Feedback policies for measurement-based quantum state manipulation

Shuangshuang Fu* and Guodong Shi*
Research School of Engineering, Australian National University, Canberra, ACT 0200, Australia

Alexandre Proutiere†
School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden

Matthew R. James†
ARC Centre for Quantum Computation and Communication Technology, Research School of Engineering, Australian National University, Canberra, ACT 0200, Australia

(Received 24 October 2014; published 18 December 2014)

In this paper, we propose feedback designs for manipulating a quantum state to a target state by performing sequential measurements. In light of Belavkin’s quantum feedback control theory, for a given set of (projective or nonprojective) measurements and a given time horizon, we show that finding the measurement selection policy that maximizes the probability of successful state manipulation is an optimal control problem for a controlled Markovian process. The optimal policy is Markovian and can be solved by dynamical programming. Numerical examples indicate that making use of feedback information significantly improves the success probability compared to classical scheme without taking feedback. We also consider other objective functionals including maximizing the expected fidelity with the target state as well as minimizing the expected arrival time. The connections and differences among these objectives are also discussed.

DOI: 10.1103/PhysRevA.90.062328
PACS number(s): 03.67.Ac

I. INTRODUCTION

One fundamental difference between classical and quantum mechanics is the unavoidable back-action of quantum measurement. On the one hand, this back-action is generally thought to be detrimental for the implementation of effective quantum control. On the other hand, it also provides us one possibility to use the change caused by the measurement as a new route to manipulate the state of the system [1,2]. A basic problem in quantum physics and engineering is how to drive a quantum system to a desired target state. There have been studies on the preparation of a given target state from a given initial state using sequential (projective or nonprojective) measurements in the last few years [3–7].

A quantum measurement $E$ is described by a collection of measurement operators

$$\{M_E(m)\}_{m \in \mathcal{Y}},$$

where $\mathcal{Y}$ is an index set for measurement outcomes and the measurement operators satisfy

$$\sum_{m \in \mathcal{Y}} M_E(m)\dagger M_E(m) = I.$$

Suppose we perform the quantum measurement $E$ on density operator $\rho$; the probability of obtaining result $m \in \mathcal{Y}$ is $\text{tr}[M_E(m)\rho M_E(m)\dagger]$, and when $m \in \mathcal{Y}$ occurs, the postmeasurement state of the quantum system becomes

$$M_E^m(\rho) = \frac{M_E(m)\rho M_E(m)\dagger}{\text{tr}[M_E(m)\rho M_E(m)\dagger]}.$$

If we are unaware of the measurement result, the unconditional state of the quantum system after the measurement can be expressed as

$$\mathcal{M}_E(\rho) = \sum_{m \in \mathcal{Y}} M_E(m)\rho M_E(m)\dagger.$$

If $\{M_E(m)\}_{m \in \mathcal{Y}}$ are orthogonal projectors, i.e., the $M_E(m)$ are Hermitian and $M_E(l)M_E(m) = \delta_{lm}M_E(m)$, $E$ is a projective measurement. The idea of quantum state manipulation using sequential measurements [3–7] is as follows. By consecutively performing the measurements $E_1, \ldots, E_N$, the unconditional state for a quantum system with initial state $\rho_0$ can be expressed as

$$\rho_N^u = \mathcal{M}_{E_N} \circ \mathcal{M}_{E_{N-1}} \circ \cdots \circ \mathcal{M}_{E_1}(\rho_0).$$

It has been shown, analytically or numerically, how to select the measurements $E_1, \ldots, E_N$ so that $\rho_N^u$ can asymptotically tend to a desired target state [3–7].

Making use of feedback information for quantum measurement and detection actually has a long history, which can be viewed as the dual problem of state manipulation. The “Dolinar’s receiver” proposes a feedback strategy for discriminating two possible quantum states with prior distribution with minimum probability of error [8]. The problem is known as the quantum detection problem and Helstrom’s bound characterizes the minimum probability of error for discriminating any two nonorthonormal states [9]. Quantum detection is to identify uncertain quantum states via projective measurements; while the considered quantum state projection is to manipulate a certain quantum state to a certain target, again via projective measurements. The Dolinar’s scheme follows a similar structure that measurement is selected based on previous measurement results on different segments of the pulse, and was recently realized experimentally [10]. See [11]
In this paper, we propose a feedback design for quantum state manipulation via sequential measurements. For a given set of measurements and a given time horizon, we show that finding the policy of measurement selections that maximizes the probability of successful state manipulation can be solved by dynamical programming. Such derivation of the optimal policy falls to Belavkin’s quantum feedback theory [1]. Numerical examples are given which indicate that the proposed feedback policy significantly improves the success probability compared to classical policy by consecutive projections without taking feedback. In particular, the probability of achieving the target state $|1\rangle$ via feedback policy reaches $0.9968$ using merely ten measurements from initial state $|0\rangle$. Other optimality criteria are also discussed such as the maximal expected fidelity and the minimal arrival time, and some connections and differences among these different criteria are also discussed.

The remainder of the paper is organized as follows. In the first part of Sec. II, we revisit a simple example of reaching $|1\rangle$ from $|0\rangle$ using sequential projective measurements [7], and show how feedback policies work under which even a little bit of feedback can make a nontrivial improvement. The rest of Sec. II devotes to a rigorous treatment for the problem definition and for finding the optimal feedback policy from classical quantum feedback theory. Numerical examples are given there. Section III investigates some other optimality criteria and finally Sec. IV concludes the paper.

II. QUANTUM STATE MANIPULATION BY FEEDBACK

A. A simple example: Why feedback?

Consider now a qubit system, i.e., a two-dimensional Hilbert space. The initial state of the quantum system is $|0\rangle|0\rangle$, and the target state is $|1\rangle|1\rangle$. Given $T \geq 2$ projective measurements from the set

$$\mathcal{E} = \{E_i, i = 1, 2, \ldots, T\},$$

where $E_i = \{|\phi_i\rangle\langle\phi_i|, |\psi_i\rangle\langle\psi_i|\}$ with

$$|\phi_i\rangle = \cos\left(\frac{\pi i}{2T}\right)|0\rangle + \sin\left(\frac{\pi i}{2T}\right)|1\rangle$$

and

$$|\psi_i\rangle = -\sin\left(\frac{\pi i}{2T}\right)|0\rangle + \cos\left(\frac{\pi i}{2T}\right)|1\rangle.$$

Note that the choice of $E_i$ follows the optimal selection given in Ref. [7].

The strategy in Refs. [6,7] is simply to perform the $T$ measurements in turn from $E_1$ to $E_T$, where the choice of these measurements has been optimized. We call it an optimal feedback-free policy. The probability of successfully driving the state from $|0\rangle$ to $|1\rangle$ in $T$ steps under this optimal feedback-free strategy is denoted by $p(T)$. We can easily calculate that $p(3) \approx 0.56$ and $p(10) \approx 0.8$.

Let $T = 3$. We next show that even a bit of measurement feedback can improve the performance of the strategy significantly. SI. After the first measurement $E_1$ has been made, perform $E_2$ if the outcome is $|\psi_1\rangle$ for the second step, and follow the optimal feedback-free policy for all other actions. Following this scheme, it turns out that the probability of arriving at $|1\rangle$ in three steps becomes around 0.66, in contrast with $p(3) \approx 0.56$ under the optimal feedback-free scheme. The improvement in the probability of success comes from the fact that a feedback decision is made based on the information of the outcome of $E_1$ so that in SI a better selection of measurement is obtained between $E_2$ and $E_3$.

B. Optimal policy from quantum feedback control

We now present the solution to the optimal policy for the considered quantum state manipulation in light of the classical work of quantum feedback control theory derived by Belavkin [1] (also see [12,13] for a thorough treatment).

Consider a quantum system whose state is described by density operators over the qubit space. Let $\mathcal{E}$ be a given finite set of measurements serving as all feasible control actions. For each $E \in \mathcal{E}$, we write

$$E = \{M_E(y)\}_{y \in \mathcal{Y}},$$

where $\mathcal{Y}$ is a finite index set of measurement outputs and $M_E(y)$ is the measurement operator corresponding to outcome $y \in \mathcal{Y}$. Time is slotted with a horizon $N \geq 1$. The initial state of the quantum system is $\rho_0$, and the target state is assumed to be, for ease of presentation, $|1\rangle|1\rangle$.

For $0 \leq k \leq N - 1$, we denote by $u_k \in \mathcal{E}$ the measurement performed at time $k$ and the postmeasurement state after $u_k$ has been performed is denoted as $\rho_{k+1}$. Let $y_k \in \mathcal{Y}$ be the outcome of $u_k$. The measurement sequence $\{u_k\}_{k=0}^{N-1}$ is selected by a policy $\pi = \{\pi_k\}_{k=0}^{N-1}$, where each $\pi_k$ takes value in the set $\mathcal{E}$ such that $u_k = \pi_k(y_0, \ldots, y_{k-1}; u_0, \ldots, u_{k-1})$ can depend on all previous selected measurements and their outcomes for all $k = 0, \ldots, N - 1$. Here for convenience we have denoted $u_{-1} = y_{-1} = 0$.

We can now express the closed-loop evolution of $\{\rho_k\}_0^N$ as

$$\rho_{k+1} = M_{u_k}(\rho_k) = \frac{M_{u_k}(y_k)\rho_kM_{u_k}^\dagger(y_k)}{\text{tr}[M_{u_k}(y_k)\rho_kM_{u_k}^\dagger(y_k)]},$$

where $k = 0, \ldots, N - 1$. The distribution of $y_k$ is given by

$$P(y_k = y | u_k, \rho_k) = \text{tr}[M_{u_k}(y_k)\rho_kM_{u_k}^\dagger(y_k)],$$

where $k = 0, \ldots, N - 1$. Clearly $\{\rho_k\}_0^N$ defines a Markov chain.

We define $J_\pi(N) := P_\pi(\rho_N = |1\rangle|1\rangle)$ as the probability of successfully manipulating the quantum state to the target density matrix $|1\rangle|1\rangle$, where $P_\pi$ is the probability measure equipped with $\pi$. We also define the cost-to-go function

$$V(t,x) = \max_{\pi} P_\pi(\rho_N = |1\rangle|1\rangle | \rho_{t-1} = x)$$

for $t = 0, 1, \ldots, N$. The theory of controlled Markovian process [15,16] then applies, and the following conclusion holds.

062328-2
Proposition 1. The cost-to-go function $V(t,x)$ satisfies the following recursion:

$$V(t,x) = \max_{u \in E} \sum_{y \in Y} P(y|u,x) V(t-1, M^y_x(x)), \quad (3)$$

where $t = 1, \ldots, N$, with boundary condition $V(0,x) = 1$ if $x = |1\rangle\langle 1|$, and $V(0,x) = 0$ otherwise. The maximum arrival probability $J_\pi(N)$ is given by $J_\pi(N) = V(N, \rho_0)$. The optimal policy $\pi^* = \{\pi^*_k\}_{k=0}^{N-1}$ is Markovian, and is given by

$$\pi^*_k(\rho_k) = \arg \max_{u \in E} \sum_{y \in Y} P(y|u,\rho_k) V(N-k-1, M^y_x(\rho_k)) \quad (4)$$

for $k = 0, \ldots, N-1$.

Proposition 1 provides a way of recursively calculating the optimal policy $\pi^* = \{\pi^*_k\}_{k=0}^{N-1}$ with $V(0,x)$ as given in Eq. (1).

C. Numerical examples

We now investigate how the size of the available measurement set $E$ influences the successful arrival probability in $N$ steps under optimal feedback. In this case, the optimal arrival probability $J_\pi(N)$ is also a function of $T$, and we therefore rewrite $J_\pi(N) = J^T_\pi(N)$.

In Fig. 2, we plot $J^T_\pi(N)$, for $T = 10, 100, 1000$, respectively. The numerical results show that as $T$ increases, the $J^T_\pi(N)$ quickly tends to a limiting curve, suggesting the existence of some fundamental upper bound on the arrival probability in $N$ steps using sequential projections from an arbitrarily large measurement set.

![FIG. 1. The probabilities of successfully reaching |1⟩ from the initial state |0⟩ using optimal feedback-free policy πn and optimal feedback policy π*, respectively.](image)

III. MORE OPTIMALITY CRITERIA

In this section, we discuss two other useful optimality criteria, to maximize the expected fidelity with the target state, or to minimize the expected time it takes to arrive at the target state.

A. Maximal expected fidelity

Fidelity measures the closeness of two quantum states. Now that our target state $|1\rangle\langle 1|$ is a pure state, we have

$$\text{tr} \sqrt{|1\rangle\langle 1| \sigma \sqrt{|1\rangle\langle 1|}} = \sqrt{|1\rangle\langle 1| \sigma |1\rangle\langle 1|}.$$

Alternatively, we can consider the following objective functional:

$$\tilde{J}_\pi(N) = \mathbb{E}_{\pi} \{ |1\rangle\langle 1| \rho, \pi, |1\rangle\langle 1| \},$$

and the goal is to find a policy that maximizes $\tilde{J}_\pi(N)$.

<table>
<thead>
<tr>
<th>$\pi^n$</th>
<th>$k = 0$</th>
<th>$k = 1$</th>
<th>$k = 2$</th>
<th>$k = 3$</th>
<th>$k = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_1$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>$E_2$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>$E_3$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>$E_4$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>$E_5$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

TABLE I. The actions using optimal feedback-free strategy $\pi^n$ to prepare the target state $|1\rangle$, starting from $|0\rangle$, for $N = 5$. Here $E_i$ represents the measurement that the policy chooses, and * means that it is not possible to be in that state at the corresponding step.
For the two objective functionals $J_\pi(N)$ and $\tilde{J}_\pi(N)$, we denote their corresponding optimal policy as $\pi^\ast(N) = (\pi_k^\ast(N))_{k=0}^{N-1}$ and $\pi^\diamondsuit(N) = (\pi_k^\diamondsuit(N))_{k=0}^{N-1}$, respectively, where the time horizon $N$ is also indicated.

Let $\pi^\ast(N-1) + E_\ast$ be the policy that follows $\pi^\ast(N-1)$ for $k = 0, \ldots, N - 2$ and takes value $E_\ast$ for $k = N - 1$. Let $\rho_k^\ast$ be the unconditional density operator at step $k$ for $k = 0, \ldots, N - 1$. The following equations hold:

$$\tilde{J}_\pi(N-1) = \mathbb{E}_\pi[\{1|\rho_k^\ast|1\}] = \text{tr}(\rho_{N-1}^\ast|1\rangle\langle1|) = \mathbb{E}_\pi(\rho_{N-1}^\ast|1\rangle\langle1|),$$

for any $\pi = (\pi_k)_{k=0}^{N-2}$, where $\pi^\ast = \pi + E_\ast = (\pi_k)_{k=0}^{N-1}$ with $\pi_{N-1} = E_\ast$. As a result, the following relation holds between the optimal policies under the two objectives $J_\pi(N)$ and $\tilde{J}_\pi(N)$.

**Proposition 2.** It holds that $\max_\pi J_\pi(N) = \max_\pi \tilde{J}_\pi(N-1)$. In fact, $\pi^\ast(N) = \pi^\diamondsuit(N-1) + E_\ast$ with $E_\ast = \{0\}|0\rangle\langle1|\{1\}$.

The intuition behind Proposition 2 is that one would expect to get as close as possible to the target state at step $N = 1$, if one tends to successfully project onto the target state at step $N$. We also know from Proposition 2 that we can solve the maximal expected fidelity problem in $N$ steps by the solutions of maximizing the arrival probability in $N + 1$ steps.

Similarly, we can also find the optimal policy $\pi^\diamondsuit$ for the objective $\tilde{J}_\pi(N)$ using dynamical programming. Define the cost-to-go function $V(k,x)$ for $\tilde{J}_\pi(N)$ as

$$V(k,x) = \max_\pi \mathbb{E}_\pi[\{1|\rho_k|1\}|\rho_k = x]$$

for $k = 0, \ldots, N$. Then $V(k,x)$ satisfies the following recursive equation:

$$V(k,x) = \max_{u \in \mathcal{Y}} \sum_{y \in \mathcal{Y}} \mathbb{P}(y|u,x)V(k+1,M_u(x)),$$

for $k = 0, \ldots, N - 1$, with terminal condition

$$V(N,x) = \text{tr}(x|1\rangle\langle1|).$$

The optimal policy $\pi^\diamondsuit$ can be obtained by solving

$$\pi_k^\diamondsuit(\rho_k) = \arg \max_{u \in \mathcal{Y}} \sum_{y \in \mathcal{Y}} \mathbb{P}(y|u,\rho_k)V(k+1,M_u(\rho_k))$$

for $k = 0, \ldots, N - 1$. The maximal expected fidelity $J_{\pi^\diamondsuit}(N) = \tilde{V}(0,\rho_0)$.  

**B. Minimal Arrival Time**

In previous discussions the deadline $N$ plays an important role in the objective functionals as well as in their solutions. We now consider the case when the deadline is flexible, and we aim to minimize the average number of steps it takes to arrive at the target state. Now the control policy is denoted as $\pi = (\pi_k)_{k=0}^\infty$, where $\pi_k$ selects a measurement from the set $\mathcal{E}$. Associated with $\pi$, we define

$$\alpha_x := \inf_k \{\rho_k = |1\rangle\langle1|\}.$$

Note that $\alpha_x$ defines a stopping time (cf. [18]) associated with the random processes $\{\rho_k\}_{k=0}^\infty$ and we assume that $\pi$ is proper in the sense that

$$\mathbb{P}_\pi(\alpha_x < \infty) = 1.$$

We continue to introduce

$$J^\pi_\pi := \mathbb{E}_\pi[\alpha_x]$$

as the objective functional, which is the expected time it takes for the quantum state to reach the target $|1\rangle\langle1|$ following policy $\pi$. Minimizing $J^\pi_\pi$ is a *stochastic shortest path problem* [19].

We introduce $\mathcal{B}_\pi(x) := \inf_k \{\rho_k = |1\rangle\langle1|\}|\rho_0 = x\}$ and

$$V^\pi(x) = \min_\pi \mathbb{E}_\pi[\alpha_x].$$

The Markovian property of $\{\rho_k\}_{k=0}^\infty$ leads to the fact that the optimal policy $\pi^\ast$ is stationary in the sense that $\pi_k = \pi^\ast(x)$ for all $k$. The following conclusion holds by directly applying the results of [19].

**Proposition 3.** The cost-to-go function $V^\pi$ satisfies the following recursion:

$$V^\pi(x) = 1 + \min_{u \in \mathcal{Y}} \sum_{y \in \mathcal{Y}} \mathbb{P}(y|u,x)V^\pi(M_u^\pi(x)).$$

---

**Fig. 2.** (Color online) The probabilities of successfully reaching $|1\rangle$ from the initial state $|0\rangle$ using different sizes of measurement set by feedback strategy.
for all \( x \neq |1\rangle\langle 1| \), with boundary condition \( V^0(|1\rangle\langle 1|) = 0 \). The optimal policy \( \pi^\sharp \) is given by

\[
\pi^\sharp(x) = \arg \min_{u \in \mathcal{E}} \sum_{y \in \mathcal{Y}} \mathbb{P}(y | u, x) V^0 \left( \mathcal{M}_u^y(x) \right).
\]

(13)

The optimal \( J^\ast_{\pi^\sharp} \) is given by \( J^\ast_{\pi^\sharp} = V^3(\rho_0) \).

Proposition 3 shows that solving Eq. (12) yields the solution to the optimal policy \( \pi^\sharp \). The intuition behind this is that when there is no deadline constraint (i.e., successful projection must be realized within a given \( N \) steps), the optimal policy depends only on what the quantum state is before the projection. Moreover, technically it cannot be guaranteed that for any given measurement set \( \mathcal{E} \), there always exists at least one policy \( \pi \) under which \( J^\ast_{\pi} \) admits a finite number. However, some straightforward calculations indicate that for the set \( \mathcal{E} \) of projective measurements given in Eq. (1), finite \( J^\ast_{\pi} \) can always be achieved for a class of policies.

C. Numerical example: Minimal arrival time

Again, consider \( T \) projective measurements from the set [7]

\[ \mathcal{E} = \{ E_i, i = 1, 2, \ldots, T \}. \]

\[ \begin{array}{cccccccccc}
 x & |0\rangle & |1\rangle & |\phi_1\rangle & |\psi_1\rangle & |\psi_2\rangle & |\psi_3\rangle & |\phi_2\rangle & |\phi_3\rangle & |\phi_4\rangle \\
 \pi^\sharp(x) & E_2 & E_3 & E_3 & E_4 & E_3 & E_4 & E_1 & E_4 & E_2 \\
\end{array} \]

In Fig. 3, we plot \( J^\ast_{\pi^\sharp}(T) \) as a function of \( T \), for \( T = 2, 3, \ldots, 30 \). Numerical calculations show that the minimized average number of steps of driving \( |0\rangle \langle 0| \) to \( |1\rangle \langle 1| \) does not depend too much on the size of the control set; it oscillates around 3.8 for control sets of reasonable size. Also, for measurement set \( \mathcal{E} \) with \( T = 5 \), we show the optimal policy \( \pi^\sharp \) in Table III.

IV. CONCLUSIONS

We have proposed feedback designs for manipulating a quantum state to a target state by performing sequential measurements. Making use of Belavkin’s quantum feedback control theory, we showed that finding the measurement selection policy that maximizes the probability of successful state manipulation is an optimal control problem which can be solved by dynamical programming for any given set of measurements and a given time horizon. Numerical examples indicate that making use of feedback information significantly improves the success probability compared to the classical scheme without taking feedback. It was shown that the probability of reaching the target state via feedback control reaches 0.9968 using merely ten steps, while classical results [6,7] suggested that optimal feedback-free strategy via consecutive measurements in turn reaches success probability when the number of steps tends to infinity. Maximizing the expected fidelity to the target state and minimizing the expected arrival time were also considered, and some connections and differences among these objectives were also discussed.

ACKNOWLEDGMENTS

We gratefully acknowledge support by the Australian Research Council Centre of Excellence for Quantum Computation and Communication Technology (Project No. CE110001027), and AFOSR Grant No. FA2386-12-1-4075.

[14] It is clear from this objective that $E_*$ must be a measurement in the set $E$ for $J_*(N)$ to be a nontrivial function if all measurements in $E$ are projective.