

Optimality of feedback control strategies for qubit purification

Author

Wiseman, Howard M, Bouten, Luc

Published

2008

Journal Title

Quantum Information Processing

DOI

[10.1007/s11128-008-0075-8](https://doi.org/10.1007/s11128-008-0075-8)

Rights statement

© 2008 Springer-Verlag. This is the author-manuscript version of this paper. Reproduced in accordance with the copyright policy of the publisher. The original publication is available at www.springerlink.com

Downloaded from

<http://hdl.handle.net/10072/22530>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Optimality of feedback control strategies for qubit purification

Howard M. Wiseman^{1,2} and Luc Bouten³

¹*Centre for Quantum Computer Technology, Australia*

²*Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111 Australia*

³*Physical Measurement and Control 266-33, California Institute of Technology, Pasadena, CA 91125*

(Dated: February 1, 2008)

Recently two papers [K. Jacobs, Phys. Rev. A **67**, 030301(R) (2003); H. M. Wiseman and J. F. Ralph, New J. Physics **8**, 90 (2006)] have derived control strategies for rapid purification of qubits, optimized with respect to various goals. In the former paper the proof of optimality was not mathematically rigorous, while the latter gave only heuristic arguments for optimality. In this paper we provide rigorous proofs of optimality in all cases, by applying simple concepts from optimal control theory, including Bellman equations and verification theorems.

PACS numbers: 03.67.-a, 03.65.Ta, 02.50.Tt, 02.50.Ey

I. INTRODUCTION

In the absence of decoherence, monitoring (that is, continuous-in-time weak measurement) of a qubit observable such as σ_z will eventually purify the qubit. However, the process of purification for finite times can be affected by applying Hamiltonian controls. Based upon results for discrete (but non-projective) measurements, Jacobs [1, 2] derived a strategy to *maximize the average purity* of the qubit at any given time. His strategy requires the controller to maintain the qubit state with $\langle \sigma_z \rangle = 0$. Wiseman and Ralph [3] considered two different goals: first to *maximize the fidelity* of the qubit with any given state at any given time, and second, to *minimize the average time* required to attain a given level of purity (or fidelity). For the first goal, they argued that Jacobs' strategy (with a trivial modification) was again optimal. For the second goal, they considered the strategy which maintains $\langle \sigma_y \rangle = \langle \sigma_z \rangle = 0$, which is a no-control strategy provided that these conditions are met initially (and that any Hamiltonian not proportional to σ_z can be nullified). They showed that this outperforms Jacobs' control strategy, and argued that it was the optimal strategy. However, none of these claims of optimality were proven using rigorous results from continuous time stochastic control theory.

Control strategies derived from heuristic arguments may well fail to attain optimality. Some examples include continuous quantum error correction in Ref. [4] and gradual discrimination of non-orthogonal states in Ref. [5] (in which Jacobs is well aware of the non-optimality). However, if a simple feedback strategy is conjectured to be optimal, then the truth of the conjecture can be demonstrated straightforwardly using standard techniques from optimal control theory — the so-called *verification theorems*. The purpose of this paper is to highlight this fact, by showing that both of the above strategies (that of Jacobs, and that of Wiseman and Ralph), are indeed optimal for the different goals as claimed. We introduce the system in Sec. II, as well as reviewing some terminology and results from optimal control theory. In Sec. III

we discuss and refine Jacob's problem, and prove the optimality of his strategy. In Sec. IV we do likewise for Wiseman and Ralph's problem, and we conclude with a brief discussion in Sec. V.

II. THE SYSTEM

In continuous measurement theory a basic role is played by the quantum filtering equation [6, 7, 8], also known as a stochastic master equation [9, 10]. This can be obtained from an underlying quantum Markov model describing the coupling of the system to an external environment. The filtering equation is a stochastic differential equation whose solution at time t is a (random) state matrix ρ_t , such that the quantity $\text{Tr}[X\rho_t]$ corresponds to the conditional expectation of an arbitrary operator X with respect to the observations up to time t . In a more traditional quantum measurement perspective (see for example Ref. [10]), ρ_t is known as the conditional system state, and represents the observer's beliefs about the system. The filtering equation also plays a fundamental role in quantum feedback control, because most control problems separate into a filtering step and a classical control problem for the filtering equation [6, 8]. That is, the optimal feedback control algorithms make use only of ρ_t and have thus been called state-based [11] and Bayesian [12] in the physics literature. In control theory such strategies are known as separated strategies.

The goal of the particular control problems studied in [1, 2, 3] is to maximize the rate of purification of the conditional state ρ_t in a suitable sense. We consider precisely the same system as in [1, 2, 3]: monitoring one operator ($\hat{\sigma}_z$) for a qubit with the ability to control the Hamiltonian. In a solid-state context this monitoring could be realized by a quantum point contact [13], and in an optical context it could arise from an off-resonant coupling of the qubit to the electromagnetic field which is observed continuously in time via homodyne detection [14]. Here we denote by $\sigma_x, \sigma_y, \sigma_z$ the well-known Pauli matrices. We assume without loss of generality that the initial state ρ_0 satisfies $\text{Tr}[\rho_0\sigma_y] = 0$. In that case we

only need a control Hamiltonian proportional to σ_y . The quantum filtering equation for this model is

$$d\rho_t = -i\Omega_t[\sigma_y/2, \rho_t] dt + (\sigma_z \rho_t \sigma_z - \rho_t) dt + (\sigma_z \rho_t + \rho_t \sigma_z - 2\text{Tr}[\sigma_z \rho_t] \rho_t) dW_t. \quad (2.1)$$

Here Ω_t is the control input and W_t is the innovations Wiener process [6, 7, 8]. It is related to the measurement record Y_t on which ρ_t is conditioned by

$$dY_t = 2\text{Tr}[\rho_t \hat{\sigma}_z] dt + dW_t. \quad (2.2)$$

Rewriting Eq. (2.1) in terms of the Bloch vector components $(x_t, y_t, z_t) = \text{Tr}[(\sigma_x, \sigma_y, \sigma_z)\rho_t]$, we obtain

$$dz_t = 2(1 - z_t^2) dW_t - \Omega_t x_t dt, \quad (2.3)$$

$$dx_t = -2x_t dt - 2z_t x_t dW_t + \Omega_t z_t dt, \quad (2.4)$$

and $y_t = 0$. We now introduce polar coordinates $z_t = R_t \cos \theta_t$ and $x_t = R_t \sin \theta_t$. It turns out to be convenient in calculations to work with $S_t = 1 - R_t^2$ instead of R_t . S_t is sometimes known as the linear entropy of the qubit. In these variables the filtering equation reads

$$dS_t = -4S_t \left\{ [1 - (1 - S_t)u_t^2] dt + \sqrt{1 - S_t} u_t dW_t \right\}, \quad (2.5)$$

$$d\theta_t = \Omega_t dt + f(\theta_t, S_t) dt + g(\theta_t, S_t) dW_t. \quad (2.6)$$

Here f and g are functions whose explicit form we will not need and we have written u_t for $\cos \theta_t$.

The equations (2.5) and (2.6) are already too complicated to be amenable to optimal control. In order to proceed, the papers [1, 2, 3] assume that, rather than controlling Ω_t , one can directly control u_t at every time to take any value in the interval $[-1, 1]$. This could be seen as a singular limit of the above system, in the sense that an infinitely strong driving field Ω_t would ostensibly allow one to set the angle θ_t to any desired value. This type of assumption is commonly used to reduce open-loop optimal control problems in NMR to explicitly computable form [17]. In the present case, it is less clear that this is a reasonable approximation, as singular controls are incompatible with the assumptions used to derive the quantum filtering equation [6, 8]. Our goal here, however, is to demonstrate optimality of the problems ultimately considered in [1, 2, 3], so we will adopt their assumptions in the following.

Before we proceed to the problems considered in [1, 2, 3], we review some terminology. A control strategy of the form $u_t = u(t, S_t)$, where $u(t, s)$ is a deterministic function, is called a *Markov* control. It is easily verified using the Itô rule that for any function $f(t, s)$ that is C^1 in t and C^2 in s , we have

$$\mathbf{E}_s(f(t, S_t)) = \mathbf{E}_s \int_0^t \left[\frac{\partial f}{\partial t'}(t', S_{t'}) + \mathcal{L}^{u'} f(t', S_{t'}) \right] dt' + f(0, s), \quad (2.7)$$

which is known as *Dynkin's formula*. Here $\mathbf{E}_s(X)$ denotes $\mathbf{E}(X|S_0 = s)$, the expectation of X given the initial condition $S_0 = s$. Moreover, for any $v \in [-1, 1]$, we have defined the differential operator \mathcal{L}^v by

$$\mathcal{L}^v f(s) = -4s \frac{\partial f(s)}{\partial s} + v^2 \mathcal{D}f(s), \quad (2.8)$$

where

$$\mathcal{D}f(s) = 8s^2(1 - s) \frac{\partial^2 f(s)}{\partial s^2} + 4s(1 - s) \frac{\partial f(s)}{\partial s}. \quad (2.9)$$

III. JACOBS' PROBLEM

Jacobs [1, 2] considered the following control problem: Fix an arbitrary terminal time $T > 0$, and find a feedback control u_t (with values in $[-1, 1]$) that maximizes the purity of the qubit at time T . Specifically, Jacobs chose a goal of minimizing the cost function $\mathbf{E}_s(S_T)$, the expected value of the random variable S_T , for $S_0 = s$. This choice of cost function was made “to keep the calculations tractable” [1]. While this choice gives a well-defined control problem with a well-behaved solution (as discussed below), it was pointed out by Wiseman and Ralph [3] that the cost function $\mathbf{E}(S_T)$ lacks a clear physical motivation. The reason is that it is a nonlinear function of the quantum state, and hence does not correspond to the expected outcome of any physical process that can be performed on the qubit.

Say the qubit is being purified for some particular purpose for which ideally the qubit would be in a particular pure state $|\phi\rangle$, as discussed in Ref. [2]. Then the cost function should be based on the performance of the qubit, as can be empirically determined independent of the preparation procedure (for example, by a second party who will impose the cost on the preparer). Any such cost function depends only upon $\mathbf{E}\langle\phi|\rho_T|\phi\rangle$, the expected fidelity of the state $\mathbf{E}[\rho_T]$ with the fiducial state $|\phi\rangle$. For example, for $\sigma_z|\phi\rangle = +|\phi\rangle$, the fidelity, which we would wish to maximize, is $[1 + \mathbf{E}(R_T \cos \theta_T)]/2$. Since we are working under the assumption that θ_t can be set arbitrarily, all that we require is to maximize $\mathbf{E}(R_T)$. That is, the best motivated cost function for qubit purification is $\mathbf{E}_s(-\sqrt{1 - S_T})$ rather than $\mathbf{E}_s(S_T)$.

We note that Jacobs' argument for the optimality of his control strategy with respect to his cost function relies upon the linearity of that cost function in S and hence does not apply to the cost function $\mathbf{E}_s(-\sqrt{1 - S_T})$. Nevertheless, below we prove rigorously that Jacobs' strategy is optimal for both cost functions, and the calculation is essentially the same in both cases. We do this by considering the general cost function

$$J[u, s] = \mathbf{E}_s(F(S_T^u)), \quad (3.1)$$

where F can be given either by $F(x) = x$ or by $F(x) = -\sqrt{1 - x}$. Here the superscript u on S_t^u means that S_t^u is propagated by Eq. (2.5) under the control strategy u .

Before presenting the theorem we use to prove that Jacob's strategy is optimal, we first give a heuristic derivation of the so-called *Bellman equation*, see e.g. [16]. This derivation provides some insight into the problem, and the Bellman equation will reappear in the theorem. In situations (unlike the present case) where one did not have a candidate optimal control strategy, the Bellman equation may also suggest such a candidate.

To derive the Bellman equation, consider a given Markov strategy u , and define the *cost-to-go* at time t given that $S_t = s$ by

$$J[u, s, t] = \mathbf{E} (F(S_T^u) | S_t = s), \quad (3.2)$$

so that $J[u, s] = J[u, s, 0]$. Now we assume that there exists an *optimal* control strategy u_t^* ; that is, a strategy that minimizes the cost function. We also assume that it is *Markov*; that is, $u_t^* = u^*(t, S_t)$, where $u^*(t, s)$ is a deterministic function. From these assumptions, we can define the *value function* to be the *optimal cost-to-go*:

$$V(t, s) = J[u^*, s, t]. \quad (3.3)$$

Since u_t^* is a strategy that minimizes the cost over the full interval $[0, T]$ it seems obvious that it will also minimize the cost-to-go over the interval $[t, T]$; that is,

$$V(t, s) = J[u^*, s, t] = \inf_u J[u, s, t], \quad (3.4)$$

where the infimum is over all Markov control strategies. This is a form of the *the principle of optimality* and can indeed be made rigorous, see e.g. [16]. Since at this point we are merely arguing heuristically, we do not need a rigorous proof. The idea behind the proof is, however, simple to explain. If a better strategy u exists on the interval $[t, T]$, we would simply use u^* up to time t and would proceed with the better strategy u from time t onwards and obtain a lower cost over the interval $[0, T]$. This contradicts optimality of u^* and hence such a u can not exist.

The principle of optimality leads to the following recursion for the value function (optimal cost-to-go)

$$V(t, s) = \inf_u \mathbf{E} (V(t', S_{t'}^u) | S_t^u = s), \quad V(T, s) = F(s), \quad (3.5)$$

where $0 \leq t \leq t' \leq T$ and the infimum is over the Markov control strategies on the interval $[t, t']$. To see this, note that $\mathbf{E}(V(t', S_{t'}^u) | S_t^u = s)$ is the cost-to-go of a control strategy that follows u on the interval $[t, t']$ and u^* on the interval $[t', T]$. If we take $t' = t + dt$, expand $V(t', S_{t'})$ according to Itô's rule, and re-arrange the terms, then we arrive at the following *Bellman equation*

$$\begin{aligned} \frac{\partial V}{\partial t}(t, s) &= - \inf_{v \in [0, 1]} \mathcal{L}^v V(t, s) \\ &= 4s \frac{\partial V}{\partial s}(t, s) - \inf_{v \in [0, 1]} v^2 \mathcal{D}V(t, s), \end{aligned} \quad (3.6)$$

with terminal condition $V(T, s) = F(s)$. Since u^* is an optimal strategy, the infimum is attained at $v = u^*(t, s)$.

Moreover, it is clear that $u^*(t, s) \in \{0, 1\}$, so that the Bellman equation can be rewritten as

$$\frac{\partial V}{\partial t} - 4s \frac{\partial V}{\partial s} + \min \{0, \mathcal{D}V\} = 0. \quad (3.8)$$

Jacob's strategy chooses $u_t = 0$ for $0 \leq t \leq T$. This corresponds to instantaneously rotating the qubit onto the x -axis at $t = 0$, and then actively maintaining it there. We now show that it yields a solution to the Bellman equation, as it should if it were an optimal strategy. For $u_t = 0$ Eq. (2.5) gives the deterministic solution $S_t = S_0 e^{-4t}$, so that $S_T = S_t e^{-4(T-t)}$. If this strategy were optimal then $V(t, s)$, the optimal cost-to-go, should be constant at $F(S_T)$, and hence should be given by

$$V(t, s) = F(s e^{-4(T-t)}). \quad (3.9)$$

For this choice it is simple to verify that for both cases $F(x) = x$ and $F(x) = -\sqrt{1-x}$, and for all $s \in [0, 1]$ and $t \in [0, T]$, that

$$\mathcal{D}V \geq 0, \quad (3.10)$$

$$\frac{\partial V}{\partial t} = 4s \frac{\partial V}{\partial s}. \quad (3.11)$$

Thus Eq. (3.9) is indeed a solution of Eq. (3.8).

We emphasize that this result does not prove that Jacob's strategy is optimal. A strategy that is optimal will give a solution to the Bellman equation, but a solution of the Bellman equation does not necessarily correspond to an optimal strategy. However, it is simple to check whether it does correspond to an optimal strategy, as the theorem below shows. Instead of assuming the existence of an optimal strategy u^* , we now consider a solution V to the Bellman equation Eq. (3.6) with terminal condition $V(T, s) = F(s)$ and we *define* a candidate optimal Markov strategy u^* by

$$u^*(t, s) \in \underset{v \in [0, 1]}{\operatorname{argmin}} \mathcal{L}^v V(t, s), \quad (3.12)$$

where the right-hand-side means the value of $v \in [0, 1]$ that minimizes $\mathcal{L}^v V(t, s)$.

Theorem 1, which is a special case of a standard result in control theory [15, Thm. 8.1], provides a set of criteria that guarantees that a particular control strategy u_t is optimal. We apply the theorem to our candidate strategy with V as in Eq. (3.9). Note that criteria 2 and 3 together are equivalent to V satisfying the Bellman equation and u being defined as in Eq. (3.12). Criterion 4 is the terminal condition for the Bellman equation. Criterion 1 is a smoothness condition on V , which from Eq. (3.9) is readily seen to hold for the two functions F we are considering. The Theorem therefore allows us to conclude that our candidate strategy is indeed optimal.

Theorem 1 (Verification theorem) *Suppose there exists a Markov control $u : [0, T] \times [0, 1] \rightarrow [-1, 1]$ and a function $V : [0, T] \times [0, 1] \rightarrow \mathbb{R}$ such that*

1. $V(t, s)$ is C^1 on $[0, T]$ and C^2 in s .
2. For all $t \in [0, T]$ and $s \in [0, 1]$,

$$\frac{\partial V}{\partial t}(t, s) + \mathcal{L}^{u(t,s)}V(t, s) = 0.$$

3. For all $t \in [0, T]$, $v \in [-1, 1]$ and $s \in [0, 1]$,

$$\frac{\partial V}{\partial t}(t, s) + \mathcal{L}^vV(t, s) \geq 0.$$

4. For all $s \in [0, 1]$, V satisfies the terminal condition

$$V(T, s) = F(s).$$

Then the Markov control strategy u is optimal, i.e. $J[u, s] \leq J[\tilde{u}, s]$ for all $s \in [0, 1]$ and any control \tilde{u}_t taking values in $[-1, 1]$, and moreover $V(0, s) = J[u, s]$.

Proof. Let \tilde{u}_t be any control strategy. Then using Dynkin's formula and the first, third and fourth conditions in the verification theorem, we obtain

$$\begin{aligned} V(0, s) &= \mathbf{E}_s(V(T, S_T)) \\ &\quad - \mathbf{E}_s \int_0^T \left[\frac{\partial V}{\partial t}(t, S_t) + \mathcal{L}^{\tilde{u}_t}V(t, S_t) \right] dt \\ &\leq \mathbf{E}_s(V(T, S_T)) = \mathbf{E}_s(F(S_T)) = J[\tilde{u}, s]. \end{aligned}$$

Using the second condition, we similarly conclude that $V(0, s) = J[u, s]$. Hence $J[u, s] \leq J[\tilde{u}, s]$ for any \tilde{u} . \square

IV. WISEMAN AND RALPH'S PROBLEM

As well as the variation on Jacob's problem discussed above, Wiseman and Ralph [3] also considered a quite different control problem. Their problem fixes some threshold $h \in (0, 1)$, and seeks a feedback control u_t that minimizes the average time at which the linear entropy first hits the threshold value of h . That is, the cost function is

$$J[u, s] = \mathbf{E}_s(\tau^u), \quad \tau^u = \inf \{t : S_t^u \leq h\}. \quad (4.1)$$

As before, the superscript u on S_t^u reminds us that S is propagated by Eq. (2.5) under the control strategy u . It is clear from the discussion at the start of the preceding section that, under our assumptions, this problem is equivalent to minimizing the time taken to prepare a given pure state (in the $y = 0$ plane) with fidelity $(1 + \sqrt{1-h})/2$.

From consideration of the stochastic evolution of $\log(1 - S_t)$, Wiseman and Ralph [3] claimed that the strategy $u = 1$ was optimal in the asymptotic ($h \ll 1$) limit. This strategy requires instantaneously rotating (if necessary) the qubit onto the z -axis at $t = 0$, then letting $\Omega_t = 0$ for $t > 0$. Here we verify that this is indeed

the optimal strategy for all h . We proceed as before, first finding the Bellman equation and exhibiting a solution using the strategy proposed by Wiseman and Ralph. Then we turn the argument around and use the solution to the Bellman equation to prove rigorously that the candidate strategy is indeed optimal.

To derive the Bellman equation, we again start by assuming the existence of an optimal Markov strategy u_t^* . The value function, the optimal cost-to-go, is given by

$$V(t, s) = \mathbf{E} \left((\tau^{u^*} - t)^+ \middle| S_t^{u^*} = s \right), \quad t \geq 0, \quad 1 \geq s \geq h. \quad (4.2)$$

Here $(\tau^{u^*} - t)^+ = \max\{0, (\tau^{u^*} - t)\}$, the time left after time t until the threshold h is first hit. Note that it follows immediately from the definition that $V(t, s) = 0$ for all $t \geq 0$ and $s \leq h$. The principle of optimality then yields the following recursion for the value function (where $t' \wedge \tau^u$ means $\min\{t', \tau^u\}$)

$$V(t, s) = \inf_u \mathbf{E} \left((t' \wedge \tau^u) - t + V(t', S_{t' \wedge \tau^u}^u) \middle| S_t^u = s \right). \quad (4.3)$$

Here $0 \leq t \leq t'$, $h \leq s \leq 1$ and the infimum is over all Markov control strategies on the interval $[t, t']$. To see this, note that $\mathbf{E} \left((t' \wedge \tau^u) - t + V(t', S_{t' \wedge \tau^u}^u) \middle| S_t^u = s \right)$ is the cost-to-go of a control strategy that follows u on the interval $[t, t']$ and u^* after time t' . If t' is smaller than τ^u then we incur a cost $t' - t$ on the interval $[t, t']$ plus the optimal cost-to-go if we are at $S_{t'}^u$ at time t' . If t' is greater than τ^u then we incur a cost $\tau^u - t'$ over the interval $[t, t']$. Note that $V(t', S_{\tau^u}) = V(t', h) = 0$.

If we take $t' = t + dt$, expand V according to Itô's rule and re-arrange the terms, then Eq. (4.3) leads to the following Bellman equation

$$-\frac{\partial V}{\partial t} = 1 + \inf_{v \in [0, 1]} \mathcal{L}^v V(t, s) \quad (4.4)$$

$$= 1 - 4s \frac{\partial V}{\partial s}(t, s) + \inf_{v \in [0, 1]} v^2 \mathcal{D}V(t, s), \quad (4.5)$$

with boundary condition $V(t, h) = 0$. As before, the minimand is linear in v^2 so this can be rewritten as

$$-\frac{\partial V}{\partial t} = 1 - 4s \frac{\partial V}{\partial s} + \min\{0, \mathcal{D}V\}. \quad (4.6)$$

Having obtained the Bellman equation, we now need to find a solution. Wiseman and Ralph [3] show by explicit calculation that for $u_t = 1$ the expected cost-to-go is

$$J[u, s, t] = \mathbf{E} \left((\tau^u - t)^+ \middle| S_t^u = s \right) \quad (4.7)$$

$$= [\gamma(h) - \gamma(s)]/4, \quad (4.8)$$

where for convenience we have defined, for $x \in [0, 1]$,

$$\gamma(x) = \sqrt{1-x} \operatorname{atanh}(\sqrt{1-x}). \quad (4.9)$$

Note that $J[u, s, t]$ is independent of t . This is to be expected, since the control strategy $u_t = 1$ is not time dependent, so the time left until we hit the threshold h ,

starting at s , does not depend on when s was actually reached. Now, if Wiseman and Ralph's proposed strategy is optimal, then from the heuristic arguments above, taking

$$V(s) = [\gamma(h) - \gamma(s)]/4 \quad (4.10)$$

should satisfy the Bellman equation Eq. (4.6).

It is easily verified that for all $s \in [h, 1]$ Eq. (4.10) satisfies

$$4s \frac{\partial V}{\partial s}(s) - \mathcal{D}V(s) = 1, \quad (4.11)$$

$$\mathcal{D}V(s) \leq 0, \quad (4.12)$$

as well as $V(h) = 0$. Thus the strategy of Wiseman and Ralph does indeed give a solution of the Bellman equation Eq. (4.6). Moreover, it is clear that the infimum in Eq. (4.6) is attained at $v = 1$. Thus if we turn the problem around and *define* a candidate optimal control strategy via Eq. (3.12) using the solution Eq. (4.10) to the Bellman equation, this gives the Wiseman-Ralph strategy $u^*(t, s) = 1$ for all $t \geq 0$ and $s \in [h, 1]$. Note that we do not need to worry about $s \in [0, h)$ —in this case $\tau = 0$, which can not be improved upon!

We emphasize again that the above argument only establishes that the Wiseman-Ralph strategy is a candidate optimal strategy. To verify that it is optimal we proceed as before, via a second verification theorem. This is another special case of a standard result in control theory [19, Thm. 3.1]. We apply the Theorem to our candidate strategy with V as in Eq. (4.10). Note again that criteria 2 and 3 together are equivalent to V satisfying the Bellman equation and u being defined as in Eq. (3.12). Criterion 4 is the terminal condition for the Bellman equation. Criterion 1 is a smoothness condition on V , which from Eq. (4.10) can be verified to hold. This Theorem therefore allows us to conclude that the Wiseman-Ralph strategy is indeed optimal.

Theorem 2 (Verification theorem) *Suppose there exist a stationary Markov control $u : [h, 1] \rightarrow [-1, 1]$ such that $\mathbf{E}_s(\tau^u) < \infty$, and a function $V : [h, 1] \rightarrow \mathbb{R}$ such that*

1. $V(s)$ is C^2 in s .
2. For all $s \in [h, 1]$, $\mathcal{L}^{u(s)}V(s) + 1 = 0$.
3. For all $v \in [-1, 1]$ and $s \in [h, 1]$, $\mathcal{L}^vV(s) + 1 \geq 0$.
4. $V(h) = 0$.

Then the stationary Markov control strategy u is optimal, i.e. $J[u, s] \leq J[\tilde{u}, s]$ for any control \tilde{u}_t taking values in $[-1, 1]$, and moreover $V(s) = J[u, s]$.

Proof. Let \tilde{u}_t be any control strategy. Just as in the proof of the previous theorem, we would like to apply Dynkin's formula, but with the *random* terminal time $\tau^{\tilde{u}}$ replacing T . This is possible only if $\mathbf{E}_s(\tau^{\tilde{u}}) < \infty$ [18]. However, note that if $\mathbf{E}_s(\tau^{\tilde{u}})$ is infinite, then it is immediate that

$$J[u, s] = \mathbf{E}_s(\tau^u) < \mathbf{E}_s(\tau^{\tilde{u}}) = J[\tilde{u}, s].$$

That is, we can safely restrict our attention to control strategies \tilde{u} for which $\mathbf{E}_s(\tau^{\tilde{u}})$ is finite. Therefore we can now safely apply Dynkin's formula using the random terminal time $\tau^{\tilde{u}}$ [18]. This gives, using the conditions of the theorem,

$$\begin{aligned} V(s) &= \mathbf{E}_s \left[V(S_{\tau^{\tilde{u}}}^{\tilde{u}}) - \int_0^{\tau^{\tilde{u}}} \mathcal{L}^{\tilde{u}_t} V(S_t^{\tilde{u}}) dt \right] \\ &\leq \mathbf{E}_s \left[V(S_{\tau^{\tilde{u}}}^{\tilde{u}}) + \int_0^{\tau^{\tilde{u}}} dt \right] = \mathbf{E}_s (V(S_{\tau^{\tilde{u}}}^{\tilde{u}}) + \tau^{\tilde{u}}). \end{aligned}$$

But note that $S_{\tau^{\tilde{u}}}^{\tilde{u}} = h$ by the definition of $\tau^{\tilde{u}}$; so

$$V(s) \leq \mathbf{E}_s (V(h) + \tau^{\tilde{u}}) = \mathbf{E}_s(\tau^{\tilde{u}}) = J[\tilde{u}, s].$$

Using the remaining (second) condition, we similarly conclude that $V(s) = J[u, s]$. Hence $J[u, s] \leq J[\tilde{u}, s]$ for any \tilde{u} . \square

V. CONCLUSION

We have highlighted a simple technique from optimal control theory to verify optimality of candidate control strategies. This paper was prompted by the absence of such arguments in the current physics literature (but see [8]). In particular, we have verified the optimality of previously proposed control strategies for two problems of current interest [1, 2, 3, 20, 21, 22] in nonlinear quantum feedback control.

Optimal control theory is notoriously difficult for nonlinear systems, and the reader might be surprised at our explicit computations. This simplicity is only made possible by the rather "violent" assumption that we can directly control $\cos \theta_t$ rather than Ω_t . A much more realistic control problem would be to find an optimal Ω_t under additional finite energy constraints on the control. Such problems, however, are not analytically tractable. Various strategies with finite controls were investigated numerically for a particular (solid-state) setting in Ref. [21], with no claims of optimality.

Acknowledgments

We thank Ramon van Handel for motivating and guiding this work, and for many stimulating discussions. HMW acknowledges support by the Australian Research Council and the State of Queensland, and LB acknowledges support by the ARO under grant number W911NF-06-1-0378.

-
- [1] K. Jacobs, Phys. Rev. A **67** 030301(R) (2003).
 - [2] K. Jacobs, Proc. SPIE **5468** 355 (2004).
 - [3] H. M. Wiseman and J. F. Ralph, New J. Phys. **8**, 90 (2006).
 - [4] C. Ahn, A. C. Doherty, and A. J. Landahl, Phys. Rev. A **65**, 042301 (2002).
 - [5] K. Jacobs, Quant. Inf. and Comp. **7**, 127 (2006).
 - [6] V. P. Belavkin, pp. 310–324 in Proceedings XXIV Karpacz winter school (R. Guelerak and W. Karwowski, eds.), World Scientific, Singapore, (1988).
 - [7] L. Bouten, M. Guta and H. Maassen, J. Phys. A **37**, 3189 (2004).
 - [8] L. Bouten and R. Van Handel, *math-ph/0511021* (2005).
 - [9] H. M. Wiseman and G. J. Milburn, Phys. Rev. A **47**, 1652 (1993).
 - [10] H. M. Wiseman, Quantum Semiclass. Opt. **8**, 205 (1996).
 - [11] A. C. Doherty, S. Habib, K. Jacobs, H. Mabuchi, S. M. Tan Phys. Rev. A **62**, 012105 (2000).
 - [12] H. M. Wiseman, S. Mancini, and J. Wang, Phys. Rev. A **66**, 013807 (2002).
 - [13] Hsi-Sheng Goan, G. J. Milburn, H. M. Wiseman, and He Bi Sun, Phys. Rev. B **63**, 125326 (2001).
 - [14] R. van Handel, J. K. Stockton, and H. Mabuchi, J. Opt. B: Quantum Semiclass. Opt. **7**, S179 (2005).
 - [15] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions* (Springer, Berlin, 2006)
 - [16] H.J. Kushner, *Introduction to stochastic control* (Holt, Rhinehart and Winston Inc., 1971)
 - [17] N. Khaneja, *et al.* J. Mag. Res. **162**, 311 (2003).
 - [18] B. Oksendal, *Stochastic Differential Equations, fifth ed.* (Springer, Berlin, 2000)
 - [19] B. Oksendal and A. Sulem, *Applied stochastic control of jump diffusions*, (Springer, Berlin, 2005)
 - [20] J. Combes and K. Jacobs, Phys. Rev. Lett. **96**, 010504 (2006)
 - [21] E. J. Griffith, C. D. Hill, J. F. Ralph, H. M. Wiseman, and K. Jacobs, Phys. Rev. B **75**, 014511 (2007)
 - [22] C. D. Hill and J. F. Ralph *quant-ph/0610156* (2006).