\text{full}}$ as the closure of the set of all real matrices less than or equal to $(Z_{\text{full}} + \delta \mathbf{1})^{-1}$ for some $\delta > 0$. The convexity of both sets is trivial. This suggests that we try to

deal with the case of a p parameter model by considering it a submodel of the full $d^2 - 1$ parameter model.

The relation between inverse information matrices for full models and submodels is complicated, but that between the information matrices themselves is simple: the information matrix for a submodel is a submatrix of the information matrix of a full model. Thus we might conjecture that for every $I \in \mathcal{I}$, there exists a $W \geq Z_{\text{full}}$ such that $I \leq (W^{-1})_{11}$, the subscript "11" indicating the submodel submatrix. However, it could be that we have positive information for the submodel parameters, but zero information for the auxiliary parameters. This would make the corresponding inverse information matrix W^{-1} for the full model undefined. This problem can be solved by approximating singular information matrices by nonsingular ones. We will prove the following theorem:

Theorem A.1. *$V^{-1} \in \mathcal{I}$ if and only if there exist real matrices $W^{(n)} > Z_{\text{full}}$, with $((W^{(n)})^{-1})_{11} = (V^{(n)})^{-1} \to V^{-1}$ as $n \to \infty$.*

In words, $\mathcal{I}$ is the closure of the set of 11 submatrices of real symmetric non-nonsingular matrices less than or equal to $(Z_{\text{full}}+\delta\mathbf{1})^{-1}$ for some $\delta > 0$. Consequently $\mathcal{I}$ is convex.

Proof. The proof will work by frequent reparametrizations of the nuisance part of the full model. By this we mean that ϕ is transformed smoothly and one-to-one into, say, ψ, in such a way that the interest component of ϕ is unaltered. Under such a transformation, the vector of symmetric logarithmic derivatives $\vec{\mu}$ transforms by premultiplication by an invertible matrix C whose 11 block is the identity and whose 12 block is zero, so the 'interest" part of $\vec{\mu}$ is unchanged. (Subject to C being nonsingular, for which it is just necessary that the 22 block is nonsingular, the 21 block of C can be arbitrary). At the same time the vector of operators $\vec{Y}$ transforms by premultiplication by the transposed inverse of C. Consequently, Z_{full} is transformed into $(C^\top)^{-1} Z_{\text{full}} C^{-1}$, $W \geq Z_{\text{full}}$ is transformed the same way, while W^{-1} is transformed into $CW^{-1}C^\top$. We therefore see that the 11 block (i.e., the submatrix corresponding to the submodel) of W^{-1} *remains invariant under reparametrization of the auxiliary or nuisance parameters.*

In the statement of the theorem the choice of parametrization of the auxiliary parameters is arbitrary, and so can be chosen in any convenient way. We take advantage of this possibility immediately, in the proof of the the forwards implication of the theorem.

Suppose $V \geq Z(\vec{X})$ for some $\vec{X}$ satisfying the usual constraints. Augment $\vec{\lambda}$ to a vector $\vec{\mu}$ of $d^2 - 1$ linearly independent elements $\mu_i \in \mathcal{L}_0^2(\rho)$

such that $\langle \mu_i, X_j \rangle_\rho = \delta_{ij}$ for all $i \leq d^2 - 1$, $j \leq p$. (The extra elements can be an arbitrary basis of the orthocomplement of the X_j, it is easy to check that together with the old elements they are linearly independent, hence because of their number, a basis). Next augment $\vec{X}$ to $\vec{Y}$, so that the the orthogonality relation, with X_j replaced by Y_j, also holds for $p < j \leq d^2 - 1$.

For square matrices A, B write $\mathrm{diag}(A, B)$ for the block diagonal matrix with A and B as diagonal blocks corresponding to interest and nuisance parts of the full model. Let $D_\epsilon = \mathrm{diag}(\mathbf{1}, \epsilon\mathbf{1})$, this is the diagonal matrix with 1's on the interest parameter part of the diagonal, ϵ's on the nuisance part.

We have $D_\epsilon Z_{\mathrm{full}} D_\epsilon \to \mathrm{diag}(Z(\vec{X}), \mathbf{0}) \leq \mathrm{diag}(V, \mathbf{0})$ as $\epsilon \to 0$. Therefore, for each $\epsilon > 0$ we can find $\delta = \delta(\epsilon) > 0$ such that $D_\epsilon Z_{\mathrm{full}} D_\epsilon < \mathrm{diag}(V, \mathbf{0}) + \delta\mathbf{1}$ and moreover such that $\delta \to 0$ as $\epsilon \to 0$. Thus for each ϵ, $Z_{\mathrm{full}} < D_\epsilon^{-1}(\mathrm{diag}(V, \mathbf{0}) + \delta\mathbf{1})D_\epsilon^{-1} = W_\epsilon$ where $((W_\epsilon)^{-1})_{11} \to V^{-1}$ as $\epsilon \to 0$.

Choosing a sequence $\epsilon_n \to 0$ as $n \to \infty$ we have found $W^{(n)} > Z_{\mathrm{full}}$ for all n with $((W^{(n)})^{-1})_{11} \to V^{-1}$ as $n \to \infty$. Going back to the original parametrization does not alter $((W^{(n)})^{-1})_{11}$ so the forwards implication of the theorem is proved.

Now for the backwards implication. Suppose I am given $W > Z_{\mathrm{full}}$, $(W^{-1})_{11} = V^{-1}$. Reparametrize the nuisance part of the full model so that $(W^{-1})_{12} = \mathbf{0}$. This does not alter $(W^{-1})_{11}$ but does alter both interest and nuisance parts of $\vec{Y}$. Denote the interest part of the transformed $\vec{Y}$ by $\vec{X}$. The inequality $W > Z_{\mathrm{full}}$ remains true after the transformation, hence $W_{11} > Z(\vec{X})$. Since W is block diagonal, we obtain from this $(W^{-1})_{11} \leq (Z(\vec{X}) + \delta\mathbf{1})^{-1}$ for some $\delta > 0$. Taking the closure completes the proof. □

Bibliography

1. Bagan, E., M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia (2006a). Optimal full estimation of qubit mixed states. *Phys. Rev. A* **73** 032301; `arXiv:quant-ph/0510158` .
2. Bagan, E., M. A. Ballester, R. D. Gill, R. Muñoz-Tapia, and O. Romero-Isart (2006b). Separable-measurement estimation of density matrices and its fidelity gap with collective protocols. *Phys. Rev. Lett.* (in print); `arXiv:quant-ph/0512177`
3. Barndorff-Nielsen, O. E., R. D. Gill, and P. E. Jupp (2003). On quantum statistical inference. *J. Roy. Statist. Soc. (B)* ***65***, 775–816. With discussion and reply by the authors. `arXiv:quant-ph/0307191`.
4. Fuchs, C. A. (1995). Distinguishability and Accessible Information in Quantum theory. PhD thesis, University of New Mexico; `arXiv:quant-ph/9601020`.
5. Gill, R. D. (2001). Asymptotics in quantum statistics. In A. W. van der Vaart,

M. de Gunst, and C. A. J. Klaassen (Eds.), *State of the Art in Probability and Statistics (Leiden, 1999)*, Volume **36** of *IMS Lecture notes Monogr. Ser.*, pp. 255–285. Beachwood, OH: Inst. Math. Statist. arXiv:math.ST/0405571.

6. Gill, R. D. and B. Y. Levit (1995). Applications of the Van Trees inequality: a Bayesian Cramér-Rao bound. *Bernoulli* ***1***, 59–79.
7. Gill, R. D. and S. Massar (2000). State estimation for large ensembles. *Phys. Rev. A* ***61***, 042312–042335. arXiv:quant-ph/9902063 .
8. Guţă, M. and J. Kahn (2006). Local asymptotic normality for qubit states. *Phys. Rev. A* **73**, 052108. arXiv:quant-ph/0512075.
9. Hayashi, M. (1998). Asymptotic estimation theory for a finite dimensional pure state model. *J. Phys. A* ***31***, 4633–4655. arXiv:quant-ph/9704041 .
10. Hayashi, M. (2003). Quantum estimation and quantum central limit theorem (in Japanese). *Sugaku* ***55***(4), 368–391. New, English translation: arXiv:quant-ph/0608198.
11. Hayashi, M. (2005) (editor). *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. Singapore: World Scientific.
12. Hayashi, M. and K. Matsumoto (2004). Asymptotic performance of optimal state estimation in quantum two level system. arXiv:quant-ph/0411073.
13. Holevo, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory*. Amsterdam: North-Holland. First appeared, in Russian, 1980.
14. Matsumoto, K. (2002). A new approach to the Cramér-Rao-type bound of the pure state model. *J. Phys. A* ***35***, 3111–3123. arXiv:quant-ph/9711008.
15. van Trees, H. (1968). *Detection, Estimation and Modulation Theory, Part 1*. New York: Wiley.

OPTIMAL QUANTUM FEEDBACK FOR CANONICAL OBSERVABLES

JOHN GOUGH

Department of Computing & Mathematics
Nottingham-Trent University, United Kingdom.
E-mail: john.gough@ntu.ac.uk

We consider the problem of optimal feedback control of a quantum system with linear dynamics undergoing continual non-demolition measurement of position and/or momentum, or both together. Specifically, we show that a stable domain of solutions for the filtered state of the system will be given by a class of randomized squeezed states and we exercise the control problem amongst these states. Bellman's principle is then applied directly to optimal feedback control of such dynamical systems and the Hamilton Jacobi Bellman equation for the minimum cost is derived. The situation of quadratic performance criteria is treated as the important special case and solved exactly for the class of relaxed states.

Keywords: Quantum Filtering, Optimal Control, Canonical Observables

1. Introduction

The advances in experimental physics over the last few decades with respect to the manipulation of individual quantum systems has renewed interest in theoretical schemes for the control of quantum systems. It is impossible to measure a quantum system with generating stochastic effects. Quantum noise was originally developed to model irreversible quantum dynamical systems, where it often played an external and secondary role, however, the realization that it could be measured, and that the results used to influence the system evolution, has had a profound effect on its physical status[1–5]. The issue we wish to address in this paper is the optimal control of a solvable model for a particle undergoing continuous non-demolition measurement of its canonical observables of position and momentum. Solvability of the associated stochastic master equation comes down to assuming that the internal dynamics is linear: that is generated by a quadratic Hamiltonian.

The theory of quantum feedback that is most familiar to physicists is

the one pioneered by Wiseman[5] for quantum optics. Here we make continual indirect measurements of a quantum system by means of some output channel, for instance a photocurrent, and feed the channel back onto system. The result is often improved convergence to and stability of a target state. As pointed out by Doherty and Jacobs[6] in their work on Kalman filtering for quantum state estimation, it is possible to broaden the outlook of what may be considered as feedback: rather than just direct feedback, one can consider suitable modifications of the output channel before feedback. In the information theoretic sense, one considers feedback to mean using the past output observations to influence present dynamical evolution: just as in classical control. Non-feedback control is also used, particularly for controlling molecular systems by laser pulses, and here algorithmic procedures exist to determine optimal control policies that achieve some desired effect with minimum energy cost[7].

The general theory of *optimal quantum feedback control*[3] sets out to determine optimal control policies for steering quantum systems so as to minimize some cost, for instance, the energy of operation. As in classical control, we must consider introducing an observer in order to obtain information about the current state and a controller to use this information to steer our system. The best estimate for the current state, conditional on the observations so far, is known as the *filtered state* and its evolution is described by a stochastic master equation which may be called the filtering equation. The ensuing problem of controlling the filtered state to meet some optimal cost criteria can then be considered as a separate problem, see for instance[8–10]. So far, this is analogous to classical control except that the noise present in the observations, which is of course quantum mechanical in origin, is no longer independent of the noise perturbing the state.

It should perhaps be remarked that the form of the filtering equation is a stochastic Schrödinger equation which falls within the category of equations unraveling completely positive Markovian semi-groups. The same class of equations has turned up in the considerations of several authors, for instance[1,11,12]: although mathematically identical, these equations differ significantly in their physical interpretation, derivation and status. Most of these treat the noise as an uncontrollable effect, however we specifically need the interpretation where it is a physically observable process and the stochastic state is just the filtered state, conditional on these observations. Experimentalists have already made the practical implementation of quantum state estimation and adaptive feedback control a reality. With this, has come new problems that have received intense interest in the physics

community[13–17].

In this paper, we wish to treat the problem of how to optimally control the quantum evolution of a system with linear free dynamics when we perform non-demolition measurement of, typically both, canonical position and momentum. Continuous position measurements on its own has been of historical importance and here the model is essentially the one considered by Ghirardi, Rimini and Weber[11], who also obtained the asymptotic form for the state. The asymptotic solution, with explicit reference to the stochastic Schrödinger equation within the Itô formulation, was first given by Diósi[19], with the full time solution given by Belavkin and Staszewski[20]. Effectively, the solution to the stochastic Schrödinger equation could be understood as an randomly parameterized squeezed state, that is, a Gaussian state with the parameters being mean position, mean momentum and a complex inverse variance. It was shown by Staszewski that the same class of states suffice for the stochastic Schrödinger equation describing simultaneous monitoring of position and momentum[21]. We re-derive this result and generalize to several dimensions.

Our main goal is optimal quantum feedback control of such a dynamical system. Bellman equations have been derived previously for the optimal cost of controlling a qubit system[8]. In fact, the general problem can be understood as a classical control problem on the space of quantum states[10] if one exploits the separation of quantum estimation component from the control component: here we may construct a, typically infinite dimensional, Hamilton Jacobi Bellman theory and are then faced with the problem of finding a sufficient parameterization of states for particular situation. In the case of non-demolition position and momentum measurements, we have that the squeezed states offer a sufficient parameterization. The quadratic performance problem is the important special case, Belavkin has previously given a arbitrary-dimensional solution to the quantum Kalman filter using the Heisenberg-Langevin approach[22] and Doherty and Jacobs[6] studied the dual problem of state estimation. Our contribution lies in deriving the optimal control through the Hamilton Jacobi Bellman principle. We would also like to mention that a similar approach has been given by Edwards and Belavkin[29] for a different set of sufficient state parameters.

2. Quantum Dynamical Programming

We shall begin by working through the Bellman principle and show how it leads to a partial differential equation, called the Hamilton-Jacobi-Bellman (HJB) equation, for the optimal cost. Our main dynamical assumption will

be that we remain within a finite parameterized family of states

$$\varrho = \rho(\theta) \tag{1}$$

where $\theta = (\theta_1, \cdots, \theta_n)^\top$ takes values in some parameter space Θ. We shall refer to θ as the set of sufficient coordinates for the problem. In particular, we shall assume that if we start in state $\varrho_t = \rho(\theta_t)$ at time t then the state evolves according to the (Itô) stochastic differential equation $d\theta = \mathbf{A}ds + \sum_\alpha \mathbf{B}_\alpha dW_\alpha$ or, more explicitly,

$$d\theta_i(s) = A_i(\theta_s, s; u_s)\, ds + \sum_\alpha B_{i,\alpha}(\theta_s, s)\, dW_\alpha(s) \tag{2}$$

where $\{u_s\}$ is a prescribed piece-wise continuous function taking values in some space $\mathcal{U}$ and the W_α are a finite number of independent Wiener processes. In other words, the dynamical evolution of the state is expressed in terms of the evolution of parameters which in this case undergo a diffusion in the parameter space Θ.

In optimal control, we wish to find a function $\{u_s\}$ which will minimize a pre-assigned cost. We refer to a given $\{u_s\}$ as a control function. The cost that we shall try to minimize will be assumed to take the form

$$\begin{aligned} J &= J[\theta, t; \{u\}] \\ &= \int_t^T \ell(\theta_s, s; u_s)\, ds + G(\theta_T). \end{aligned} \tag{3}$$

Here we have a "Lagrangian" ℓ that depends on the current state parameter θ_s, the current time s, and the current control parameter u_s. The time integral is from the initial time t to a fixed terminal time T: we have implicitly taken $0 < t < T$. In the definition, θ_s denoted the solution to the SDE (2) with initial condition being that we start at parameter value θ at time t. Finally we have an additional cost on termination $G(\theta_T)$, known as the bequest cost in control theory.

The actual cost J will vary from one experimental trial to another, and must be thought of as a random variable depending on the stochastic process $\{\theta_s : t \leq s \leq T\}$. The aim of this section is to evaluate the minimum average cost over all possible control policies, which we denote as

$$S(\theta, t) = \inf_{\{u\}} \overline{J[\theta; t; \{u\}]}. \tag{4}$$

Here we denote the average with a bar, and seek an infimum rather than a minimum just in case the optimal cost may only be attained as a limit.

2.1. *Bellman Optimality Principle*

Let us take $t < t + \Delta t < T$, Bellman's optimality principle is the simple observation that optimality in a given time interval implies optimality, upon restriction, to any sub-interval. However, this principle has far reaching consequences. Let $\Delta t > 0$, then restricting the optimal control problem from $[t, T]$ to $[t + \Delta t, T]$, we have that

$$S(\theta, t) = \inf_{\{u\}} \int_t^{t+\Delta t} \ell(\theta_s; s; u_s)\, ds + J[\theta_{t+\Delta t}, t + \Delta t; \{u\}]. \tag{5}$$

Note that $\theta_{t+\Delta t}$ will be random, which is why it too is averaged. We shall write $\theta_{t+\Delta t} = \theta + \Delta\theta_t$ where $\Delta\theta_t$ can be approximated by the Itô differential () above. We first note that

$$\int_t^{t+\Delta t} \ell(\theta_s, s; u_s)\, ds = \ell(\theta_t, t; u_t)\ \Delta t + o(\Delta t)$$

up to terms that are small of order in Δt. Likewise, assuming that S will be sufficiently differentiable,

$$\begin{aligned}
&S(\theta_{t+\Delta t}, t + \Delta t) \\
&= S(\theta, t) + \frac{\partial S}{\partial t}\Delta t + \frac{\partial S}{\partial \theta}^{\top} \Delta\theta_t + \frac{1}{2}\Delta\theta_t^{\top} \left.\frac{\partial^2 S}{\partial \theta^2}\right|_0 \Delta\theta_t + o(\Delta t) \\
&= S(\theta, t) + \left\{ \frac{\partial S}{\partial t} + \sum_i A_i \frac{\partial S}{\partial \theta_i} + \frac{1}{2}\sum_{i,j} \sigma_{ij} \frac{\partial^2 S}{\partial \theta_i \partial \theta_j} \right\} \Delta t \\
&\quad + \sum_i \sum_\alpha B_{i,\alpha} \frac{\partial S}{\partial \theta_i}\, \Delta W_\alpha(t) + o(\Delta t)
\end{aligned}$$

where we use the discrete Itô rule $\Delta W_\alpha(t).\Delta W_\beta(t) = \delta_{\alpha\beta}\,\Delta t + o(\Delta t)$ and introduce the diffusion matrix

$$\sigma_{ij} = \sum_\alpha B_{i,\alpha} B_{j,\alpha}.$$

The Bellman principle of optimality[26], see also[27] for instance, states that if $\{u_s^*\}$ is an optimal control function exercised over the time interval $t \le s \le T$ for a given start state at time t, then if we operate this policy up to time $t + \Delta t$ the remaining component of the control function will be optimal for the control problem over $t + \Delta t \le s \le T$ with initial condition now being that we start at the current (random) state at current time $t + \Delta T$. If we assume the existence of such an optimal control, then, within the above approximations with $\Delta t \to 0^+$, we are led to the partial differential

equation (Hamilton Jacobi Bellman equation, or just Bellman equation) for S

$$0 = \frac{\partial S}{\partial t} + \mathcal{H}\left(\theta, \frac{\partial S}{\partial \theta}, t\right) + \frac{1}{2}\sum_{i,j} \sigma_{ij} \frac{\partial^2 S}{\partial \theta_i \partial \theta_j}, \tag{6}$$

where we introduce

$$\mathcal{H}(\theta, \mathbf{I}, t) := \inf_{u \in \mathcal{U}} \left\{ \mathbf{I}^\top \mathbf{A}(\theta, t; u) + \ell(\theta, t; u) \right\}. \tag{7}$$

The Legendre transform (7) involves taking the infimum over the point value $u = u_t$ only. The transform parameters $\mathbf{I} = (I_1, \cdots, I_n)$ may be called the co-parameters. It should perhaps be stressed that the derivation of this equation is entirely classical. The equation is to be solved subject to the terminal condition

$$\lim_{t \to T^-} S(\theta, t) = G(\theta). \tag{8}$$

2.2. *Stochastic Schrödinger Equation*

We now consider a quantum system evolving with free Hamiltonian $H = H(u)$ while undergoing continual diffusive interaction with several independent apparatuses, each coupling to the system in a Markovian manner with coupling operator L_α for the α-th apparatus. (The $\{L_\alpha\}$ do not generally need to be either commuting or self-adjoint.)

The conditioned state of the system, ψ_t,continually updated using the output of the apparatuses, will then satisfy a stochastic Schrödinger equation of the type[3,18],

$$\begin{aligned} |d\psi_t\rangle = {} & \frac{1}{i\hbar} H(u_t) |\psi_t\rangle \, dt - \frac{1}{2} \sum_\alpha \left(L_\alpha^\dagger L_\alpha - 2\lambda_\alpha(t) L_\alpha + \lambda_\alpha^2(t) \right) |\psi_t\rangle \, dt \\ & + \sum_\alpha \left(L_\alpha - \lambda_\alpha(t) \right) |\psi_t\rangle \, dW_\alpha(t). \end{aligned} \tag{9}$$

where $\lambda_\alpha(t) = \mathrm{Re}\langle \psi_t | L_\alpha \psi_t \rangle$ and $\{W_\alpha\}$ is a multi-dimensional Wiener process with $dW_\alpha(t)\, dW_\beta(t) = \delta_{\alpha\beta}\, dt$. This equation was first derived in the context of filtering by Belavkin where the apparatuses are separate Bose fields. The processes $W_\alpha(t)$ are innovations processes: if $Y_\alpha(t)$ is the integrated measurement process from the α-th apparatus, then the corresponding innovations process is determined through the equation

$$dW_\alpha(t) = dY_\alpha(t) - \lambda_\alpha(t)\, dt. \tag{10}$$

Loosely speaking, the innovations give the difference between the observations and our expectations.

We are painfully aware that this type of equation is generic and has been derived in a variety of contexts by several authors: see[18] for references. We are forced to add the disclaimer that our interest resides solely in the interpretation of ψ_t as conditioned state and $W_\alpha(t)$ as the innovations coming from the observations. No other interpretation is applicable to control theory!

Let $\varrho_t = |\psi_t\rangle\langle\psi_t|$ be the von Neumann density matrix corresponding to the vector state. It obeys the SDE

$$d\varrho_t = w(\varrho_t; u_t)\, dt + \sum_\alpha \sigma_\alpha(\varrho_t)\, dW_\alpha(t), \tag{11}$$

with

$$w(\varrho; u) = i[\varrho, H(u)] + \sum_\alpha \left\{ L_\alpha \varrho L_\alpha^\dagger - \frac{1}{2}\varrho L_\alpha^\dagger L_\alpha - \frac{1}{2} L_\alpha^\dagger L_\alpha \varrho \right\},$$
$$\sigma_\alpha(\varrho) = L_\alpha \varrho + \varrho L_\alpha^\dagger - 2\lambda_\alpha \varrho.$$

Note that $\lambda_\alpha \equiv \frac{1}{2}\mathrm{tr}\{\varrho(L_\alpha + L_\alpha^\dagger)\}$ and so the SDE is nonlinear in ϱ.

A simple example is given by taking the qubit. Here the states are parameterized as $\varrho(\theta) = \frac{1}{2}(1 + \theta.\sigma)$ where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices and $\theta = (\theta_x, \theta_y, \theta_z)$ with Θ being the Bloch sphere. With fixed choices of operators $H(\cdot)$, L_α and control function $\{u_s\}$ we obtain a diffusion process θ_s on the Bloch sphere. The Hamilton-Jacobi-Bellman equation theory for this problem has been treated in[8].

3. Control of Canonically Observables

The stochastic Schrödinger equation for measurement of canonically conjugate observables, $\hat{q}$ and $\hat{p}$, is given by

$$|d\psi_t\rangle = \left(\frac{1}{i\hbar} H - \frac{\kappa}{4}(\hat{q} - \langle\hat{q}\rangle_t)^2 - \frac{\tilde{\kappa}}{4}(\hat{p} - \langle\hat{p}\rangle_t)^2 \right) |\psi_t\rangle\, dt$$
$$+ \sqrt{\frac{\kappa}{2}}(\hat{q} - \langle\hat{q}\rangle_t)|\psi_t\rangle\, dW_t + \sqrt{\frac{\tilde{\kappa}}{2}}(\hat{p} - \langle\hat{p}\rangle_t)|\psi_t\rangle\, d\tilde{W}_t. \tag{12}$$

Here we have two independent apparatuses and we have set $L_1 = \sqrt{\frac{\kappa}{2}}\hat{q}$ and $L_2 = \sqrt{\frac{\tilde{\kappa}}{2}}\hat{p}$ with the corresponding innovations being denoted as

$W_1(t) = W_t$ and $W_2(t) = \tilde{W}_t$. It has been derived from first principles by Scott and Milburn[24]. They considered a discrete time model with simultaneous measurement of position and momentum by separate apparatuses, and considered the continuous time limit of progressively more imprecise and frequent measurements.

The equation involves the expectations $\langle \hat{q} \rangle_t = \langle \psi_t | \hat{q}\, \psi_t \rangle$ and $\langle \hat{p} \rangle_t = \langle \psi_t | \hat{p}\, \psi_t \rangle$. Here the constants κ $\mathrm{m}^{-2}\mathrm{s}^{-1}$ and $\tilde{\kappa}$ $\mathrm{N}^{-2}\mathrm{s}^{-3}$ are positive and describe the measurement strength for the two apparatuses. In general, κ and $\tilde{\kappa}$ has units of inverse variance of position, respectively momentum, per unit time. In[25], the limiting procedure was revisited and, as an alternative to increasingly imprecise measurements, one could use increasingly weak interaction between the apparatuses and the system. The scaling between the imprecision of measurement, or weakness of interaction with the apparatus, and the rate at which the discrete measurements is made must be such as to allow a general central limit effect to take place. In principle, it is possible, to set up the apparatuses to obtain desired values of κ and $\tilde{\kappa}$.

The purpose of[24] was to consider nonlinear dynamics, however, we shall only deal with quadratic Hamiltonians of the type $H = H(u)$ with $u = (u_1, u_2) \in \mathbb{R}^2$ and

$$H(u) = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{q}^2 + \frac{1}{2}\zeta\left(\hat{q}\hat{p} + \hat{p}\hat{q}\right) + u_1\hat{q} + u_2\hat{p}. \tag{13}$$

It was shown by Staszweski[21] that there exists a set of sufficient parameters to describe the dynamical evolution of the state for this system.

3.1. *Squeezed States*

Let $L^2(\mathbb{R})$ be the Hilbert space of square integrable functions of position coordinate x with standard Schrödinger representation of the canonical observables $\hat{q}$ and $\hat{p}$. A state $\langle \cdot \rangle$ is said to be Gaussian if we have

$$\langle \exp\{ir\hat{q} + is\hat{p}\} \rangle = \exp\left\{ ir\bar{q} + is\bar{p} - \frac{1}{2}\left(C_{qq}r^2 + 2C_{qp}rs + C_{pp}s^2\right)\right\} \tag{14}$$

where we have the interpretation that the mean values of the position and the momentum are $\langle \hat{q} \rangle = \bar{q}$ and $\langle \hat{p} \rangle = \bar{p}$ respectively while the covariance matrix is given by

$$C_{XY} = \frac{1}{2}\langle XY + YX \rangle - \langle X \rangle \langle Y \rangle,$$

that is, C_{qq} is the variance of $\hat{q}$, C_{pp} is the variance of $\hat{p}$, and $C_{qp} = \frac{1}{2}\langle \hat{q}\hat{p} + \hat{p}\hat{q} \rangle - \langle \hat{p} \rangle \langle \hat{q} \rangle$ is the covariance of $\hat{q}$ and $\hat{p}$.

We now introduce a wave function $\psi(\theta)$, parameterized by $\theta = (\bar{q}, \bar{p}, \eta)$, where we have real numbers $\bar{q}, \bar{p}$ and a complex number $\eta = \eta' + i\eta''$ with $\eta' > 0$, and taking the form

$$\langle x|\psi(\bar{q},\bar{p},\eta)\rangle = \left(\frac{\eta'}{2\pi}\right)^{1/4} \exp\left\{-\frac{\eta}{4}(x-\bar{q})^2 + i\frac{\bar{p}}{\hbar}x\right\}. \tag{15}$$

When η is real ($\eta'' = 0$), the vectors are just the well-known coherent states[28], otherwise they describe squeezed states The distribution of the canonical variables in the state $\psi(\bar{q}, \bar{p}, \eta)$ is Gaussian with characteristic function (see appendix A)

$$\langle \exp\{ir\hat{q} + is\hat{p}\}\rangle_{\bar{q},\bar{p},\eta} = \exp\left\{ir\bar{q} + is\bar{p} - \frac{1}{2}\left(C_{qq}r^2 + 2C_{qp}rs + C_{pp}s^2\right)\right\}, \tag{16}$$

where

$$C_{qq} = \frac{1}{\eta'},\ C_{qp} = -\frac{\hbar\eta''}{2\eta'},\ C_{pp} = \frac{\hbar^2}{4}\left(\eta' + \frac{\eta''^2}{\eta'}\right). \tag{17}$$

3.2. *Evolution of Parameters*

We might reasonably expect that, for the filtering equation (12) with the quadratic Hamiltonian above, we will remain with the class of squeezed states. That is,

$$\psi_t \equiv \psi(\bar{q}_t, \bar{p}_t, \eta_t). \tag{18}$$

This is indeed the case, and the dynamical evolution of the parameters is given by the system of equations (see appendix B)

$$d\bar{q} = \left(\frac{1}{m}\bar{p} + \zeta\bar{q} + u_2\right)dt + \sqrt{2\kappa}C_{qq}\,dW + \sqrt{2\tilde{\kappa}}C_{qp}\,d\tilde{W}, \tag{19}$$

$$d\bar{p} = -\left(m\omega^2\bar{q} + \zeta\bar{p} + u_1\right)dt + \sqrt{2\kappa}C_{qp}\,dW + \sqrt{2\tilde{\kappa}}C_{pp}\,d\tilde{W}, \tag{20}$$

$$\frac{d\eta}{dt} = 2\left(\kappa + i\frac{m\omega^2}{\hbar}\right) - 2\zeta\eta - \frac{1}{2}\left(\tilde{\kappa}\hbar^2 + i\frac{\hbar}{m}\right)\eta^2. \tag{21}$$

We note that the last equation is a deterministic Riccati equation and so the (co)-variances evolve in a non-random manner. The first two are the linear in the parameters $\bar{q}, \bar{p}$ however note that noise coefficients depend on η.

3.3. *Asymptotic States*

The Riccati equation (20) is to be solved in the half plane $\eta' > 0$ of physical solutions and has the unique, globally attractive, fixed point η_∞ in that region. In the special case of a harmonic oscillator ($\zeta = 0$) the fixed point is given by

$$\eta_\infty^{\mathrm{osc}} = \frac{2}{\hbar} \sqrt[+]{\frac{\hbar\kappa + im\omega^2}{\hbar\tilde{\kappa} + im^{-1}}}. \tag{22}$$

(Here $\sqrt[+]{\cdot}$ denotes the complex root having positive real part.)

We may actually achieve a coherent state (η_∞ real) in the limit state if we tune the measurement strengths such that $\kappa \equiv m^2\omega^2\,\tilde{\kappa}$. In this case, $\eta_\infty^{\mathrm{osc}} \equiv \dfrac{2m\omega}{\hbar}$, corresponding to a coherent state with position uncertainty $\sigma_\infty^{\mathrm{osc}} = \sqrt{\dfrac{\hbar}{2m\omega}}$. Otherwise the limit state will be squeezed. We should remark that $\sqrt{\dfrac{\kappa}{\tilde{\kappa}}}$ corresponds to the squeezing parameter s introduced in[24] to describe the bias in favor of the $\hat{q}$ or $\hat{p}$ coupling. We may say that the measurement strengths are balanced for the oscillator when $s = m\omega$ leading to a limit coherent state.

3.4. *The Optimal Control Problem*

The Hamilton Jacobi Bellman equation for this problem is, with $S = S(t; \bar{q}, \bar{p}, \eta)$,

$$\begin{aligned} 0 = \frac{\partial S}{\partial t} + \mathcal{H}\left(t; \bar{q}, \bar{p}, \eta; \frac{\partial S}{\partial \bar{q}}, \frac{\partial S}{\partial \bar{p}}\right) + \frac{\partial S}{\partial \eta'}\frac{d\eta'}{dt} + \frac{\partial S}{\partial \eta''}\frac{d\eta''}{dt}. \\ + \frac{\partial^2 S}{\partial \bar{q}^2}\left[\kappa C_{qq}^2 + \tilde{\kappa} C_{qp}^2\right] + 2\frac{\partial^2 S}{\partial \bar{q}\partial \bar{p}}\sqrt{\kappa\tilde{\kappa}}\left[C_{qq} + C_{pp}\right] C_{qp} + \frac{\partial^2 S}{\partial \bar{p}^2}\left[\kappa C_{qp}^2 + \tilde{\kappa} C_{pp}^2\right] \end{aligned} \tag{23}$$

with terminal condition $\lim_{t\to T^-} S(t; \bar{q}, \bar{p}, \eta) = g(\bar{q}, \bar{p}, \eta)$. With the shorthand $x = (\bar{q}, \bar{p})^\top$, $I^\top = (I_q, I_p)$ and $u = (u_1, u_2)^\top$, we may write

$$\mathcal{H}(t; \bar{q}, \bar{p}, \eta; I_q, I_p) := \min_{u\in\mathbb{R}^2} \left\{ I^\top (F_t x + M_t u) + \ell(t, u, x) \right\}$$

where

$$F_t = \begin{pmatrix} \zeta & m^{-1} \\ -m\omega^2 & -\zeta \end{pmatrix}, \quad M_t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In principle, once a minimizing solution $u^*_\alpha = u_\alpha(t;\bar{q},\bar{p},\eta)$ is known, it may be used as a Markov control for closed loop feedback: that is, the control policies are taken as these functions of the current state parameters.

The Bellman equations arising in quantum feedback control have so far proved to be highly nonlinear and prohibitively hard to solve as a rule. Our equation is no exception, however, the nonlinearities are in due to the η variable. We remark that if we assume that we start off in a state relaxed at the equilibrium value $\eta = \eta_\infty$, then the coefficients of the η', η'' derivatives vanish exactly, and we may take the covariances C_{qq}, C_{qp} and C_{pp} at their relaxed values determined by the asymptotic value η_∞. As the relaxation time is typically small, we may justify this for large times T in comparison. This ignores any η-transient contribution to the cost, but at least opens up the possibility of solving the Bellman equation and finding optimal Markov control policies. We give the fundamental class of interest, quadratic performance criteria, next.

3.5. *Quantum Linear Stochastic Regulator*

We consider the following quadratic control problem not involving any costs on the η parameter. In particular, we make the assumption that the starting state is an asymptotic state ($\eta = \eta_\infty$) and so we ignore η as a variable. We set and take the specific choices

$$\begin{aligned} \ell(x,t,u) &= \frac{1}{2}x^\top A_t x + \frac{1}{2}u^\top Q_t u, \\ G(x) &= \frac{1}{2}x^\top R x, \end{aligned} \tag{24}$$

where P_t, Q_t and R are 2×2 symmetric matrices with Q_t being invertible. The control problem is now essentially the same as the classical stochastic regulator[27]. In this case the $\mathcal{H}$-function is

$$\mathcal{H}(t,x,I) = \frac{1}{2}x^\top P_t x + I^\top F_t x + \min_u \left\{ \frac{1}{2}u^\top Q_t u + I^\top M_t u \right\}$$

with the minimum attained at

$$u^* = -Q_t^{-1} M^\top I$$

and we find

$$\mathcal{H}(t,x,I) = \frac{1}{2}x^\top P_t x + I^\top F_t x - \frac{1}{2} I^\top M_t Q_t^{-1} M^\top I$$

Seeking an η-independent solution, the Bellman equation reduces to

$$0 = \frac{\partial S}{\partial t} + \mathcal{H}(t,x,\nabla S) + \frac{1}{2}K_{ij}\frac{\partial^2 S}{\partial x_i \partial x_j}.$$

Here K is the matrix of the second order coefficients and these will be determined by the covariances (16) at the asymptotic value η_∞. As is well known[27], the solution takes the form

$$S(t,x) \equiv \frac{1}{2} x^\top \Sigma_t x + a_t \tag{25}$$

where Σ_t satisfies the matrix Riccati equation

$$\frac{d\Sigma_t}{dt} = -\Sigma_t F_t - F^\top \Sigma_t + \Sigma_t M_t Q_t^{-1} M^\top \Sigma_t - P_t, \qquad \Sigma_T = R, \tag{26}$$

while a_t satisfies

$$\frac{da_t}{dt} = -tr\left\{K\Sigma_t\right\}, \qquad a_T = 0. \tag{27}$$

The optimal control policy is therefore given by

$$u^*(x,t) = -Q_t^{-1} M^\top \nabla S = Q_t^{-1} M^\top \Sigma_t x. \tag{28}$$

4. Several Dimensions

If we now have d degrees of freedom, leading to canonical variables $\hat{\mathbf{q}} = (\hat{q}_1, \cdots, \hat{q}_d)$ and $\hat{\mathbf{p}} = (\hat{p}_1, \cdots, \hat{p}_d)$ with non-zero commutators $[\hat{q}_\mu, \hat{p}_\nu] = i\hbar\delta_{\mu\nu}$. The appropriate class of vector to consider is the multidimensional squeezed state with parametrization $\theta = (\bar{\mathbf{q}}, \bar{\mathbf{p}}, \eta)$

$$\langle \mathbf{x} | \psi(\theta) \rangle = \left(\frac{\det \eta'}{(2\pi)^d} \right)^{\frac{1}{4}} \exp\left\{ -\frac{1}{4} (\mathbf{x} - \bar{\mathbf{q}})^\top \eta (\mathbf{x} - \bar{\mathbf{q}}) + \frac{i}{\hbar} \bar{\mathbf{p}}^\top \mathbf{x} \right\}$$

where this time $\bar{\mathbf{q}}, \bar{\mathbf{p}} \in \mathbb{R}^d$ and $\eta = \eta' + i\eta''$ with η', η'' real symmetric $d \times d$ matrices with η' positive definite (and hence invertible).

The characteristic function for the state is

$$\begin{aligned} &\left\langle \exp\left\{ i\mathbf{r}^\top \hat{\mathbf{q}} + i\mathbf{s}^\top \hat{\mathbf{p}} \right\} \right\rangle_{\bar{\mathbf{q}}, \bar{\mathbf{p}}, \eta} \\ &= \exp\left\{ i\mathbf{r}^\top \bar{\mathbf{q}} + i\mathbf{s}^\top \bar{\mathbf{p}} - \frac{1}{2} \left(\mathbf{r}^\top C_{qq} \mathbf{r} + 2\mathbf{r}^\top C_{qp} \mathbf{s} + \mathbf{s}^\top C_{pp} \mathbf{s} \right) \right\}, \end{aligned}$$

where we now encounter the matrices

$$\begin{aligned} C_{qq} &= (\eta')^{-1}, \; C_{qp} = -\frac{\hbar}{4} \left(\eta'' (\eta')^{-1} + (\eta')^{-1} \eta'' \right), \\ C_{pp} &= \frac{\hbar^2}{4} \left(\eta' + \eta'' (\eta')^{-1} \eta'' \right). \end{aligned}$$

Some comments are in order. The proof of is actually a straightforward generalization of the $d = 1$ case in appendix A: let R be the positive

square root of η' and consider the canonical transform $\tilde{q}_k = \sum_j R_{kj}\hat{q}_j, \tilde{p}_k = \sum_j \hat{p}_j R^{-1}_{jk}$ with creation/annihilation operators $\tilde{a}^{\pm}_i = \frac{1}{2}\tilde{q}_i \pm \frac{1}{i\hbar}\tilde{p}_i$, so that we have the CCR $\left[\tilde{a}^-_j, \tilde{a}^+_k\right] = \delta_{jk}$, the remaining manipulations are then an easy multi-dimensional version of what we had before with the $\tilde{a}^{\pm}_j$ in place of the single pair $a^{\pm}$. Next we remark that we have $d^2 - d$ degrees of freedom in the pair of matrices η' and η'': this is readily seen to be the correct number needed to specify the covariances of a state $\langle\cdot\rangle$ since it is required only to specify the values $\left\langle \tilde{a}^+_j \tilde{a}^-_k \right\rangle$ and $\left\langle \tilde{a}^+_j \tilde{a}^+_k \right\rangle$ for $j \leq k$ as all other second order expectations can be deduced from the CCR or by taking conjugates.

We consider monitoring each of the canonical variables with a separate independent apparatus, the corresponding stochastic Schrödinger equation will then be

$$|d\psi_t\rangle = \left(\frac{1}{i\hbar}H - \sum_\alpha \frac{\kappa_\alpha}{4}\left(\hat{q}_\alpha - \langle\hat{q}_\alpha\rangle_t\right)^2 - \sum_\alpha \frac{\tilde{\kappa}_\alpha}{4}\left(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle_t\right)^2\right)|\psi_t\rangle\, dt$$
$$+ \sum_\alpha \sqrt{\frac{\kappa_\alpha}{2}}\left(\hat{q}_\alpha - \langle\hat{q}_\alpha\rangle_t\right)|\psi_t\rangle\, dW_\alpha(t) + \sum_\alpha \sqrt{\frac{\tilde{\kappa}_\alpha}{2}}\left(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle_t\right)|\psi_t\rangle\, d\tilde{W}_\alpha(t) \tag{29}$$

with the obvious interpretations.

We again take a quadratic Hamiltonians of the type $H = H(u)$ with $u = (\mathbf{u}_1, \mathbf{u}_2) \in \mathbb{R}^d \times \mathbb{R}^d$ and

$$H(u) = \sum_{\mu,\nu=1}^{d}\left(\frac{1}{2}\hat{p}_\mu \Gamma_{\mu\nu}\hat{p}_\nu + \frac{1}{2}\hat{q}_\mu\tilde{\Gamma}_{\mu\nu}\hat{q}_\nu\right) + \mathbf{u}_1^\top\hat{\mathbf{q}} + \mathbf{u}_2^\top\hat{\mathbf{p}}. \tag{30}$$

This leads to the following evolution of the state parameters

$$d\bar{\mathbf{q}} = (\Gamma\bar{\mathbf{p}} + \mathbf{u}_2)\, dt + \sum_\alpha \sqrt{2\kappa_\alpha}C(\hat{\mathbf{q}}, \hat{q}_\alpha)\, dW_\alpha + \sum_\alpha \sqrt{2\tilde{\kappa}_\alpha}C(\hat{\mathbf{q}}, \hat{p}_\alpha)\, d\tilde{W}_\alpha,$$
$$d\bar{\mathbf{p}} = -\left(\tilde{\Gamma}\bar{\mathbf{q}} + \mathbf{u}_1\right) dt + \sum_\alpha \sqrt{2\kappa_\alpha}C(\hat{\mathbf{p}}, \hat{q}_\alpha)\, dW_\alpha + \sum_\alpha \sqrt{2\tilde{\kappa}_\alpha}C(\hat{\mathbf{p}}, \hat{p}_\alpha)\, d\tilde{W}_\alpha,$$
$$\frac{d\eta}{dt} = 2\left(\kappa + \frac{i}{\hbar}\tilde{\Gamma}\right) - \frac{1}{2}\hbar^2\left(\tilde{\kappa} + \frac{i}{\hbar}\Gamma^{-1}\right)\eta^2$$

where we introduce the diagonal matrices $\kappa = diag(\kappa_1, \cdots, \kappa_d)$ and $\tilde{\kappa} = diag(\tilde{\kappa}_1, \cdots, \tilde{\kappa}_d)$.

The condition that the steady state be coherent is then

$$\tilde{\Gamma}\Gamma = diag\left(\frac{\kappa_1}{\tilde{\kappa}_1}, \cdots, \frac{\kappa_d}{\tilde{\kappa}_d}\right)$$

which may, of course, not always be realized.

The control problem for the multi-dimensional quantum stochastic regulator is then just the trivial extension of the $d = 1$ case considered in th last section.

5. Appendix

5.1. *Derivation of the Characteristic Function* ($d = 1$)

Coherent states may be constructed from creation/annihilation operators $a^{\pm} = \frac{1}{2}\sqrt{\eta'}\hat{q} \pm \frac{1}{i\hbar\sqrt{\eta'}}\hat{p}$ by identifying $\psi(\bar{q}, \bar{p}, \eta')$ as the eigenstate of a^- with eigenvalue $\alpha = \frac{1}{2}\sqrt{\eta'}\bar{q} - \frac{1}{i\hbar\sqrt{\eta'}}\bar{p}$. In particular, if Ω denotes the zero-eigenstate of a^- then

$$\psi(\bar{q}, \bar{p}, \eta') = D_\alpha \, \Omega$$

where $D_\alpha = \exp\{\alpha a^+ - \alpha^* a^-\}$ is a Weyl displacement unitary. Next observe that we may obtain squeezed states from coherent states by the simple application of a unitary transformation:

$$\psi(\bar{q}, \bar{p}, \eta' + i\eta'') \equiv V \, \psi(\bar{q}, \bar{p}, \eta')$$

with $V = \exp\left\{-\frac{i}{4}\eta''(\hat{q} - \bar{q})^2\right\}$. (This transformation is, in fact, linear canonical.) We may introduce new canonical variables $\hat{q}'$ and $\hat{p}'$ by $\hat{q}' = V^\dagger \hat{q} V \equiv \hat{q}$ and $\hat{p}' = V^\dagger \hat{p} V = \hat{p} - \frac{1}{2}\hbar\eta''(\hat{q} - \bar{q})$. We note that $\exp\{ir\hat{q} + is\hat{p}\} = D_z$ where $z = -\frac{1}{2}\hbar\sqrt{\eta'}s + i\frac{1}{\sqrt{\eta'}}r$ and

$$V^\dagger D_z V = \exp\{ir\hat{q}' + is\hat{p}'\} = D_w \, e^{\frac{1}{2}i\hbar\eta''\bar{q}s}$$

where $w = -\frac{1}{2}\hbar\sqrt{\eta'}s + i\frac{1}{\sqrt{\eta'}}\left(r - \frac{1}{2}\hbar\eta''s\right)$. Using well-known properties for Weyl displacement operators[28] and their Ω-state averages, we find

$$\begin{aligned}\langle \exp\{ir\hat{q} + is\hat{p}\}\rangle_{\bar{q},\bar{p},\eta} &= \langle \Omega | D_\alpha^\dagger V^\dagger D_z V D_\alpha \, \Omega\rangle \\ &= \langle \Omega | D_\alpha^\dagger D_w D_\alpha \, \Omega\rangle \, e^{\frac{1}{2}i\hbar\eta''\bar{q}s} \\ &= e^{w\alpha^* - w^*\alpha - \frac{1}{2}|w|^2} e^{\frac{1}{2}i\hbar\eta''\bar{q}s}\end{aligned}$$

and substituting in for α and w gives the required result.

We may say that the canonical variables are Weyl independent for a given quantum state $\langle \cdot \rangle$, not necessarily pure, if we have the following

factorization

$$\langle \exp\{ir\hat{q}+is\hat{p}\}\rangle = \langle \exp\{ir\hat{q}\}\rangle \, \langle \exp\{is\hat{p}\}\rangle$$

for all real r and s. If the state possesses moments to all orders, then Weyl independence means that symmetrically (Weyl) ordered moments factor according to $\langle : f(\hat{q})\, g(\hat{p}) : \rangle = \langle f(\hat{q})\rangle \langle g(\hat{p})\rangle$, for all polynomials f, g. By inspection, we see that coherent states leave the canonical variables Gaussian and Weyl-independent. However, the $\eta'' \neq 0$ squeezed states do not have this property.

5.2. *Parameter Evolution*

Let $\langle X\rangle_t = \langle \psi_t | X | \psi_t\rangle$, for a general operator X, then we have the following stochastic differential equation

$$d\langle X\rangle = \left\{ \frac{1}{i\hbar}\langle [X,H]\rangle - \frac{\kappa}{4}\langle [[X,\hat{q}],\hat{q}]\rangle - \frac{\tilde{\kappa}}{4}\langle [[X,\hat{p}],\hat{p}]\rangle \right\} dt$$
$$+\sqrt{\frac{\kappa}{2}}\left(\langle X\hat{q}+\hat{q}X\rangle - \langle \hat{q}\rangle\langle X\rangle\right) dW_t + \sqrt{\frac{\tilde{\kappa}}{2}}\left(\langle X\hat{p}+\hat{p}X\rangle - \langle \hat{p}\rangle\langle X\rangle\right) d\tilde{W}_t. \tag{31}$$

For $X = \hat{q}, \hat{p}$, we find (suppressing the t-dependencies)

$$d\langle \hat{q}\rangle = \left(\frac{1}{m}\langle \hat{p}\rangle + \zeta\langle \hat{q}\rangle + u_2\right) dt + \sqrt{2\kappa}C(\hat{q},\hat{q})\, dW_t + \sqrt{2\tilde{\kappa}}C(\hat{q},\hat{p})\, d\tilde{W}_t,$$
$$d\langle \hat{p}\rangle = -\left(m\omega^2\langle \hat{q}\rangle + \zeta\langle \hat{p}\rangle + u_1\right) dt + \sqrt{2\kappa}C(\hat{q},\hat{p})\, dW_t + \sqrt{2\tilde{\kappa}}C(\hat{p},\hat{p})\, d\tilde{W}_t.$$

where $C(\hat{q},\hat{q}) = \langle \hat{q}^2\rangle - \langle \hat{q}\rangle^2, C(\hat{p},\hat{p}) = \langle \hat{p}^2\rangle - \langle \hat{p}\rangle^2$, and $C(\hat{q},\hat{p}) = \frac{1}{2}\langle \hat{p}\hat{q}+\hat{q}\hat{p}\rangle - \langle \hat{p}\rangle\langle \hat{q}\rangle$.

We now make the ansatz that the state $\psi_t = \psi(\bar{q}_t, \bar{p}_t, \eta_t)$. Let r, s be fixed real parameters and set $D = \exp\{ir\hat{q}+is\hat{p}\}$. We shall investigate the evolution through the characteristic function

$$G_t = \langle \psi_t | D | \psi_t\rangle = \langle D\rangle_t .$$

Observing that $[D,\hat{q}] = \hbar s D$, $[D,\hat{p}] = -\hbar r D$, we find that we can reduce all the expectations in (32) with $X = D$ down to just combinations of $\langle \hat{q}D\rangle, \langle D\hat{q}\rangle, \langle \hat{p}D\rangle$ and $\langle D\hat{p}\rangle$. The Baker Campbell Hausdorff formula, $e^{ir\hat{q}+is\hat{p}} = e^{\frac{1}{2}irs\hbar}e^{ir\hat{q}}e^{is\hat{p}} = e^{-\frac{1}{2}irs\hbar}e^{is\hat{p}}e^{ir\hat{q}}$, allows us to compute that

$$\langle \hat{q}D\rangle = e^{\frac{1}{2}irs\hbar}\frac{1}{i}\frac{\partial}{\partial r}\left(e^{-\frac{1}{2}irs\hbar}G\right) = \left(\bar{q} + i\left(C_{qq}r + C_{qp}s\right) + \frac{1}{2}s\hbar\right)G,$$

and likewise

$$\langle D\hat{q}\rangle = \left(\bar{q} + i\left(C_{qq}r + C_{qp}s\right) - \frac{1}{2}s\hbar\right)G,$$
$$\langle \hat{p}D\rangle = \left(\bar{p} + i\left(C_{qp}r + C_{pp}s\right) + \frac{1}{2}r\hbar\right)G,$$
$$\langle D\hat{p}\rangle = \left(\bar{p} + i\left(C_{qp}r + C_{pp}s\right) - \frac{1}{2}r\hbar\right)G.$$

Hence

$$dG = \frac{ir}{m}\left\{\bar{p} + i\left(C_{qp}r + C_{pp}s\right)\right\}G\,dt - is\hbar\frac{m\omega^2}{\hbar}\left\{\bar{q} + i\left(C_{qq}r + C_{qp}s\right)\right\}G\,dt$$
$$+\zeta\left(ir\bar{q} - is\bar{p} + C_{pp}s^2 - r^2C_{qq}\right)G\,dt + \left(-iu_1s + iu_2r - \frac{\kappa\hbar^2s^2}{4} - \frac{\tilde{\kappa}\hbar^2r^2}{4}\right)G\,dt$$
$$+ i\sqrt{2\kappa}\left(C_{qq}r + C_{qp}s\right)G\,dW + i\sqrt{2\tilde{\kappa}}\left(C_{qp}r + C_{pp}s\right)G\,d\tilde{W}.$$

Under our ansatz (17), we should also have, by the Itô rule,

$$\begin{aligned} dG &= \frac{\partial G}{\partial \bar{q}}d\bar{q} + \frac{\partial G}{\partial \bar{p}}d\bar{p} + \frac{1}{2}\frac{\partial^2 G}{\partial \bar{q}^2}\left(d\bar{q}\right)^2 + \frac{\partial^2 G}{\partial \bar{q}\partial \bar{p}}\left(d\bar{q}d\bar{p}\right) + \frac{1}{2}\frac{\partial^2 G}{\partial \bar{p}^2}\left(d\bar{p}\right)^2 \\ &\quad + \frac{\partial G}{\partial \eta'}d\eta' + \frac{\partial G}{\partial \eta''}d\eta'' \\ &= irGd\bar{q} + isGd\bar{p} - \frac{1}{2}r^2G\left(d\bar{q}\right)^2 - rsG\left(d\bar{q}d\bar{p}\right) - \frac{1}{2}s^2G\left(d\bar{p}\right)^2 \\ &\quad + \left(\frac{1}{2\eta'^2}r^2 - \frac{\hbar\eta''}{2\eta'}rs - \frac{\hbar^2}{8}\left(1 - \frac{\eta''^2}{\eta'^2}\right)s^2\right)d\eta' \\ &\quad + \left(\frac{\hbar}{2\eta'}rs - \frac{1}{4}\frac{\hbar^2\eta''}{\eta'}s^2\right)d\eta''. \end{aligned} \tag{32}$$

Equating the coefficients of (29) and (30) gives the system of equations

$$r, d\bar{q} = \left(\frac{1}{m}\bar{p} + \zeta\bar{q} + u_2\right)dt + \sqrt{2\kappa}C_{qq}\,dW + \sqrt{2\tilde{\kappa}}C_{qp}\,d\tilde{W},$$
$$s, d\bar{p} = -\left(\hbar\frac{m\omega^2}{\hbar}\bar{q} + \zeta\bar{p} + +u_1\right)dt + \sqrt{2\kappa}C_{qp}\,dW + \sqrt{2\tilde{\kappa}}C_{pp}\,d\tilde{W},$$
$$r^2, \left(d\bar{q}\right)^2 - \frac{1}{\eta'^2}d\eta' = \frac{2}{m}C_{qp}dt + 2\zeta C_{qq}dt + \frac{\tilde{\kappa}\hbar^2}{2}\,dt,$$
$$s^2, \left(d\bar{p}\right)^2 + \frac{\hbar^2}{4}\left(1 - \frac{\eta''^2}{\eta'^2}\right)d\eta' + \frac{1}{2}\frac{\hbar^2\eta''}{\eta'}\,d\eta'' = -2\hbar\frac{m\omega^2}{\hbar}C_{qp}\,dt + \frac{\kappa\hbar^2}{2}\,dt - 2\zeta C_{pp}dt,$$
$$rs, \left(d\bar{q}d\bar{p}\right) + \frac{\hbar\eta''}{2\eta'^2}\,d\eta' - \frac{\hbar}{2\eta'}\,d\eta'' = \frac{1}{m}C_{pp}\,dt - \hbar\frac{m\omega^2}{\hbar}\,C_{qq}dt.$$

The first two of these agree exactly with (18,19), while the next three are entirely consistent with the pair of real equations

$$\begin{cases} \dfrac{d}{dt}\eta' = 2\kappa + \dfrac{\hbar}{m}\eta'\eta'' - 2\zeta\eta' - \frac{1}{2}\tilde{\kappa}\hbar^2\left(\eta'^2 - \eta''^2\right), \\ \dfrac{d}{dt}\eta'' = 2\frac{m\omega^2}{\hbar} - 2\zeta\eta'' - \dfrac{\hbar}{2m}\left(\eta'^2 - \eta''^2\right) - \tilde{\kappa}\hbar^2\eta'\eta'', \end{cases} \tag{33}$$

which, together, they are equivalent to the single complex Riccati equation (20).

Acknowledgment: The author is indebted to Luc Bouten for a critical reading of the article, and to Simon Edwards for pointing out reference[29].

Bibliography

1. V.P. Belavkin, Phys. Lett. A **140**, 355 (1989)
2. M.J. Collett and C.W. Gardiner, Phys. Rev. A **30**, 1386 (1984)
3. V.P. Belavkin, Rep. Math. Phys., **43**, 405 (1999)
4. C.M. Caves, G.J. Milburn, Phys. Rev. A **36**, 5543 (1987)
5. H.M. Wiseman, Phys. Rev. A **49**, 2133 (1994)
6. A.C. Doherty and K. Jacobs, Phys. Rev. A, **60**, 2700 (1999)
7. R.S. Judson and H. Rabitz, Phys. Rev. Lett. **68**, 1500-1503 (1992)
8. L. Bouten, S C. Edwards, V.P. Belavkin, J. Phys. B: At. Mol. Opt. Phys. **38**, 151 (2005)
9. M.R. James, J. Opt. B: Quantum Semiclass. Opt. **7**, S198-S207 (2005)
10. J. Gough, V.P. Belavkin, O.G. Smolyanov, J. Opt. B: Quantum Semiclass. Opt. **7**, S237-S244 (2005)
11. G.C. Ghirardi, A. Rimini and T. Weber, Phys. Rev. D, **34**, 470 (1986)
12. L. Diósi, Phys. Lett. A, **129**, 419 (1988)
13. A.C. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. Tan, Phys. Rev. A, **62**:012105 (2000)
14. M. Armen, J. Au, J. Stockton, A. Doherty, and H. Mabuchi, Phys. Rev. A **89**:133602, (2002)
15. J. Geremia, J. Stockton, A.C. Doherty, and H. Mabuchi, Phys. Rev. Lett. **91**:250801 (2003)
16. R. van Handel, J, Stockton, and H. Mabuchi, IEEE Trans. Automat. Control **50** (6), 768-780 (2005)
17. M.R. James, Risk-sensitive optimal control of quantum systems, Phys. Rev. A **69**: 032108, (2004)
18. L. Bouten, M. Guţă, H. Maassen, J. Phys. A: Math. Gen., **37**: 3189, (2004)
19. L. Diósi, Phys. Lett. A. **132**, number 5, 233 (1988)
20. V.P. Belavkin and P. Staszewski, Phys. Lett. A, **140**, no. 7,8, 359-362 (1989)
21. P. Staszewski, *Quantum mechanics of continuously observed systems*, Nicholas Copernicus University Press (1993)

22. V.P. Belavkin, Commun.Math. Phys. **146**, 611-635 (1992)
23. J. Kupsch, I.-O. Stamatescu. in *Decoherence and the appearance of a classical world in quantum theory* Editors D. Guilini et al., Springer (1997)
24. A.J. Scott, G.J. Milburn, Phys. Rev. A, Vol. **63**, 042101 (2001)
25. J. Gough and A. Sobolev, Phys. Rev. A, **69**, 032107 (2004)
26. R. Bellman, *Dynamic Programming*, Princeton University Press (1957)
27. M.H.A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall Publishers (1977)
28. W.H. Louisell, *Quantum Statistical Properties of Radiation*, Wiley Classics Library (1990)
29. S.C. Edwards, V.P. Belavkin, quant-ph\0506018

FEEDBACK CONTROL OF QUANTUM SYSTEMS

MATTHEW R. JAMES*

Department of Engineering, Australian National University, Canberra, ACT 0200, Australia
E-mail: Matthew.James@anu.edu.au.

This paper presents an overview of some recent work in quantum feedback control, featuring optimal control, optimal risk-sensitive control, and robust control. We highlight the important pioneering contributions made by V.P. Belavkin in laying the foundations for quantum feedback control, discuss its importance, and give a little history.

1. Introduction

It is both a pleasure and a privilege to participate in this conference *Quantum Probability, Information, and Control 2006* in honor of the occasion of the 60th birthday of Slava Belavkin. Interestingly, this celebration is occurring at an important juncture in the history of technology, where we are witnessing the birth of a new generation of technologies that exploit quantum mechanics. Slava Belavkin, over a period of more than 30 years has shown considerable foresight by focusing his research efforts on laying foundations for this new generation of technologies,[1–4,6,7]. His research contributions in quantum filtering and feedback control are fundamental and have far reaching implications.

Feedback control is a critical enabler for technological development. From its origins in steam engine governors, through applications in electronics, aerospace, robotics, telecommunications and elsewhere, the use of feedback control has been essential in shaping our modern world. In the 20th century, quantum technology, through semiconductor physics and microchips, made possible the information age. New developments in quantum technology, which include quantum information and computing, precise metrology, atom lasers, and quantum electromechanical systems, further ex-

*Research supported by the Australian Research Council

ploit quantum phenomena and hold significant promise for the future. Slava Belavkin, to the best of my knowledge, was the first person to formulate and solve feedback control and filtering problems for systems governed by the laws of quantum physics. These results provide a crucially important conceptual framework and mathematical tools that will assist in the creation of these new quantum technologies.

In this paper I briefly describe some aspects of my recent work that have been influenced by Slava Belavkin's research. Before doing so, I say a few words about what quantum control is, why it is important, and its history.

2. What is Quantum Control?

2.1. *From Classical to Quantum*

In the 18th century, James Watt used mechanical governors to regulate the speed of his steam engines, Figure 1.
This is a classic example of feedback control which is used to maintain engine speed at a desired value. The governor senses any deviation in actual engine speed from the desired value, and makes adjustments to either increase or decrease speed, as appropriate. Feedback control was essential for the stable operation of the steam engines, and was a critical enabling technology for these machines which powered the industrial revolution. Steam engines, are of course macroscopic systems described by classical physics, and control engineering has been founded on classical models.

At this point in time, it is beginning to be possible to monitor and manipulate objects at the nanoscale. One can realistically contemplate controlling single atoms, for example, Figure 2.

At this scale, the laws of quantum physics are needed, and in fact provide a significant new resource for technological exploitation. This is where Slava Belavkin's work on quantum filtering and feedback control will have a significant impact. While it is difficult to speculate at this juncture of the enormous possibilities ahead of us, the currently envisaged quantum technologies, including those mentioned earlier, demonstrate considerable potential for quantum technology.

Quantum control, then, is the control of physical systems whose behavior is dominated by the laws of quantum physics. The importance of quantum control was highlighted by Dowling and Milburn in 2003[12]:

> "The development of the general principles of quantum control theory is an essential task for a future quantum technology."

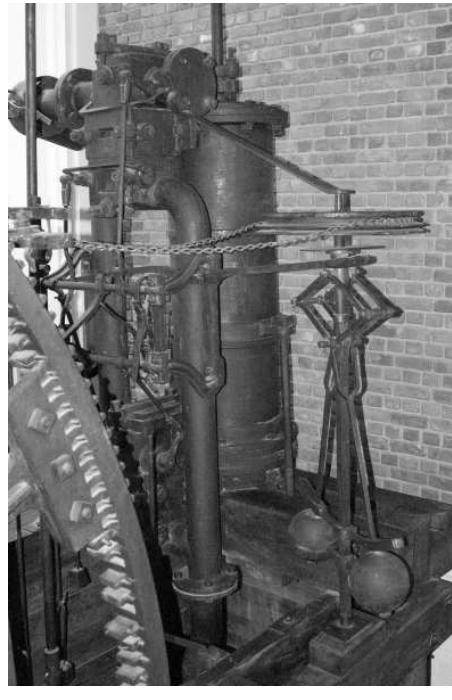


Fig. 1. Boulton and Watt steam engine, 1788, showing the mechanical governor (metal ball mechanism) [located at London Science Museum].

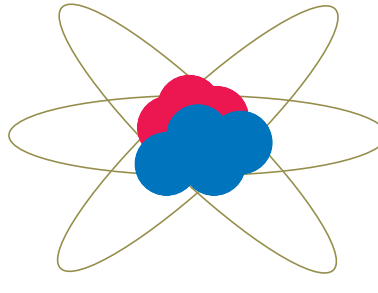


Fig. 2. Model of an atom.

It was also foreseen by Feynman in his famous *Plenty of Room at the Bottom* lecture[13], when he said:

> "What I want to talk about is the problem of manipulating and *controlling* things on a small scale"(emphasis added).

The study of quantum control therefore has *practical value*, but it also has *fundamental value*, since it includes the study of measurement and classical-quantum interactions in feedback loops that were not considered in the early days of quantum mechanics.

2.2. *Types of Quantum Control*

In control engineering, the system to be controlled is often called the "plant" (certainly an appropriate descriptor for a steam engine), while the system responsible for effecting the desired control is called the "controller" (in the case of the steam engine, the controller is the mechanical governor). In *open loop control*, control actions are predetermined and no feedback is involved, Figure 3. There is a one-way flow of information from the controller to the quantum system being controlled.

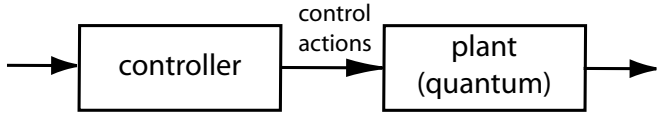


Fig. 3. Open loop control.

In *closed loop control*, control actions depend on information gained as the system is operating, Figure 4. Here, there is a two-way flow of information between the quantum system and the controller. Closed loop means feedback, just as used in Watt's steam engines.

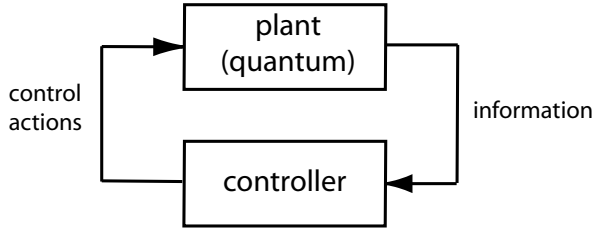


Fig. 4. Closed loop feedback control.

There are two types of quantum feedback. In quantum feedback with

measurement, the quantum system is monitored and the classical measurement results are used by the controller (which is a classical system, e.g. classical electronics) to determine the control signal applied to the quantum system, Figure 5.

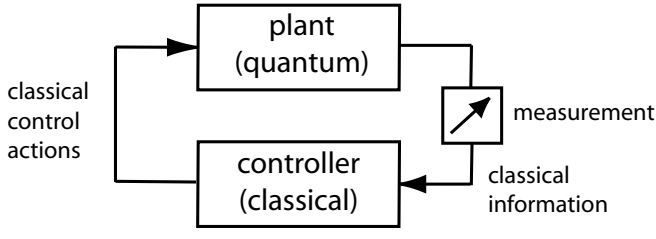


Fig. 5. Closed loop measurement feedback control.

It is also possible to use another quantum system as the controller, Figure 6. This type of feedback *does not use measurement*, and the information flowing in the loop is fully quantum. However, up to the present time most research in quantum feedback control has used measurement and classical controllers.

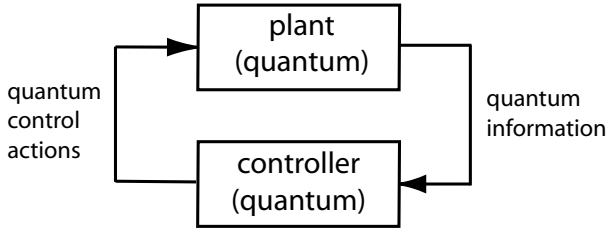


Fig. 6. Closed loop feedback control with no measurement.

2.3. *Some History*

To the best of my knowledge, Slava Belavkin was the first to publish results concerning quantum feedback control (late 1970's). However, there were independent pioneers in the physics community in the 1990s including Wiseman, Milburn, Doherty and Jacobs, who also made fundamental con-

tributions,[11,23]. A brief chronology of quantum feedback control is shown in Table 1.

Table 1. Quantum Feedback Chronology. These developments took place along with progress in modeling open quantum systems and measurement theory (not shown).

Late 1970's	Belavkin	linear Gaussian filtering and control
Early 1980's	Belavkin	optimal control using quantum operations
Late 1980's	Belavkin	optimal filtering and control using quantum stochastic differential equations
Early 1990's	Wiseman, Milburn	quantum optical measurement feedback
Late 1990's	Doherty, Jacobs, Mabuchi, *et al* Korotkov	LQG optimal control, experiments, solid state
2000's	Ahn, Belavkin, D' Helon, Doherty, Edwards, Gough, Bouten, van Handel, James, Kimura, Lloyd, Thomsen, Petersen, Schwab, Wiseman, Yanigisawa, Yamamoto and others	optimal control, Lyapunov control, robust control, applications, experiments

3. Optimal Control Using Quantum Langevin Equations

Quantum Langevin equations (QLE) provide a general framework for describing open quantum systems and contain considerable physical information. QLEs are expressed in terms of quantum stochastic differential equations (QSDE) are very well suited for control engineering. Quantum operation models arise naturally after suitable conditioning.

We consider the problem of controlling a quantum system or plant S (e.g. an atom) that is interacting with an environment, specifically, an electromagnetic field A. We suppose that the evolution of S can be influenced by control variables u that enter the Hamiltonian $H(u)$ for S. The field A is measured continuously in time, say by homodyne photodetection, and thanks to the interaction, the field contains information about the system S. The results $y(t)$, $0 \leq t \leq T$, of these measurements form a classical (i.e. non-quantum) signal that is available to the controller $\mathbf{K}$, a classical system which processes this information to produce the (classical) control actions $u(t)$, $0 \leq t \leq T$. The control in general is allowed to be a causal function of

the measurement trajectory. The problem we discuss is to find a controller $\mathbf{K}$ so that it minimizes a cost function $J(\mathbf{K})$ defined below.

3.1. *Model*

In order to present the quantum model for this problem (see[9,15,20]), we need to consider how observables of the system S evolve with time. These observables are represented as operators on a Hilbert space h. More generally we consider "system operators" X, which are linear operators on h, and we denote by $\mathscr{B}$ the von Neumann algebra of such operators. The electromagnetic field is represented by a family of operators $A(t) \in \mathscr{W}$, $0 \leq t \leq T$, on a Fock space F, where $\mathscr{W}$ denotes the von Neumann algebra of bounded linear operators on F. $A(t)$ models quantum white noise. States are normalized positive linear functionals that act on algebras of operators. Let ρ denote an arbitrary state on $\mathscr{B}$ and let ϕ be the vacuum state on $\mathscr{W}$. The total interacting system is described by the quantum probability space $(\mathscr{B} \otimes \mathscr{W}, \rho \otimes \phi)$. In what follows we will write $\mathbb{P} = \rho \otimes \phi$. In this model, system operators $X \in \mathscr{B}$ evolve in time according to

$$X(t) = j_t(X) = U^*(t)XU(t),$$

where the unitary $U(t)$ satisfies

$$dU_t = \Big\{LdA_t^* - L^*dA_t - \frac{1}{2}L^*Ldt - iH(u(t))dt\Big\}U_t, \qquad U_0 = I.$$

By Ito's rule, the QLE is)

$$dj_t(X) + j_t(\mathcal{L}^{u(t)}(X))dt + j_t([X,L])dB^*(t) + j_t([L^*,X])dB(t),$$

where

$$\mathcal{L}^u(X) = \frac{i}{\hbar}[H(u),X] + L^*XL - \frac{1}{2}(L^*LX + XL^*L), \tag{1}$$

and the commutator is defined as usual by $[A,B] = AB - BA$. Here, $L \in \mathscr{B}$ is a system operator, and we have denoted Hilbert space adjoints with the symbol *.

The measurements are described by

$$dY(t) = j_t(L + L^*)dt + dQ(t), \tag{2}$$

where $Q(t) = A(t) + A^*(t)$. The measurements generate a filtration of commutative von Neumann algebras $\mathscr{Y}_t = \mathrm{vN}\{Y_s,\ s \leq t\}$, so Y is equivalent to a classical process, which can be measured via homodyne detection. We also write $\mathscr{C}_t = \mathrm{vN}\{Z_s = A_s + A_s^*,\ s \leq t\}$, which is also commutative, and

$$\mathscr{Y}_t = U_t^*\mathscr{C}_tU_t.$$

Thanks to the unitary evolution, the non-demolition condition

$$[X, Z_s] = [X_t, Y_s] = 0, \quad 0 \leq s \leq t$$

holds, so that X belongs to the commutant of $\mathscr{C}_t$ and $j_t(X)$ belongs to the commutant of $\mathscr{Y}_t$. Thus conditional expectations

$$\mathbb{P}[X|\mathscr{C}_t], \quad \mathbb{P}[X_t|\mathscr{Y}_t]$$

are well-defined, and so filtering is possible.

With respect to the state $\mathbb{P}$, the process $Z_t = A_t + A_t^*$ is equivalent to a standard Wiener process. Furthermore, with respect to the state $\mathbb{P}_T^0$ defined by

$$\mathbb{P}_T^0[X] = \mathbb{P}[U_T X U_T^*], \ X \in \mathscr{B} \otimes \mathscr{W},$$

the measurement process Y_t is equivalent to a standard Wiener process (cf. *reference measure* in classical filtering).

A *controller* is a causal function from measurement data to control signals:

$$u(t) = \mathbf{K}(t, y_{[0,t]})$$

The coefficients of the stochastic differential equations are now adapted to the measurement filtration $\mathscr{Y}_t$. The filtering theory continues to apply,[8].

3.2. *Optimal Control*

The optimal control problem is specified by cost operators $C_1(u)$, C_2 (self-adjoint and non-negative), which define the performance integral

$$\int_0^T C_1(t)dt + C_2(T),$$

where $C_1(t) = j_t(u, C_1(u(t))$, $C_2(t) = j_t(u, C_2)$. We wish to minimize the expected cost

$$J(\mathbf{K}) = \mathbb{P}[\int_0^T C_1(t)dt + C_2(T)]$$

over all controllers $\mathbf{K}$.

To solve this problem, we express this cost in terms of filtered quantities. To this end, define $\tilde{U}_t$ by

$$d\tilde{U}_t = \Big\{ L(dA_t^* + dA_t) - \frac{1}{2} L^* L dt - iH(u(t))dt \Big\} \tilde{U}_t, \qquad \tilde{U}_0 = I.$$

Then $\tilde{U}_t^* X \tilde{U}_t$ belongs to the commutant of $\mathscr{C}_t$ (for $X \in \mathscr{B}$) and the unnormalized filter state

$$\sigma_t(X) = U_t^* \mathbb{P}[\tilde{U}_t^* X \tilde{U}_t | \mathscr{C}_t] U_t$$

is well-defined. It evolves according to

$$d\sigma_t(X) = \sigma_t\big(\mathcal{L}^{u(t)}(X)\big)dt + \sigma_t(L^*X + XL)dY_t.$$

This is the form of the *Belavkin quantum filter* we use,[5]. It is analogous to the *Duncan-Mortensen-Zakai* equation of classical filtering. σ_t is an *information state* (control theory terminology),[19].

We next derive the following representation in terms of σ_t:

$$\begin{aligned}
J(\mathbf{K}) &= \mathbb{P}[\int_0^T \tilde{U}_t^* C_1(u(t)) \tilde{U}_t dt + \tilde{U}_T^* C_2 \tilde{U}_T] \\
&= \mathbb{P}[\int_0^T \mathbb{P}[\tilde{U}_t^* C_1(u(t)) \tilde{U}_t | \mathscr{C}_t] dt + \mathbb{P}[\tilde{U}_T^* C_2 \tilde{U}_T | \mathscr{C}_T]] \\
&= \mathbb{P}[\int_0^T U_t U_t^* \mathbb{P}[\tilde{U}_t^* C_1(u(t)) \tilde{U}_t | \mathscr{C}_t] U_t U_t^* dt + U_T U_T^* \mathbb{P}[\tilde{U}_T^* C_2 \tilde{U}_T | \mathscr{C}_T] U_T U_T^*] \\
&= \mathbb{P}_T^0[\int_0^T \sigma_t(C_1(u(t)))dt + \sigma_T(C_2)].
\end{aligned}$$

This last expression is equivalent to classical expectation with respect to Wiener measure:

$$J(\mathbf{K}) = \mathbf{E}^0[\int_0^T \sigma_t(C_1(u(t)))dt + \sigma_T(C_2)].$$

The fact that the information state σ_t is computable from dynamics driven by the measured data means that the methods of dynamic programming are applicable. Define the *value function*

$$S(\sigma, t) = \inf_{\mathbf{K}} \mathbf{E}_{\sigma,t}^0[\int_t^T \sigma_s(C_1(u(s)))ds + \sigma_T(C_2)]$$

which quantifies the optimal cost to go from a current state σ at time t. The *dynamic programming principle* states that

$$S(\sigma, t) = \inf_{\mathbf{K}} \mathbf{E}_{\sigma,t}^0[\int_t^s \sigma_r(C_1(u(r)))dr + S(\sigma_s, s)]$$

At least formally, $S(\sigma, t)$ satisfies the *Hamilton-Jacobi-Bellman* equation

$$\begin{aligned}
\frac{\partial}{\partial t} S(\sigma, t) + \inf_{u \in \mathbf{U}} \{\mathscr{L}^u S(\sigma, t) + C_1(u)\} &= 0 \\
S(\sigma, T) &= \sigma(C_2)
\end{aligned}$$

where $\mathscr{L}^u$ is the Markov generator of σ_t (for fixed value $u \in \mathbf{U}$).

The optimal controller can be constructed as follows (verification theorem). Suppose we have a solution $S(\sigma, t)$ of the HJB equation. Define

$$\mathbf{u}^\star(\sigma, t) = \arg\min_u \{\mathscr{L}^u S(\sigma, t) + C_1(u)\}$$

This defines the optimal feedback controller:

$$\mathbf{K}^\star \ : \ \begin{cases} d\sigma_t(X) = \sigma_t\big(\mathcal{L}^{u(t)}(X)\big)dt + \sigma_t(L^*X + XL)dY_t \\ u(t) \quad\quad = \mathbf{u}^\star(\sigma_t, t) \end{cases}$$

This controller has the *separation structure*, with filter dynamics the Belavkin quantum filter for the information state σ_t.

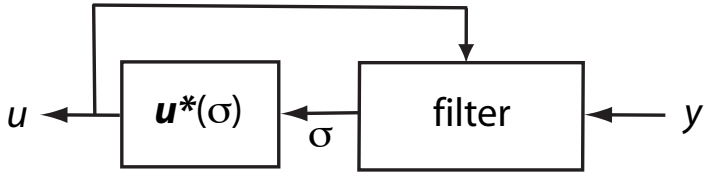


Fig. 7. Separation structure of the optimal controller.

3.3. *Another Type of Cost - Risk-Sensitive*

The cost function considered above is additive in nature, and generalizes the Linear Quadratic Gaussian (LQG) criterion of classical control theory. It is also possible to consider multiplicative costs such as the risk-sensitive criterion

$$J^\mu(\mathbf{K}) = \mathbb{P}[R^*(T)e^{\mu C_2(T)}R(T)],$$

where $R(t)$ is the time-ordered exponential defined by

$$\frac{dR(t)}{dt} = \frac{\mu}{2}C_1(t)R(t), \quad R(0) = I,$$

i.e.

$$R(t) = T\exp\left(\frac{\mu}{2}\int_0^t C_1(s)ds\right).$$

Here, $\mu > 0$ is a risk parameter. This generalizes the classical Linear Exponential Quadratic Gaussian (LEQG) cost function

$$J^\mu(\mathbf{K}) = \mathbf{E}[\exp\{\mu(\int_0^t C_1(u(t), t)dt + C_2(T))\}],$$

see[16,21].

In general, it does not appear possible to solve this problem using the unnormalized conditional state σ_t. Accordingly, we introduce a *risk-sensitive information state*

$$\sigma_t^\mu(X) = U_t^* \mathbb{P}[\tilde{V}_t^* X \tilde{V}_t | \mathscr{C}_t] U_t, \ X \in \mathscr{B},$$

where $\tilde{V}_t$ in the commutant of $\mathscr{C}_t$ is given by

$$d\tilde{V}_t = \{L(dA_t^* + dA_t) - \frac{1}{2}L^*L - \frac{i}{\hbar}H + \frac{\mu}{2}C_1(u(t))\}\tilde{V}_t.$$

We then have the representation

$$J^\mu(\mathbf{K}) = \mathbb{P}_T^0[\sigma_T^\mu(e^{\mu C_2})],$$

which facilitates dynamic programming.

The Hamilton-Jacobi-Bellman equation for this problem is

$$\frac{\partial}{\partial t}S^\mu(\sigma,t) + \inf_{u\in\mathbf{U}}\{\mathscr{L}^{\mu,u}S^\mu(\sigma,t)\} = 0$$
$$S^\mu(\sigma,T) = \sigma(e^{\mu C_2})$$

where $\mathscr{L}^{\mu,u}$ is the Markov generator of σ_t^μ (for fixed value $u \in \mathbf{U}$). The optimal risk-sensitive controller can be obtained as follows. Suppose we have a solution $S^\mu(\sigma,t)$ of the risk-sensitive HJB equation. Define

$$\mathbf{u}^{\mu,\star}(\sigma,t) = \arg\min_u\{\mathscr{L}^{\mu,u}S^\mu(\sigma,t)\}$$

Then the optimal risk-sensitive feedback controller is given by

$$\mathbf{K}^{\mu,\star} \ : \ \begin{cases} d\sigma_t^\mu(X) = \sigma_t^\mu\big(\mathcal{L}^{\mu,u(t)}(X)\big)dt + \sigma_t^\mu(L^*X + XL)dY_t \\ u(t) \qquad\quad = \mathbf{u}^{\mu,\star}(\sigma_t^\mu,t) \end{cases}$$

where

$$\mathcal{L}^{\mu,u}(X) = \mathcal{L}^u(X) + \frac{\mu}{2}(C_1(u)X + XC_1(u))$$

Note the inclusion of the cost observable in the modified Lindblad. This controller also has the *separation structure*, with filter dynamics the *modified* Belavkin quantum filter for the risk-sensitive information state σ_t^μ.

The risk-sensitive problem is of interest for at least two reasons:

(1) *Robustness.* Risk-sensitive controllers have the practical benefit that they can cope with uncertainty better than standard, see Figure 8.
(2) *Fundamentals of quantum mechanics.* The risk-sensitive information state can be viewed as a subjective state that includes knowledge and purpose, extending the Copenhagen interpretation in the feedback context.

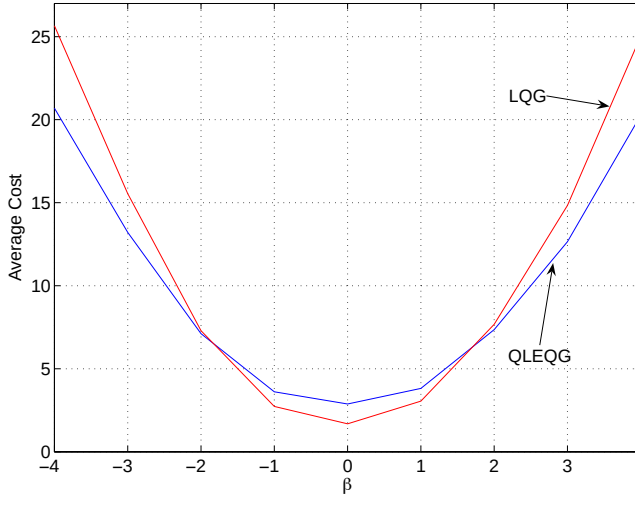


Fig. 8. Performance (vertical axis) of the LQG (standard) and LEQG (risk-sensitive) controllers under uncertainty,[22]. The vertical axis shows the integral of quadratic costs function averaged with respect to the true model. The horizontal axis is the uncertainty parameter β. When $\beta = 0$, the nominal and actual models coincide, and the QLQG has lower cost, as expected (QLQG is defined to minimize this cost). However, as β increases, we see that the QLEQG controller achieves better performance, indicated by the curve with smaller concavity. The QLEQG controller's performance degrades less rapidly and QLQG with increasing uncertainty. This is consistent with expectations for a robust controller: good performance under nominal conditions, and acceptable performance in other than nominal conditions.

4. H^∞ Control for Linear Stochastic Quantum Systems

4.1. *Background*

"H^∞" refers to the Hardy space which provides the setting for a frequency domain approach to robust control system design (initiated by Zames,[25]). Robustness refers to the ability of a control system to tolerate uncertainty, noise and disturbances, to some extent at least. Feedback is fundamental to this, and in fact is the basic reason for feedback (without noise or uncertainty, open loop would suffice).

Even when individual components are stable, feedback interconnections need not be, see Figure 9, as they may contain active elements.

The *small gain theorem* asserts stability of the feedback loop if the loop

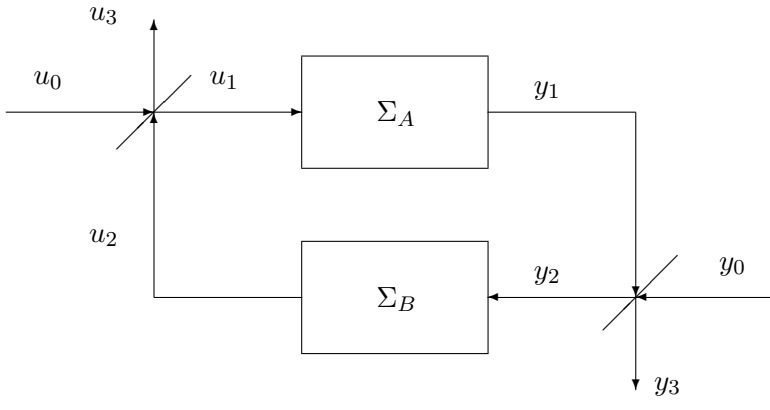


Fig. 9. A quantum feedback circuit.

gain is less than one:

$$g_A g_B < 1$$

see[10,24].

Stability is quantified in a mean-square sense as follows, Figure 10, and is captured mathematically by the inequality

$$\| \beta_y \|_t^2 \leq \mu + \lambda t + g^2 \| , \beta_u \|_t^2$$

where the "signals" are of the form

$$dU(t) = \beta_u(t)dt + dB_u(t)$$
$$dY(t) = \beta_y(t)dt + dB_y(t)$$

and the norm is defined by

$$\| \beta \|_t = \sqrt{\int_0^t \langle |\beta(s)|^2 \rangle ds}.$$

4.2. *The Problem*

Given a system (plant), find another system (controller) so that the gain from w to z is small. This is one way of reducing the effect of uncertainty or environmental influences, see[17,25].

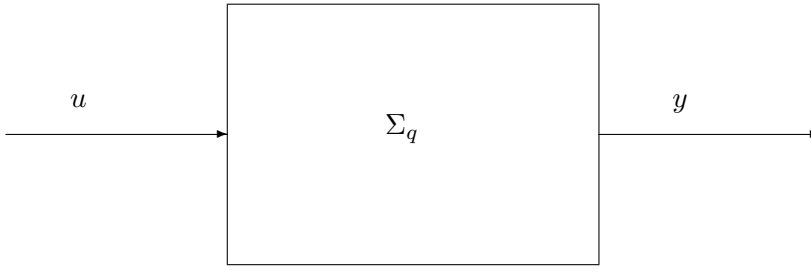


Fig. 10. A quantum system with inputs and outputs.

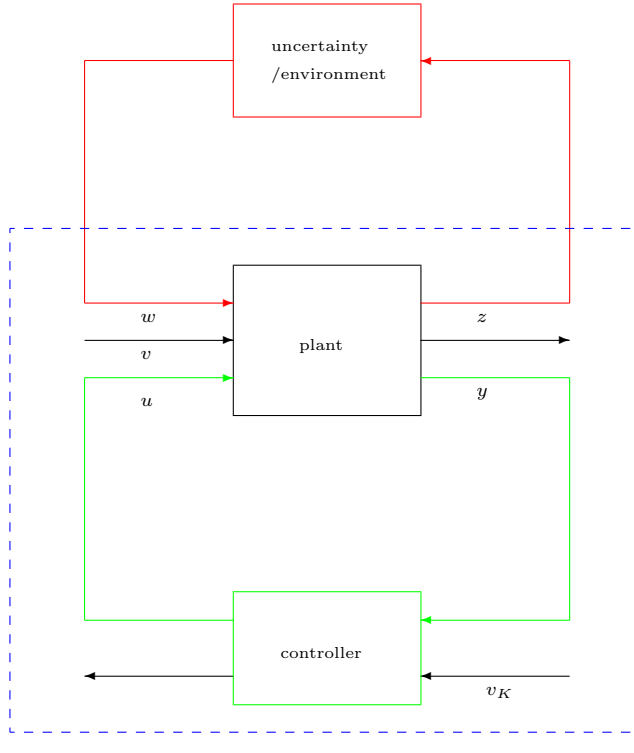


Fig. 11. A quantum system (the plant) in a feedback loop showing the controller and external influences.

The plant is described by the linear QLEs

$$dx(t) = Ax(t)dt + B_0dv(t) + B_1dw(t) + B_2du(t);\ x(0) = x;$$
$$dz(t) = C_1x(t)dt + D_{12}du(t);$$
$$dy(t) = C_2x(t)dt + D_{20}dv(t) + D_{21}dw(t),$$

as is the controller

$$d\xi(t) = A_K\xi(t)dt + B_{K1}dv_K(t) + B_K dy(t)$$
$$du(t) = C_K\xi(t)dt + B_{K0}dv_K(t),$$

where the coefficient matrices (A_K, etc) and the noise v_K are to be determined. The signals: w, u, v, z, y, v_K are semimartingales, e.g.

$$dw(t) = \beta_w(t)dt + d\tilde{w}(t),$$

where $\tilde{w}(t)$ is the noise which is assumed white Gaussian with Ito table

$$d\tilde{w}(t)d\tilde{w}^T(t) = F_{\tilde{w}}dt,$$

and $F_{\tilde{w}}$ is non-negative Hermitian (as in Belavkin's work).

Under some assumptions[17,18], then roughly speaking:

(1) If the closed loop system regarded as an operator $w \mapsto z$ has gain less than g then there exists solutions X and Y to the algebraic Riccati equations

$$(A - B_2E_1^{-1}D_{12}^TC_1)^TX + X(A - B_2E_1^{-1}D_{12}^TC_1)$$
$$+X(B_1B_1^T - g^2B_2E_1^{-1}B_2')X + g^{-2}C_1^T(I - D_{12}E_1^{-1}D_{12}^T)C_1 = 0;$$

$$(A - B_1D_{21}^TE_2^{-1}C_2)Y + Y(A - B_1D_{21}^TE_2^{-1}C_2)$$
$$+Y(g^{-2}C_1^TC_1 - C_2^TE_2^{-1}C_2)Y + B_1(I - D_{21}^TE_2^{-1}D_{21})B_1^T = 0.$$

satisfying stabilizability conditions and XY has spectral radius less than one.

(2) Conversely, if there exists solutions X, Y of these Riccati equations satisfying stabilizability conditions and XY has spectral radius less than one, then the controller defined by

$$A_K = A + B_2C_K - B_KC_2 + (B_1 - B_KD_{21})B_1^TX;$$
$$B_K = (I - YX)^{-1}(YC_2^T + B_1D_{21}^T)E_2^{-1};$$
$$C_K = -E_1^{-1}(g^2B_2^TX + D_{12}^TC_1).$$

and an arbitrary choice of v_K, B_{K1}, B_{K0}, achieves a closed loop with gain less than g.

Note: Physical realizability may impose conditions on the controller noise terms v_K, B_{K1}, B_{K0}, see[17].

4.3. *Examples*

In[17,18] the following example is considered. It is based on an optical cavity resonantly coupled to three optical channels v, w, u as in Figure 12.

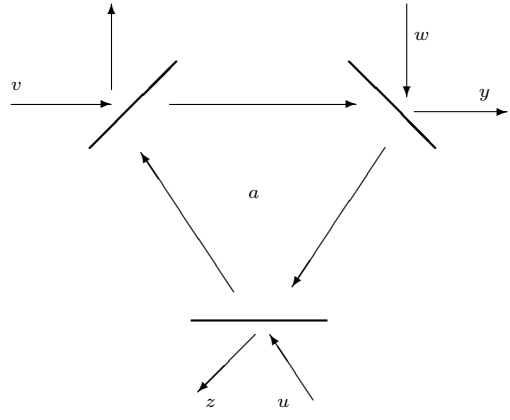


Fig. 12. An optical cavity (plant).

The annihilation operator a for this cavity system (representing a standing wave) evolves in time according to the equations

$$\begin{aligned} da &= -\frac{\gamma}{2}a\,dt - \sqrt{\kappa_1}\,dV - \sqrt{\kappa_2}\,dW - \sqrt{\kappa_3}\,dU \\ dZ &= \sqrt{\kappa_3}\,a\,dt + dU \\ dY &= \sqrt{\kappa_2}\,a\,dt + dW. \end{aligned} \tag{3}$$

It is required that $\gamma = \kappa_1+\kappa_2+\kappa_3$. In the quadrature notation of (3), $x_1(t) = q(t) = a(t) + a^*(t)$, $x_2(t) = p(t) = (a(t) - a^*(t)/i$, $V(t) = (v_1(t) + iv_2(t))/2$, $W(t) = (w_1(t) + iw_2(t))/2$, $U(t) = (u_1(t) + iu_2(t))/2$ and $\gamma = 3\kappa$. The quantum noises v, $\tilde{w}$ have Ito Hermitian matrices F_v, $F_{\tilde{w}}$ equal to

$$F_0 = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}. \tag{4}$$

This leads to a system of the form (3) with the following system matrices:

$$\begin{aligned} A &= -\frac{\gamma}{2}I; \\ B_0 &= \sqrt{\kappa_1}I; \quad B_1 = -\sqrt{\kappa_2}I; \quad B_2 = -\sqrt{\kappa_3}I; \\ C_1 &= \sqrt{\kappa_3}I; \quad D_{12} = I; \\ C_2 &= \sqrt{\kappa_2}I; \quad D_{21} = I. \end{aligned} \tag{5}$$

In this model, the boson commutation relation $[a, a^*] = 1$ holds (here, $[A, B] = AB - BA$).

The total cavity decay rate was chosen to be $\gamma = 3$ and the coupling coefficients $\kappa_1 = 2.6$, $\kappa_2 = 0.2$, $\kappa_3 = 0.2$. With a disturbance attenuation constant of $g = 0.1$, it was found that the Riccati equations (3) and (3) have stabilizing solutions satisfying the assumptions mentioned above; namely

$$X = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}; \quad Y = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

It follows from the main results that if a controller of the form (3) is applied to this system with matrices A_K, B_K, C_K defined as in (3) then the resulting closed loop system will be strictly bounded real with disturbance attenuation g. These matrices are given by

$$\begin{aligned} A_K &= \begin{bmatrix} -1.1 & 0 \\ 0 & -1.1 \end{bmatrix}; \quad B_K = -0.447I; \\ C_K &= -0.447I. \end{aligned} \tag{6}$$

The controller (3), (6) can be implemented with another optical cavity with annihilation operator a_K (with quadratures $\xi_1 = q_K = a_K + a_K^*$, $\xi_2 = p_K = (a_K - a_K^*)/i$, $\xi = (q_K, p_K)^T$) and a 180° phase shift. This system evolves according to the equations

$$\begin{aligned} da_K &= -\frac{\gamma_K}{2} a_K dt - \sqrt{\kappa_{K1}}\, dV_{K1} \\ &\quad -\sqrt{\kappa_{K2}}\, dV_{K2} - \sqrt{\kappa_{K3}}\, dY \\ dU &= -\sqrt{\kappa_{K1}}\, a_K dt - dV_{K1} \end{aligned} \tag{7}$$

where $\kappa_{K1} = 0.2$, $\kappa_{K2} = 1.8$, $\kappa_{K3} = 0.2$, and $\gamma_K = 2.2$. This gives

$$B_{K0} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix};$$

$$B_{K1} = \begin{bmatrix} -0.447 & 0 & -1.342 & 0 \\ 0 & -0.447 & 0 & -1.342 \end{bmatrix},$$

where

$$v_K(t) = \begin{bmatrix} v_{K11}(t) \\ v_{K12}(t) \\ v_{K21}(t) \\ v_{K22}(t) \end{bmatrix}$$

and $V_{K1} = (v_{K11} + iv_{K12})/2$, $V_{K2} = (v_{K21} + iv_{K22})/2$. The non-zero Ito products for this noise process are $dV_{K1}(t)dV_{K1}^*(t) = dt$ and

$dV_{K2}(t)dV_{K2}^*(t) = dt$, which defines the Hermitian matrix F_{V_K}. The closed loop system is illustrated in Figure 13.

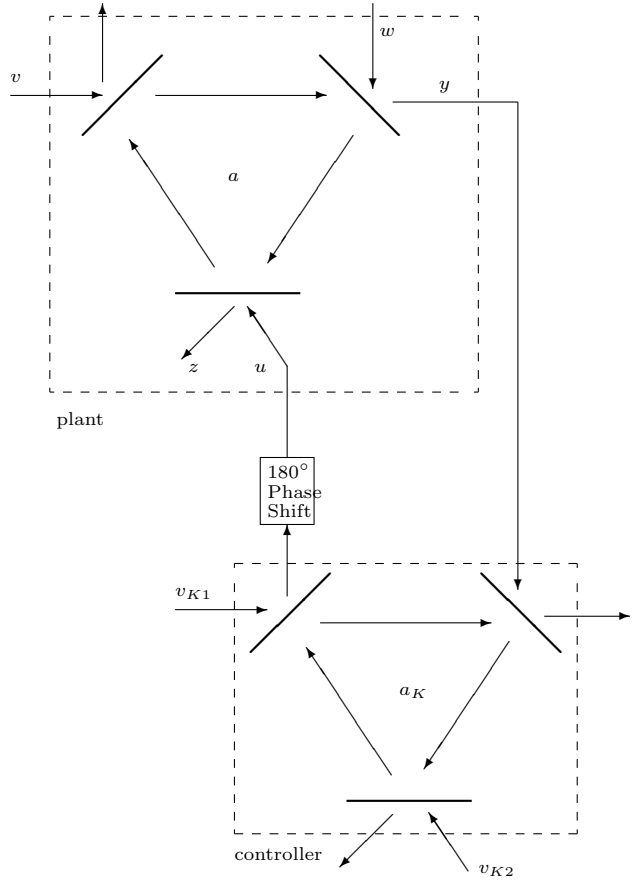


Fig. 13. An optical cavity (plant) controlled by another optical cavity (controller).

Note that additional quantum noises have been included in the controller construction. These are not determined by the Riccati equations, but instead they are required to ensure that a_K satisfies the boson commutation relation $[a_K, a_K^*] = 1$ (in terms of quadratures, $[q_K, p_K] = 2i$), a distinctively non-classical feature of quantum optical models, see[14].

It is also possible to construct classical controllers for this problem, when an appropriate measurement of the optical field y is made,[17]. More detailed

results concerning physical reaalization are provided in[17].

4.4. *Discussion*

These results provide the beginning of a robust control theory for quantum systems. The controllers themselves may be quantum or classical. It is important to note that the controller may need quantum noise inputs - this broadens the concept of controller, like randomization in classical optimal control.

5. Conclusion

We have sketched some recent work in quantum control that I have been involved with. Slava Belavkin's work was a crucial foundation. We discussed the practical and foundational importance of quantum control. Quantum feedback controllers may themselves be quantum, and may require additional quantum noise. There is an important and exciting future for "feedback control of quantum systems".

Bibliography

1. A. Barchielli and V.P. Belavkin. Measurements continuous in time and *a posteriori* states in quantum mechanics. *J. Phys. A: Math. Gen.*, 24:1495–1514, 1991.
2. V. P. Belavkin. Optimal measurement and control in quantum dynamical systems. preprint 411, Institute of Physics, Nicolaus Copernicus University, Torun, 1979.
3. V.P. Belavkin. On the theory of controlling observable quantum systems. *Automation and Remote Control*, 44(2):178–188, 1983.
4. V.P. Belavkin. Nondemolition measurements, nonlinear filtering, and dynamic programming of quantum stochastic processes. In A. Blaquiere, editor, *Modelling and Control of Systems in Engineering, Quantum Mechanics, Economics, and Biosciences*, pages 245–265, New York, 1988. Springer Verlag.
5. V.P. Belavkin. Quantum continual measurements and a posteriori collapse on CCR. *Commun. Math. Phys.*, 146:611–635, 1992.
6. V.P. Belavkin. Quantum stochastic calculus and quantum nonlinear filtering. *J. Multivariate Analysis*, 42:171–201, 1992.
7. L. Bouten, S. Edwards, and V.P. Belavkin. Bellman equations for optimal feedback control of qubit states. *J. Phys. B: At. Mol. Opt. Phys.*, 38:151–160, 2005.
8. L. Bouten and R. Van Handel. On the separation principle of quantum control. math-ph/0511021 2005.
9. L. Bouten, R. Van Handel, and M.R. James. An introduction to quantum filtering. math.OC/06011741 2006.

10. C. D'Helon and M.R. James. Stability, gain, and robustness in quantum feedback networks. *Phys. Rev. A.*, 73:053803, 2006.
11. A.C. Doherty and K. Jacobs. Feedback-control of quantum systems using continuous state-estimation. *Phys. Rev. A*, 60:2700, 1999. quant-ph/9812004.
12. J.P. Dowling and G.J. Milburn. Quantum technology: The second quantum revolution. 2003.
13. R.P. Feynman. Plenty of room at the bottom. Pasadena, CA, December 1959. Americal Physical Society, California Institute of Technology.
14. C.W. Gardiner and P. Zoller. *Quantum Noise*. Springer, Berlin, 2000.
15. A.S. Holevo. *Statistical Structure of Quantum Theory*. Springer, Berlin, 2001.
16. M.R. James. A quantum Langevin formulation of risk-sensitive optimal control. *J. Optics B: Semiclassical and Quantum, special issue on quantum control*, to appear 2005.
17. M.R. James, H. Nurdin, and I.R. Petersen. H^∞ control of linear quantum systems. quant-ph/0703150 2007.
18. M.R. James and I.R. Petersen. H^∞ control of linear quantum systems. In *Proc. IEEE CDC*, 2006.
19. P.R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, Englewood Cliffs, NJ, 1986.
20. K.R. Parthasarathy. *An Introduction to Quantum Stochastic Calculus*. Birkhauser, Berlin, 1992.
21. P. Whittle. Risk-sensitive linear/ quadratic/ Gaussian control. *Advances in Applied Probability*, 13:764–777, 1981.
22. S.D. Wilson, C. D'Helon, A.C. Doherty, and M.R. James. Quantum risk-sensitive control. In *Proc. IEEE CDC*, December 2006.
23. H. Wiseman and G.J. Milburn. Quantum theory of field-quadrature measurements. *Phys. Rev. A*, 47(1):642–663, 1993.
24. G. Zames. On the input-output stability of time-varying nonlinear feedback systems Part I: Conditions derived using concepts of loop gain, conicity, and positivity. *IEEE Trans. Automat. Control*, 11(2):228–238, 1966.
25. G. Zames and B. Francis. Feedback, minimax sensitivity, and optimal robustness. *IEEE Trans. Automat. Contr.*, AC-28:585–601, 1983.

LOCAL ASYMPTOTIC NORMALITY AND OPTIMAL ESTIMATION FOR D-DIMENSIONAL QUANTUM SYSTEMS

JONAS KAHN

Université Paris-Sud 11, Département de Mathématiques, Bât 425, 91405 Orsay Cedex, France
E-mail: jokahn@clipper.ens.fr

MĂDĂLIN GUŢĂ

School of Mathematics , University of Nottingham, University Park, Nottingham, NG7 2RD, U.K.
E-mail: madalin.guta@nottingham.ac.uk

We give a short account of local asymptotic normality for d-dimensional quantum systems. As application, we present a two step measurement procedure which is asymptotically optimal for state estimation.
Roughly speaking local asymptotic normality means that the family $\rho_{\theta/\sqrt{n}}^{\otimes n}$ of joint states of n identically prepared quantum systems approaches in a statistical sense a family of Gaussian state ϕ_θ of an algebra of canonical commutation relations. The convergence holds for all "local parameters" $\theta = (u_i, \zeta_{jk})$ corresponding to changes in the eigenvalues and respectively off-diagonal elements of a fixed diagonal state ρ_0. The states of the limit model are products between a $(d-1)$-dimensional classical normal distribution $N(\overrightarrow{u}, V)$, and $d(d-1)/2$ independent Gaussian states $G(\zeta_{jk}, \beta_{jk})$ of quantum harmonic oscillators. The latter is obtained by displacing a thermal equilibrium state with inverse temperature β by $\zeta_{jk} \in \mathbb{C}$.

Keywords: local asymptotic normality, optimal state estimation

1. Introduction

Quantum statistics deals with problems of statistical inference arising in quantum mechanics. The first significant results in this area appeared in the seventies[1–3,9] and tackled issues such as quantum Cramér-Rao bounds for unbiased estimators, optimal estimation for families of states possessing

a group symmetry, estimation of Gaussian states, optimal discrimination between non-commuting states. A good deal of this pioneering work was done by Slava Belavkin[4–8] to whom we dedicate this paper on the occasion of his 60's birthday.

The more recent theoretical advances[10–13] are closely related to the rapid development of quantum information and quantum engineering, and are often accompanied by practical implementations[14–16]. In quantum optics a measurement method called quantum homodyne tomography[17–19] allows the estimation with arbitrary precision[20,21] of the state of a monochromatic beam of light, by repeatedly measuring a sufficiently large number of identically prepared beams[22–24].

An important topic in quantum statistics is optimal estimation of an unknown state using the results of measurements performed on n quantum systems, identically prepared in that state[25–34]. In the case of two dimensional systems, or qubits, the problem has been solved explicitly in the context of Bayesian inference with invariant priors and figure of merit (risk) based on the fidelity distance between states[33]. However the method used there does not work for more general priors, loss functions or higher dimensions. In the pointwise approach, Hayashi and Matsumoto[31] showed that the Holevo bound[9] for the variance of locally unbiased estimators can be achieved asymptotically, and provided a sequence of measurements with this property. Their results, building on earlier work[35,36], indicate for the first time the emergence of a Gaussian limit in the problem of optimal state estimation for qubits. The extension to d-dimensional case is analysed by Matsumoto in.[43]

In[37,38] we gave a precise formulation of this fact and showed that we deal with the quantum generalization of an important concept in mathematical statistics called *local asymptotic normality*. The idea of approximating a sequence of statistical models by a family of Gaussian distributions appeared in[39], and was fully developed by Le Cam[40] who coined the term "local asymptotic normality". Among the many applications we mention its role in asymptotic optimality theory and in proving the asymptotic normality of certain estimators such as the maximum likelihood estimator.

For qubits, local asymptotic normality means roughly the following[37,38]: for large n the model described by n identically prepared qubits is asymptotically equivalent to a model consisting of pairs of classical Gaussian random variables and Gaussian states of a quantum harmonic oscillator.

For a more precise formulation let us parametrize the qubit states by their Bloch vectors $\rho(\overrightarrow{r}) = \frac{1}{2}(\mathbf{1} + \overrightarrow{r}\,\overrightarrow{\sigma})$ where $\overrightarrow{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. The neighborhood of the state ρ_0 with $\overrightarrow{r_0} = (0, 0, 2\mu - 1)$ and $1/2 < \mu < 1$, is a three-dimensional ball parametrized by the deviation $u \in \mathbb{R}$ of diagonal elements and $\zeta \in \mathbb{C}$ of the off-diagonal ones

$$\rho_\theta = \begin{pmatrix} \mu + u & \zeta^* \\ \zeta & 1 - \mu - u \end{pmatrix}, \qquad \theta = (u, \zeta). \tag{1}$$

Consider now n identically prepared qubits whose individual states are in a neighborhood of ρ_0 of size $1/\sqrt{n}$, so that their joint state is $\rho_\theta^n := \left[\rho_{\theta/\sqrt{n}}\right]^{\otimes n}$. We would like to understand the structure of the family (statistical experiment)

$$\mathcal{Q}_n := \{\rho_\theta^n : \|\theta\| \le C\}, \tag{2}$$

as a whole, more precisely what is its asymptotic behavior as $n \to \infty$?

For this we consider a quantum harmonic oscillator with position and momentum operators $\mathbf{Q}$ and $\mathbf{P}$ acting on $L^2(\mathbb{R})$ and satisfying the commutation relations $[\mathbf{Q}, \mathbf{P}] = i\mathbf{1}$. We denote by $\{|n\rangle, n \ge 0\}$ the eigenbasis of the number operator and define the thermal equilibrium state at inverse temperature β

$$G(\beta) = (1 - e^{-\beta}) \sum_{k=0}^{\infty} e^{-k\beta} |k\rangle\langle k|, \qquad e^{-\beta} = \frac{1-\mu}{\mu},$$

which has centered Gaussian distributions for both $\mathbf{Q}$ and $\mathbf{P}$ with variance $1/(4\mu - 2) > 1/2$. We define a family of displaced thermal equilibrium states

$$G(\zeta, \beta) := D(\zeta/\sqrt{2\mu - 1})\, G(\beta)\, D(\zeta/\sqrt{2\mu - 1})^*, \tag{3}$$

where $D(\zeta) := \exp(\zeta a^* - \zeta a)$ is the unitary displacement operator with $\zeta \in \mathbb{C}$. Additionally we consider a classical *Gaussian shift* model consisting of the family of normal distributions $N(u, \mu(1-\mu))$ with unknown center u and fixed variance. The classical-quantum statistical experiment to which we alluded above is

$$\mathcal{R} := \{\phi_\theta := N(u, \mu(1-\mu)) \otimes G(\zeta, \beta) : \|\theta\| \le C\} \tag{4}$$

where the unknown parameters $\theta = (u, \zeta)$ are the same as those of $\mathcal{Q}_n$.

Theorem 1.1. *Let $\mathcal{Q}_n$ be the quantum statistical experiment* (2) *and let $\mathcal{R}$ be the classical-quantum experiment* (4)*. Then for each n there exist*

quantum channels (normalized completely positive maps)

$$T_n : M\left(\mathbb{C}^{2^n}\right) \to L^1(\mathbb{R}) \otimes \mathcal{T}(L^2(\mathbb{R})),$$
$$S_n : L^1(\mathbb{R}) \otimes \mathcal{T}(L^2(\mathbb{R})) \to M\left(\mathbb{C}^{2^n}\right),$$

with $\mathcal{T}(L^2(\mathbb{R}))$ *the trace-class operators, such that*

$$\lim_{n\to\infty} \sup_{\|\theta\|\leq C} \|\phi_\theta - T_n\left(\rho_\theta^n\right)\|_1 = 0,$$
$$\lim_{n\to\infty} \sup_{\|\theta\|\leq C} \|\rho_\theta^n - S_n\left(\phi_\theta\right)\|_1 = 0,$$

for an arbitrary constant $C > 0$.

The local asymptotic normality theorem show that from a statistical point of view the joint qubits states are asymptotically indistinguishable from the limit Gaussian system. A consequence of this insight is that one can design optimal state estimators, and even propose a realistic measurement set-up for this purpose[38]. The local nature of the result is not a limitation but rather the correct normalization of the parameters with $n \to \infty$. Indeed as n grows we have more information about the state and we can easily pin it down to a region of size slightly larger that $1/\sqrt{n}$ by performing rough measurements on a small proportion of the systems. In a second stage we can use more sophisticated techniques to estimate the state within the local neighborhood of the first level estimator, and it is here where we use results on local asymptotic normality.

We stress that the convergence of experiments is defined at the level of quantum statistical models, it looks at the asymptotic structure of the model and it is not tied to a particular estimation problem. A mathematical analysis of the theory of convergence of quantum statistical experiments has been initiated in[41] to which we refer for more details.

In this paper we give a short account of the extension of the previous results to d-dimensional systems and we refer to[42] for the rather lengthy details of the proofs. The local neighborhood of a d-dimensional state can be parametrized in a similar fashion to (1). We need a (d$-$1)-dimensional real vector $\overrightarrow{u}$ for changes in the diagonal elements, and $d(d-1)/2$ complex variables $\zeta_{j,k}$ for the off-diagonal elements, with $j < k$. The limit experiment will consist of pairs of $(d-1)$-dimensional classical Gaussian variables $N(\overrightarrow{u}, V(\mu))$ with center $\overrightarrow{u}$, and a set of quantum Gaussian states $G(\zeta_{j,k}, \beta_{j,k})$, one for each off-diagonal matrix entry, and independent of each

other. An important technical issue relevant for optimal estimation, is that the local asymptotic normality holds in a "growing local neighborhood" $\|\theta\| \leq n^{\eta}$ where $\eta < 1/6$ (see Theorem 4.1) rather than the fixed ball $\|\theta\| < C$ as in Theorem 1.1.

The paper is structured as follows. In section 2 we familiarize with the concepts of classical and quantum statistical experiments, and learn how to compare their "statistical information content". In section 3 we explain the key concept of local asymptotic normality and its use in mathematical statistics. Its quantum version for d-dimensional systems is discussed in section 4, and in section 5 we derive an asymptotically optimal estimation procedure for completely unknown states with loss function equal to the square of the L^2 distance between states.

2. Classical and quantum statistical experiments

Suppose we are given some random data X with probability distribution P over the measure space $(\mathcal{X}, \Sigma_{\mathcal{X}})$, and we know that the probability distribution belongs to some family $\{P_\theta : \theta \in \Theta\}$ where the parameter θ is unknown. Statistical inference deals with the question of how to use the available data X in order to draw conclusions about some properties of θ. We will call the family

$$\mathcal{E} := (P_\theta : \theta \in \Theta), \tag{5}$$

a statistical experiment or model over $(\mathcal{X}, \Sigma_{\mathcal{X}})$[40].

In quantum statistics the data is replaced by a quantum system prepared in a state ϕ which belongs to a family $\{\phi_\theta : \theta \in \Theta\}$ of states over an algebra of observables. In order to make inference about θ one first has to measure the system, and then apply statistical techniques to draw conclusions from the data consisting of the measurement outcomes. An important difference with the classical case is that the experimenter has the possibility to choose the measurement set-up M, and each set-up will lead to a different classical model $\{P_\theta^{(M)} : \theta \in \Theta\}$, where $P_\theta^{(M)}$ is the distribution of outcomes when performing the measurement M on the system prepared in state ϕ_θ. Typically, the posterior state of the quantum system does not carry any more information about θ, hence it is important to choose beforehand the most informative measurement for each particular decision problem.

The guiding idea of this paper is to investigate the structure of the family

of quantum states

$$\mathcal{Q} := (\phi_\theta : \theta \in \Theta),$$

which will be called a *quantum statistical experiment*. We will show that in an important asymptotic set-up, namely that of a large number of identically prepared systems, the joint state can be approximated by a multidimensional quantum Gaussian state, for *all* possible preparations of the individual systems. This will bring a drastic simplification in the problem of optimal estimation for d-dimensional quantum systems, which will then be solved in the asymptotic framework.

2.1. *Classical and quantum randomizations*

Any procedure for data processing has a corresponding *Markov kernel*. Suppose we are given a random variable X taking values in $(\mathcal{X}, \Sigma_\mathcal{X})$ and we want to produce a "decision" $y \in \mathcal{Y}$ based on the data X. The space $\mathcal{Y}$ may be for example the parameter space Θ in the case of estimation, or just the set $\{0,1\}$ in the case of testing between two hypotheses. For every value $x \in \mathcal{X}$ we choose y randomly with probability distribution given by $K_x(dy)$. Assuming that $K : \mathcal{X} \times \Sigma_\mathcal{Y} \to [0,1]$ is measurable with respect to x for all fixed $A \in \Sigma_\mathcal{Y}$, we can regard K as a map from probability distributions over $(\mathcal{X}, \Sigma_\mathcal{X})$ to probability distributions over $(\mathcal{Y}, \Sigma_\mathcal{Y})$ with

$$K(P)(A) = \int K_x(A) P(dx), \quad A \in \Sigma_\mathcal{Y}. \tag{6}$$

A *statistic* $S : \mathcal{X} \to \mathcal{Y}$ is a particular example of such a procedure, where K_x is simply the delta measure at $S(x)$. Another particular case is that of a *randomized statistic* $R : \mathcal{X} \times [0,1] \to \mathcal{Y}$ where an additional independent variable U uniformly distributed over $[0,1]$ is used to compute the decision $R(X,U)$.

Besides taking a decision, there is another important reason why one would like to apply such treatment to the data, namely to summarize it in a more convenient and informative way for future purposes as illustrated in the following simple example. Consider n independent identically distributed random variables $X_1, \ldots, X_n$ with values in $\{0,1\}$ and distribution $P_\theta := (1-\theta, \theta)$ with $\theta \in \Theta := (0,1)$. The associated statistical experiment is

$$\mathcal{E}_n := (P_\theta^n : \theta \in \Theta).$$

It is easy to see that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is an unbiased estimator of θ and moreover it is a *sufficient statistic* for $\mathcal{E}_n$, i.e. the conditional distribution

$P_\theta^n(\cdot|\bar{X}_n = \bar{x})$ does not depend on θ! In other words the dependence on θ of the total sample $(X_1, X_2, \ldots, X_n)$ is completely captured by the statistic $\bar{X}_n$ which can be used as such for any statistical decision problem concerning $\mathcal{E}_n$. If we denote by $\bar{P}_\theta^n$ the distribution of $\bar{X}_n$ then the experiment

$$\bar{\mathcal{E}}_n = (\bar{P}_\theta^n : \theta \in \Theta),$$

is statistically equivalent to $\mathcal{E}_n$. To convince ourselves that $\bar{X}_n$ does contain the same statistical information as $(X_1, \ldots, X_n)$, we show that we can obtain the latter from the former by means of a randomized statistic. Indeed for every fixed value $\bar{x}$ of $\bar{X}_n$ there exists a measurable function

$$f_{\bar{x}} : [0,1] \to \{0,1\}^n,$$

such that the distribution of $f_{\bar{x}}(U)$ is $P_\theta^n(\cdot|\bar{X}_n = \bar{x})$. In other words

$$\lambda(f_{\bar{x}}^{-1}(x_1, \ldots, x_n)) = P_\theta^n(x_1, \ldots, x_n|\bar{X}_n = \bar{x}),$$

where λ is the Lebesgue measure on $[0,1]$. Then $F(\bar{X}_n, U) := f_{\bar{X}_n}(U)$, has distribution P_θ^n.

To summarize, statistics, randomized statistics and Markov kernels, are ways to transform the available data for a specific purpose. The Markov kernel K defined in (6) maps the experiment $\mathcal{E}$ of equation (5) into the experiment

$$\mathcal{F} := \{Q_\theta : \theta \in \Theta\},$$

over $(\mathcal{Y}, \Sigma_\mathcal{Y})$ with $Q_\theta = K(P_\theta)$. For mathematical convenience it is useful to represent such transformations in terms of linear maps between linear spaces.

Definition 2.1. A positive linear map

$$T_* : L^1(\mathcal{X}, \Sigma_\mathcal{X}, P) \to L^1(\mathcal{Y}, \Sigma_\mathcal{Y}, Q)$$

is called a *stochastic operator* or *transition* if $\|T_*(g)\|_1 = \|g\|_1$ for every $g \in L^1_+(\mathcal{X})$.

Definition 2.2. A positive linear map

$$T : L^\infty(\mathcal{Y}, \Sigma_\mathcal{Y}, Q) \to L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P)$$

is called a *Markov operator* if $T\mathbf{1} = \mathbf{1}$, and if for any $f_n \downarrow 0$ in $L^\infty(\mathcal{Y})$ we have $Tf_n \downarrow 0$.

A pair (T_*, T) as above is called a dual pair if

$$\int fT(g)dP = \int T_*(f)gdQ,$$

for all $f \in L^1(\mathcal{X})$ and $g \in L^\infty(\mathcal{Y})$. It is a theorem that for any stochastic operator T there exists a unique dual Markov operator T and vice versa.

What is the relation between Markov operators and Markov kernels ? Roughly speaking, any Markov kernel defines a Markov operator when we restrict to families of dominated probability measures (see Theorem 2.1). Let us assume that all distributions P_θ of the experiment $\mathcal{E}$ defined in (5) are absolutely continuous with respect to a fixed probability distribution P, such that there exist densities $p_\theta := dP_\theta/dP : \mathcal{X} \to \mathbb{R}_+$. Such an experiment is called *dominated* and in concrete situations this condition is usually satisfied. Moreover, P can be chosen to be a countable convex combination of P_θ's and such that $\{P_\theta\} \sim P$, i.e. for any $A \in \Sigma$, $P(A) = 0$ if and only if $P_\theta(A) = 0$ for all θ (see Lemma 20.3 in[44]).

Theorem 2.1. *Let $\mathcal{E} := (P_\theta : \theta \in \Theta)$ be an experiment over $(\mathcal{X}, \Sigma_\mathcal{X})$ and $\mathcal{F} := (Q_\theta : \theta \in \Theta)$ an experiment over $(\mathcal{Y}, \Sigma_\mathcal{Y})$ and suppose that both $\mathcal{E}, \mathcal{F}$ are dominated, with $\{P_\theta\} \sim P$ and $\{Q_\theta\} \sim Q$. Let K be the Markov kernel defined in* (6) *such that*

$$K(P_\theta) = Q_\theta.$$

Then there exists a Markov operator $T : L^\infty(\mathcal{Y}, \Sigma_\mathcal{Y}, Q) \to L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P)$ such that

$$Q_\theta = P_\theta \circ T, \qquad \forall \theta. \tag{7}$$

When the probability distributions of two experiments are related to each other as in (7), we say that $\mathcal{F}$ is a *randomization* of $\mathcal{E}$. From the duality between T and T_* we obtain an equivalent characterization in terms of the stochastic operator $T_* : L^1(\mathcal{X}, \Sigma_\mathcal{X}, P) \to L^1(\mathcal{Y}, \Sigma_\mathcal{Y}, Q)$ such that

$$T_*(dP_1^\theta/dP) = dQ^\theta/dQ, \qquad \forall \theta\,.$$

Hence, the previous theorem shows that the concept of randomization is weaker than that of Markov kernel transformation. Conversely, under the additional condition that $(\mathcal{Y}, \Sigma_\mathcal{Y})$ is locally compact space with countable base and Borel σ-field, it can be shown that any randomization can be implemented by a Markov kernel[44].

What is the analogue of randomizations in the quantum case ? In the language of operator algebras $L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P)$ is a commutative von Neumann algebra and $L^1(\mathcal{X}, \Sigma_\mathcal{X}, P)$ is the space of (densities of) *normal* linear functionals on it. The stochastic operator T_* is the classical version of *quantum channel*, i.e. a completely positive normalized (trace-preserving) map

$$T_* : \mathcal{A}_* \to \mathcal{B}_*$$

where $\mathcal{A}_*, \mathcal{B}_*$ are the spaces of normal states on the von Neumann algebra $\mathcal{A}$ and respectively $\mathcal{B}$. Any normal state ϕ on $\mathcal{A}$ has a density ρ with respect to the trace such that $\phi(A) = \mathrm{Tr}(\rho A)$ for all $A \in \mathcal{A}$. The dual of T_* is

$$T : \mathcal{B} \to \mathcal{A},$$

which is a unital completely positive map and has the property that $T(\phi)(b) = \phi(T(b))$ for all $b \in \mathcal{B}$ and $\phi \in \mathcal{A}_*$. We interpret such quantum channels as possible physical transformations from input to output states.

A particular class of channels is that of measurements. In this case the input is the state of a quantum system described by an algebra $\mathcal{A}$, and the output is a probability distribution over the space of outcomes $(\mathcal{X}, \Sigma_\mathcal{X})$. Any measurement is described by a positive linear map

$$M : L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P) \to \mathcal{A},$$

which is completely specified by the image of characteristic functions of measurable sets, also called *positive operator valued measure* (POVM). This map $M : \Sigma_\mathcal{X} \to \mathcal{A}$ has following properties

(1) Positive: $M(A) \geq 0, \qquad \forall\, A \in \Sigma_\mathcal{X}$;
(2) Countably additive: $\sum_{i=1}^\infty M(A_i) = M(\cup_i A_i), \quad A_i \cap A_j = \emptyset, i \neq j$;
(3) Normalized: $M(\mathcal{X}) = \mathbf{1}$.

The corresponding channel acting on states is a positive map $M_* : \mathcal{A}_* \to L^1(\mathcal{X}, \Sigma_\mathcal{X}, P)$ given by

$$M(\phi)(A) = \phi(M(A)) = \mathrm{Tr}(\rho M(A)),$$

where ρ is the density matrix of ϕ. By applying the channel M to the quantum statistical experiment consisting of the family of states $\mathcal{Q} = (\phi_\theta : \theta \in \Theta)$ on $\mathcal{A}$ we obtain a classical statistical experiment

$$\mathcal{Q}_M := \{M(\phi_\theta) : \theta \in \Theta\},$$

over the outcomes space $(\mathcal{X}, \Sigma_\mathcal{X})$.

As in the classical case, quantum channels can be seen as ways to compare quantum experiments. The first steps in this direction were made by Petz[45,46] who developed the theory of *quantum sufficiency* dealing with the problem of characterizing when a sub-algebra of observables contains the same statistical information about a family of states, as the original algebra. More generally, two experiments $\mathcal{Q} := (\mathcal{A}, \phi_\theta : \theta \in \Theta)$ and $\mathcal{R} := (\mathcal{B}, \psi_\theta : \theta \in \Theta)$ are called *statistically equivalent* if there exist channels $T : \mathcal{A} \to \mathcal{B}$ and $S : \mathcal{B} \to \mathcal{A}$ such that

$$\psi_\theta \circ T = \phi_\theta \qquad \text{and} \qquad \phi_\theta \circ S = \psi_\theta.$$

As consequence, for any measurement $M : L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P) \to \mathcal{A}$ there exists a measurement $T \circ M : L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P) \to \mathcal{A}$ such that the resulting classical experiments coincide $\mathcal{Q}_M = \mathcal{R}_{T\circ M}$. Thus for any statistical problem, and any procedure concerning the experiment $\mathcal{Q}$ there exists a procedure for $\mathcal{R}$ with the same risk (average cost), and vice versa.

2.2. *The Le Cam distance and its statistical meaning*

We have seen that two experiments are statistically equivalent when they can be transformed into each other be means of quantum channels. When this cannot be done exactly, we would like to have a measure of how close the two experiments are when we allow any channel transformation. We define the *deficiency* of $\mathcal{R}$ with respect to $\mathcal{Q}$ as

$$\delta(\mathcal{R}, \mathcal{Q}) = \inf_T \sup_\theta \|\phi_\theta - \psi_\theta \circ T\| \tag{8}$$

where the infimum is taken over all channels $T : \mathcal{A} \to \mathcal{B}$. The norm distance between two states on $\mathcal{A}$ is defined as

$$\|\phi_1 - \phi_2\| := \sup\{|\phi_1(a) - \phi_2(a)| : a \in \mathcal{A}, \|a\| \leq 1\},$$

and for $\mathcal{A} = \mathcal{B}(\mathcal{H})$ it is equal to $\mathrm{Tr}(|\rho_1 - \rho_2|)$, where ρ_i is the density matrix of the state ϕ_i. When $\delta(\mathcal{R}, \mathcal{Q}) = 0$ we say that $\mathcal{R}$ is more informative than $\mathcal{Q}$. Note that $\delta(\mathcal{R}, \mathcal{Q})$ is not symmetric but satisfies a triangle inequality of the form $\delta(\mathcal{R}, \mathcal{Q}) + \delta(\mathcal{Q}, \mathcal{T}) \geq \delta(\mathcal{R}, \mathcal{T})$. By symmetrizing we obtain a proper distance over the space of equivalence classes of experiments, called Le Cam's distance[40]

$$\Delta(\mathcal{Q}, \mathcal{R}) := \max\left(\delta(\mathcal{Q}, \mathcal{R}),\, \delta(\mathcal{R}, \mathcal{Q})\right).$$

What is the statistical meaning of the Le Cam distance ? We will show that if $\delta(\mathcal{R}, \mathcal{Q}) \leq \epsilon$ then for any statistical decision problem with loss function

between 0 and 1, any measurement procedure for $\mathcal{Q}$ can be matched by a measurement procedure for $\mathcal{R}$ whose risk will be at most ϵ larger than the previous one.

A decision problem is specified by a *decision space* $(\mathcal{X}, \Sigma_{\mathcal{X}})$ and a *loss function* $W_\theta : \mathcal{X} \to [0,1]$ for each $\theta \in \Theta$. We are given a quantum system prepared in the state $\phi_\theta \in \mathcal{A}_*$ with unknown parameter $\theta \in \Theta$ and would like to perform a measurement with outcomes in $\mathcal{X}$ such that the expected value of the loss function is small. Let

$$M : L^\infty(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{A},$$

be such a measurement, and $P_\theta^{(M)} = \phi_\theta \circ M$, then the *risk* at θ is

$$R(M, \theta) := \int_{\mathcal{X}} W_\theta(x) P_\theta^{(M)}(dx).$$

Since the point θ is unknown one would like to obtain a small risk over all possible realizations

$$R_{max}(M) = \sup_{\theta \in \Theta} R(M, \theta).$$

The *minimax risk* is then $R_{minmax} := \inf_M R_{max}(M)$. In the Bayesian framework one considers a prior distribution π over Θ and then averages the risk with respect to π

$$R_\pi(M) = \int_\Theta R(M, \theta) \pi(d\theta).$$

The optimal risk in this case is $R_\pi := \inf_M R_\pi(M)$.

Coming back to the experiments $\mathcal{Q}$ and $\mathcal{R}$ we will compare their achievable risks for a given decision problem as above. Consider the measurement $N : L^\infty(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{B}$ given by $N = T \circ M$ where $T : \mathcal{A} \to \mathcal{B}$ is the channel which achieves the infimum in (8). Then

$$\begin{aligned} R(N, \theta) &= \int_{\mathcal{X}} W(\theta, x) P_\theta^{(N)}(dx) = \psi_\theta(T \circ M(W_\theta)) \\ &\leq \|\psi_\theta \circ T - \phi_\theta\| + \phi_\theta(M(W_\theta)) \leq \delta(\mathcal{R}, \mathcal{Q}) + R(M, \theta), \end{aligned}$$

where we have used the fact that $0 \leq W_\theta \leq 1$.

Lemma 2.1. *For every achievable risk $R(M, \theta)$ for $\mathcal{Q}$ there exists a measurement $N : L^\infty(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{B}$ for $\mathcal{R}$ such that*

$$R(N, \theta) \leq R(M, \theta) + \delta(\mathcal{R}, \mathcal{Q}).$$

3. Local asymptotic normality in statistics

In this section we describe the notion of local asymptotic normality and its significance in statistics[40,44,47,48]. Suppose that we observe a sample $X_1, \ldots, X_n$ with X_i taking values in a measurable space $(\mathcal{X}, \Sigma_{\mathcal{X}})$ and assume that X_i are independent, identically distributed with distribution P_θ indexed by a parameter θ belonging to an open subset $\Theta \subset \mathbb{R}^m$. The full sample is a single observation from the product P_θ^n of n copies of P_θ on the sample space (Ω^n, Σ^n). Local asymptotic normality means that for large n such statistical experiments can be approximated by Gaussian experiments after a suitable reparametrization. Let θ_0 be a fixed point and define a local parameter $u = \sqrt{n}(\theta - \theta_0)$ characterizing points in a small neighborhood of θ_0, and rewrite P_θ^n as $P^n_{\theta_0+u/\sqrt{n}}$ seen as a distribution depending on the parameter u. Local asymptotic normality means that for large n the experiments

$$\left(P_{\theta_0+u/\sqrt{n}} : u \in \mathbb{R}^m\right) \qquad \text{and} \qquad \left(N(u, I_{\theta_0}^{-1}) : u \in \mathbb{R}^m\right),$$

have the same statistical properties when the models $\theta \mapsto P_\theta$ are sufficiently "smooth". The point of this result is that while the original experiment may be difficult to analyze, the limit one is a tractable *Gaussian shift* experiment in which we observe a single sample from the normal distribution with unknown mean u and fixed variance matrix $I_{\theta_0}^{-1}$. Here

$$[I_{\theta_0}]_{ij} = \mathbb{E}_{\theta_0}\left[\dot{\ell}_{\theta_0,i}\dot{\ell}_{\theta_0,j}\right],$$

is the Fisher information matrix at θ_0, with $\dot{\ell}_{\theta,i} := \partial \log p_\theta / \partial\theta_i$ the score function and p_θ is the density of P_θ with respect to some measure P.

There exist two formulations of the result depending on the notion of convergence which one uses. In this paper we only discuss the *strong* version based on convergence with respect to the Le Cam distance, and we refer to[48] for another formulation using the so called weak convergence (convergence in distribution of finite dimensional marginals of the likelihood ratio process), and to[41] for its generalization to quantum statistical experiments.

Before formulating the theorem we will explain what sufficiently smooth means. The least restrictive condition is that p_θ is *differentiable in quadratic mean*, i.e. there exists a measurable function $\ell_\theta : \mathcal{X} \to \mathbb{R}$ such that as $u \to 0$

$$\int \left[p_{\theta+u}^{1/2} - p_\theta^{1/2} - u^t \dot{\ell}_\theta p_\theta^{1/2}\right]^2 dP \to 0.$$

Note that $\dot{\ell}_\theta$ must still be interpreted as score function since under some regularity conditions we have $\partial p_\theta^{1/2}/\partial\theta_i = \frac{1}{2}(\partial \log p_\theta/\partial\theta_i)p_\theta^{1/2}$.

Theorem 3.1. *Let $\mathcal{E} := (P_\theta : \theta \in \Theta)$ be a statistical experiment with $\Theta \subset \mathbb{R}^d$ and $P_\theta \ll P$ such that the map $\theta \to p_\theta$ is differentiable in quadratic mean. Define*

$$\mathcal{E}_n = (P^n_{\theta_0+u/\sqrt{n}} : \|u\| \le C), \qquad \mathcal{F} = (N(u, I_0) : \|u\| \le C),$$

with I_0 the Fisher information matrix of $\mathcal{E}$ at point θ_0, and C a positive constant. Then $\Delta(\mathcal{E}_n, \mathcal{F}) \to 0$. In other words, there are sequences of randomizations T_n and S_n such that:

$$\lim_{n\to\infty} \sup_{\|u\|\le C} \left\| T_n(P^n_{\theta_0+u/\sqrt{n}}) - N(u, I_0) \right\| = 0$$

$$\lim_{n\to\infty} \sup_{\|u\|\le C} \left\| P^n_{\theta_0+u/\sqrt{n}} - S_n(N(u, I_0)) \right\| = 0.$$

Remark. Note that the statement of the Theorem is much more powerful than the Central Limit Theorem which shows convergence to a Gaussian distribution at a single point θ_0. Indeed local asymptotic normality states that the convergence is *uniform* around the point θ_0, and moreover the variance of the limit Gaussian is fixed whereas the variance obtained from the Central Limit Theorem depends on the point θ. Additionally, the randomization transforming the data $(X_1, \ldots, X_n)$ into the Gaussian variable is the same for all $\theta = \theta_0 + u/\sqrt{n}$ and thus does not require the a priori knowledge of θ.

Remark. Local asymptotic normality is the basis of many important results in asymptotic optimality theory and explains the asymptotic normality of certain estimators such as the maximum likelihood estimator. The quantum version introduced in the next section plays a similar role for the case of quantum statistical model. In[38] an asymptotically optimal estimation strategy for qubits was derived from a the qubit version of local asymptotic normality as presented below.

Example 3.1. Let $P_\mu = (\mu_1, \ldots, \mu_d)$ be a probability distribution with unknown parameters $(\mu_1, \ldots, \mu_{d-1}) \in \mathbb{R}_+^{d-1}$ satisfying $\mu_i > 0$ and $\sum_{i\le d-1} \mu_i < 1$. The Fisher information at a point μ is

$$I(\mu)_{ij} = \sum_{k=1}^{d-1} \mu_k(\delta_{ik}\mu_i^{-1} \cdot \delta_{jk}\mu_j^{-1}) + (1 - \sum_{l=1}^{d-1} \mu_l)^{-1} = \delta_{ij}\mu_i^{-1} + (1 - \sum_{l=1}^{d-1} \mu_l)^{-1},$$

and its inverse is

$$V(\mu)_{ij} := [I(\mu)^{-1}]_{ij} = \delta_{ij}\mu_i - \mu_i\mu_j. \tag{9}$$

Thus the limit experiment in this case is $\mathcal{F} := (N(u, V(\mu)) : u \in \mathbb{R}^{d-1}, \|u\| \leq C)$.

This experiment will appear again in Theorem 4.1, as the classical part of the limit Gaussian shift experiment. Let us consider as loss function the square of the ℓ^2 distance $\|\mu - \nu\|_2^2 = \sum_{i \leq d}(\mu_i - \nu_i)^2$, then in the limit experiment this corresponds to

$$W(u, v) = \sum_{i=1}^{d-1}(u_i - v_i)^2 + (\sum_{i=1}^{d-1}(u_i - v_i))^2.$$

The optimal estimator of u for this loss function is the data itself $\hat{u} := X \sim N(u, V(\mu))$ and the risk is independent of u

$$R = \sum_{i=1}^{d-1}\mu_i(1-\mu_i) + \sum_{i=1}^{d-1}\mu_i(1-\mu_i) - \sum_{1 \leq i \neq j \leq d-1}\mu_i\mu_j = \sum_{i=1}^{d}\mu_i(1-\mu_i), \tag{10}$$

where the last sum contains d terms and we used the fact that $\mu_d = 1 - \sum_{i \leq d-1}\mu_i$.

4. Local asymptotic normality in quantum statistics

In this section we will present the main result of the paper, that of local asymptotic normality for d-dimensional quantum systems, which means roughly the following: the sequence $\mathcal{Q}_n$ of experiments consisting of joint states φ^n of n identical quantum systems prepared independently in the same state φ, converges to a limit experiment $\mathcal{R}$ which is described by a family of Gaussian states on an algebra of canonical commutation relations.

Consider a d-dimensional quantum system whose state φ is described by its density matrix $\rho \in M(\mathbb{C}^d)$. The joint state of n identically prepared systems is given by $\rho^{\otimes n} \in M(\mathbb{C}^{d^n})$. As our theory will be local in nature, we first parametrize around one particular faithful state

$$\rho_0 = \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mu_d \end{bmatrix} \qquad \text{with } \mu_1 > \mu_2 > \cdots > \mu_d > 0, \tag{11}$$

which for technical reasons is chosen to have different eigenvalues. The states in a neighborhood of ρ_0 are of the form

$$\rho_\theta = \begin{bmatrix} \mu_1 + u_1 & \zeta_{1,2}^* & \dots & \zeta_{1,d}^* \\ \zeta_{1,2} & \mu_2 + u_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \zeta_{d-1,d}^* \\ \zeta_{1,d} & \dots & \zeta_{d-1,d} & \mu_d - \sum_{i=1}^{d-1} u_i \end{bmatrix} \qquad \text{with } u_i \in \mathbb{R}, \zeta_{j,k} \in \mathbb{C}. \tag{12}$$

We will investigate the properties of experiments

$$\mathcal{Q}_n := (\rho_{\theta/\sqrt{n}}^{\otimes n} : \theta \in \Theta_n)$$

consisting of n systems, each one prepared in a state $\rho_{\theta/\sqrt{n}}$ situated in a local neighborhood of ρ_0, as it was done in the classical case. The local parameter $\theta = (\overrightarrow{u}, \overrightarrow{\zeta})$ belongs to a neighborhood Θ_n of the origin of $\mathbb{R}^{d-1} \times \mathbb{C}^{d(d-1)/2}$ which is allowed to grow slowly with n in a way that will be made precise later. Before stating the main result we will study the limit of the sequence $\mathcal{Q}_n$.

4.1. *Quantum Gaussian shift experiment*

In this section we describe the limit experiment appearing in the local asymptotic normality Theorem 4.1. It contains a classical part described by a $(d-1)$-dimensional Gaussian shift experiment similar to the one appearing in Theorem 3.1, and a quantum part described by a $d(d-1)/2$-dimensional *quantum Gaussian shift experiment* which will be analyzed in more detail below. The classical part corresponds to changes in the diagonal parameters $\overrightarrow{u} = (u_1, \dots, u_{d-1})$ of ρ_θ. The quantum part is a product of Gaussian states of $d(d-1)/2$ quantum harmonic oscillators, the displacement of each state being related to one of the off-diagonal elements ζ_{jk} of ρ_θ.

Le us recall that a harmonic oscillator is described by the canonical observables $\mathbf{Q}$ and $\mathbf{P}$ acting on wavefunctions in $\mathcal{H} := L^2(\mathbb{R})$ as

$$(\mathbf{Q}\psi)(q) = q\psi(q), \qquad \text{and} \qquad (\mathbf{P}\psi)(q) = -i\frac{d\psi}{dq},$$

and satisfying the Heisenberg commutation relations $[\mathbf{Q}, \mathbf{P}] = i\mathbf{1}$. The *Fock basis* is a distinguished orthonormal basis in $\mathcal{H}$ consisting of eigenvectors $(|e_k\rangle : k \in \mathbb{N})$ of the number operator $\mathbf{N} := (\mathbf{Q}^2 + \mathbf{P}^2 - \mathbf{1})/2$. In quantum optics a monochromatic mode of light is described by a harmonic oscillator

with canonical variables the electric and magnetic fields. Alternatively to the density matrix ρ the states of light can be described through a variety of two-dimensional functions such as the Wigner function and the Husimi function[49]. The latter is given by

$$H_\rho(z) = \langle z|\rho|z\rangle,$$

where the vectors

$$|z\rangle = e^{-\|z\|^2/2} \sum_{k=0}^{\infty} \frac{z^k}{\sqrt{k!}} |e_k\rangle, \qquad z \in \mathbb{C},$$

describe the laser, or coherent states of intensity $|z|^2$. The Gaussian states of the harmonic oscillator are those states for which the distributions of all linear combinations of $\mathbf{Q}$ and $\mathbf{P}$ are Gaussian. We will be interested in displaced thermal equilibrium states $G(\zeta, \beta)$ for which the variance in any direction is the same and the corresponding Husimi functions are

$$H_{\zeta,\beta}(z) = \langle z|G(\zeta,\beta)|z\rangle := (1 - e^{-\beta}) \exp\left(-(1 - e^{-\beta})\|z - \zeta\|^2\right),$$

for any $\zeta \in \mathbb{C}$ and $0 < \beta \leq \infty$. Physically, the parameter β is an inverse temperature and $G(0, \beta)$ is a thermal equilibrium state. A one-dimensional quantum gaussian shift experiment is defined as a family of displaced thermal equilibrium states with known temperature and unknown displacement

$$\mathcal{R} = (G(\zeta, \beta) : \zeta \in \mathbb{C}).$$

This experiment is well known in quantum statistics and has been largely investigated[2,4,9]. In particular it is known that the optimal mean square estimation procedure is the coherent or heterodyne measurement, whose POVM elements are the rank one operators $\frac{1}{\pi}|z\rangle\langle z|$ yielding an outcome with distribution proportional to the Husimi function $H_{\zeta,\beta}(z)$! We are then back to a classical two-dimensional Gaussian shift experiment and the best estimator of ζ is the outcome z. If we consider the loss function $W(\zeta, z) = |\zeta - z|^2$ then the risk of the estimator is independent of ζ

$$R = \frac{1}{\pi} \int H_{0,\beta}(z)|z|^2 d^2z = (1 - e^{-\beta})^{-1}. \tag{13}$$

4.2. *The main theorem*

We are now ready to state the main result of the paper.

Theorem 4.1. *Let $0 \leq \eta < 1/6$ be fixed and define $\Theta_n \subset \mathbb{R}^{d-1} \times \mathbb{C}^{d(d-1)/2}$ to be the ball or radius n^η centered at the origin, where the norm of $\theta =$*

$(\overrightarrow{u}, \overrightarrow{\zeta})$ *is given by* $\|\theta\|^2 = \|\overrightarrow{u}\|^2 + \|\overrightarrow{\zeta}\|^2$. *Define the sequences of quantum statistical experiments*

$$\mathcal{Q}_n := (\rho_{\theta/\sqrt{n}}^{\otimes n} : \theta \in \Theta_n), \qquad \mathcal{R}_n := (\phi_\theta : \theta \in \Theta_n),$$

with ρ_θ *defined as in* (12) *and*

$$\phi_\theta := N(\overrightarrow{u}, V(\mu)) \otimes \bigotimes_{j<k} G\left(\frac{\zeta_{j,k}}{\sqrt{\mu_j - \mu_k}}, \ln\left(\frac{\mu_j}{\mu_k}\right)\right) \tag{14}$$

where $V(\mu)$ *is a the covariance matrix defined in* (9). *Then the sequences of experiments* $\mathcal{Q}_n$ *and* $\mathcal{R}_n$ *are asymptotically equivalent, more precisely there are channels* T_n *and* S_n *such that, for any* $\epsilon > 0$*:*

$$\begin{aligned} \sup_{\|\theta\| \le n^\eta} \left\| T_n(\rho_{\theta/\sqrt{n}}^{\otimes n}) - \phi_\theta \right\|_1 &= O(n^{-1/4+3\eta/2+\epsilon}) \to 0 \\ \sup_{\|\theta\| \le n^\eta} \left\| \rho_{\theta/\sqrt{n}}^{\otimes n} - S_n(\phi_\theta) \right\|_1 &= O(n^{-1/4+3\eta/2+\epsilon}) \to 0 \end{aligned} \tag{15}$$

Moreover the error term depends smoothly on the eigenvalues of ρ_0.

Remark. Asymptotically the sequence $\mathcal{Q}_n$ separates into two independent experiments, a classical one with unknown parameter $\overrightarrow{u} \in \mathbb{R}^{d-1}$ and a quantum one with unknown parameters $\overrightarrow{\zeta} \in \mathbb{C}^{d(d-1)/2}$. The fact that changes in diagonal elements are described asymptotically by a classical experiment is not surprising considering that the diagonal density matrices commute with each other. The variance of the limit Gaussian is precisely that of Example 3.1. On the other hand, for each off-diagonal element there exists a $SU(2)$ subgroup of rotations in a two-dimensional plane which converges to displacements of Gaussian states of a quantum oscillator as in the case of qubits[37,38].

Remark. The channels T_n and S_n can be constructed explicitly using the $SU(d)$ representation theory[42]. Specifically a highest weight vector of each representation is identified with the vacuum, and the actions of the different rising operators correspond *approximately* to that of creation operators of oscillators.

Remark. We allow the parameter space Θ_n to grow with n. Nevertheless notice that the neighborhood around ρ_0 is still shrinking as $n^{1/2}$ dominates $n^{1/6}$. The theorem is still local in nature. However this "large domain" is the stepping stone to prove global results as we will see in the next section.

5. Two-step strategy for state estimation

Though Theorem 4.1 is local, the fact that the domain $\|\theta\| \leq n^{\eta}$ $(0 < \eta < 1/6)$ increases with n, makes it almost as good as global. We will illustrate this with the problem of estimating the state $\rho \in M(\mathbb{C}^d)$ given n identically prepared copies of the system. Details of the construction in the case of qubits can be found in[38].

First step. Measure a small proportion of the n copies, say $n^{1-\epsilon}$ to produce an estimator $\tilde{\rho}_n$ of ρ. In this first step we have localized the state within a small region around $\tilde{\rho}_n$ of the size $n^{-1/2+\epsilon}$ with arbitrarily small $\epsilon > 0$[38]. We can then rotate the remaining $\tilde{n} := n - n^{1-\epsilon}$ states such that $\tilde{\rho}_n$ becomes diagonal and we denote $\rho_0 = \tilde{\rho}_n$. Since the true state ρ is in a neighborhood of size $n^{-1/2+\epsilon} < \tilde{n}^{-1/2}\tilde{n}^{1/6-\epsilon}$ of ρ_0, we can apply the results of Theorem 4.1 on local asymptotic normality around ρ_0. In particular, the state ρ can be identified with ρ_θ parametrized as in (12).

Second step. On the remaining $\tilde{n}$ systems we apply the channel $T_{\tilde{n}}$. Since asymptotically the quantum experiment $\mathcal{Q}_n$ can be identified with the Gaussian limit $\mathcal{R}_n$, we employ Lemma 2.1 to construct an optimal estimator for the state $\rho = \rho_\theta$ by performing the optimal measurement for estimating the local parameters $\theta = (\overrightarrow{u}, \overrightarrow{\zeta})$ on the limit experiment. A technical difficulty arises however due to the fact that the loss function in the case of estimating the displacement is not bounded as required in Lemma 2.1, thus one has to estimate precisely the different contributions in the total risk of the two step strategy.

5.1. *Risk of the procedure*

We consider a loss function equal to the square of the L^2 operator norm, which in our parametrization is given by

$$\|\rho_{\theta^a} - \rho_{\theta^b}\|_2^2 = \sum_{i=1}^{d-1}(u_i^a - u_i^b)^2 + \left(\sum_{i=1}^{d-1} u_i^a - u_i^b\right)^2 + 2\sum_{j<k}|\zeta_{j,k}^a - \zeta_{j,k}^b|^2. \quad (16)$$

The following inequality gives an upper bound to the risk at ρ in terms of

three contributions:

$$\mathbb{E}_\rho\left[\|\rho-\hat{\rho}_n\|_2^2\right]\leq$$
$$2P_\rho\left[\|\rho-\tilde{\rho}_n\|_2>n^{-1/2+\epsilon}\right]+$$
$$\mathbb{E}_\rho\left[\|T_n(\rho_{\theta/\sqrt{n}}^{\otimes n})-\phi_\theta\|_1\,\middle|\,\|\rho-\tilde{\rho}_n\|_2<n^{-1/2+\epsilon}\right]\cdot n^{-1}\cdot\sup_{\substack{\|\theta\|<n^\epsilon\\ \|\hat{\theta}_n\|<2n^\epsilon}}\|\theta-\hat{\theta}_n\|_2^2+$$
$$n^{-1}\cdot\mathbb{E}_{\phi_\theta}\left[\|\theta-\hat{\theta}_n\|_2^2\right].$$

Below we will explain the origin of each term on the right side of the inequality and provide an upper bound for it.

5.1.1. *Error from the first step measurement*

The first term corresponds to the risk of having the first rough estimate very far from the truth, in which case ρ could fall outside the validity region of local asymptotic normality. Since we consider a loss function equal to the square of the L^2 distance, the maximum error of the final estimator is 2 as both our estimate and the unknown ρ are states, so that $\|\rho-\hat{\rho}_n\|_2^2\leq 2$. Thus the expected loss associated to having a very bad first estimate is upper bounded by the double of the its probability. It is not difficult to show that there exists a simple rough estimator for which the probability of the event $\|\tilde{\rho}_n-\rho\|_2>n^{-1/2+\epsilon}$ is $o(n^{-1})$[38]. This can be understood from the fact that the average error is of order $n^{-1/2+\epsilon/2}\ll n^{-1/2+\epsilon}$, for $\epsilon>0$.

5.1.2. *Error from the use of local asymptotic normality*

The second term corresponds to the loss incurred by transforming the $\tilde{n}$ systems through the channel $T_{\tilde{n}}$. Since the difference between n and $\tilde{n}$ is negligible compared with n it will not play any role in the computations and we can safely revert to the notation n for the number of systems in the second step of the procedure.

We obtain the state $T_n(\rho_{\theta/\sqrt{n}}^{\otimes n})$ which is asymptotically close to ϕ_θ but does not coincide with it. The following simple inequality shows that we can replace $T_n(\rho_{\theta/\sqrt{n}}^{\otimes n})$ by ϕ_θ at the cost of an additional contribution to the risk. Let P,Q be two probability distributions with densities p,q and let

$0 \leq W \leq M$ be a loss function, then

$$\mathbb{E}_P[W] \leq \mathbb{E}_Q[W] + M\|p - q\|_1.$$

We can now condition on $\|\rho - \tilde{\rho}_n\|_2 \leq n^{-1/2+\epsilon}$ since the bigger errors have been taken into account through the first error term. Additionally, the estimator $\hat{\theta}_n$ can be chosen to be bounded by $2n^\epsilon$ since we know that the true parameter θ is bounded by n^ϵ. Hence the error of the estimator $\|\hat{\theta}_n - \theta\|_2$ is less than $3n^\epsilon$. Moreover, we can bound the norm distance between $T_n(\rho_{\theta/\sqrt{n}}^{\otimes n})$ and ϕ_θ using Theorem 4.1 with $\eta > \epsilon$. The bound (15) does not depend on $\tilde{\rho}_n$, as we know that $\tilde{\rho}_n$ is in a neighbourhood of ρ, so that its eigenvalues are bounded away from each other. Finally we get that the norm distance is less than $n^{-1/4+3/2\eta+\epsilon}$, hence the second error term is $O(n^{-1+2\epsilon-1/4+3/2\eta+\epsilon'}) = o(n^{-1})$ provided that $2\epsilon + 3/2\eta + \epsilon < 1/4$.

5.1.3. *Error from the optimal measurement in the limit experiment*

The third term is the risk of the optimal estimation procedure in the limit experiment and dominates the previous two error contributions.

Since the limit experiment separates into a classical and a quantum part and the loss function is the sum of contributions corresponding to the two sets of parameters, the risk is the sum of the risks for the classical and the quantum part. The quantum part splits further into a sum of risks for estimating each of the parameters ζ_{jk}. Let us recall that the optimal measurement of the displacement of a Gaussian state is the coherent measurement described in (13) and its risk is $R = (1 - e^{-\beta})^{-1}$. The risk of estimating ζ_{kj} based on the states (14) is

$$R_{jk} = \left(1 - e^{-\ln \frac{\mu_j}{\mu_k}}\right)^{-1} \cdot (\mu_j - \mu_k) = \mu_j.$$

The risk of the classical experiment has been calculated in (10)

$$R_{cl} = \sum_{i=1}^{d} \mu_i(1 - \mu_i).$$

The total risk is the sum of the classical and quantum parts according to the loss function (16)

$$R = \sum_{i=1}^{d} \mu_i(1 - \mu_i) + 2\sum_{j<k} \mu_j.$$

5.2. *Asymptotic minimax bound*

We have thus proven the following asymptotic result.

Theorem 5.1. *For any ρ with eigenvalues $0 < \mu_1 < \dots \mu_d < 1$ the risk of the estimator $\hat{\rho}_n$ satisfies*

$$\lim_{n\to\infty} n\,\mathbb{E}_\rho\|\rho - \hat{\rho}_n\|_2^2 = \sum_{i=1}^{d} \mu_i(1-\mu_i) + 2\sum_{j<k} \mu_j.$$

Moreover the sequence of estimators $\hat{\rho}_n$ is asymptotically optimal in the minimax sense.

Theorem 5.2. *For any sequence of estimation strategies $\hat{\sigma}_n$, and for any ρ with eigenvalues $0 < \mu_1 < \dots \mu_d < 1$ we have*

$$n \limsup_{n\to\infty} \sup_{\|\sigma-\rho\|<n^{-1/2+\epsilon}} \mathbb{E}_\sigma\|\hat{\sigma}_n - \sigma\|_2^2 \geq \sum_{i=1}^{d} \mu_i(1-\mu_i) + 2\sum_{j<k} \mu_j.$$

The factor n is inserted because typically the risk is of the order $1/n$ and the optimization is about obtaining the smallest constant factor possible. The inequality says that one cannot find an estimator which performs better that $\hat{\rho}_n$ over a ball of size $n^{-1/2+\epsilon}$ centered at ρ, even if one has the knowledge that the true state belongs to that ball! Thus the proposed estimator saturated the minimax bound

$$R_{minmax}(\rho) = \sum_{i=1}^{d} \mu_i(1-\mu_i) + 2\sum_{j<k} \mu_j.$$

The proof of this bound makes use of the other direction in local asymptotic normality, namely the inverse channel S_n. The idea is that if there was a better estimation strategy, then we could find a way of estimating modes of light better than the heterodyne measurement. The proof of a similar result for qubits can be found in.[38]

Bibliography

1. C. W. Helstrom, *Journal of Statistical Physics* **1**, 231 (1969).
2. H. P. Yuen and M. Lax, *IEEE Trans. Inform. Theory* **19**, p. 740 (1973).
3. H. Yuen, R. Kennedy and M. Lax, *IEEE Trans. Inform. Theory* **21**, 125 (1975).

4. V. P. Belavkin, *Radio Eng. Electron. Physics* **17**, 527 (1972).
5. V. P. Belavkin, *Problems of Information Transmission* **4**, 241 (1973).
6. V. P. Belavkin, *Radio Eng. Electron. Physics* **20**, 1177 (1975).
7. V. P. Belavkin, *Stochastics* **1**, 315 (1975).
8. V. P. Belavkin, *Theor. Math. Phys.* **26**, 213 (1976).
9. A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, 1982).
10. M. Hayashi (ed.), *Asymptotic theory of quantum statistical inference: selected papers* (World Scientific, 2005).
11. M. Hayashi, *Quantum Information* (Springer-Verlag, Berlin Heidelberg, 2006).
12. M. G. A. Paris and J. Řeháček (eds.), *Quantum State Estimation* 2004.
13. O. E. Barndorff-Nielsen, R. D. Gill and P. E. Jupp, *J. R. Statist. Soc. B* **65**, 775 (2003).
14. M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty and H. Mabuchi, *Phys. Rev. Lett.* **89**, p. 133602 (2002).
15. T. Hannemann, D. Reiss, C. Balzer, W. Neuhauser, P. E. Toschek and C. Wunderlich, *Phys. Rev. A* **65**, 050303 (2002).
16. G. A. Smith, A. Silberfarb, I. H. Deutsch and P. S. Jessen, *Phys. Rev. Lett.* **97**, 180403 (2006).
17. K. Vogel and H. Risken, *Phys. Rev. A* **40**, 2847 (1989).
18. G. M. D'Ariano, U. Leonhardt and H. Paul, *Phys. Rev. A* **52**, R1801 (1995).
19. U. Leonhardt, M. Munroe, T. Kiss, T. Richter and M. G. Raymer, *Optics Communications* **127**, 144 (1996).
20. L. Artiles, R. D. Gill, and M. Guţă, *J. Royal Statist. Soc. B (Methodological)* **67**, 109 (2005).
21. C. Butucea, M. Guţă and L. Artiles, Minimax and adaptive estimation of the Wigner function in quantum homodyne tomography with noisy data, *Ann. Statist.* **35**, 465 (2007).
22. D. T. Smithey, M. Beck, M. G. Raymer and A. Faridani, *Phys. Rev. Lett.* **70**, 1244 (1993).
23. S. Schiller, G. Breitenbach, S. F. Pereira, T. Müller and J. Mlynek, *Phys. Rev. Lett.* **77**, 2933 (1996).
24. A. Zavatta, S. Viciani and M. Bellini, *Science* **306**, 660 (2004).
25. S. Massar and S. Popescu, *Phys. Rev. Lett.* **74**, 1259 (1995).
26. J. I. Cirac, A. K. Ekert and C. Macchiavello, *Phys. Rev. Lett.* **82**, p. 4344 (1999).
27. G. Vidal, J. I. Latorre, P. Pascual and R. Tarrach, *Phys. Rev. A* **60**, p. 126 (1999).
28. R. D. Gill and S. Massar, *Phys. Rev. A* **61**, p. 042312 (2000).
29. M. Keyl and R. F. Werner, *Phys. Rev. A* **64**, p. 052311 (2001).
30. E. Bagan, M. Baig and R. Munoz-Tapia, *Phys. Rev. Lett.* **89**, p. 277904 (2002).
31. M. Hayashi and K. Matsumoto, Asymptotic performance of optimal state estimation in quantum two level system, quant-ph/0411073.
32. M. Hayashi and K. Matsumoto, Statistical model with measurement degree

of freedom and quantum physics, in *Asymptotic theory of quantum statistical inference: selected papers*, ed. M. Hayashi (World Scientific, 2005) pp. 162–170. (English translation of a paper in Japanese published in Surikaiseki Kenkyusho Kokyuroku, vol. 35, pp. 7689-7727, 2002.).
33. E. Bagan, M.A. Ballester, R.D. Gill, A. Monras and R. Munõz-Tapia, *Phys. Rev. A* **73**, p. 032301 (2006).
34. R. D. Gill, Asymptotic information bounds in quantum statistics, quant-ph/0512443, to appear in Annals of Statistics.
35. M. Hayashi, *Bulletin of the Mathematical Society of Japan* **55**, 368 (2003), (in Japanese; Translated into English in quant-ph/0608198).
36. M. Hayashi, presentations at maphysto and quantop workshop on quantum measurements and quantum stochastics, aarhus, 2003, and special week on quantum statistics, isaac newton institute for mathematical sciences, cambridge, 2004.
37. M. Guţă and J. Kahn, *Phys. Rev. A* **73**, p. 052108 (2006).
38. M. Guţă, B. Janssens and J. Kahn, Optimal estimation of qubit states with continuous time measurements, *Commun. Math. Phys.* **277**, 127 (2008).
39. A. Wald, *Statistical Decision Functions* (John Wiley & Sons, New York, 1950).
40. L. Le Cam, *Asymptotic Methods in Statistical Decision Theory* (Springer Verlag, New York, 1986).
41. M. Guţă and A. Jençová, Local asymptotic normality in quantum statistics, *Commun. Math. Phys.* **276**, 341 (2007)
42. J. Kahn, M. Guţă, Local asymptotic normality for d-dimensional quantum states, arXiv:0804.3876 [quant-ph].
43. K. Matsumoto, unpublished manuscript.
44. H. Strasser, *Mathematical Theory of Statistics* (De Gruyter, Berlin, New York, 1985).
45. D. Petz, *Commun. Math. Phys.* **105**, 123 (1986).
46. D. Petz and A. Jencova, *Commun. Math. Phys.* **263**, 259 (2006).
47. E. Torgersen, *Comparison of Statistical Experiments* (Cambridge University Press, 1991).
48. A. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, 1998).
49. U. Leonhardt, *Measuring the Quantum State of Light* (Cambridge University Press, 1997).

PART C

Quantum Measurements and Information

This page intentionally left blank

INFORMATION GAIN IN QUANTUM CONTINUAL MEASUREMENTS

ALBERTO BARCHIELLI

Politecnico di Milano, Dipartimento di Matematica,
Piazza Leonardo da Vinci 32, I-20133 Milano, Italy,
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano
E-mail: Alberto.Barchielli@polimi.it

GIANCARLO LUPIERI

Università degli Studi di Milano, Dipartimento di Fisica,
Via Celoria 16, I-20133 Milano, Italy,
and Istituto Nazionale di Fisica Nucleare, Sezione di Milano
E-mail: Giancarlo.Lupieri@mi.infn.it

Inspired by works on information transmission through quantum channels, we propose the use of a couple of mutual entropies to quantify the efficiency of continual measurement schemes in extracting information on the measured quantum system. Properties of these measures of information are studied and bounds on them are derived.

Keywords: Quantum continual measurement; Quantum trajectories; Quantum/classical mutual entropies; Marginals of states on product algebras; Bounds on information

1. Quantum measurements and entropies

We speak of quantum continual measurements when a quantum system is taken under observation with continuity in time and the output is not a single random variable, but rather a stochastic process[1,2]. The aim of this paper is to quantify, by means of entropic quantities, the effectiveness of a continual measurement in extracting information from the underlying quantum system.

Various types of entropies and bounds on informational quantities can be introduced and studied in connection with continual measurements[3–5]. In particular, in Ref. 5 the point of view was the one of information transmission: the quantum system is a channel in which some information is

encoded at an initial time; the continual measurement represents the decoding apparatus. In this paper, instead, we consider the quantum system in itself, not as a transmission channel, and we propose and study a couple of mutual entropies giving two indexes of how good is the continual measurement in extracting information about the quantum system.

1.1. *Algebras, states, entropies*

From now on $\mathcal{H}$ will be a separable complex Hilbert space, the space where our quantum system lives.

1.1.1. *Von Neumann algebras and normal states*

A normal state on $\mathcal{L}(\mathcal{H})$ (bounded linear operators on $\mathcal{H}$) is identified with a statistical operator, $\mathcal{T}(\mathcal{H})$ and $\mathcal{S}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$ are the trace-class and the space of the statistical operators on $\mathcal{H}$, respectively.

Let $(\Omega, \mathcal{F}, Q)$ be a measure space, where Q is a σ-finite measure. We consider the W^*-algebras $L^\infty(\Omega, \mathcal{F}, Q)$ and $L^\infty\big(\Omega, \mathcal{F}, Q; \mathcal{L}(\mathcal{H})\big) \simeq L^\infty(\Omega, \mathcal{F}, Q) \otimes \mathcal{L}(\mathcal{H})$. Let us note that a normal state on $L^\infty(\Omega, \mathcal{F}, Q)$ is a probability density with respect to Q, while a normal state σ on $L^\infty\big(\Omega, \mathcal{F}, Q; \mathcal{L}(\mathcal{H})\big)$ is a measurable function $\omega \mapsto \sigma(\omega) \in \mathcal{T}(\mathcal{H})$, $\sigma(\omega) \geq 0$, such that $\mathrm{Tr}\{\sigma(\omega)\}$ is a probability density with respect to Q.

1.1.2. *Relative entropy*

The general definition of the relative entropy $S(\Sigma|\Pi)$ for two states Σ and Π is given in Ref. 6; here we give only some particular cases of the general definition.

Let us consider two quantum states σ, $\tau \in \mathcal{S}(\mathcal{H})$ and two classical states q_k on $L^\infty(\Omega, \mathcal{F}, Q)$ (two probability densities with respect to Q). The von Neumann entropy, the quantum relative entropy and the classical one are

$$S_{\mathrm{q}}(\tau) := -\,\mathrm{Tr}\{\tau \ln \tau\}, \qquad S_{\mathrm{q}}(\sigma\|\tau) = \mathrm{Tr}\{\sigma(\ln\sigma - \ln\tau)\}, \tag{1}$$

$$S_{\mathrm{c}}(q_1\|q_2) = \int_\Omega Q(\mathrm{d}\omega)\, q_1(\omega) \ln \frac{q_1(\omega)}{q_2(\omega)}\,. \tag{2}$$

Let us consider now two normal states σ_k on $L^\infty\big(\Omega, \mathcal{F}, Q; \mathcal{L}(\mathcal{H})\big)$ and set $q_k(\omega) := \mathrm{Tr}\{\sigma_k(\omega)\}$, $\varrho_k(\omega) := \sigma_k(\omega)/q_k(\omega)$ (these definitions hold where the denominators do not vanish and are completed arbitrarily where the

denominators vanish). Then, the relative entropy is

$$\begin{aligned} S(\sigma_1\|\sigma_2) &= \int_\Omega Q(\mathrm{d}\omega)\,\mathrm{Tr}\left\{\sigma_1(\omega)\big(\ln\sigma_1(\omega)-\ln\sigma_2(\omega)\big)\right\} \\ &= S_{\mathrm{c}}(q_1\|q_2)+\int_\Omega Q(\mathrm{d}\omega)\,q_1(\omega)S_{\mathrm{q}}\big(\varrho_1(\omega)\|\varrho_2(\omega)\big). \end{aligned} \tag{3}$$

We are using a subscript "c" for classical entropies, a subscript "q" for purely quantum ones and no subscript for general entropies, eventually of a mixed character. Having used the natural logarithm in these definitions, the entropies are in *nats*. To obtain entropies in *bits* one has to divide by $\ln 2$.

The following result is very useful ([6] Corollary 5.20 and Eq. (5.22)).

Proposition 1.1. *Let* $\Pi_1\otimes\Pi_2$ *and* Σ_{12} *be normal states of the tensor product von Neumann algebra* $\mathcal{M}_1\otimes\mathcal{M}_2$ *and let* $\Sigma_i=\Sigma_{12}\big|_{\mathcal{M}_i}$, $i=1,2$. *Then,*

$$\begin{aligned} S(\Sigma_{12}\|\Pi_1\otimes\Pi_2) &= S(\Sigma_1\|\Pi_1)+S(\Sigma_{12}\|\Sigma_1\otimes\Pi_2) \\ &= S(\Sigma_1\|\Pi_1)+S(\Sigma_2\|\Pi_2)+S(\Sigma_{12}\|\Sigma_1\otimes\Sigma_2) \end{aligned} \tag{4}$$

The quantity $S(\Sigma_{12}\|\Sigma_1\otimes\Sigma_2)$ is the relative entropy of a state with respect to its marginals; this is what we call *mutual entropy*.

1.2. *Instruments and channels*

1.2.1. *Channels*

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two W^*-algebras. A linear map Λ^* from $\mathcal{M}_2$ to $\mathcal{M}_1$ is said to be a *channel* ([6] p. 137) if it is completely positive, unital (i.e. identity preserving) and normal (or, equivalently, weakly* continuous).

Due to the equivalence[7] of w*-continuity and existence of a preadjoint Λ, a *channel* is equivalently defined by: Λ is a norm-one, completely positive linear map from the predual $\mathcal{M}_{1*}$ to the predual $\mathcal{M}_{2*}$. Let us note also that Λ maps normal states on $\mathcal{M}_1$ into normal states on $\mathcal{M}_2$.

A key result which follows from the convexity properties of the relative entropy is *Uhlmann monotonicity theorem* ([6], Theor. 1.5 p. 21), which implies that channels decrease the relative entropy.

Theorem 1.1. *If* Σ *and* Π *are two normal states on* $\mathcal{M}_1$ *and* Λ *is a channel from* $\mathcal{M}_{1*}\to\mathcal{M}_{2*}$, *then* $S(\Sigma\|\Pi)\geq S(\Lambda[\Sigma]\|\Lambda[\Pi])$.

1.2.2. *Instruments and POV measures*

The notion of instrument is central in quantum measurement theory; an instrument gives the probabilities and the state changes[8,9].

Let $(\Omega, \mathcal{F})$ be a measurable space. An *instrument* $\mathcal{I}$ is a map valued measure such that (i) $\mathcal{I}(F)$ is a completely positive, linear, bounded operator on $\mathcal{T}(\mathcal{H})$, $\forall F \in \mathcal{F}$, (ii) $\mathcal{I}(\Omega)$ is trace preserving, (iii) for every countable family $\{F_i\}$ of disjoint sets in $\mathcal{F}$ one has $\sum_i \mathrm{Tr}\,\{a\,\mathcal{I}(F_i)[\rho]\} = \mathrm{Tr}\,\{a\,\mathcal{I}\left(\bigcup_i F_i\right)[\rho]\}$, $\forall \rho \in \mathcal{T}(\mathcal{H})$, $\forall a \in \mathcal{L}(\mathcal{H})$.

The map $F \mapsto \mathcal{I}(F)^*[\mathbb{1}]$ turns out to be a positive operator valued (POV) measure (the observable associated with the instrument $\mathcal{I}$). For every $\rho \in \mathcal{S}(\mathcal{H})$ the map $F \mapsto P_\rho(F) := \mathrm{Tr}\{\mathcal{I}(F)[\rho]\}$ is a probability measure: the probability that the result of the measurement be in F when the pre-measurement state is ρ. Moreover, given the result F, the post-measurement state is $\big(P_\rho(F)\big)^{-1}\mathcal{I}(F)[\rho]$.

1.2.3. *The instrument as a channel*

Given an instrument $\mathcal{I}$ with value space $(\Omega, \mathcal{F})$ it is always possible to find a σ-finite measure on $(\Omega, \mathcal{F})$ (or even a probability measure), such that all the probabilities P_ρ, $\rho \in \mathcal{S}(\mathcal{H})$, are absolutely continuous with respect to Q.

Theorem 1.2 (Ref. 10, Theorem 2). *Let $\mathcal{I}$ be an instrument on the trace-class of a complex separable Hilbert space $\mathcal{H}$ with value space $(\Omega, \mathcal{F})$ and let Q be a σ-finite measure on $(\Omega, \mathcal{F})$ such that $\mathrm{Tr}\{\mathcal{I}(\bullet)[\rho]\} \ll Q$, $\forall \rho \in \mathcal{S}(\mathcal{H})$. Then, there exists a unique channel $\Lambda_{\mathcal{I}}$ from $\mathcal{T}(\mathcal{H})$ into $L^1\big(\Omega, \mathcal{F}, Q; \mathcal{T}(\mathcal{H})\big)$ such that*

$$\mathbb{E}_Q\left[f\,\mathrm{Tr}\,\{a\,\Lambda_{\mathcal{I}}[\rho]\}\right] = \int_\Omega f(\omega)\,\mathrm{Tr}\,\{a\,\mathcal{I}(\mathrm{d}\omega)[\rho]\} \qquad (5)$$
$$\forall \rho \in \mathcal{T}(\mathcal{H}), \quad \forall a \in \mathcal{L}(\mathcal{H}), \quad \forall f \in L^\infty(\Omega, \mathcal{F}, Q).$$

Viceversa, a channel Λ from $\mathcal{T}(\mathcal{H})$ into $L^1\big(\Omega, \mathcal{F}, Q; \mathcal{T}(\mathcal{H})\big)$ defines a unique instrument $\mathcal{I}$ by

$$\mathcal{I}(F)[\rho] = \mathbb{E}_Q\left[1_F \Lambda[\rho]\right], \qquad \forall \rho \in \mathcal{T}(\mathcal{H}), \quad \forall F \in \mathcal{F}\,. \qquad (6)$$

1.2.4. *A posteriori states*

When $\rho \in \mathcal{S}(\mathcal{H})$, then $\Lambda_{\mathcal{I}}[\rho]$ is a normal state on $L^\infty\big(\Omega, \mathcal{F}, Q; \mathcal{L}(\mathcal{H})\big)$. Let us normalize the positive trace-class operators $\Lambda_{\mathcal{I}}[\rho](\omega)$ by setting

$$\pi_\rho(\omega) := \begin{cases} \left(\mathrm{Tr}_{\mathcal{H}}\left\{\Lambda_{\mathcal{I}}[\rho](\omega)\right\}\right)^{-1} \Lambda_{\mathcal{I}}[\rho](\omega) & \text{if } \mathrm{Tr}_{\mathcal{H}}\left\{\Lambda_{\mathcal{I}}[\rho](\omega)\right\} > 0 \\ \tilde{\rho} \quad \big(\tilde{\rho} \in \mathcal{S}(\mathcal{H}), \text{ fixed}\big) & \text{if } \mathrm{Tr}_{\mathcal{H}}\left\{\Lambda_{\mathcal{I}}[\rho](\omega)\right\} = 0 \end{cases} \tag{7}$$

Then, we have

$$\int_F \pi_\rho(\omega) P_\rho(\mathrm{d}\omega) = \mathcal{I}(F)[\rho]\,, \quad \forall F \in \mathcal{F}\,, \qquad \text{(Bochner integral)}. \tag{8}$$

According to Ozawa[11], π_ρ is a family *of a posteriori states* for the instrument $\mathcal{I}$ and the pre-measurement state ρ. The interpretation is that $\pi_\rho(\omega)$ is the state just after the measurement to be attributed to the quantum system if the result of the measurement has been exactly ω.

Let us note that $p_\rho := \mathrm{Tr}\left\{\Lambda_{\mathcal{I}}[\rho]\right\}$ and $\overline{\pi}_\rho := \int_\Omega P_\rho(\mathrm{d}\omega)\pi_\rho(\omega) = \mathcal{I}(\Omega)[\rho]$ are the marginals of the state $\Lambda_{\mathcal{I}}[\rho]$ on the algebras $L^\infty(\Omega, \mathcal{F}, Q)$ and $\mathcal{L}(\mathcal{H})$, respectively. Then, $S\left(\Lambda_{\mathcal{I}}[\rho] \| p_\rho \overline{\pi}_\rho\right)$ is a first example of a mutual entropy. From Eqs. (3) and (1) we get

$$\begin{aligned} S\left(\Lambda_{\mathcal{I}}[\rho] \| p_\rho \overline{\pi}_\rho\right) &= \int_\Omega S_{\mathrm{q}}\big(\pi_\rho(\omega) \| \overline{\pi}_\rho\big) P_\rho(\mathrm{d}\omega) \\ &= S_{\mathrm{q}}(\overline{\pi}_\rho) - \int_\Omega S_{\mathrm{q}}\big(\pi_\rho(\omega)\big) P_\rho(\mathrm{d}\omega). \end{aligned} \tag{9}$$

Quantities like this one are used in quantum information transmission and are known as Holevo capacities or χ-quantities[12–14]; Eq. (9) gives the χ-quantity of the *ensemble* of states $\{P_\rho, \pi_\rho\}$.

2. Continual measurements

Quantum continual measurement theory can be formulated in different equivalent ways. To construct our entropic measures of efficiency, we need two approaches to continual measurements: the one based on positive operator valued measures, instruments, quantum channels[1,5,15] and the one based on classical stochastic differential equations (SDE's), known also as quantum trajectory theory[2,4,16].

The SDE approach to continual measurements is based on a couple of stochastic equations, a linear one for random trace-class operators and a non-linear one for random statistical operators. The two equations are linked by a change of normalization and a change of probability measure. Both equations have a Hilbert space formulation, particularly suited for

numerical computations. We shall use a simplified version of SDE's for continual measurements as presented in [2].

2.1. *The linear equation*

Let $H(t), L_l(t), R_j(t), V_k^r(t), J_k(t)$ be bounded operators on $\mathcal{H}$; their time dependence is taken to be continuous from the left and with limits from the right in the strong topology. The indices k, l, j take a finite number of values; the index r can take infinitely many values, but in this case the series $\sum_r V_k^r(t)^* V_k^r(t)$ is strongly convergent. Let the operator $H(t)$ be self-adjoint, $H(t) = H(t)^*$, and let us define $(\forall \rho \in \mathcal{T}(\mathcal{H}))$

$$\mathcal{J}_k(t)[\rho] := \sum_r V_k^r(t) \rho V_k^r(t)^* \,, \tag{10a}$$

$$J_k(t) := \mathcal{J}_k(t)^*[\mathbb{1}] = \sum_r V_k^r(t)^* V_k^r(t) \,, \tag{10b}$$

$$\mathcal{L}(t) := \mathcal{L}_0(t) + \mathcal{L}_1(t) + \mathcal{L}_2(t), \tag{10c}$$

$$\mathcal{L}_0(t)[\rho] := -\mathrm{i}[H(t), \rho] + \sum_l \left(L_l(t) \rho L_l(t)^* - \frac{1}{2} \left\{ L_l(t)^* L_l(t), \rho \right\} \right), \tag{10d}$$

$$\mathcal{L}_1(t)[\rho] := \sum_j \left(R_j(t) \rho R_j(t)^* - \frac{1}{2} \left\{ R_j(t)^* R_j(t), \rho \right\} \right), \tag{10e}$$

$$\mathcal{L}_2(t)[\rho] := \sum_k \left(\mathcal{J}_k(t)[\rho] - \frac{1}{2} \left\{ J_k(t), \rho \right\} \right). \tag{10f}$$

By $[\ ,\ ]$ we denote the commutator and by $\{\ ,\ \}$ the anticommutator.

Then, we introduce a probability space $(\Omega, \mathcal{F}, Q)$ where the Poisson processes $N_k(t)$, of intensity λ_k, and the standard (continuous) Wiener processes $W_j(t)$ are defined. All the processes are assumed to be independent from the other ones. We introduce also the two-times natural filtration of such processes:

$$\mathcal{F}_t^s = \sigma\{W_j(u) - W_j(s),\, N_k(v) - N_k(s),\, u, v \in [s,t],\, j,k = 1, \ldots\}. \tag{11}$$

Having all these ingredients, we can introduce the linear equation of continual measurement theory, for the a trace-class valued process σ_t:

$$\begin{aligned} \mathrm{d}\sigma_t = \mathcal{L}(t)[\sigma_{t_-}]\mathrm{d}t + \sum_j \left(R_j(t)\sigma_{t_-} + \sigma_{t_-} R_j(t)^* \right) \mathrm{d}W_j(t) \\ + \sum_k \left(\frac{1}{\lambda_k} \mathcal{J}_k(t)[\sigma_{t_-}] - \sigma_{t_-} \right) \left(\mathrm{d}N_k(t) - \lambda_k \mathrm{d}t \right). \end{aligned} \tag{12}$$

The initial condition is taken to be a non-random statistical operator: $\sigma_0 \equiv \sigma_{0_-} \in \mathcal{S}(\mathcal{H})$.

The notation σ_{t_-} means that, in case there is a jump in the noise at time t, the value just before the jump σ_{t_-} of σ has to be taken. More precisely, if the augmented natural filtration of the noises is considered, the solution can be taken to be continuous from the right and with limits from the left and σ_{t_-} is just the limit from the left. We prefer not to add the null sets to the natural filtration and by σ_t we mean some $\mathcal{F}_t^0$-adapted version of the solution.

Properties of the solution. Let us consider now, for $0 \leq s \leq t$, the von Neumann algebra $L^\infty\big(\Omega, \mathcal{F}_t^s, Q; \mathcal{L}(\mathcal{H})\big) \simeq L^\infty(\Omega, \mathcal{F}_t^s, Q) \otimes \mathcal{L}(\mathcal{H})$ (cf. Section 1.1.1) and let us give a name to the set of normal states on this algebra:

$$\mathcal{S}_t^s := \left\{ \tau \in L^1\big(\Omega, \mathcal{F}_t^s, Q; \mathcal{T}(\mathcal{H})\big) : \tau(\omega) \geq 0\,,\ \int_\Omega \mathrm{Tr}\{\tau(\omega)\} Q(\mathrm{d}\omega) = 1 \right\}. \tag{13}$$

- First of all, it is possible to prove that $\sigma_t \in \mathcal{S}_t^0$; we can say that the solution at time t of Eq. (12) is a kind of quantum/classical state.
- The marginals of σ_t are (cf. Section 1.2.4):
  - The probability density $p_t := \mathrm{Tr}\{\sigma_t\}$. The probability measure $p_t(\omega)Q(\mathrm{d}\omega)$ will be the physical probability.
  - The *a priori state at time t* $\eta_t := \mathbb{E}_Q[\sigma_t] \in \mathcal{S}(\mathcal{H})$. This is the state to be attributed at time t to the system when no selection is done and the result of the measurement has not been taken into account.
- Moreover, we define the random *a posteriori state at time t* $\rho_t = \dfrac{1}{p_t}\sigma_t$. This is the state to be attributed at time t to the system known the result of the measurement up to t.

Note that $p_0 = 1$, $\rho_0 = \sigma_0 = \eta_0$, $\eta_{t_-} = \eta_t = \eta_{t_+}$.

2.2. *Physical probabilities*

A very important property of Eq. (12) is that p_t is a mean one Q-martingale, which implies that

$$P_t(\mathrm{d}\omega) := p_t(\omega)Q(\mathrm{d}\omega)\big|_{\mathcal{F}_t^0} \tag{14}$$

is a consistent family of probabilities, i.e., if $0 \leq t < T$, $P_T(F) = P_t(F)$, $\forall F \in \mathcal{F}_t^0$. These are taken as physical probabilities.

From Eq. (12) we have that p_t satisfies the Doléans equation

$$\mathrm{d}p_t = p_{t_-}\left\{\sum_j m_j(t)\,\mathrm{d}W_j(t) + \sum_k \left(\frac{\mu_k(t)}{\lambda_k} - 1\right)\left(\mathrm{d}N_k(t) - \lambda_k\,\mathrm{d}t\right)\right\}, \quad (15)$$

where

$$m_j(t) = \mathrm{Tr}\left\{(R_j(t) + R_j(t)^*)\,\rho_{t_-}\right\}, \qquad \mu_k(t) = \mathrm{Tr}\left\{J_k(t)\rho_{t_-}\right\}. \quad (16)$$

The solution of this equation, with $p_0 = 1$, is

$$p_t = \exp\left\{\sum_j\left[\int_0^t m_j(s)\,\mathrm{d}W_j(s) - \frac{1}{2}\int_0^t m_j(s)^2\,\mathrm{d}s\right] + \sum_k\left[\int_0^t \ln\frac{\mu_k(s)}{\lambda_k}\,\mathrm{d}N_k(s) + \int_0^t(\lambda_k - \mu_k(s))\,\mathrm{d}s\right]\right\}. \quad (17)$$

Remark 2.1.

(1) The output of the continual measurement is the set of processes $W_j(t)$, $N_k(t)$, $0 \le t \le T$, under the physical probability P_T; T is a completely arbitrary large time. By the consistency of the probabilities (14), P_T can be substituted by P_t in any expectation involving $\mathcal{F}_t^0$-measurable random variables (for $t < T$).

(2) By Girsanov theorem and its generalizations for situations with jumps, we have that, under the physical probability, the processes

$$\widehat{W}_j(t) = W_j(t) - \int_0^t m_j(s)\,\mathrm{d}s \quad (18)$$

are independent, standard Wiener processes and $N_k(t)$ is a counting process of stochastic intensity $\mu_k(t)\mathrm{d}t$.

(3) Expressions for the moments of the outputs can be given; in particular we have the mean values

$$\mathbb{E}_{P_t}\left[W_j(t)\right] = \int_0^t n_j(s)\,\mathrm{d}s\,, \qquad \mathbb{E}_{P_t}\left[N_k(t)\right] = \int_0^t \nu_k(s)\,\mathrm{d}s\,, \quad (19)$$

where

$$n_j(t) = \mathrm{Tr}\left\{(R_j(t) + R_j(t)^*)\,\eta_t\right\} = \mathbb{E}_{P_t}[m_j(t)], \quad (20a)$$

$$\nu_k(t) = \mathrm{Tr}\left\{J_k(t)\eta_t\right\} = \mathbb{E}_{P_t}[\mu_k(t)]. \quad (20b)$$

2.3. *The non-linear SDE*

Under the physical law P_T, the a posteriori states ρ_t satisfy the non-linear SDE

$$\mathrm{d}\rho_t = \mathcal{L}(t)[\rho_{t_-}]\mathrm{d}t + \sum_j \left(R_j(t)\rho_{t_-} + \rho_{t_-}R_j(t)^* - m_j(t)\rho_{t_-}\right)\mathrm{d}\widehat{W}_j(t) + \sum_k \left(\frac{1}{\mu_k(t)}\,\mathcal{J}_k(t)[\rho_{t_-}] - \rho_{t_-}\right)\left(\mathrm{d}N_k(t) - \mu_k(t)\mathrm{d}t\right). \quad (21)$$

Let us stress that for the a priori states we have

$$\eta_t = \mathbb{E}_Q[\sigma_t] = \mathbb{E}_{P_t}[\rho_t] \quad (22)$$

and that they satisfy the *master equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\eta_t = \mathcal{L}(t)[\eta_t]. \quad (23)$$

2.4. *The fundamental matrix and the instruments*

To apply the notions of Section 1 to continual measurements, we need to see how such a theory is connected to instruments and channels[2–5]. This is done by introducing the fundamental matrix Λ_t^s of (12). This operator is defined by stipulating that $\Lambda_t^s[|u_i\rangle\langle u_j|]$ satisfies (12) with initial condition $\Lambda_s^s[|u_i\rangle\langle u_j|] = |u_i\rangle\langle u_j|$, where $\{u_i\ i = 1, \ldots\}$ is a c.o.n.s. in $\mathcal{H}$. It turns out that Λ_t^s is a channel from $\mathcal{T}(\mathcal{H})$ into $L^1\big(\Omega, \mathcal{F}_t^s, Q; \mathcal{T}(\mathcal{H})\big)$, or, by trivial ampliation, from $L^1\big(\Omega, \mathcal{F}_s^r, Q; \mathcal{T}(\mathcal{H})\big)$ into $L^1\big(\Omega, \mathcal{F}_t^r, Q; \mathcal{T}(\mathcal{H})\big)$, $0 \le r \le s \le t$. Then, we have

$$\Lambda_t^s[\sigma_s] = \sigma_t\,, \qquad \Lambda_t^s = \Lambda_t^u \circ \Lambda_u^s\,, \quad 0 \le s \le u \le t\,. \quad (24)$$

The instrument associated to this channel is

$$\mathcal{I}_t^s(F)[\rho] = \mathbb{E}_Q\left[1_F \Lambda_t^s[\rho]\right] \equiv \int_F \Lambda_t^s(\omega)[\rho]Q(\mathrm{d}\omega), \qquad \forall F \in \mathcal{F}_t^s. \quad (25)$$

The time evolution of the quantum states is the one generated by $\mathcal{L}(t)$ and we have

$$\mathcal{U}(t,s)[\rho] = \mathcal{I}_t^s(\Omega)[\rho] = \mathbb{E}_Q\left[\Lambda_t^s[\rho]\right], \quad (26)$$

$$\mathcal{U}(t,s)[\eta_s] = \eta_t\,, \qquad \mathcal{U}(t,s) = \mathcal{U}(t,u) \circ \mathcal{U}(u,s)\,, \quad 0 \le s \le u \le t\,. \quad (27)$$

According to the definitions of Section 1.2.4, the random statistical operator ρ_t is the a posteriori state for the instrument $\mathcal{I}_t^0$ and the pre-measurement state $\rho_0 \equiv \eta_0$.

Another important property is

$$\mathbb{E}_Q[\sigma_t|\mathcal{F}_t^s] = \Lambda_t^s[\eta_s] \in \mathcal{S}_t^s\,. \tag{28}$$

Indeed, by the first of (24) and the fact that Λ_t^s is $\mathcal{F}_t^s$-measurable, we have $\mathbb{E}_Q[\sigma_t|\mathcal{F}_t^s] = \Lambda_t^s\big[\mathbb{E}_Q[\sigma_s|\mathcal{F}_t^s]\big]$. By the fact that all the noises have independent increments, we have that σ_s is independent from $\mathcal{F}_t^s$ and $\mathbb{E}_Q[\sigma_s|\mathcal{F}_t^s] = \mathbb{E}_Q[\sigma_s] = \eta_s$. This gives Eq. (28).

3. Mutual entropies and information gains

3.1. *The information embedded in the a posteriori states*

The quantity σ_t is a state on $L^\infty\big(\Omega, \mathcal{F}_t^0, Q; \mathcal{L}(\mathcal{H})\big) = L^\infty(\Omega, \mathcal{F}_t^0, Q) \otimes \mathcal{L}(\mathcal{H})$ and its marginals on $L^\infty(\Omega, \mathcal{F}_t^0, Q)$ and $\mathcal{L}(\mathcal{H})$ are p_t and η_t, respectively. The mutual entropy $S(\sigma_t\|p_t\eta_t)$ is the "information" contained in the joint state with respect to the product of these marginals; more explicitly we have (compare with (9))

$$S(\sigma_t\|p_t\eta_t) = \int_\Omega P_t(\mathrm{d}\omega)\,\mathrm{Tr}\left\{\rho_t(\omega)\big(\ln\rho_t(\omega) - \ln\eta_t\big)\right\}$$

and we can write

$$S(\sigma_t\|p_t\eta_t) = \mathbb{E}_{P_t}[S_{\mathrm{q}}(\rho_t\|\eta_t)] = S_{\mathrm{q}}(\eta_t) - \mathbb{E}_{P_t}[S_{\mathrm{q}}(\rho_t)]. \tag{29}$$

This mutual entropy is a sort of quantum information embedded by the measurement in the a posteriori states. When the measurement is not informative, we have $\rho_t(\omega) = \eta_t$ and $S(\sigma_t\|p_t\eta_t) = 0$. It is zero also if for any reason it happens that η_t is a pure state. For instance, if $\mathcal{U}(t,0)$ has a unique equilibrium state which is pure, then $\lim_{t\to+\infty} S(\sigma_t\|p_t\eta_t) = 0$ even if the measurement is "good".

Let us note that from Eq. (29) we have the bound

$$S(\sigma_t\|p_t\eta_t) \le S_{\mathrm{q}}(\eta_t). \tag{30}$$

When the von Neumann entropy of the a priori state is not zero, an instantaneous index of "goodness" of the measurement could be $S(\sigma_t\|p_t\eta_t)\big/S_{\mathrm{q}}(\eta_t)$, while a "cumulative" index could be $\int_0^T \frac{S(\sigma_t\|p_t\eta_t)}{S_{\mathrm{q}}(\eta_t)}\,\mathrm{d}t$.

3.2. *A classical continual information gain*

3.2.1. *Product densities*

Let us consider any time s in the time interval $(0,t)$ and let us decompose the von Neumann algebra $L^\infty(\Omega, \mathcal{F}_t^0, Q)$ as $L^\infty(\Omega, \mathcal{F}_t^0, Q) =$

$L^\infty(\Omega, \mathcal{F}_s^0, Q) \otimes L^\infty(\Omega, \mathcal{F}_t^s, Q)$. Now, the density p_t can be seen as a state on $L^\infty(\Omega, \mathcal{F}_t^0, Q)$ and we can consider its marginals p_s^0 and p_t^s on the two factors $L^\infty(\Omega, \mathcal{F}_s^0, Q)$ and $L^\infty(\Omega, \mathcal{F}_t^s, Q)$, respectively. These marginals are given by

$$p_s^0 = \mathbb{E}_Q[p_t|\mathcal{F}_s^0], \qquad p_t^s = \mathbb{E}_Q[p_t|\mathcal{F}_t^s]. \tag{31}$$

By using the fact that $\{p_t,\, t \geq 0\}$ is a martingale and by taking the trace of Eq. (28), we get

$$p_s^0 = p_s\,, \qquad p_t^s = \mathrm{Tr}\{\Lambda_t^s[\eta_s]\}\,. \tag{32}$$

By comparing the last equality with $p_t = \mathrm{Tr}\{\sigma_t\} = \mathrm{Tr}\{\Lambda_t^0[\eta_0]\}$, we see that p_t^s is similar to p_t, but with s as initial time, instead of 0, and with η_s as initial state, instead of η_0. By this remark and Eq. (17), we get

$$p_t^s = \exp\Bigg\{\sum_j \bigg[\int_s^t m_j(u;s)\,\mathrm{d}W_j(u) - \frac{1}{2}\int_s^t m_j(u;s)^2\,\mathrm{d}u\bigg] + \sum_k \bigg[\int_s^t \ln\frac{\mu_k(u;s)}{\lambda_k}\,\mathrm{d}N_k(u) + \int_s^t (\lambda_k - \mu_k(u;s))\,\mathrm{d}u\bigg]\Bigg\}, \tag{33}$$

where

$$m_j(t;s) = \mathrm{Tr}\left\{(R_j(t) + R_j(t)^*)\,\rho_{t_-}^s\right\}, \qquad \mu_k(t;s) = \mathrm{Tr}\left\{J_k(t)\rho_{t_-}^s\right\}, \tag{34}$$

$$\rho_t^s = \frac{1}{p_t^s}\,\Lambda_t^s[\eta_s]. \tag{35}$$

The random state ρ_t^s is the a posteriori state for the instrument $\mathcal{I}_t^s$ and the pre-measurement state η_s; it satisfy the non-linear SDE (21).

Then, we can consider the mutual entropy $S_{\mathrm{c}}(p_t \| p_s^0 p_t^s)$. But the significance of this quantity is dubious, because the time s is completely arbitrary and, moreover, we could divide the time interval in more pieces. For instance, we can take the decomposition $L^\infty(\Omega, \mathcal{F}_t^0, Q) = L^\infty(\Omega, \mathcal{F}_r^0, Q) \otimes L^\infty(\Omega, \mathcal{F}_s^r, Q) \otimes L^\infty(\Omega, \mathcal{F}_t^s, Q)$ and we recognize that $p_r^0 p_s^r p_t^s$ is the product of the marginals of p_t related to this decomposition. Taking a finer generic partition of $(0,t)$ with $t_0 = 0$ and $t_n = t$, we recognize that $\prod_{j=1}^n p_{t_j}^{t_{j-1}}$ is again a product of marginals of p_t. To eliminate arbitrariness, let us consider finer and finer partitions and let us go to a continuous product of marginals.

Let us note that we have

$$\lim_{s\uparrow t} m_j(t;s) = n_j(t), \qquad \lim_{s\uparrow t} \mu_k(t;s) = \nu_k(t), \qquad \text{a.s.}$$

Then, for an infinitesimal interval we get

$$p^s_{s+\mathrm{d}s} = \exp\Big\{\sum_j \Big[n_j(s)\,\mathrm{d}W_j(s) - \frac{1}{2}\,n_j(s)^2\mathrm{d}s\Big] + \sum_k \Big[\frac{\nu_k(s)}{\lambda_k}\mathrm{d}N_k(s) + (\lambda_k - \nu_k(s))\,\mathrm{d}s\Big]\Big\} \quad (36)$$

and, so, the following density q_t is the continuous product of marginals of p_t:

$$q_t = \exp\Big\{\sum_j \Big[\int_0^t n_j(s)\,\mathrm{d}W_j(s) - \frac{1}{2}\int_0^t n_j(s)^2\,\mathrm{d}s\Big] + \sum_k \Big[\int_0^t \ln\frac{\nu_k(s)}{\lambda_k}\,\mathrm{d}N_k(s) + \int_0^t (\lambda_k - \nu_k(s))\,\mathrm{d}s\Big]\Big\}. \quad (37)$$

Notice that $n_j(t)$ and $\nu_k(t)$ are deterministic functions. Under the probability $q_T(\omega)Q(\mathrm{d}\omega)$, the processes $W_j(t) - \int_0^t n_j(s)\,\mathrm{d}s$ are independent, standard Wiener processes and $N_k(t)$ is a Poisson process of time dependent intensity $\nu_k(t)$.

Under $q_T(\omega)Q(\mathrm{d}\omega)$, the processes W_j, N_k have independent increments as under Q (so they can be interpreted as noises), but the means have been changed and made equal to the means they have under P_T.

The fact that it is possible to consider a "continuous product of marginals" is not so unexpected; indeed, the theory of continual measurements is connected to infinite divisibility[15].

We have already seen that the marginals of p_t with respect to the decomposition of the time interval $(0,t)$ into $(0,s)$ and (s,t) are $p^0_s = p_s$ and p^s_t given by Eq. (33). The analogous marginals for q_t are $q^0_s = q_s$ and

$$q^s_t = \exp\Big\{\sum_j \Big[\int_t^u n_j(s)\,\mathrm{d}W_j(s) - \frac{1}{2}\int_t^u n_j(s)^2\,\mathrm{d}s\Big] + \sum_k \Big[\int_t^u \ln\frac{\nu_k(s)}{\lambda_k}\,\mathrm{d}N_k(s) + \int_t^u (\lambda_k - \nu_k(s))\,\mathrm{d}s\Big]\Big\} = \frac{q_t}{q_s}\,. \quad (38)$$

3.2.2. *The classical mutual entropy $S_{\mathrm{c}}(p_t\|q_t)$*

The density q_t is no more dependent on some arbitrary choice of intermediate times and the measure $q_T(\omega)Q(\mathrm{d}\omega)$ has a distinguished role and can be considered as a reference measure. So, we can introduce the relative entropy

$$S_{\mathrm{c}}(p_t\|q_t) = \mathbb{E}_{P_t}\left[\ln\frac{p_t}{q_t}\right].$$

Being q_t a product of marginals of p_t, this quantity is a mutual entropy and, being q_t the finest product of marginals, we can interprete $S_{\rm c}(p_t\|q_t)$ as a measure of the classical information on the measured system extracted in the time interval $(0,t)$. Other reasons can be given to reinforce this interpretation.

By Eqs. (17), (33), (37), (38) we have $p_t^0 = p_t$, $q_t^0 = q_t$, $q_u = q_t q_u^t$. By Proposition 1.1 or by direct computation, we get

$$S_{\rm c}(p_t\|q_t) - S_{\rm c}(p_s\|q_s) = S_{\rm c}(p_t\|p_s q_t^s), \qquad 0 \le s \le t\,. \tag{39}$$

Firstly, by the positivity of relative entropies, this equation says that

$$0 \le S_{\rm c}(p_s\|q_s) \le S_{\rm c}(p_t\|q_t), \tag{40}$$

i.e. that $S_{\rm c}(p_t\|q_t)$ is non negative and not decreasing in time, as should be for a measure of an information gain in time. Moreover, the increment of information in the time interval (s,t) can be written as

$$S_{\rm c}(p_t\|p_s q_t^s) = \mathbb{E}_Q\left[p_s\,\mathbb{E}_Q\left[\frac{p_t}{p_s}\ln\frac{p_t/p_s}{q_t/q_s}\middle|\mathcal{F}_s^0\right]\right]. \tag{41}$$

This expression can be interpreted as a *conditional relative entropy* (Ref. 17 pp. 22–23). The quantity $\mathbb{E}_Q\left[\frac{p_t}{p_s}\ln\frac{p_t/p_s}{q_t/q_s}\middle|\mathcal{F}_s^0\right]$ has the same structure as $S_{\rm c}(p_t\|q_t)$, but it refers to the interval (s,t) and it is constructed with the conditional densities. We can say that Eq. (39) expresses in a consistent way a kind of "additivity property" of our measure of information.

Having the explicit exponential forms of the densities p_t and q_t, we can compute the explicit expression of the information gain.

Proposition 3.1. *The explicit expression of the classical mutual entropy* $S_{\rm c}(p_t\|q_t)$ *is*

$$S_{\rm c}(p_t\|q_t) = \frac{1}{2}\sum_j \int_0^t \mathrm{Var}_{P_t}[m_j(s)]\mathrm{d}s + \sum_k \int_0^t \mathbb{E}_{P_t}\left[\mu_k(s)\ln\frac{\mu_k(s)}{\nu_k(s)}\right]\mathrm{d}s\,. \tag{42}$$

Proof. By Eqs. (17) and (37) we get

$$\ln\frac{p_t}{q_t} = \sum_j \Big[\int_0^t \big(m_j(s)-n_j(s)\big)\mathrm{d}W_j(s) - \frac{1}{2}\int_0^t \big(m_j(s)^2-n_j(s)^2\big)\mathrm{d}s\Big]$$
$$+\sum_k\Big[\int_0^t \ln\frac{\mu_k(s)}{\nu_k(s)}\big(\mathrm{d}N_k(s)-\lambda_k\,\mathrm{d}s\big)+\int_0^t\Big(\lambda_k\ln\frac{\mu_k(s)}{\nu_k(s)}-\mu_k(s)+\nu_k(s)\Big)\,\mathrm{d}s\Big]$$
$$=\sum_j\Big[\int_0^t\big(m_j(s)-n_j(s)\big)\big(\mathrm{d}W_j(s)-m_j(s)\mathrm{d}s\big)+\frac{1}{2}\int_0^t\big(m_j(s)-n_j(s)\big)^2\,\mathrm{d}s\Big]$$
$$+\sum_k\Big[\int_0^t\ln\frac{\mu_k(s)}{\nu_k(s)}\big(\mathrm{d}N_k(s)-\mu_k(s)\,\mathrm{d}s\big)$$
$$+\int_0^t\mu_k(s)\Big(\frac{\nu_k(s)}{\mu_k(s)}-\ln\frac{\nu_k(s)}{\mu_k(s)}-1\Big)\,\mathrm{d}s\Big].$$

By point 2 in Remark 2.1, the first term in the j sum and the first term in the k sum have zero mean under P_T (or under P_t, by consistency). Therefore, Eq. (42) follows by taking the P_t-mean of $\ln p_t/q_t$ and by taking into account Eqs. (20). □

Remark 3.1.

(1) By (20b) and Jensen inequality applied to the convex function $x\ln x$, we have that both integrands in formula (42) are non-negative and, so, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}S_{\mathrm{c}}(p_t\|q_t) = \frac{1}{2}\sum_j \mathrm{Var}_{P_t}[m_j(t)] + \sum_k \mathbb{E}_{P_t}\Big[\mu_k(t)\ln\frac{\mu_k(t)}{\nu_k(t)}\Big] \geq 0. \quad (43)$$

The positivity of this time derivative follows also from Eq. (40).

(2) By the properties of relative entropy $\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_t\|\eta_t)] = 0$ is equivalent to $\rho_t = \eta_t$, P_T-a.s. By Eqs. (20), (43), this last relation implies the vanishing of the quantity (43). So, we have

$$\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_t\|\eta_t)] = 0 \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}S_{\mathrm{c}}(p_t\|q_t) = 0\,. \quad (44)$$

(3) From Eqs. (16), (20), (43) we see that

- if $R_j(t) + R_j(t)^* \propto \mathbb{1}$, then $\mathrm{Var}_{P_t}[m_j(t)] = 0$,
- if $J_k(t) \propto \mathbb{1}$, then $\ln\frac{\mu_k(t)}{\nu_k(t)} = 0$.

This says that when both conditions hold for all j and k, no information is extracted from the system, whatever the initial state is.

3.3. *A quantum/classical mutual entropy*

The two mutual entropies introduced in Sections 3.1 and 3.2.2 can be obtained from a unique mutual entropy

$$S(\sigma_t\|q_t\eta_t) = \int_\Omega Q(\mathrm{d}\omega)\,\mathrm{Tr}\left\{\sigma_t(\omega)\big(\ln\sigma_t(\omega) - \ln q_t(\omega)\eta_t\big)\right\}. \tag{45}$$

Indeed, by Proposition 1.1 or by direct computation, we get

$$S(\sigma_t\|q_t\eta_t) = S(\sigma_t\|p_t\eta_t) + S_{\mathrm{c}}(p_t\|q_t) = \mathbb{E}_{P_t}[S_{\mathrm{q}}(\rho_t\|\eta_t)] + S_{\mathrm{c}}(p_t\|q_t). \tag{46}$$

4. An upper bound on the increments of $S_{\mathrm{c}}(p_t\|q_t)$

4.1. *The main bound*

By Proposition 1.1 and Eqs. (17), (33), (37), (38), the increment of information in the time interval (t,u) can be expressed as

$$S_{\mathrm{c}}(p_u\|q_u) - S_{\mathrm{c}}(p_t\|q_t) = S_{\mathrm{c}}(p_u\|p_tp_u^t) + S_{\mathrm{c}}(p_u^t\|q_u^t). \tag{47}$$

Lemma 4.1. *For $0 \le t \le u$, we have the bound*

$$0 \le S_{\mathrm{c}}\left(p_u\|p_tp_u^t\right) \le \mathbb{E}_{P_u}\left[S_{\mathrm{q}}\left(\rho_t\|\eta_t\right) - S_{\mathrm{q}}\left(\rho_u\|\rho_u^t\right)\right]. \tag{48}$$

Proof. Consider the mutual entropy $S(\sigma_t\|p_t\eta_t)$ introduced in Section 3.1 and apply to both states the channel Λ_u^t. By Theorem 1.1 and the definition (35) we get the inequality

$$\begin{aligned}\mathbb{E}_{P_t}\left[S_{\mathrm{q}}\left(\rho_t\|\eta_t\right)\right] &= S\left(\sigma_t\|p_t\eta_t\right) \ge S\left(\Lambda_u^t[\sigma_t]\|\Lambda_u^t[p_t\eta_t]\right) = S\left(\sigma_u\|p_t\sigma_u^t\right)\\ &= S\left(p_u\rho_u\|p_tp_u^t\rho_u^t\right) = \mathbb{E}_{P_u}\left[\mathrm{Tr}\left\{\rho_u\left(\ln p_u + \ln\rho_u - \ln\left(p_tp_u^t\right) - \ln\rho_u^t\right)\right\}\right]\\ &= S_{\mathrm{c}}\left(p_u\|p_tp_u^t\right) + \mathbb{E}_{P_u}\left[S_{\mathrm{q}}\left(\rho_u\|\rho_u^t\right)\right],\end{aligned}$$

and this gives (48). □

Apart from the different notations, Eq. (48) is the bound (29) in Ref. 5.

From Eqs. (36) and (38) we get immediately

$$\lim_{u\downarrow t}\frac{S_{\mathrm{c}}\left(p_u^t\|q_u^t\right)}{u-t} = 0\,. \tag{49}$$

Then, the second summand in the expression (47) of the increment of information becomes negligible with respect to the first when $u \downarrow t$. Therefore, from Lemma 4.1 we have the following theorem.

Theorem 4.1 (The bound on the derivative of $S_{\rm c}(p_t\|q_t)$). *The following bound holds:*

$$0 \leq \frac{\mathrm{d}}{\mathrm{d}t}\, S_{\rm c}(p_t\|q_t) \leq -\frac{\mathrm{d}}{\mathrm{d}u}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_u\|\rho_u^t)]\Big|_{u=t^+} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_t)] - \frac{\mathrm{d}}{\mathrm{d}u}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_u^t)]\Big|_{u=t^+}. \quad (50)$$

Remark 4.1. We already saw in Remark 3.1 that $\mathbb{E}_{P_T}[S_{\rm q}(\rho_t\|\eta_t)] = 0$ is equivalent to $\rho_t = \eta_t$, P_T-a.s.; but this implies $\rho_u = \rho_u^t$, P_T-a.s., because in this case these two quantities, which satisfy the same equation, have the same initial condition at time t. Therefore we have $\mathbb{E}_{P_T}[S_{\rm q}(\rho_u\|\rho_u^t)] = 0$, $\forall u \geq t$, and

$$\mathbb{E}_{P_T}[S_{\rm q}(\rho_t\|\eta_t)] = 0 \quad \Rightarrow \quad -\frac{\mathrm{d}}{\mathrm{d}u}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_u\|\rho_u^t)]\Big|_{u=t^+} = 0. \quad (51)$$

4.2. *Explicit computation of the bound*

All the derivatives can be elaborated and from Eq. (50) we get the following explicit form of the difference between the bound and the time derivative in which we are interested in.

Proposition 4.1. *By computation of all the terms appearing in Eq.* (50) *we get*

$$\begin{aligned}
0 \leq{}& \frac{\mathrm{d}}{\mathrm{d}t}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_t)] - \frac{\mathrm{d}}{\mathrm{d}u}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_u^t)]\Big|_{u=t^+} - \frac{\mathrm{d}}{\mathrm{d}t}\, S_{\rm c}(p_t\|q_t) \\
={}& \sum_k \mathbb{E}_{P_T}\left[\,\mathrm{Tr}\left\{J_k(t)\rho_t\left(\ln\rho_t - \ln\eta_t\right) - \mathcal{J}_k(t)[\rho_t]\left(\ln\mathcal{J}_k(t)[\rho_t] - \ln\mathcal{J}_k(t)[\eta_t]\right)\right\}\right] \\
&+ \frac{1}{2}\sum_j \mathbb{E}_{P_T}\left[\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{\eta_t}{u+\eta_t}\left(R_j(t)+R_j(t)^*\right)\frac{\eta_t}{u+\eta_t}\left(R_j(t)+R_j(t)^*\right)\right.\right. \\
&\qquad \left.\left. - \frac{\rho_t}{u+\rho_t}\left(R_j(t)+R_j(t)^*\right)\frac{\rho_t}{u+\rho_t}\left(R_j(t)+R_j(t)^*\right)\right\}\right] \\
&+ \sum_l \mathbb{E}_{P_T}\left[\mathrm{Tr}\left\{L_l(t)\eta_t\big[L_l(t)^*,\,\ln\eta_t\big] - L_l(t)\rho_t\big[L_l(t)^*,\,\ln\rho_t\big]\right\}\right]. \quad (52)
\end{aligned}$$

Proof. Let us start with the term $\frac{\mathrm{d}}{\mathrm{d}u}\, \mathbb{E}_{P_T}[S_{\rm q}(\rho_u^t)]\Big|_{u=t^+}$. By recalling that ρ_u^t satisfies in u the non-linear SDE with initial condition η_t at $u = t$ and that $\eta_t + \mathcal{L}(t)[\eta_t]\mathrm{d}t = \eta_{t+\mathrm{d}t}$, we get

$$\rho_{t+\mathrm{d}t}^t - \eta_{t+\mathrm{d}t} = \sum_j A_j(t)\mathrm{d}\check{W}_j(t) + \sum_k \left(\tau_k(t) - \eta_t\right)\left(\mathrm{d}N_k(t) - \nu_k(t)\mathrm{d}t\right),$$

where

$$A_j(t) := R_j(t)\eta_t + \eta_t R_j(t)^* - n_j(t)\eta_t\,, \qquad \tau_k(t) := \frac{1}{\nu_k(t)}\,\mathcal{J}_k(t)[\eta_t],$$

$$\mathrm{d}\check{W}_j(t) := \mathrm{d}W_j(t) - n_j(t)\mathrm{d}t\,.$$

By setting also

$$B(t) := -\frac{1}{2}\sum_k \{J_k(t) - \nu_k(t), \eta_t\} + \mathcal{L}_0(t)[\eta_t] + \mathcal{L}_1(t)[\eta_t],$$

we can write

$$\eta_{t+\mathrm{d}t} = \left(1 - \sum_k \nu_k(t)\mathrm{d}t\right)\eta_t + \sum_k \nu_k(t)\tau_k(t)\mathrm{d}t + B(t)\mathrm{d}t\,.$$

Moreover, by the properties of the increments of the counting processes, we have

$$\rho^t_{t+\mathrm{d}t}\mathrm{d}N_k(t) = \tau_k(t)\mathrm{d}N_k(t),$$

$$\left(1 - \sum_k \mathrm{d}N_k(t)\right)\rho^t_{t+\mathrm{d}t} = \left(1 - \sum_k \mathrm{d}N_k(t)\right)\left(\eta_t + B(t)\mathrm{d}t + \sum_j A_j(t)\,\mathrm{d}\check{W}_j(t)\right).$$

By putting these things all together and by using the rules of stochastic calculus, we get

$$\rho^t_{t+\mathrm{d}t}\ln\rho^t_{t+\mathrm{d}t} - \eta_t\ln\eta_t = \sum_k \left[\tau_k(t)\ln\tau_k(t) - \eta_t\ln\eta_t\right]\mathrm{d}N_k(t)$$

$$+\eta_t\left[\ln\left(\eta_t + B(t)\mathrm{d}t + \sum_j A_j(t)\,\mathrm{d}\check{W}_j(t)\right) - \ln\eta_t\right] + \sum_j A_j(t)\ln\eta_t\,\mathrm{d}\check{W}_j(t)$$

$$+ B(t)\mathrm{d}t\,\ln\eta_t + \sum_j A_j(t)\left[\ln\left(\eta_t + A_j(t)\,\mathrm{d}\check{W}_j(t)\right) - \ln\eta_t\right]\mathrm{d}\check{W}_j(t).$$

It exists a nearly obvious and very useful integral representation of the logarithm of an operator ([6] p. 51):

$$\ln A = \int_0^{+\infty}\left(\frac{1}{1+t} - \frac{1}{t+A}\right)\mathrm{d}t.$$

By iterating this formula we get also

$$\ln(A+B)-\ln A=\int_0^{+\infty}\frac{1}{t+A}\,B\,\frac{1}{t+A+B}\,\mathrm{d}t$$
$$=\int_0^{+\infty}\frac{1}{t+A}\,B\,\frac{1}{t+A}\left(1-B\,\frac{1}{t+A+B}\right)\mathrm{d}t\,.$$

These two formulae and stochastic calculus rules allow to write

$$\mathbb{E}_{P_T}\left[\mathrm{Tr}\left\{\rho^t_{t+\mathrm{d}t}\ln\rho^t_{t+\mathrm{d}t}-\eta_t\ln\eta_t\right\}\right]=\sum_k\left[S_{\mathrm{q}}(\eta_t)-S_{\mathrm{q}}\big(\tau_k(t)\big)\right]\nu_k(t)\mathrm{d}t$$
$$-\,\mathrm{d}t\sum_j\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{\eta_t}{(u+\eta_t)^2}\,A_j(t)\,\frac{1}{u+\eta_t}\,A_j(t)\right\}$$
$$+\,\mathrm{d}t\sum_j\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{1}{u+\eta_t}\,A_j(t)\,\frac{1}{u+\eta_t}\,A_j(t)\right\}$$
$$+\,\mathrm{d}t\,\mathrm{Tr}\left\{B(t)\left(\ln\eta_t+\int_0^{+\infty}\frac{\eta_t}{(u+\eta_t)^2}\,\mathrm{d}u\right)\right\}.$$

By computing the integral we get

$$\mathrm{Tr}\left\{B(t)\left(\ln\eta_t+\int_0^{+\infty}\frac{\eta_t}{(u+\eta_t)^2}\,\mathrm{d}u\right)\right\}=\mathrm{Tr}\left\{B(t)\left(\ln\eta_t+\mathbb{1}\right)\right\}$$
$$=\mathrm{Tr}\left\{B(t)\ln\eta_t\right\}=\sum_k\mathrm{Tr}\left\{\left[\nu_k(t)-J_k(t)\right]\eta_t\ln\eta_t\right\}$$
$$+\sum_j\mathrm{Tr}\left\{\left[R_j(t)\eta_tR_j(t)^*-R_j(t)^*R_j(t)\eta_t\right]\ln\eta_t\right\}$$
$$+\sum_l\mathrm{Tr}\left\{\left[L_l(t)\eta_tL_l(t)^*-L_l(t)^*L_l(t)\eta_t\right]\ln\eta_t\right\}$$

and by using the integration by parts with $\frac{1}{(u+\eta_t)^2}=-\frac{\mathrm{d}}{\mathrm{d}u}\frac{1}{u+\eta_t}$ we have also

$$\sum_j\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{1}{u+\eta_t}\,A_j(t)\,\frac{1}{u+\eta_t}\,A_j(t)-\frac{\eta_t}{(u+\eta_t)^2}\,A_j(t)\,\frac{1}{u+\eta_t}\,A_j(t)\right\}$$
$$=\sum_j\int_0^{+\infty}\mathrm{d}u\;\mathrm{Tr}\left\{A_j(t)\,\frac{1}{(u+\eta_t)^2}\,A_j(t)\,\frac{u}{u+\eta_t}\right\}$$
$$=\sum_j\int_0^{+\infty}\mathrm{d}u\;\mathrm{Tr}\left\{A_j(t)\,\frac{\eta_t}{(u+\eta_t)^2}\,A_j(t)\,\frac{1}{u+\eta_t}\right\}$$
$$=\frac{1}{2}\sum_j\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{1}{u+\eta_t}\,A_j(t)\,\frac{1}{u+\eta_t}\,A_j(t)\right\}.$$

From the previous formulae we get

$$
\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}u}\,\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_u^t)]\Big|_{u=t^+} &= \sum_k \mathrm{Tr}\left\{\mathcal{J}_k(t)[\eta_t]\ln\frac{\mathcal{J}_k(t)[\eta_t]}{\nu_k(t)} - J_k(t)\eta_t\ln\eta_t\right\} \\
&\quad + \sum_j \mathrm{Tr}\left\{R_j(t)\eta_t\big[R_j(t)^*,\ln\eta_t\big]\right\} + \sum_l \mathrm{Tr}\left\{L_l(t)\eta_t\big[L_l(t)^*,\ln\eta_t\big]\right\} \\
+\frac{1}{2}\sum_j\Bigg(\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\Bigg\{ & R_j(t)^*\frac{\eta_t}{u+\eta_t}R_j(t)^*\frac{\eta_t}{u+\eta_t} + \frac{\eta_t}{u+\eta_t}R_j(t)\frac{\eta_t}{u+\eta_t}R_j(t) \\
& + \frac{2}{u+\eta_t}R_j(t)\frac{{\eta_t}^2}{u+\eta_t}R_j(t)^*\Bigg\} - n_j(t)^2\Bigg).
\end{aligned}
$$

But we have

$$
\begin{aligned}
&\mathrm{Tr}\left\{R_j(t)\eta_t\big[R_j(t)^*,\ln\eta_t\big]\right\} + \int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{1}{u+\eta_t}R_j(t)\frac{{\eta_t}^2}{u+\eta_t}R_j(t)^*\right\} \\
&\quad = \int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{1}{u+\eta_t}R_j(t)\frac{{\eta_t}^2}{u+\eta_t}R_j(t)^* - \frac{1}{u+\eta_t}R_j(t)\eta_t R_j(t)^*\right. \\
&\qquad\qquad \left. + \frac{1}{u+\eta_t}R_j(t)\frac{\eta_t}{u+\eta_t}R_j(t)^*\right\} \\
&= \int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{-\frac{u}{u+\eta_t}R_j(t)\frac{\eta_t}{u+\eta_t}R_j(t)^* + \frac{1}{u+\eta_t}R_j(t)\frac{\eta_t}{u+\eta_t}R_j(t)^*\right\} \\
&\quad = \int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{\eta_t}{u+\eta_t}R_j(t)\frac{\eta_t}{u+\eta_t}R_j(t)^*\right\}.
\end{aligned}
$$

Then, we have the final expression

$$
\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}u}\,\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_u^t)]\Big|_{u=t^+} &= \sum_k \mathrm{Tr}\left\{\mathcal{J}_k(t)[\eta_t]\ln\frac{\mathcal{J}_k(t)[\eta_t]}{\nu_k(t)} - J_k(t)\eta_t\ln\eta_t\right\} \\
+\frac{1}{2}\sum_j\Bigg(\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\Bigg\{ & \frac{\eta_t}{u+\eta_t}\left(R_j(t)+R_j(t)^*\right)\frac{\eta_t}{u+\eta_t}\left(R_j(t)+R_j(t)^*\right)\Bigg\} \\
& - n_j(t)^2\Bigg) + \sum_l \mathrm{Tr}\left\{L_l(t)\eta_t\big[L_l(t)^*,\ln\eta_t\big]\right\}. \quad (53)
\end{aligned}
$$

Analogously we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_t)] = -\sum_k \mathbb{E}_{P_T}\left[\mathrm{Tr}\left\{\mathcal{J}_k(t)[\rho_t]\ln\frac{\mathcal{J}_k(t)[\rho_t]}{\mu_k(t)} - J_k(t)\rho_t\ln\rho_t\right\}\right]$$
$$-\frac{1}{2}\sum_j \mathbb{E}_{P_T}\left[\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{\frac{\rho_t}{u+\rho_t}\left(R_j(t)+R_j(t)^*\right)\frac{\rho_t}{u+\rho_t}\left(R_j(t)+R_j(t)^*\right)\right\}\right.$$
$$\left. - m_j(t)^2\right] - \sum_l \mathbb{E}_{P_T}\left[\mathrm{Tr}\left\{L_l(t)\rho_t\left[L_l(t)^*,\,\ln\rho_t\right]\right\}\right]. \quad (54)$$

By (53), (54), (43) we get the statement of the Proposition. □

Corollary 4.1. *A sufficient condition to have the equality in the main bound*

$$\frac{\mathrm{d}}{\mathrm{d}t}\,S_{\mathrm{c}}(p_t\|q_t) = \frac{\mathrm{d}}{\mathrm{d}t}\,\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_t)] - \frac{\mathrm{d}}{\mathrm{d}u}\,\mathbb{E}_{P_T}[S_{\mathrm{q}}(\rho_u^t)]\Big|_{u=t^+} \quad (55)$$

is to have P_T-a.s. in ω ($T \geq t$), $\forall r,k,j,l$,

$$[V_k^r(t),\rho_t(\omega)] = 0\,,\quad [R_j(t)+R_j(t)^*,\rho_t(\omega)] = 0\,,\quad [L_l(t),\rho_t(\omega)] = 0\,. \quad (56)$$

Proof. By the commutation relations (56) we get

$$J_k(t)\rho_t\left(\ln\rho_t - \ln\eta_t\right) - \mathcal{J}_k(t)[\rho_t]\left(\ln\mathcal{J}_k(t)[\rho_t] - \ln\mathcal{J}_k(t)[\eta_t]\right)$$
$$= J_k(t)\rho_t\left(\ln\rho_t - \ln\eta_t - \ln J_k(t)\rho_t + \ln J_k(t)\eta_t\right) = 0$$

and the first term in Eq. (52) vanishes.

By Eq. (56) also the last term in Eq. (52) is zero because it explicitly involves vanishing commutators.

Finally, let us consider one of the terms in the j-sum. We have

$$\mathbb{E}_{P_T}\left[\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{-R_j(t)^*\frac{\rho_t}{u+\rho_t}R_j(t)^*\frac{\rho_t}{u+\rho_t}\right\}\right]$$
$$= -\,\mathbb{E}_{P_T}\left[\int_0^{+\infty}\mathrm{d}u\,\mathrm{Tr}\left\{R_j(t)^{*2}\frac{{\rho_t}^2}{(u+\rho_t)^2}\right\}\right]$$
$$= -\,\mathbb{E}_{P_T}\left[\mathrm{Tr}\left\{R_j(t)^{*2}\rho_t\right\}\right] = -\,\mathrm{Tr}\left\{R_j(t)^{*2}\eta_t\right\}.$$

The opposite result comes out from the corresponding term with η and the j-sum vanishes too. □

Acknowledgments

Work supported by the *European Community's Human Potential Programme* under contract HPRN-CT-2002-00279, QP-Applications.

Bibliography

1. A. Barchielli, L. Lanz, G. M. Prosperi, *Statistics of continuous trajectories in quantum mechanics: Operation valued stochastic processes*, Found. Phys. **13** (1983) 779–812.
2. A. Barchielli, V. P. Belavkin, *Measurements continuous in time and a posteriori states in quantum mechanics*, J. Phys. A: Math. Gen. **24** (1991) 1495–1514.
3. A. Barchielli, *Entropy and information gain in quantum continual measurements*, in P. Tombesi and O. Hirota (eds.), *Quantum Communication, Computing, and Measurement 3* (Kluwer, New York, 2001) pp. 49–57; quant-ph/0012115.
4. A. Barchielli, G. Lupieri, *Instrumental processes, entropies, information in quantum continual measurements*, in O. Hirota (ed.), *Quantum Information, Statistics, Probability* (Rinton, Princeton, 2004) pp. 30–43; Quantum Inform. Compu. **4** (2004) 437–449; quant-ph/0401114.
5. A. Barchielli, G. Lupieri, *Entropic bounds and continual measurements*, to appear in Quantum Probability Series QP-PQ, World Scientific; quant-ph/0511090.
6. M. Ohya, D. Petz, *Quantum Entropy and Its Use* (Springer, Berlin, 1993).
7. J. Dixmier, *Les Algèbres d'Opérateurs dans l'Espace Hilbertien* (Gauthier-Villars, Paris, 1957).
8. E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
9. M. Ozawa, *Quantum measuring processes of continuous observables*, J. Math. Phys. **25** (1984) 79–87.
10. A. Barchielli, G. Lupieri, *Instruments and mutual entropies in quantum information theory*, to appear in Banach Center Publications; quant-ph/0412116.
11. M. Ozawa, *Conditional probability and a posteriori states in quantum mechanics*, Publ. R.I.M.S. Kyoto Univ. **21** (1985) 279–295.
12. A. S. Holevo, *Some estimates for the amount of information transmittable by a quantum communication channel*, Probl. Inform. Trans. **9** no. 3 (1973) 177–183 (Engl. transl.: 1975).
13. A. S. Holevo, M. E. Shirokov, *Continuous ensembles and the χ-capacity of infinite-dimensional channels*, quant-ph/0408176 (2004).
14. A. Barchielli, G. Lupieri, *Quantum measurements and entropic bounds on information transmission*, Quantum Information and Computation **6** (2006) 16–45; quant-ph/0505090.
15. A. Barchielli, A. S. Holevo, G. Lupieri, *An analogue of Hunt's representation theorem in quantum probability*, J. Theor. Probab. **6** (1993) 231–265.
16. A. Barchielli, A. S. Holevo, *Constructing quantum measurement processes via classical stochastic calculus*, Stoch. Proc. Appl. **58** (1995) 293–317.
17. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

NOISY QUTRIT CHANNELS

A. CHĘCIŃSKA*

Instytut Fizyki Teoretycznej, Uniwersytet Warszawski,
Warszawa 00-681, Poland
** E-mail: achec@fuw.edu.pl*
www.fuw.edu.pl

K. WÓDKIEWICZ†

Instytut Fizyki Teoretycznej, Uniwersytet Warszawski,
Warszawa 00-681, Poland
Department of Physics and Astronomy, University of New Mexico,
800 Yale Boulevard NE, Albuquerque, New Mexico 87131, USA
† E-mail: wodkiew@fuw.edu.pl
http://www.fuw.edu.pl, http://panda.unm.edu/

We present an analysis of spontaneous emission in a 3-level atom as an example of a qutrit state under the action of noisy quantum channels. We choose a 3-level atom with V–configuration to be the qutrit state. Gell-Mann matrices and a generalized Bloch vector (8-dimensional) are used to describe the qutrit density operator. Using the time-evolution equations of atomic variables we find the Kraus representation of spontaneous emission quantum channel (SE channel). Furthermore, we consider a generalized Werner state of two qutrits and investigate the separability condition. We give similar analysis of spontaneous emission for qubit channels. The influence of spontaneous emission on the separability of Werner states for qutrit and qubit states is compared.

Keywords: qutrit; quantum channel; spontaneous emission; Werner state; separability.

1. Introduction

Quantum engineering, teleportation and the idea of constructing quantum computers were brought into focus of recent, widespread, scientific research. Consequently, investigation of quantum phenomena, such as entanglement, has been intensified[1,2]. One of the possible ways is to investigate atomic systems, their entanglement and behaviour under action of quantum channels. Since these might be useful, it is desirable to quantify the influence of channels on state separability and entanglement. The state and channel are

very general concepts, hence we would like to focus on particular examples of a spontaneous emission channel involving a quantum state described by a 3-level atom. Spontaneous emission is a proccess that might destroy mutual entanglement. The majority of papers have been devoted to qubit states. It is the purpose of this paper to concentrate on a qutrit state, which has a physical realization in form of a three-level atom in the presence of spontaneous emission. Such a system is an example of qutrit channel with noise. We compare the qutrit case with the qubit, in other words, three-level and two-level atoms under the action of spontaneous emission channel.
The paper is structured as follows: we present description of (single) qutrit states (analogy to Bloch formalism), we choose 3-level atoms with V–configuration to be the qutrit state; equations for atomic evolution in presence of spontaneous emission are given, we present channel mathematics, focusing on Kraus representation; Kraus operators for spontaneous emission channel (SE channel) are evaluated; analysis of two qutrit state is brought into discussion and generalized Werner state is described together with separability condition; the action of spontaneous emission channel on two qutrit state is investigated; we give Kraus representation for analogous qubit channel and analyze Werner state of two qubits; we compare both cases: qutrit and qubit separability behaviuor under influence of spontaneous emission and recognize that, depending on channel characteristic, the qutrit systems can preserve entanglement longer.

2. Qutrit state

The concept of a qubit[3,4] *i.e.*, a quantum state living in two dimensional Hilbert space is used as a basic building block of Quantum Information. Within the framework of atomic physics two-level atom is the simplest physical realization of a qubit[3]. Qubit and qubit channels have been investigated comprehensively[3,4,12] and generalization to N-dimensional cases involving qudits has been studied[13–15], though less exhaustively. From physical point of view the use of more complex atomic structures might be advantageous[10], therefore three-level atoms, qutrits, might deserve more interest and studies. To give their mathematical description, we note that the Bloch formalism of a two-level atom is based on the $SU(2)$ generators given by Pauli matrices, as a basis for qubit density operator[5,6] :

$$\rho_{qubit} = \frac{1}{2}(\mathbb{I} + \vec{b} \cdot \vec{\sigma}), \tag{1}$$

where $\vec{b}$ is a three dimensional (real) Bloch vector[13]. The mathematical description of a qutrit density operator involves in a natural way the $SU(3)$

generators, called the Gell-Mann matrices λ_i[6,16] :

$$\rho = \frac{1}{3}(\mathbb{I} + \sqrt{3}\vec{n} \cdot \vec{\lambda}), \tag{2}$$

where $\vec{n}$ is a real eight dimensional generalized Bloch vector.
Qutrit states belong to three-dimensional complex Hilbert space $\mathcal{H}^{(3)}$; pure qutrit states correspond to vectors that satisfy[16] :

$$\vec{n} \cdot \vec{n} = 1, \ \vec{n} * \vec{n} = \vec{n}, \tag{3}$$

where $(\vec{A} * \vec{B})_k \equiv d_{klm} A_l B_m$, with d_{klm} being a totally symmetric tensor associated with the $SU(3)$[6].
These two conditions define a generalized Bloch sphere for qutrits, in analogy to Bloch qubit sphere. In general, pure state can be parameterized as follows (in the atomic basis $|1\rangle, |2\rangle, |3\rangle$):

$$|\Psi\rangle = \sin\frac{\xi}{2}\cos\frac{\theta}{2}|1\rangle + e^{\imath\phi_{12}}\sin\frac{\xi}{2}\sin\frac{\theta}{2}|2\rangle + e^{\imath\phi_{13}}\cos\frac{\xi}{2}|3\rangle. \tag{4}$$

Orthogonal states in $\mathcal{H}^{(3)}$ do not correspond to opposite points on $\mathcal{S}^7$ (the seven-dimensional unit sphere in $\mathcal{R}^8$), but to points of maximum opening angle of $\frac{2\pi}{3}$. Distribution of points on $\mathcal{S}^7$ that represent physical states, the generalized Bloch sphere, is highly nontrivial[5,6,17].

3. 3-level atoms

We will consider a particular physical realization of a qutrit state, namely 3-level atoms. There are three configurations of 3-level atoms[18] - we choose the so called V–configuration in which the only allowed transitions $|2\rangle \longrightarrow |1\rangle$ and $|3\rangle \longrightarrow |1\rangle$ are depicted on *Fig. 1*:

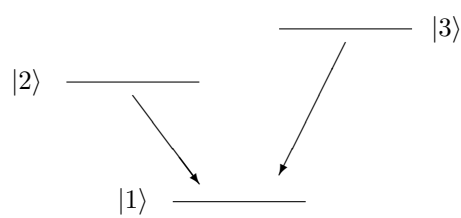


Fig. 1. Transitions allowed in 3-level atom with V–configuration.

The Bloch vector $\bar{n}$ can be expressed by atomic populations and dipole transitions:

$$\begin{array}{llll} n_1 = & \frac{\sqrt{3}}{2}(d_1^* + d_1), & n_2 = \frac{\sqrt{3}\imath}{2}(d_1^* - d_1), \\ n_4 = & \frac{\sqrt{3}}{2}(d_2^* + d_2), & n_5 = \frac{\sqrt{3}\imath}{2}(d_2^* - d_2), \\ n_6 = & \frac{\sqrt{3}}{2}(d_3^* + d_3), & n_7 = \frac{\sqrt{3}\imath}{2}(d_3^* - d_3), \\ n_3 = & \frac{\sqrt{3}}{2}(1 - 2p_2 - p_3), & n_8 = \frac{1}{2}(1 - 3p_3). \end{array} \tag{5}$$

(recall that $d_1 : |1\rangle \to |2\rangle,\ d_2 : |1\rangle \to |3\rangle,\ d_3 : |2\rangle \to |3\rangle$.)

Evolution in the presence of spontaneous emission

Spontaneous emission is a dissipative process, in which the atom is coupled to the electromagnetic vacuum. The dissipative evolution of the atomic variables in the presence of spontaneous emission is characterized by two Einstein coefficients A_1 and A_2 that describe the irreversible depopulation from excited states. This corresponds to the following Bloch equations with time independent coefficients:

$$\frac{d}{dt}\vec{n}(t) = \mathcal{M}\vec{n}(0) + \vec{m}_0, \tag{6}$$

where $\mathcal{M}$ is a (almost) diagonal matrix with only one non-diagonal entry and $\vec{m}_0$ is a translation.
The solution to this equation is of the form:

$$\vec{n}(t) = \mathcal{T} \cdot \vec{n}(0) + \vec{n}_0(t), \tag{7}$$

$$\mathcal{T} = \begin{pmatrix} e^{-\frac{A_1 t}{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{-\frac{A_1 t}{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{-A_1 t} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}}(e^{-A_2 t} - e^{-A_1 t}) \\ 0 & 0 & 0 & e^{-\frac{A_2 t}{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{-\frac{A_2 t}{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{-\frac{(A_1+A_2)t}{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & e^{-\frac{(A_1+A_2)t}{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & e^{-A_2 t} \end{pmatrix}$$

where $\vec{n}_0(t)$ is a translation dependant on time:

$$\vec{n}_0(t)=\begin{pmatrix} 0 \\ 0 \\ \frac{1}{2\sqrt{3}}(3-e^{-A_2t}-2e^{-A_1t}) \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{2}(1-e^{-A_2t}) \end{pmatrix}$$

Hence, density operator representing the state of an atom in presence of spontaneous emission is of the form:

$$\rho(t) = \frac{1}{3}\left(\mathbb{I} + \sqrt{3}(\mathcal{T}\vec{n}(0) + \vec{n}_0(t)) \cdot \vec{\lambda}\right). \tag{8}$$

4. Completely positive maps and Kraus representation

Channel acting on a density operator maps density operators into density operators[3,12,19] :

$$\Phi : \rho_{in} \ \mapsto \rho_{out} \tag{9}$$

It is well known that the channel transformation Φ is described by a completely positive map (CPM)[12]. The simplest way to describe a channel is by means of an operator-sum representation[19,20] :

$$\rho_{out} = \Phi(\rho_{in}) = \sum_i \mathcal{K}_i \rho_{in} \mathcal{K}_i^\dagger, \tag{10}$$

where $\mathcal{K}_i$ are Kraus operators that satisfy normalization condition:

$$\sum_i \mathcal{K}_i^\dagger \mathcal{K}_i = \mathbb{I}. \tag{11}$$

From the Bloch equations, we can calculate the action of spontaneous emission channel (*SE* channel) on the V-atom in terms of the operator-sum representation. In this case the set of corresponding Kraus operators is as follows:

$$\mathcal{K}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\frac{A_1 t}{2}} & 0 \\ 0 & 0 & e^{-\frac{A_2 t}{2}} \end{pmatrix}, \quad \mathcal{K}_1 = \begin{pmatrix} 0 & \sqrt{1-e^{-A_1 t}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{K}_2 = \begin{pmatrix} 0 & 0 & \sqrt{1-e^{-A_2 t}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

5. Influence of SE channel on state separability

5.1. *Generalized Werner state for two qutrits*

We generalize our discussion to the situation in which we have two qutrits. The generalized Werner state describing two qutrits labelled by A and B is of the form[5] :

$$\rho_\epsilon = \frac{1-\epsilon}{9}\mathbb{I}^A \otimes \mathbb{I}^B + \epsilon|\Psi^{AB}\rangle\langle\Psi^{AB}|, \tag{12}$$

where $0 \leq \epsilon \leq 1$. The state is a convex combination of a maximally mixed state and a pure state. The aim is to characterize the values of parameter ϵ for which ρ_ϵ is separable[8,9] (meaning it can be represented as an ensamble of product states).
We will consider only Werner states that consist of a specific pure state, namely:

$$|\Psi\rangle = \frac{1}{\sqrt{3}}(|1^A\rangle \otimes |1^B\rangle + |2^A\rangle \otimes |2^B\rangle + |3^A\rangle \otimes |3^B\rangle). \tag{13}$$

To investigate the separability condition on ϵ we follow the discussion given in[5]. Therefore we represent density operator ρ_ϵ in the basis of the Gell-Mann matrices $\{\lambda_\alpha\}_{\alpha=0}^8$ enriched by $\lambda_0 = \sqrt{\frac{2}{3}}\mathbb{I}$ (with $tr(\lambda_\alpha\lambda_\beta) = 2\delta_{\alpha\beta}$):

$$\rho_\epsilon = \frac{1}{9}c_{\alpha\beta}\lambda_\alpha^A \otimes \lambda_\beta^B,\ c_{\alpha\beta} = \frac{9}{4}tr\{\rho\lambda_\alpha^A \otimes \lambda_\beta^B\}. \tag{14}$$

In this form, the state is characterized by the $c_{\alpha\beta}$ coefficients ($\alpha, \beta \in \{0, ..., 8\}$). It turns out[5] that the condition for the state to be separable is of the form:

$$4\epsilon = \frac{1}{3}\sum_{j=1}^{8}|c_{jj}| \leq 1. \tag{15}$$

Hence we obtain:

$$\epsilon \leq \frac{1}{4}. \tag{16}$$

5.2. *SE channel action on qutrits*

Action of the channel Φ on ρ_ϵ changes the coefficients:

$$\Phi : c_{\alpha\beta} \ \mapsto c_{\alpha\beta}(t), \tag{17}$$

Therefore, the condition on the ϵ to produce a separable state becomes time dependant.
Consider the channel that alters only one subsystem (for instance A):

$$\Phi_1(\rho_\epsilon) = \sum_{i=0}^{2} (\mathcal{K}_i^A \otimes \mathbb{I}^B)\rho_\epsilon(\mathcal{K}_i^A \otimes \mathbb{I}^B)^\dagger = \frac{1}{9} c_{\alpha\beta}(t)\lambda_\alpha \otimes \lambda_\beta. \tag{18}$$

In this case the separability condition is:

$$s_{\mathrm{qt}}(t) \equiv \frac{\epsilon}{8}\left(2e^{-\frac{1}{2}A_1 t} + 2e^{-\frac{1}{2}A_2 t} + 2e^{-\frac{1}{2}(A_1+A_2)t} + e^{-A_1 t} + e^{-A_2 t}\right) \leq \frac{1}{4}, \tag{19}$$

where obviously, for $t = 0$ we have:

$$s_{\mathrm{qt}}(0) = \epsilon \leq \frac{1}{4}. \tag{20}$$

The function $s_{\mathrm{qt}}(t)$ is shown on *Fig. 2*. It is clear, that even initially maximally entangled state (meaning $\epsilon = 1$) becomes separable eventually. Two qutrit SE channels with various parameters A_1, A_2 are compared.

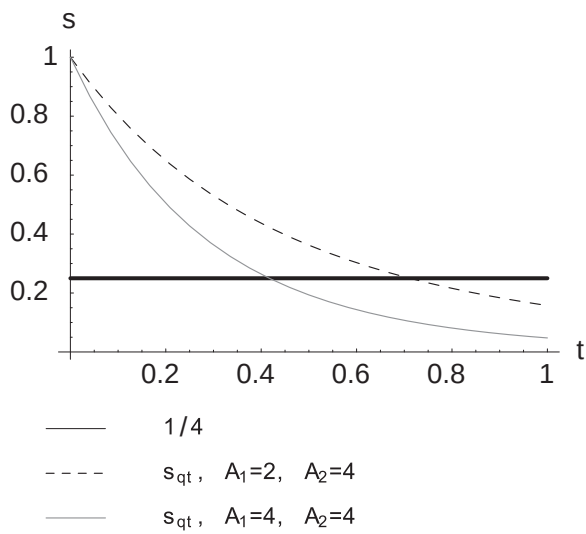


Fig. 2. The function $s_{\mathrm{qt}}(t)$ (*Eq. 19*) for $\epsilon = 1$. Region below $s = \frac{1}{4}$ corresponds to separable states.

In the following we shall compare the decoherence of a qutrit with a similar decoherence for a qubit.

5.3. *SE channel and Werner state for qubits*

The Kraus representation for spontaneous emission channel for qubits is as follows (A is the Einstein coefficient):

$$\mathcal{K}_0 = \begin{pmatrix} 1 & 0 \\ 0 & e^{-\frac{At}{2}} \end{pmatrix}, \ \mathcal{K}_1 = \sqrt{1-e^{-At}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{21}$$

Whereas the Werner state describing two non-interacting qubits is of the form[21] ($\alpha, \beta \in \{0,1,2,3\}$, $\sigma_0 = \mathbb{I}$):

$$\rho_\epsilon = \frac{1-\epsilon}{4}\mathbb{I}^A \otimes \mathbb{I}^B + \epsilon |\Psi^{AB}\rangle\langle\Psi^{AB}| = \frac{1}{4} d_{\alpha\beta} \sigma_\alpha^A \otimes \sigma_\beta^B, \tag{22}$$

where:

$$|\Psi^{AB}\rangle = \frac{1}{\sqrt{2}}(|1\rangle^A \otimes |2\rangle^B + |2\rangle^A \otimes |1\rangle^B). \tag{23}$$

The state ρ_ϵ is separable when[5,8] :

$$\epsilon \leq \frac{1}{3}. \tag{24}$$

5.4. *SE channel altering qubit separability*

In analogy to what has been done before, we consider SE channel altering only one qubit, with (two) Kraus operators of the form: $\mathcal{K}_i^A \otimes \mathbb{I}^B$. Initial qubit state has the form:

$$\rho_\epsilon = \frac{1}{4} d_{\alpha\beta} \sigma_\alpha^A \otimes \sigma_\beta^B. \tag{25}$$

Action of the channel changes coefficients of expansion:

$$\Phi(\rho_\epsilon) = \rho_\epsilon(t) = \sum_{i=1,2} \mathcal{K}_i^A \otimes \mathbb{I}^B \rho_\epsilon (\mathcal{K}_i^A \otimes \mathbb{I}^B)^\dagger = \frac{1}{4} d_{\alpha\beta}(t) \sigma_\alpha^A \otimes \sigma_\beta^B. \tag{26}$$

And the separability condition leads to the inequality:

$$s_{\mathrm{qb}}(t) \equiv \frac{\epsilon}{3}(2e^{-\frac{At}{2}} + e^{-At}) \leq \frac{1}{3}, \tag{27}$$

with initial condition: $s_{\mathrm{qb}}(0) = \epsilon$.
Function $s_{qb}(t)$ is depicted on *Fig. 3*. Again, it is clear that any state becomes eventually separable. Time in which maximally entangled state becomes separable can be calculated and depends on the value of parameter A.

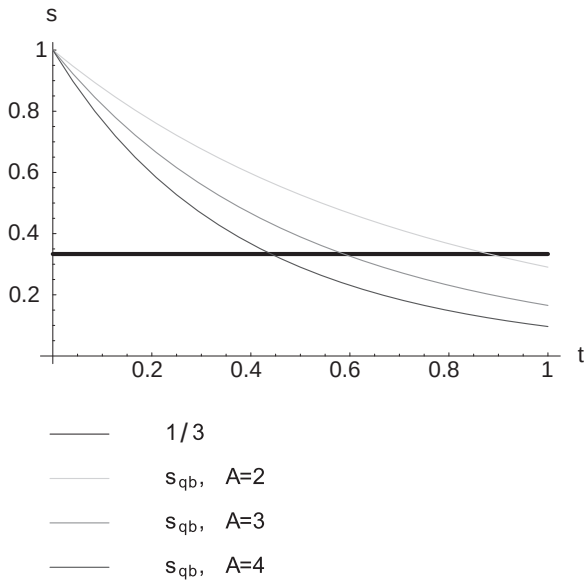


Fig. 3. Function $s_{qb}(t)$ (*Eq. 27*) for $\epsilon = 1$. Region below $s = \frac{1}{3}$ corresponds to separable states.

5.5. *Comparison of qubit and qutrit states under the action of SE channels*

Knowing how spontaneous channel acts on both qutrit and qubit states we can compare these two cases in order to state whether qutrit or qubit Werner states preserve entanglement longer. In *Fig. 4* we show the comparison of the SE channel action. Points of intersection of $s_{\mathrm{qt}} = \frac{1}{4}$ and $s_{\mathrm{qb}} = \frac{1}{3}$, associated with time needed to reach separability, are crucial to this analysis. The proper choice of parameters A, A_1, A_2 can lead us to the case in which qutrit entanglement is stronger or *vice versa*. The function $s_{\mathrm{qt}}(t)$ is symmetric with respect to the change $A_1 \leftrightarrow A_2$, the relative value of A parameter (with respect to A_1, A_2) in $s_{\mathrm{qb}}(t)$ influences the result of qubit–qutrit comparison.

6. Summary

We have presented an example of a qutrit state, namely a 3-level atom in the V–configuration, and its evolution under the action of spontaneous emission channel. Separability of two qutrit states is, obviously, influenced by spontaneous emission. This influence is governed by channel parameters.

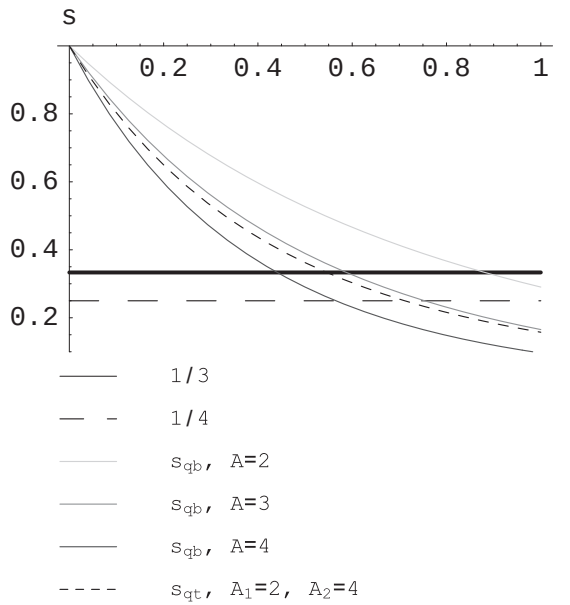


Fig. 4. Comparison of functions $s_{\mathrm{qb}}(t)$ and $s_{\mathrm{qt}}(t)$ (*Eqs. 27,19*) for initial condition $\epsilon = 1$.

As a consequence, Werner qutrit states can preserve entanglement longer than Werner qubit states - depending on relative values of qubit and qutrit channel parameters. This result might be of some experimental importance when it comes to use of N-level atoms and multipartite entanglement. We plan to investigate further examples of qutrit channels and their influence on state separability. We aim as well at general description of qutrit channels with respect to complete positivity.

Acknowledgements

This paper was supported by a Polish MNiSW grant No. 1P03B13730.

Bibliography

1. M. Razavi, J. H. Shapiro, *Phys.Rev. A* **73**, 042303 (2006).
2. P. P. Munhoz, *et al.*, *J.Phys.B: At. Mol. Opt. Phys.* **38**, 3875-3884 (2005).
3. S. Daffer, J. McIver, K. Wodkiewicz, *Phys.Rev. A*, **67**, 062312 (2003).
4. M. A. Nielsen, I. L. Chuang, in *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

5. C. Caves, G. Milburn, *Optics Communications*, **179**, (2000).
6. Arvind, K. S. Mallesh, N. Mukunda, *J.Phys. A: Math.Gen.* **30**, 2417-2431 (1997).
7. K. G. H. Vollbrecht, R. F. Werner, *Phys.Rev. A* **64**, 062307 (2001).
8. A. Peres, *Phys. Rev. Lett.* **77**, 1413 (1996).
9. M. Horodecki, P. Horodecki, R. Horodecki, *Phys. Lett. A* **223**, 1–8 (1996).
10. D. Kaszlikowski, D. Gosal, E. J. Ling, L. C. Kwek, M.Żukowski, *Phys. Rev. A* **66**, 032103 (2002).
11. Guo-Qiang Zhu, Xuean Zhao, You-Quan Li, *Eur. Phys. J. B* **44**, 359-364 (2005).
12. M. B. Ruskai, S. Szarek, E. Werner, *Lin. Alg. Appl.* **347**, 159 (2002).
13. G. Kimura, *Phys. Lett. A* **314**, 339 (2003).
14. H. F. Hofmann, S. Takeuchi, *Phys. Rev. A* **68**, 032103 (2003).
15. A. Cabello, *Phys. Rev. Lett.* **89**, 100402 (2002).
16. A. B. Klimov, L. L. Sanchez–Soto, H. de Guise, G. Bjork, *J. Phys. A: Math. Gen.* **37** (2004).
17. H. Sommers, K. Życzkowski, *J. Phys. A* **36**, 10083-10100 (2003).
18. J. H. Eberly, F. T. Hioe, *Phys.Rev. A*, **25**, 4 (1982).
19. M. A. Nielsen, B. Schumacher, *Phys.Rev. A* **54**, 2629 (1996).
20. K. Kraus, in *States, Effects, and Operations*, Lecture Notes in Physics (Springer-Verlag, 1983).
21. R. F. Werner, *Phys.Rev. A* **40**, (1989)

ADDITIVITY OF ENTANGLED CHANNEL CAPACITY GIVEN QUANTUM INPUT STATES

VIACHESLAV P. BELAVKIN and XIANHUA DAI

School of Mathematical Sciences, University of Nottingham
University Park, Nottingham NG7 2RD, UK

An elementary introduction into algebraic approach to unified quantum information theory and operational approach to quantum entanglement as generalized encoding is given. After introducing compound quantum state and two types of informational divergences, namely, Araki-Umegaki (a-type) and of Belavkin-Staszewski (b-type) quantum relative entropic information, this paper treats two types of quantum mutual information via entanglement and defines two types of corresponding quantum channel capacities as the supremum via the generalized encodings. It proves the additivity property of quantum channel capacities via entanglement, which extends the earlier results of V. P. Belavkin to products of arbitrary quantum channels for quantum relative entropy of any type.

1. Introduction

Unlike classical channels, quantum channels can have several different capacities (e.g. for sending classical information or quantum information, one-way or two-way communication, prior or via entanglement, etc.). Well, the problem of characterizing in general the capacity of a noisy quantum channel is unsolved although several attempts have been made to define a quantum analog of Shannon mutual information (see the conceptions of coherent information[1] or von Neumann mutual entropy[2,3,4]). Unfortunately most of these attempts do not give a satisfactory solution because the defined quantities fail to preserve such important property of classical informational capacity as additivity and some do not have even monotonicity property. This paper is based on the operational entanglement approach to quantum channel capacity suggested in[5–8] which is free of the above difficulties due to the enlargement of the class of input encodings, including the encodings via entanglement for one-way communication.

Quantum entanglement is a uniquely quantum mechanical resource that

plays a key role, along with the celebrating paper[9] of Einstein-Podolsky-Rosen, in many of the most interesting applications of quantum information and quantum computation, such as, Quantum entanglement is extensively used in teleporting an unknown quantum state via dual classic and Einstein-Podolsky-Rosen channels in the subject of quantum teleportation[10] quantum cryptography was investigated based on Bell's theorem[11] quantum noiseless coding theorem appeared[12] as a quantum analogous of Shannon's classical noiseless coding theorem. This paper will concentrate on application of quantum entanglement to quantum source entropy and quantum channel capacity in the subject of quantum information.

Recently tremendous effort has been made to better understand the properties of quantum entanglement as a fundamental resource of nature. Although there is as yet no complete understanding and proof of physical realizability of quantum entanglement for quantum technologies, a theoretical progress has been made in understanding this strange property of quantum mechanics, for example, mathematical aspects of quantum entanglement are extensively studied: V. P. Belavkin[13] described the dynamical procedure of quantum entanglement in terms of transpose-completely positive maps in the subject of quantum decoherence and stochastic filtering theory; V. P. Belavkin and M. Ohya[78] initiated mathematical study of quantum entanglement as truly quantum couplings from an operational view in algebraic approach; Peter Levay[14] investigated geometry of quantum entanglement for two qubits (quantum entanglement of two qubits corresponds to the twisting of the bundle); R. Penrose[15] treated quantum entanglement via spinor representation in the subject of mathematical physics; Peter Levay[16] investigated twistor geometry of quantum entanglement for three qubits still in mathematical physics. This paper will follow[5–8] to treat with quantum entanglement in algebraic approach.

Using the operational treatment of entanglement as "true quantum" encoding, V. P. Belavkin and M. Ohya[5–8] introduced quantum conditional entropy of the entangled compound state related to the product of marginal states which is positive and obeys all natural properties of the classical conditional entropy as the relative conditional/unconditional entropy of a compound state. They studied its relation to the mutual information as the informational divergence (relative informational entropy) of the compound state with respect to the product of its marginal states in the sense of Lindblad, Araki and Umegaki.[17,18,19] This quantum mutual information leads to an entropy bound of quantum mutual information and quantum channel capacity via entanglement (entanglement-assisted quantum capacity intro-

duced in[34]), which considered the mutual information of input-output state of quantum channel. Also V. P. Belavkin and P. Staszewski[20] investigated C*-algebraic generalization of relative and conditional entropy including two types of quantum relative entropy, such as Araki-Umegaki type and Belavkin-Staszewski type, and even more general informational divergencies which meet natural axiomatic properties of relative information were studied in quantum information.[21]

Based on the combination of these two original ideas, after introducing compound quantum state and two types of quantum relative entropy, namely Araki-Umegaki type and Belavkin-Staszewski type, this paper treats two types of quantum mutual information via entanglement in algebraic approach and corresponding quantum channel capacities via entanglement in operational approach. It proves additivity property of quantum channel capacities via entanglement, which extends the results of V. P. Belavkin[56] to products of arbitrary quantum channel and to quantum relative entropy of not only Araki-Umegaki type but also Belavkin-Staszewski type.

The rest of this paper is organized as follows: section two and three introduce related notion of quantum mechanics, such as quantum state and quantum entanglement respectively; section four introduces two types of quantum relative entropy via entanglement; section five introduces quantum channel capacity via entanglement and show additivity of quantum channel capacity via entanglement; final section contributes to conclusion and further problems.

2. Quantum States in Algebraic Approach

This subsection is a brief mathematical review of Quantum State in Quantum Mechanics in a discrete algebraic approach. Anyone can turn for general physical review,[22] or for mathematical foundations of Quantum Mechanics,[23] or for a brief review of Quantum Mechanics Principles[24,25] in quantum information and computation.

In order to keep a closer link with classical information theory, we will allow for a possibility of having classical-quantum combined systems described in what follows by discrete non-commutative W*-algebras $\mathcal{A} = (\mathcal{A}_i)$ represented by block-diagonal matrices $A = [A(i)\delta_j^i]$ with arbitrary uniformly bounded operators $A(i) \in \mathcal{A}_i$ on some separable Hilbert spaces $\mathcal{G}_i$.

Let $\mathcal{H}$ denote the separable Hilbert space of a quantum system, and $\mathcal{L}(\mathcal{H})$ denote the algebra of all linear bounded operators on $\mathcal{H}$, with a decomposable subalgebra $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H})$ of elements $B \in \mathcal{B}$ of the block-diagonal

form $B = [B(j)\delta^i_j]$, where $B(j) \in \mathcal{L}(\mathcal{H}_j)$, corresponding to an orthogonal decomposition $\mathcal{H} = \bigoplus_j \mathcal{H}_j$. Note that any such algebra is weakly closed in $\mathcal{L}(\mathcal{H})$, i.e. is a W^*-algebra having a predual space $\mathcal{B}_*$, which can be identified with the trace class subspace of $\mathcal{B}$ with respect to the pairing

$$\langle \varsigma | B \rangle = \sum_j \mathrm{Tr}_{\mathcal{H}_j}[\varsigma(j)^\dagger B(j)] = \mathrm{Tr}_{\mathcal{H}}[B\varsigma^\dagger],$$

where $\varsigma(j) \in \mathcal{B}_j$ are such operators in $\mathcal{H}_j$ that $\mathrm{Tr}_{\mathcal{H}}\sqrt{\varsigma^\dagger\varsigma} < \infty$ and $\mathrm{Tr}_{\mathcal{H}}$ is the standard trace on $\mathcal{B}$ normalized on one dimensional projectors $P_\psi = \psi\psi^\dagger$ for $\psi \in \mathcal{H}_j$. We now remind the definition of quantum normal state.

Definition 2.1. A bounded linear functional $\sigma : \mathcal{B} \to \mathbb{C}$ of the form $\sigma(B) = \mathrm{Tr}_{\mathcal{H}}[B\varsigma]$ for a $\varsigma = \varsigma^\dagger \in \mathcal{B}_*$ is called the state on $\mathcal{B}$ if it is positive for any positive operator $B \in \mathcal{B}$ and normalized $\sigma(I) = 1$ for the identity operator I in $\mathcal{B}$. The operator ς, uniquely defined as a positive trace one operator on $\mathcal{H}$, is called density operator of the state σ.

Let $\mathcal{G}$ be another separable Hilbert space and χ be a Hilbert-Schmidt operator from $\mathcal{G}$ to $\mathcal{H}$ defining a decomposition $\varsigma = \chi\chi^\dagger$ of the state density with the adjoint operator $\chi^\dagger$ from $\mathcal{H}$ to $\mathcal{G}$. We now equip $\mathcal{G}$ with an isometric involution $J = J^\dagger, J^2 = I$, the complex conjugation on $\mathcal{G}$,

$$J\Sigma_k\lambda_k\zeta_k = \Sigma_k\bar{\lambda}_k J\zeta_k, \forall\lambda_k \in \mathbb{C}, \zeta_k \in \mathcal{G}, \tag{1}$$

defining an isometric transposition $\widetilde{A} = JA^\dagger J = \overline{A}^\dagger$ on the algebra $\mathcal{L}(\mathcal{G})$, where $\overline{A} = JAJ$. A normal state $\rho : \mathcal{A} \to \mathbb{C}$ on the algebra $\mathcal{A} \subseteq \mathcal{L}(\mathcal{G})$ is called real (or equivalently symmetric) if its density is real, $\overline{\varrho} = \varrho$ (or equivalently symmetric, $\widetilde{\varrho} = \varrho$). Given a state, J can be always chosen in such a way that $\varrho = \overline{\varrho}$ as it was done in,[5–7] but here we fix J but not ϱ, and in general we will not assume that $\varrho = \widetilde{\varrho}$. Instead, we may assume that the transposition leaves invariant the decomposable subalgebra $\mathcal{A} \subseteq \mathcal{L}(\mathcal{G})$ such that $\overline{\mathcal{A}} := J\mathcal{A}J = \mathcal{A}$, however from the notational and operational point of view, it is preferable to distinguish the algebra $\mathcal{A}$ from the transposed algebras $\widetilde{\mathcal{A}} = \{\widetilde{A} : A \in \mathcal{A}\} = \overline{\mathcal{A}}$.

Lemma 2.1. *([5–8]) Any normal state ρ on $\mathcal{A} \subseteq \mathcal{L}(\mathcal{G})$ can be expressed as*

$$\rho(A) = \mathrm{Tr}_{\mathcal{H}}[\chi\widetilde{A}\chi^\dagger] = \mathrm{Tr}_{\mathcal{H}}[A\varrho], \tag{2}$$

where the density operator $\varrho \in \mathcal{A}_*$ is uniquely defined by $\widetilde{\varrho} = \chi^\dagger\chi = \overline{\varrho}$ iff $\chi^\dagger\chi \in \widetilde{\mathcal{A}}$.

Thus we have an operational expression $\rho(A) = \langle \chi\overline{A}\chi^\dagger | I \rangle$ of quantum normal state, which is called standard in the case $\mathcal{G} = \mathcal{H}$ and $\chi = \sqrt{\varsigma}$, in

which case $\varrho = \overline{\varsigma}$. *Generally* χ *is named as the amplitude operator, or simply amplitude given by a vector* $\chi = \psi \in \mathcal{H}$ *with* $\psi^\dagger \psi = \|\psi\|^2 = 1$ *in the case of one dimensional* $\mathcal{G} = \mathbb{C}$, *corresponding to the pure state* $\sigma(B) = \psi^\dagger B \psi$, *where* $\chi^\dagger$ *is the functional* $\psi^\dagger$ *from* $\mathcal{H}$ *to complex field* $\mathbb{C}$.

Remark 2.1. The amplitude operator χ is unique up to a unitary transform in $\mathcal{H}$ as a probability amplitude satisfying the conditions $\chi^\dagger \chi \in \widetilde{\mathcal{A}}$ such that $\varrho = \overline{\chi^\dagger \chi}$ is positive decomposable trace one operator $\varrho = \oplus_i \varrho(i)$ with the components $\varrho(i) \in \mathcal{L}(\mathcal{G}_i)$ normalized as

$$\mathrm{Tr}_{\mathcal{G}_i} \varrho(i) = k(i) \geq 0, \sum_i k(i) = 1. \tag{3}$$

Therefore we can identify the predual space $\mathcal{A}_*$ with the direct sum $\bigoplus \mathcal{T}(\mathcal{G}_i) \subseteq \mathcal{A}$ of the Banach spaces $\mathcal{T}(\mathcal{G}_i)$ of trace class operators in $\mathcal{G}_i$.

Note that we denote the probability operators $P_{\mathcal{A}} = \varrho \in \mathcal{A}_*$, $P_{\mathcal{B}} = \varsigma \in \mathcal{B}_*$ as trace densities of the states ρ, σ defined as the expectations on the algebras $\mathcal{A}, \mathcal{B}$ respectively by the variations of Greek letters ρ, σ which are also used[8] for the transposed (contravariant) density operators $\widetilde{\varrho} \equiv \rho = \overline{\varrho}, \widetilde{\varsigma} \equiv \sigma = \overline{\varsigma}$ with respect to the bilinear pairings $\rho(A) = \langle A, \rho \rangle \equiv \langle \overline{\rho} | A \rangle$, $\sigma(B) = \langle B, \sigma \rangle \equiv \langle \overline{\sigma} | B \rangle$.

We now define an entangled state ω on the W^*-tensor product algebra $\mathcal{A} \otimes \mathcal{B}$ of bounded operators on the Hilbert product space $\mathcal{G} \otimes \mathcal{H}$ by

$$\mathrm{Tr}_{\mathcal{G}}[\widetilde{A} \chi^\dagger B \chi] = \omega(A \otimes B) = \mathrm{Tr}_{\mathcal{H}}[\chi \widetilde{A} \chi^\dagger B]. \tag{4}$$

Obviously ω can be uniquely extended by linearity to a normal state on the algebra $\mathcal{A} \otimes \mathcal{B}$ generated by all the linear combinations $C = \Sigma_k \lambda_k A_k \otimes B_k$ such that $\omega(C^\dagger C) = \mathrm{Tr}_{\mathcal{G}}[X^\dagger X] \geq 0$, where $X = \Sigma_k \lambda_k B_k \chi \tilde{A}_k$, and $\omega(I \otimes I) = \mathrm{Tr}[\chi^\dagger \chi] = 1$.

Remark 2.2. The state (4) is pure on $\mathcal{L}(\mathcal{G} \otimes \mathcal{H})$, since it is given by an amplitude $\psi \in \mathcal{G} \otimes \mathcal{H}$ defined as $(\zeta \otimes \eta)^\dagger \psi = \eta^\dagger \chi J \zeta$, $\forall \zeta \in \mathcal{G}, \eta \in \mathcal{H}$, with the states ρ on $\mathcal{A}$ and σ on $\mathcal{B}$ as the marginals of ω:

$$\sigma(B) = \omega(I \otimes B) = \mathrm{Tr}_{\mathcal{H}}[B\varsigma],\ \rho(A) = \omega(A \otimes I) = \mathrm{Tr}_{\mathcal{G}}[\widetilde{A}\varrho]. \tag{5}$$

Therefore, we call the state ω defined above as a pure entanglement state for $\mathcal{A} = \mathcal{L}(\mathcal{G})$, $\mathcal{B} = \mathcal{L}(\mathcal{H})$.

More general, mixed entangled states for $\mathcal{A} = \mathcal{L}(\mathcal{G})$, $\mathcal{B} = \mathcal{L}(\mathcal{H})$ can be obtained by using a stochastic amplitude operator $\chi : \mathcal{G} \to \mathcal{F} \otimes \mathcal{H}$.

Given an amplitude operator $\upsilon : \mathcal{F} \longrightarrow \mathcal{G} \otimes \mathcal{H}$ on a Hilbert space $\mathcal{F}$ into the tensor product Hilbert space $\mathcal{G} \otimes \mathcal{H}$ such that $\varpi := \upsilon \upsilon^\dagger \in \mathcal{A} \otimes \mathcal{B}$ and

$\mathrm{Tr}_{\mathcal{F}}[\upsilon^\dagger\upsilon]=1$, we define a compound state $\omega:\mathcal{A}\otimes\mathcal{B}\longrightarrow\mathbb{C}$ as

$$\omega(A\otimes B)=\mathrm{Tr}_{\mathcal{F}}[\upsilon^\dagger(A\otimes B)\upsilon]=\mathrm{Tr}[(A\otimes B)\varpi]. \tag{6}$$

Lemma 2.2. *([5–8]) Any compound state (6) can be achieved via an entanglement χ as*

$$\mathrm{Tr}_{\mathcal{G}}[\widetilde{A}\chi^\dagger(I\otimes B)\chi]=\omega(A\otimes B)=\mathrm{Tr}_{\mathcal{F}\otimes\mathcal{G}}[\chi\widetilde{A}\chi^\dagger(I\otimes B)], \tag{7}$$

with $\omega(A\otimes I)=\mathrm{Tr}_{\mathcal{G}}[A\varrho]$, $\omega(I\otimes B)=\mathrm{Tr}_{\mathcal{H}}[B\varsigma]$, $\widetilde{\varrho}=\chi^\dagger\chi$ and $\varsigma=\mathrm{Tr}_{\mathcal{F}}[\chi\chi^\dagger]$, where χ is an operator $\mathcal{G}\longrightarrow\mathcal{F}\otimes\mathcal{H}$ with $\mathrm{Tr}_{\mathcal{F}}[\chi\mathcal{A}\chi^\dagger]\subset\mathcal{B}$, $\chi^\dagger(I\otimes B)\chi\subset\mathcal{A}$. Moreover, the operator χ is uniquely defined by $\widetilde{\chi}U=\upsilon$, where

$$(\zeta\otimes\eta)^\dagger\widetilde{\chi}\xi=(J\xi\otimes\eta)^\dagger\chi J\zeta,\ \ \forall\xi\in\mathcal{F},\zeta\in\mathcal{G},\eta\in\mathcal{H}, \tag{8}$$

up to a unitary transformation U of the minimal space $\mathcal{F}=\mathrm{rank}\upsilon^\dagger$ equipped with an isometric involution J.

Note that we have used the invariance of trace under the transposition such that $\mathrm{Tr}_{\mathcal{G}}[\widetilde{\varrho}]=\mathrm{Tr}_{\mathcal{G}}[\varrho]$.

3. Entanglement as Quantum Operation

Quantum entanglement is iron to the classical world's bronze age. Quantum entanglement are recently researched extensively, such as Peter Levay[14] via geometric method, Penrose[15] and Peter Levay[16] via spinor and twistor representation, Belavkin[5–8] via algebraic approach. We now follow[7,8,26] for entangled state.

Let us write the entangled state as

$$\omega(A\otimes B)=\mathrm{Tr}_{\mathcal{H}}[B\pi^*(A)]=\mathrm{Tr}_{\mathcal{G}}[A\pi(B)], \tag{9}$$

where the operator $\pi^*(A)=\mathrm{Tr}_{\mathcal{F}}[\chi\widetilde{A}\chi^\dagger]\in\mathcal{B}$, bounded by $\|A\|\varsigma\in\mathcal{B}_*$, is in the predual space $\mathcal{B}_*=\mathcal{T}(\mathcal{H})$ of $\mathcal{B}$ for any $A\in\mathcal{G}$, and

$$\pi(B)=J\chi^\dagger(I\otimes B^\dagger)\chi J=\tilde{\chi}(I\otimes\widetilde{B})\bar{\chi}, \tag{10}$$

with $\widetilde{B}$ defined by isometric involution in $\mathcal{H}$ as $\widetilde{B}=JB^+J$, is in $\mathcal{A}_*$ as a trace-class operator in $\mathcal{G}$, bounded by $\|B\|\varsigma\in\mathcal{A}_*$.

The dual linear maps π and π^* in (9), $\pi^{**}=\pi$, with respect to the standard pairing $\langle A|A\rangle=\mathrm{Tr}[A^*A]$, are both positive, but in general not completely positive but transpose-completely positive maps, with $\pi^*(I)=\varsigma$, $\pi(I)=\varrho$.

Remark 3.1. For the entangled state $\omega(A\otimes B)=\mathrm{Tr}[(A\otimes B)\varpi]$, in terms of the compound density operator $\varpi=\upsilon\upsilon^{\dagger}$, the entanglements π and π^{*} can be written as

$$\pi(B)=\mathrm{Tr}_{\mathcal{H}}[(I\otimes\widetilde{B})\varpi],\pi^{*}(A)=\mathrm{Tr}_{\mathcal{G}}[(\widetilde{A}\otimes I)\varpi]. \tag{11}$$

Definition 3.1. ([5–8]) The transpose-completely positive map $\pi:\mathcal{B}\rightarrow\mathcal{A}_{*}$, (or its dual map $\pi^{*}:\mathcal{A}\rightarrow\mathcal{B}_{*}$), normalized as $\mathrm{Tr}_{\mathcal{G}}[\pi(I)]=1$ (or, equivalently, $\mathrm{Tr}_{\mathcal{H}}[\pi^{*}(I)]=1$) is called the quantum entanglement of the state $\sigma(B)=\mathrm{Tr}_{\mathcal{H}}[\pi(B)]$ to a state on $\mathcal{A}$ described by the density operator $\varrho=\pi(I)$ (or of $\rho(A)=\mathrm{Tr}_{\mathcal{G}}[\pi^{*}(A)]$ to $\varsigma=\pi^{*}(I)$).

We call the standard entanglement $\pi=\pi_{q}$ for $(\mathcal{B},\sigma)$ the entanglement to $\varrho=\widetilde{\varsigma}$ on $\mathcal{A}=\tilde{\mathcal{B}}$ by

$$\pi_{q}(B)=\varrho^{1/2}\widetilde{B}\varrho^{1/2},B\in\tilde{\mathcal{B}}. \tag{12}$$

Obviously $\pi_{q}^{*}(A)=\varsigma^{1/2}\widetilde{A}\varsigma^{1/2}$, where $\varsigma=\widetilde{\varrho}$, and $\pi_{q}^{*}=\pi_{q}$ iff $\mathcal{B}=\tilde{\mathcal{B}}$ and $\varsigma=\widetilde{\varsigma}$.

The standard entanglement defines the standard compound state

$$\omega_{q}(A\otimes B)=\mathrm{Tr}_{\mathcal{H}}[B\varsigma^{1/2}\widetilde{A}\varsigma^{1/2}]=\mathrm{Tr}_{\mathcal{H}}[A\varrho^{1/2}\widetilde{B}\varrho^{1/2}]. \tag{13}$$

Theorem 3.1. *Every entanglement π on $\mathcal{B}$ to the state $\varrho\in\mathcal{A}_{*}$ has a decomposition*

$$\pi(B)=\sqrt{\varrho}\widetilde{\Pi(B)}\sqrt{\varrho}\equiv\pi_{q}(\Pi(B)), \tag{14}$$

where Π is a normal completely positive map $\mathcal{B}\rightarrow\tilde{\mathcal{A}}$ normalized to the identity operator at least on the minimal Hilbert subspace supporting density operator $\widetilde{\varrho}$. This decomposition is unique by the condition $\Pi(I)=E_{\widetilde{\varrho}}$, where $E_{\widetilde{\varrho}}\in\widetilde{\mathcal{A}}$ is the orthoprojector on this minimal Hilbert subspace $\tilde{\mathcal{G}}_{\varrho}\subseteq\mathcal{G}$.

Proof. Π can be found as a solution to the linear equation

$$\widetilde{\varrho}^{1/2}\Pi(B)\widetilde{\varrho}^{1/2}\equiv\widetilde{\pi(B)}\ \ \forall B\in\mathcal{B} \tag{15}$$

which is unique if ϱ and therefore $\widetilde{\varrho}$ is not degenerate:

$$\Pi(B)=\widetilde{\varrho}^{-1/2}\widetilde{\pi(B)}\widetilde{\varrho}^{-1/2}. \tag{16}$$

If ϱ is degenerate, we should consider the Hilbert subspace $\mathcal{G}_{\widetilde{\varrho}}=E_{\widetilde{\varrho}}\mathcal{G}$ given by the minimal orthoprojector $E_{\widetilde{\varrho}}\in\tilde{\mathcal{A}}$ supporting the state $\widetilde{\rho}(A)=\rho(\widetilde{A})$ on the transposed algebra $\tilde{\mathcal{A}}$ such that $\widetilde{\rho}\,(E_{\widetilde{\varrho}})=1$. □

4. Quantum Mutual Information via Entanglement

Quantum mutual information is extensively researched in the past starting from Belavkin and Stratonovich[26] and more recently by Belavkin and Ohya[5] Belavkin and Ohya[17] Benjamin Schumacher and Michael D. Westmoreland.[27] Belavkin and Ohya[78] introduced quantum mutual information as the von Neumann negaentropy $\mathcal{R}(\varpi) = -\mathcal{S}(\varpi)$ of the entangled compound state related to negaentropy $\mathcal{R}(\varrho\otimes\varsigma) = -\mathcal{S}(\varrho\otimes\varsigma)$ of the product of marginal states, i.e. as the relative negaentropy $\mathcal{R}^{(\mathrm{a})}(\varpi : \varphi) = -\mathcal{S}^{(\mathrm{a})}(\varpi : \varphi)$, in the sense of Lindblad, Araki and Umegaki relative entropy[17,18,19] with respect to $\varphi = \varrho \otimes \varsigma$. Cerf and Adami[2] discussed mutual quantum information entropy and its subadditivity property via entropy diagram.

Note that we prefer to use in what is following the term "information" for negaentropy, leaving the term "entropy" for the opposite quantities like relative negainformation $\mathcal{S}^{(\mathrm{a})}(\varpi : \varphi) = -\mathcal{R}^{(\mathrm{a})}(\varpi : \varphi)$, which coincides with usual von Newmann entropy $\mathcal{S}(\varpi)$ if it is taken with respect to the trace $\phi = \mathrm{Tr}$.

We now follow[7,8,26] to define quantum mutual information via quantum entanglement.

Definition 4.1. Relative quantum information of Araki-Umegaki type to compound state ω on the algebra $\mathcal{A} \otimes \mathcal{B}$, (or information divergence of the state ω with respect to a reference state ϕ) is defined by the density operator ϖ, φ of these states ω and ϕ as

$$\mathcal{R}^{(\mathrm{a})}(\varpi : \varphi) = \mathrm{Tr}[\varpi(\ln \varpi - \ln \varphi)]. \tag{17}$$

This quantity is used in most definitions of quantum relative information. However unlike the classical case, this is not only possible choice for informational divergence of the states ω and ϕ, and it does not relate explicitly the informational divergence to the Radon-Nikodym type (RN) density $\varpi_\phi = \varphi^{-1/2}\varpi\varphi^{-1/2}$ of the state ω with respect to ϕ as in the classical case.

Another quantum relative information (of Belavkin-Staszewski type[28]) was introduced[21] as

$$\mathcal{R}^{(\mathrm{b})}(\varpi : \varphi) = \mathrm{Tr}[\varpi \ln(\varphi^{-1}\varpi)], \tag{18}$$

where $\varpi \ln(\varphi^{-1}\varpi) = \ln(\varpi\varphi^{-1})\varpi$ is understood as the Hermitian operator

$$\varpi^{1/2} \ln(\varpi^{1/2}\varphi^{-1}\varpi^{1/2})\varpi^{1/2} = \upsilon \ln(\upsilon^\dagger\varphi^{-1}\upsilon)\upsilon^\dagger. \tag{19}$$

This relative information can be explicitly written in terms of the RN density ϖ_ϕ as $\mathcal{R}^{(\mathrm{b})}(\varpi : \varphi) = \phi(r(\varpi_\phi))$, where $r(\varpi_\phi) = \varpi_\phi \ln \varpi_\phi$.

Ohya and Petz[28] were able to show that, in finite dimensions and faithful states, the Belavkin-Staszewski information divergence based on quantum relative information of Belavkin and Staszewski type gives better distinction of ϖ and φ in the sense that it is greater than relative quantum information of Araki-Umegaki type, and that it satisfies the following important property.

Lemma 4.1. *Given a normal completely positive unital map* $\mathrm{K} : \mathcal{M} \to \mathcal{M}^0$*, if* $\omega = \omega_0\mathrm{K}, \phi = \phi_0\mathrm{K}$*, then for both relative informations,*

$$\mathcal{R}(\varpi : \varphi) \leq \mathcal{R}(\varpi_0 : \varphi_0). \tag{20}$$

Generally this is called monotonicity property of relative information, which is well known for Araki-Umegaki type[17,29] while it is less known that Belavkin-Staszewski type also satisfies all axioms for quantum relative entropy including this inequality. Of course it is worth mathematically proving this inequality of Belavkin-Staszewski type in the most general case.

Definition 4.2. We define the mutual quantum information $\mathcal{I}_{\mathcal{A},\mathcal{B}}(\pi) = \mathcal{I}_{\mathcal{B},\mathcal{A}}(\pi^*)$ of both types in a compound state ω achieved by a quantum entanglement $\pi : \mathcal{B} \to \mathcal{A}_*$, or by $\pi^* : \mathcal{A} \to \mathcal{B}_*$ with

$$\rho(A) = \omega(A \otimes I) = \mathrm{Tr}_{\mathcal{G}}[A\varrho], \sigma(B) = \omega(I \otimes B) = Tr_{\mathcal{H}}[B\varsigma] \tag{21}$$

as the relative information of each type of the state ω on $\mathcal{M} = \mathcal{A} \otimes \mathcal{B}$ with the respect to the product state $\phi = \rho \otimes \sigma$:

$$\mathcal{I}^{(\mathrm{a})}_{\mathcal{A},\mathcal{B}}(\pi) = \mathrm{Tr}[\varpi(\ln \varpi - \ln(\varrho \otimes I) - \ln(I \otimes \varsigma))]. \tag{22}$$

$$\mathcal{I}^{(\mathrm{b})}_{\mathcal{A},\mathcal{B}}(\pi) = \mathrm{Tr}[\varpi \ln((\varrho \otimes \varsigma)^{-1}\varpi)]. \tag{23}$$

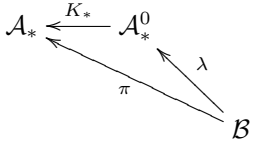
The definition of mutual quantum entropy for Araki-Umegaki type can be found.[5–8] Note that $\mathcal{I}^{(\mathrm{a})}_{\mathcal{A},\mathcal{B}}(\pi) \leq \mathcal{I}^{(\mathrm{b})}_{\mathcal{A},\mathcal{B}}(\pi)$ as it follows from Ohya and Petz.[28]

The following inequality for Araki-Umegaki type can also be found.[5–8] Similarly this inequality for Belavkin-Staszewski type holds.

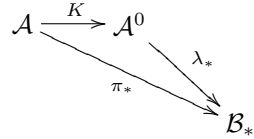
Theorem 4.1. *Let* $\lambda : \mathcal{B} \to \mathcal{A}^0_*$ *be an entanglement of the state* $\sigma(B) = \mathrm{Tr}[\lambda(B)]$ *to* $(\mathcal{A}^0, \rho^0)$ *with* $\mathcal{A}^0 \subseteq \mathcal{L}(\mathcal{G}_0)$*,* $\varrho^0 = \lambda(I)$ *on* $\mathcal{B}$*, and* $\pi = \mathrm{K}_*\lambda$ *be entanglement to the state* $\rho = \rho^0\mathrm{K}$ *on* $\mathcal{A} \subseteq \mathcal{G}$ *defined as the composition of* λ *with the predual operator* $\mathrm{K}_* : \mathcal{A}^0_* \to \mathcal{A}_*$ *normal completely positive unital map* $\mathrm{K} : \mathcal{A} \to \mathcal{A}^0$*. Then for both mutual quantum informations, the following monotonicity holds*

$$\mathcal{I}_{\mathcal{A},\mathcal{B}}(\pi) \leq \mathcal{I}_{\mathcal{A}^0,\mathcal{B}}(\lambda). \tag{24}$$

Proof. This follows from the commutativity of the following diagrams:



Commutative diagram for entanglement π



Dual commutative diagram for entanglement π_*

Applying the monotonicity property of the relative information on $\mathcal{M} = \mathcal{A} \otimes \mathcal{B}$ with respect to the predual map $\varpi_0 \mapsto (\mathrm{K}_* \otimes \mathrm{Id})(\varpi_0)$ corresponding to $\omega_0 \mapsto \omega_0(\mathrm{K} \otimes \mathrm{Id})$ as the ampliation $\mathrm{K} \otimes \mathrm{Id}$ of a normal completely positive unital map $\mathrm{K} : \mathcal{A} \to \mathcal{A}^0$. □

Definition 4.3. The maximal quantum mutual information $\mathcal{J}_{\tilde{\mathcal{B}},\mathcal{B}}(\pi_q)$ for both types as the supremum

$$H_{\mathcal{B}}(\varsigma) = \sup_{\pi^*(I)=\varsigma} \mathcal{I}_{\mathcal{B},\mathcal{A}}(\pi^*) = \mathcal{J}_{\mathcal{B},\tilde{\mathcal{B}}}(\pi_q^*) \tag{25}$$

over all entanglements π^* of any $(\mathcal{A}, \rho)$ to $(\mathcal{B}, \sigma)$ is achieved on $\mathcal{A}^0 = \tilde{\mathcal{B}}$, $\varrho^0 = \tilde{\varsigma}$ by the standard quantum entanglement $\pi_q^*(A) = \varsigma^{1/2}\widetilde{A}\varsigma^{1/2}$ for a fixed $\sigma(B) = \mathrm{Tr}_{\mathcal{H}}[B\varsigma]$, which is named as entangled, or true quantum entropy of each type of the state σ.

This definition for Araki-Umegaki type can be found.[5–8]

Definition 4.4. We call the positive difference

$$H_{\mathcal{B}|\mathcal{A}}(\pi) = H_{\mathcal{B}}(\varsigma) - \mathcal{I}_{\mathcal{A},\mathcal{B}}(\pi) \tag{26}$$

entangled (or true quantum) conditional entropy respectively of each type on $\mathcal{B}$ with respect to $\mathcal{A}$.

This definition for Araki-Umegaki type can be found.[5–8] Obviously, the conditional mutual quantum entropies of both types are positive, unlike the "conditional entropies" considered for example.[24]

5. Entangled Channel Capacity and its Additivity

Entanglement-assisted quantum capacity, or entangled quantum capacity is extensively researched recently, such as entangled quantum capacity[5–8] and entanglement-assisted quantum capacity[34] Generally C. H. Bennett, P. W. Shor, J. A. Smolin and A. V. Thapliyal[34] defined entanglement-assisted capacity of quantum channel via a common framework, we now

discuss quantum channel capacity via entanglement via mutual quantum information entropy.

Let $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H})$ be the W^*-algebra of operators in a (not necessarily finite dimensional unitary) Hilbert space $\mathcal{H}$. Generally we denote the set of states, i.e. positive unit trace operators in $\mathcal{B}(\mathcal{H})$ by $\mathcal{S}(\mathcal{H})$, the set of all m-dimensional projections by $\mathcal{P}_m(\mathcal{H})$ and the set of all projections by $\mathcal{P}(\mathcal{H})$.

Definition 5.1. A quantum channel Λ is a normal unital completely positive linear map (UCP) of $\mathcal{B}$ into the same or another algebra $\mathcal{B}^0 \subseteq \mathcal{B}(\mathcal{H}^0)$.

These maps admit the Kraus decomposition, which is usually written in terms of the dual map $\Lambda^* : \mathcal{B}^0_* \to \mathcal{B}_*$ as $\Lambda^*(\varsigma^0) = \sum_k A_k \varsigma^0 A_k^* \equiv \Lambda_*(\varsigma^0)$ (W. F. Stinespring[29] G. Lindblad[30] A. S. Holevo,[28]), $\Lambda(B) = \sum_k A_k^* B A_k$, for A_k are operators $\mathcal{H}^0 \to \mathcal{H}$ satisfying $\sum_k A_k^* A_k = I^0$. For example, quantum noiseless channel in the case $\mathcal{B} = \mathcal{L}(\mathcal{H})$, $\mathcal{B}^0 = \mathcal{L}(\mathcal{H}^0)$ is described by a single isometric operator $Y : \mathcal{H}^0 \to \mathcal{H}$ as $\Lambda(B) = Y^* B Y$. See for example[29,30] for the simple cases $\mathcal{B} = \mathcal{L}(\mathcal{H})$, $\dim(\mathcal{H}) < \infty$.

A noisy quantum channel sends input pure states $\sigma_0 = \rho_0$ on the algebra $\mathcal{B}^0 = \mathcal{L}(\mathcal{H}^0)$ into mixed states described by the output densities $\varsigma = \Lambda^*(\varsigma^0)$ on $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H})$ given by the predual $\Lambda_* = \Lambda^* \mid \mathcal{B}^0_*$ to the normal completely positive unital map $\Lambda : \mathcal{B} \to \mathcal{B}^0$ which can always be written as

$$\Lambda(B) = \mathrm{Tr}_{\mathcal{F}_+}[Y^\dagger B Y]. \tag{27}$$

Here Y is a linear operator from $\mathcal{H}^0 \otimes \mathcal{F}_+$ to $\mathcal{H}$ with $\mathrm{Tr}_{\mathcal{F}_+}[Y^\dagger Y] = I$, and $\mathcal{F}_+$ is a separable Hilbert space of quantum noise in the channel. Each input mixed state σ^0 is transmitted into an output state $\sigma = \sigma^0 \Lambda$ given by the density operator

$$\Lambda^*(\varsigma^0) = Y(\varsigma^0 \otimes I_+)Y^\dagger \in \mathcal{B}_* \tag{28}$$

for each density operator $\varsigma^0 \in \mathcal{B}^0_*$, the identity operator $I_+ \in \mathcal{F}_+$.

We follow[5–8] to denote K_q the set of all normal transpose-completely positive maps $\kappa : \mathcal{A} \to \mathcal{B}^0$ with any probe algebra $\mathcal{A}$, normalized as $\mathrm{Tr}[\kappa(I)] = 1$, and $\mathcal{K}_q(\varsigma^0)$ be the subset of $\kappa \in \mathcal{K}_q$ with $\kappa(I) = \varsigma^0$. We take the standard entanglement π^0_q on $(\mathcal{B}^0, \sigma^0) = \big(\mathcal{A}_0, \rho^0\big)$, where $\rho_0(A_0) = \mathrm{Tr}[A_0 \varrho_0]$ given by the density operator $\varrho_0 = \varsigma^0$, and denote by K a normal unital completely positive map $\mathcal{A} \to \mathcal{A}^0 = \widetilde{\mathcal{A}}_0$ that decomposes κ as $\kappa(A) = \varrho_0^{1/2} \widetilde{\mathrm{K}(A)} \varrho_0^{1/2}$. It defines an input entanglement $\kappa^* = \mathrm{K}_* \pi^0_q$ on the input of quantum channel as transpose-completely positive map on $\mathcal{A}_0 = \mathcal{B}^0$ into $\mathcal{A}_*$ normalized to $\varrho = \mathrm{K}_* \varrho^0$, $\varrho^0 = \widetilde{\varrho}_0$.

The channel Λ transmits this input entanglement as a true-quantum encoding into the output entanglement $\pi = \mathrm{K}_*\pi_q^0\Lambda \equiv \mathrm{K}_*\lambda$ mapping $\mathcal{B}$ via the channel Λ into $\mathcal{A}_*$ with $\pi(I) = \varrho$. The mutual entangled information, transmitted via the channel for quantum encoding κ, is therefore $\mathcal{J}_{\mathcal{A},\mathcal{B}}(\kappa^*\Lambda) = \mathcal{J}_{\mathcal{A},\mathcal{B}}(\mathrm{K}_*\pi_q^0\Lambda) = \mathcal{J}_{\mathcal{A},\mathcal{B}}(\mathrm{K}_*\lambda)$, where $\lambda = \pi_q^0\Lambda$ is the standard input entanglement $\pi_q^0(B) = \varsigma_0^{1/2}\widetilde{B}\varsigma_0^{1/2}$ with $\varsigma_0 = \widetilde{\varsigma}^0$, transmitted via the channel Λ.

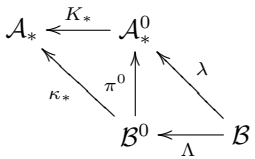
Lemma 5.1. *Given a quantum channel $\Lambda : \mathcal{B} \to \mathcal{B}^0$, and an input state σ^0 on $\mathcal{B}^0$, the entangled input-output quantum information capacity via a channel $\Lambda : \mathcal{B} \to \mathcal{B}^0$ as the supremum over the set $\mathcal{K}_q(\varsigma^0)$ including true-quantum encodings κ achieves the maximal value*

$$\mathcal{J}(\varsigma^0, \Lambda) = \sup_{\kappa \in \mathcal{K}_q(\varsigma^0)} (\kappa^*\Lambda) = \mathcal{I}_{\mathcal{A}^0,\mathcal{B}}(\lambda), \tag{29}$$

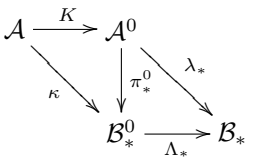
where $\lambda = \pi_q^0\Lambda$ is given by the corresponding extremal input entanglement π_q^0 mapping $\mathcal{B}^0 = \tilde{\mathcal{A}}^0$ into $\mathcal{A}^0 = \tilde{\mathcal{B}}^0$ with $\mathrm{Tr}[\pi_q(B)] = \sigma^0(B)$ for all $B \in \mathcal{B}^0$.

Note that this Lemma for Araki-Umegaki type can be found.[5–8]

The following definition uses commutativity of diagrams:



Commutative diagram for quantum channel Λ with standard entanglement $\pi^0 = \pi_q^0$ for $\mathcal{A} = \widetilde{\mathcal{B}}^0$



Dual commutative diagram for quantum channel Λ with standard entanglement π_*^0 for $\mathcal{A}_*^0 = \widetilde{\mathcal{B}}_*^0$

Definition 5.2. Given a quantum channel $\Lambda : \mathcal{B} \to \mathcal{B}^0$, and a input state σ^0 on $\mathcal{B}^0$, we can define the input-output entangled information capacity as the maximal mutual quantum information

$$\mathcal{J}(\varsigma^0, \Lambda) = \mathcal{I}_{\mathcal{B}^0,\mathcal{B}}(\pi_q^0\Lambda) \tag{30}$$

for input standard entanglement of the state ς^0 to the state $\varrho^0 = \widetilde{\varsigma}^0$.

Note that this definition for Araki-Umegaki type can be found in.[5–8] Thus we have at least two types of such mutual quantum entropy, and

obviously, $\mathcal{J}^{(\mathrm{a})}(\varsigma^0,\Lambda) \leq \mathcal{J}^{(\mathrm{b})}(\varsigma^0,\Lambda)$ with input product state $\rho_0^{\otimes} = \otimes_{i=1}^n \rho_0^i$ corresponding to the states $\rho_0^i = \sigma_i^0$ on $\mathcal{B}_i^0$.

Here and below for notational simplicity we implement the agreements $\mathcal{A}_0^i = \mathcal{B}_i^0$, $\rho_0^i = \sigma_i^0$, $\mathcal{A}_0^{\otimes} = \otimes_{i=1}^n \mathcal{B}_i^0$, $\rho_0^{\otimes} = \otimes_{i=1}^n \sigma_i^0$ such that $\varsigma_0^{\otimes} = \otimes_{i=1}^n \varrho_i^0$ is transposed input state $\widetilde{\varrho}_0^{\otimes} = \otimes_{i=1}^n \widetilde{\varsigma}_i^0$ on $\mathcal{B}_0^{\otimes} = \otimes_{i=1}^n \mathcal{A}_i^0$ with $\widetilde{\mathcal{B}}_i^0 = \mathcal{A}_i^0 \equiv \mathcal{B}_0^i = \widetilde{\mathcal{A}}_0^i$, $\widetilde{\varsigma}_i^0 = \varrho_i^0 \equiv \varsigma_0^i = \widetilde{\varrho}_0^i$.

Let Λ_i be channels respectively from the algebra $\mathcal{B}_i$ on $\mathcal{H}_i$ to $\mathcal{B}_i^0$ on $\mathcal{H}_i^0$ for $i = 1, 2, ..., n$, and let $\Lambda^{\otimes} = \otimes_{i=1}^n \Lambda_i$ be their tensor product.

We now show the additivity property of this entangled input-output quantum information capacity under a given input state, using monotonicity property (as indicated[5–8] for Araki-Umegaki type).

Theorem 5.1. *Let $\Lambda^{\otimes}$ be product channel from the algebra $\mathcal{B}^{\otimes} = \otimes_{i=1}^n \mathcal{B}_i$ to $\mathcal{A}_0^{\otimes} = \otimes_{i=1}^n \mathcal{A}_0^i$, and let $\rho_0^{\otimes} = \otimes_{i=1}^n \rho_0^i$ be the tensor product of input states σ_0^i on $\mathcal{B}_0^i$, then*

$$\mathcal{J}(\varrho_0^{\otimes}, \Lambda^{\otimes}) = \sum_{i=1}^n \mathcal{J}(\varrho_0^i, \Lambda_i). \tag{31}$$

Proof. Take $\Lambda_{i*} : \mathcal{B}_{i*}^0 \to \mathcal{B}_{i*}$, and $\varrho_0^i \in \mathcal{B}_{i*}^0$, $\varsigma_i = \Lambda_{i*}(\varrho_0^i) \in \mathcal{B}_{i*}$, and $\mathrm{K}_*^{(n)} : \mathcal{A}_*^{\otimes} \to \mathcal{A}_*^{(n)}$, where $\mathcal{A}_{0*}^{\otimes} = \otimes_{i=1}^n \mathcal{B}_{i*}^0$, but $\mathcal{A}_*^{(n)}$ is predual to a general, not necessarily product algebra $\mathcal{A}^{(n)} \subseteq \mathcal{L}(\mathcal{G}^{(n)})$. For $\pi^{(n)} = \mathrm{K}_*^{(n)} \pi_q^{0\otimes} \Lambda^{\otimes}$, below we consider quantum mutual information $\mathcal{I}_{\mathcal{A}^{(n)},\mathcal{B}^{\otimes}}(\pi^{(n)})$ as relative quantum entropy

$$\mathcal{R}((\mathrm{K}_*^{(n)} \otimes \Lambda_*^{\otimes})\widetilde{\varpi}_0^{\otimes} : \mathrm{K}_*^{(n)}(\varsigma_0^{\otimes}) \otimes \Lambda_*^{\otimes}(\varrho_0^{\otimes})), \tag{32}$$

where $\widetilde{\varpi}_0^{\otimes} = \otimes_{i=1}^n \widetilde{\varpi}_0^i$ is the density operator of the standard compound state $\otimes_{i=1}^n \omega_0^i$ with $\omega_0^i(A_i \otimes B_i) = \varpi_i^0(A_i \otimes B_i) = \mathrm{Tr}[B_i \sqrt{\varrho_i^0}\widetilde{A}_i \sqrt{\varrho_i^0}]$ for $A_i \in \widetilde{\mathcal{B}}_i^0, B_i \in \mathcal{B}_i^0$, corresponding to $\varsigma_i^0 = \varrho_0^i$.

Applying monotonicity property (Lemma 3) of quantum relative entropy to the probe system $(\mathcal{G}^{(n)}, \mathcal{A}^{(n)})$ for this given ϱ_0^i and Λ_i, we obtain

$$\mathcal{R}((\mathrm{K}_*^{(n)} \otimes \Lambda_*^{\otimes})\widetilde{\varpi}_0^{\otimes} : \mathrm{K}_*^{(n)}(\varsigma_0^{\otimes}) \otimes \Lambda_*^{\otimes}(\varrho_0^{\otimes})) \tag{33}$$

$$\leq \mathcal{R}((\mathrm{Id}^{\otimes} \otimes \Lambda^{\otimes})\widetilde{\varpi}_0^{\otimes} : \mathrm{Id}^{\otimes}(\varsigma_0^{\otimes}) \otimes \Lambda_*^{\otimes}(\varrho_0^{\otimes})) \tag{34}$$

$$= \sum_{i=1}^n \mathcal{R}((\mathrm{Id} \otimes \Lambda_{i*})(\widetilde{\varpi}_0) : \mathrm{Id}(\varsigma_0^i) \otimes \Lambda_{i*}(\varrho_0^i), \tag{35}$$

where $\varsigma_0^i = \varrho_i^0 = \widetilde{\varrho}_0^i$, $\varrho_0^i = \varsigma_i^0 = \widetilde{\varsigma}_0^i$.

The suprema over $\mathrm{K}^{(n)}$ is achieved on $\mathrm{K}^{(n)} = \mathrm{Id}^{\otimes}$ identically mapping $\mathcal{A}^{(n)} = \otimes_{i=1}^n \mathcal{A}_0^i$ to $\mathcal{B}_{0*}^{\otimes} = \otimes_{i=1}^n \mathcal{B}_0^i$, where $\mathcal{B}_0^i = \widetilde{\mathcal{B}}_i^0$, coinciding with such $\mathcal{A}^{(n)}$ due to $\mathcal{A}_0^i = \widetilde{\mathcal{B}}_i^0$.

Thus $\mathcal{J}(\varrho_0^{\otimes}, \Lambda^{\otimes}) = \sum_{i=1}^n \mathcal{J}(\varrho_0^i, \Lambda_i)$. □

Definition 5.3. Given a normal unital completely positive map $\Lambda : \mathcal{B} \to \mathcal{A}$, the suprema

$$C_q(\Lambda) = \sup_{\kappa \in \mathcal{K}_q} \mathcal{I}_{\mathcal{A},\mathcal{B}}(\kappa^* \Lambda) = \sup_{\varsigma^0} \mathcal{J}(\varsigma^0, \Lambda) \tag{36}$$

is called the quantum channel capacity via entanglement, or q-capacity.

Note that this definition for Araki-Umegaki type can be found in,[5–8] there we have two types of entangled channel capacities, and obviously $C_q^{(\mathrm{a})}(\Lambda) \le C_q^{(\mathrm{b})}(\Lambda)$.

Lemma 5.2. *Let $\Lambda(B) = Y^\dagger B Y$ be a unital completely positive map $\Lambda : \mathcal{B} \to \mathcal{B}^0$ describing a quantum deterministic channel by an isometry $Y : \mathcal{H}^0 \to \mathcal{H}$. Then*

$$\mathcal{J}(\varsigma^0, \Lambda) = H_{\mathcal{B}_0}(\varsigma^0), \tag{37}$$

$$C_q(\Lambda) = \ln \dim \mathcal{B}^0. \tag{38}$$

Note that this Lemma for Araki-Umegaki type can be found.[5–8]

Let $\Lambda^{\otimes}$ be product channel from the algebra $\mathcal{B}^{\otimes} = \otimes_{i=1}^n \mathcal{B}_i$ to $\mathcal{A}_0^{\otimes} = \otimes_{i=1}^n \mathcal{B}_i^0$. The additivity problem for quantum channel capacity via entanglement is if it is true that

$$\mathcal{C}_q(\Lambda^{\otimes}) = \sum_{i=1}^n \mathcal{C}_q(\Lambda_i). \tag{39}$$

We now still follow the idea of[5–8] to give a proof of this additivity property via operational approach using monotonicity property (as indicated[5–8] for Araki-Umegaki type).

Theorem 5.2. *Let $\Lambda^{\otimes}$ be product channel from the algebra $\mathcal{B}^{\otimes} = \otimes_{i=1}^n \mathcal{B}_i$ to $\mathcal{A}_0^{\otimes} = \otimes_{i=1}^n \mathcal{B}_i^0$, then*

$$\mathcal{C}_q(\Lambda^{\otimes}) = \sum_{i=1}^n \mathcal{C}_q(\Lambda_i). \tag{40}$$

Proof. It simply follows from the additivity (31). Indeed,

$$C_q(\Lambda^{\otimes}) = \sup_{\kappa\in\mathcal{K}_q^{(n)}} \mathcal{I}_{\mathcal{A}^{(n)},\mathcal{B}}(\kappa^*\Lambda^{\otimes}) = \sup_{\varrho_0^{\otimes}} \mathcal{J}(\varrho_0^{\otimes},\Lambda^{\otimes}) = \sup_{\varrho_0^{\otimes}} \sum_{i=1}^{n} \mathcal{J}(\varrho_0^i,\Lambda_i) \tag{41}$$

Therefore by further taking suprema over $\varrho_0^{\otimes}$ as over independently for each $i = 1, 2, ..., n$, thus we have

$$\mathcal{C}_q(\Lambda^{\otimes}) = \sum_{i=1}^{n} \sup_{\varrho_0^{\otimes}} \mathcal{J}(\varrho_0^i,\Lambda_i) = \sum_{i=1}^{n} \mathcal{C}_q(\Lambda_i), \tag{42}$$

which is the additivity property of entangled quantum channel capacity due to encodings via entanglement obviously. □

Remark 5.1. Note that there is no such additivity for the Holevo capacity for a arbitrary channel $\Lambda : \mathcal{B} \to \mathcal{B}^0$. Indeed, this smaller, semiclassical capacity is defined as the supremum

$$C_d(\Lambda) = \sup_{\kappa\in\mathcal{K}_d} \mathcal{I}_{\mathcal{A},\mathcal{B}}(\kappa^*\Lambda) \tag{43}$$

over the smaller class $\mathcal{K}_d \subseteq \mathcal{K}_q$ of the diagonal[5–8] (semiclassical) encodings $\kappa : \mathcal{A} \to \mathcal{B}_*^0$ corresponding to only diagonal (Abelian) algebras $\mathcal{A}$. This supremum cannot in general be achieved on the standard entanglement of $\mathcal{A}^0 = \widetilde{\mathcal{B}}^0 \equiv \mathcal{B}_0$ if $\mathcal{A}^0$ is non Abelian corresponding to the non Abelian input algebra $\mathcal{B}^0$. Therefore the supremum $\mathcal{C}_d(\Lambda^{\otimes}) \leq \sum_{i=1}^{n} \mathcal{C}_d(\Lambda_i)$ can be achieved not on a product Abelian algebra $\mathcal{A}^{(n)}$ as is was in the true quantum case where we could take $\mathcal{A}^{(n)} = \otimes_{i=1}^{n} \mathcal{B}_0^i$ with non Abelian $\mathcal{B}_0^i = \widetilde{\mathcal{B}}_i^0$.

6. Conclusion

So far, continuing in this paper research on quantum channel capacity for one-way communication via entanglement following,[5–8,21] we treated two types of quantum mutual information via entanglement in algebraic approach and corresponding quantum channel capacities via entanglement in operational approach. Using monotonicity property of quantum mutual information of a- and b-type introduced[5,21] we proved additivity property of quantum channel capacities via entanglement, therefore extending the results of V. P. Belavkin[5,6] to products of arbitrary quantum channel to quantum relative entropy of both Araki-Umegaki type and Belavkin-Staszewski type.

As written in the introduction, quantum channel capacities can have several different formulations when considering to send classical information or quantum information, one-way or two-way communication, prior or via entanglement, etc. in the form of different constraints on the encoding class $\mathcal{K}$. Anyway general quantum channel capacity with different constraints is still a big open and challenging research problem in quantum information theory. Much more open problems can be found[25] . There we anticipate some research on quantum channel capacity for two-way communication and prior or via entanglement, i.e. trading communication and entanglement for quantum channel capacity.

Another natural problem in this direction is to compare true quantum capacities in quantity for some interesting quantum channels with other smaller capacities under constraints, such as Holevo capacity, entanglement-assistant capacity, etc., and find for which channels they coincide.

The third natural problem in this direction is to consider quantum mutual information via entanglement and corresponding quantum channel capacities via entanglement for γ type since[21] studied this third and more general quantum relative entropy in quantum information, which also meet more natural axiomatic properties of relative entropy.

Tracing through the original research ideas on quantum mutual information via entanglement and corresponding quantum channel capacities via entanglement, it is easy to find that the maximal quantum information and capacity is achieved on maximal (standard) entanglement for a given state and on the absolutely maximal entanglement (which exists only in finite dimensions) without constraint on the input state.

Generally how to access those capacities, using physically implementable operations for encodings and decodings, such as in this direction of quantum channel capacity for one-way communication via entanglement, is of course an open problem in quantum information and quantum computation.

All those problems wait forthcoming papers in the future.

Bibliography

1. B. Schumacher, Sending entanglement through noisy quantum channels, *Phys. Rev.* **A 54**, 2614-2628 (1996).
2. N. Cerf and G. Adami, Von Neumann capacity of noisy quantum channels, *Phys. Rev.* **A 56**, 3470-3483 (1997).
3. C. H. Bennett, P. W. Shor, J. A. Smolin and A. V. Thapliyal, Entanglement assisted classical capacity of noisy quantum channels, *Phys. Rev. Lett.* **83** 3081-3084 (1999).
4. C. H. Bennett, P. W. Shor, J. A. Smolin and A. V. Thapliyal, Entanglement-

Assisted Capacity of a Quantum Channel and the Reverse Shannon Theorem, *quant-ph/0106052*.
5. V. P. Belavkin, On Entangled Quantum Capacity, *Quantum Communication, Computing, and Measurement* . **3. Kluwer/Plenum** 325-333 (2001).
6. V. P. Belavkin, On Entangled Information and Quantum Capacity, *Open Sys. and Information Dyn.* **8**1-18 (2001).
7. V. P. Belavkin, M. Ohya, Quantum Entropy and Information in Discrete Entangled States, *Infinite Dimensional Analysis, Quantum Probability And Related Topics.* **4** 137-160 (2001).
8. V. P. Belavkin, M. Ohya, Entanglement, Quantum Entropy and Mutual Information, *Proc. R. Soc. Lond.* **A 458** 209 - 231 (2002).
9. A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.* **47** 777-780 (1935).
10. C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. K. Wootters, Teleporting an unknown quantum state via dual classic and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.* **70** 1895-1899 (1993).
11. A. Ekert, Quantum cryptography based on Bell 's Theorem, *Phys. Rev. Lett.* **67** 661-663 (1990).
12. R. Jozsa and B. Schumacher, A new proof of the quantum noiseless coding theorem, *J. Mod. Opt.* **41** 2343-2350 (1994).
13. V. P. Belavkin, Quantum stochastic positive evolutions: characterization, construction, dilation, *Commun. Math. Phys.* **184** 533-566 (1997).
14. Peter Levay, The geometry of entanglement: metrics, connections and the geometric phase, *J. Phys.* **A37** 1821-1842 (2004).
15. R. Penrose, *Report in NCG workshop in Newton Institute* (Cambridge, Sept., 2006).
16. Peter Levay, The twistor geometry of three-qubit entanglement, *quant-ph/0403060*.
17. G. Lindblad, Entropy, Information and Quantum Measurements, *Comm. In Math. Phys.* **33** 305-322 (1973).
18. H. Araki, Relative Entropy of states of von Neumann Algebras, *Publications RIMS, Kyoto University* **11** 809 (1976).
19. H. Umegaki, *Kodai Math. Sem. Rep.* **14** 59 (1962).
20. V. P. Belavkin and P. Staszewski, C*-algebraic generalization of relative entropy and entropy, *Ann. Inst. Henri Poincare* **37** Sec. A 51-58 (1982).
21. S. J. Hammersley and V. P. Belavkin, Information Divergence for Quantum Channels, Infinite Dimensional Analysis, *Quantum Information and Computing* World Scientific, Quantum Probability and White Noise Analysis, VXIX 149-166 (2006).
22. L. D. Landau, and E. M. Lifschitz, *Quantum Mechanics (Non-Relativistic Theory)*, 3rd ed. (Oxford, England: Pergamon Press, 1977).
23. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, 1955).
24. Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, 2000).

25. Peter Shor, Quantum Information Theory: Results and Open Problems, *Geom. Funct. Anal., Special Volume-GAFA2000*, 816-838 (2000).
26. V. P. Belavkin and R. L. Stratonovich, Optimization of Quantum Information Processing Maximizing Mutual Information, *Radio Eng. Electron. Phys.* **19**, 1349 (1973).
27. Benjamin Schumacher, and Michael D. Westmorel, Quantum mutual information and the one-time pad, *quant-ph/0604207*.
28. M. Ohya and D. Petz, *Quantum Entropy and its Use*, (Springer-Verlag, Berlin, 1993).
29. W. F. Stinespring, *Proc. Amer. Math. Soc.***6** 211 (1955).
30. A. S. Holevo, Quantum coding theorems, *Russian Math. Surveys.* **53**, 1295-1331 (1998).
31. G. Lindblad, Quantum entropy and quantum measurements, *Lect. Notes Phys.* **378**, 71-80, (Springer-Verlag, Berlin, 1991).

CLASSICAL CODING AND THE CAUCHY-SCHWARZ INEQUALITY

BAS JANSSENS

Mathematisch Instituut, Utrecht University, Postbus 80.010, 3508 TA Utrecht, The Netherlands
E-mail: janssens@math.uu.nl

In classical coding, a single quantum state is encoded into classical information. Decoding this classical information in order to regain the original quantum state is known to be impossible. However, one can attempt to construct a state which comes as close as possible. We give bounds on the smallest possible trace distance between the original and the decoded state which can be reached. We give two approaches to the problem: one starting from Keyl and Werner's no-cloning theorem[3], and one starting from an operator-valued Cauchy-Schwarz inequality.

1. Introduction

There exist results which are possible in classical physics, yet forbidden in quantum mechanics. For instance, it is generally impossible to perform a joint measurement on two observables, or to clone one quantum state to two identical ones. This yields a class of problems: exactly how close can one approximate these desired, yet impossible results?

We examine this problem in the specific case of 'classical coding'. Impossibility of classical coding is the following statement: if a single quantum state is encoded into classical information, then it is impossible, starting from this classical information, to reconstruct the original quantum state.

In order to quantify how far one is removed from the ideal situation, we define Δ to be the worst case trace distance between the original and the decoded state. The impossibility of classical coding is then equivalent to $\Delta \neq 0$. The aim is now to investigate which values of Δ can be reached. We will do this in two different ways.

- The heart of this paper is the operator-valued Cauchy-Schwarz in-

equality, described in section 4. We will use it to obtain the bound $\Delta \geq (3 - \sqrt{5})/4$.
- We will also attack the same problem using a no-cloning theorem[3], which will yield the superior bound $\Delta \geq 1/3$.

The Cauchy-Schwarz method was used earlier in[1] for the three quantum impossibilities called 'no measurement without disturbance', 'no measurement without decoherence' and 'no joint measurement'. In contrast with the case considered here, the bounds in these three cases did turn out to be sharp. The purpose of this paper is mainly to illustrate the flexibility of the Cauchy-Schwarz method.

The article is structured as follows: in section 2, a short introduction to quantum probability theory is given. In section 3, the problem of classical coding is formulated in a rigorous mathematical fashion. In section 4, the Cauchy-Schwarz inequality is given, and applied to the problem of classical coding. This yields the bound $\Delta \geq (3-\sqrt{5})/4$. Section 5 provides the superior bound $\Delta \geq 1/3$, based on the no-cloning theorem. A short discussion of the results then follows in section 6.

2. Quantum Mechanics

A quantum mechanical system is described by a von Neumann algebra $\mathcal{A}$ of bounded operators on a Hilbert space $\mathcal{H}$, usually the algebra $B(\mathcal{H})$ of all bounded operators. Its state space is formed by the normalized density matrices $\mathcal{S}(\mathcal{A}) = \{\rho \in \mathcal{A}\,;\, \rho \geq 0, \mathbf{tr}(\rho) = 1\}$. With the system in state $\rho \in \mathcal{S}(\mathcal{A})$, observation of a (Hermitean) observable $A \in \mathcal{A}$ is postulated to yield the average value $\mathbf{tr}(\rho A)$.

2.1. *Completely Positive maps*

The natural notion of a map between von Neumann-algebras is that of a completely positive (or CP for short) map.

Definition 2.1. Let $\mathcal{A}$ and $\mathcal{B}$ be von Neumann algebras. A weakly continuous map $T : \mathcal{B} \to \mathcal{A}$ is called *Completely Positive* (or CP for short) if it is linear, normalized (i.e. $T(I) = I$), positive (i.e. $T(X^\dagger X) \geq 0$ for all $X \in \mathcal{B}$) and if moreover the extension $Id_n \otimes T : M_n \otimes \mathcal{B} \to M_n \otimes \mathcal{A}$ is positive for all $n \in \mathbb{N}$, where M_n is the algebra of complex $n \times n$-matrices.

Its dual $T^* : \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$, defined by the requirement $\mathbf{tr}(T^*(\rho)X) = \mathbf{tr}(\rho T(X))\ \forall\, X \in \mathcal{B}$, has a direct physical interpretation as an operation

between quantum systems. T is positive, linear and normalized. This is equivalent to T^* being an affine map $\mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$. That is, each state $\rho \in \mathcal{S}(\mathcal{A})$ is again mapped to a state $T^*(\rho) \in \mathcal{S}(\mathcal{B})$, and for all $p \in [0,1]$ and $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{A})$, we have $pT^*(\rho_1) + (1-p)T^*(\rho_2) = T^*(p\rho_1 + (1-p)\rho_2)$. This expresses the *stochastic equivalence principle*: a system which is in state ρ_1 with probability p and in state ρ_2 with probability $(1-p)$ cannot be distinguished from a system in state $p\rho_1 + (1-p)\rho_2$.

It is possible to extend the systems $\mathcal{A}$ and $\mathcal{B}$ under consideration with another system M_n, on which the operation acts trivially. Due to *complete* positivity, states in $\mathcal{S}(M_n \otimes \mathcal{A})$ are once again mapped to states in $\mathcal{S}(M_n \otimes \mathcal{B})$. This is of course a crucial property if one seeks to interpret T^* as a physical map: the mere act of coupling the system $\mathcal{A}$ to another system M_n may never cause negative probabilities. Surprisingly enough, there exist linear maps which are positive, but not completely positive. It is generally believed that *any* operation which can be physically implemented on a quantum system is described by a CP-map.

2.2. *Probability Spaces*

We are interested in quantum operations which take as input a quantum state, and yield a classical probability distribution as output.

In order to encompass this in our framework of CP-Maps and von Neumann algebras, we identify a classical probability space with a commutative von Neumann algebra. (See[4].) A classical probability space $(\Omega, \Sigma, \mathbb{P})$ gives rise to $\mathcal{A} := L^\infty(\Omega, \Sigma, \mathbb{P})$, the set of bounded measurable functions on Ω up to the equivalence $\sim$, with $f \sim g$ if $f = g$ almost surely. We think of these as the *random variables*. Since $f \in \mathcal{A}$ acts on the Hilbert space $L^2(\Omega)$ by multiplication, we can regard $\mathcal{A}$ as a commutative subalgebra of $B(L^2(\Omega))$, which turns out to be weakly closed. The probability measure $\mathbb{P}$ of course induces the expectation $\mathbb{E}(f) = \int_\Omega f(\omega)\mathbb{P}(d\omega)$ of a random variable. This yields a state $\rho_\mathbb{P}$ by requiring $\mathbf{tr}(\rho_\mathbb{P} M_f) := \mathbb{E}(f)$.

In short, a classical probability space corresponds to a commutative von Neumann algebra, and a probability measure corresponds to a state on that algebra. An operation C^* which maps quantum states in $\mathcal{S}(B(\mathcal{H}))$ to classical probability distributions in $\mathcal{S}(\mathcal{A})$ must therefore be the dual of a CP-map $C : \mathcal{A} \to B(\mathcal{H})$.

For example, the direct measurement of an observable (= bounded Hermitean operator) X. With each Hermitean operator, there is associated a projection valued measure $P(dx)$, such that $X = \int_{\mathbf{Spec}(X)} xP(dx)$. The measurement of X is then represented by the CP-map $L^\infty(\mathbf{Spec}(X)) \to$

$B(\mathcal{H})$ defined by $f \mapsto \int_{\mathbf{Spec}(X)} f(x)P(dx)$. In the dual (Schrödinger) picture, we then see that a state $\rho \in \mathcal{S}(B(\mathcal{H}))$ induces the probability distribution $\mathbb{P}(dx) = \mathbf{tr}(\rho P(dx))$ on the spectrum of X. More generally, each CP-map $L^\infty(\Omega) \to B(\mathcal{H})$ corresponds to a Positive Operator Valued Measure, but we will not need this here.

Similarly, an operation D^* which maps classical probability distributions into quantum states must be the dual of a CP-map $D : B(\mathcal{H}) \to \mathcal{A}$.

3. Classical Coding

By *classical coding*, we mean the following procedure. First, classical information is extracted from a quantum system. This is described by the coding map $C : \mathcal{A} \to B(\mathcal{H})$, with $\mathcal{A} = L^\infty(\Omega)$ for some space Ω. Then, on the basis of this classical information, the original state is reconstructed as well as possible by means of some decoding procedure. This is described by a CP-map $D : B(\mathcal{H}) \to \mathcal{A}$. In the dual picture, this gives

$$\mathcal{S}(B(\mathcal{H})) \stackrel{C^*}{\to} \mathcal{S}(\mathcal{A}) \stackrel{D^*}{\to} \mathcal{S}(B(\mathcal{H}))\,.$$

The coding procedure is flawless iff every state is reconstructed perfectly, i.e. iff $C \circ D : B(\mathcal{H}) \to B(\mathcal{H})$ is the identity.

It is well known[5,8] that perfect classical coding is impossible. So let us investigate how close we can come to perfection. Define Δ to be the maximum difference between input and output probability on a single event P, i.e.

$$\Delta := \sup_{\rho,P} |\mathbf{tr}(D^*C^*(\rho)P) - \mathbf{tr}(\rho P)|,$$

where ρ runs over $\mathcal{S}(B(\mathcal{H}))$, and P runs over the projections in $B(\mathcal{H})$. We give two reformulations of this definition (see[6]). The first is $\Delta = \sup_\rho |D(\rho, D^*C^*(\rho))|$, with $D(\rho,\tau)$ the trace distance or Kolmogorov distance $D(\rho,\tau) = \frac{1}{2}\mathbf{tr}(|\rho - \tau|)$ between ρ and τ. The second is $\Delta = \sup_B \|B - CD(B)\|$, where B runs over the positive operators $0 \leq B \leq I$. The latter is most convenient, and we will use it in the remainder of the article.

We remark that Δ quantifies the quality of the coding procedure: a large value of Δ corresponds to a poor classical coding procedure, a small value of Δ corresponds to a good one. Furthermore, $\Delta = 0$ if and only if all states are encoded perfectly. We now investigate how close to zero Δ can come.

4. The Cauchy-Schwarz Method

We start with a lemma which, in all its simplicity, is the cornerstone of a veritable zoo of quantum inequalities. (See[1,2]. The special case $(A,A) = 0 \Rightarrow (A,B) = 0$ is older, and due to Werner[8].)

Lemma 4.1 (Cauchy-Schwarz). *Let $\mathcal{A}$, $\mathcal{B}$ be von Neumann algebras, and let $(\,\cdot\,,\,\cdot\,) : \mathcal{A}\times\mathcal{A} \to \mathcal{B}$ be a positive semidefinite sesquilinear form. That is, it is linear in the second argument, $(A,B)^\dagger = (B,A)$ for all $A, B \in \mathcal{A}$, and $(A,A) \geq 0$ for all $A \in \mathcal{A}$. Then $\|\Re(A,B)\|^2 \leq \|(A,A)\|\|(B,B)\|$ and $\|\Im(A,B)\|^2 \leq \|(A,A)\|\|(B,B)\|$ for all $A, B \in \mathcal{A}$.*

Proof. For $A, B \in \mathcal{A}$ and $\lambda \in \mathbb{C}$, we have (with $\Re X := (X + X^\dagger)/2$ the 'real' and $\Im X := (X - X^\dagger)/2i$ the 'imaginary' part):

$$0 \leq (A - \lambda B, A - \lambda B) = (A,A) - 2\Re\lambda(A,B) + |\lambda|^2(B,B). \qquad (1)$$

If $(A,A) = 0$ and $(B,B) = 0$, the lemma follows immediately from (1). If not, assume that $(B,B) \neq 0$, exchanging the roles of A and B if necessary. Choose $\lambda = \pm\|\Re(A,B)\|/\|(B,B)\|$, so that (1) becomes the operator inequality

$$\pm 2\Re(A,B)\|\Re(A,B)\|/\|(B,B)\| \leq (A,A) + (B,B)\|\Re(A,B)\|^2/\|(B,B)\|^2.$$

In particular, the spectrum of the operator on the l.h.s. is contained in $[-\|R\|, \|R\|]$, with R the r.h.s. operator. Thus $2\|\Re(A,B)\|^2/\|(B,B)\| \leq \|R\|$, and $\|R\| \leq \|(A,A)\| + \|\Re(A,B)\|^2/\|(B,B)\|$. This yields $\|\Re(A,B)\|^2 \leq \|(A,A)\|\|(B,B)\|$, as required. Since $\Re(iA,B) = \Im(A,B)$ holds, we also have $\|\Im(A,B)\|^2 \leq \|(A,A)\|\|(B,B)\|$. □

For example, each CP-map $T : \mathcal{A} \to \mathcal{B}$ induces a positive semidefinite sesquilinear form by $(A,B)_T := T(A^\dagger B) - T(A)^\dagger T(B)$. Indeed, according to Stinespring's theorem (see[7]), we can assume without loss of generality that T is of the form $T(A) = V^\dagger A V$, with $\|V\| \leq 1$. Then $(A,A)_T = V^\dagger A^\dagger (I - VV^\dagger)AV = (\sqrt{(I - VV^\dagger)}AV)^\dagger \sqrt{(I - VV^\dagger)}AV \geq 0$. The sesquilinearity is clear.

We use the Cauchy-Schwarz inequality to obtain a the bound on Δ for any classical coding procedure. (The proof is inspired by the 'impossibility of classical coding' version in[5].)

Theorem 4.1. *Let $\mathcal{A}$ be a commutative von Neumann-algebra, and let $B(\mathcal{H})$ be the algebra of bounded operators on a Hilbert space $\mathcal{H}$ of dimension > 1. Let $C : \mathcal{A} \to B(\mathcal{H})$ and $D : \mathcal{B}(\mathcal{H}) \to \mathcal{A}$ be CP-maps. Let $\Delta := \sup\{\|B - C \circ D(B)\| : 0 \leq B \leq I\}$. Then $\Delta \geq (3 - \sqrt{5})/4$.*

Proof. Take two orthogonal vectors $\psi, \phi \in \mathcal{H}$, and define X to be the projection on ψ, and Y the projection on $(\psi + \phi)/\sqrt{2}$. We have $\|[X, Y]\| = 1/2$. Since $\mathcal{A}$ is Abelian, we have $D(X)D(Y) = D(Y)D(X)$, and we can write

$$[X, Y] = [X, Y] - CD([X, Y]) + \tag{2}$$
$$C\left(D(XY) - D(X)D(Y)\right) - \tag{3}$$
$$C\left(D(YX) - D(Y)D(X)\right) .$$

We will bound the r.h.s. in terms of Δ. Remembering that the l.h.s. is at least $1/2$ in norm will then yield a minimum value on Δ.

We start with (2). Like any antihermitean operator, $[X, Y]$ can be written as $[X, Y] = i(A_+ - A_-)$, with $0 \leq A_\pm \leq \|[X, Y]\| I = \frac{1}{2} I$. Therefore, we have $\|[X, Y] - CD([X, Y])\| \leq \Delta \|A_+\| + \Delta \|A_-\| \leq \Delta$.

We then proceed with (3). Consider the positive semidefinite sesquilinear form $(X, Y) := C\left(D(X^\dagger Y) - D(X)^\dagger D(Y)\right)$, in terms of which the expression (3) equals $2i\Im(X, Y)$. It is positive as the concatenation of $(\,\cdot\,,\,\cdot\,)_D : B(\mathcal{H}) \times \mathcal{B}(\mathcal{H}) \to \mathcal{A}$ and the positive map $C : \mathcal{A} \to B(\mathcal{H})$.

According to Lemma 4.1, we have

$$\|2i\Im(X, Y)\| \leq 2\sqrt{\|(X, X)\|\|(Y, Y)\|}.$$

Now $(X, X) = (X, X)_{CD} - (D(X), D(X))_C \leq (X, X)_{CD}$, and a similar expression holds for Y.

Since $X^\dagger X = X$, we have $(X, X)_{CD} = CD(X)(I - CD(X))$. Since $\|X - CD(X)\| \leq \Delta$, and X has spectrum $\{0, 1\}$, the spectrum of $CD(X)$ is contained in $[0, \Delta] \cup [1 - \Delta, 1]$. (Recall that $0 \leq CD(X) \leq I$.) The spectrum of $CD(X)(I - CD(X))$ therefore lies within $[0, \Delta(1 - \Delta)]$, so that $\|(X, X)_{CD}\| \leq \Delta(1 - \Delta)$. The same holds for Y of course. Putting this together, we bound (3) as $\|2i\Im(X, Y)\| \leq 2\Delta(1 - \Delta)$.

We conclude that $1/2 = \|[X, Y]\| \leq \Delta + 2\Delta(1 - \Delta)$, or $(\Delta - 3/4)^2 \leq 5/16$. Thus $\Delta \geq (3 - \sqrt{5})/4$, which was to be proven. □

5. The No-Cloning Method

It is easy to obtain a better bound from the no-cloning theorem of Keyl and Werner[3]. The idea is, that each classical coding procedure yields a cloning machine. All one has to do is to 'decode' the classical information an arbitrary amount M of times, rather than just once.

To be more explicit, suppose that we are in the finite setting: $\mathcal{H} = \mathbb{C}^d$ is a finite-dimensional Hilbert space, and $\mathcal{A} = L^\infty(\Omega)$ with $\Omega = \{1, 2, \ldots, n\}$.

The 'diagonal' map $\Omega \to \Omega^M : i \mapsto (i, \ldots, i)$ induces the 'classical cloning' map $K : \mathcal{A}^{\otimes M} \to \mathcal{A}$, i.e. $(Kf)(i) := f(i, \ldots, i)$. (Note that for e.g. $\Omega = \mathbb{R}$, a cloning map poses difficulties.)

The composition $T := C \circ K \circ D^{\otimes M}$, mapping $B(\mathbb{C}^d)^{\otimes M}$ to $B(\mathbb{C}^d)$, is a so-called $1 \to M$ cloner. By construction of T, it is clear that $T(I \otimes \ldots \otimes B \otimes \ldots \otimes I) = CD(B)$. The main result of[3] then says that $\sup_{0 \leq B \leq I} \|T(I \otimes \ldots \otimes B \otimes \ldots \otimes I) - B\| \geq \frac{(M-1)}{M} \frac{d-1}{d+1}$. Since $M \in \mathbb{N}^+$ was arbitrary, this implies $\Delta \geq \frac{d-1}{d+1}$. The best possible coding occurs for $d = 2$, when $\Delta \geq \frac{1}{3}$.

6. Discussion

The number $(3-\sqrt{5})/4 \approx 0.19$, obtained from the Cauchy-Schwarz inequality, is inferior to the $\frac{1}{3} \approx 0.33$ from the no-cloning theorem, in usefulness as well as in the standard order on $\mathbb{R}$.

Although it yields inferior results for classical coding, the Cauchy-Schwarz method is simpler. At least taking into account the fairly heavy machinery needed in[3]. It is also very flexible. For any kind of quantum impossibility, ranging from 'no classical coding' to 'no measurement without disturbance', and from 'decoherence after measurement' to 'no joint measurement', the Cauchy-Schwarz inequality yields a bound which quantifies how far one is removed from the ideal, impossible situation. (See[1,2].)

Surprisingly enough, in the cases of 'no measurement without decoherence' 'no measurement without disturbance' and 'no joint measurement', this bound becomes sharp in the sense that there exist quantum operations which reach it. Furthermore, as far as I am aware, these three bounds cannot be obtained in any other way.

In conclusion, the Cauchy-Schwarz method yields quantitative bound for a variety of quantum impossibilities, including classical coding. Although this is not the case for classical coding, many of these bounds turn out to be sharp.

Bibliography

1. B. Janssens, '*Unifying Decoherence and the Heisenberg Principle*', `www.arxiv.org/quant-ph/0606093`.
2. B. Janssens, '*Quantum Measurement, a Coherent Description*', master's thesis, `www.arxiv.org/quant-ph/0503009`.
3. M. Keyl, R. Werner, '*Optimal Cloning of Pure States, Judging Single Clones*', J.Math.Phys. **40** 3283–3299, (1999).

4. J.D.M. Maassen, *'Quantum Probability Applied to the Damped Harmonic Oscillator'*, `www.arxiv.org/quant-ph/0411024`, (2004).
5. J.D.M. Maassen, *'Quantum Probability, Quantum Information and Quantum Computing'*, `www.math.kun.nl/medewerkers/maassen`, (2004).
6. M. Nielsen, I. Chuang, *'Quantum Computation and Quantum Information'*, Cambridge University Press, (2000).
7. W.F. Stinespring, *'Positive Functions on C^*-algebras'*, Proc. Amer. Math. Soc. **6**, 211–216, (1955).
8. R. Werner, *'Quantum Information Theory – an Invitation'*, Springer Tracts in Modern Physics, **173**, 14–57, (2001).

NOTE ON INFORMATION TRANSMISSION IN QUANTUM SYSTEMS

NOBORU WATANABE

Department of Information Sciences,
Tokyo University of Science
Noda City, Chiba, 278-8510, Japan
E-mail: watanabe@is.noda.tus.ac.jp

Dedicated to Professor Viacheslav P. Belavkin on his 60th birthday

The quantum entropy was introduced by von Neumann around 1932, which describes the amont of information of the quantum state itself. It was extended by Ohya for C*-systems. The quantum relative entropy was first defined by Umegaki for σ-finite von Neumann algebras, which denotes a certain difference between two states. It was extended by Araki and Uhlmann for general von Neumann algebras and *-algebras, respectively. The development of communication theory is closely connected with study of entropy theory. The mutual entropy denotes the amount of information correctly transmitted from the input system to the output system through the channel. The classical mutual entropy is defined by using the joint probability measure between the input and output systems, but there does not exist the joint states in the quantum systems. Ohya introduced the quantum mutual entropy by defining the compund states instead of the joint probability measures in the classical communication systems. Recently, Shor and Bennet et al formulated two mutual entropy-type measures for discussing the coding theorem in quatum communication processes. These mutual entropy-type measures are defined by using the entropy exchange relating the entanglement of the states. Belavkin and Ohya introduced the most poweful tool to measure the degree of entanglement.
In this paper, three mutual entropy-type measures are compared and one can conclude that the Ohya mutual entropy is the most suitable one for discussing the information transmission in quantum systems.

1. Introduction

In communication theory, there are two kind of important measures. One is called an entropy measuring the mount of information of a state itself and another is called a mutual entropy (information) measuring the mount of information correctly transmitted from an input system to an output

system through a channel. The signal of the input system is sent through a channel. In classical communication theory, the mutual entropy was defined by using the joint probability distribution between the input and output systems. The (semi-classical) mutual entropies for classical input and quantum output were studied by several researchers[8,9]. In fully quantum system, there does not exist the joint probability distribution in general. Instead of the joint probability distribution, Ohya introduced a compound state and defined Ohya mutual entropy[11] based on quantum relative entropy of Umegaki[25] in 1983, he extended it[12] to general quantum systems by using the relative entropy of Araki[2] and Uhlmann[26]. The channel is represented mathematically by a mapping from the input state system to the output state space. One of the example of the quantum channel is the quantum teleportation by using the quantum entanglement state of the input and out put systems. In[3], Belavkin and Ohya introduced the measure of entanglment for the states, which is one of the most effective tool for measuring the degree of entanglement of the quantum states.

Shor[24] and Bennet et al[4,5,22,23] took the entropy exchange and defined the mutual entropy-type measures so called the coherent entropy and the Lindblad entropy to discuss a sort of coding theorem for communication processes.

In this paper, we compare with mutual entropy-type measures and show one of the results for quantum capacity for the attenuation channel.

2. Quantum Channels

The concept of channel has been carried out an important role in the progress of the quantum communication theory. Here we review the definition of the quantum channels.

Let $\mathcal{H}_1, \mathcal{H}_2$ be complex separable Hilbert spaces of an input and an output systems, respectively, and let $\mathbf{B}(\mathcal{H}_k)$ be the set of all bounded linear operators on $\mathcal{H}_k$. We denote the set of all density operators on $\mathcal{H}_k \quad (k = 1, 2)$ by

$$\mathfrak{S}(\mathcal{H}_k) \equiv \{\rho \in \mathbf{B}(\mathcal{H}_k)\,;\, \rho \geq 0, tr\rho = 1\}\,. \tag{1}$$

A map Λ^* from the quantum input system to the quantum output system is called a **(fully) quantum channel**.

(1) Λ^* is called a **linear channel** if it satisfies the affine property, i.e.,

$$\sum_k \lambda_k = 1 \left(\forall \lambda_k \geq 0\right) \Rightarrow \Lambda^* \left(\sum_k \lambda_k \rho_k\right) = \sum_k \lambda_k \Lambda^* \left(\rho_k\right), \forall \rho_k \in S\left(\mathcal{H}_1\right).$$

2. $\Lambda^* : \mathfrak{S}\left(\mathcal{H}_1\right) \to \mathfrak{S}\left(\mathcal{H}_2\right)$ is called a **completely positive (CP) channel** if its dual map Λ satisfies

$$\sum_{j,k=1}^{n} B_j^* \Lambda \left(A_j^* A_k\right) B_k \geq 0 \tag{2}$$

for any $n \in \mathbb{N}$, any $B_j \in \mathbf{B}\left(\mathcal{H}_1\right)$ and any $A_k \in \mathbf{B}\left(\mathcal{H}_2\right)$, where the dual map $\Lambda : \mathbf{B}\left(\mathcal{H}_2\right) \to \mathbf{B}\left(\mathcal{H}_1\right)$ of $\Lambda^* : \mathfrak{S}\left(\mathcal{H}_1\right) \to \mathfrak{S}\left(\mathcal{H}_2\right)$ satisfies $tr\rho\Lambda\left(A\right) = tr\Lambda^*\left(\rho\right)A$ for any $\rho \in \mathfrak{S}\left(\mathcal{H}_1\right)$ and any $A \in \mathbf{B}\left(\mathcal{H}_2\right)$.

2.1. *Attenuation channel*

In particular, an attenuation channel introduced in[11] is one of the most inportaint model for discussing the information transmission in quantum optical communication.

Let us consider the communication processes including noise and loss systems. Let $\mathcal{K}_1, \quad \mathcal{K}_2$ be the complex separable Hilbert spaces for the noise and the loss systems, respectively. The quantun communication channel

$$\Lambda_0^*(\rho) \equiv tr_{\mathcal{K}_2}\pi_0^*\left(\rho \otimes \xi_0\right), \quad \xi_0 \equiv |0\rangle\langle 0| \text{ and } \pi_0^*\left(\cdot\right) \equiv V_0\left(\cdot\right)V_0^* \tag{3}$$

is called the attenuation channel, where $|0\rangle\langle 0|$ is vacuum state in $\mathcal{H}_1$ and V_0 is a linear mapping from $\mathcal{H}_1 \otimes \mathcal{K}_1$ to $\mathcal{H}_2 \otimes \mathcal{K}_2$ given by

$$V_0\left(|n\rangle \otimes |0\rangle\right) \equiv \sum_{j=0}^{n} C_j^n |j\rangle \otimes |n-j\rangle, \quad C_j^n = \sqrt{\frac{n!}{j!(n-j)!}}\alpha^j \overline{\beta}^{n-j} \tag{4}$$

for any $|n\rangle$ in $\mathcal{H}_1$ and α,β are complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$. $\eta = |\alpha|^2$ is the transmission rate of the channel. $\pi_0^*$is called a beam splittings, which means that one beam comes and two beams appear after passing throughπ_0^*. This attenuation channel is generalized by Ohya and Watanabe such as noisy optical channel[18,19]. After that, Accardi and Ohya[1] reformulated it by using liftings, which is the dual map of the transition expectation by mean of Accardi. It contains the concept of beam splittings, which is extended by Fichtner, Freudenberg and Libsher[7] concerning the mappings on generalized Fock spaces. For the attenuation channel Λ_0^*, one can obtain the following theorem:

Theorem 2.1. *The attenuation channel* Λ_0^* *is described by*

$$\Lambda_0^*(\rho) = \sum_{i=0}^{\infty} O_i V_0 Q \rho Q^* V_0^* O_i^*, \tag{5}$$

where $Q \equiv \sum_{l=0}^{\infty} (|y_l\rangle \otimes |0\rangle) \langle y_l|$, $O_i \equiv \sum_{k=0}^{\infty} |z_k\rangle (\langle z_k| \otimes \langle i|)$, $\{|y_l\rangle\}$ *is a CONS in*$\mathcal{H}_1$, $\{|z_k\rangle\}$ *is a CONS in* $\mathcal{H}_2$ *and* $\{|i\rangle\}$ *is the set of number states in* $\mathcal{K}_2$.

3. Ohya $\mathcal{S}$-Mixing Entropy

The quantum entropy was introduced by von Neumann around 1932[10], which is defined by

$$S(\rho) \equiv -tr\rho \log \rho$$

for any density operators ρ in $S(\mathcal{H}_1)$. It denotes the amount of information of the quantum state ρ. It was extended by Ohya[13] for general quantum systems as follows.

Let $(\mathcal{A}, S(\mathcal{A}))$ be a C*-system. The entropy of a state $\varphi \in \mathcal{S}$ seen from the reference system, a weak$*$-compact convex subset of the whole state space $S(\mathcal{A})$ on the C*-algebra $\mathcal{A}$, was introduced by Ohya, which is called a Ohya $\mathcal{S}$-mixing entropy. This Ohya $\mathcal{S}$-mixing entropy contains von Neumann's entropy and classical entropy as special cases.

Every state $\varphi \in \mathcal{S}$ has a maximal measure μ pseudosupported on ex$\mathcal{S}$ (extreme points in $\mathcal{S}$) such that

$$\varphi = \int_{\mathrm{ex}\mathcal{S}} \omega d\mu. \tag{6}$$

The measure μ giving the above decomposition is not unique unless $\mathcal{S}$ is a Choquet simplex, so that we denote the set of all such measures by $M_\varphi(\mathcal{S})$. Take

$$D_\varphi(\mathcal{S}) \equiv \left\{ \mu \in M_\varphi(\mathcal{S});\ \exists\{\mu_k\} \subset \mathbb{R}^+ \text{ and } \{\varphi_k\} \subset \mathrm{ex}\mathcal{S} \text{ s.t.} \right.$$
$$\left. \sum_k \mu_k = 1,\ \mu = \sum_k \mu_k \delta(\varphi_k) \right\}, \tag{7}$$

where $\delta(\varphi)$ is the delta measure concentrated on $\{\varphi\}$. Put

$$H(\mu) = -\sum_k \mu_k \log \mu_k \tag{8}$$

for a measure$\mu \in D_\varphi(\mathcal{S})$.

Ohya $\mathcal{S}$-mixing entropy of a general state $\varphi \in \mathcal{S}$ w.r.t. $\mathcal{S}$ is defined by

$$S^{\mathcal{S}}(\varphi) = \begin{cases} \inf\{H(\mu); \mu \in D_\varphi(\mathcal{S})\} & (D_\varphi(\mathcal{S}) \neq \varnothing) \\ \infty & (D_\varphi(\mathcal{S}) = \varnothing) \end{cases} \tag{9}$$

When $\mathcal{S}$ is the total space $S(\mathcal{A})$, we simply denote $S^{\mathcal{S}}(\varphi)$ by $S(\varphi)$. This entropy (mixing $\mathcal{S}$-entropy) of a general state φ satisfies the following properties[13].

Theorem 3.1. *When $\mathcal{A} = \mathbf{B}(\mathcal{H})$ and $\alpha_t = Ad(U_t)$ (i.e., $\alpha_t(A) = U_t^* A U_t$ for any$A \in \mathcal{A}$) with a unitary operator U_t, for any state φ given by $\varphi(\cdot) = tr\rho\cdot$ with a density operator ρ, the following facts hold:*

(1) $S(\varphi) = -tr\rho\log\rho$.

(2) If φ is an α-invariant faithful state and every eigenvalue of ρ is non-degenerate, then $S^{I(\alpha)}(\varphi) = S(\varphi)$, where $I(\alpha)$ is the set of all α-invariant faithful states.

(3) If $\varphi \in K(\alpha)$, then $S^{K(\alpha)}(\varphi) = 0$, where $K(\alpha)$ is the set of all KMS states.

Theorem 3.2. *For any $\varphi \in K(\alpha)$, we have*

(1) $S^{K(\alpha)}(\varphi) \leq S^{I(\alpha)}(\varphi)$.

(2) $S^{K(\alpha)}(\varphi) \leq S(\varphi)$.

This Ohya $\mathcal{S}$-mixing entropy gives a measure of the uncertainty observed from the reference system $\mathcal{S}$ so that it has the following merits: Even if the total entropy $S(\varphi)$ is infinite, $S^{\mathcal{S}}(\varphi)$ is finite for some $\mathcal{S}$, hence it explains a sort of symmetry breaking in $\mathcal{S}$. Other similar properties as $S(\rho)$ hold for $S^{\mathcal{S}}(\varphi)$. This entropy can be applied to characterize normal states and quantum Markov chains in von Neumann algebras.

The relative entropy for two general states φ and ψ was introduced by Araki and Uhlmann and their relation is studied by Donald, Petz and Hiai et al.

4. Quantum Relative Entropy

4.1. *Umegaki's definition*

Let $\mathbf{B}(\mathcal{H})$ be the set of all bounded linear operators on a Hilbert space $\mathcal{H}$ and ρ, σ be density operators on $\mathcal{H}$. The Umegaki's relative entropy[25] with respect to ρ and σ is defined by

$$S(\rho, \sigma) \equiv \begin{cases} tr\rho(\log\rho - \log\sigma) & (\text{when } \overline{ran\rho} \subset \overline{ran\sigma}) \\ \infty & (\text{otherwise}) \end{cases} \tag{10}$$

It represents a certain difference between two quantum states ρ, σ. There were several trials to extend the relative entropy to more general quantum systems and apply it to some other fields[2,13,14,26].

4.2. *Araki's definition*

Let N be σ-finite von Neumann algebra acting on a Hilbert space $\mathcal{H}$ and φ, ψ be normal states on N given by $\varphi(\cdot) = \langle x, \cdot x\rangle$ and $\psi(\cdot) = \langle y, \cdot y\rangle$ with $x, y \in \mathcal{K}$ (a positive natural cone). The operator $S_{x,y}$ is defined by

$$S_{x,y}(Ay + z) = s^N(y)A^*x,\ A \in N,\ s^{N'}(y)z = 0, \tag{11}$$

on the domain $\mathfrak{N}y + (I - s^{\mathfrak{N}'}(y))\mathcal{H}$, where $s^{\mathfrak{N}}(y)$ is the projection from $\mathcal{H}$ to $\{\mathfrak{N}'y\}^-$, the $\mathfrak{N}$ -support of$y$. Using this $S_{x,y}$, the relative modular operator $\Delta_{x,y}$ is defined as $\Delta_{x,y} = (S_{xy})^* \overline{S_{x,y}}$, whose spectral decomposition is denoted by $\int_0^\infty \lambda de_{x,y}(\lambda)$ ($\overline{S_{x,y}}$ is the closure of $S_{x,y}$). Then the Araki relative entropy[2] is given by

$$S(\psi, \varphi) = \begin{cases} \int_0^\infty \log \lambda d\,\langle y, e_{x,y}\,(\lambda)\, y\rangle & (\psi \ll \varphi) \\ \infty & (\text{otherwise}) \end{cases}, \tag{12}$$

where $\psi \ll \varphi$ means that $\varphi(A^*A) = 0$ implies $\psi(A^*A) = 0$ for $A \in \mathfrak{N}$.

4.3. *Uhlmann's definition*

Let $\mathcal{L}$ be a complex linear space and p, q be two seminorms on$\mathcal{L}$. Moreover, let $H(L)$ be the set of all positive hermitian forms α on $\mathcal{L}$ satisfying $|\alpha(x, y)| \leq p(x)q(y)$ for all$x, y \in \mathcal{L}$. Then the quadratical mean $QM(p, q)$ of p and q is defined by

$$QM(p, q)(x) = \sup\{\alpha(x, x)^{1/2};\ \ \alpha \in H(\mathcal{L})\},\ \ x \in \mathcal{L}, \tag{13}$$

and there exists a function $p_t(x)$ of $t \in [0, 1]$ for each $x \in \mathcal{L}$ satisfying the following conditions:

(1) For any$x \in \mathcal{L}$, $p_t(x)$ is continuous in t,
(2) $p_{1/2} = QM(p, q)$,
(3) $p_{t/2} = QM(p, p_t)$,
(4) $p_{(t+1)/2} = QM(p_t, q)$.

This seminorm p_t is denoted by $QI_t(p, q)$ and is called the quadratical interpolation from p to$q$. It is shown that for any positive hermitian formsα, β, there exists a unique function $QF_t(\alpha, \beta)$ of $t \in [0, 1]$ with values in the set $H(\mathcal{L})$ such that $QF_t(\alpha, \beta)(x, x)^{1/2}$ is the quadratical interpolation

from $\alpha(x,x)^{1/2}$ to $\beta(x,x)^{1/2}$. The relative entropy functional $S(\alpha,\beta)(x)$ of α and β is defined as

$$S(\alpha,\beta)(x) = -\liminf_{t\to 0}\frac{1}{t}\left\{QF_t(\alpha,\beta)(x,x) - \alpha(x,x)\right\} \tag{14}$$

for $x \in \mathcal{L}$. Let $\mathcal{L}$ be a *-algebra $\mathcal{A}$ and φ, ψ be positive linear functionals on $\mathcal{A}$ defining two hermitian forms φ^L, ψ^R such as $\varphi^L(A,B) = \varphi(A^*B)$ and $\psi^R(A,B) = \psi(BA^*)$.

The Uhlmanns relative entropy[26] of φ and ψ is defined by

$$S(\psi,\varphi) = S(\psi^R,\varphi^L)(I). \tag{15}$$

5. Ohya Mutual Entropy for Genaral Quantum Systems

The classical mutual entropy is determined by an input state and a channel, so that we denote the quantum mutual entropy with respect to the input state φ and the quantum channel Λ^* by $I(\varphi;\Lambda^*)$. This quantum mutual entropy $I(\varphi;\Lambda^*)$ should satisfy the following three conditions:

(1) The quantum mutual entropy is well-matched to the von Neumann entropy. That is, if a channel is trivial, i.e., Λ^* = identity map, then the mutual entropy equals to the von Neumann entropy: $I(\varphi;id) = S(\varphi)$.

(2) When the system is classical, the quantum mutual entropy reduces to classical one.

(3) Shannon's fundamental inequality $0 \le I(\varphi;\Lambda^*) \le S(\varphi)$ is held.

In order to define such a quantum mutual entropy, we need the quantum relative entropy and the joint state, which is called a compound state, describing the correlation between an input state φ and the output state $\Lambda^*\varphi$ through a channel Λ^*. For $\varphi \in \mathcal{S} \subset S(\mathcal{A})$ and $\Lambda^* : S(\mathcal{A}) \to S(\overline{\mathcal{A}})$, the compound states are defined by

$$\Phi_\mu^{\mathcal{S}} = \int_{\mathcal{S}} \omega \otimes \Lambda^*\omega d\mu \tag{16}$$

and

$$\Phi_0 = \varphi \otimes \Lambda^*\varphi. \tag{17}$$

The first compound state, which is called a Ohya compund state, generalizes the joint probability in classical dynamical system and it exhibits the correlation between the initial state φ and the final state $\Lambda^*\varphi$.

Ohya mutual entropy w.r.t. $\mathcal{S}$ and μ is

$$I_\mu^{\mathcal{S}}(\varphi;\Lambda^*) = S(\Phi_\mu^{\mathcal{S}},\Phi_0) \tag{18}$$

and Ohya mutual entropy[13] w.r.t. $\mathcal{S}$ is defined by

$$I^{\mathcal{S}}(\varphi;\Lambda^*) = \limsup_{\varepsilon\to 0} \left\{ I^{\mathcal{S}}_{\mu}(\varphi;\Lambda^*); \mu \in F^{\varepsilon}_{\varphi}(\mathcal{S}) \right\}, \tag{19}$$

where

$$F^{\varepsilon}_{\varphi}(\mathcal{S}) = \begin{cases} \left\{ \mu \in D_{\varphi}(\mathcal{S}); S^{\mathcal{S}}(\varphi) \leq H(\mu) \leq S^{\mathcal{S}}(\varphi) + \varepsilon < +\infty \right\} & \left(S^{\mathcal{S}}(\varphi) < +\infty\right) \\ M_{\varphi}(\mathcal{S}) & \left(S^{\mathcal{S}}(\varphi) = +\infty\right) \end{cases} \tag{20}$$

The following fundamental inequality is satisfied for almost all physical cases[14].

$$0 \leq I^{\mathcal{S}}(\varphi;\Lambda^*) \leq S^{\mathcal{S}}(\varphi) \tag{21}$$

In the case that the C*-algebra is $\mathbf{B}(\mathcal{H})$ and S is the set of all density operators, the above Ohya mutual entropy goes to

$$I(\rho;\Lambda^*) = \sup \left\{ \sum_n S(\Lambda^* E_n, \Lambda^* \rho),\ \rho = \sum_n \lambda_n E_n \right\}, \tag{22}$$

where ρ is a density operator (state), $S(\Lambda^* E_n, \Lambda^* \rho)$is Umegaki's relative entropy and $\rho = \sum_{\sum} nE_n$ is Schatten-von Neumann (one dimensional spectral) decomposition. As was mentioned above, it satisfies the Shannon's type inequality as follows:$0 \leq I(\rho,\Lambda^*) \leq \min\{S(\rho), S(\Lambda^*\rho)\}$. It is easily shown that we can take orthogonal decomposition instead of the Schatten-von Neumann decomposition[21].

5.1. *Semi-classical mutual entropy*

When the input system is classical, the state φ is a probability distribution and the Schatten-von Neumann decomposition is unique with delta measures δ_n such that $\varphi = \sum_n \lambda_n \delta_n$. In this case we need to code the classical state φ by a quantum state ψ, whose process is a quantum coding described by a channel Γ^* such that $\Gamma^*\delta_n = \psi_n$ (quantum state) and$\psi \equiv \Gamma^*\varphi = \sum_n \lambda_n \psi_n$. Then Ohya mutual entropy $I(\varphi;\Lambda^* \circ \Gamma^*)$ becomes Holevo's one, that is,

$$I(\varphi;\Lambda^* \circ \Gamma^*) = S(\Lambda^*\psi) - \sum_n \lambda_n S(\Lambda^*\psi_n) \tag{23}$$

when $\sum_n \lambda_n S(\Lambda^*\psi_n)$ is finite. These Ohya mutual entropy (ME) are completely quantum, namely, it represents the information transmission from a quantum input to a quantum output. The quantum system is described by a noncommutative structure. The classical system is expressed by a commutative construction. In the mathematical point of view, the commutative

systems are contained in the noncommutative framework. One can obtain the following diagram.

$$\begin{array}{ccccc} & \swarrow & \text{Semi-classical ME} & \longleftarrow & \text{Ohya ME} \\ \text{Shannon's ME} & & & & \uparrow \\ & \nwarrow & \text{ME (GKY)} & \longleftarrow & \text{Ohya ME for GQS} \end{array}$$

6. Quantum Mutual Entropy Type Measures

Recently Shor[24] and Bennet et al[4,5] took the coherent entropy and defined the mutual type entropy to discuss a sort of coding theorem for quantum communication. In this section, we compare these mutual types entropy.

Let us discuss the entropy exchange[22]. For a stateρ, a channel Λ^* is denoted by using an operator valued measure $\{A_j\}$ such as

$$\Lambda^*(\cdot) \equiv \sum\nolimits_j A_j^* \cdot A_j, \tag{24}$$

which is called a Stinespring-Sudarshan-Kraus form. Then one can define a matrix $W = (W_{ij})_{i,j}$ with

$$W_{ij} \equiv trA_i^* \rho A_j, \tag{25}$$

by which the entropy exchange is defined by

$$S_e(\rho, \Lambda^*) = -trW \log W. \tag{26}$$

By using the entropy exchange, two mutual type entropies are defined as follows:

$$I_C(\rho; \Lambda^*) \equiv S(\Lambda^* \rho) - S_e(\rho, \Lambda^*), \tag{27}$$

$$I_L(\rho; \Lambda^*) \equiv S(\rho) + S(\Lambda^* \rho) - S_e(\rho, \Lambda^*). \tag{28}$$

The first one is called the coherent entropy $I_C(\rho; \Lambda^*)$[23] and the second one is called the Lindblad entropy $I_L(\rho; \Lambda^*)$[5]. By comparing these mutual entropies for quantum information communication processes, we have the following theorem[20] :

Theorem 6.1. *Let $\{A_j\}$ be a projection valued measure with $dimA_j = 1$. For arbitrary state ρ and the quantum channel $\Lambda^*(\cdot) \equiv \sum_j A_j \cdot A_j^*$, one has*

(1) $0 \leq I(\rho; \Lambda^*) \leq \min\{S(\rho), S(\Lambda^* \rho)\}$ *(Ohya mutual entropy),*
(2) $I_C(\rho; \Lambda^*) = 0$ *(coherent entropy),*
(3) $I_L(\rho; \Lambda^*) = S(\rho)$ *(Lindblad entropy).*

For the attenuation channel Λ_0^*, one can obatain the following theorems[20] :

Theorem 6.2. *For any state* $\rho = \sum_n \lambda_n |n\rangle \langle n|$ *and the attenuation channel* Λ_0^* *with* $|\alpha|^2 = |\beta|^2 = \frac{1}{2}$, *one has*

(1) $0 \leq I(\rho; \Lambda_0^*) \leq \min\{S(\rho), S(\Lambda_0^* \rho)\}$ *(Ohya mutual entropy),*

(2) $I_C(\rho; \Lambda_0^*) = 0$ *(coherent entropy),*

(3) $I_L(\rho; \Lambda_0^*) = S(\rho)$ *(Lindblad entropy).*

Theorem 6.3. *For the attenuation channel* Λ_0^* *and the input state* $\rho = \lambda |0\rangle \langle 0| + (1-\lambda) |\theta\rangle \langle \theta|$, *we have*

(1) $0 \leq I(\rho; \Lambda_0^*) \leq \min\{S(\rho), S(\Lambda_0^* \rho)\}$ *(Ohya mutual entropy),*

(2) $-S(\rho) \leq I_C(\rho; \Lambda_0^*) \leq S(\rho)$ *(coherent entropy),*

(3) $0 \leq I_L(\rho; \Lambda_0^*) \leq 2S(\rho)$ *(Lindblad entropy).*

Therem 6.3 shows that the coherent entropy $I_C(\rho; \Lambda_0^*)$ takes a minus value for $|\alpha|^2 < |\beta|^2$ and the Lindblad entropy $I_L(\rho; \Lambda_0^*)$ is grater than the von Neumann entropy of the input state ρ for $|\alpha|^2 > |\beta|^2$. From these theorems, the coherent entropy and the Lindblad entropy can be used for the measure of the tendency of entanglement of the quantum states, but these are not fit for measuring the amount of information correctly transmitted from the input system to the output system through the quantum channel. Therefore Ohya mutual entropy $I(\rho; \Lambda^*)$ only satisfies the above three conditions of the mutual entropy for information transmission in quantum systems. One can conclude that the Ohya mutual entropy is a most suitable one as quantum mutual entropy including the classical mutual entropy.

7. Quantum Capacity

The capacity means the ability of the information transmission of the channel, which is used as a measure for construction of channels. The fully quantum capacity is formulated by taking the supremum of the fully quantum mutual entropy with respect to a certain subset of the initial state space. The capacity of purely quantum channel was studied in[15–18].

Let $\mathcal{S}$ be the set of all input states satisfying some physical conditions. Let us consider the ability of information transmission for the quantum channelΛ^*. The answer of this question is the capacity of quantum channel Λ^* for a certain set $\mathcal{S} \subset S(\mathcal{H}_1)$ defined by

$$C_q^{\mathcal{S}}(\Lambda^*) \equiv \sup\{I(\rho; \Lambda^*); \rho \in \mathcal{S}\}. \tag{29}$$

When $\mathcal{S} = S(\mathcal{H}_1)$, the capacity of quantum channel Λ^* is denoted by

$C_q(\Lambda^*)$. Then the following theorem for the attenuation channel was proved in.[20]

Theorem 7.1. *For a subset* $\mathcal{S}_n \equiv \{\rho \in S(\mathcal{H}_1); \dim s(\rho) = n\}$, *the capacity of the attenuation channel* Λ_0^* *satisfies*

$$C_q^{\mathcal{S}_n}(\Lambda_0^*) = \log n,$$

where $s(\rho)$ *is the support projection of* ρ.

When the mean energy of the input state vectors $\{|\tau\theta_k\rangle\}$ can be taken infinite, i.e., $\lim_{\tau\to\infty}|\tau\theta_k|^2 = \infty$, the above theorem tells that the quantum capacity for the attenuation channel Λ_0^* with respect to $\mathcal{S}_n$ becomes $\log n$. It is a natural result, however it is impossible to take the mean energy of input state vector infinite. Therefore we have to compute the quantum capacity

$$C_q^{\mathcal{S}_e}(\Lambda^*) = \sup\{I(\rho;\Lambda^*); \rho \in \mathcal{S}_e\} \tag{30}$$

under some constraint $\mathcal{S}_e \equiv \{\rho \in \mathcal{S}; E(\rho) < e\}$ on the mean energy $E(\rho)$ of the input state ρ. In,[12,15] we also considered the **pseudo-quantum capacity** $C_{pq}^{\mathcal{S}_e}(\Lambda^*)$ defined by

$$C_{pq}^{\mathcal{S}_e}(\Lambda^*) = \sup\{I_p(\rho;\Lambda^*); \rho \in \mathcal{S}_e\} \tag{31}$$

with the **pseudo-mutual entropy** $I_{pq}(\rho;\Lambda^*)$

$$I_{pq}(\rho;\Lambda^*) = \sup\left\{\sum_k \lambda_k S(\Lambda^*\rho_k, \Lambda^*\rho); \rho = \sum_k \lambda_k\rho_k, \text{finite decomposition}\right\}, \tag{32}$$

where the supremum is taken over all finite decompositions instead of all orthogonal pure decompositions for purely quantum mutual entropy.

Bibliography

1. Accardi, L., and Ohya, M., Compond channnels, transition expectation and liftings, Appl. Math, Optim., **39**, 33-59 (1999).
2. H. Araki, Relative entropy for states of von Neumann algebras, Publ. RIMS Kyoto Univ. **11**, 809–833, 1976.
3. Belavkin, V. P. and Ohya, M., Quantum entanglements and entangled mutual entropy, Proc. R. Soc. Lond. A., **458**, 209-231 (2002).
4. Barnum, H., Nielsen, M.A., and Schumacher, B.W., Information transmission through a noisy quantum channel, Physical Review A, **57**, No.6, 4153-4175 (1998).

5. Bennett, C.H., Shor, P.W., Smolin, J.A., and Thapliyalz, A.V., Entanglement-Assisted Capacity of a Quantum Channel and the Reverse Shannon Theorem, quant-ph/0106052.
6. I. Csiszar, I - divergence geometry of probability distributions and minimization problems, Ann. Probab. **3**, 146–158, 1975.
7. Fichtner, K.H., Freudenberg, W., and Liebscher, V., Beam splittings and time evolutions of Boson systems, Fakultat fur Mathematik und Informatik, Math/ Inf/96/ 39, Jena, 105 (1996).
8. Holevo, A.S., Some estimates for the amount of information transmittable by a quantum communication channel (in Russian)}, Problemy Peredachi Informacii, **9**, 3-11 (1973).
9. Ingarden, R.S., Kossakowski, A., and Ohya, M., Information Dynamics and Open Systems, Kluwer, 1997.
10. von Neumann, J., Die Mathematischen Grundlagen der Quantenmechanik, Springer-Berlin, 1932.
11. Ohya, M., On compound state and mutual information in quantum information theory, IEEE Trans. Information Theory, **29**, 770-774 (1983).
12. M. Ohya, Some aspects of quantum information theory and their applications to irreversible processes, Rep. Math. Phys. **27**, 19–47, 1989.
13. Ohya, M., Some aspects of quantum information theory and their applications to irreversible processes, Rep. Math. Phys., **27**, 19-47 (1989).
14. Ohya, M., and Petz, D., Quantum Entropy and its Use, Springer, Berlin, 1993.
15. Ohya, M., Petz, D., and Watanabe, N., On capacity of quantum channels, Probability and Mathematical Statistics, **17**, 179-196 (1997).
16. Ohya, M., Petz, D., and Watanabe, N., Numerical computation of quantum capacity, International Journal of Theoretical Physics, **37**, No.1, 507-510 (1998).
17. Ohya, M., and Watanabe, N., Quantum capacity of noisy quantum channel, Quantum Communication and Measurement, **3**, 213-220 (1997).
18. Ohya, M., and Watanabe, N., Foundatin of Quantum Communication Theory (in Japanese), Makino Pub. Co., 1998.
19. Ohya, M., and Watanabe, N., Construction and analysis of a mathematical model in quantum communication processes, Electronics and Communications in Japan, Part 1, **68**, No.2, 29-34 (1985).
20. Ohya, M., and Watanabe, N., Comparison of mutual entropy - type measures, TUS preprint.
21. Schatten, R., Norm Ideals of Completely Continuous Operators, Springer-Verlag, 1970.
22. Schumacher, B.W., Sending entanglement through noisy quantum channels, Physical Review A, **54**, 2614 (1996).
23. Schumacher, B.W., and Nielsen, M.A., Quantum data processing and error correction, Physical Review A, **54**, 2629 (1996).
24. Shor, P., The quantum channel capacity and coherent information, Lecture Notes, MSRI Workshop on Quantum Computation, 2002.
25. Umegaki, H., Conditional expectations in an operator algebra IV (entropy

and information), Kodai Math. Sem. Rep., **14**, 59-85 (1962).
26. A.Uhlmann, Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in interpolation theory, Commun. Math. Phys., **54**, 21–32, 1977.

This page intentionally left blank

AUTHOR INDEX