

Quantum filtering using POVM measurements

Ram A. Somaraju, Alain Sarlette and Hugo Thienpont

Abstract—The objective of this work is to develop a recursive, discrete time quantum filtering equation for a system that interacts with a probe, on which measurements are performed according to the Positive Operator Valued Measures (POVMs) framework. POVMs are the most general measurements one can make on a quantum system and although in principle they can be reformulated as projective measurements on larger spaces, for which filtering results exist, a direct treatment of POVMs is more natural and can simplify the filter computations for some applications. Hence we formalize the notion of strongly commuting (Davies) instruments which allows one to develop joint measurement statistics for POVM type measurements. This allows us to prove the existence of conditional POVMs, which is essential for the development of a filtering equation. We demonstrate that under generally satisfied assumptions, knowing the observed probe POVM operator is sufficient to uniquely specify the quantum filtering evolution for the system.

I. INTRODUCTION

The theory of filtering considers the estimation of the system state from noisy and/or partial observations (see e.g. [1]). For quantum systems, filtering theory was initiated in the 1980s by Belavkin in a series of papers [2], [3], [4], [5]. Belavkin makes use of the operational formalism of Davies [6], which is a precursor to the theory of quantum filtering. He has also realized that due to the unavoidable back-action of quantum measurements, the theory of filtering plays a fundamental role in quantum feedback control (see e.g. [3], [5]). The theory of quantum filtering was independently developed in the physics community, particularly in the context of quantum optics, under the name of quantum trajectory theory [7], [8], [9], [10].

The basic model used to derive filtering equations for a quantum system uses a system-probe interaction. A quantum system, whose state needs to be estimated, is made to interact with a probe and the state of the system becomes entangled with that of the probe. After this interaction, an observable is measured on the probe and this measurement outcome is used to estimate the state of the system. The commutativity of

any system observable with any probe observable is used to develop a recursive Markov filtering equation for the system observables (see e.g. [11], [12] for an excellent tutorial).

Suppose \mathcal{H}_S is the Hilbert space corresponding to the system whose state needs to be estimated and \mathcal{H}_P is the Hilbert space of the probe. According to the classical von Neumann definition, any probe observable is a self-adjoint operator Q on \mathcal{H}_P ; the measurement of such an observable results in an outcome that is (stochastically) an eigenvalue of Q , and the probe state after measurement gets projected onto the corresponding eigenspace of Q . As far as we are aware, all discussions on quantum filtering theory so far have assumed that the probe undergoes such a von Neumann measurement, also called projective measurement or Projection Valued Measure (PVM). However, a more modern treatment of quantum measurement theory shows that the most general possible quantum measurements that one can perform are the so-called Positive Operator Valued Measures¹ (POVMs), of which von Neumann measurements are merely a special case where all the operators are commuting projections [6], [13]. See Section II for a brief overview of POVMs.

POVMs on \mathcal{H}_P can be reformulated as the restriction to \mathcal{H}_P of a PVM on a larger space. However, there is no canonical PVM that corresponds uniquely to a given POVM. This is closely related to the fact that the state of a quantum system after a POVM measurement is not uniquely determined as a function of the POVM. To remedy the latter situation, Davies [6, Ch.3] has shown that one can associate a (non-unique) instrument to any POVM, which determines a completely positive map that specifies the state after measurement conditioned on the measurement outcome. However, there is again no canonical instrument that corresponds to a given POVM. Therefore, it is impossible to uniquely specify the post-measurement state for a given POVM measurement outcome unless the instrument associated with the measurement is known. As stated by Nielsen and Chuang [14, p. 91], “POVMs are best viewed as [...] providing the simplest means by which one can study general measurement statistics, without the necessity for knowing the post-measurement state.” As such they are a minimal description of quantum measurements, so one can hope that the POVM formalism leads to more concise and fast filtering equations, suited for (possibly analog) implementation in real-time quantum feedback experiments.

There are other reasons to develop a POVM-based filtering theory that shortcuts the lift to PVMs in larger spaces. For

Ram A. Somaraju and Hugo Thienpont are with the Brussels Photonics Team, Dept. of Applied Physics and Photonics, Vrije Universiteit Brussel, Pleinlaan 2-1050 Brussels, Belgium. Alain Sarlette is with the SYSTeMS research group, Faculty of Engineering and Architecture, Ghent University, Technologiepark Zwijnaarde 914, 9052 Zwijnaarde(Ghent), Belgium. a.somaraju@gmail.com, alain.sarlette@ugent.be, hthienpo@b-phot.org

R. A. Somaraju and H.Thienpont would like to thank Methusalem project IPARC@VUB for financial support as well as the BELSPO IAP project Photonics@BE. A.Sarlette and R. A. Somaraju are members of the BELSPO IAP project DYSCO.

This paper presents research results funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

¹The terminology is not standard and other terms such as Positive Operator Measures or generalized measurements are also used.

one, once the POVM theory is available it can be more natural to use, as several practical measurement setups are based on POVMs, such as e.g. approximate position/momentum measurements [6, Ch. 3] or phase measurements [13], [15]. In some infinite-dimensional situations there can even be conceptual barriers to a PVM viewpoint. Indeed, with phase measurements after much investigation there is still no universal agreement on an acceptable PVM [16]. Finally, regarding system identification, the POVM associated with any experimental setup can (at least conceptually) be directly deduced from measurement outcomes; in contrast, in order to ascertain the associated instrument it is necessary to analyze the post-measurement state of the system (for more details see Section II-A). This is not generally feasible in practical experimental setups where often the measurements destroy the quantum state (e.g. photo-detection) and/or non-measurably alter it due to interaction with the environment.

In this paper, we develop a discrete time filtering equation for the system state conditioned on POVM measurements performed on a probe. After reviewing the POVM formalism (Section II), we provide a general theory about (strongly) commuting instruments that are associated to POVMs (Section III). We then generalize the filtering framework of [11], [12] from PVMs to POVMs. First Section IV-B illustrates POVM-specific difficulties in the conditional expectation approach. Section IV-C then defines conditional probabilities for POVMs. In the setup consisting of a probe coupled to a target system, any (physically reasonable) instrument associated with a POVM acting only on the probe, strongly commutes with any instrument associated with a POVM acting only on the target system. In Section V we show how this allows to define a filtering equation for the system state conditioned on probe POVM measurement outcomes. This filtering equation is only a function of the observed probe POVM and does not depend on the associated instrument nor other POVM elements.

II. REVIEW: POVMs AND ASSOCIATED INSTRUMENTS

The POVM formalism is a standard part of most modern quantum information textbooks. We briefly review it here and refer the interested reader to [6, Ch. 3] for more details.

Consider a quantum system with Hilbert space \mathcal{H} , i.e. the system state is given by a density operator ρ which is a unit-trace nonnegative self-adjoint linear operator on \mathcal{H} . We use $*$ to denote the adjoint. Denote by $\mathcal{L}(\mathcal{H})$ the set of linear operators on \mathcal{H} , by $\mathcal{L}_+(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ the set of self-adjoint nonnegative linear operators, and by $\mathcal{S}(\mathcal{H}) \subset \mathcal{L}_+(\mathcal{H})$ the set of all possible density operators (i.e. non-negative trace class operators of unit trace). Standard textbook treatment of quantum measurements assumes that any physically measurable quantity \hat{A} is associated to a self-adjoint operator $A : \mathcal{H} \rightarrow \mathcal{H}$. Because A is self-adjoint, we have the spectral decomposition²

$$A = \sum_{\omega \in \Omega} \omega P_\omega \quad (1)$$

²For clarity of explanation, we assume that A has a discrete spectrum. The discussion easily generalizes to the continuous spectrum situation.

where Ω is the set of eigenvalues of A and P_ω is the eigenprojection corresponding to eigenvalue ω . Starting with a system in state ρ , according to von Neumann's measurement postulates we have:

- 1) any measurement of the observable A gives some outcome $\omega \in \Omega$ with probability $\text{Tr}\{\rho P_\omega\}$, and
- 2) after measurement outcome ω , the state of the system becomes

$$\rho' = \frac{P_\omega \rho P_\omega}{\text{Tr}\{P_\omega \rho P_\omega\}}.$$

The first postulate can be thought of as: Ω is the set of all possible measurement outcomes of an experimental setup (\hat{A}) and to each $\omega \in \Omega$, one assigns a projection P_ω in \mathcal{H} such that $\text{Tr}\{\rho P_\omega\}$ is the probability of measuring ω . This motivates the following generalization.

Definition 2.1: [6, Def 3.1.1] Let Ω be a set, \mathcal{F} a σ -field of subsets of Ω , and \mathcal{H} a Hilbert space. Then a \mathcal{H} -valued Positive Operator Valued Measure (POVM) on Ω is a map $\hat{A} : \mathcal{F} \rightarrow \mathcal{L}_+(\mathcal{H})$; $E \rightarrow \hat{A}(E) = \hat{A}_E$ such that

- 1) $\hat{A}(E) \geq \hat{A}(\emptyset) = 0$ for all $E \in \mathcal{F}$;
- 2) For any countable, mutually disjoint collection $\{E_n\} \subset \mathcal{F}$ we have $\hat{A}(\bigcup_n E_n) = \sum_n \hat{A}(E_n)$ where the series convergence on the right is in the weak operator topology;
- 3) $\hat{A}(\Omega) = \mathbb{I}_{\mathcal{H}}$, the identity operator on \mathcal{H} .

A POVM that corresponds to a physical experiment has a simple interpretation. The set Ω is the sample space corresponding to experimental outcomes so that the σ -field \mathcal{F} consists of the set of all events. The POVM \hat{A} and a state ρ on \mathcal{H} induce a measure $\mu_{\rho, \hat{A}}(\cdot) = \text{Tr}\{\rho \hat{A}(\cdot)\}$ on Ω so that $\mu_{\rho, \hat{A}}(E)$ gives the probability of event $E \in \mathcal{F}$. We use the notation $\hat{A} \in E$ to denote the event that the measurement of POVM \hat{A} resulted in a value in $E \in \mathcal{F}$.

Note that here Ω can have any general structure. This allows one to describe measurement apparatuses with outcomes that are physically e.g. multi-dimensional or on a manifold topology like the circle or sphere, which is not possible with standard von Neumann measurements. The latter are indeed equivalent to a special case of POVMs called *Projection Valued Measures (PVMs)*, which require that Ω is a closed subset of the real line and that the range of \hat{A} only consists of commuting projections. The unique correspondence between PVMs (\hat{A}) and self-adjoint operators describing von Neumann measurements (A) is obtained through (1) by setting $P_\omega = \hat{A}(\omega)$ for each $\omega \in \Omega$.

It has been shown that any POVM on \mathcal{H} can be viewed as the restriction to \mathcal{H} of a PVM on a larger Hilbert space.

Theorem 2.1: [6, Th.9.3.2] Let Ω be a compact metrizable space with Borel field \mathcal{F} , and \hat{A} a POVM taking values in $\mathcal{L}_+(\mathcal{H})$. Then there exists a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a PVM $\hat{A}^{\mathcal{K}} : \mathcal{F} \rightarrow \mathcal{L}_+(\mathcal{K})$ such that if P is the orthogonal projection from \mathcal{K} onto \mathcal{H} then $\hat{A}(E)$ is the restriction of $P \hat{A}^{\mathcal{K}}(E) P$ to \mathcal{H} for all Borel sets E .

In principle, the existing filtering theory for PVMs [11], [12] thus covers the needs of POVM-based filtering, modulo a proper lift of the Hilbert space. However, the latter is not

unique and for reasons explained in Section I, it makes sense to look for a POVM theory that does not build on its reduction from a PVM.

To generalize the second measurement postulate, Davies [6] introduces the notion of an instrument as a complement to POVMs. In the following we denote by $\text{CP}(\mathcal{H})$ the set of all completely positive (CP) maps³ $\mathbf{A} : \rho \mapsto \mathbf{A}(\rho)$ on states $\rho \in \mathcal{S}(\mathcal{H})$.

Definition 2.2: [6, Def 4.1.1] Let Ω be a set, \mathcal{F} a σ -field of subsets of Ω , and \mathcal{H} a Hilbert space. Then a \mathcal{H} -valued instrument on Ω is a map $\hat{\mathbf{A}} : \mathcal{F} \rightarrow \text{CP}(\mathcal{H})$; $E \rightarrow \hat{\mathbf{A}}_E(\cdot)$ such that

- 1) $\hat{\mathbf{A}}_E \geq \hat{\mathbf{A}}_\emptyset = 0$ for all $E \in \mathcal{F}$;
- 2) For any countable, mutually disjoint collection $\{E_n\} \subset \mathcal{F}$ we have $\hat{\mathbf{A}}_{(\cup_n E_n)} = \sum_n \hat{\mathbf{A}}_{E_n}$ where the series convergence on the right is in the weak operator topology;
- 3) $\text{Tr} \left\{ \hat{\mathbf{A}}_\Omega(\rho) \right\} = \text{Tr} \left\{ \rho \right\}$, for all $\rho \in \mathcal{S}(\mathcal{H})$.

If an experiment is set up so that the outcomes take values in some set Ω , then for a quantum system prepared in state $\rho \in \mathcal{S}(\mathcal{H})$ the measurement postulates for an instrument write:

- 1) an outcome in the set $E \subset \Omega$ is obtained with probability $\mathbb{P}(E) = \text{Tr} \left\{ \hat{\mathbf{A}}_E(\rho) \right\}$;
- 2) the state of the system conditioned on a measurement outcome in set E is $\hat{\mathbf{A}}_E(\rho)/\mathbb{P}(E)$.

Theorem 2.2: [6, Th.3.1.3, Th.9.2.3] If $\hat{\mathbf{A}}$ is an instrument then for all $E \in \mathcal{F}$, there exists a (non-unique) countable set $\{A_n(E)\}_{n \in \mathbb{N}} \subset \mathcal{L}(\mathcal{H})$ such that

$$\hat{\mathbf{A}}_E(\rho) = \sum_{n \in \mathbb{N}} A_n(E) \rho A_n(E)^*. \quad (2)$$

Moreover, there exists a unique POVM \hat{A} on Ω associated to $\hat{\mathbf{A}}$ such that for all $E \subset \Omega$ and $\rho \in \mathcal{S}(\mathcal{H})$ we have:

$$\text{Tr} \left\{ \hat{\mathbf{A}}_E(\rho) \right\} = \text{Tr} \left\{ \rho \hat{A}_E \right\}. \quad (3)$$

In particular, this unique POVM is given by

$$\hat{A}(E) = \sum_{n \in \mathbb{N}} A_n(E)^* A_n(E) \text{ for all } E \in \mathcal{F}.$$

Theorem 2.1 implies that one can also always construct an instrument corresponding to a given POVM. However, there is no unique nor canonical way to choose the instrument without further information about the physical system.

A. System identification of POVMs

Consider an experimental setup corresponding to unknown instrument $\hat{\mathbf{A}}$, associated to POVM \hat{A} . In order to experimentally determine \hat{A} , one can initialize the system being measured in some state $\rho = |\phi\rangle\langle\phi|$ and for any $E \in \mathcal{F}$, the probability of measurement outcome in set E is

$$\mathbb{P}[\hat{A} \in E] = \langle\phi| \hat{A}(E) |\phi\rangle.$$

³CP maps are the most general quantum evolutions, see [6, Sec.9.2] and [14, Ch.8] for a discussion. The definition given below for instruments is different from that given in [6], wherein the assumption of complete positivity is replaced by positivity; see the discussion in [6, Sec.9.2].

With sufficiently many experimental outcomes, one can estimate the classical probability distribution $\langle\phi| \hat{A}(E) |\phi\rangle$ over all $E \subset \Omega$; doing this for different $|\phi\rangle$ and using polarization then allows to calculate $\hat{A}(E)$ itself. In order to ascertain the instrument $\hat{\mathbf{A}}$ however, we must have access to the state $\hat{\mathbf{A}}_E(\rho)$ e.g. performing a state tomography experiment. This is often impractical in experimental setups.

Remark 2.1: In the following sections, we use instruments in theoretical developments and to examine general properties of the measurement settings we consider. The goal is however to show that our final filtering equation uses POVM data only.

B. Notation

In the remainder of the paper we will use $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \dots$ to denote instruments and \hat{A}, \hat{B}, \dots to denote the POVMs corresponding to these instruments. Also, if an instrument $\hat{\mathbf{A}}$ corresponds to a von Neumann measurement, then we denote by A the associated self-adjoint operator. We will use the terms PVM and self-adjoint operator interchangeably.

III. COMMUTING INSTRUMENTS

The classical development of filtering equations builds on joint probabilities, which are not obvious in the quantum context. Therefore the central idea in [11] is that in order to define a conditional expectation of two self-adjoint operators, the two operators must commute with each other. The von Neumann measurement postulates imply that ‘joint’ measurement statistics can be defined only in this situation.

We now wish to generalize the filtering framework of [11], [12] towards POVMs and for this we need to understand when it is possible to measure two POVMs simultaneously. In fact this depends on the commutativity of the instruments used to implement the POVMs.

Definition 3.1: Suppose $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are two measure spaces and \mathcal{H} is some Hilbert space. Then two \mathcal{H} -valued instruments, $\hat{\mathbf{A}}_1 : \mathcal{F}_1 \rightarrow \text{CP}(\mathcal{H})$ and $\hat{\mathbf{A}}_2 : \mathcal{F}_2 \rightarrow \text{CP}(\mathcal{H})$ are said to *strongly commute* if for all $E_1 \in \mathcal{F}_1$ and $E_2 \in \mathcal{F}_2$, there exist sequences of operators $\{A_m^1(E_1) : m \in N_1\}$ and $\{A_m^2(E_2) : m \in N_2\}$ in $\mathcal{L}(\mathcal{H})$ such that:

- the instruments write (cf. Theorem 2.2)

$$\hat{\mathbf{A}}_{i,(E_i)}(\rho) = \sum_{m \in N_i} A_m^i(E_i) \rho A_m^i(E_i)^* \quad \forall \rho, \quad i = 1, 2 \quad (4)$$

- for all $m \in N_1$ and $n \in N_2$ the commutator $[\cdot, \cdot]$ gives:

$$[A_m^1(E_1), A_n^2(E_2)] = [A_m^1(E_1), A_n^2(E_2)^*] = 0. \quad (5)$$

Remark 3.1: For the special case of PVMs, $N_i = 1$ for all i and the commutativity condition of Def. 3.1 is clearly equivalent to the commutativity of the associated self-adjoint operators A_i . Note that in general, an instrument does not strongly commute with itself (cf. proof of Theorem 4.3.1 in [6]).

Now we consider the composition of instruments — the filtering application will involve one (actual) instrument on the probe and one (hypothetical, expressing our goal-variable) on the target system.

Theorem 3.1: [6, Th.3.4.2] Suppose $\hat{\mathbf{A}}_i$, $i = 1, 2, \dots, n$ are instruments on some compact metrizable Ω_i with Borel field \mathcal{F}_i . Then there exists a unique ‘‘joint’’ instrument $\hat{\mathbf{A}}$ on $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$ such that for all $E_i \in \mathcal{F}_i$ and ρ we have:

$$\hat{\mathbf{A}}_{E_1 \times E_2 \times \dots \times E_n}(\rho) = \hat{\mathbf{A}}_{n, (E_n)} \circ \dots \circ \hat{\mathbf{A}}_{2, (E_2)} \circ \hat{\mathbf{A}}_{1, (E_1)}(\rho).$$

We now prove the first result of this paper.

Theorem 3.2: Suppose $\hat{\mathbf{A}}_i$, $i = 1, 2, \dots, p$ are instruments on some compact metrizable Ω_i with Borel field \mathcal{F}_i and \hat{A}_i are the corresponding POVMs. If the $\hat{\mathbf{A}}_i$ are pairwise strongly commutative, then the POVM \hat{A} corresponding to the joint instrument $\hat{\mathbf{A}}$ is uniquely determined by the POVMs \hat{A}_i , according to

$$\hat{A}(E_1 \times E_2 \times \dots \times E_p) = \hat{A}_1(E_1)\hat{A}_2(E_2) \dots \hat{A}_p(E_p).$$

Moreover, the POVMs are mutually commutative, that is $[\hat{A}_i(E_i), \hat{A}_j(E_j)] = 0$ for all E_i, E_j , $i \neq j$.

Proof: We first prove the result for $p = 2$. Select some events $E_1 \in \mathcal{F}_1$, $E_2 \in \mathcal{F}_2$ and let $E = E_1 \times E_2$. From Definition 3.1 construct sequences of operators $\{A_n^1 : n \in N_1\}$ and $\{A_n^2 : n \in N_2\}$ satisfying (4),(5), where for notational convenience we have written A_n^i for $A_n^i(E_i)$, with $i = 1, 2$. Then for all $\rho \in \mathcal{S}(\mathcal{H})$, we have

$$\begin{aligned} \text{Tr} \left\{ \rho \hat{A}_E \right\} &= \text{Tr} \left\{ \hat{\mathbf{A}}_E(\rho) \right\} \\ \text{(Th.3.1)} &= \text{Tr} \left\{ \hat{\mathbf{A}}_{2, (E_2)} \circ \hat{\mathbf{A}}_{1, (E_1)}(\rho) \right\} \\ \text{(Th.2.2)} &= \text{Tr} \left\{ \sum_{m \in N_1, n \in N_2} A_n^2 A_m^1 \rho A_m^1 {}^* A_n^2 {}^* \right\} \\ \text{(trace property)} &= \text{Tr} \left\{ \rho \sum_{m \in N_1, n \in N_2} A_m^1 {}^* A_n^2 {}^* A_n^2 A_m^1 \right\} \\ \text{(Def.3.1, (5))} &= \text{Tr} \left\{ \rho \sum_{m \in N_1} A_m^1 {}^* A_m^1 \sum_{n \in N_2} A_n^2 {}^* A_n^2 \right\} \\ &= \text{Tr} \left\{ \rho \hat{A}_1 \hat{A}_2 \right\} = \text{Tr} \left\{ \rho \hat{A}_2 \hat{A}_1 \right\}. \quad (6) \end{aligned}$$

To see how the recursive argument works, consider $p = 3$. The above leads to $\text{Tr} \left\{ \rho \hat{A}_E \right\} = \text{Tr} \left\{ \hat{\mathbf{A}}_{1, (E_1)}(\rho) \hat{A}_2 \hat{A}_3 \right\}$. Choosing representations (2) such that instruments $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$ commute, we get $\text{Tr} \left\{ \rho \hat{A}_E \right\} = \text{Tr} \left\{ \hat{\mathbf{A}}_{1, (E_1)}(\rho \hat{A}_2) \hat{A}_3 \right\}$. Now choosing representations (2) such that $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_3$ commute, we get the result. ■

IV. CONDITIONING WITH RESPECT TO A POVM

We are now in a position to define the conditional POVM associated with two strongly commuting instruments. Recall that we wish to find an expression for the conditional POVM that is expressed only in terms of the POVMs and not in terms of the instruments themselves (cf. Remark 2.1).

A. Basic definitions

In this section, let \mathcal{H} be a Hilbert space, Ω_A and Ω_B be two compact metrizable sets with Borel algebras \mathcal{F}_A and

\mathcal{F}_B , respectively, and \hat{A} and \hat{B} be two \mathcal{H} -valued POVMs on Ω_A and Ω_B corresponding to the instruments $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$.

We denote ρ a state on \mathcal{H} and introduce the semi-norm $\|\cdot\|_\rho = |\text{Tr} \{\cdot \rho\}|$ on the set of bounded operators on \mathcal{H} . Two operators Γ_1, Γ_2 on \mathcal{H} are said to be ρ -equivalent (written $\Gamma_1 \equiv \Gamma_2$) if $\|\Gamma_1 - \Gamma_2\|_\rho = 0$. If $(\Omega, \mathcal{F}, \mu)$ is a probability space then two functions $f, g : \Omega \rightarrow \mathcal{L}^+$ are said to be ρ -equivalent if $\|f(x) - g(x)\|_\rho = 0$ for μ -a.e. $x \in \Omega$.

If f is any measurable function on Ω_A , then the integral of f with respect to \hat{A} over a set $E \in \mathcal{F}_A$ is defined by:⁴

$$\int_{\omega \in E} f(\omega) \text{Tr} \left\{ \rho \hat{A}(d\omega) \right\} \triangleq \int_{\omega \in E} f(\omega) d\mu_{\rho, \hat{A}}(\omega)$$

where the measure $\mu_{\rho, \hat{A}}(\cdot) \triangleq \text{Tr} \left\{ \rho \hat{A}(\cdot) \right\}$ is a probability measure on $(\Omega_A, \mathcal{F}_A)$.

B. Conditional Expectation

Following [12], one approach to conditioning and filtering is through the conditional expectation. When a vector space is associated to the POVM measurement results Ω_A , we can compute the *expectation value of a POVM \hat{A} in state ρ* by:

$$\mathbb{E}_\rho[\hat{A}] = \int_{\omega \in \Omega} \omega \text{Tr} \left\{ \rho \hat{A}(d\omega) \right\}.$$

If $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are strongly commuting instruments, we can set $\Omega = \Omega_A \times \Omega_B$ with the product σ -field and define $\hat{A}\hat{B}$ the product POVM as in Theorem 3.2. On the product space Ω we can then define two classical random variables $\alpha : (\omega_A, \omega_B) \mapsto \omega_A$ and $\beta : (\omega_A, \omega_B) \mapsto \omega_B$, and we know from the classical Kolmogorov theory that the conditional expectation $\mathbb{E}[\alpha|\beta] : \Omega \rightarrow \Omega_A$ exists and is a (a.e.) unique random variable that is measurable with respect to $\mathcal{F}_\beta \equiv \mathcal{F}_B$. Defining \mathcal{G}_A the σ -algebra of subsets of Ω_A generated by $\mathbb{E}[\alpha|\beta]$,⁵ we can lift this to a well-defined *conditional expectation POVM*:

$$\hat{\mathbb{E}}_{\hat{A}|\hat{B}}(E) = \hat{B} \left(\mathbb{E}[\alpha|\beta]^{-1}(E) \right) \quad \text{for any } E \in \mathcal{G}_A. \quad (7)$$

This basically attributes to event E the \hat{B} -POVM element associated to the union of β for which $\mathbb{E}[\alpha|\beta] \in E$.

However, unlike in the PVM case [12], $\hat{\mathbb{E}}_{\hat{A}|\hat{B}}$ is not trivial to use for filtering purposes. Indeed, because in general $\hat{B}(E_1)\hat{B}(E_2) \neq 0$ even for disjoint events E_1, E_2 , it is not clear how to input an actual measurement result for \hat{B} into this expression.

C. Conditional POVMs

We now consider an approach that is motivated from classical conditional probability.

Theorem 4.1: Let $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ as defined in Section IV-A be strongly commuting instruments. There exists a (not necessarily unique) map $\hat{P} : \mathcal{F}_A \times \Omega_B \rightarrow \mathcal{L}^+$ such that:

⁴It should also be possible to define this integral as a Stieltjes integral over the measurable space $\mathcal{L}_+(\mathcal{H})$ with the measure induced by the POVM.

⁵Thus, ‘‘ \mathcal{G}_A is as coarse-grained as \mathcal{F}_B or coarser’’, in the sense that it is the σ -algebra generated by an \mathcal{F}_B -measurable function.

1) For all $E \in \mathcal{F}_A$ and $F \in \mathcal{F}_B$, we have

$$\hat{A}(E)\hat{B}(F) \equiv \int_F \hat{P}(E, \omega) \text{Tr} \left\{ \rho \hat{B}(d\omega) \right\}.$$

2) For $\mu_{\rho, \hat{B}}$ -almost-every $\omega \in \Omega_B$ and every mutually disjoint sequence $E_1, E_2, \dots \in \mathcal{F}_A$ we have the countable additivity condition

$$\hat{P} \left(\bigcup_{n=1}^{\infty} E_n, \omega \right) \equiv \sum_{n=1}^{\infty} \hat{P}(E_n, \omega).$$

3) $\hat{P}(\Omega_A, \omega) \equiv \mathbb{I}_{\mathcal{H}}$ for $\mu_{\rho, \hat{B}}$ -a.e. $\omega \in \Omega_B$.

If \hat{P}' is another such map then \hat{P} and \hat{P}' are ρ -equivalent.

Definition 4.1: We call $\hat{P}_{A|B} \triangleq \hat{P}$ the *conditional POVM of \hat{A} given \hat{B}* . Then $\text{Tr} \left\{ \rho \hat{P}_{A|B}(E, \omega) \right\} = \mathbb{P}(\hat{A} \in E | \hat{B} \in \{\omega\})$ gives the probability that a measurement of \hat{A} is in E when knowing that ω is obtained from a \hat{B} measurement.

Proof: The proof follows [18] and uses the notion of a lifting. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and denote $E \sim_{\mu} F$ if $E \in \mathcal{F}$ and $F \in \mathcal{F}$ differ by a μ -measure zero set. A lifting of μ is a map $\phi: \mathcal{F} \rightarrow \mathcal{F}$ such that

- 1) $\phi(E) \sim_{\mu} E$.
- 2) $E \sim_{\mu} F$ implies $\phi(E) = \phi(F)$.
- 3) $\phi(E \cup F) = \phi(E) \cup \phi(F)$ and $\phi(E \cap F) = \phi(E) \cap \phi(F)$.
- 4) $\phi(\emptyset) = \emptyset$ and $\phi(\Omega) = \Omega$.

Denote $\hat{C}(E, F) = \hat{A}(E)\hat{B}(F)$ the joint POVM on $\Omega_A \times \Omega_B$, see Theorem 3.2, and let $\mu_{\rho, \hat{A}}$, $\mu_{\rho, \hat{B}}$ and $\mu_{\rho, \hat{C}}$ be the associated probability measures on Ω_A , Ω_B and $\Omega_A \times \Omega_B$.

Because Ω_A is a complete metric space we can apply Theorem 6.6.6 in [17] and this implies there exists a regular conditional probability $\mathbb{P}: \mathcal{F}_A \times \Omega_B \rightarrow [0, 1]$ which is \mathcal{F}_B -measurable for any $E \in \mathcal{F}_A$ and satisfies

- 1) $\mu_{\rho, \hat{C}}(E, F) = \int_F \mathbb{P}(E, \omega) d\mu_{\rho, \hat{B}}(\omega)$ for all $E \in \mathcal{F}_A$ and $F \in \mathcal{F}_B$.
- 2) For every mutually disjoint sequence $E_1, E_2, \dots \in \mathcal{F}_A$ and $\mu_{\rho, \hat{B}}$ -a.e. ω we have

$$\mathbb{P} \left(\bigcup_{n=1}^{\infty} E_n, \omega \right) \equiv \sum_{n=1}^{\infty} \mathbb{P}(E_n, \omega).$$

3) For $\mu_{\rho, \hat{B}}$ -a.e. ω we have $\mathbb{P}(\Omega_A, \omega) = 1$.

Following [18], fix $E \in \mathcal{F}_A$, denote S the support of $\mathbb{P}(E, \cdot)$ in Ω_B and let ϕ be a lifting for μ restricted to the measurable subsets of S . Consider Π be the collection of partitions of S such that $\pi = \{S_1, \dots, S_n\} \in \Pi$ implies $\phi(S_i) = S_i \neq \emptyset$ for all i . Such Π always exists and is a directed set under refinement [18]. Define

$$\begin{aligned} \mathbb{P}_{\pi}(E, \cdot) &= \sum_i \frac{\mu_{\rho, \hat{C}}(E, S_i)}{\mu_{\rho, \hat{B}}(S_i)} \chi_{S_i}(\cdot) \quad \text{and} \\ \hat{P}_{\pi}(E, \cdot) &= \sum_i \frac{\hat{C}(E, S_i)}{\mu_{\rho, \hat{B}}(S_i)} \chi_{S_i}(\cdot) \end{aligned} \quad (8)$$

where χ_{S_i} is the characteristic function of the set S_i . From the definition of \hat{C} , we have $\text{Tr} \left\{ \rho \hat{P}_{\pi} \right\} = \mathbb{P}_{\pi}$ and by [18, Lemma 4.3] the limit exists and gives

$$\lim_{\pi} \text{Tr} \left\{ \rho \hat{P}_{\pi}(E, \cdot) \right\} = \lim_{\pi} \mathbb{P}_{\pi}(E, \cdot) = \mathbb{P}(E, \cdot). \quad (9)$$

The set \mathcal{L}_1^+ of positive operators on \mathcal{H} with norm less than 1 is a closed and bounded subset of $\mathcal{L}(\mathcal{H})$. Therefore, it is compact in the weak-* topology⁶ by the Banach-Alaoglu theorem (see e.g. [19, Sec 3.15]), so there exists a (not necessarily unique) positive operator $\hat{P}(E, \cdot)$ such that

$$\lim_{\pi}^* \hat{P}_{\pi}(E, \cdot) = \hat{P}(E, \cdot).$$

By definition of weak-* convergence we have $\lim_{\pi} \text{Tr} \left\{ \rho \hat{P}_{\pi}(E, \cdot) \right\} = \text{Tr} \left\{ \rho \hat{P}(E, \cdot) \right\}$ and by (9) we get $\text{Tr} \left\{ \rho \hat{P}(E, \cdot) \right\} = \mathbb{P}(E, \cdot)$. The proof now follows from the properties of \mathbb{P} . ■

Definition 4.2: The conditional expectation of \hat{A} knowing that $\hat{B} \in \{\omega_B\}$ is given by

$$\mathbb{E}_{\rho}[\hat{A} | \hat{B} \in \{\omega_B\}] = \int_{\omega_A \in \Omega_A} \omega_A \text{Tr} \left\{ \rho \hat{P}_{A|B}(d\omega_A, \omega_B) \right\}.$$

V. FILTERING WITH POVMS

Our measurement model is motivated from the discrete-time model used in [12] for filtering using PVMs. We consider a system with Hilbert space \mathcal{H}_S and a probe consisting of a sequence of subsystems $n = 1, 2, \dots$, each with Hilbert space $\mathcal{H}_n = \mathcal{H}$. So the probe is described on a Hilbert space $\mathcal{H}_P = \otimes_{n=1}^{\infty} \mathcal{H}_n$; the combined state space of the probe and system is written $\mathcal{H}_{tot} = \mathcal{H}_S \otimes \mathcal{H}_P$. In the practical setup, we will assume that system and probe are initially separated, and at consecutive times $n = 1, 2, \dots$ the system undergoes an interaction with probe subsystem \mathcal{H}_n , which is then measured by a POVM. We want to know the evolution of ρ_n^S , the system state after n interactions and probe measurements, conditioned on the latter's outcomes.

We therefore suppose that Ω_n is a compact metrizable space⁷ for $n = 1, 2, \dots$, let \mathcal{F}_n a σ -field on Ω_n and $\mathbf{B}_n: \mathcal{F}_n \rightarrow \text{CP}(\mathcal{H}_n)$ an instrument with corresponding POVM B_n . Also, let $(\Omega_A, \mathcal{F}_A)$ be some measure space and $\hat{A}: \mathcal{F}_A \rightarrow \text{CP}(\mathcal{H}_S)$ any system instrument with corresponding POVM A . Note that we do not associate the Ω_n to a vector space; this allows to consider probe measurements with results on manifolds, like the circle for phase measurements.

To consider the ‘active’ part while the sequence of interactions progresses, we set $\mathcal{H}_n = \mathcal{H}_S$ for $n = 0$ and recursively define $\mathcal{H}_n = \mathcal{H}_{n-1} \otimes \mathcal{H}_n$ for $n \geq 1$; similarly define $\Omega_n = \Omega_A \times \Omega_1 \times \dots \times \Omega_n$. Consider the initial state on \mathcal{H}_{tot} ,

$$\rho_0^{tot} = \rho_0^S \otimes_{n=1}^{\infty} \rho_n^P.$$

We suppose that between time steps n and $n+1$, the system interacts with the probe according to a unitary evolution operator U_n on $\mathcal{H} \otimes \mathcal{H}_n$, i.e. it interacts only with subsystem n of the probe. After this unitary evolution, the POVM B_n is measured to be some $\omega_n \in \Omega_n$. With a slight abuse of notation, let $U_n = \prod_{i=1}^n U_n$, a unitary on \mathcal{H}_n .

⁶A sequence $A_n \in \mathcal{L}^+$ converges to $A \in \mathcal{L}$ in the weak-* topology if and only if $\text{Tr} \left\{ \rho A_n \right\}$ converges to $\text{Tr} \left\{ \rho A \right\}$ for all trace class operators ρ . We write this convergence $\lim_n^* A_n = A$.

⁷If Ω_n is not compact then we can simply consider the 1-point compactification of Ω_n [6, p.12].

It seems physically reasonable to assume that the instrument associated in this setup to the \mathcal{H}_{tot} -POVM \hat{B}_n acts non-trivially only on \mathcal{H}_n , i.e. its operator-sum representation has elements $\{B_i^n\}_i$ of the form

$$B_i^n = \mathbb{I}_{\mathcal{H}_S} \left(\bigotimes_{m=1}^{n-1} \mathbb{I}_{\mathcal{H}} \right) \otimes b_i^n \left(\bigotimes_{m=n+1}^{\infty} \mathbb{I}_{\mathcal{H}} \right).$$

One formal argument for this is that the measurement generally acts at a place away from the system, and often also away from the other (e.g. travelling or already destroyed) probe subsystems. Similarly, when speaking of system properties through the \mathcal{H}_{tot} -POVM \hat{A} , it makes sense to assume that the associated instrument \hat{A} acts nontrivially only on \mathcal{H}_S . Therefore the set of instruments \hat{A} , \hat{B}_n for $n = 1, 2, \dots$ would strongly commute. Then the same commutations apply to the evolved instruments where the effect of n interactions is described in the Heisenberg picture,

$$\hat{A}(n) \triangleq U_n \hat{A} U_n^* \quad \text{and} \quad \hat{B}_i(n) \triangleq U_n \hat{B}_i U_n^*.$$

Therefore, the conditional probability of $\hat{A}(n)$ with respect to $\hat{B}_1(n) \hat{B}_2(n) \dots \hat{B}_n(n)$ is well-defined according to Theorems 3.2 and 4.1: there exists a function

$$\hat{P}_n^A(E, \omega_n) \triangleq \hat{P}_{\hat{A}(n) | \hat{B}_1(n) \hat{B}_2(n) \dots \hat{B}_n(n)}(E, \omega_n)$$

such that $\text{Tr} \left\{ \hat{P}_n^A(E, \omega_n) \rho_0^{\text{tot}} \right\}$ gives the probability of \hat{A} -events, knowing the outcomes $\omega_n \triangleq \omega_1, \dots, \omega_n$ of the n first probe measurements after interactions with the system. From (8), the map $\Phi_{\rho_0^{\text{tot}}} : \hat{A} \rightarrow \mathbb{E}_{\rho_0^{\text{tot}}}[\hat{A} | \omega_n]$ is linear and real-valued.⁸ Thus by Gleason's theorem, at least for finite-dimensional \mathcal{H}_S , there exists a unique density operator, which we identify as the *post-measurement state* $\rho_n^S(\omega_n)$, such that $\Phi_{\rho_0^{\text{tot}}}(\hat{A}) = \text{Tr} \left\{ \hat{A} \rho_n^S(\omega_n) \right\}$ for all \hat{A} . Algebraic computations based on property 1) of Theorem 4.1 then lead to the well-known expression:

$$\rho_n^S(\omega_n) = \frac{\text{Tr}_P \left\{ U_n \rho_0^{\text{tot}} U_n^* \hat{B}(\omega_n) \right\}}{\text{Tr} \left\{ U_n \rho_0^{\text{tot}} U_n^* \hat{B}(\omega_n) \right\}},$$

where $\text{Tr}_P \{ \cdot \}$ is the partial trace over \mathcal{H}_P . Note that, thanks to the strong commutativity condition, this expression depends on the POVM but not on the associated instrument. Moreover, in this expression only the POVM element associated to the actually observed ω_n is needed, irrespective of the other potentialities completing the POVM. Interesting implications of this are best illustrated by the following example.

Qubit phase: Consider a probe composed of qubits, $\mathcal{H} = \text{span}_{\mathbb{C}} \{ |0\rangle, |1\rangle \} \cong \mathbb{C}^2$, on which we apply the POVM with M elements $d = 0, \dots, M-1$,

$$\hat{B}_{\mathcal{H}}(d) = \frac{1}{M} (|0\rangle + e^{2i\pi d/M} |1\rangle) (\langle 0| + e^{-2i\pi d/M} \langle 1|).$$

Then the effect on the system of a detection result e.g. $\omega_1 = 0$ is the same, whether $M = 2$ (projective measurement), or $M = 3, 4, \dots$ or any larger number. It seems legitimate to attribute the same effect to the continuous POVM limit.

⁸Note that Φ is unique since all possible \hat{P} are ρ_0^{tot} -equivalent.

VI. CONCLUSION AND FUTURE WORK

In this paper we show how quantum filtering can be performed in the POVM setting. We formalize a notion of strongly commuting instruments, which gives a sufficient condition to define joint measurement statistics in terms of the associated POVMs only, without explicitly depending on the instruments. We introduce the notion of conditional-expectation-POVM for measurements with commuting instruments, and highlight that it can be inappropriate for filtering, unlike in the PVM case. We then show that the notion of conditional probabilities can be defined for strongly commuting instruments. On that basis, for a system-probe model, we analyze filtering of the system state conditioned on POVM probe measurements, and highlight its general properties. A future goal is to apply these ideas to derive a filtering equation for discrete-time *phase* measurements, a quintessential example of a POVM-type measurement where the associated instrument is not known. On a more theoretical mode, we also should explore *necessary* conditions for the strong commutation of two instruments.

REFERENCES

- [1] A. Bensoussan, *Stochastic Control of Partially Observable Systems*. Cambridge University Press, 1992.
- [2] V. P. Belavkin, "Quantum filtering of Markov signals with white quantum noise," *Radiotekhnika i Elektronika*, vol. 25, pp. 1445–1453, 1980.
- [3] V. P. Belavkin, "Theory of the control of observable quantum-systems," *Automation and Remote Control*, vol. 44, p. 178, 1983.
- [4] V. P. Belavkin, *Nondemolition stochastic calculus in Fock space and nonlinear filtering and control in quantum systems*, ser. Stochastic methods in mathematics and physics. XXIV Karpacz winter school: World Scientific, Singapore, 1988, pp. 310–324.
- [5] V. P. Belavkin, "Quantum stochastic calculus and quantum nonlinear filtering," *Journal of Multivariate Analysis*, vol. 42, no. 2, pp. 171–201, 1992.
- [6] E. B. Davies, *Quantum theory of open systems*. IMA, 1976.
- [7] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement*. Cambridge University Press, 1992.
- [8] S. Haroche and J. Raimond, *Exploring the Quantum: Atoms, Cavities, and Photons (Oxford Graduate Texts)*. Oxford University Press, 2006.
- [9] C. W. Gardiner and P. Zoller, *Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics*. Springer, 2004, vol. 56.
- [10] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control*. Cambridge University Press, 2010.
- [11] L. Bouten, R. Van Handel, and M. James, "An introduction to quantum filtering," *SIAM Journal on Control and Optimization*, vol. 46, no. 6, pp. 2199–2241, 2007.
- [12] L. Bouten, R. Van Handel, and M. R. James, "A discrete invitation to quantum filtering and feedback control," *SIAM review*, vol. 51, no. 2, pp. 239–316, 2009.
- [13] J. H. Shapiro and S. R. Shepard, "Quantum phase measurement: A system-theory perspective," *Phys. Rev. A*, vol. 43, pp. 3795–3818, 1991.
- [14] M. Nielsen and I. Chuang, *Quantum computation and quantum information*. Cambridge university press, 2010.
- [15] D. Berry, "Adaptive Phase Measurements," Ph.D. dissertation, University of Queensland, Feb. 2002.
- [16] R. Lynch, "The quantum phase problem: a critical review," *Physics Reports*, vol. 256, no. 6, pp. 367–436, 1995.
- [17] R. B. Ash, *Real analysis and probability*. Academic Press, 1972.
- [18] J. Kupka, "Radon-Nikodym theorems for vector valued measures," *Transactions of the American Mathematical Society*, vol. 169, pp. 197–217, 1972.
- [19] W. Rudin, *Functional analysis*. 1973. McGraw-Hill, New York, 1973.