Risk-Sensitive Optimal Control of Quantum Systems

M.R. James
Department of Engineering,
Australian National University,
Canberra, ACT 0200, Australia.
Matthew.James@anu.edu.au
quant-ph/0302136
To appear, Phys. Rev. A
(Dated: November 3, 2003)

The importance of feedback control is being increasingly appreciated in quantum physics and applications. This paper describes the use of optimal control methods in the design of quantum feedback control systems, and in particular the paper formulates and solves a risk-sensitive optimal control problem. The resulting risk-sensitive optimal control is given in terms of a new unnormalized conditional state, whose dynamics include the cost function used to specify the performance objective. The risk-sensitive conditional dynamic equation describes the evolution of our *knowledge* of the quantum system tempered by our *purpose* for the controlled quantum system. Robustness properties of risk-sensitive controllers are discussed, and an example is provided.

PACS numbers: 03.65.Ta,02.30.Yy

I. INTRODUCTION

Optimal control theory provides a systematic approach to control system design that is widely used. A cost function is formulated by the designer that encodes the desired performance of the system as its minimum, and then the cost is minimized to obtain the desired controller. Perhaps the most famous example is Kalman's *linear quadratic Gaussian (LQG)* regulator problem, where the cost criterion is an average of an integral,

$$J^{LQG} = \mathbf{E}[\sum_{k=0}^{M-1}(x_k' P x_k + u_k' Q u_k + x_M' P_M x_M)] \tag{1}$$

where x_k and u_k are respectively state and control variables (vectors), and P, Q and P_M are weighting matrices [28]. The cost criterion (1) is an example of what is sometimes called a *risk-neutral* criterion. The state (or phase space) variable x_k is part of the model of the classical physical system being controlled. In general, the controller has only partial access to state information, with measurements corrupted by noise. Kalman's optimal LQG feedback controller is an explicit function of the conditional state and covariance. It is dynamic, since the conditional state and covariance evolve in time via the Kalman Filter (see, e.g., [1], [22]), and the Kalman filter does not involve the cost function in any way; it gives the optimal mean square state estimate independently of any control objective. Interestingly, the function giving the optimal feedback control is the same as for an analogous problem with full state information, viz. multiplication by a gain matrix determined by solving a Riccati equation. Kalman's optimal LQG controller is the paradigm example of the so-called *separation structure*, where the controller is decomposed into an estimation part (filtering) and a control part, as illustrated in Figure 1.

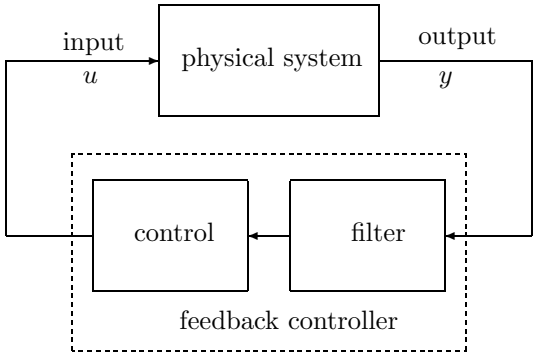

FIG. 1: Feedback controller showing the separation structure.

Over the past 20 or so years, another type of optimal control problem has generated considerable interest, viz. *linear exponential quadratic Gaussian (LEQG)* optimal control, or *risk-sensitive* optimal control, [6, 20, 25]. In this average of exponential of integral problem, the cost is of the form

$$J^{LEQG} = \mathbf{E}[\exp\Big(\sum_{k=0}^{M-1}(x_k' P x_k + u_k' Q u_k) + x_M' P_M x_M\Big)]. \tag{2}$$

In this case the optimal feedback control is an explicit function of a dynamical quantity closely related to the conditional state and covariance, but given by dynamics that include terms from the cost

function. It also has a separation structure, though in this case the filter depends on the cost function used to specify the performance objective, and is a modification of the Kalman Filter. One of the major reasons for the interest in the risk-sensitive problem is its close connections to *robust control* and minimax games, [13, 18, 21]. Robust control concerns the desire to design controllers that are robust with respect to uncertainty, such as model errors and exogenous disturbances, [19]. Robustness properties of risk-sensitive controllers are described in [15].

Risk-neutral, risk-sensitive and other stochastic control problems have been considered for problems with a finite number of states, see, e.g. [2, 8, 16, 22]. After an analysis of an example of a machine replacement problem [8], the authors concluded that for that problem the risk-neutral controller was more *aggressive* than risk-sensitive and related minimax controllers.

Suppose we wish to control a quantum physical system using real time feedback via a non-quantum feedback system (say using a digital computer) in some optimal fashion. If one were to do this using a standard cost criterion, say one analogous to Kalman's (LQG) regulator problem (risk-neutral), then one would find that the optimal control is a function of the conditional (selective) state (a density operator), as is well known, see, e.g. [4, 10, 11, 24]. The conditional state is the solution of a stochastic master equation that describes the evolution of our *knowledge* of the system. This stochastic master equation is used in two ways: (i) as the model of the quantum physical system, taking into account the effect of the measurements, and (ii) as the dynamics of the filter in the optimal controller, Figure 1.

The purpose of this paper is to consider the *risk-sensitive* optimal control of quantum physical systems. The quantum systems are modelled by stochastic master equations for the conditional state. The risk-sensitive criterion is one of a class of multiplicative cost functions. The optimal solution for this class of problems has a separation structure, Figure 1, where the filter describes the evolution of an *unnormalized* conditional state via a modified stochastic master equation that contains the cost function used to specify the performance objective. The optimal control is a function of this unnormalized conditional state. It is important to note that, in contrast to the risk-neutral case described above, the states and dynamics for the quantum physical model and the filter are not the same. Indeed, the unnormalized conditional dynamic equation used in the filter describes the evolution of our *knowledge* of the quantum system tempered by our *purpose* for the controlled quantum system. This type of extension of the conditional dynamics appears to be new to quantum physics, and may merit further investigation. We emphasize that the unnormalized conditional state is defined only in the context of the risk-sensitive and multiplicative control objectives considered here, where it is used in a specific feedback situation. Again, we emphasize that (i) the model of the quantum physical system is the standard stochastic master equation for the conditional state, and (ii) the filter is described by a modified stochastic master equation for an unnormalized conditional state; this modified equation includes terms from the cost function.

This paper is organized as follows. In section II we carefully describe the model we use for the controlled quantum system. Then in section III we summarize some relevant results for a risk-neutral optimal control problem, and make some comments on the feedback solution. Section IV contains the formulation and dynamic programming solution to the risk-sensitive and related multiplicative cost optimal control problems, together with a brief discussion of robustness. The ideas are illustrated by a simple example of a two-state system with feedback. Further developments, applications and examples will be given in subsequent papers.

Acknowledgement. The author wishes to thank Professor I.R. Petersen for many valuable discussions.

II. THE CONTROLLED QUANTUM SYSTEM

A. Controlled State Transfer

We consider a controlled quantum physical system with *inputs* u and *outputs* y. The inputs represent signals or actions that are applied to the system, such as voltages, forces, or light pulses. The outputs are signals that result from repeated measurements of observable quantities, such as position, spin, etc. We will assume, for simplicity, that the measurements are *discrete* valued. It is sometimes useful to denote the range of input and output values by $\mathbf{U}$ and $\mathbf{Y}$ respectively.

The *state* of the quantum system is described by a density operator ω[29]. This state evolves in time as a result of a variety of factors including the underlying unitary evolution, interaction with the environment, the effect of repeated measurements, and feedback control actions. Since measurements are made, and the outcomes are used to determine control actions in a feedback context, we are inter-

ested in the *selective* or *conditional* evolution of the states. As an example [3, 5, 27], a range of conditional evolutions can be described by an Ito-type *stochastic master equations (SME)* of the form

$$d\omega = \mathcal{L}[\omega]dt + \mathcal{M}[\omega]dW \tag{3}$$

for suitable (super) operators $\mathcal{L}$ and $\mathcal{M}$ (which may depend on the control u). Here, dW represents an Ito-type Brownian motion (Wiener process) increment, called an innovation, related to the measured output value y by

$$dy = \mathrm{tr}\{\mathcal{N}[\omega]\}dt + dW \tag{4}$$

for a suitable (super) operator $\mathcal{N}$. If we denote by ρ the expected value of ω with respect to W (or y), we obtain the *master equation*, frequently encountered in the analysis of open systems:

$$\dot{\rho} = \mathcal{L}[\rho]. \tag{5}$$

It is conceptually and technically simpler to work in *discrete time*, and so we will do so in this paper. Effectively, we will be using a model for sampled-data feedback control of quantum systems. In this model, measurements are made and control actions are applied at discrete time instants t_k[30], called sample times. Continuous time models are of considerable importance, and will be considered elsewhere.

The discrete time model we use for the quantum system is defined in terms of a (super) operator $\Gamma(u, y)$ [31] that depends on the control input u and the output measurement y. The idea is that if the quantum system is in state ω_k at time k, and at this time the control value u_k is applied, a measurement outcome y_{k+1} will be recorded, and the system will transfer to a new state ω_{k+1}. The probability of y_{k+1} is $p(y_{k+1}|u_k, \omega_k)$, where

$$p(y|u, \omega) = \langle \Gamma(u, y)\omega, I\rangle. \tag{6}$$

Here, we have used the notation

$$\langle \omega, B\rangle = \mathrm{tr}[B\omega] \tag{7}$$

to specify the (expected) value of an observable B when the system is in state ω. The operator $\Gamma(u, y)$ is assumed to be normalized, i.e.

$$\sum_{y\in\mathbf{Y}} \langle \Gamma(u, y)\omega, I\rangle = \langle \omega, I\rangle = 1$$

so that $p(y|u, \omega)$ is a *probability distribution*, since it satisfies $\sum_y p(y|u, \omega) = 1$.

Selective or conditional evolution means that the new state ω_{k+1} depends on the value of the measurement y_{k+1}, and we write this dependance as follows:

$$\omega_{k+1} = \Lambda_\Gamma(u_k, y_{k+1})\omega_k, \tag{8}$$

where

$$\Lambda_\Gamma(u, y)\omega = \frac{\Gamma(u, y)\omega}{p(y|u, \omega)}. \tag{9}$$

Equation (8) is a discrete time *stochastic master equation (SME)*, and can be viewed, e.g., as the result of integrating an equation of the form (3) over one time step (after substituting for dW in terms of dy).

We denote the average of the conditional state ω_k with respect to the measurements by ρ_k. If u_k is a deterministic (non-random) input signal, then ρ_k satisfies the *master equation*

$$\rho_{k+1} = \sum_{y\in\mathbf{Y}} \Gamma(u_k, y)\rho_k. \tag{10}$$

Equation (8) constitutes our model of the quantum system. For further information on this framework of operator valued measures and quantum operations, see [4, 9, 17, 23]. We now give some examples.

Example II.1 We define the controlled transfer $\Gamma(u, y)$ by interleaving open system dynamics and imperfect orthogonal measurements. The open system dynamics are modelled by a quantum operation

$$\mathcal{E}^u\omega = \sum_b E_b^u \omega E_b^{u\,\dagger} \tag{11}$$

where the controlled operators E_b^u satisfy $\sum_b E_b^{u\,\dagger} E_b^u = I$ for all inputs u. Closed systems are described by the unitary evolution operation $\mathcal{E}^u\omega = T^u\omega T^{u\,\dagger}$, where for each input value u, T^u is a unitary operator.

The imperfect measurements are modelled as follows. Let A be a self-adjoint operator with discrete nondegenerate spectrum $\mathrm{spec}(A)$. For $a \in \mathrm{spec}(A)$ an eigenvalue of A let $|a\rangle$ denote the normalized eigenvector, and let $P_a = |a\rangle\langle a|$ denote the projection onto the eigenspace of A ($P_a|\psi\rangle = \langle a|\psi\rangle|a\rangle$). Perfect measurements would correspond to $y = a$; however, to reflect the presence of measurement noise in applications we will assume that when a measurement occurs on the quantum system, the values a and associated projections occur in the usual (perfect) way, but that knowledge of the outcomes is corrupted by sensor noise so that the controller (or any observing device or person) measures a value y. The measurement y is a random variable, related to the outcomes a via probability kernels $q(y|a)$, the probability of y given that a occurred. The kernels have the property that $\sum_y q(y|a) = 1$ for all a. In the case of perfect measurements, $q(y|a) = 1$ if $y = a$, and $q(y|a) = 0$ if $y \neq a$.

The operator $\Gamma(u,y)$ is given by

$$\Gamma(u,y)\omega = \sum_{a,b} q(y|a) P_a E_b^u \omega E_b^{u\dagger} P_a \qquad (12)$$

and the adjoint is given by

$$\Gamma^\dagger(u,y)B = \sum_{a,b} q(y|a) E_b^{u\dagger} P_a B P_a E_b^u \qquad (13)$$

where B is an observable. The expressions in this example can be derived using standard techniques of quantum operations and discrete time filtering based on Bayes' Rule (see, e.g., [17, Chapter 2.2], [23, Chapter 8], [9], [4], [22, Chapter 6], [1, Chapter 7]). □

Example II.2 (Two-state system.) We now describe a specific instance of Example II.1, viz. a two-state system and measurement device, where it is desired to use feedback control to put the system into a given state. The example is inspired by a simple quantum feedback example [26, Section 1.3] and an example in stochastic control concerning a machine replacement problem [8, 16].

In [26, Section 1.3], a particle beam is passed through a Stern-Gerlach device, which results in one beam of particles in the up state, and one beam in the down state. The beam of particles in the up state is subsequently left alone, while the beam in the down state is subject to a further device which will result in a change of spin direction from down to up. The final outcome of this feedback arrangement is that all particles are in the up state. Analogous feedback configurations can be constructed using other physical systems, e.g. light and polarization measurement.

In what follows we extend the general features of this example to accommodate repeated noisy measurements. Physically, the noisy measurements might arise from imperfectly separated beams, where a proportion of each beam contaminates the other, and/or from interference or noise affecting sensors. The example was chosen because the risk-neutral and risk-sensitive problems can be solved explicitly. Hence the example provides a concrete illustration of some ideas concerning quantum feedback control. More substantial examples and applications will be considered elsewhere.

The pure states of the system are of the form

$$|\psi\rangle = c_{-1}|-1\rangle + c_1|1\rangle \;\equiv\; \begin{pmatrix} c_{-1} \\ c_1 \end{pmatrix}.$$

The states $|-1\rangle$ and $|1\rangle$ are eigenstates of the observable

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad (14)$$

corresponding to ideal measurement values $a = -1$ and $a = 1$. It is desired to put the system into the state

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ or } |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

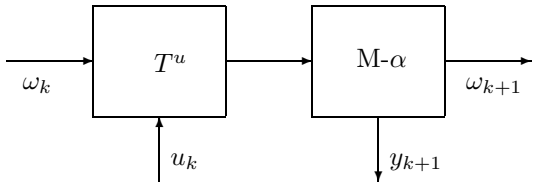

FIG. 2: Two-state system example showing the controlled unitary operator T^u and the noisy measurement device M-α with error probability α.

We define a controlled transfer operator $\Gamma(u,y)$ as the following physical process, Figure 2. First apply a unitary transformation T^u, where the control value $u = 0$ means do nothing, while $u = 1$ means to flip the states (quantum not gate), i.e.

$$T^u = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } u = 0 \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{if } u = 1. \end{cases}$$

We then make an imperfect measurement corresponding to the observable A. We model this by an ideal device (e.g. Stern-Gerlach, beam splitter) with projection operators

$$P_{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; P_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

followed by a memoryless channel with error probability kernels

$$\begin{aligned} q(-1|-1) &= 1-\alpha \\ q(-1|1) &= \alpha \\ q(1|-1) &= \alpha \\ q(1|1) &= 1-\alpha \end{aligned}$$

where $0 \le \alpha \le 1$ is the probability of a measurement error (cf. [23, Figure 8.1]).

The controlled transfer operator is therefore (from (12))

$$\begin{aligned} \Gamma(u,y)\omega = {} & q(y|-1) P_{-1} T^u \omega T^{u\,\dagger} P_{-1} \\ & + q(y|1) P_1 T^u \omega T^{u\,\dagger} P_1. \end{aligned}$$

In this example, the control u can take the values 0 or 1, and output y has values 0 or 1 ($\mathbf{U} = \{0,1\}$), $\mathbf{Y} = \{0,1\}$).

If we write a general density matrix as

$$\omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12}^* & \omega_{22} \end{pmatrix}, \tag{15}$$

then the controlled operators $\Gamma(u, y)$ are given explicitly by

$$\begin{aligned} \Gamma(0,-1)\omega &= \begin{pmatrix} (1-\alpha)\omega_{11} & 0 \\ 0 & \alpha\omega_{22} \end{pmatrix} \\ \Gamma(0,1)\omega &= \begin{pmatrix} \alpha\omega_{11} & 0 \\ 0 & (1-\alpha)\omega_{22} \end{pmatrix} \\ \Gamma(1,-1)\omega &= \begin{pmatrix} (1-\alpha)\omega_{22} & 0 \\ 0 & \alpha\omega_{11} \end{pmatrix} \\ \Gamma(1,1)\omega &= \begin{pmatrix} \alpha\omega_{22} & 0 \\ 0 & (1-\alpha)\omega_{11} \end{pmatrix} \end{aligned}$$

This example is continued in stages in the remainder of the paper (Examples II.3, III.1, III.3, IV.7). □

B. Feedback Control

In the above description of the quantum system (8), we have not described how the controls u_k are determined by the measurements y_k via a feedback controller K. We now do this.

Feedback controllers should be *causal*, i.e., the current control value u_k cannot depend on future values of the measurements $y_{k+1}, y_{k+2}, \ldots$. On a time interval $0 \le k \le M-1$ this is expressed as follows:

$$K = \{K_0, K_1, \ldots, K_{M-1}\}$$

where

$$\begin{aligned} u_0 &= K_0 \\ u_1 &= K_1(y_1) \\ u_2 &= K_2(y_1, y_2) \\ &\text{etc.} \end{aligned}$$

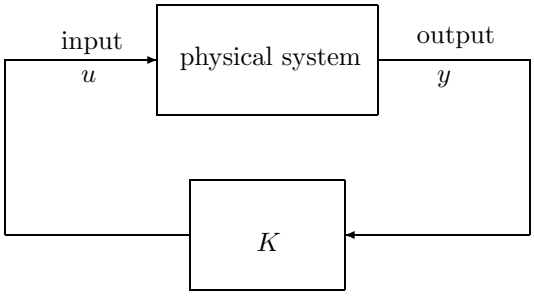

FIG. 3: Feedback control of quantum system showing a general feedback controller K.

To simplify notation, we often write sequences $u_{k_1}, u_{k_1+1}, \ldots, u_{k_2}$ as u_{k_1,k_2}. Then we can write $u_k = K_k(y_{1,k})$. A controller K can be restricted to subintervals $k \le j \le M$ by fixing (or omitting) the first arguments in the obvious way. We denote by $\mathcal{K}$ the class of all such feedback controllers.

A feedback controller K in closed loop with the quantum system, Figure 3, operates as follows. The given initial state ω_0 and controller K are sufficient to define random sequences of states $\omega_{0,M}$, inputs $u_{0,M-1}$ and outputs $y_{1,M}$ over a given time interval $0 \le k \le M$ iteratively as follows. The control value u_0 is determined by K_0 (no observations are involved yet), and it is applied to the quantum system, which responds by selecting y_1 at random according to the distribution $p(y_1|u_0, \omega_0)$. This then determines the next state ω_1 via (8). Next u_1 is given by $K_1(y_1)$, and applied to the system. This process is repeated until the final time.

The controller K therefore determines controlled stochastic processes ω_k, u_k and y_k on the interval $0 \le k \le M$. Expectation with respect to the associated probability distribution is denoted $\mathbf{E}^K_{\omega_0,0}$. The state sequence ω_k is a *controlled Markov process*.

One way a controller K can be constructed is using a function

$$u_k = \mathbf{u}(\omega_k, k)$$

where ω_k is given by (8) with initial state ω_0. This controller is denoted $K^{\mathbf{u}}_{\omega_0}$. The SME equation (8) forms part of this controller, viz. its dynamics, and must be implemented with suitable technology (e.g. digital computer). Controllers of this type are said to have a *separation structure*, where the controller can be decomposed into an estimation part (i.e. filtering via (8)) and a control part (i.e. the function $\mathbf{u}$). We will see in section III that the optimal risk-neutral controller is of this form (Figure 6). In section IV, the optimal risk-sensitive controller also has a separation structure, but the filter used is different (Figure 7). The separation structure arises naturally from the dynamic programming techniques, as we shall see.

Example II.3 (Two-state system with feedback, Example II.2 continued.) We consider a particular feedback controller $\bar{K}$ for a time horizon $M = 2$ defined by

$$u_0 = \bar{K}_0 = 0, \quad u_1 = \bar{K}_1(y_1) = \begin{cases} 0 & \text{if } y_1 = 1 \\ 1 & \text{if } y_1 = -1. \end{cases} \tag{16}$$

We apply $\bar{K}$ to the system with initial pure state

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}|-1\rangle + \frac{1}{\sqrt{2}}|1\rangle, \text{ or } \omega_0 = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{17}$$

ω_0	p_1	ω_1	p_2	ω_2
ω_0			$2\alpha(1-\alpha)$	$\omega_2^{(0,-1),(1,-1)}$
	$\frac{1}{2}$	$\omega_1^{(0,-1)}$	$\alpha^2+(1-\alpha)^2$	$\omega_2^{(0,-1),(1,1)}$
	$\frac{1}{2}$	$\omega_1^{(0,1)}$	$2\alpha(1-\alpha)$	$\omega_2^{(0,1),(0,-1)}$
			$\alpha^2+(1-\alpha)^2$	$\omega_2^{(0,1),(0,1)}$
$\rho_0=\omega_0$		ρ_1		ρ_2

TABLE I: State evolution under the controller $\bar{K}$.

The result is shown in Table II.3, which displays the resulting conditional states

$$\begin{aligned}\omega_1^{(u_0,y_1)} &= \Lambda_\Gamma(u_0,y_1)\omega_0,\\ \omega_2^{(u_0,y_1),(u_1,y_2)} &= \Lambda_\Gamma(u_1,y_2)\omega_1^{(u_0,y_1)}\end{aligned}$$

and the associated probabilities. Explicitly, the terms shown in Table II.3 are:

$$p_1 = p(y_1|u_0,\omega_0), \quad p_2 = p(y_2|u_1,\omega_1)$$

$$\omega_1^{(0,-1)} = \begin{pmatrix}(1-\alpha) & 0\\ 0 & \alpha\end{pmatrix}, \quad \omega_1^{(0,1)} = \begin{pmatrix}\alpha & 0\\ 0 & (1-\alpha)\end{pmatrix}$$

$$\omega_2^{(0,-1),(1,-1)} = \omega_2^{(0,1),(0,-1)} = \tfrac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

$$\begin{aligned}\omega_2^{(0,-1),(1,1)} &= \omega_2^{(0,1),(0,1)}\\ &= \frac{1}{\alpha^2+(1-\alpha)^2}\begin{pmatrix}\alpha^2 & 0\\ 0 & (1-\alpha)^2\end{pmatrix}.\end{aligned}$$

Also shown are the non-selective states:

$$\rho_0 = \omega_0$$

$$\begin{aligned}\rho_1 &= p(-1|u_0,\omega_0)\omega_1^{(0,-1)} + p(1|u_0,\omega_0)\omega_1^{(0,1)}\\ &= \tfrac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\end{aligned}$$

$$\begin{aligned}\rho_2 = {} & p(-1|1,\omega_1^{(0,-1)})\omega_2^{(0,-1),(1,-1)}\\ & +p(1|1,\omega_1^{(0,-1)})\omega_2^{(0,-1),(1,1)}\\ & +p(-1|0,\omega_1^{(0,1)})\omega_2^{(0,1),(0,-1)}\\ & +p(1|0,\omega_1^{(0,1)})\omega_2^{(0,1),(0,1)}\\ = {} & \tfrac{1}{2}\begin{pmatrix}\alpha^2+\alpha(1-\alpha) & 0\\ 0 & \alpha(1-\alpha)+(1-\alpha)^2\end{pmatrix}.\end{aligned} \tag{18}$$

At time $k=0$ the control $u=0$ is applied. If $y_1=-1$ is observed, as a result of the imperfect measurement, the system moves to the state $\omega_1^{(0,-1)}$. Since $y_1=-1$, the controller $\bar{K}$ (16) gives $u_1=1$. This results in the states $\omega_2^{(0,-1),(1,-1)}$ or $\omega_2^{(0,-1),(1,1)}$, depending on the outcome of the second measurement y_2. If, on the other hand, $y_1=1$ is observed, the system moves to the state $\omega_1^{(0,1)}$. Since $y_1=1$, the controller $\bar{K}$ (16) gives $u_1=0$, and hence $\omega_2^{(0,1),(0,-1)}$ or $\omega_2^{(0,1),(0,1)}$, again depending on the outcome of the second measurement y_2. This is illustrated in Figure 4.

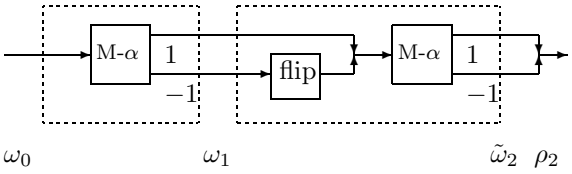

FIG. 4: Physical realization of the two stages of the two-state system with feedback using controller $\bar{K}$. Due to the merging of the beams in the second stage, we have the intermediate state $\tilde{\omega}_2 = \frac{1}{2}\omega_2^{(0,-1),(1,-1)} + \frac{1}{2}\omega_2^{(0,1),(0,-1)}$ if $y_2=-1$ (with probability $2\alpha(1-\alpha)$), or $\tilde{\omega}_2 = \frac{1}{2}\omega_2^{(0,-1),(1,1)} + \frac{1}{2}\omega_2^{(0,1),(0,1)}$ if $y_2=1$ (with probability $\alpha^2+(1-\alpha)^2$).

These results are consistent with [26, Section 1.3]. Indeed, when $\alpha=0$ (perfect measurements), the feedback system terminates in the desired pure state $\rho_2 = |1\rangle\langle 1|$. The role of feedback control is clearly demonstrated here. With imperfect measurements, $0<\alpha<1$, the system terminates in the mixed state ρ_2 given by (18), with the degree of mixing (indicating the expected degradation in performance) depending on the measurement error probability parameter α:

$$\begin{aligned}\mathrm{tr}\rho_2^2 &= (\alpha^2+\alpha(1-\alpha))^2 + (\alpha(1-\alpha)+(1-\alpha)^2)^2\\ &< 1 \ \text{ if } 0<\alpha<1\\ &= 1 \text{ if } \alpha = 0.\end{aligned}$$

□

III. RISK-NEUTRAL CONTROL

In this section we summarize dynamic programming results for a well-known type of finite time horizon optimal control problem, [4, 22]. The optimal control problem discussed here can be considered to be a prototype problem illustrating measurement feedback in the quantum context. The dynamic programming methods used in this paper for solving the optimal control problems are standard, and the reader is referred to the literature for further information, see, e.g. [1, 7, 22].

We define a *cost function* to be a non-negative observable $L(u)$ that can depend on the control u.

The cost function encodes the designer's control objective. We also use a non-negative observable N to define a cost for the final state.

Example III.1 (Two-state system with feedback, Example II.3 continued.) To set up the cost function $L(u)$ to reflect our objective of regulating the system to the desired pure state $|1\rangle$, we define

$$X = \frac{1}{2}(A - 1.I) = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

where A is the observable corresponding to the projective measurement (14). We note that the expected value of X^2 is

$$\begin{aligned} \langle 1|X^2|1\rangle &= \mathrm{tr}[X^2|1\rangle\langle 1|] &= 0 \\ \langle -1|X^2|-1\rangle &= \mathrm{tr}[X^2|-1\rangle\langle -1|] &= 1 \end{aligned}$$

which gives zero cost to the desired state, and nonzero cost to the undesired state. We shall also introduce a cost of control action, as follows:

$$c(u) = \begin{cases} 0 & \text{if } u = 0 \\ p & \text{if } u = 1 \end{cases}$$

where $p > 0$. This gives zero cost for doing nothing, and a nonzero cost for the flip operation. Thus we define the cost function to be

$$L(u) = X^2 + c(u)I \tag{19}$$

and the cost for the final state is defined to be

$$N = X^2.$$

This modifies our earlier objective of putting the system into the desired state by including a penalty for control action. □

Let $M > 0$ be a positive integer indicating a finite time interval $k = 0, \ldots, M$. Given a sequence of control values $u_{0,M-1} = u_0, \ldots, u_{M-1}$ and measurements $y_{1,M} = y_1, \ldots, y_M$, define the *risk-neutral cost functional*

$$J_{\omega,0}(K) = \mathbf{E}^K_{\omega,0}[\sum_{i=0}^{M-1} \langle \omega_i, L(u_i)\rangle + \langle \omega_M, N\rangle], \tag{20}$$

where ω_i, $i = 0, \ldots, M$ is the solution of the system dynamics (8) with initial state $\omega_0 = \omega$ under the action of a controller K. This is an appropriate quantum generalization of the classical LQG cost (1). The objective is to minimize this functional over all measurement feedback controllers $K \in \mathcal{K}$.

Following [4] it is convenient to rewrite the cost functional (20). For each k, given a sequence of control values $u_{k,M-1} = u_k, \ldots, u_{M-1}$ and measurements $y_{k+1,M} = y_{k+1}, \ldots, y_M$, define a random sequence of observables Q_k by the recursion ([4, equation (3.1)])

$$\begin{aligned} Q_k &= \Gamma^\dagger(u_k, y_{k+1})Q_{k+1} + L(u_k), \quad 0 \le k \le M-1 \\ Q_M &= N \end{aligned} \tag{21}$$

When useful, we write

$$Q_k = Q_k(u_{k,M-1}, y_{k+1,M})$$

to indicate dependence on the input and outputs. Q_k may be called a *cost observable*. The cost functional (20) is given by

$$\begin{aligned} &J_{\omega,0}(K) \\ &= \sum_{y_{1,M} \in \mathbf{Y}^M} \langle \omega, Q_0(K(y_{1,M})_{0,M-1}, y_{1,M})\rangle \end{aligned} \tag{22}$$

Here and elsewhere we use abbreviations of the form

$$K(y_{1,M})_{0,M-1} = (K_0, K_1(y_1), \ldots, K_{M-1}(y_{1,M-1}))$$

Remark III.2 The cost observable Q_k given by (21) and the expression in (22) is analogous to the familiar Heisenberg picture used in quantum physics. It is very natural from the point of view of dynamic programming, and indeed (20) and (22) are related by iterating (21). Here is the first step:

$$\begin{aligned} \langle \omega_0, Q_0\rangle &= \langle \omega_0, \Gamma^\dagger(u_0, y_1)Q_1 + L(u_0)\rangle \\ &= \langle \omega_0, L(u_0)\rangle + \langle \Gamma(u_0, y_1)\omega_0, Q_1\rangle \\ &= \langle \omega_0, L(u_0)\rangle + \langle \omega_1, Q_1\rangle p(y_1|u_0, \omega_0) \end{aligned}$$

where $\omega_1 = \Lambda_\Gamma(u_0, y_1)\omega_0$ and $p(y_1|u_0, \omega_0)$ is given by (6). □

The key idea of dynamic programming is to look at the current state at a current time $0 \le k \le M-1$ and to optimize the remaining cost from the current time to the final time. This leads to an iterative solution. Accordingly, we define, for each $0 \le k \le M$, the cost to go incurred by a controller K (restricted to $k \le l \le M-1$) to be

$$\begin{aligned} &J_{\omega,k}(K) \\ &= \sum_{y_{k+1,M} \in \mathbf{Y}^{M-k}} \langle \omega, Q_k(K(y_{k+1,M})_{k,M-1}, y_{k+1,M})\rangle \end{aligned} \tag{23}$$

The dynamic programming equation associated with this risk-neutral problem is

$$\begin{aligned} V(\omega, k) &= \inf_{u \in \mathbf{U}} \{\langle \omega, L(u)\rangle \\ &\qquad + \sum_{y \in \mathbf{Y}} V(\Lambda_\Gamma(u, y)\omega, k+1)p(y|u, \omega)\}, \\ V(\omega, M) &= \langle \omega, N\rangle \end{aligned} \tag{24}$$

where $0 \leq k \leq M-1$. This is the fundamental equation from which optimality or otherwise of a controller can be determined.

Let V be the solution to the dynamic programming equation (24). Then for any controller $K \in \mathcal{K}$ we have

$$V(\omega, k) \leq J_{\omega,k}(K). \tag{25}$$

If we assume in addition that a minimizer

$$\begin{aligned}\mathbf{u}^*(\omega, k) \ \in \operatorname{argmin}_{u \in \mathbf{U}} \{ & \langle \omega, L(u) \rangle \\ & + \sum_{y \in \mathbf{Y}} V(\Lambda_\Gamma(u, y)\omega, k+1) p(y|u, \omega)) \}\end{aligned} \tag{26}$$

exists[32] for all ω, $0 \leq k \leq M-1$, then the separation structure controller $K^{\mathbf{u}^*}_{\omega_0}$ (recall section II B) defined by (26) is optimal, i.e.

$$J_{\omega_0,0}(K^{\mathbf{u}^*}_{\omega_0}) = V(\omega_0, 0) \leq J_{\omega_0,0}(K) \tag{27}$$

for all $K \in \mathcal{K}$.

Example III.3 (Two-state system with feedback, Example III.1 continued.) We solve the dynamic programming equation (24) and determine the optimal feedback controls as follows. For $k = M = 2$ we have

$$V(\omega, 2) = \langle \omega, X^2 \rangle = \omega_{11}$$

and hence for $k = 1$

$$V(\omega, 1) = \omega_{11} + \min[V_0(\omega, 1), V_1(\omega, 1)]$$

where where $V_0(\omega, 1), V_1(\omega, 1)$ are given in Appendix A. Hence we obtain

$$\mathbf{u}^*(\omega, 1) = \begin{cases} 0 & \text{if } V_0(\omega, 1) \leq V_1(\omega, 1) \\ 1 & \text{if } V_0(\omega, 1) > V_1(\omega, 1). \end{cases}$$

At time $k = 0$ we have

$$V(\omega, 0) = \omega_{11} + \min[V_0(\omega, 0), V_1(\omega, 0)]$$

where $V_0(\omega, 0), V_1(\omega, 0)$ are given in Appendix A, which gives

$$\mathbf{u}^*(\omega, 0) = \begin{cases} 0 & \text{if } V_0(\omega, 0) \leq V_1(\omega, 0) \\ 1 & \text{if } V_0(\omega, 0) > V_1(\omega, 0). \end{cases}$$

The optimal risk-neutral feedback controller is given by

$$u_0 = K^{\mathbf{u}^*}_{\omega_0,0} = \mathbf{u}^*(\omega_0, 0), \ \ u_1 = K^{\mathbf{u}^*}_{\omega_0,1}(y_1) = \mathbf{u}^*(\omega_1, 1)$$

where $\omega_1 = \Lambda_\Gamma(u_0, y_1)\omega_0$. *Note that the control u_1 depends on y_1 through the conditional state ω_1 (separation structure).* A physical implementation

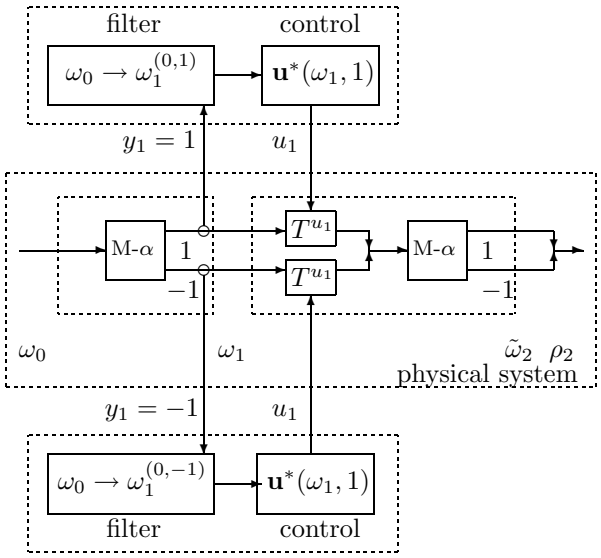

FIG. 5: Physical realization of the two stages of the two-state system with feedback using the optimal risk-neutral controller $K^{\mathbf{u}^*}_{\omega_0}$ (with ω_0 given by (17), we have $u_0 = \mathbf{u}^*(\omega_0, 0) = 0$, $u_1 = \mathbf{u}^*(\omega_1, 1)$).

of the quantum system with optimal risk-neutral feedback is shown in Figure 5.

Let's consider the special case $\alpha = 0$ and $p = 0$, with initial state (17). We then find that $V_0(\omega_0, 0) = V_1(\omega_0, 0) = 0.5$, and hence we take $\mathbf{u}^*(\omega_0, 0) = 0$; i.e. $u_0 = 0$.

Next, if $y_1 = -1$ is observed, we have $\omega_1 = |-1\rangle\langle -1|$, $V_0(\omega_1, 1) = 1$ and $V_1(\omega_1, 1) = 0$. Hence we take $\mathbf{u}^*(\omega_1, 1) = 1$, i.e. $u_1 = 1$. However, if $y_1 = 1$ is observed, we have $\omega_1 = |1\rangle\langle 1|$, $V_0(\omega_1, 1) = 0$ and $V_1(\omega_1, 1) = 1$; and hence we take $\mathbf{u}^*(\omega_1, 1) = 0$, i.e. $u_1 = 0$. In either case we achieve the desired state $\rho_2 = \omega_2 = |1\rangle\langle 1|$.

This action is the same as that seen before for the controller $\bar{K}$. The same controller is obtained for $0 < \alpha < 0.5$ and $p = 0$, but ω_2 will be a mixed state. If $p \neq 0$ the optimal controller $K^{\mathbf{u}^*}_{\omega_0}$ will result in control actions that in general differ from those of $\bar{K}$. □

Remark III.4 Note that the optimal risk-neutral controller $K^{\mathbf{u}^*}_{\omega_0}$ determined by (26) feeds back the conditional state ω_k, given by the SME (8), in accordance with its separation structure, Figure 6. We note that the conditional state ω_k obtained from (8) provides the optimal means for calculating estimates of observables (in the sense of minimum mean square error), viz. $\langle \omega_k, B \rangle$, and can be regarded as the optimal filter in this sense. This means that from the point of view of optimal risk-neutral control, the best thing to do is to make use of the optimal filter (8), a dynamical quantity that

contains *knowledge* of the quantum system, as obtained by the controller through the measurement process embedded in Γ. □

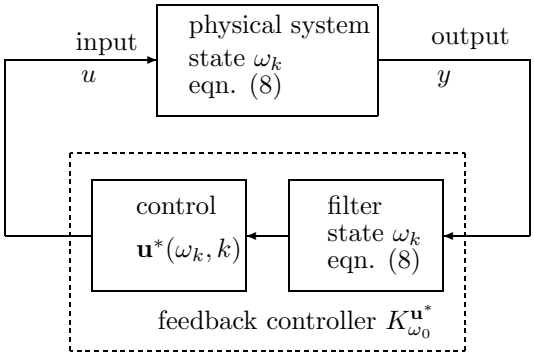

FIG. 6: Optimal risk-neutral controller $K_{\omega_0}^{\mathbf{u}^*}$ showing separation structure and states of the physical system ω_k and filter ω_k.

Remark III.5 A second remark we wish to make here concerns the well-known concept in control engineering of *dual control*, [22, Chapter 6.8]. This concept relates to the dual function of the measurement feedback controller K, viz. (i) to alter the future evolution of the system, and (ii) to alter the future values of the available information. The optimal choice of K takes both of these factors into account. □

Remark III.6 The final remark for this section concerns *feedback* and *robustness*. Feedback is the most important concept in control engineering, and has a long history going back at least to the mechanical governors used to regulate the speed of steam engines. Feedback is used to compensate for disturbances and uncertainty, and feedback loops typically enjoy a robustness margin (e.g. gain margin and phase margin in classical control engineering), a measure of this compensation ability. Note that in the absence of disturbances and uncertainty, feedback is completely unnecessary and control can be achieved by a prescribed open loop controller. However, in reality both quantum and classical systems are subject to disturbances and uncertainty, e.g. (i) the influence of an environment, (ii) model error due to approximation and unknown parameters, and (iii) imprecise measurements. In the quantum context, there is the further complication [12] that as a consequence of the act of measurement randomness is introduced, and this could potentially reduce control effectiveness. Measurement feedback control of quantum systems is fundamentally a stochastic control problem containing non-classical characteristics. These considerations underscore the importance of feedback control and the need for robustness when controlling quantum systems. Robustness issues will be taken up again in section IV (see Example IV.8). □

IV. MULTIPLICATIVE COSTS AND RISK-SENSITIVE CONTROL

We turn now to the risk-sensitive optimal control problem, the main object of this paper. The *risk-sensitive cost functional* we consider, a quantum generalization of LEQG (2), is

$$J_{\omega,0}^{\mu}(K) = \mathbf{E}_{\omega,0}[\prod_{k=0}^{M-1} \langle \omega_k, e^{\mu L(u_k)} \rangle \langle \omega_M, e^{\mu N} \rangle] \tag{28}$$

where $\mu > 0$ is a positive risk parameter, $L(u)$ is a cost function (as defined in section III), and N is a non-negative observable. The conditional states ω_k are given by the quantum system model (8).

Remark IV.1 The risk-sensitive cost (28), by use of the exponential function, gives heavy weight to large values of the cost functions in the exponents. A system controlled by a controller minimizing this cost is not likely to experience large values of these quantities. Risk-sensitive controllers are known to enjoy some robustness properties against uncertainty in the model and external disturbances, see [15] and Example IV.8. □

The primary goal in this section is to find the optimal controller for the risk-sensitive cost functional (28). As noted, this cost functional is defined in terms of the conditional state ω_k of the quantum system (8). However, in order to solve this optimization problem, we need to express the cost functional in a manner that facilitates the use of optimal control methods. As in the classical LEQG case, this requires the introduction of a new state, which in general is *unnormalized*. To define this new unnormalized state, for which we use the notation $\hat{\omega}$ to distinguish such states from normalized states ω, we need to use possibly nonlinear operators (observables) B and (super) operators R. These nonlinear operators allow us to formulate and solve a general class of multiplicative cost optimal control problems for quantum systems[33].

Our risk-sensitive and multiplicative cost functionals can be defined in terms of *(super) operator valued costs* $R(u)$ that satisfy the real multiplicative homogeneity property

$$R(u) r\hat{\omega} = rR(u)\hat{\omega} \tag{29}$$

for any real number r and any $\hat{\omega}$, u. The risk-sensitive problem corresponds to particular choices of operator valued cost, Example IV.2. However, the fundamental equations in this section are valid for any operator valued cost $R(u)$ satisfying (29). Note that operator valued costs $R(u)$ are not in general quantum operations (because linearity and the inequality $\langle R(u)\hat{\omega}, I\rangle \leq \langle \hat{\omega}, I\rangle$ need not hold in general).

Example IV.2 We give two examples of operator valued costs.

(i) A specific linear form for $R(u)$ is

$$R(u)\hat{\omega} = \sum_c Z_c(u)\hat{\omega}Z_c(u) \tag{30}$$

where $\{Z_c(u)\}$ is a family of cost functions (section III). The adjoint $R^\dagger(u)$ acts on observables B via

$$R^\dagger(u)B = \sum_c Z_c(u)BZ_c(u), \tag{31}$$

and thereby defines a linear functional on unnormalized states by

$$\langle \hat{\omega}, R^\dagger(u)B\rangle = \sum_c \langle Z_c(u)\hat{\omega}Z_c(u), B\rangle \tag{32}$$

(we have written this explicitly to facilitate comparison with (34) below).

(ii) An operator valued cost $R(u)$ corresponding to the risk-sensitive cost (28) can be defined as follows. Let $L(u)$ be a cost function, and $\mu > 0$. Then set

$$R(u)\hat{\omega} = \frac{\langle \hat{\omega}, e^{\mu L(u)}\rangle}{\langle \hat{\omega}, 1\rangle}\hat{\omega}. \tag{33}$$

Note that $R(u)$ is nonlinear, but satisfies the real multiplicative homogeneity condition (29). The adjoint operator $R^\dagger(u)$ applied to an operator B is a nonlinear functional of $\hat{\omega}$ given by

$$\langle \hat{\omega}, R^\dagger(u)B\rangle = \frac{\langle \hat{\omega}, e^{\mu L(u)}\rangle}{\langle \hat{\omega}, 1\rangle}\langle \hat{\omega}, B\rangle \tag{34}$$

(cf. (32) above).

The relationship between $R(u)$ and the risk-sensitive cost (28) will be explained in Example IV.4 below.

□

Given an operator valued cost $R(u)$, we shall find it convenient to introduce an operator $\Gamma_R(u, y)$ defined by

$$\Gamma_R(u, y) = \Gamma(u, y)R(u). \tag{35}$$

In general, Γ_R is not normalized:

$$\sum_{y\in\mathbf{Y}} \langle \Gamma_R(u, y)\hat{\omega}, I\rangle = \langle R(u)\hat{\omega}, I\rangle \neq \langle \hat{\omega}, I\rangle.$$

The operator Γ_R will be used to define a new state evolution as follows. Define an operator $\Lambda_{\Gamma,R}$ by

$$\Lambda_{\Gamma,R}(u, y)\hat{\omega} = \frac{\Gamma_R(u, y)\hat{\omega}}{p_R(y|u, \hat{\omega})} \tag{36}$$

where

$$p_R(y|u, \hat{\omega}) = \frac{\langle \Gamma_R(u, y)\hat{\omega}, I\rangle}{\langle R(u)\hat{\omega}, I\rangle}. \tag{37}$$

In general, the state $\Lambda_{\Gamma,R}(u, y)\hat{\omega}$ is *unnormalized*. However, $p_R(y|u, \hat{\omega})$ is a probability distribution, since it is easy to check that

$$\sum_{y\in\mathbf{Y}} p_R(y|u, \hat{\omega}) = 1.$$

However, we point out that

$$\langle \Lambda_{\Gamma,R}(u, y)\hat{\omega}, I\rangle = \langle R(u)\hat{\omega}, I\rangle. \tag{38}$$

This unnormalized state transition operator arises in the dynamic programming equation, as we shall see below.

Associated with the operator $\Lambda_{\Gamma,R}$ is the dynamics

$$\hat{\omega}_{k+1} = \Lambda_{\Gamma,R}(u_k, y_{k+1})\hat{\omega}_k, \tag{39}$$

where y_{k+1} is distributed according to the probability distribution $p_R(y_{k+1}|u_k, \hat{\omega}_k)$ given by (37). This is a controlled Markov chain, with *unnormalized* states $\hat{\omega}_k$. It is a modified stochastic master equation corresponding to the operator Γ_R. Under the action of a controller $K \in \mathcal{K}$ the stochastic process $\hat{\omega}_k$ is determined by (39) and $u_k = K_k(y_{1,k})$.

The separation structure controller in this case takes the following form. Given a function $\hat{\mathbf{u}}(\hat{\omega}, k)$ and initial state $\hat{\omega}_0$ we define a controller $K^{\hat{\mathbf{u}}}_{\hat{\omega}_0} \in \mathcal{K}$ by

$$u_k = \hat{\mathbf{u}}(\hat{\omega}_k, k)$$

where $\hat{\omega}_k$ is given by (39), $0 \leq k \leq M$, with initial condition $\hat{\omega}_0$.

Let $M > 0$ be a positive integer indicating a finite time interval $k = 0, \ldots, M$. For each k, given a sequence of control values $u_{k,M-1} = u_k, \ldots, u_{m-1}$ and measurement values $y_{k+1,M} = y_{k+1}, \ldots, y_M$, define random cost observables G_k by the recursion

$$\begin{aligned} G_k &= R^\dagger(u_k)\Gamma^\dagger(u_k, y_{k+1})G_{k+1}, \\ G_M &= F \end{aligned} \tag{40}$$

where $0 \leq k \leq M-1$ and F is a non-negative linear observable. It is evident that G_k is real multiplicative homogeneous if G_{k+1} is (recall (29)).

We next define the *multiplicative* cost functional

$$J^{\mu}_{\hat{\omega},0}(K) = \sum_{y_{1,M} \in \mathbf{Y}^M} \langle \hat{\omega}, G_0(K(y_{1,M})_{0,M-1}, y_{1,M}) \rangle \tag{41}$$

where $K \in \mathcal{K}$ is a measurement feedback controller.

Lemma IV.3 *The cost functional $J^{\mu}_{\hat{\omega},0}(K)$ defined by (41) is given by the alternate expression*

$$J^{\mu}_{\hat{\omega},0}(K) = \mathbf{E}^K_{\hat{\omega},0}[\langle \hat{\omega}_M, F \rangle] \tag{42}$$

where $\hat{\omega}_i$, $i = k, \ldots, M$ is the solution of the recursion (39) with initial state $\hat{\omega}_0 = \hat{\omega}$ under the action of the controller K.

Proof. We have

$$\begin{aligned} \langle \hat{\omega}_0, G_0 \rangle &= \langle \hat{\omega}_0, R^{\dagger}(u_0)\Gamma^{\dagger}(u_0, y_1)G_1 \rangle \\ &= \langle R(u_0)\hat{\omega}_0, \Gamma(u_0, y_1)^{\dagger} G_1 \rangle \\ &= \langle \Gamma(u_0, y_1)R(u_0)\hat{\omega}_0, G_1 \rangle \\ &= \langle \hat{\omega}_1, G_1) \rangle p_R(y_1 | u_0, \hat{\omega}_0) \end{aligned}$$

where $\hat{\omega}_1 = \Lambda_{\Gamma,R}(u_0, y_1)\hat{\omega}_0$ and $p_R(y_1|u_0, \hat{\omega}_0)$ is given by (37). Iterating in this way we see that (41) and (42) are equivalent. These properties use the real multiplicative homogeneity property of G_k. □

Example IV.4 (Continuation of Example IV.2 (ii).) We now show that when $R(u)$ is given by (33), and

$$F = e^{\mu N},$$

where N is a non-negative linear observable, the multiplicative cost functional $J^{\mu}_{\hat{\omega},0}(K)$ defined by (41) equals the risk-sensitive cost functional (28).

Proceeding as in the proof of Lemma IV.3 we have

$$\begin{aligned} &\langle \hat{\omega}_0, G_0 \rangle \\ &= \langle \hat{\omega}_0, R^{\dagger}(u_0)\Gamma^{\dagger}(u_0, y_1)G_1 \rangle \\ &= \langle \Gamma(u_0, y_1)R(u_0)\hat{\omega}_0, G_1 \rangle \\ &= \langle \Gamma(u_0, y_1)\hat{\omega}_0, G_1 \rangle \frac{\langle \hat{\omega}_0, e^{\mu L(u_0)} \rangle}{\langle \hat{\omega}_0, 1 \rangle} \\ &= \frac{\langle \Gamma(u_0, y_1)\hat{\omega}_0, G_1 \rangle / \langle \hat{\omega}_0, 1 \rangle}{\langle \Gamma(u_0, y_1)\hat{\omega}_0, 1 \rangle / \langle \hat{\omega}_0, 1 \rangle} \langle \hat{\omega}_0, e^{\mu L(u_0)} \rangle \cdot \\ &\qquad \cdot \frac{\langle \Gamma(u_0, y_1)\hat{\omega}_0, 1 \rangle}{\langle \hat{\omega}_0, 1 \rangle} \\ &= \langle \Lambda_{\Gamma}(u_0, y_1)\bar{\hat{\omega}}_0, G_1 \rangle \langle \hat{\omega}_0, e^{\mu L(u_0)} \rangle p(y_1 | u_0, \bar{\hat{\omega}}_0) \end{aligned}$$

where $\bar{\hat{\omega}}_0 = \hat{\omega}_0 / \langle \hat{\omega}_0, 1 \rangle$, $\Lambda_{\Gamma}(u, y)$ is defined by (9) and $p(y|u, \omega)$ is defined by (6). Now if $\hat{\omega}_0 = \omega_0$ is normalized, with $\langle \omega_0, 1 \rangle = 1$, then we have shown that

$$\begin{aligned} &\langle \omega_0, G_0 \rangle \\ &= \langle \Lambda_{\Gamma}(u_0, y_1)\omega_0, G_1 \rangle \langle \omega_0, e^{\mu L(u_0)} \rangle p(y_1 | u_0, \bar{\omega}_0) \\ &= \langle \omega_1, G_1 \rangle \langle \omega_0, e^{\mu L(u_0)} \rangle p(y_1 | u_0, \omega_0) \end{aligned}$$

where $\omega_1 = \Lambda_{\Gamma}(u_0, y_1)\omega_0$ is the normalized state evolving according to the quantum system model (8). Note that $\langle \omega_1, 1 \rangle = 1$. Continuing in this way we see that (41) equals the risk-sensitive cost functional (28), using Lemma IV.3.

It can also be checked that ω_k and $\hat{\omega}_k$ are related simply via

$$\hat{\omega}_k = \prod_{i=0}^{k-1} \langle \omega_i, e^{\mu L(u_i)} \rangle \omega_k. \tag{43}$$

□

To solve the optimal control problem for the cost functional (41), we define the cost to go

$$\begin{aligned} &J^{\mu}_{\hat{\omega},k}(K) \\ &= \sum_{y_{k+1,M} \in \mathbf{Y}^{M-k}} \langle \hat{\omega}, G_k(K(y_{k+1,M})_{k,M-1}, y_{k+1,M}) \rangle \end{aligned} \tag{44}$$

and the corresponding dynamic programming equation

$$\begin{aligned} W(\hat{\omega}, k) &= \inf_{u \in \mathbf{U}} \{ \sum_{y \in \mathbf{Y}} W(\Lambda_{\Gamma,R}(u, y)\hat{\omega}, k+1) \\ &\qquad \cdot p_R(y | u, \hat{\omega}) \}, \\ W(\hat{\omega}, M) &= \langle \hat{\omega}, F \rangle \end{aligned} \tag{45}$$

where $0 \leq k \leq M-1$.

Theorem IV.5 *Let $W(\hat{\omega}, k)$, $0 \leq k \leq M$, be the solution of the dynamic programming equation (45). (i) Then for any $K \in \mathcal{K}$ we have*

$$W(\hat{\omega}, k) \leq J^{\mu}_{\hat{\omega},k}(K) \tag{46}$$

(ii) Assume in addition that the minimizer

$$\begin{aligned} &\hat{\mathbf{u}}^*(\hat{\omega}, k) \\ &\in \operatorname*{argmin}_{u \in \mathbf{U}} \{ \sum_{y \in \mathbf{Y}} W(\Lambda_{\Gamma,R}(u, y)\hat{\omega}, k+1) p_R(y | u, \hat{\omega}) \} \end{aligned} \tag{47}$$

exists for all $\hat{\omega}$, $0 \leq k \leq M-1$. Then the separation structure controller $K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0}$ defined by (47) is optimal for problem (41), i.e. $J^{\mu}_{\hat{\omega}_0,0}(K) \geq J^{\mu}_{\hat{\omega}_0,0}(K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0})$ for all $K \in \mathcal{K}$.

Proof. We prove part (i) by induction. Let $K \in \mathcal{K}$. For $k = M$, we have

$$W(\hat{\omega}, M) = \langle \hat{\omega}, F \rangle = J^{\mu}_{\hat{\omega},M}(K)$$

so (46) holds for $k = M$. Next, we assume (46) holds for $k+1$, i.e.

$$W(\hat{\omega}, k+1) \leq J^{\mu}_{\hat{\omega},k+1}(K) \tag{48}$$

Now by (45), (48) and (37)

$$\begin{aligned}
&W(\hat{\omega}, k) \\
&\leq \sum_{y_{k+1}\in\mathbf{Y}} W(\Lambda_{\Gamma,R}(u_k, y_{k+1})\hat{\omega}, k+1) p_R(y_{k+1}|u_k, \hat{\omega}) \\
&\leq \sum_{y_{k+1}\in\mathbf{Y}} J^{\mu}_{\Lambda_{\Gamma,R}(u_k,y_{k+1})\hat{\omega},k+1}(K) p_R(y_{k+1}|u_k, \hat{\omega}) \\
&= \sum_{y_{k+1}\in\mathbf{Y}} \sum_{y_{k+2,M}\in\mathbf{Y}^{M-(k+1)}} \langle \Lambda_{\Gamma,R}(u_k, y_{k+1})\hat{\omega}, G_{k+1}\rangle \cdot \\
&\qquad\qquad \cdot p_R(y_{k+1}|u_k, \hat{\omega}) \\
&= \sum_{y_{k+1}\in\mathbf{Y}} \sum_{y_{k+2,M}\in\mathbf{Y}^{M-(k+1)}} \langle \hat{\omega}, \Gamma^{\dagger}_R(u_k, y_{k+1}) G_{k+1}\rangle \\
&= J^{\mu}_{\hat{\omega},k}(K)
\end{aligned}$$

as required.

Part (ii) follows from the proof of part (i), with $k = 0$, since at every step we have equality and so

$$W(\hat{\omega}_0, 0) = J^{\mu}_{\hat{\omega}_0,0}(K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0})$$

Hence $J^{\mu}_{\hat{\omega}_0,0}(K) \geq J^{\mu}_{\hat{\omega}_0,0}(K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0})$ for all $K \in \mathcal{K}$. The real multiplicative homogeneity property of G_k has been used here also. □

Remark IV.6 Note that the optimal multiplicative cost/risk-sensitive controller $K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0}$ determined by (47) feeds back the unnormalized conditional state $\hat{\omega}_k$, given by the modified SME (39), Figure 7. This means that from the point of view of optimal risk-sensitive or multiplicative control, the best thing to do involves use of a dynamical quantity that not only contains *knowledge* (as measured by the controller) of the quantum system, but also contains information about the *purpose* of the controller. This should be contrasted with the risk-neutral case, section III. Note in particular that the modified SME (39) is no longer the optimal filter from the point of view of seeking the best estimate of observables (c.f. Remark III.4). Further, the concept of dual control (Remark III.5) has greater weight here, since the cost $R(u)$ appears explicitly in the controller dynamics (39)—while the optimal multiplicative cost/risk-sensitive controller has a separation structure, in the sense of a decomposition into a dynamical filter part and static control part, the task of estimation is not separated from the task of control, Figure 7. We emphasize that the multiplicative cost/risk-sensitive conditional state is defined only in the context of these specific control objectives, where they are used in specific feedback situations. □

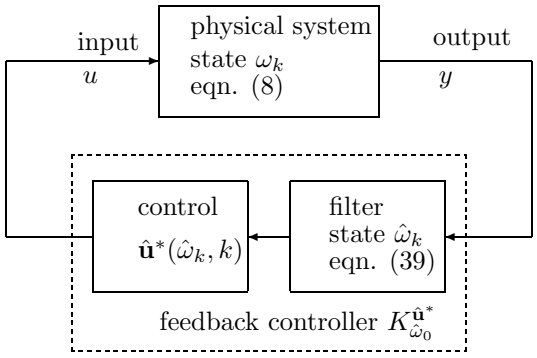

FIG. 7: Optimal multiplicative/risk-sensitive controller $K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0}$ showing separation structure and states of the physical system ω_k and filter $\hat{\omega}_k$.

Example IV.7 (Two-state system with feedback, Example III.3 continued.) We now consider the risk-sensitive optimal control problem for the two-state example, with operator valued cost $R(u)$ defined by (33) where $L(u)$ is given by (19). If we write the density matrix as (15), then

$$R(u)\hat{\omega} = \frac{e^{\mu}\hat{\omega}_{11} + \hat{\omega}_{22}}{\hat{\omega}_{11} + \hat{\omega}_{22}} e^{\mu c(u)} \begin{pmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} \\ \hat{\omega}^*_{12} & \hat{\omega}_{22} \end{pmatrix}.$$

The risk-sensitive controlled transfer operators $\Gamma_R(u, y)$ are defined by

$$\begin{aligned}
\Gamma_R(u, y)\hat{\omega} &= \Gamma(u, y) R(u)\hat{\omega} \\
&= \tfrac{\langle \hat{\omega}, e^{\mu L(u)}\rangle}{\langle \hat{\omega}, 1\rangle} \Gamma(u, y)\hat{\omega}.
\end{aligned}$$

Explicitly, we have

$$\begin{aligned}
\Gamma_R(0,-1)\hat{\omega} &= \tfrac{e^{\mu}\hat{\omega}_{11}+\hat{\omega}_{22}}{\hat{\omega}_{11}+\hat{\omega}_{22}} \begin{pmatrix} (1-\alpha)\hat{\omega}_{11} & 0 \\ 0 & \alpha\hat{\omega}_{22} \end{pmatrix} \\
\Gamma_R(0,1)\hat{\omega} &= \tfrac{e^{\mu}\hat{\omega}_{11}+\hat{\omega}_{22}}{\hat{\omega}_{11}+\hat{\omega}_{22}} \begin{pmatrix} \alpha\hat{\omega}_{11} & 0 \\ 0 & (1-\alpha)\hat{\omega}_{22} \end{pmatrix} \\
\Gamma_R(1,-1)\hat{\omega} &= \tfrac{e^{\mu}\hat{\omega}_{11}+\hat{\omega}_{22}}{\hat{\omega}_{11}+\hat{\omega}_{22}} e^{\mu p} \begin{pmatrix} (1-\alpha)\hat{\omega}_{22} & 0 \\ 0 & \alpha\hat{\omega}_{11} \end{pmatrix} \\
\Gamma_R(1,1)\hat{\omega} &= \tfrac{e^{\mu}\hat{\omega}_{11}+\hat{\omega}_{22}}{\hat{\omega}_{11}+\hat{\omega}_{22}} e^{\mu p} \begin{pmatrix} \alpha\hat{\omega}_{22} & 0 \\ 0 & (1-\alpha)\hat{\omega}_{11} \end{pmatrix}.
\end{aligned}$$

The dynamic programming equation (45) is solved and the optimal feedback controls are found as follows. First, for $k = M = 2$ we have

$$W(\hat{\omega}, 2) = \langle \hat{\omega}, e^{\mu X^2}\rangle = e^{\mu}\hat{\omega}_{11} + \hat{\omega}_{22}$$

and then for $k = 1$

$$W(\hat{\omega}, 1) = \min[W_0(\hat{\omega}, 1), W_0(\hat{\omega}, 1)]$$

where $W_0(\hat{\omega}, 1)$ and $W_1(\hat{\omega}, 1)$ are given in Appendix A, and

$$\hat{\mathbf{u}}^*(\omega, 1) = \begin{cases} 0 & \text{if } W_0(\omega, 1) \leq W_1(\omega, 1) \\ 1 & \text{if } W_0(\omega, 1) > W_1(\omega, 1). \end{cases}$$

Next, for $k = 0$,

$$W(\hat{\omega}, 0) = \min[W_0(\hat{\omega}, 0), W_0(\hat{\omega}, 0)]$$

where $W_0(\hat{\omega}, 0)$ and $W_1(\hat{\omega}, 0)$ are given in Appendix A, and

$$\hat{\mathbf{u}}^*(\omega, 0) = \begin{cases} 0 & \text{if } W_0(\omega, 0) \le W_1(\omega, 0) \\ 1 & \text{if } W_0(\omega, 0) > W_1(\omega, 0). \end{cases}$$

The optimal feedback controller is given by

$$u_0 = K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0,0} = \hat{\mathbf{u}}^*(\hat{\omega}_0, 0), \;\; u_1 = K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0,1}(y_1) = \hat{\mathbf{u}}^*(\omega_1, 1)$$

where $\hat{\omega}_1 = \Lambda_{\Gamma,R}(u_0, y_1)\hat{\omega}_0$. *Again, we see the separation structure, where here the control u_1 depends on y_1 through the unnormalized conditional state $\hat{\omega}_1$.* A physical implementation of this controller is shown in Figure 8.

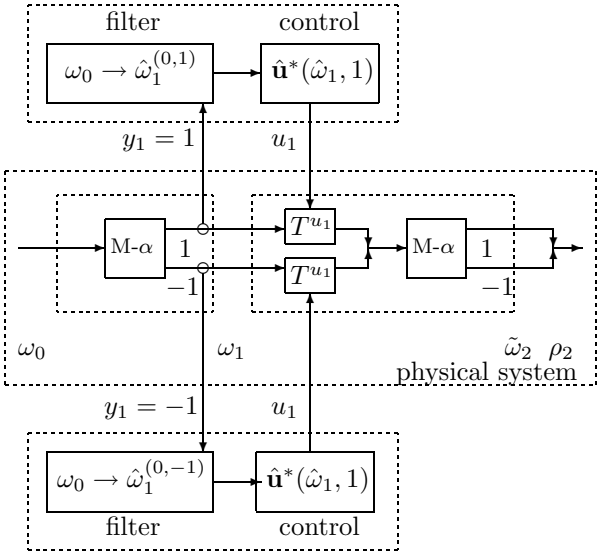

FIG. 8: Physical realization of the two stages of the two-state system with feedback using the optimal risk-sensitive controller $K^{\hat{\mathbf{u}}^*}_{\hat{\omega}_0}$ (with $\hat{\omega}_0 = \omega_0$ given by (17), we have $u_0 = \hat{\mathbf{u}}^*(\omega_0, 0) = 0$, $u_1 = \hat{\mathbf{u}}^*(\hat{\omega}_1, 1)$).

To see the effect of the risk-sensitive controller, consider the initial state $\hat{\omega}_0 = \omega_0$ given by (17), and parameter values $\alpha = 0.25$, $\mu = 2$. We find that $W_0(\hat{\omega}_0, 0) = W_1(\hat{\omega}_0, 0)$, and hence we take $u_0 = \hat{\mathbf{u}}^*(\hat{\omega}_0, 0) = 0$.

If $y_1 = -1$ is measured, we find that

$$\hat{\omega}_1 = \begin{pmatrix} 3.1459 & 0 \\ 0 & 1.04863 \end{pmatrix},$$
$$\omega_1 = \begin{pmatrix} 0.75 & 0 \\ 0 & 0.25 \end{pmatrix} \text{ with prob. } 0.5,$$

and

$$u_1 = \hat{\mathbf{u}}^*(\hat{\omega}_1, 1) = \begin{cases} 1 & \text{if } 0 \le p \le 0.4 \\ 0 & \text{if } p > 0.4 \end{cases} \qquad (49)$$

If, on the other hand, $y_1 = 1$ is measured, we find that

$$\hat{\omega}_1 = \begin{pmatrix} 1.04863 & 0 \\ 0 & 3.1459 \end{pmatrix},$$
$$\omega_1 = \begin{pmatrix} 0.25 & 0 \\ 0 & 0.75 \end{pmatrix} \text{ with prob. } 0.5,$$

and

$$u_1 = \hat{\mathbf{u}}^*(\hat{\omega}_1, 1) = 0$$

for any value of p. When the control cost $p = 0.2$, the final (non-selective) state is given by

$$\rho_2 = \begin{pmatrix} 0.25 & 0 \\ 0 & 0.75 \end{pmatrix} = 0.25|-1\rangle\langle -1| + 0.75|1\rangle\langle 1|$$

This state does not equal the desired pure state $|1\rangle\langle 1|$, a reflection of the level of measurement uncertainty $\alpha = 0.25$ and the presence of a non-zero control penalty.

To compare with the risk-neutral version of this problem, we find that the threshold in (49) for the risk-neutral problem is $p = 0.75$. This means that for a larger range of values of the control cost p, the risk-neutral controller will be active, i.e. select $u = 1$ than is the case for the risk-sensitive controller. This is consistent with the description of the example in [8, 16] where the risk-neutral controller is more aggressive than the risk-sensitive controller. □

We conclude with an example which indicates the likely robustness properties of the risk-sensitive controller and the relationship between the risk-neutral and risk-sensitive problems.

Example IV.8 We consider the risk-sensitive cost functional (28), where where the operator valued cost $R(u)$ has the form (33), and $F = e^{\mu N}$.

Robustness. To describe the robustness properties of the risk-sensitive controller, we follow [15] and make use of the following general convex duality formula (see, e.g. [14, Chapter 1.4]):

$$\log \mathbf{E}_{\mathbf{P}}[e^f] = \sup_{\mathbf{Q}}\{\mathbf{E}_{\mathbf{Q}}[f] - RE(\mathbf{Q} \parallel \mathbf{P})\} \qquad (50)$$

where $\mathbf{P}$ and $\mathbf{Q}$ are probability distributions[34], and where the relative entropy is defined by (see, e.g., [23, Chapter 11])

$$RE(\mathbf{Q} \parallel \mathbf{P}) = \mathbf{E}_{\mathbf{Q}}[\log \frac{d\mathbf{Q}}{d\mathbf{P}}].$$

To apply formula (50), we proceed as follows. Let Γ_{nom} be the nominal operator used for design of the optimal risk-sensitive controller, here denoted

$\hat{K}^*_{nom}$. Together, Γ_{nom} and $\hat{K}^*_{nom}$ determine a probability distribution, here denoted $\mathbf{P}_{nom}$. In reality, the nominal Γ_{nom} need not equal the operator for the "true" system, denoted Γ_{true}. The controller $\hat{K}^*_{nom}$ is applied to the true system, resulting in a probability distribution $\mathbf{P}_{true}$[35].

We write $\mu = 1/\gamma^2$, and apply (50) to obtain the following inequality ($\mathbf{P} = \mathbf{P}_{nom}$, $\mathbf{Q} = \mathbf{P}_{true}$):

$$\begin{aligned}
&\gamma^2 \log \mathbf{E}_{\mathbf{P}_{true}}[\textstyle\prod_{k=0}^{M-1} \langle \omega_k, e^{\mu L(u_k)}\rangle \langle \omega_M, e^{\mu N}\rangle] \\
&\geq \mathbf{E}_{\mathbf{P}_{true}}[\gamma^2 \log(\textstyle\prod_{k=0}^{M-1} \langle \omega_k, e^{\mu L(u_k)}\rangle \langle \omega_M, e^{\mu N}\rangle)] \\
&\quad -\gamma^2 RE(\mathbf{P}_{true} \parallel \mathbf{P}_{nom}) \\
&= \mathbf{E}_{\mathbf{P}_{true}}[\gamma^2(\textstyle\sum_{k=0}^{M-1} \log \langle \omega_k, e^{\mu L(u_k)}\rangle \\
&\qquad + \log \langle \omega_M, e^{\mu N}\rangle)] \\
&\qquad\qquad -\gamma^2 RE(\mathbf{P}_{true} \parallel \mathbf{P}_{nom}) \\
&\geq \mathbf{E}_{\mathbf{P}_{true}}[\textstyle\sum_{k=0}^{M-1} \langle \omega_k, L(u_k)\rangle + \langle \omega_M, N\rangle)] \\
&\quad -\gamma^2 RE(\mathbf{P}_{true} \parallel \mathbf{P}_{nom})
\end{aligned}$$

This implies the important bound:

$$\begin{aligned}
&J^{rn}_{\mathbf{P}_{true}}(\bar{K}^*_{nom}) \\
&\quad \leq \gamma^2 \log J^{rs,\gamma^2}_{\mathbf{P}_{nom}}(\hat{K}^*_{nom}) + \gamma^2 RE(\mathbf{P}_{true} \parallel \mathbf{P}_{nom})
\end{aligned} \tag{51}$$

The LHS of (51) is the risk-neutral cost criterion (20), evaluated using the true system model $\mathbf{P}_{true}$ and the controller $\hat{K}^*_{nom}$ designed using the nominal model $\mathbf{P}_{nom}$. Inequality (51) bounds this cost by two terms, the first term is related to the optimal risk-sensitive cost (28), while the second is the relative entropy term, which is a measure of the "distance" between the true and nominal systems. The number $\gamma^2 = 1/\mu > 0$ is a "robustness gain" parameter, which we would like to be as small as possible for maximum robustness, as in H^∞ robust control, [19], where the relative entropy term is a measure of the "energy" in the disturbance or uncertainty. This shows that the risk-sensitive controller enjoys good performance, as measured by the risk-neutral criterion, under nominal conditions ($\mathbf{P}_{true} = \mathbf{P}_{nom}$), and acceptable performance in other than nominal conditions ($\mathbf{P}_{true} \neq \mathbf{P}_{nom}$), as implied by the bound. In summary, risk-sensitive controllers enjoy enhanced robustness (recall Remark III.6).

Relationship between the risk-neutral and risk-sensitive value functions. We indicate briefly how the results of [21, Theorem 5.5] apply in the present context. Indeed, the reader may check that for small $\mu > 0$ one has

$$\frac{1}{\mu} \log \frac{\langle \hat{\omega}, \exp(\mu N)\rangle}{\langle \hat{\omega}, 1\rangle} \approx \frac{\langle \hat{\omega}, N\rangle}{\langle \hat{\omega}, 1\rangle}$$

This suggests the relation

$$\lim_{\mu \downarrow 0} \frac{1}{\mu} \log \frac{W(\hat{\omega}, k)}{\langle \hat{\omega}, 1\rangle} = \frac{V(\hat{\omega}, k)}{\langle \hat{\omega}, 1\rangle}, \tag{52}$$

which says that a logarithmic risk-sensitive optimal cost tends to the optimal risk-neutral cost as the parameter $\mu \to 0$, as might be expected. □

APPENDIX A: FORMULAS FOR THE TWO-STATE SYSTEM WITH FEEDBACK EXAMPLE

The following quantities were used in the solution of the risk-neutral problem, Example III.3:

$$V_0(\omega, 1) = \omega_{11}, \quad V_1(\omega, 1) = \omega_{22} + p$$

$$\begin{aligned}
V_0(\omega, 0) = \omega_{11} &+ \min[\alpha\omega_{11}, p + \omega_{22} - \alpha\omega_{22}] \\
&+ \min[\omega_{11} - \alpha\omega_{11}, p + \alpha\omega_{22}]
\end{aligned}$$

$$\begin{aligned}
V_1(\omega, 0) = p + \alpha\omega_{11} + \omega_{22} - \alpha\omega_{22} \\
+ \min[\alpha\omega_{11}, p + \omega_{22} - \alpha\omega_{22}] \\
+ \min[p + \alpha\omega_{11}, \omega_{22} - \alpha\omega_{22}]
\end{aligned}$$

The following quantities were used in the solution of the risk-sensitive problem, Example IV.7:

$$W_0(\hat{\omega}, 1) = \frac{\left(e^{\mu}\,\hat{\omega}_{11} + \hat{\omega}_{22}\right)^2}{\hat{\omega}_{11} + \hat{\omega}_{22}}, \quad W_1(\hat{\omega}, 1) = \frac{e^{\mu\,p}\left(\hat{\omega}_{11}\,\hat{\omega}_{22} + e^{2\,\mu}\,\hat{\omega}_{11}\,\hat{\omega}_{22} + e^{\mu}\left(\hat{\omega}_{11}^2 + \hat{\omega}_{22}^2\right)\right)}{\hat{\omega}_{11} + \hat{\omega}_{22}}$$

$$\begin{aligned}
W_0(\hat{\omega}, 0) \;= \min[&-\left(\tfrac{\left(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left((-1+\alpha)\,e^{\mu}\,\hat{\omega}_{11}-\alpha\,\hat{\omega}_{22}\right)^2}{\left(\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left((-1+\alpha)\,\hat{\omega}_{11}-\alpha\,\hat{\omega}_{22}\right)}\right), \\
&-\left(\tfrac{e^{\mu\,p}\left(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left(-((-1+\alpha)\,\alpha\,\hat{\omega}_{11}\,\hat{\omega}_{22})-(-1+\alpha)\,\alpha\,e^{2\,\mu}\,\hat{\omega}_{11}\,\hat{\omega}_{22}+e^{\mu}\left((-1+\alpha)^2\,\hat{\omega}_{11}^2+\alpha^2\,\hat{\omega}_{22}^2\right)\right)}{\left(\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left((-1+\alpha)\,\hat{\omega}_{11}-\alpha\,\hat{\omega}_{22}\right)}\right)] \\
&+\min([\tfrac{\left(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left(\alpha\,e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22}-\alpha\,\hat{\omega}_{22}\right)^2}{\left(\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left(\alpha\,\left(\hat{\omega}_{11}-\hat{\omega}_{22}\right)+\hat{\omega}_{22}\right)}, \\
&\tfrac{e^{\mu\,p}\left(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left(e^{\mu}\,\hat{\omega}_{22}^2+\alpha\,\hat{\omega}_{22}\,\left(\hat{\omega}_{11}+e^{2\,\mu}\,\hat{\omega}_{11}-2\,e^{\mu}\,\hat{\omega}_{22}\right)+\alpha^2\,\left(-(\hat{\omega}_{11}\,\hat{\omega}_{22})-e^{2\,\mu}\,\hat{\omega}_{11}\,\hat{\omega}_{22}+e^{\mu}\,\left(\hat{\omega}_{11}^2+\hat{\omega}_{22}^2\right)\right)\right)}{\left(\hat{\omega}_{11}+\hat{\omega}_{22}\right)\left(\alpha\,\left(\hat{\omega}_{11}-\hat{\omega}_{22}\right)+\hat{\omega}_{22}\right)}]
\end{aligned}$$

$$W_1(\hat{\omega},0) = \min[-(\frac{e^{\mu\,p}\,(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22})\,(\hat{\omega}_{11}-\alpha\,\hat{\omega}_{11}+\alpha\,e^{\mu}\,\hat{\omega}_{22})^2}{(wr11+\hat{\omega}_{22})\,((-1+\alpha)\,\hat{\omega}_{11}-\alpha\,\hat{\omega}_{22})}),$$
$$-(\frac{e^{2\,\mu\,p}\,(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22})\,(-((-1+\alpha)\,\alpha\,\hat{\omega}_{11}\,\hat{\omega}_{22})-(-1+\alpha)\,\alpha\,e^{2\,\mu}\,\hat{\omega}_{11}\,\hat{\omega}_{22}+e^{\mu}\,((-1+\alpha)^2\,\hat{\omega}_{11}^2+\alpha^2\,\hat{\omega}_{22}^2))}{(\hat{\omega}_{11}+\hat{\omega}_{22})\,((-1+\alpha)\,\hat{\omega}_{11}-\alpha\,\hat{\omega}_{22})})]$$
$$+\min[\frac{e^{\mu\,p}\,(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22})\,(e^{\mu}\,\hat{\omega}_{22}+\alpha\,(\hat{\omega}_{11}-e^{\mu}\,\hat{\omega}_{22}))^2}{(\hat{\omega}_{11}+\hat{\omega}_{22})\,(\alpha\,(\hat{\omega}_{11}-\hat{\omega}_{22})+\hat{\omega}_{22})},$$
$$\frac{e^{2\,\mu\,p}\,(e^{\mu}\,\hat{\omega}_{11}+\hat{\omega}_{22})\,(e^{\mu}\,\hat{\omega}_{22}^2+\alpha\,\hat{\omega}_{22}\,(\hat{\omega}_{11}+e^{2\,\mu}\,\hat{\omega}_{11}-2\,e^{\mu}\,\hat{\omega}_{22})+\alpha^2\,(-(\hat{\omega}_{11}\,\hat{\omega}_{22})-e^{2\,\mu}\,\hat{\omega}_{11}\,\hat{\omega}_{22}+e^{\mu}\,(\hat{\omega}_{11}^2+\hat{\omega}_{22}^2)))}{(\hat{\omega}_{11}+\hat{\omega}_{22})\,(\alpha\,(\hat{\omega}_{11}-\hat{\omega}_{22})+\hat{\omega}_{22})}]$$

[1] K.J. Astrom. *Introduction to Stochastic Control Theory*. Academic Press, New York, 1970.

[2] J.S. Baras and M.R. James. Robust and risk-sensitive for finite state machines and hidden markov models. *J. Math. Systems, Estimation and Control*, 7:371–374, 1997.

[3] A. Barchielli and V.P. Belavkin. Measurements continuous in time and *a posteriori* states in quantum mechanics. *J. Phys. A: Math. Gen.*, 24:1495–1514, 1991.

[4] V.P. Belavkin. On the theory of controlling observable quantum systems. *Automation and Remote Control*, 44(2):178–188, 1983.

[5] V.P. Belavkin. Quantum noise, bits and jumps: Uncertainties, decoherence, trajectories and filtering. *Progress in Quantum Electronics*, 25(1):3–50, 2001.

[6] A. Bensoussan and J.H. van Schuppen. Optimal control of partially observable stochastic systems with an exponential-of-integral performance index. *SIAM Journal on Control and Optimization*, 23:599–613, 1985.

[7] D.P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Boston, 1995.

[8] S.P. Coraluppi and S.I. Marcus. Risk-sensitive and minimax control of discrete-time, finite-state markov decision processes. *Automatica*, 35, 1999.

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

[10] A.C. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S.M Tan. Quantum feedback control and classical control theory. *Phys. Rev. A*, 62:012105, 2000. quant-ph/9912107.

[11] A.C. Doherty and K. Jacobs. Feedback-control of quantum systems using continuous state-estimation. *Phys. Rev. A*, 60:2700, 1999. quant-ph/9812004.

[12] A.C. Doherty, K. Jacobs, and G. Jungman. Information, disturbance and hamiltonian quantum feedback control. *quant-ph/0006013*, 2000.

[13] J.C. Doyle, K. Glover, P.P. Khargonekar, and B. Francis. State-space solutions to the standard H_2 and H_∞ control problems. *IEEE Transactions on Automatic Control*, 34(8):831–847, 1989.

[14] P. Dupuis and R.S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York, 1997.

[15] P. Dupuis, M.R. James, and I.R. Petersen. Robust properties of risk-sensitive control. *Math. Control, Systems and Signals*, 13:318–332, 2000.

[16] E. Fernandez-Gaucherand and S.I. Marcus. Risk-sensitive optimal control of hidden markov models: Structural results. *IEEE Trans. Automatic Control*, 42(10):1418 –1422, 1997.

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

[18] K. Glover and J.C. Doyle. State-space formulae for all stabilizing controllers that satisfy an H_∞ norm bound and relations to risk-sensitivity. *Systems and Control Letters*, pages 167–172, 1988.

[19] M. Green and D.J.N. Limebeer. *Linear Robust Cotrol*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

[20] D.H. Jacobson. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control*, 18(2):124–131, 1973.

[21] M.R. James, J.S. Baras, and R.J. Elliott. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, 39:780–792, 1994.

[22] P.R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, Englewood Cliffs, NJ, 1986.

[23] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[24] F. Verstraete, A.C. Doherty, and H. Mabuchi. Sensitivity optimization in quantum parameter estimation. *Phys. Rev. A*, 64(3):032111, 2001. quant-ph/0104116.

[25] P. Whittle. Risk-sensitive linear/ quadratic/ Gaussian control. *Advances in Applied Probability*, 13:764–777, 1981.

[26] H. Wiseman. *Quantum Trajectories and Feedback*. PhD thesis, University of Queensland, Brisbane, Australia, 1994.

[27] H. Wiseman and G.J. Milburn. Squeezing via feedback. *Phys. Rev. A*, 49(2):1350–1366, 1994.

[28] If x is a column vector, the notation x' indicates the transpose, a row vector.

[29] Without further qualification, we use the term state to refer to a positive self-adjoint operator normalized to have trace equal to one.

[30] For example, $t_k = k\Delta t$, where Δt is the time between samples, and k is an integer indicating discrete time values.

[31] The operator $\Gamma(u,y)$ should preserve self-adjointness and positivity.

[32] The notation $\mathrm{argmin}_{u\in\mathbf{U}}\,f(u)$ means the subset of

values from $\mathbf{U}$ minimizing f.
[33] We denote the value of B at $\hat{\omega}$ by $\langle \hat{\omega}, B \rangle$, extending the notation (7). The (generalized) adjoint $R^\dagger$ of R is defined by $\langle \hat{\omega}, R^\dagger B \rangle = \langle R\hat{\omega}, B \rangle$ for all $\hat{\omega}$ and all B.
[34] We require that $\mathbf{Q}$ is absolutely continuous with respect to $\mathbf{P}$.
[35] We assume that the true distribution $\mathbf{P}_{true}$ is absolutely continuous with respect to the nominal distribution $\mathbf{P}_{nom}$.