# On the Value Functions of the Optimal Quadratic Regulation Problem for Discrete-Time Switched Linear Systems

Wei Zhang and Jianghai Hu
School of Electrical and Computer Engineering,
Purdue University, West Lafayette, IN 47907, USA.
Email:{zhang70,jianghai@purdue.edu}

*Abstract*— **In this paper, we study the value functions associated with the discrete-time LQR problem for switched linear systems (DLQRS). Some important properties of the value functions and value iterations are derived. In particular, we will show that under some mild conditions, the family of the finite-horizon value functions of the DLQRS problem is homogeneous (of degree 2), uniformly bounded over the unit ball, and converges exponentially fast to the corresponding infinite-horizon value function. More importantly, the exponential convergence rate of the value iteration is characterized analytically in terms of the subsystem matrices. The properties derived in this paper are not only of theoretical importance, but also crucial in the analysis and design of various optimal and suboptimal control strategies for DLQRS problems.**

## I. Introduction

Switched systems arise naturally in many engineering fields, such as power electronics [1], [2], embedded systems [3], [4], manufacturing [5], and communication networks [6], etc. Incorporating the switching behavior in the model and controller structures offers much greater freedom and infinitely more possibilities for capturing complex system dynamics, achieving stabilization and improving the overall performance of the feedback systems. On the other hand, the generality of switched systems has also posed many challenges in their formal analysis and design [7]. This paper studies one of these challenges, namely, the optimal control of switched systems.

Compared with the traditional optimal control problems, a distinctive feature of the optimal control of switched systems is its freedom in selecting the mode sequence and switching instants. For a fixed mode sequence, variational approach can usually be applied to derive certain gradient based algorithms for optimizing the corresponding switching instants [8], [9]. However, to further find the best mode sequence becomes a discrete optimization problem and is believed to be NP hard in general. Recent research attention ([10], [11], [12]) has been focused on the optimal control of discrete-time switched linear systems with quadratic cost functions, which contains many interesting properties of the general optimal control problems of switched systems, but at the same time allows efficient ways for optimizing the mode sequences. Such problems can be viewed as an extension of the classical discrete-time LQR problems to the context of the switched linear systems, and are thus referred to as the DLQRS problems. It has been proved in our previous paper [13] that the finite-horizon value function associated with the DLQRS problems can be written as a pointwise minimum of a finite set of quadratic functions. This simple structure of the value function has also inspired an efficient numerical algorithm for solving the finite-horizon DLQRS problems. Although the algorithm scales very well with respect to the control horizon $N$, its complexity may still grow out of hand when $N$ is extremely large or infinite. To extend our algorithm to the cases with large or infinite horizons, a deeper understanding on the sequence of the value functions, generated by the Bellman iteration of the DLQRS problem, is necessary.

The main contribution of this paper is the derivation of various interesting and important properties of the value functions of the DLQRS problems. In particular, it is proved that under some mild conditions, the family of the finite-horizon value functions of the DLQRS problem is homogeneous (of degree 2) and uniformly bounded over the unit ball, and converges exponentially fast to the corresponding infinite-horizon value function. More importantly, the exponential convergence rate of the value iterations is characterized analytically in terms of the subsystem matrices. The convergence of the value iterations is of great importance for developing suboptimal control strategies. If the cost function is discounted by a constant factor strictly less than 1, the stationary policy generated by the converged value function is suboptimal [14]. Even in the undiscounted case, the optimal finite-horizon policy, with a sufficiently large horizon, can be used repetitively to construct a suboptimal infinite-horizon control strategy. The formal discussion of the suboptimal controls will appear in the upcoming paper [15]. The properties derived in this paper provide the theoretical foundation and preparations for the future studies of various aspects of the large or infinite-horizon DLQRS problems.

This paper is organized as follows. The DLQRS problem is formulated in Section II. The structure of its value functions is briefly reviewed in Section III. Various important properties of the value functions are derived in Section IV. Some concluding remarks are given in Section V.

## II. Problem Formulation

Consider the discrete-time switched linear system defined as:

$$x(t+1) = A_{v(t)}x(t) + B_{v(t)}u(t),\ t = 0, \dots, N-1, \quad (1)$$

This work was partially supported by the National Science Foundation under Grant CNS-0643805

where $x(t) \in \mathbb{R}^n$ is the continuous state, $v(t) \in \mathbb{M} \triangleq \{1, \ldots, M\}$ is the discrete control or switching signal, and $u(t) \in \mathbb{R}^p$ is the continuous control. The combination of the two types of controls, $(u(\cdot), v(\cdot))$, will be referred to as the *hybrid control sequence*. For each $i \in \mathbb{M}$, $A_i$ and $B_i$ are constant matrices of appropriate dimension, and the pair $(A_i, B_i)$ is called a subsystem of (1). This switched linear system is time invariant in the sense that the set of available subsystems $\{(A_i, B_i)\}_{i=1}^M$ is independent of time $t$. We assume that there is no internal forced switchings, i.e., the system can stay at or switch to any mode at any time instant. A *state feedback control strategy* of system (1) at time $t$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^p \times \mathbb{M}$ and is denoted by $\xi_t(x) \triangleq (\mu_t(x), \nu_t(x))$. A sequence of feedback control strategies over the whole time horizon $[0, N]$ constitutes a *feedback policy*: $\pi_N \triangleq \{\xi_0, \xi_1, \ldots, \xi_{N-1}\}$. If the system is driven by a feedback policy $\pi_N$, then its hybrid control sequence will be given by $u(t) = \mu_t(x(t))$ and $v(t) = \nu_t(x(t))$.

In this paper, the terminal cost function $\psi(x)$ and the running cost function $L(x, u, v)$ are assumed to be in the following quadratic forms:

$$\psi(x) = x^T Q_f x, \quad L(x, u, v) = x^T Q_v x + u^T R_v u,$$

where $Q_f = Q_f^T \succeq 0$ is the terminal state weight, and $Q_v = Q_v^T \succeq 0$ and $R_v = R_v^T \succ 0$ are the running weights for the state and the control for subsystem $v \in \mathbb{M}$, respectively. Starting from $x(0) = z$, the overall objective function to be minimized over the time horizon $[0, N]$ is defined as

$$J_N(z, u, v) = \psi(x(N)) + \sum_{j=0}^{N-1} L(x(j), u(j), v(j))). \tag{2}$$

When the control horizon is infinite, the final cost will never be incurred. Thus, the infinite-horizon objective function under the hybrid control sequence $(u, v)$ is defined as:

$$J_\infty(z, u, v) = \sum_{j=0}^{\infty} L(x(j), u(j), v(j)).$$

For a fixed switching sequence $v(\cdot) = \{v(0), v(1), \ldots\}$, system (1) becomes a linear time-varying system whose state transition matrix from time $i$ to time $j \geq i$ is given by:

$$\phi_{i,j}(v) = \begin{cases} \prod_{k=j}^{i-1} A_{v(k)}, & i \geq j+1, \\ I_{n\times n}, & i = j, \end{cases}$$

where $I_{n\times n}$ denotes the identity matrix of dimension $n$. With this definition, the zero-input response with initial state $x(0) = z$ and switching signal $v(\cdot)$ is simply $x_{zi}(t; v) = \phi_{t,0}(v)z$, $t \geq 0$. Similarly, for a given hybrid control sequence $(u(\cdot), v(\cdot))$ the zero-state response can be written as

$$x_{zs}(t; u, v) = \sum_{k=0}^{t-1} \phi_{t,k}(v) B_{v(k)} u_k, \quad t \geq 1. \tag{3}$$

With these notations, $J_\infty$ can be written as:

$$\begin{aligned} J_\infty(z, u, v) = z^T &\left( \sum_{j=0}^{\infty} \phi_{j,0}(v)^T Q_{v(j)} \phi_{j,0}(v) \right) z \\ &+ \sum_{j=1}^{\infty} x_{zs}(j; u, v)^T Q_{v(j)} x_{zs}(j; u, v) \\ &+ \left( 2 \sum_{j=1}^{\infty} x_{zs}(j; u, v)^T Q_{v(j)} \phi_{j,0}(v) \right) z \\ &+ \sum_{j=0}^{\infty} u(j)^T R_{v(j)} u(j). \end{aligned} \tag{4}$$

The discrete-time LQR problem for the switched linear system (1), referred to as DLQRS problem hereby, can be formulated as follows:

*Problem 1 (DLQRS problem):* For an arbitrary initial state $z$ and a time horizon $[0, N]$, with the possibility of $N$ being infinite, find the hybrid control sequence $(u, v)$ that minimizes $J_N(z, u, v)$ subject to the dynamic equation (1).

## III. Review of Existing Results

A common way of solving Problem 1 is the dynamic programming approach. When $N$ is finite, the value function $V_{t,N} : \mathbb{R}^n \to \mathbb{R}$ is defined as:

$$V_{t,N}(z) = \min_{\substack{v(j), u(j), \\ t \leq j \leq N-1}} \Big\{ \psi(x(N)) + \sum_{j=t}^{N-1} L(x(j), u(j), v(j)) \Big| \text{ subject to eq. (1) with } x(t) = z \Big\}, \tag{5}$$

for each time $t \in \{0, 1, \ldots, N\}$. The function $V_{t,N}(z)$ thus defined is the minimum cost-to-go starting from state $z$ at time $t$. The minimum cost of the DLQRS problem is simply $V_{0,N}(z)$. Due to the time-invariant nature of the switched system (1), its value function depends only on the number of remaining time steps. In the rest of this paper, when no ambiguity arises, we will denote by $V_k(z)$ the value function at time $t = N - k$ when there are $k$ time steps left, i.e., $V_k(z) \triangleq V_{N-k,N}(z)$. To emphasize its substantial difference from the finite-horizon value function, the *infinite-horizon value function* is specially denoted by $V^*(z)$, i.e.,

$$V^*(z) = \inf_{u,v} J_\infty(z, u, v). \tag{6}$$

It turns out that the value function of the DLQRS problem has a strong connection with the Riccati equation and the Kalman gain arising in the classical LQR problem of linear systems. Denote by $\mathcal{A}$ the set of all positive semi-definite (p.s.d.) matrices. For any matrix $P \in \mathcal{A}$, define the Riccati mapping $\rho_i : \mathcal{A} \to \mathcal{A}$ associated with subsystem $(A_i, B_i)$, $i \in \mathbb{M}$ as:

$$\rho_i(P) = Q_i + A_i^T P A_i - A_i^T P B_i (R_i + B_i^T P B_i)^{-1} B_i^T P A_i. \tag{7}$$

Let $\mathcal{H}_0 = \{Q_f\}$ be an initial set consisting of only the matrix $Q_f$. Define the set $\mathcal{H}_k$ for $k \geq 0$ iteratively as

$$\mathcal{H}_{k+1} = \rho_{\mathbb{M}}(\mathcal{H}_k) \triangleq \{\rho_i(P_k) : i \in \mathbb{M} \text{ and } P_k \in \mathcal{H}_k\}. \quad (8)$$

In other words, the matrices in $\mathcal{H}_{k+1}$ are the images of the matrices in $\mathcal{H}_k$ under the Riccati mappings of the $M$ subsystems. Denote by $|\mathcal{H}_k|$ the number of distinct matrices in $\mathcal{H}_k$. Then it can be easily seen that $|\mathcal{H}_0| = 1$ and $|\mathcal{H}_k| \leq M^k$ for any $k \geq 0$. It has been proved in [13] that the finite-horizon value function of the DLQRS problem is a pointwise minimum of a finite set of quadratic functions as described in the following theorem.

*Theorem 1 ( [13]):* When $N$ is finite, the value function for the DLQRS problem at time $N-k$, i.e., with $k$ time steps left, is

$$V_k(z) = \min_{P \in \mathcal{H}_k} z^T P z. \quad (9)$$

Furthermore, for $k \geq 0$, if we define

$$(P_k^*(z), i_k^*(z)) = \underset{(P \in \mathcal{H}_k, i \in \mathbb{M})}{\arg\min}\; z^T \rho_i(P) z, \quad (10)$$

then the optimal feedback strategy at state $z$ and time $t = N-(k+1)$ is $\xi_t^*(z) = (\mu_t^*(z), \nu_t^*(z))$, where $\mu_t^*(z) = -K_{i_k^*(z)}(P_k^*(z))z$ and $\nu_t^*(z) = i_k^*(z)$. Here, $K_i(P)$ is the Kalman gain for subsystem $i$ with matrix $P$, i.e.,

$$K_i(P) \triangleq (R_i + B_i^T P B_i)^{-1} B_i^T P A_i. \quad (11)$$

.

*Remark 1:* The special structure of $V_k(z)$ is crucial in solving the finite-horizon DLQRS problems. Although in the worst case, $|\mathcal{H}_k|$ grows exponentially fast, in terms of computing the value function (9), only a small portion of the matrices in $\mathcal{H}_k$, which give rise to the minimum of (9) for at least one $z \in \mathbb{R}^n$, should be kept. Thus, if we remove all the other redundant matrices after each iteration of (8), we can obtain a new sequence of reduced sets $\{\hat{\mathcal{H}}_k\}$, which can be dramatically smaller than $\{\mathcal{H}_k\}$ but defines exactly the same value functions $\{V_k(z)\}$. In this way, the optimal control strategy can be efficiently computed using the reduced sets $\{\hat{\mathcal{H}}_k\}$. Refer to [13] for more details about the efficient solutions of the finite-horizon DLQRS problems.

Unlike the finite-horizon case, the explicit form of the infinite-horizon value function $V^*(z)$ can not be easily obtained. Nevertheless, as will be discussed in the next section, many important properties of $V^*(z)$ can be still inferred based on its definition (6) and (4).

## IV. Properties of the Value Functions

In this section, we will derive various important properties of the infinite-horizon value function $V^*(z)$ and the family of the finite-horizon value functions $\{V_k(z)\}_{k\geq 0}$. These properties are not only of theoretical importance, but also crucial in developing suboptimal control strategies for large or infinite horizon DLQRS problems.

### A. Homogeneity

*Lemma 1 (Homogeneity):* $V_k(z)$ and $V^*(z)$ are homogeneous in the sense that $V_k(\lambda z) = \lambda^2 V_k(z)$ and $V^*(\lambda z) = \lambda^2 V^*(z)$, for any $z \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.

*Proof:* The homogeneity of $V_k$ follows directly from (9). To prove the homogeneity of $V^*(z)$, let $\lambda$ be an arbitrary real number. By the definition of zero-state response in (3), we have $x_{zs}(t;\lambda u, v) = \lambda x_{zs}(t; u, v)$. Then it follows easily from (4) that $J_\infty(\lambda z, \lambda u, v) = \lambda^2 J_\infty(z, u, v)$ (the equality still holds if one side is infinite). Let $z$ be an arbitrary point in $\mathbb{R}^n$. If $V^*(z) = \infty$, then $V^*(\lambda z)$ must also be infinite. The desired equality holds. If $V^*(z) < \infty$, then for any $\epsilon \in \mathbb{R}$, there exists a control sequence $(u_z, v_z)$ such that $J_\infty(z, u_z, v_z) \leq V^*(z) + \epsilon$. Thus $V^*(\lambda z) \leq J_\infty(\lambda z, \lambda u_z, v_z) = \lambda^2 J_\infty(z, u_z, v_z) \leq \lambda^2 V^*(z) + \lambda^2 \epsilon$. Since $\epsilon$ can be arbitrarily small, we must have $V^*(\lambda z) \leq \lambda^2 V^*(z)$. Let $\delta = \lambda^2 V^*(z) - V^*(\lambda z)$. If $\delta > 0$, then there exists a control strategy $(\hat{u}_z, \hat{v}_z)$ such that $J_\infty(\lambda z, \lambda \hat{u}_z, \hat{v}_z) \leq V^*(\lambda z) + \delta/2 < \lambda^2 V^*(z)$. Hence, $J_\infty(z, \hat{u}_z, \hat{v}_z) = \frac{1}{\lambda^2} J_\infty(\lambda z, \lambda \hat{u}_z, \hat{v}_z) < V^*(z)$, which contradicts the optimality of $V^*(z)$. Thus $\delta = 0$ and the desired result is proved. ■

The properties of the value functions presented in the rest of this section are based on the following stabilizability condition of the switched system (1).

(A1) At least one subsystem is stabilizable.

### B. Boundedness

*Proposition 1:* Denote by $\|\cdot\|$ the 2-norm of a given matrix or vector. Under assumption (A1), there must exist a finite constant $\beta$ such that $V_k(z) \leq \beta \|z\|^2$, for all $k \in \mathbb{Z}^+$ and $z \in \mathbb{R}^n$. Furthermore, if the stabilizable subsystem is $(A_i, B_i)$ and $F$ is any feedback gain for which $\bar{A}_i \triangleq A_i - B_i F$ is stable, then one possible choice of $\beta$ is given by:

$$\beta = n\|Q_f\| + \|Q_i + F^T R_i F\| \cdot \left( \sum_{j=0}^{\infty} \|\bar{A}_i^j\|^2 \right), \quad (12)$$

*Proof:* Suppose subsystem $(A_i, B_i)$ is stabilizable. Let $\{P_k^{(i)}\}_{k=0}^\infty$ be the sequence of matrices generated by the Riccati mapping using only subsystem $i$, i.e., $P_{k+1}^{(i)} = \rho_i(P_k^{(i)})$ with $P_0^{(i)} = Q_f$. Since the switched system (1) can stay in subsystem $(A_i, B_i)$ all the time, its value function must be no greater than the value function of the LQR problem with only one subsystem $(A_i, B_i)$, i.e., $V_k(z) \leq z^T P_k^{(i)} z$ for all $k \in \mathbb{Z}^+$ and $z \in \mathbb{R}^n$. Thus, it suffices to show that the $\beta$ given in (12) is an upper bound of the 2-norm of all the matrices in $\{P_k^{(i)}\}_{k=0}^\infty$. Let $F$ be a feedback gain for which $\bar{A}_i \triangleq (A_i - B_i F)$ is stable. Define $\{\tilde{P}_k^{(i)}\}_k^\infty$ iteratively by

$$\tilde{P}_{k+1}^{(i)} = Q_i + \bar{A}_i^T \tilde{P}_k^{(i)} \bar{A}_i + F^T R_i F, \text{ with } \tilde{P}_0^{(i)} = Q_f. \quad (13)$$

In the above equation, if we let $F = K_i(\tilde{P}_k^{(i)})$ for each $k$, then $\tilde{P}_k^{(i)}$ will coincide with $P_k^{(i)}$. In other words, $\tilde{P}_k^{(i)}$ defines the quadratic energy cost of using the stable feedback gain $F$ instead of the time-dependent optimal Kalman gain in the $k$-horizon LQR problem. By a standard result of the Riccati equation theory (Theorem 2.1 in [16]), we have $P_k^{(i)} \preceq \tilde{P}_k^{(i)}$ for all $k \geq 0$. Thus, it suffices to show $\|\tilde{P}_k^{(i)}\|_2 \leq \beta$ for each $k \geq 0$. By (13), we have

$$
\begin{aligned}
\tilde{P}_k^{(i)} &= P_0^{(i)} + \sum_{j=1}^{k} (P_j^{(i)} - P_{j-1}^{(i)}) \\
=& P_0^{(i)} + \sum_{j=0}^{k-1} (\bar{A}_i^T)^j (P_1^{(i)} - P_0^{(i)}) (\bar{A}_i)^j \\
=& Q_f + \sum_{j=0}^{k-1} (\bar{A}_i^T)^{j+1} Q_f (\bar{A}_i)^{j+1} \\
&+ \sum_{j=0}^{k-1} (\bar{A}_i^T)^j (Q_i - Q_f + F^T R_i F)(\bar{A}_i)^j \\
\preceq& (\bar{A}_i^T)^k Q_f (\bar{A}_i)^k \\
&+ \sum_{j=1}^{\infty} (\bar{A}_i^T)^j (Q_i + F^T R_i F)(\bar{A}_i)^j
\end{aligned}
$$

Thus, $\|P_k^{(i)}\| \leq n\|Q_f\| + \|Q_i + F^T R_i F\| \left(\sum_{j=0}^{\infty} \|\bar{A}_i^j\|^2\right)$. Note that the geometric series formula does not apply here, as the 2-norm of a stable matrix may not be strictly less than 1 in general. However, by a well known fact that $\lim_{j\to\infty} \|\bar{A}^j\|^{1/j} = \rho(\bar{A}) < 1$, where $\rho(\cdot)$ denotes the spectrum radius of a given matrix, we know that $\|\bar{A}_i^j\| < 1$ for all large $j$. Therefore, the sum of the series does converge and the proposition is proved. ■

### C. Exponential Convergence

The convergence of the value iteration cannot be guaranteed by only the assumption (A1). The following condition is required.

$$\text{(A2)} \quad Q_i \succ 0 \quad \text{for each } i \in \mathbb{M}.$$

The rest of this subsection is devoted to proving that under assumptions (A1) and (A2), the finite-horizon value functions $V_k(z)$ converges to $V^*(z)$ exponentially fast. The result will be proved through a series of lemmas. We first introduce some notations used throughout the subsequent discussions. Denote by $\sigma_{\min}(\cdot)$ and $\sigma_{\max}(\cdot)$, respectively, the smallest and the largest singular values of a given matrix. Define $\sigma_Q^- = \min_{i\in\mathbb{M}}\{\sigma_{min}(Q_i)\}$ and $\sigma_A^+ = \max_{i\in\mathbb{M}}\{\sigma_{\max}(A_i)\}$. For $0 \leq j \leq k$, denote by $x_{z,k}^*(j)$ the optimal trajectory originating from $z$ at time $t = N - k$. Here, $j$ denotes the relative time during the remaining horizon $[N-k, N]$, i.e., $t = j + N - k$. Let $(u_{z,k}^*(j), v_{z,k}^*(j))$ be the corresponding hybrid control sequence.

*Lemma 2:* Let $k_1$ and $k_2$ be positive integers such that $k_1 > k_2$. For any $z \in \mathbb{R}^n$, the following inequality holds:

$$
\begin{aligned}
&V_{k_1-k_2}(x_{z,k_1}^*(k_2)) - \psi(x_{z,k_1}^*(k_2)) \\
\leq& V_{k_1}(z) - V_{k_2}(z) \\
\leq& V_{k_1-k_2}(x_{z,k_2}^*(k_2)) - \psi(x_{z,k_2}^*(k_2))
\end{aligned} \tag{14}
$$

*Proof:* Denote by $z_1$ the final state value of the optimal trajectory originating from $z$ at time $N - k_2$, i.e., $z_1 = x_{z,k_2}^*(k_2)$. Define

$$
\tilde{x}(j) = \begin{cases} x_{z,k_2}^*(j), & j \leq k_2 \\ x_{z_1,k_1-k_2}^*(j-k_2), & k_2 < j \leq k_1 \end{cases} \tag{15}
$$

As shown in Fig. 1, $\tilde{x}(\cdot)$ is obtained by concatenating trajectory $x_{z_1,k_1-k_2}^*(\cdot)$ to the end of $x_{z,k_2}^*(\cdot)$. Let $\tilde{u}(\cdot)$ and $\tilde{v}(\cdot)$ be the controls corresponds to $\tilde{x}$. Then by the definition of the value function, we have

$$
\begin{aligned}
&V_{k_1}(z) \leq \sum_{j=1}^{k_1-1} L(\tilde{x}(j), \tilde{u}(j), \tilde{v}(j)) + \psi(\tilde{x}(k_1)) \\
=& \sum_{j=0}^{k_2-1} L(x_{z,k_2}^*(j), u_{z,k_2}^*(j), v_{z,k_2}^*(j)) \\
+& \sum_{j=0}^{k_1-k_2-1} L(x_{z_1,k_1-k_2}^*(j), u_{z_1,k_1-k_2}^*(j), v_{z_1,k_1-k_2}^*(j)) \\
&+ \psi(x_{z_1,k_1-k_2}^*(k_1-k_2)) \\
=& V_{k_2}(z) - \psi(x_{z,k_2}^*(k_2)) + V_{k_1-k_2}(z_1) \\
=& V_{k_2}(z) - \psi(x_{z,k_2}^*(k_2)) + V_{k_1-k_2}(x_{z,k_2}^*(k_2))
\end{aligned} \tag{16}
$$

Equation (16) describes exactly the second inequality in (14). To prove the first inequality in (14), define $\hat{x}(t) = x_{z,k_1}^*(j)$ for $0 \leq j \leq k_2$ and let $(\hat{u}(\cdot), \hat{v}(\cdot))$ be the corresponding hybrid control sequence. Since $\hat{x}(\cdot)$ can be viewed as a system trajectory originating from $z$ at time $N - k_2$, we have

$$
\begin{aligned}
V_{k_2}(z) &\leq \sum_{j=0}^{k_2-1} L(\hat{x}(j), \hat{u}(j), \hat{v}(j)) + \psi(\hat{x}(k_2)) \\
&= \sum_{j=0}^{k_2-1} L(x_{z,k_1}^*(j), u_{z,k_1}^*(j), v_{z,k_1}^*(j)) + \psi(x_{z,k_1}^*(k_2)) \\
&= V_{k_1}(z) - V_{k_1-k_2}(x_{z,k_1}^*(k_2)) + \psi(x_{z,k_1}^*(k_2)),
\end{aligned} \tag{17}
$$

where the last step follows from the Bellman's principle of optimality, namely, any segment of an optimal trajectory must be the optimal trajectory joining the two end points of the segment. The desired result follows from (16) and (17). ■

For the switched linear system considered in this paper, the $k$-horizon value function $V_k(z)$ may not be monotone with respect to $k$. Nevertheless, by Lemma 2, the difference between two value functions $V_{k_1}(z)$ and $V_{k_2}(z)$ can be bounded by some quadratic functions of $x_{z,k_1}^*(k_2)$ and $x_{z,k_2}^*(k_2)$. In the next lemma, we will prove that $x_{z,k}^*(j) \to 0$ as $k \geq j \to \infty$. This will guarantee that by choosing large $k_1$ and $k_2$, the upper and lower bounds in (14) can be made arbitrarily small. The convergence of the value iteration can thus be easily proved.

*Lemma 3:* Under assumptions (A1) and (A2), for any $j = 0, 1, \ldots, k-1$, the optimal trajectory originating from

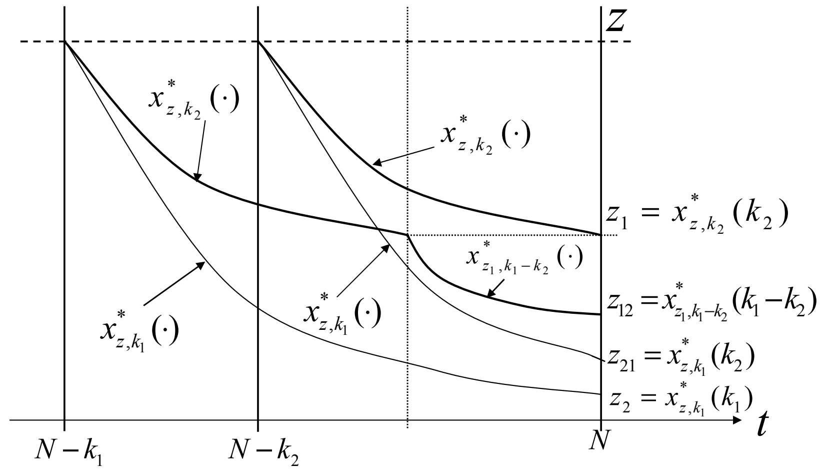


Fig. 1. Illustrating the proof of Lemma 2

$z$ at time $N-k$, namely, $x^*_{z,k}(\cdot)$, satisfies the following inequalities:

$$\|x^*_{z,k}(j)\|_2^2 \le \frac{\beta}{\sigma_Q^-}\gamma^j\|z\|_2^2,$$

$$\|x^*_{z,k}(k)\|_2^2 \le \frac{\beta(\sigma_A^+)^2}{\sigma_Q^-}\gamma^{k-1}\|z\|_2^2 \tag{18}$$

$$\text{with} \quad \gamma = \left(\frac{1}{1+\sigma_Q^-/\beta}\right).$$

where $\beta$ is the finite positive real number introduced in Proposition 1

*Proof:* For simplicity, for $j = 0, 1, \cdots, k$, define $\tilde{x}(j) \triangleq x^*_{z,k}(j)$ and $\tilde{V}_{k-j} \triangleq V_{k-j}(x^*_{z,k}(j))$. Denote by $\tilde{u}(\cdot)$ and $\tilde{v}(\cdot)$ the optimal controls corresponding to $x^*_{z,k}(\cdot)$. For $j = 1, \ldots, k$, we have

$$\begin{aligned}
&\tilde{V}_{k-(j-1)} - \tilde{V}_{k-j}\\
=&L(\tilde{x}(j-1), \tilde{u}(j-1), \tilde{v}(j-1))\\
\ge&\tilde{x}(j-1)^T Q_{\tilde{v}(j-1)}\tilde{x}(j-1)\\
\ge&\sigma_Q^-\|\tilde{x}(j-1)\|_2^2\\
\ge&\sigma_Q^-/\beta\tilde{V}_{k-(j-1)}\\
\ge&\sigma_Q^-/\beta\tilde{V}_{k-j}.
\end{aligned} \tag{19}$$

By (19), we have $\tilde{V}_{k-j} \le \frac{1}{1+\sigma_Q^-/\beta}\tilde{V}_{k-(j-1)}$ for $j = 1, \cdots, k$. Hence, $\tilde{V}_{k-j} \le \left(\frac{1}{1+\sigma_Q^-/\beta}\right)^j \tilde{V}_k$. Obviously, for $j \le k-1$, $\tilde{V}_{k-j} \ge \sigma_Q^-\|\tilde{x}(j)\|_2^2$. Thus,

$$\|\tilde{x}(j)\|^2 \le \frac{1}{\sigma_Q^-}\tilde{V}_1 \le \frac{1}{\sigma_Q^-}\left(\frac{1}{1+\sigma_Q^-/\beta}\right)^j\tilde{V}_k \tag{20}$$

$$\le \frac{\beta}{\sigma_Q^-}\left(\frac{1}{1+\sigma_Q^-/\beta}\right)^j\|z\|_2^2 \tag{21}$$

$$= \frac{\beta}{\sigma_Q^-}\gamma^j\|z\|_2^2. \tag{22}$$

For $j = k$, by Theorem 1, we have that $\tilde{x}(k) = (A_i - B_iK_i(P))\cdot\tilde{x}(k-1)$ for some $i \in \mathbb{M}$ and $P \in \mathcal{H}^*$. It follows from (11) that

$$\begin{aligned}
B_iK_i(P) &= B_i(R_i + B_i^TPB_i)^{-1}B_i^TPA_i\\
&\preceq B_i(B_i^TPB_i)^{-1}B_i^TPA_i = A_i.
\end{aligned}$$

Thus,

$$\begin{aligned}
\|x(k)\|_2^2 &= \|(A_i - B_iK_i(P))\cdot\tilde{x}(k-1)\|_2^2\\
&\le \|A_i\|_2^2\cdot\|\tilde{x}(k-1)\|_2^2 = (\sigma_A^+)^2\cdot\|\tilde{x}(k-1)\|_2^2\\
&\le \frac{(\sigma_A^+)^2\beta}{\sigma_Q^-}\gamma^{k-1}\|z\|_2^2.
\end{aligned}$$

■

Now we are ready to prove the main result of this subsection.

*Theorem 2:* Under assumptions (A1) and (A2), $V_k(z)$ converges exponentially fast to $V_\infty(z)$ for each $z \in \mathbb{R}^n$ as $k \to \infty$. Furthermore, the convergence is uniform over the unit ball in $\mathbb{R}^n$ and the difference between the value functions at time step $N-k_1$ and $N-k_2$ is bounded above by

$$|V_{k_1}(z) - V_{k_2}(z)| \le \alpha\gamma^{k_2}\|z\|^2, \tag{23}$$

where $\alpha = \max\{1, \frac{(\sigma_A^+)^2}{\gamma}\}\cdot\frac{(\beta+\lambda_f^+)\beta}{\sigma_Q^-}$.

*Proof:* By Lemma 3, for any $z \in \mathbb{R}^n$ and $k_1 > k_2$, we have $\|x^*_{z,k_2}(k_2)\|^2 \le \frac{(\sigma_A^+)^2\beta}{\sigma_Q^-\gamma}\gamma^{k_2}\|z\|^2$ and $\|x^*_{z,k_1}(k_2)\|^2 \le \frac{\beta}{\sigma_Q^-}\gamma^{k_2}\|z\|^2$. Hence,

$$V_{k_1-k_2}(x^*_{z,k_2}(k_2)) \le \beta\|x^*_{z,k_2}(k_2)\|^2 \le \frac{(\sigma_A^+)^2\beta^2}{\sigma_Q^-\gamma}\gamma^{k_2}\|z\|^2,$$

$$\psi(x^*_{z,k_2}(k_2)) \le \lambda_f^+\|x^*_{z,k_2}(k_2)\|^2 \le \frac{\lambda_f^+(\sigma_A^+)^2\beta}{\sigma_Q^-\gamma}\gamma^{k_2}\|z\|^2,$$

$$V_{k_1-k_2}(x^*_{z,k_1}(k_2)) \le \beta\|x^*_{z,k_1}(k_2)\|^2 \le \frac{\beta^2}{\sigma_Q^-}\gamma^{k_2}\|z\|^2,$$

$$\psi(x^*_{z,k_1}(k_2)) \le \lambda_f^+\|x^*_{z,k_1}(k_2)\|^2 \le \frac{\lambda_f^+\beta}{\sigma_Q^-}\gamma^{k_2}\|z\|^2.$$

Thus, by Lemma 2, the difference between the value functions at time $N - k_1$ and $N - k_2$ is bounded above by

$$|V_{k_1}(z) - V_{k_2}(z)| \leq \max\{1, \frac{(\sigma_A^+)^2}{\gamma}\} \cdot \frac{(\beta + \lambda_f^+)\beta}{\sigma_Q^-}\gamma^{k_2}\|z\|^2.$$

Since $\gamma < 1$ and the upper bound in the above equation is independent of $k_1$, the value function converges exponentially fast for each fixed $z$. In addition, the convergence is obviously uniform for all $z$ in the unit ball. ∎

*Remark 2:* Assumptions (A1) and (A2) together imply the exponential convergence of the value iteration. It is well known ([14]) that the limiting function $V_\infty(z)$, even exists, may not coincide with the infinite-horizon value function $V^*(z)$. However, it can be proved that it is indeed the case for the DLQRS problem under assumptions (A1) and (A2). The proof is omitted here, but can be found in [15].

*Remark 3:* Some convergence results of general value iterations can be found in [14], [10]. Compared with the previous work, our convergence result derived specially for the DLQRS problem has several distinctions. Firstly, it allows nonzero terminal cost, which is especially important for finite-horizon DLQRS problem. Secondly, its conditions are expressed in terms of the system matrices rather than the infinite-horizon value function, and thus become much easier to verify. Finally, as indicated in Theorem 2, the convergence rate can be approximated using the system matrices. Thus, for a given tolerance on the optimal cost, an upper bound of the required number of iterations can be computed before the actual computation starts. This provides an efficient means to stop the value iterations with guaranteed suboptimal performance.

## V. Conclusion

Some important properties of the value functions of the DLQRS problems are derived in this paper. In particular, we have proved that under some mild conditions, the family of the value functions generated by the Bellman iteration is homogeneous of degree 2 and is uniformly bounded over the unit ball. More importantly, we have also proved that the finite-horizon value functions converges exponentially fast to the corresponding infinite-horizon value function under the additional assumption that each state weighting matrix $Q_i$ is positive definite. The convergence rate is also characterized analytically in terms of the subsystem matrices. Future research will focus on using these properties to develop infinite-horizon control strategies for the DLQRS problems with guaranteed suboptimal performance.

## References


[1] T. Geyer, G. Papafotiou, and M. Morari. Model Predictive Control in Power Electronics: A Hybrid Systems Approach. In *IEEE Conference on Decision and Control*, Seville, Spain, December 2005.

[2] A.G. Beccuti, G. Papafotiou, and M. Morari. Optimal Control of the Boost dc-dc Converter. In *IEEE Conference on Decision and Control*, Seville, Spain, December 2005.

[3] W. Zhang, J. Hu, and Y.-H. Lu. Optimal Power Modes Scheduling Using Hybrid Systems. In *Proceedings of the American Control Conference*, New York City, NY, July 2007.

[4] W. Zhang and J. Hu. Dynamic buffer management using optimal control of hybrid systems. *Automatica*, (7):1831–1840, Jul 2008.

[5] C. G. Cassandras, D. L. Pepyne, and Y. Wardi. Optimal control of a class of hybrid systems. *IEEE Transactions on Automatic Control*, 46(3):398–415, 2001.

[6] J. Hespanha, S. Bohacek, K. Obraczka, and J. Lee. Hybrid modeling of TCP congestion control. In Maria Domenica Di Benedetto and Alberto Sangiovanni-Vincentelli, editors, *Hybrid Systems: Computation and Control*, number 2034 in LNCS, pages 291–304. Springer-Verlag, Berlin, March 2001.

[7] Daniel Liberzon. *Switching in Systems and Control*. Systems & Control: Foundations & Applications. Birkhauser, Boston, 2003.

[8] X. Xu and P.J. Antsaklis. Optimal control of switched systems based on parameterization of the switching instants. *IEEE Transactions on Automatic Control*, 49(1):2–16, 2004.

[9] M. Egerstedt, Y. Wardi, and F. Delmotte. Optimal control of switching times in switched dynamical systems. *IEEE Transactions on Automatic Control*, 51(1):110–115, 2006.

[10] B. Lincoln and A. Rantzer. Relaxing dynamic programming. *IEEE Transactions on Automatic Control*, 51(8):1249–1260, August 2006.

[11] A. Rantzer. Relaxed dynamic programming in switching systems. *IEE Proceedings - Control Theory & Applications*, 153(5):567 – 574, September 2006.

[12] F. Borrelli, M. Baotic, A. Bemporad, and M. Morari. Dynamic programming for constrained optimal control of discrete-time linear hybrid systems. *Automatica*, 41:1709–1721, October 2005.

[13] W. Zhang and J. Hu. On optimal quadratic regulation for discrete-time switched liear systems. In Magnus Egerstedt and Bud Mishra, editors, *Hybrid Systems: Computation and Control*, LNCS. Springer-Verlag, St. Louis, MO, Apr 2008. to appear.

[14] D.P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, 2 edition, 2001.

[15] W. Zhang and J. Hu. Optimal quadratic regulation for linear switched systems. 2007. to be submitted for journal publication (available upon request).

[16] P. E. Caines and D. Q. Mayne. On the discrete time matrix riccati equation of optimal control. *Internation Journal of Control*, 12(5):785–794, 1970.