Design and stability of discrete-time quantum filters with measurement imperfections

Abhinav Somaraju, Igor Dotsenko, Clément Sayrin, Pierre Rouchon

Design and Stability of Discrete-Time Quantum Filters with Measurement Imperfections

Abhinav Somaraju Igor Dotsenko Clement Sayrin Pierre Rouchon

***Abstract*—This work considers the theory underlying a discrete-time quantum filter recently used in a quantum feedback experiment. It proves that this filter taking into account decoherence and measurement errors is optimal and stable. We present the general framework underlying this filter and show that it corresponds to a recursive expression of the least-square optimal estimation of the density operator in the presence of measurement imperfections. By measurement imperfections, we mean in a very general sense unread measurement performed by the environment (decoherence) and active measurement performed by non-ideal detectors. However, we assume to know precisely all the Kraus operators and also the detection error rates. Such recursive expressions combine well known methods from quantum filtering theory and classical probability theory (Bayes' law). We then demonstrate that such a recursive filter is always stable with respect to its initial condition: the fidelity between the optimal filter state (when the initial filter state coincides with the real quantum state) and the filter state (when the initial filter state is arbitrary) is a sub-martingale.**

I. Introduction

The theory of filtering considers the estimation of the system state from noisy and/or partial observations (see, e.g., [5]). For quantum systems, filtering theory was initiated in the 1980s by Belavkin in a series of papers [1], [2], [3], [4] (also see the tutorial papers [7], [6] for a more recent introduction). Belavkin makes use of the operational formalisms of Davies [9], 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. [4], [2]). The theory of quantum filtering was independently developed in the physics community, particularly in the context of quantum optics, under the name of Quantum Measurement Theory [8], [12], [11], [16].

Most of this theory has been developed for continuous-time systems and little emphasis has been given to measurement imprecisions and their explicit impact on the filter design and time-recursive equations. To our knowledge, the problem of designing a quantum filter in the presence of classical measurement imperfections has not been examined in the discrete time setting. In this paper, we focus on this issue and propose a systematic method to derive quantum filters taking into account several detection error rates.

In [11, Sec. 2.2.2], the authors discuss how the state of a quantum system evolves after a single imprecise measurement. In [10], a recursive quantum state estimation with measurement imperfections has been considered. In [14], such a quantum state estimate has been used in a quantum feedback experiment that stabilizes photon-number states of a quantized field mode, trapped in a super-conducting cavity. We prove here that such estimates are in fact optimal since they coincide with the conditional expectation of the quantum state (density matrix) knowing the past detections, the error rates and the initial quantum state.

Section II describes the structure of a genuine quantum measurement model including detection error rates which is a straightforward generalization of the models considered in [11], [10], [14]. This model may be used in situations with partial knowledge of all the quantum jumps and also measurement errors of the jumps that are detected. However, they assume to know precisely all the Kraus operators and also the detection error rates.

Section III is devoted to the first result in this paper summarized in Theorem III.1: the conditional expectation of the quantum state knowing the past detections and the initial state obeys a recursive equation in each discrete time-step. This recursive equation is given in (3) and depends explicitly on the error rates. The proof of Theorem III.1 shows that such recursive equation may be derived by a simple application of Bayes' law.

In section IV, we prove that the quantum filter defined in Theorem III.1 is stable versus its initial conditions: the fidelity between the optimal estimate conditioned on the initial state of the system being known and a second estimate in which the initial state is unknown is a sub-martingale. This stability result combines Theorem III.1 and [13]. Note that stability does not imply convergence, in general. For convergence results in the continuous-time case see, e.g., [15] and the references therein.

In section V, we describe in detail the Kraus operators and error rates modeling the discrete-time quantum system considered in the quantum feedback experiment [14]: the quantum filter used in the feedback loop corresponds precisely to the recursive equation (3) given by Theorem III.1; according to Theorem IV.1, this filter tends to forget its initial condition.

This work has been partially supported by the ANR under the projects CQUID, QUSCO-INCA and EPOQ2, and by EU under the IP project AQUTE and the ERC project DECLIC

Abhinav Somaraju, Brussels Photonics Team, Dept. of Applied Physics and Photonics, Vrije Universiteit Brussel, Pleinlaan 2-1050 Brussels, Belgium. a.somaraju@gmail.com

Igor Dotsenko, Laboratoire Kastler Brossel, École Normale Supérieure, CNRS, Université P. et M. Curie, 24 rue Lhomond, 75231 Paris Cedex 05, France igor.dotsenko@lkb.ens.fr

Clement Sayrin, Laboratoire Kastler Brossel, École Normale Supérieure, CNRS, Université P. et M. Curie, 24 rue Lhomond, 75231 Paris Cedex 05, France clement.sayrin@lkb.ens.fr

Pierre Rouchon, Mines ParisTech, Centre Automatique et Systèmes, Unité Mathématiques et Systèmes, 60 Bd Saint-Michel, 75272 Paris cedex 06, France pierre.rouchon@mines-paristech.fr

A. Acknowledgments

The authors thank Michel Brune, Serge Haroche, Mazyar Mirrahimi and Jean-Michel Raimond for useful discussions and advices.

II. Measurement Model

In this section we discuss the model describing repeated and imperfect measurements on a quantum system. Such modeling including decoherence-induced quantum jumps and measurement errors is a direct generalization of the one proposed in [10] and used in real-time for the quantum feedback experiments reported in [14] (also see [11]). We initially consider the case of a single ideal measurement and then develop the model to consider imperfect and repeated measurements. The final model is described in Subsection II-C.

A. Ideal Case

Let $\mathcal{H}$ be the system's Hilbert space with ρ_1 a density matrix denoting the initial state of the system at step $k = 1$. We consider the evolution $\rho_1 \mapsto \rho_2$ of such a quantum system under discrete-time quantum jumps (see e.g. [12, Ch. 4] or [11, Ch. 2]).

Consider a set of Kraus operators $M_q : \mathcal{H} \to \mathcal{H}$, $q \in \{1, 2 \ldots, m^{\mathtt{id}}\}$ that satisfy $\sum_{q=1}^{m^{\mathtt{id}}} M_q^\dagger M_q = \mathbf{I}$. Here we assume that there are $m^{\mathtt{id}} \in \mathbb{N}$ possible quantum jumps and $\mathbf{I}$ is the identity operator on $\mathcal{H}$. The superscript id stands for an abbreviation of ideal. Consider an ideal world with full access to all quantum jumps via a complete and ideal set of jump detectors. When the quantum jump indexed by q is detected, the state of the system changes to

$$\rho_2 = \mathcal{M}_q(\rho_1) \triangleq \frac{M_q \rho_1 M_q^\dagger}{\mathbb{P}[q]}. \tag{1}$$

Moreover,

$$\mathbb{P}[q] = \mathrm{Tr}\left\{M_q \rho_1 M_q^\dagger\right\} \tag{2}$$

is the probability to detect jump q, knowing the state ρ_1. We now consider the case of realistic experiments with possible measurement errors.

B. Realistic Case (with imprecise measurements)

We consider that the ideal detection of the jump q corresponds to an ideal measure outcome $\mu^{\mathtt{id}} = q$. We denote by random variable $\mu^{\mathtt{id}} \in \{1, 2, \ldots, m^{\mathtt{id}}\}$ this outcome provided by ideal sensors. We assume that realistic sensors provide an outcome $\mu^{\mathtt{rl}}$ that is a random variable in the set $\{1, 2, \ldots, m^{\mathtt{rl}}\}$.

We assume that, with a known probability, an ideal measurement outcome $\mu^{\mathtt{id}}$ occurs effectively whereas the realistic sensors detect an outcome $\mu^{\mathtt{rl}}$. The correlations between the events $\mu^{\mathtt{id}} = q$ and $\mu^{\mathtt{rl}} = p$ are modeled by classical probabilities through a stochastic matrix $\eta \in \mathbb{R}^{m^{\mathtt{rl}} \times m^{\mathtt{id}}}$:

$$\eta_{p,q} = \mathbb{P}[\mu^{\mathtt{rl}} = p | \mu^{\mathtt{id}} = q].$$

It gives the probability that the real sensors detect $\mu^{\mathtt{rl}} = p$ given the ideal sensors would detect $\mu^{\mathtt{id}} = q$, for $p \in \{1, \ldots, m^{\mathtt{rl}}\}$, $q \in \{1, \ldots, m^{\mathtt{id}}\}$. Since $\eta_{p,q} \geq 0$ and for each q, $\sum_{p=1}^{m^{\mathtt{rl}}} \eta_{p,q} = 1$, the matrix $\eta = (\eta_{p,q})$ is a left stochastic matrix.

C. Realistic Experiment with Repeated Measurements

Consider the case of a sequence of discrete-time measurements. We denote by ρ_k the state of the system at discrete time-step k. Also suppose $M_{q;k}$ is the Kraus operator corresponding to the k^{th} ideal measurement for $q \in \{1, 2, \ldots, m^{\mathtt{id}}\}$. Note that we allow for a different set of Kraus operators $M_{.;k}$ for different time-steps k. One can also consider $m^{\mathtt{id}}$ and $m^{\mathtt{rl}}$ be dependent on k.

Similar to the previous subsection, we denote by $\mu_k^{\mathtt{rl}} \in \{1, \ldots, m^{\mathtt{rl}}\}$ and $\mu_k^{\mathtt{id}} \in \{1, \ldots, m^{\mathtt{id}}\}$, the random variables corresponding to the k^{th} realistic and ideal outcomes, respectively. Therefore,

$$\mathbb{E}[\rho_{k+1} | \rho_k, \mu_k^{\mathtt{id}} = q] = \mathcal{M}_{q;k}(\rho_k) \triangleq \frac{M_{q;k}\rho_k M_{q;k}^\dagger}{\mathrm{Tr}\left\{M_{q;k}\rho_k M_{q;k}^\dagger\right\}}$$

and $\mathbb{P}\left[\mu_k^{\mathtt{id}} = q \mid \rho_k\right] = \mathrm{Tr}\left\{M_{q;k}\rho_k M_{q;k}^\dagger\right\}$.

Also, we assume $\eta^k \in \mathbb{R}^{m^{\mathtt{rl}} \times m^{\mathtt{id}}}$, the stochastic matrix determining the probability of error, can depend on the discrete-time step k. In particular, we have

$$\mathbb{P}[\mu_k^{\mathtt{rl}} = p | \mu_k^{\mathtt{id}} = q] = \eta_{p,q}^k,$$

for $p \in \{1, \ldots, m^{\mathtt{rl}}\}$, $q \in \{1, \ldots, m^{\mathtt{id}}\}$.

III. Recursive Equation for the optimal filter

We wish to obtain a recursive equation for the optimal estimate $\hat{\rho}_{k+1}$ of the state ρ_{k+1} knowing initial value ρ_1 and real measurement outcomes $\mu_1^{\mathtt{rl}}, \ldots, \mu_k^{\mathtt{rl}}$. This optimal estimate $\hat{\rho}_k$ is defined as

$$\hat{\rho}_k = \mathbb{E}[\rho_k | \rho_1, \mu_1^{\mathtt{rl}}, \ldots, \mu_{k-1}^{\mathtt{rl}}].$$

The following theorem says that we can ignore the original state ρ_k and only consider $\hat{\rho}_k$ that is shown to be the state of a Markov process.

Theorem III.1. *The optimal estimate $\hat{\rho}_k$ satisfies the following recursive equation*

$$\hat{\rho}_{k+1} = \frac{\sum_{q=1}^{m^{\mathtt{id}}} \eta_{p_k,q}^k M_{q;k}\hat{\rho}_k M_{q;k}^\dagger}{\mathrm{Tr}\left\{\sum_{q=1}^{m^{\mathtt{id}}} \eta_{p_k,q}^k M_{q;k}\hat{\rho}_k M_{q;k}^\dagger\right\}}, \tag{3}$$

if $\mu_k^{\mathtt{rl}} = p_k$. Moreover, we have

$$\mathbb{P}\left[\mu_k^{\mathtt{rl}} = p_k \middle| \rho_1, \mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1}\right] = \mathrm{Tr}\left\{\sum_{q=1}^{m^{\mathtt{id}}} \eta_{p_k,q}^k M_{q;k}\hat{\rho}_k M_{q;k}^\dagger\right\}. \tag{4}$$

Remark III.1. *The division in (3) by the R.H.S of (4) could appear problematic when this denominator vanishes.*

Nevertheless, if we assume that the real measurements are $\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_k^{\mathtt{rl}} = p_k$, *then*

$$\mathbb{P}\Big[\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1} \Big| \rho_1\Big] > 0$$

and

$$\mathbb{P}\Big[\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_k^{\mathtt{rl}} = p_k \Big| \rho_1\Big] > 0$$

(otherwise such measurement outcomes are not possible). Consequently,

$$\mathbb{P}\Big[\mu_k^{\mathtt{rl}} = p_k \Big| \rho_1, \mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1}\Big] = \frac{\mathbb{P}\Big[\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1}, \mu_k^{\mathtt{rl}} = p_k \Big| \rho_1\Big]}{\mathbb{P}\Big[\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1} \Big| \rho_1\Big]}$$

cannot vanish. Thus recurrence (3) *is always well defined because we have* (4).

Remark III.2 (Markov property of the filter)**.** *Equations* (3) *and* (4) *tell us that the joint-process* $(\mu_k^{\mathtt{rl}}, \hat{\rho}_k)$ *is a Markov process and therefore the statistics of the measurement process* $\mu_k^{\mathtt{rl}}$ *may be determined using* $\hat{\rho}_k$. *This in particular implies that we may use Monte Carlo methods to simulate the observation process* $\mu_k^{\mathtt{rl}}$ *only using* $\hat{\rho}_k$ *independent of the actual state* ρ_k *and measurement history* $\mu_1^{\mathtt{rl}}, \ldots, \mu_{k-1}^{\mathtt{rl}}$.

Proof: In this proof we use the following notation for ease of presentation: we use $\mu_\iota^{\mathtt{id}} = q_\iota$ and $\mu_\iota^{\mathtt{rl}} = p_\iota$ to denote the set of events $\{\mu_1^{\mathtt{id}} = q_1, \mu_2^{\mathtt{id}} = q_2, \ldots, \mu_k^{\mathtt{id}} = q_k\}$ and $\{\mu_1^{\mathtt{rl}} = p_1, \mu_2^{\mathtt{rl}} = p_2, \ldots, \mu_k^{\mathtt{rl}} = p_k\}$, respectively. For instance, using this notation, we have

$$\mathbb{P}\Big[\mu_\iota^{\mathtt{id}} = q_\iota \Big| \rho_1, \mu_\iota^{\mathtt{rl}} = p_\iota\Big] \triangleq \mathbb{P}\Big[\mu_1^{\mathtt{id}} = q_1, \ldots, \mu_k^{\mathtt{id}} = q_k \Big| \rho_1, \mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_k^{\mathtt{rl}} = p_k\Big].$$

Assume that the values measured by the real detector are $p_1 = \mu_1^{\mathtt{rl}}, \ldots, p_k = \mu_k^{\mathtt{rl}}$. Then we have the optimal estimate

$$\hat{\rho}_{k+1} = \sum_{q_1,\ldots,q_k=1}^{m^{\mathtt{id}}} \mathbb{P}\Big[\mu_\iota^{\mathtt{id}} = q_\iota \Big| \rho_1, \mu_\iota^{\mathtt{rl}} = p_\iota\Big] \mathcal{M}_{q_k;k}(\ldots(\mathcal{M}_{q_1;1}(\rho_1))\ldots), \tag{5}$$

where

$$\mathcal{M}_{q_k;k}(\cdots(\mathcal{M}_{q_1;1}(\rho_1))\cdots) = \frac{M_{q_k;k}\cdots M_{q_1;1}\rho_1 M_{q_1;1}^\dagger \cdots M_{q_k;k}^\dagger}{\mathrm{Tr}\left\{M_{q_k;k}\cdots M_{q_1;1}\rho_1 M_{q_1;1}^\dagger \cdots M_{q_k;k}^\dagger\right\}}. \tag{6}$$

Using Bayes law, we have for each $(q_1, \ldots, q_k)$,

$$\mathbb{P}\Big[\mu_\iota^{\mathtt{id}} = q_\iota \Big| \rho_1, \mu_\iota^{\mathtt{rl}} = p_\iota\Big] \mathbb{P}\Big[\mu_\iota^{\mathtt{rl}} = p_\iota \Big| \rho_1\Big] = \mathbb{P}\Big[\mu_\iota^{\mathtt{rl}} = p_\iota \Big| \rho_1, \mu_\iota^{\mathtt{id}} = q_\iota\Big] \mathbb{P}\Big[\mu_\iota^{\mathtt{id}} = q_\iota \Big| \rho_1\Big], \tag{7}$$

where

$$\mathbb{P}\Big[\mu_\iota^{\mathtt{id}} = q_\iota \Big| \rho_1\Big] = \mathrm{Tr}\left\{M_{q_k;k}\ldots M_{q_1;1}\rho_1 M_{q_1;1}^\dagger \ldots M_{q_k;k}^\dagger\right\}$$

and

$$\mathbb{P}\Big[\mu_\iota^{\mathtt{rl}} = p_\iota \Big| \rho_1, \mu_\iota^{\mathtt{id}} = q_\iota\Big] = \mathbb{P}\Big[\mu_\iota^{\mathtt{rl}} = p_\iota \Big| \mu_\iota^{\mathtt{id}} = q_\iota\Big] = \eta_{p_1,q_1}^1 \cdots \eta_{p_k,q_k}^k.$$

Summing (7) over all $(q_1, \ldots, q_k)$ gives:

$$\mathbb{P}\Big[\mu_\iota^{\mathtt{rl}} = p_\iota \Big| \rho_1\Big] = \sum_{s_1,\ldots,s_k=1}^{m^{\mathtt{id}}} \eta_{p_1,s_1}^1 \ldots \eta_{p_k,s_k}^k \times \mathrm{Tr}\left\{M_{s_k;k}\ldots M_{s_1;1}\rho_1 M_{s_1;1}^\dagger \ldots M_{s_k;k}^\dagger\right\}. \tag{8}$$

Consequently, we have

$$\mathbb{P}\Big[\mu_\iota^{\mathtt{id}} = q_\iota \Big| \rho_1, \mu_\iota^{\mathtt{rl}} = p_\iota\Big] = \frac{\eta_{p_1,q_1}^1 \ldots \eta_{p_k,q_k}^k \mathrm{Tr}\left\{M_{q_k;k}\ldots M_{q_1;1}\rho_1 M_{q_1;1}^\dagger \ldots M_{q_k;k}^\dagger\right\}}{\mathrm{Tr}\left\{\sum_{s_1,\ldots,s_k=1}^{m^{\mathtt{id}}} \eta_{p_1,s_1}^1 \ldots \eta_{p_k,s_k}^k M_{s_k;k}\ldots M_{s_1;1}\rho_1 M_{s_1;1}^\dagger \ldots M_{s_k;k}^\dagger\right\}}.$$

Injecting the above relation into (5) and using (6) yields

$$\hat{\rho}_{k+1} = \frac{\sum_{q_1,\ldots,q_k} \eta_{p_1,q_1}^1 \ldots \eta_{p_k,q_k}^k M_{q_k;k}\cdots M_{q_1;1}\rho_1 M_{q_1;1}^\dagger \cdots M_{q_k;k}^\dagger}{\mathrm{Tr}\left\{\sum_{q_1,\ldots,q_k} \eta_{p_1,q_1}^1 \ldots \eta_{p_k,q_k}^k M_{q_k;k}\cdots M_{q_1;1}\rho_1 M_{q_1;1}^\dagger \ldots M_{q_k;k}^\dagger\right\}}. \tag{9}$$

It is then clear that $\hat{\rho}_{k+1}$ can be calculated from ρ_1 in a recursive manner according to

$$\hat{\rho}_{k+1} = \frac{\sum_q \eta_{p_k,q}^k M_{q;k}\hat{\rho}_k M_{q;k}^\dagger}{\mathrm{Tr}\left\{\sum_q \eta_{p_k,q}^k M_{q;k}\hat{\rho}_k M_{q;k}^\dagger\right\}},$$

$$\vdots$$

$$\hat{\rho}_2 = \frac{\sum_q \eta_{p_1,q}^1 M_{q;1}\hat{\rho}_1 M_{q;1}^\dagger}{\mathrm{Tr}\left\{\sum_q \eta_{p_1,q}^1 M_{q;1}\hat{\rho}_1 M_{q;1}^\dagger\right\}},$$

where the intermediate states correspond to optimal estimates from step 2 to k: $\hat{\rho}_j = \mathbb{E}[\rho_j | \rho_1, \mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{j-1}^{\mathtt{rl}} = p_{j-1}]$, $j = 2, \ldots, k$. The recursive relation (3) is thus proved.

We now prove (4). In the following, we set ordered product

$$\vec{\mathbf{M}}_j \triangleq M_{q_j;j}\cdots M_{q_2;2}\cdot M_{q_1;1}$$

for $j = 1, \ldots, k$.

We have

$$\mathbb{P}\Big[\mu_k^{\mathtt{rl}} = p_k \Big| \rho_1, \mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1}\Big] = \frac{\mathbb{P}\Big[\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_k^{\mathtt{rl}} = p_k \Big| \rho_1\Big]}{\mathbb{P}\Big[\mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1} \Big| \rho_1\Big]}.$$

This fraction can be computed using (8):

$$\mathbb{P}\Big[\mu_k^{\mathtt{rl}} = p_k \Big| \rho_1, \mu_1^{\mathtt{rl}} = p_1, \ldots, \mu_{k-1}^{\mathtt{rl}} = p_{k-1}\Big] = \sum_{q_1,\ldots,q_k} \eta_{p_1,q_1}^1 \ldots \eta_{p_k,q_k}^k \mathrm{Tr}\left\{\vec{\mathbf{M}}_k \rho_1 \vec{\mathbf{M}}_k^\dagger\right\} \times \left(\sum_{q_1,\ldots,q_{k-1}} \eta_{p_1,q_1}^1 \ldots \eta_{p_{k-1},q_{k-1}}^{k-1} \mathrm{Tr}\left\{\vec{\mathbf{M}}_{k-1}\rho_1 \vec{\mathbf{M}}_{k-1}^\dagger\right\}\right)^{-1}.$$

According to (9) with $k-1$ instead of k, we have

$$\hat{\rho}_k = \sum_{q_1,\dots,q_{k-1}} \eta^1_{p_1,q_1}\cdots\eta^k_{p_{k-1},q_{k-1}} \vec{\mathbf{M}}_{k-1}\rho_1\vec{\mathbf{M}}^\dagger_{k-1}\times$$
$$\mathrm{Tr}\left\{\sum_{q_1,\dots,q_{k-1}} \eta^1_{p_1,q_1}\cdots\eta^{k-1}_{p_{k-1},q_{k-1}}\mathrm{Tr}\left\{\vec{\mathbf{M}}_{k-1}\rho_1\vec{\mathbf{M}}^\dagger_{k-1}\right\}\right\}^{-1}.$$

Since

$$\sum_{q_1,\dots,q_k}\eta^1_{p_1,q_1}\cdots\eta^k_{p_k,q_k}\mathrm{Tr}\left\{\vec{\mathbf{M}}_k\rho_1\vec{\mathbf{M}}^\dagger_k\right\} = \sum_{q_k}\eta^k_{p_k,q_k}\times$$
$$\mathrm{Tr}\left\{M_{q_k;k}\left(\sum_{q_1,\dots,q_{k-1}}\eta^1_{p_1,q_1}\cdots\eta^{k-1}_{p_{k-1},q_{k-1}}\vec{\mathbf{M}}_{k-1}\rho_1\vec{\mathbf{M}}^\dagger_{k-1}\right)M^\dagger_{q_k;k}\right\},$$

we get finally (4).

■

IV. Stability with respect to initial conditions

Assume that we do not have access to the real initial state ρ_1. We cannot compute the optimal estimate $\hat{\rho}_k$. We can still use the recurrence formula (3) based on the real measurement outcomes $(\mu^{\mathtt{rl}}_j = p_j)_{j=1,\dots,k-1}$ to propose an estimation $\rho^{\mathtt{e}}_k$ of ρ_k. We will prove below that this estimation procedure is stable in the sense that the fidelity between $\hat{\rho}_k$ and $\rho^{\mathtt{e}}_k$ is non-decreasing in average whatever the initial condition $\rho^{\mathtt{e}}_1$ is.

For ease of notation we set

$$M_{p,q;k} = \sqrt{\eta^k_{p,q}}M_{q;k},$$
$$\mathcal{M}_{p;k}(\rho) = \frac{\sum_{q=1}^{m^{\mathtt{id}}} M_{p,q;k}\rho M^\dagger_{p,q;k}}{\sum_{q=1}^{m^{\mathtt{id}}}\mathrm{Tr}\left\{M_{p,q;k}\rho M^\dagger_{p,q;k}\right\}} \tag{10}$$

for any $k\geq 1$, $p\in\{1,\dots,m^{\mathtt{rl}}\}$ and $q\in\{1,\dots,m^{\mathtt{id}}\}$. The sets

$$\mathcal{S}_{p;k} \triangleq \{M_{p,1;k},\dots,M_{p,m^{\mathtt{id}};k}\}$$

for $p=1,\dots,m^{\mathtt{rl}}$ form a partition of

$$\mathcal{S}_k \triangleq \{M_{p,q;k}\mid p=1,\dots,m^{\mathtt{rl}},\ q=1,\dots,m^{\mathtt{id}}\}.$$

Using this notation, the recursive Equation (3) defines a coarse-grained Markov chain in the sense of [13].

If $\mu^{\mathtt{rl}}_k = p$, we define $\rho^{\mathtt{e}}_k$ recursively for $k\geq 1$ as follows

$$\rho^{\mathtt{e}}_{k+1} = \mathcal{M}_{p;k}(\rho^{\mathtt{e}}_k). \tag{11}$$

Such a recursive formula is valid as soon as $\sum_q \mathrm{Tr}\left\{M_{p,q;k}\rho^{\mathtt{e}}_k M^\dagger_{p,q;k}\right\} > 0$, which is automatically satisfied when $\rho^{\mathtt{e}}_k$ is of full rank. $\rho^{\mathtt{e}}_{k+1}$ is indeterminate when $\sum_q \mathrm{Tr}\left\{M_{p,q;k}\rho^{\mathtt{e}}_k M^\dagger_{p,q;k}\right\} = 0$. But using the continuity arguments developed at the end of the appendix we can give a value for $\rho^{\mathtt{e}}_{k+1}$ in the following way: for each $\rho^{\mathtt{e}}_k$, consider the set of density operators

$$\left(\mathcal{M}_{p;k}(\rho^{\mathtt{e}}_{k,\epsilon})\right)_{p=1,\dots,m^{\mathtt{rl}}},$$

where $\epsilon>0$ and $\rho^{\mathtt{e}}_{k,\epsilon} = (\rho^{\mathtt{e}}_k+\epsilon\mathbf{I})/\mathrm{Tr}\{\rho^{\mathtt{e}}_k+\epsilon\mathbf{I}\}$. Since $\rho^{\mathtt{e}}_{k,\epsilon}$ is positive definite, $\left(\mathcal{M}_{p;k}(\rho^{\mathtt{e}}_{k,\epsilon})\right)_{p=1,\dots,m^{\mathtt{rl}}}$ are well defined and admit a limit point when ϵ tends to 0^+ ($\mathcal{H}$ is of finite dimension here). Take for each $\rho^{\mathtt{e}}_k$ such a limit point $(\rho^{\mathtt{e}}_{k+1,p})_{p=1,\dots,m^{\mathtt{rl}}}$. Set $\rho^{\mathtt{e}}_{k+1} = \rho^{\mathtt{e}}_{k+1,p}$ when $\mu^{\mathtt{rl}}_k = p$. If $\sum_q \mathrm{Tr}\left\{M_{p,q;k}\rho^{\mathtt{e}}_k M^\dagger_{p,q;k}\right\} > 0$ then $\rho^{\mathtt{e}}_{k+1,p}$ coincides necessarily with $\mathcal{M}_{p;k}(\rho^{\mathtt{e}}_k)$ and we recover (11). During the proof of Theorem IV.1, we will use only recurrence (11) having in mind that, when $\sum_q \mathrm{Tr}\left\{M_{p,q;k}\rho^{\mathtt{e}}_k M^\dagger_{p,q;k}\right\} = 0$ we have to use $\rho^{\mathtt{e}}_{k+1} = \rho^{\mathtt{e}}_{k+1,p}$.

If the initial state of the system $\rho^{\mathtt{e}}_1$ coincides with ρ_1, then $\rho^{\mathtt{e}}_k$ coincides with the optimal estimate $\hat{\rho}_k$ of the state ρ_k from Theorem III.1. In fact, once the initial states $\hat{\rho}_1$ and $\rho^{\mathtt{e}}_1$ are given, the process $(\hat{\rho}_k)$ and $(\rho^{\mathtt{e}}_k)$ are driven by the same stochastic process $(\mu^{\mathtt{rl}}_{k-1})$, itself driven by the combination of the original quantum process of state (ρ_k) with the classical process associated to the left stochastic matrices (η^k) governing detection errors.

The following theorem shows that the fidelity between $\hat{\rho}_k$ and $\rho^{\mathtt{e}}_k$ is non-decreasing in average.

Theorem IV.1. *Suppose $\rho^{\mathtt{e}}_k$ satisfies the recursive relation (11) with an arbitrary initial density operator $\rho^{\mathtt{e}}_1$. Then the fidelity between $\hat{\rho}_k$ and $\rho^{\mathtt{e}}_k$ defined by*

$$F(\hat{\rho}_k,\rho^{\mathtt{e}}_k) \triangleq \left(\mathrm{Tr}\left\{\sqrt{\sqrt{\hat{\rho}_k}\rho^{\mathtt{e}}_k\sqrt{\hat{\rho}_k}}\right\}\right)^2.$$

is a submartingale in the following sense:

$$\mathbb{E}\left[F(\hat{\rho}_{k+1},\rho^{\mathtt{e}}_{k+1})\middle|\hat{\rho}_1,\dots,\hat{\rho}_k,\rho^{\mathtt{e}}_1,\dots,\rho^{\mathtt{e}}_k\right]\geq F(\hat{\rho}_k,\rho^{\mathtt{e}}_k).$$

Proof: Denote $F_k = F(\hat{\rho}_k,\rho^{\mathtt{e}}_k)$. We have

$$\begin{aligned}
&\mathbb{E}\left[F_{k+1}\middle|\hat{\rho}_1,\dots,\hat{\rho}_k,\rho^{\mathtt{e}}_1,\dots,\rho^{\mathtt{e}}_k\right]\\
&= \mathbb{E}\left[F_{k+1}\middle|\hat{\rho}_1,\rho^{\mathtt{e}}_1,\mu^{\mathtt{rl}}_1=p_1,\dots,\mu^{\mathtt{rl}}_{k-1}=p_{k-1}\right]\\
&= \sum_{p=1}^{m^{\mathtt{rl}}}\mathbb{P}\left[\mu^{\mathtt{rl}}_k=p\middle|\hat{\rho}_1,\rho^{\mathtt{e}}_1,\mu^{\mathtt{rl}}_1=p_1,\dots,\mu^{\mathtt{rl}}_{k-1}=p_{k-1}\right]\times\\
&\qquad\mathbb{E}\left[F_{k+1}\middle|\hat{\rho}_1,\rho^{\mathtt{e}}_1,\mu^{\mathtt{rl}}_1=p_1,\dots,\mu^{\mathtt{rl}}_{k-1}=p_{k-1},\mu^{\mathtt{rl}}_k=p\right].
\end{aligned}$$

The conditional probabilities appearing in this sum are given by (4) and the conditional expectations read

$$\mathbb{E}\left[F_{k+1}\middle|\hat{\rho}_1,\rho^{\mathtt{e}}_1,\mu^{\mathtt{rl}}_1=p_1,\dots,\mu^{\mathtt{rl}}_{k-1}=p_{k-1},\mu^{\mathtt{rl}}_k=p\right]$$
$$= F\left(\mathcal{M}_{p;k}(\hat{\rho}_k)\ ,\ \mathcal{M}_{p;k}(\rho^{\mathtt{e}}_k)\right)$$

since, once $\hat{\rho}_1$ and $\rho^{\mathtt{e}}_1$ and $\mu^{\mathtt{rl}}_1=p_1,\dots,\mu^{\mathtt{rl}}_{k-1}=p_{k-1}$ and $\mu^{\mathtt{rl}}_k=p$ are given, $\hat{\rho}_{k+1}=\mathcal{M}_{p;k}(\hat{\rho}_k)$ and $\rho^{\mathtt{e}}_{k+1}=\mathcal{M}_{p;k}(\rho^{\mathtt{e}}_k)$. Thus we have

$$\begin{aligned}
&\mathbb{E}\left[F_{k+1}\middle|\hat{\rho}_1,\dots,\hat{\rho}_k,\rho^{\mathtt{e}}_1,\dots,\rho^{\mathtt{e}}_k\right]=\\
&\sum_{p=1}^{m^{\mathtt{rl}}}\left(\sum_q\mathrm{Tr}\left\{M_{p,q;k}\hat{\rho}_k M^\dagger_{p,q;k}\right\}\right)\times\\
&F\left(\frac{\sum_q M_{p,q;k}\hat{\rho}_k M^\dagger_{p,q;k}}{\sum_q\mathrm{Tr}\left\{M_{p,q;k}\hat{\rho}_k M^\dagger_{p,q;k}\right\}}\ ,\ \frac{\sum_q M_{p,q;k}\rho^{\mathtt{e}}_k M^\dagger_{p,q;k}}{\sum_q\mathrm{Tr}\left\{M_{p,q;k}\rho^{\mathtt{e}}_k M^\dagger_{p,q;k}\right\}}\right).
\end{aligned}$$

The fact that $\mathbb{E}\left[F_{k+1}\middle|\hat{\rho}_1,\ldots,\hat{\rho}_k,\rho^{\mathtt{e}}_1,\ldots,\rho^{\mathtt{e}}_k\right] \geq F_k$ is then a direct consequence of equation (12) given in appendix with $r = m^{\mathtt{rl}}$, $s = m^{\mathtt{rl}}m^{\mathtt{id}}$, index j corresponding to p, index i to (p,q), operators L_i to $M_{p,q;k}$, density operators ρ and σ to $\hat{\rho}_k$ and $\rho^{\mathtt{e}}_k$, respectively.

■

V. Example: Quantum Filter for the Photon-Box

This section considers, as a key illustration, the quantum filter design in [14] to estimate in real-time the state ρ of a quantized electro-magnetic field. Since this filter admits exactly the recursive form of Theorem III.1, Theorem IV.1 applies and thus, this filter is proved here to be stable versus its initial condition.

The actual experiment under consideration uses quantum non-demolition measurements [12] to estimate the state of the quantized field trapped in a superconducting microwave cavity. Circular Rydberg atoms are sent at discrete time intervals to perform partial measurements of the photon number. Atoms are subsequently detected either in their excited (e) or ground (g) state. The outcomes of these measurements are then used to estimate the state of the cavity field, thanks to the quantum filter described below. This estimation is eventually used to calculate the amplitude of classical fields injected in the cavity in order to bring the field closer to a predefined target state. The interested reader is directed to [10] and [14] for further details of the experimental setup and results obtained.

The Hilbert space $\mathcal{H}$ of the cavity is, up to some finite photon number truncation (around 10), the Fock space with basis $\{|n\rangle\}_{n\geq 0}$, each $|n\rangle$ being the Fock state associated to exactly n photons (photon-number state). The annihilation operator $\mathbf{a} : \mathcal{H} \mapsto \mathcal{H}$ is defined by $\mathbf{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle$ for $n \geq 1$ and $\mathbf{a}\,|0\rangle = 0$. Its Hermitian conjugate $\mathbf{a}^\dagger$ is the creation operator satisfying $\mathbf{a}\,|n\rangle = \sqrt{n+1}\,|n+1\rangle$, for all $n \geq 0$. The photon-number operator (energy operator) is $\mathbf{N} = \mathbf{a}^\dagger\mathbf{a}$ which satisfies $\mathbf{N}\,|n\rangle = n\,|n\rangle$, for all $n \geq 0$. Recall the commutation $[\mathbf{a},\mathbf{a}^\dagger] = \mathbf{I}$.

In [14] the following imperfections have been considered:

- atomic preparation efficiency characterized by $P_a(n_a) \geq 0$, the probability to have $n_a \in \{0,1,2\}$ atom(s) interact with the cavity: $P_a(0)+P_a(1)+P_a(2) = 1$.
- detection efficiency characterized by $\epsilon_d \in [0,1]$, the probability that the detector detects an atom when it is present.
- state detection error rate $\eta_g \in [0,1]$ (resp. η_e) probability of erroneous state assignation to e (resp. g) when the atom collapses in g (resp. e).

The original state ρ is subject to $m^{\mathtt{id}} = 3\times 7$ possible quantum jumps and the available sensors (atomic detector) admits only $m^{\mathtt{rl}} = 6$ possibilities. We begin by introducing some operators that are used to describe these quantum jumps:

$$D_\alpha = e^{\alpha\mathbf{a}^\dagger - \alpha^*\mathbf{a}}, \quad L_{no} = \sqrt{P_a(0)}\;\mathbf{I}, \quad L_g = \sqrt{P_a(1)}\cos\phi_{\mathbf{N}},$$
$$L_e = \sqrt{P_a(1)}\sin\phi_{\mathbf{N}}, \quad L_{gg} = \sqrt{P_a(2)}\cos^2\phi_{\mathbf{N}},$$
$$L_{ge} = L_{eg} = \sqrt{P_a(2)}\cos\phi_{\mathbf{N}}\sin\phi_{\mathbf{N}}, \quad L_{ee} = \sqrt{P_a(2)}\sin^2\phi_{\mathbf{N}},$$
$$L_o = 1 - \tfrac{\epsilon(1+2n_{th})}{2}\mathbf{N} - \tfrac{\epsilon n_{th}}{2}\mathbf{I}, \quad L_+ = \sqrt{\epsilon(1+n_{th})}\mathbf{a},$$
$$L_- = \sqrt{\epsilon n_{th}}\mathbf{a}^\dagger,$$

where $\phi_{\mathbf{N}} = \frac{\phi_0(\mathbf{N}+1/2)+\phi_R}{2}$ and $0 < \epsilon, n_{th} \ll 1$, ϕ_0, ϕ_R are real experimental parameters. The unitary displacement operator D_α corresponds to the control input $\alpha \in \mathbb{C}$, depending on the sampling step k. The operators L_o, L_+ and L_- correspond to the interaction of the cavity-field with its environment (decoherence due to mirrors and thermal photons):

1) L_o corresponds to no photon jump;
2) L_+ corresponds to the capture of one thermal photon by the cavity-field;
3) L_- corresponds to one photon lost from the cavity-field.

Since $L_o^\dagger L_o + L_+^\dagger L_+ + L_-^\dagger L_- = \mathbf{I} + O(\epsilon^2)$ and ϵ is small, we consider in the sequel that (L_o, L_+, L_-) are associated to an effective Kraus map $L_o\rho L_o^\dagger + L_+\rho L_+^\dagger + L_-\rho L_-^\dagger$.

The operators L_{no}, L_g, L_e, L_{gg}, L_{ge}, L_{eg} and L_{ee} correspond to the jump induced by the collapse of possible crossing atom(s) having interacted with the cavity-field:

1) L_{no} - no atom in the atomic sample;
2) L_g - one atom having interacted with the cavity-field and collapsed to the atomic ground state during the detection process;
3) L_e - one atom having interacted with the cavity-field and collapsed to the atomic excited state during the detection process;
4) L_{gg} - two atoms having interacted with the cavity-field, both having collapsed to g;
5) L_{ge} - two atoms having interacted with the cavity-field, the first one having collapsed to g and the second to e;
6) L_{eg} - two atoms having interacted with the cavity-field, the first one having collapsed to e and the second to g;
7) L_{ee} - two atoms having interacted with the cavity-field, both having collapsed to e.

For each control input α, we have a total of $m^{\mathtt{id}} = 3\times 7$ Kraus operators. The jumps are labeled by $q = (q^a, q^c)$ with $q^a \in \{no, g, e, gg, ge, eg, ee\}$ labeling atom related jumps and $q^c \in \{o, +, -\}$ cavity decoherence jumps. The Kraus operators associated to such q are $M_q = L_{q^c}D_\alpha L_{q^a}$. So, for instance, with the control input α_k at step k:

- the Kraus operator corresponding to one atom collapsing in ground state, $q^a = g$, and one photon destroyed by mirrors, $q^c = -$, reads $M^k_{(g,-)} = L_- D_{\alpha_k} L_g$.
- the Kraus operator corresponding to two atoms, the first one collapsing to g, the second one to e, $q^a = ge$, and one thermal photon being caught between the two mirrors, $q^c = +$, reads $M^k_{(ge,+)} = L_+ D_{\alpha_k} L_{ge}$.

One can check that, for any value of α, these 21 operators define a Kraus map (using the assumption that $L_o^\dagger L_o + L_+^\dagger L_+ + L_-^\dagger L_- \approx \mathbf{I}$).

$p \setminus q$	(no, q^c)	(g, q^c)	(e, q^c)	(gg, q^c)	(ee, q^c)	(ge, q^c) or (eg, q^c)
no	1	$1-\epsilon_d$	$1-\epsilon_d$	$(1-\epsilon_d)^2$	$(1-\epsilon_d)^2$	$(1-\epsilon_d)^2$
g	0	$\epsilon_d(1-\eta_g)$	$\epsilon_d\eta_e$	$2\epsilon_d(1-\epsilon_d)(1-\eta_g)$	$2\epsilon_d(1-\epsilon_d)\eta_e$	$\epsilon_d(1-\epsilon_d)(1-\eta_g+\eta_e)$
e	0	$\epsilon_d\eta_g$	$\epsilon_d(1-\eta_e)$	$2\epsilon_d(1-\epsilon_d)\eta_g$	$2\epsilon_d(1-\epsilon_d)(1-\eta_e)$	$\epsilon_d(1-\epsilon_d)(1-\eta_e+\eta_g)$
gg	0	0	0	$\epsilon_d^2(1-\eta_g)^2$	$\epsilon_d^2\eta_e^2$	$\epsilon_d^2\eta_e(1-\eta_g)$
ge	0	0	0	$2\epsilon_d^2\eta_g(1-\eta_g)$	$2\epsilon_d^2\eta_e(1-\eta_e)$	$\epsilon_d^2((1-\eta_g)(1-\eta_e)+\eta_g\eta_e)$
ee	0	0	0	$\epsilon_d^2\eta_g^2$	$\epsilon_d^2(1-\eta_e)^2$	$\epsilon_d^2\eta_g(1-\eta_e)$

TABLE I
STOCHASTIC MATRIX $\eta_{p,q}$.

We have only $m^{\text{rl}} = 6$ real detection possibilities $p \in \{no, g, e, gg, ge, ee\}$ corresponding respectively to no detection, a single detection in g, a single detection in e, a double detection both in g, a double detection one in g and the other in e, and a double detection both in e. A double detection does not distinguish two atoms. The entries $\eta_{p,q}$ of the stochastic matrix describing the imperfect detections corrupted by errors are independent of α and given by Table I. It relies on error rate parameters η_g, η_e and ϵ_d in $[0, 1]$ and assume no error correlation between atoms. So for instance the probability that there is a single atom in the sample, which collapses to the ground state and is in fact detected by the experimental sensors to be in the ground state is given by $\eta_{g,(g,q^c)} = \epsilon_d(1-\eta_g)$. Note that each column in the table sums up to one, since $\eta_{p,q}$ is a left stochastic matrix.

VI. Conclusion

In this paper, we derive a recursive expression for the optimal estimate of a quantum system's state from imperfect, discrete-time measurements. The optimality of this recursive expression is proven by a simple application of Bayes' lemma and quantum measurement postulates. Such a filter is shown to satisfy a Markov property and thus can be used for quantum control as shown in [14] or to run Monte Carlo simulations of the measurement trajectories. We also demonstrate that this filter is stable with respect to initial conditions in the sense of Theorem IV.1. In the future, the continuous-time version of these results will be investigated.

References

[1] V. P. Belavkin. Quantum filtering of markov signals with white quantum noise. *Radiotechnika i Electronika*, 25:1445–1453, 1980.

[2] V. P. Belavkin. Theory of the control of observable quantum-systems. *Automation and Remote Control*, 44:178, 1983.

[3] V. P. Belavkin. Nondemolition stochastic calculus in fock space and nonlinear filtering and control in quantum systems. In R. Guelerak and W. Karwowski, editors, *XXIV Karpacz winter school*,, Stochastic methods in mathematics and physics, World Scientific, Singapore, pages 310–324, 1988.

[4] V. P. Belavkin. Quantum stochastic calculus and quantum nonlinear filtering. *J. Multivar. Anal.*, 42:171–201, 1992.

[5] A. Bensoussan. *Stochastic Control of Partially Observable Systems*. Cambridge University Press, 1992.

[6] Luc Bouten, Ramon Van Handel, and Matthew R. James. An introduction to quantum filtering. *SIAM Journal on Control and Optimization*, 46(6):2199–2241, 2007.

[7] Luc Bouten, Ramon van Handel, and Matt James. A discrete invitation to quantum filtering and feedback control. *SIAM Review 51*, 51:239–316, 2009.

[8] Vladimir B. Braginsky and Farid Ya. Khalili. *Quantum Measurement*. Cambridge University Press, Cambridge, 1992.

[9] E. B. Davies. *Quantum Theory of Open Systems*. Academic Press, London New-York San Francisco, 1976.

[10] I. Dotsenko, M. Mirrahimi, M. Brune, S. Haroche, J.-M. Raimond, and P. Rouchon. Quantum feedback by discrete quantum non-demolition measurements: towards on-demand generation of photon-number states. *Physical Review A*, 80:013805–013813, 2009.

[11] C. W. Gardiner and P. Zoller. *Quantum Noise*. Springer, Berlin, 2000.

[12] S. Haroche and J.M. Raimond. *Exploring the Quantum: Atoms, Cavities and Photons*. Oxford University Press, 2006.

[13] P. Rouchon. Fidelity is a sub-martingale for discrete-time quantum filters. *to appear in IEEE Transaction of Automatic Control*, 2011. preprint arXiv:1005.3119v3.

[14] C Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, Th. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, J.M. Raimond, and S. Haroche. Real-time quantum feedback prepares and stabilizes photon number states. *Nature*, 477(7362), 2011.

[15] R. van Handel. The stability of quantum markov filters. *Infin. Dimens. Anal. Quan- tum Probab. Relat. Top.*, 12:153–172, 2009.

[16] Howard M. Wiseman and Gerard J. Milburn. *Quantum Measurement and Control*. Cambridege University Press, 2009.

Appendix

Consider a set of s operators $(L_i)_{i=1,\dots,s}$ on the finite dimensional Hilbert space $\mathcal{H}$, such that $L_i \neq 0$ for all i and $\sum_{i=1}^s L_i^\dagger L_i = \mathbf{I}$. Take a partition of $\{1, \dots, s\}$ into $r \geq 1$ non-empty sub-sets $(\mathcal{P}_j)_{j=1,\dots,r}$. Then, for any semi-definite positive operators ρ and σ on $\mathcal{H}$ with unit traces, the following inequality proved in [13] holds true:

$$F(\rho,\sigma) \leq \sum_{j=1}^{r} \text{Tr}\left\{\Sigma_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\}\times F\left(\frac{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger}{\text{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\}}, \frac{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger}{\text{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger\right\}}\right), \quad (12)$$

where $F(\rho,\sigma) = F(\sigma,\rho)$ is the fidelity between σ and ρ defined as

$$F(\rho,\sigma) = \left(\text{Tr}\left\{\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right\}\right)^2.$$

When, in the above sum, a denominator $\text{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\}$ depending on ρ vanishes, the sum remains still valid since the corresponding value of F is multiplied by the same vanishing factor and F is bounded since between 0 and 1. This is no more true when a denominator $\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger\right\}$ depending on σ vanishes. In this case the above inequality should be interpreted in the following way relying on a continuity argument. For each σ, define

$$\mathcal{Z}_\sigma = \left\{ j \;\middle|\; \mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger\right\} = 0 \right\} \subset \{1,\ldots,r\}.$$

For almost all density operators σ, $\mathcal{Z}_\sigma = \emptyset$. Take $\epsilon > 0$ and consider the positive definite density operator $\sigma^\epsilon = \frac{\sigma+\epsilon\mathbf{I}}{\mathrm{Tr}\{\sigma+\epsilon\mathbf{I}\}}$. For $j \in \{1,\ldots,r\}$, $\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma^\epsilon L_i^\dagger\right\} > 0$ since each $L_i \neq 0$. In particular, the set of density operators

$$\left(\frac{\sum_{i\in\mathcal{P}_j} L_i\sigma^\epsilon L_i^\dagger}{\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma^\epsilon L_i^\dagger\right\}}\right)_{j\in\mathcal{Z}_\sigma}$$

admits at least an accumulation point when ϵ tends to 0^+ ($\mathcal{H}$ of finite dimension implies that the set of density operators is compact). Denote by $(\sigma_j^+)_{j\in\mathcal{Z}_\sigma}$, such an accumulation point where each σ_j^+ is a density operator. Since for any $\epsilon > 0$, inequality (12) holds true for ρ and σ^ϵ, F is continuous, and for any $j \notin \mathcal{Z}_\sigma$, $\frac{\sum_{i\in\mathcal{P}_j} L_i\sigma^\epsilon L_i^\dagger}{\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma^\epsilon L_i^\dagger\right\}}$ tends to $\frac{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger}{\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger\right\}}$, we have by continuity

$$F(\rho,\sigma) \geq \sum_{j\in\mathcal{Z}_\sigma} \mathrm{Tr}\left\{\Sigma_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\} F\left(\frac{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger}{\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\}} \;,\; \sigma_j^+\right)$$
$$+ \sum_{j\notin\mathcal{Z}_\sigma} \mathrm{Tr}\left\{\Sigma_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\} F\left(\frac{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger}{\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\rho L_i^\dagger\right\}} \;,\; \frac{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger}{\mathrm{Tr}\left\{\sum_{i\in\mathcal{P}_j} L_i\sigma L_i^\dagger\right\}}\right).$$

Such continuity argument show that we can extend inequality (12) via the accumulation value(s) σ_j^+ when $\mathcal{Z}_\sigma$ is not empty.