ON THE SEPARATION PRINCIPLE OF QUANTUM CONTROL

LUC BOUTEN AND RAMON VAN HANDEL

Abstract. 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.

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 [Bel83, Bel88] it is only recently that this has become a feasible technology, as demonstrated by recent laboratory experiments in quantum optics [AAS$^+$02, GSM04]. 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 [HP84]. 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 [Bel92, BGM04, BV05] 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 work was supported in part by the NSF under contract number PHY-0456720, and in part by the ARO under contract number DAAD19-03-1-0073.

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 function (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

$$d\pi_t(X) = u(t)\, \pi_t(i[H,X])\, dt + \pi_t(\mathcal{L}[X])\, dt \\ + (\pi_t(XL+L^*X) - \pi_t(L+L^*)\pi_t(X))\, (dY_t - \pi_t(L+L^*)\, dt).$$

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 [VSM05, MV05].

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 indeed 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 [BV05] which is inspired by the approach of Zakai [Zak69] in classical nonlinear filtering. Our treatment of the separation theorem is directly inspired by the classic papers of Wonham [Won68] and Segall [Seg77]. A fully technical account of the results in this article will be presented in [BV06], 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 [BV05] 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 $\mathbf{P}\ll\mu$.[1] Thus any self-adjoint element of $\mathcal{A}$ represents a bounded random 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

[1]We will also denote the *-isomorphism as $\iota:(\mathcal{A},\mathbb{P})\to L^\infty(\Omega,\Sigma,\mu,\mathbf{P})$; this means that the null sets are quotiented w.r.t. the measure μ, whereas $\mathbf{P}\ll\mu$ is the image of the state $\mathbb{P}$.

of (commuting) observations to be performed in every realization of the experiment, then many computations can be reduced to classical probability theory.

To define the notion of a conditional expectation in quantum probability, we now 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 [Tak71]. 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 [BV05].

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 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 2.2 (Bayes formula [BV05])**.** *Let $(\mathcal{A}, \mathbb{P})$ be a quantum probability space and $\mathcal{C} \subset \mathcal{A}$ be a commutative von Neumann algebra. Furthermore, 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 we have $\mathbb{Q}(X|\mathcal{C}) = \mathbb{P}(V^*XV|\mathcal{C}) \,/\, \mathbb{P}(V^*V|\mathcal{C})$ 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 [BV06] 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 [BV05].

A completely general theory can also be set up, e.g. [Bel92], 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 [HP84, Bia95, Mey93, Par92] 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

$$\begin{aligned} dj_t(X) = {} & j_t(i[H_t, X] + L_t^* X L_t - \tfrac{1}{2}\{L_t^* L_t X + X L_t^* L_t\}) \, dt \\ & + j_t(S_t^*[X, L_t]) \, dA_t^* + j_t([L_t^*, X] S_t) \, dA_t + j_t(S_t^* X S_t - X) \, d\Lambda_t, \end{aligned} \tag{2.1}$$

$$\begin{aligned} dY_t = {} & \lambda_t \, d\Lambda_t + j_t\left(S_t^*(\alpha_t + \lambda_t L_t)\right) dA_t^* + j_t\left((\alpha_t^* + \lambda_t L_t^*) S_t\right) dA_t \\ & + j_t\left(\lambda_t L_t^* L_t + \alpha_t^* L_t + \alpha_t L_t^*\right) dt. \end{aligned} \tag{2.2}$$

This "system-theoretic" description in terms of the system-observations pair (2.1) and (2.2) 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 $s \le t \in [0,T]$. 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 immediately. □

To obtain an explicit representation for the filtering equation we are inspired by the classical reference probability method of Zakai [Zak69]. The idea of Zakai is to introduce a change of measure such that under the new (reference) measure the observation process is a martingale which is independent of the system observables, which significantly simplifies the calculation of conditional expectations. The Bayes formula then relates the conditional expectations under the reference measure and the actual measure. We will follow a similar route and make use of the quantum Bayes formula of Lemma 2.2. 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 2.2. 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 [Hol91]. We state it here in the following form.

Lemma 2.3. *Let $C_t, D_t, F_t, G_t, \tilde{C}_t, \tilde{F}_t$ be bounded processes, and let*

$$dV_t = \{C_t\, d\Lambda_t + D_t\, dA_t^* + F_t\, dA_t + G_t\, dt\}V_t,$$
$$d\tilde{V}_t = \{\tilde{C}_t\, d\Lambda_t + D_t\, dA_t^* + \tilde{F}_t\, dA_t + G_t\, dt\}\tilde{V}_t, \quad V_0 = \tilde{V}_0.$$

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 $\langle V_t\, v \otimes \Phi, X V_t\, v \otimes \Phi\rangle = \langle \tilde{V}_t\, v \otimes \Phi, X \tilde{V}_t\, v \otimes \Phi\rangle$ 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. [Mey93])

$$\|(V_t - \tilde{V}_t)\, v \otimes \Phi\|^2 = \int_0^t \langle (V_s - \tilde{V}_s)\, v \otimes \Phi, (G_s + G_s^*)(V_s - \tilde{V}_s)\, v \otimes \Phi\rangle\, ds$$
$$+ \int_0^t \langle D_s(V_s - \tilde{V}_s)\, v \otimes \Phi, D_s(V_s - \tilde{V}_s)\, v \otimes \Phi\rangle\, ds.$$

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

$$\langle (V_s - \tilde{V}_s)\, v \otimes \Phi, (G_s + G_s^*)(V_s - \tilde{V}_s)\, v \otimes \Phi\rangle \le \|K_s^+(V_s - \tilde{V}_s)\, v \otimes \Phi\|^2.$$

But as $G_s + G_s^*$ is a bounded process, so is K_s^+ and we have

$$\|K_s^+(V_s - \tilde{V}_s)\, v \otimes \Phi\|^2 \le \left[\sup_{t\in[0,T]} \|K_t^+\|^2\right] \|(V_s - \tilde{V}_s)\, v \otimes \Phi\|^2.$$

Thus we obtain

$$\|(V_t - \tilde{V}_t)\, v \otimes \Phi\|^2 \le C \int_0^t \|(V_s - \tilde{V}_s)\, v \otimes \Phi\|^2\, ds$$

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 note that the former also covers some cases with counting observations if $\lambda_t \neq 0$.)

Theorem 2.4 (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 2.5 (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*

$$E_t = I + \int_0^t \lambda_s^{-1}(1-\alpha_s) E_s\, dA_s^* - \int_0^t \lambda_s^{-1}(1-\alpha_s^*) E_s\, dA_s - \frac{1}{2}\int_0^t \lambda_s^{-2} |1-\alpha_s|^2 E_s\, ds$$

and let $C_t = E_t^* Z_t E_t$*,* $\mathcal{C}_t = \mathrm{vN}(C_s : 0 \le s \le t) = E_t^* \mathcal{Z}_t E_t$*, 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* $\sigma_t(X) = U_t^* E_t \mathbb{P}(V_t^* X V_t | \mathcal{C}_t) E_t^* U_t$*,* $X \in \mathcal{B}$*.*

Proof of Theorem 2.4. We begin by using a general transformation property of the quantum conditional expectation. Let U be a unitary operator and define a new state $\mathbb{Q}(X) = \mathbb{P}(U^* X U)$. Then $\mathbb{P}(U^* X U | U^* \mathcal{C} U) = U^* \mathbb{Q}(X|\mathcal{C}) U$ 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 2.3, 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 $\mathbb{Q}(X) = \mathbb{P}(U_t^* X U_t)$. Evidently $\pi_t(X) = U_t^* \mathbb{Q}(X|\mathcal{Z}_t) U_t$.

We would now like to apply Lemma 2.2 to $\mathbb{Q}(X|\mathcal{Z}_t)$. Note that by Lemma 2.3 we obtain $\mathbb{P}(V_t^* X V_t) = \mathbb{P}(U_t^* X U_t) = \mathbb{Q}(X)$. 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 2.2 we obtain $\mathbb{Q}(X|\mathcal{Z}_t) = \mathbb{P}(V_t^* X V_t | \mathcal{Z}_t)/\mathbb{P}(V_t^* V_t | \mathcal{Z}_t)$. The statement of the Theorem follows immediately. □

Proof of Theorem 2.5. The proof proceeds along the same lines as the proof of Theorem 2.4, except that in order to apply Lemma 2.3 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 $\mathbb{Q}(X) = \mathbb{P}(U_t^* E_t X E_t^* U_t)$, 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 $\mathbb{P}(V_t^* X V_t) = \mathbb{Q}(X)$ by Lemma 2.3 and by application of the quantum Itô rules to $E_t^* U_t$. The statement of the Theorem follows from Lemma 2.2. □

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 2.4. Using the quantum Itô rules, we can write

$$V_t^*XV_t = X + \int_0^t V_s^*(i[H_s,X] + L_s^*XL_s - \tfrac{1}{2}\{L_s^*L_sX + XL_s^*L_s\})V_s\,ds$$
$$+ \int_0^t V_s^*(\alpha_s^{-1}XL_s + \alpha_s^{-1*}L_s^*X + (\alpha_s^*\alpha_s)^{-1}\lambda_sL_s^*XL_s)V_s\,dZ_t$$

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

$$\mathbb{P}(V_t^*XV_t - X|\mathcal{Z}_t) = \int_0^t \mathbb{P}(V_s^*(i[H_s,X] + L_s^*XL_s - \tfrac{1}{2}\{L_s^*L_sX + XL_s^*L_s\})V_s|\mathcal{Z}_s)\,ds$$
$$+ \int_0^t \mathbb{P}(V_s^*(\alpha_s^{-1}XL_s + \alpha_s^{-1*}L_s^*X + (\alpha_s^*\alpha_s)^{-1}\lambda_sL_s^*XL_s)V_s|\mathcal{Z}_s)\,dZ_t,$$

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; another application of the quantum Itô rules gives

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s,X] + L_s^*XL_s - \tfrac{1}{2}\{L_s^*L_sX + XL_s^*L_s\})\,ds$$
$$+ \int_0^t \sigma_s(\alpha_s^{-1}XL_s + \alpha_s^{-1*}L_s^*X + (\alpha_s^*\alpha_s)^{-1}\lambda_sL_s^*XL_s)\,dY_t$$

where dY_t is given by (2.2). This is the noncommutative counterpart of the linear filtering equation of classical nonlinear filtering theory. Applying a similar procedure to Theorem 2.5 yields a linear 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 linear 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. In this section we introduce the notion of a controlled QSDE, and show that this gives rise to a controlled linear 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; and

(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 [BV06] that Z_t is well defined and that U_t has a unique unitary solution (see also [Mey93, Hol96, GLW01] 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

$$\begin{aligned} dj_t(X) = {} & j_t(i[H_t, X] + L_t^* X L_t - \tfrac{1}{2}\{L_t^* L_t X + X L_t^* L_t\}) \, dt \\ & + j_t(S_t^* [X, L_t]) \, dA_t^* + j_t([L_t^*, X] S_t) \, dA_t + j_t(S_t^* X S_t - X) \, d\Lambda_t, \end{aligned} \tag{3.1}$$

$$\begin{aligned} dY_t = {} & \Xi_t \, d\Lambda_t + j_t \left(S_t^* (\Upsilon_t + \Xi_t L_t) \right) dA_t^* + j_t \left((\Upsilon_t^* + \Xi_t L_t^*) S_t \right) dA_t \\ & + j_t \left(\Xi_t L_t^* L_t + \Upsilon_t^* L_t + \Upsilon_t L_t^* \right) dt, \end{aligned} \tag{3.2}$$

which is simply the controlled counterpart of (2.1), (2.2). 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 3.2 (Hamiltonian feedback)**.** Consider the controlled quantum flow

$$dU_t = (L \, dA_t^* - L^* \, dA_t - \tfrac{1}{2} L^* L \, dt - i u_t(Z_{s \le t}) H \, dt) \, U_t, \qquad Z_t = A_t + A_t^*.$$

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

$$\begin{aligned} dj_t(X) = {} & j_t([X, L]) \, dA_t^* + j_t([L^*, X]) \, dA_t \\ & + j_t(L^* X L - \tfrac{1}{2}(L^* L X + X L^* L)) \, dt + u_t(Y_{s \le t}) \, j_t(i[H, X]) \, dt \end{aligned}$$

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. [VSM05].

Example 3.3 (Coherent feedback)**.** The controlled quantum flow defined by

$$dU_t = ((L+u_t(Z_{s\le t})I)\,dA_t^* - (L^*+u_t(Z_{s\le t})^*I)\,dA_t - \tfrac{1}{2}(L^*L+u_t(Z_{s\le t})^*u_t(Z_{s\le t})I+2u_t(Z_{s\le t})L^*)\,dt)\,U_t,$$

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 $u_t(Z_{s\le t})$ [Bou04]. As in the previous example, the control becomes a function of the observations when we transform to the system-theoretic description.

Example 3.4 (Adaptive measurement)**.** In this scenario we choose an uncontrolled Hudson-Parthasarathy equation $dU_t = (L\,dA_t^* - L^*\,dA_t - \frac{1}{2}L^*L\,dt - iH\,dt)\,U_t$, 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 the observations

$$dY_t = dZ_t + \left[j_t(L)\,e^{iu_t(Y_{s\le t})} + j_t(L^*)\,e^{-iu_t(Y_{s\le t})}\right]\,dt.$$

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. [Wis95, AAS$^+$02].

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 2.3 extend readily to the controlled case, and it is straightforward to extend the proofs of Theorems 2.4 and 2.5 to prove the following statements.

Theorem 3.5 (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*

$$V_t = I + \int_0^t (\Xi_s\,d\Lambda_s + \Upsilon_s\,dA_s^* + \Upsilon_s^*\,dA_s)\,\Upsilon_s^{-1}L_sV_s - \int_0^t (iH_s + \tfrac{1}{2}L_s^*L_s)V_s\,ds.$$

Then V_t is affiliated to $\mathscr{Z}_t'$ for every $t\in[0,T]$, and we have the representation $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$ with $\sigma_t(X) = U_t^*\mathbb{P}(V_t^*XV_t|\mathscr{Z}_t)U_t$, $X\in\mathcal{B}\otimes\mathscr{Z}_t$.

Theorem 3.6 (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*

$$E_t = I + \int_0^t \Xi_s^{-1}(1-\Upsilon_s)E_s\,dA_s^* - \int_0^t \Xi_s^{-1}(1-\Upsilon_s^*)E_s\,dA_s - \frac{1}{2}\int_0^t \Xi_s^{-2}|1-\Upsilon_s|^2E_s\,ds$$

where $|X|^2 = X^*X$. Let $C_t = E_t^*Z_tE_t$, $\mathcal{C}_t = \mathrm{vN}(C_s : 0\le s\le t) = E_t^*\mathscr{Z}_tE_t$, 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

$$V_t = I + \int_0^t (E_s^*\Xi_sE_s\,d\Lambda_s + dA_s^* + dA_s)\,E_s^*(L_s - \Xi_s^{-1}(1-\Upsilon_s))E_sV_s - \int_0^t E_s^*(iH_s + \tfrac{1}{2}L_s^*L_s + \tfrac{1}{2}\Xi_s^{-2}|1-\Upsilon_s|^2 - \Xi_s^{-1}(1-\Upsilon_s^*)L_s)E_sV_s\,ds.$$

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* $\sigma_t(X) = U_t^* E_t \mathbb{P}(V_t^* E_t^* X E_t V_t | \mathcal{C}_t) E_t^* U_t$*,* $X \in \mathcal{B} \otimes \mathcal{Z}_t$*.*

We can now obtain controlled filtering equations for $\sigma_t(\cdot)$ in a recursive form. The following statements are readily verified using the quantum Itô rules.

Corollary 3.7 (linear filtering equation, diffusive case)**.** *Suppose that* Υ_t *has a bounded inverse for all* $t \in [0,T]$*. Then* $\sigma_t(\cdot)$ *of Theorem 3.5 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 (3.2) and* $\pi_t(X) = \sigma_t(X)/\sigma_t(I)$*.*

Corollary 3.8 (linear filtering equation, counting case)**.** *Suppose that* Ξ_t *has a bounded inverse for all* $t \in [0,T]$*. Then* $\sigma_t(\cdot)$ *of Theorem 3.6 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 (3.2) 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 linear filtering equation

$$d\sigma_t(X) = u(t)\, \sigma_t(i[H, X])\, dt + \sigma_t(L^* X L - \tfrac{1}{2}\{L^* L X + X L^* L\})\, dt + \sigma_t(X L + L^* X)\, dY_t.$$

In closed loop the model is given by the controlled quantum flow of Example 3.2, and by Corollary 3.7 we obtain the controlled linear filtering equation

$$d\sigma_t(X) = u(Y_{s \le t})\, \sigma_t(i[H, X])\, dt \\ + \sigma_t(L^* X L - \tfrac{1}{2}\{L^* L X + X L^* L\})\, dt + \sigma_t(X L + L^* X)\, dY_t.$$

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 3.7 and 3.8. 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 3.7 and 3.8. 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 linear filtering equation at that time.

In order for this to be sensible, we have to show that the controlled linear 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 linear filtering equations of Corollaries 3.7 and 3.8 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)**.** *Let $dZ_t = \Xi_t\,d\Lambda_t + \Upsilon_t\,dA_t^* + \Upsilon_t^*\,dA_t$ 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. for all $s \le t \in [0,T]$ we have $\mathbb{P}(\overline{Z}_t|\mathcal{Y}_s) = \overline{Z}_s$.

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 (3.2). □

The innovations process is the starting point for martingale-based approaches to (quantum) filtering [Bel92, BGM04]. 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 [LS01], 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 [LS01]. 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 linear filtering equation of Corollary 3.7 to a classical stochastic differential equation. This gives

$$\begin{aligned}\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ + \int_0^t \sigma_{s-}(\Upsilon_{s-}^{-1} X L_{s-} + \Upsilon_{s-}^{-1*} L_{s-}^* X + (\Upsilon_{s-}^* \Upsilon_{s-})^{-1} \Xi_{s-}\, L_{s-}^* X L_{s-})\, dY_s\end{aligned}$$

and similarly, we obtain the classical equivalent of Corollary 3.8

$$\begin{aligned}\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ + \int_0^t \sigma_{s-}(\Xi_{s-}\, L_{s-}^* X L_{s-} + \Upsilon_{s-}^* X L_{s-} + \Upsilon_{s-} L_{s-}^* X - \Xi_{s-}^{-1}(1 - \Upsilon_{s-}^* \Upsilon_{s-}) X) \times \\ (dY_s - \tilde{\Xi}_s^{-1}(1 - \tilde{\Upsilon}_s^* \tilde{\Upsilon}_s)\, ds)\end{aligned}$$

where now the stochastic integrals are classical Itô integrals with respect to the semimartingale Y_t [Pro04] (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 [Pro04] 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, pure diffusion case). *Suppose that $\Xi_t = 0$ and Υ_t has a bounded inverse for all $t \in [0, T]$. Then $\pi_t(\cdot)$ satisfies with respect to the semimartingale observations Y_t the Itô equation*

$$\begin{aligned}\pi_t(X) = \mathbb{P}(X) + \int_0^t \pi_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ + \int_0^t \left\{\pi_s(\Upsilon_s^{-1} X L_s + \Upsilon_s^{-1*} L_s^* X) - \pi_s(\Upsilon_s^{-1} L_s + \Upsilon_s^{-1*} L_s^*)\, \pi_s(X)\right\} d\overline{Z}_s\end{aligned}$$

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

$$\rho_t = \rho_0 + \int_0^t \Big\{ -i[\tilde{H}_s, \rho_s] + \tilde{L}_s \rho_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \rho_s + \rho_s \tilde{L}_s^* \tilde{L}_s) \Big\}\, ds + \int_0^t \Big\{ \tilde{\Upsilon}_s^{-1} \tilde{L}_s \rho_s + \tilde{\Upsilon}_s^{-1*} \rho_s \tilde{L}_s^* - \mathrm{Tr}[(\tilde{\Upsilon}_s^{-1} \tilde{L}_s + \tilde{\Upsilon}_s^{-1*} \tilde{L}_s^*)\rho_s]\, \rho_s \Big\}\, d\overline{Z}_s.$$

Proposition 4.3 (nonlinear filtering equation, pure jump case)**.** *Suppose that Ξ_t has a bounded inverse for all $t \in [0,T]$. Then $\pi_t(\cdot)$ satisfies with respect to the semimartingale observations Y_t the Itô equation*

$$\pi_t(X) = \mathbb{P}(X) + \int_0^t \pi_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds + \int_0^t \left\{ \frac{\pi_{s-}((\Upsilon_{s-} + \Xi_{s-} L_{s-})^* X (\Upsilon_{s-} + \Xi_{s-} L_{s-}))}{\pi_{s-}((\Upsilon_{s-} + \Xi_{s-} L_{s-})^* (\Upsilon_{s-} + \Xi_{s-} L_{s-}))} - \pi_{s-}(X) \right\} \tilde{\Xi}_{s-}^{-1}\, d\overline{Z}_s$$

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

$$\rho_t = \rho_0 + \int_0^t \Big\{ -i[\tilde{H}_s, \rho_s] + \tilde{L}_s \rho_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \rho_s + \rho_s \tilde{L}_s^* \tilde{L}_s) \Big\}\, ds + \int_0^t \left\{ \frac{(\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \rho_{s-} (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^*}{\mathrm{Tr}[(\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \rho_{s-} (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^*]} - \rho_{s-} \right\} \tilde{\Xi}_{s-}^{-1}\, d\overline{Z}_s.$$

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 than in the continuous case. 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

$$[\tilde{Y}, \tilde{Y}]_t = \tilde{Y}_t^2 - 2\int_0^t \tilde{Y}_{s-}\, d\tilde{Y}_s = \tilde{Y}_t$$

by applying the quantum Itô rules to (3.2) and using the spectral theorem. The quadratic variation $[\tilde{Y}, \tilde{Y}]_t$ 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 [Pro04], 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 linear filtering equation for the pure jump case in terms of $\tilde{Y}_t$. This gives

$$\sigma_t(X) = \mathbb{P}(X) + \int_0^t \sigma_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ + \int_0^t \sigma_{s-}((\Upsilon_{s-} + \Xi_{s-} L_{s-})^* X (\Upsilon_{s-} + \Xi_{s-} L_{s-}) - X)\, (d\tilde{Y}_s - \tilde{\Xi}_s^{-2}\, ds)$$

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 [Pro04]

$$f(V_t) - f(V_0) = \int_0^t f'(V_{s-})\, dV_s + \sum_{0<s\le t} \{f(V_s) - f(V_{s-}) - f'(V_{s-})\,(V_s - V_{s-})\}$$

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 can thus express the correction term in the change of variables formula as

$$\sum_{0<s\le t} \big\{\sigma_s(I)^{-1} - \sigma_{s-}(I)^{-1} + \sigma_{s-}(I)^{-2}\,(\sigma_s(I) - \sigma_{s-}(I))\big\} = \\ \int_0^t (\pi_{s-}(B_{s-}^* B_{s-}) + \pi_{s-}(B_{s-}^* B_{s-})^{-1} - 2)\, \sigma_{s-}(I)^{-1}\, d\tilde{Y}_s$$

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

$$\pi_t(X) = \mathbb{P}(X) + \int_0^t \pi_s(i[H_s, X] + L_s^* X L_s - \tfrac{1}{2}\{L_s^* L_s X + X L_s^* L_s\})\, ds \\ + \int_0^t \left\{ \frac{\pi_{s-}(B_{s-}^* X B_{s-})}{\pi_{s-}(B_{s-}^* B_{s-})} - \pi_{s-}(X) \right\} (d\tilde{Y}_s - \tilde{\Xi}_s^{-2}\, \pi_s(B_s^* B_s)\, ds).$$

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 linear 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

$$\tau_t = \rho_0 + \int_0^t \left\{ -i[\tilde{H}_s, \tau_s] + \tilde{L}_s \tau_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \tau_s + \tau_s \tilde{L}_s^* \tilde{L}_s) \right\} ds + \int_0^t \left\{ \tilde{\Upsilon}_s^{-1} \tilde{L}_s \tau_s + \tilde{\Upsilon}_s^{-1*} \tau_s \tilde{L}_s^* \right\} dY_s,$$

in the pure diffusion case, and in the pure jump case τ_t must satisfy

$$\tau_t = \rho_0 + \int_0^t \left\{ -i[\tilde{H}_s, \tau_s] + \tilde{L}_s \tau_s \tilde{L}_s^* - \tfrac{1}{2}(\tilde{L}_s^* \tilde{L}_s \tau_s + \tau_s \tilde{L}_s^* \tilde{L}_s) \right\} ds + \int_0^t \left\{ (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \tau_{s-} (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^* - \tau_{s-} \right\} \times \tilde{\Xi}_{s-}^{-1} (dY_s - \tilde{\Xi}_s^{-1} (1 - \tilde{\Upsilon}_s^* \tilde{\Upsilon}_s)\, ds).$$

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 [Pro04]. 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 [Pro04] 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 the process $\bar{\Pi}_t$ 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

$$\bar{\Pi}_t = 1 + \int_0^t \left\{ \mathrm{Tr}[(\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-})^* (\tilde{\Upsilon}_{s-} + \tilde{\Xi}_{s-} \tilde{L}_{s-}) \bar{\rho}_{s-}] - 1 \right\} \bar{\Pi}_{s-} \times \tilde{\Xi}_{s-}^{-1} (dY_s - \tilde{\Xi}_s^{-1} (1 - \tilde{\Upsilon}_s^* \tilde{\Upsilon}_s)\, ds)$$

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. $\rho_t = \bar{\rho}_t$. □

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 (and hence also the same von Neumann 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 and there is no innovations problem.

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 [LS01] 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 [Pro04]. 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 control 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 control strategy on the interval $[0,t]$, and similarly by $\mu_{[t}$ the control 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 5.4. 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 5.5 (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 [Won68] and for finite-state Markov processes by Segall [Seg77]. 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 5.6 (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. [Kus71]. Taking the limit as the time step goes to zero gives the infinitesimal form of this recursion, i.e. Bellman's functional equation

$$-\frac{\partial V}{\partial t}(t,\theta) = \min_{u\in\mathscr{B}_t}\left\{\mathscr{L}(u)V(t,\theta) + \mathrm{Tr}[\theta\,\tilde C_t^u]\right\},\qquad t\in[0,T],\ \theta\in\mathcal{S}(\mathcal{B})$$

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

$$u_t^\mu(\theta) = \operatorname*{argmin}_{u\in\mathscr{B}_t}\left\{\mathscr{L}(u)V(t,\theta) + \mathrm{Tr}[\theta\,\tilde C_t^u]\right\},\qquad t\in[0,T],\ \theta\in\mathcal{S}(\mathcal{B}).$$

In these equations $\mathscr{L}(u)$ 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 [Pro04] 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 5.7 (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*
$$\frac{\partial V}{\partial t}(t,\theta)+\mathscr{L}(u^\mu_t(\theta))V(t,\theta)+\mathrm{Tr}\left[\theta\,\tilde{C}_t^{u^\mu_t(\theta)}\right]=0.$$

(3) *For all $t\in[0,T]$, $u\in\mathscr{B}_t$ and $\theta\in\mathcal{S}(\mathcal{B})$, the function V satisfies*
$$\frac{\partial V}{\partial t}(t,\theta)+\mathscr{L}(u)V(t,\theta)+\mathrm{Tr}[\theta\,\tilde{C}^u_t]\ge 0.$$

(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 $\mathscr{C}$, 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 $V(t,\rho^\mu_t)$ 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 $\theta=\rho^\mu_t$ and the terminal condition (4) into Definition 5.6. This gives
$$W[\mu](t)=\mathbb{E}_{\mathbf{P}^\mu}\left(V(T,\rho^\mu_T)-\int_t^T\left\{\frac{\partial V}{\partial s}(s,\rho^\mu_s)+[\mathscr{L}(u^\mu_s(\rho^\mu_s))V](s,\rho^\mu_s)\right\}ds\,\middle|\,\Sigma^\mu_t\right).$$
The purpose of condition (1) is to ensure that we can apply the Itô change of variables formula to $V(t,\rho^\mu_t)$. This gives
$$W[\mu](t)=\mathbb{E}_{\mathbf{P}^\mu}\left(V(t,\rho^\mu_t)+\int_t^T G^{\mu,V}_{s-}(\rho^\mu_{s-})\,d\overline{Z}_s\,\middle|\,\Sigma^\mu_t\right).$$
But by a fundamental property of stochastic integrals [Pro04], the stochastic integral of the bounded process $G^{\mu,V}_{t-}(\rho^\mu_{t-})$ against the square-integrable martingale $\overline{Z}_t$ is itself a martingale. Hence the conditional expectation of the second term vanishes, and as ρ^μ_t is Σ^μ_t-measurable we obtain immediately
$$W[\mu](t)=V(t,\rho^\mu_t).$$
In particular, since $\rho^\mu_0=\rho$, we find that $V(0,\rho)$ equals the expected total cost
$$V(0,\rho)=\mathbb{P}(J[\mu]).\tag{5.1}$$
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 can verify that
$$V(t,\rho^\psi_t)=\mathbb{E}_{\mathbf{P}^\psi}\left(V(T,\rho^\psi_T)-\int_t^T\left\{\frac{\partial V}{\partial s}(s,\rho^\psi_s)+[\mathscr{L}(\psi_s)V](s,\rho^\psi_s)\right\}ds\,\middle|\,\Sigma^\psi_t\right).$$
Using condition (4) and the inequality (3) with $\theta=\rho^\psi_t$ and $u=\psi_t$, we obtain
$$V(t,\rho^\psi_t)\le\mathbb{E}_{\mathbf{P}^\psi}\left(\pi^\psi_T(C_T)+\int_t^T\pi^\psi_s(C^\psi_s)\,ds\,\middle|\,\Sigma^\psi_t\right)=W[\psi](t).$$

But $\rho_0^\psi = \rho_0^\mu = \rho$, so together with Equation (5.1) 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 5.8. 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 simplicity 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. [ØS05]).

Another option is to rewrite the cost directly in terms of the linear 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 linear filtering equation. This is particularly useful, for example, in treating the risk-sensitive control problem [Jam05].

Finally, a class of interesting quantum control problems can be formulated using the theory of quantum stopping times [PS87]; this gives rise to optimal stopping problems and impulse control problems in the quantum context. Such control problems are explored in [Van06] using a similar argument to the one used above.

References

[AAS+02] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi, *Adaptive homodyne measurement of optical phase*, Phys. Rev. Lett. **89** (2002), 133602.

[Bel83] V. P. Belavkin, *Theory of the control of observable quantum systems*, Autom. remote control **44** (1983), 178–188.

[Bel88] ———, *Nondemolition stochastic calculus in Fock space and nonlinear filtering and control in quantum systems*, Proceedings XXIV Karpacz winter school (R. Guelerak and W. Karwowski, eds.), Stochastic methods in mathematics and physics, World Scientific, Singapore, 1988, pp. 310–324.

[Bel92] ———, *Quantum stochastic calculus and quantum nonlinear filtering*, J. Multivar. Anal. **42** (1992), 171–201.

[BGM04] L. Bouten, M. I. Guţă, and H. Maassen, *Stochastic Schrödinger equations*, J. Phys. A **37** (2004), 3189–3209.

[Bia95] P. Biane, *Calcul stochastique noncommutatif*, Lectures on Probability Theory, Saint-Flour, 1993 (P. Bernard, ed.), Lecture Notes in Mathematics, vol. 1608, Springer, Berlin, 1995, pp. 1–96.

[Bou04] L. Bouten, *Filtering and control in quantum optics*, Ph.D. thesis, Radboud Universiteit Nijmegen, arXiv:quant-ph/0410080, 2004.

[BV05] L. Bouten and R. Van Handel, *Quantum filtering: a reference probability approach*, preprint, arXiv:math-ph/0508006, California Institute of Technology, 2005.

[BV06] ———, *Controlled quantum stochastic processes*, 2006, in preparation.

[GLW01] D. Goswami, J. M. Lindsay, and S. J. Wills, *A stochastic Stinespring theorem*, Math. Ann. **319** (2001), 647–673.

[GSM04] J. M. Geremia, J. K. Stockton, and H. Mabuchi, *Real-time quantum feedback control of atomic spin-squeezing*, Science **304** (2004), 270–273.

[Hol91] A. S. Holevo, *Quantum stochastic calculus*, J. Soviet Math. **56** (1991), 2609–2624, Translation of Itogi Nauki i Tekhniki, ser. sovr. prob. mat. 36, 3–28, 1990.

[Hol96] ———, *Exponential formulae in quantum stochastic calculus*, Proc. Roy. Soc. Edinburgh Sect. A **126** (1996), 375–389.

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

[Jam05] M. R. James, *A quantum Langevin formulation of risk-sensitive optimal control*, J. Opt. B Quantum Semiclass. Opt. **7** (2005), S198–S207.

[Kus71] H. Kushner, *Introduction to stochastic control*, Holt, Rinehart and Winston, Inc, New York, 1971.

[LS01] R. S. Liptser and A. N. Shiryaev, *Statistics of random processes I: General theory*, second ed., Springer-Verlag, 2001.

[Mey93] P.-A. Meyer, *Quantum probability for probabilists*, Springer, Berlin, 1993.

[MV05] M. Mirrahimi and R. Van Handel, *Stabilizing feedback controls for quantum systems*, preprint, arXiv:quant-ph/0510066, California Institute of Technology, 2005.

[ØS05] B. Øksendal and A. Sulem, *Applied stochastic control of jump diffusions*, Universitext, Springer-Verlag, Berlin, 2005.

[Par92] K. R. Parthasarathy, *An introduction to quantum stochastic calculus*, Birkhäuser, Basel, 1992.

[Pro04] P. E. Protter, *Stochastic integration and differential equations*, second ed., Springer-Verlag, 2004.

[PS87] K. R. Parthasarathy and K. B. Sinha, *Stop times in Fock space stochastic calculus*, Probab. Theory Related Fields **75** (1987), 317–349.

[Seg77] A. Segall, *Optimal control of noisy finite-state Markov processes*, IEEE Trans. Automat. Control **AC-22** (1977), 179–186.

[Tak71] M. Takesaki, *Conditional expectations in von Neumann algebras*, J. Funct. Anal. **9** (1971), 306–321.

[Van06] R. Van Handel, *Stopping times in quantum control*, 2006, in preparation.

[VSM05] R. Van Handel, J. K. Stockton, and H. Mabuchi, *Feedback control of quantum state reduction*, IEEE Trans. Automat. Control **50** (2005), 768–780.

[Wis95] H. M. Wiseman, *Adaptive phase measurements of optical modes: Going beyond the marginal q distribution*, Phys. Rev. Lett. **75** (1995), no. 25, 4587–4590.

[Won68] W. M. Wonham, *On the separation theorem of stochastic control*, SIAM J. Control **6** (1968), 312–326.

[Zak69] M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrsch. th. verw. Geb. **11** (1969), 230–243.

Luc Bouten, Physical Measurement and Control 266-33, California Institute of Technology, Pasadena, CA 91125, USA

E-mail address: `bouten@its.caltech.edu`

Ramon van Handel, Physical Measurement and Control 266-33, California Institute of Technology, Pasadena, CA 91125, USA

E-mail address: `ramon@its.caltech.edu`