On the Value Functions of the Discrete-Time Switched LQR Problem

Wei Zhang, Jianghai Hu, and Alessandro Abate

***Abstract*—In this paper, we derive some important properties for the finite-horizon and the infinite-horizon value functions associated with the discrete-time switched LQR (DSLQR) problem. It is proved that any finite-horizon value function of the DSLQR problem is the pointwise minimum of a finite number of quadratic functions that can be obtained recursively using the so-called *switched Riccati mapping*. It is also shown that under some mild conditions, the family of the finite-horizon value functions is homogeneous (of degree 2), is uniformly bounded over the unit ball, and converges exponentially fast to the infinite-horizon value function. The exponential convergence rate of the value iterations is characterized analytically in terms of the subsystem matrices.**

***Index Terms*—Discrete-time switched LQR (DSLQR).**

I. INTRODUCTION

Optimal control of switched systems is a challenging problem that has received much research attention in recent years [1]–[4]. Compared with traditional optimal control problems [5], the distinctive feature of the optimal control of switched systems lies in the possibility of selecting the mode sequence and the switching instants. For a fixed mode sequence, a variational approach can be applied to derive certain gradient-based algorithms for optimizing the corresponding switching instants [1], [2]. However, finding the best mode sequence is a discrete optimization problem and is believed to be NP hard in general [2]. Recent research attention ([3], [4], [6]) has been focused on the optimal control problem of discrete-time switched linear systems with quadratic cost functions, which contains most of the interesting properties of the optimal control problem for general switched systems, while at the same time allows for efficient approaches to optimize the mode sequences. This optimal control problem can be viewed as an extension of the classical discrete-time LQR problem to the context of the switched linear systems, and are thus referred to as the discrete-time switched LQR (DSLQR) problem.

This paper studies several interesting properties of the finite-horizon and the infinite-horizon value functions associated with the DSLQR problem. It is shown that any finite-horizon value function of the DSLQR problem is the pointwise minimum of a finite number of quadratic functions that can be obtained recursively using the so-called *switched Riccati mapping*. Explicit expressions are also derived for the optimal switching-control law and the optimal continuous-control law, both of which are of state-feedback form and are homogeneous over the state space. In addition, the optimal continuous-control law is shown to be piecewise linear with different optimal feedback gains within different homogeneous regions of the state space. Although other researchers have also suggested a piecewise affine structure for the optimal control law ([4], [7], [8]), the analytical expression of the optimal feedback gain and in particular its connection with the Riccati equation of the classical LQR problem have not yet been explicitly presented.

Furthermore, several other interesting properties of the value functions are derived. It is proved that, under some mild conditions, the family of the finite-horizon value functions of the DSLQR problem is homogeneous (of degree 2), is uniformly bounded over the unit ball, and converges exponentially fast to the infinite-horizon value function. Finally, the exponential convergence rate of the value iteration is characterized analytically in terms of the subsystem matrices. These properties are not only of theoretical importance, but also play a crucial role in the design and analysis of various efficient algorithms for solving the DSLQR problem. Some preliminary algorithms developed based on the properties derived in this paper can be found in [9] and [10].

This paper is organized as follows. The DSLQR problem is formulated in Section II. Its value function is characterized analytically in Section III. Various interesting properties of the value function are derived in Section IV. The concluding remarks are given in Section V.

Notation: In this paper, n, p and M are some arbitrary finite positive integers, $\mathbb{Z}^+$ denotes the set of nonnegative integers, $\mathbb{M} \triangleq \{1, \ldots, M\}$ is a set of subsystem indices, I_n is the $n \times n$ identity matrix, $\|\cdot\|$ denotes the induced 2-norm in $\mathbb{R}^n$, $\mathcal{A}$ denotes the set of all the positive semidefinite (p.s.d.) matrices, $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and the largest eigenvalues, respectively, of a given p.s.d. matrix.

II. PROBLEM FORMULATION

Consider the discrete-time switched linear system described by

$$x(t+1) = A_{v(t)}x(t) + B_{v(t)}u(t), \quad t \in T_N \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the continuous state, $v(t) \in \mathbb{M} \triangleq \{1, \ldots, M\}$ is the switching control that determines the discrete mode, $u(t) \in \mathbb{R}^p$ is the continuous control and $T_N \triangleq \{0, \ldots, N-1\}$ is the control horizon with length N (possibly infinite). The sequence of pairs $\{(u(t), v(t))\}_{t=0}^{N-1}$ is called the *hybrid control sequence*. For each $i \in \mathbb{M}$, A_i, and B_i are constant matrices of appropriate dimensions and the pair (A_i, B_i) is called a subsystem. 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. At each time $t \in T_N$, denote by $\xi_{t,N} \triangleq (\mu_{t,N}, \nu_{t,N}) : \mathbb{R}^n \to \mathbb{R}^p \times \mathbb{M}$ the *(state-feedback) hybrid-control law* of system (1), where $\mu_{t,N} : \mathbb{R}^n \to \mathbb{R}^p$ is called the *(state-feedback) continuous-control law* and $\nu_{t,N} : \mathbb{R}^n \to \mathbb{M}$ is called the *(state-feedback) switching-control law*. A sequence of hybrid-control laws over the horizon T_N constitutes an *N-horizon feedback policy*: $\pi_N \triangleq \{\xi_{0,N}, \xi_{1,N}, \ldots, \xi_{N-1,N}\}$. If system (1) is driven by a feedback policy π_N, then the closed-loop system is given by

$$x(t+1) = A_{\nu_{t,N}(x(t))}x(t) + B_{\nu_{t,N}(x(t))}\mu_{t,N}(x(t)). \tag{2}$$

For a given initial state $x(0) = z \in \mathbb{R}^n$, the performance of the feedback policy π_N can be measured by the following cost functional:

$$J_{\pi_N}(z) = \psi(x(N)) + \sum_{t=0}^{N-1} L(x(t), \mu_{t,N}(x(t)), \nu_{t,N}(x(t)))$$

where $\psi : \mathbb{R}^n \to \mathbb{R}^+$ and $L : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{M} \to \mathbb{R}^+$ are the *terminal cost function* and the *running cost function*, respectively, given by

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

where $Q_f = Q_f^T \succeq 0$ is the terminal-state weighting matrix, and $Q_v = Q_v^T \succeq 0$ and $R_v = R_v^T \succ 0$ are the running weighting matrices for the state and the control, respectively, for subsystem $v \in \mathbb{M}$. When

Manuscript received March 18, 2009; revised July 14, 2009. This work was supported in part by the National Science Foundation under Grant CNS-0643805. Recommended by Associate Editor P. Shi.

W. Zhang and J. Hu are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906 USA (e-mail: zhang70@purdue.edu; jianghai@purdue.edu).

A. Abate is with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands (e-mail: a.abate@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2009.2031574

the control horizon N is infinite, the terminal cost will never be incurred and the objective function, which might be unbounded, becomes

$$J_{\pi_\infty}(z) = \sum_{t=0}^{\infty} L(x(t), \mu_{t,\infty}(x(t)), \nu_{t,\infty}(x(t))). \quad (3)$$

Denote by Π_N the set of all admissible N-horizon policies. The discrete-time switched LQR problem (DSLQR) is formulated below.

1) Problem 1 (DSLQR Problem): For a given initial state $z \in \mathbb{R}^n$ and a possibly infinite positive integer N, find the N-horizon policy $\pi_N \in \Pi_N$ that minimizes $J_{\pi_N}(z)$ subject to (2).

To solve Problem 1, for each time $t \in T_N$, we define the value function $V_{t,N} : \mathbb{R}^n \to \mathbb{R}$ as shown in (1)

$$V_{t,N}(z) = \inf_{\substack{u(j)\in\mathbb{R}^p, v(j)\in\mathbb{M} \\ t\le j\le N-1}} \left\{ \psi(x(N)) + \sum_{j=t}^{N-1} L(x(j),u(j),v(j)) \right| , \quad \text{subject to eq. (1) with } x(t) = z \Bigg\}$$

The $V_{t,N}(z)$ so defined is the minimum cost-to-go starting from state z at time t. The minimum cost for the DSLQR problem with initial state $x(0) = z$ 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, i.e.,

$$V_{t,N}(z) = V_{t+m,N+m}(z)$$

for all $z \in \mathbb{R}^n$ and all integers $m \ge -t$. In the rest of this paper, when no ambiguity arises, we will denote by $V_k(z) \triangleq V_{N-k,N}(z)$ and $\xi_k \triangleq \xi_{N-k,N}$ the value function and the hybrid-control law, respectively, at time $t = N-k$ when there are k time steps left. With the new notations, the N-horizon policy π_N can also be written as $\pi_N = \{\xi_N, \dots, \xi_1\}$. For any positive integer k, the control law ξ_k can be thought of as the first step of a k-horizon policy.

By a standard result of Dynamic Programming [11], for any finite integer N, the value function V_N can be obtained recursively using the one-stage *value iteration*

$$V_{k+1}(z) = \inf_{u,v}\{L(z,u,v) + V_k(A_v z + B_v u)\}, \quad \forall\, z \in \mathbb{R}^n$$

with initial condition $V_0(z) = \psi(z), \forall\, z \in \mathbb{R}^n$. Denote by $V_\infty(\cdot)$ the pointwise limit (whenever it exists) of the sequence of functions $\{V_k(\cdot)\}_{k=0}^{\infty}$ generated by the value iterations. It is well known [11, ch. 3] that even if $V_\infty(z)$ exists, it may not always coincide with the infinite-horizon value function. To emphasize its substantial difference from the finite-horizon value function, the infinite-horizon value function is specially denoted by $V^*(z)$, i.e., $V^*(z) = \inf_{\pi_\infty \in \Pi_\infty} J_{\pi_\infty}(z)$.

III. Analytical Characterization of the Finite-Horizon Value Function

When $M = 1$, the DSLQR problem reduces to the classical LQR problem. Denote by (A, B, Q, R) the system and weighting matrices associated with this simple instance. It is well known that when N is finite, the value functions of this LQR problem are of the following quadratic form:

$$V_k(z) = z^T P_k z, \quad k = 0, \dots, N \quad (4)$$

where $\{P_k\}_{k=0}^N$ is a sequence of positive semidefinite (p.s.d.) matrices satisfying the difference Riccati equation (DRE) ([12])

$$P_{k+1} = Q + A^T P_k A - A^T P_k B \left(R + B^T P_k B\right)^{-1} B^T P_k A \quad (5)$$

with initial condition $P_0 = Q_f$. Some results of the classical LQR problem are summarized in the following lemma.

Lemma 1 ([13],[14]): Let $\{P_k\}_{k=0}^N$ be generated by the DRE (5), then:

1) for each $k = 0, \dots, N-1$, if $P_k \in \mathcal{A}$, then $P_{k+1} \in \mathcal{A}$;
2) if (A, B) is stabilizable, then $V_k(z) \to V^*(z)$ for all $z \in \mathbb{R}^n$ as $k \to \infty$;
3) let $Q = C^T C$. If (A, B) stabilizable and (C, A) detectable, then the optimal trajectory of the LQR problem is exponentially stable.

In general, when $M \ge 2$, the value function $V_k(z)$ is no longer of a simple quadratic form as in (4). Nevertheless, the notion of the DRE can be generalized to the Switched LQR problem. The DRE (5) can be viewed as a mapping from $\mathcal{A}$ to $\mathcal{A}$ depending on the matrices (A, B, Q, R). We call this mapping the *Riccati Mapping* and denote by $\rho_i : \mathcal{A} \to \mathcal{A}$ the Riccati Mapping of subsystem $i \in \mathbb{M}$, i.e.,

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

Definition 1: Let $2^{\mathcal{A}}$ be the power set of $\mathcal{A}$. The mapping $\rho_{\mathbb{M}} : 2^{\mathcal{A}} \to 2^{\mathcal{A}}$ defined by

$$\rho_{\mathbb{M}}(\mathcal{H}) = \{\rho_i(P) : i \in \mathbb{M} \text{ and } P \in \mathcal{H}\}, \quad \forall\, \mathcal{H} \in 2^{\mathcal{A}}$$

is called the *Switched Riccati Mapping* (SRM) associated with Problem 1.

In words, the SRM maps a *set* of p.s.d. matrices to another *set* of p.s.d. matrices and each matrix in $\rho_{\mathbb{M}}(\mathcal{H})$ is obtained by taking the classical Riccati mapping of some matrix in $\mathcal{H}$ through some subsystem $i \in \mathbb{M}$.

Definition 2: The sequence of sets $\{\mathcal{H}_k\}_{k=0}^N$ generated iteratively by $\mathcal{H}_{k+1} = \rho_{\mathbb{M}}(\mathcal{H}_k)$ with initial condition $\mathcal{H}_0 = \{Q_f\}$ is called the *Switched Riccati Sets* (SRSs) of Problem 1.

The SRSs always start from a singleton set $\{Q_f\}$ and evolve according to the SRM. For any finite N, the set $\mathcal{H}_N$ consists of at most M^N p.s.d. matrices. An important fact about the DSLQR problem is that its value functions are completely characterized by the SRSs.

Theorem 1: For $k = 0, \dots, N$, the value function for the DSLQR problem at time $N - k$, i.e., with k time steps left, is

$$V_k(z) = \min_{P\in\mathcal{H}_k} z^T P z. \quad (6)$$

Furthermore, for $z \in \mathbb{R}^n$ and $k = 1, \dots, N$, if we define

$$(P_k^*(z), i_k^*(z)) = \underset{(P\in\mathcal{H}_{k-1}, i\in\mathbb{M})}{\arg\min} z^T \rho_i(P) z \quad (7)$$

then the optimal hybrid-control law at state z and time $t = N - k$ is $\xi_k^*(z) = (\mu_k^*(z), \nu_k^*(z))$, where $\mu_k^*(z) = -K_{i_k^*(z)}(P_k^*(z))\, z$ and $\nu_k^*(z) = i_k^*(z)$. Here, $K_i(P)$ is the optimal-feedback gain for subsystem i with matrix P, i.e.,

$$K_i(P) \triangleq \Big(R_i + B_i^T P B_i\Big)^{-1} B_i^T P A_i. \quad (8)$$

Proof: The theorem is proved by induction. It is obvious that for $k = 0$ the value function is $V_0(z) = z^T Q_f z$, satisfying (6). Now suppose (6) holds for some $k \le N-1$, i.e., $V_k(z) = \min_{P\in\mathcal{H}_k} z^T P z$. We shall show that it is also true for $k+1$. By the principle of dynamic programming and noting that $V_k(\cdot)$ represents the value function at time $N-k$, the value function at time $N-(k+1)$ can be recursively computed as (9), shown at the bottom of the next page. Since the quantity inside the bracket is quadratic in u, the optimal u^* can be easily found to be

$$u^* = -\Big(R_i + B_i^T P B_i\Big)^{-1} B_i^T P A_i z = -K_i(P) z \quad (10)$$

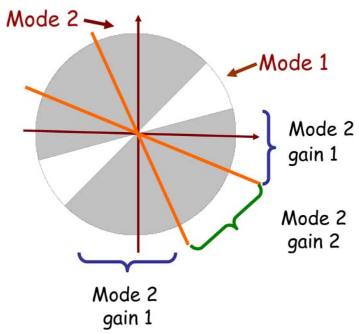

Fig. 1. Typical optimal decision regions of a two-switched system, where mode 1 is optimal within the white region and mode 2 is optimal within the gray region. The optimal mode region is divided into smaller homogeneous regions, each of which corresponds to a different optimal-feedback gain.

where $K_i(P)$ is the matrix defined in (8). Substituting u^* into (9), we obtain $V_{k+1}(z) = \min_{i\in\mathbb{M},P\in\mathcal{H}_k} z^T\rho_i(P)z$. Observing that $\{\rho_i(P) : i \in \mathbb{M}, P \in \mathcal{H}_k\} = \rho_{\mathbb{M}}(\mathcal{H}_k) = \mathcal{H}_{k+1}$, we have $V_{k+1}(z) = \min_{P\in\mathcal{H}_{k+1}} z^T P z$. In addition, let $P_k^*(z)$ and $i_k^*(z)$ be defined as in (7). Then it can easily be seen from the above derivation that $\left(-K_{i_{k+1}^*(z)}\left(P_{k+1}^*(z)\right)z, i_{k+1}^*(z)\right)$ is the optimal decision at time $N-(k+1)$ that achieves the minimum cost $V_{k+1}(z)$. ∎

Remark 1: Theorem 1 is not a trivial variation of the results in [4] and [15], which deal with piecewise affine systems, where the mode sequence $v(t)$ is determined by the evolution of the continuous state instead of being a decision variable independent of the continuous state as in the present DSLQR problem.

Remark 2: The piecewise quadratic structure of the value function has also been suggested in [3] for the infinite-horizon DSLQR problem. Compared with [3], the contribution of Theorem 1 lies in the explicit characterization of the value function in terms of the SRM and its connection to the optimal-feedback gain and the Riccati equation of the classical LQR problem.

Compared with the classical LQR problem, the value function of the DSLQR problem is no longer a single quadratic function; it becomes the pointwise minimum of a finite number of quadratic functions. At each time step, instead of having a single optimal-feedback gain for the entire state space, the optimal state feedback gain becomes state dependent. Furthermore, the minimizer $(P_k^*(z), i_k^*(z))$ of (7) is radially invariant, indicating that at each time step all the points along the same radial direction have the same optimal hybrid-control law. These properties are illustrated in Fig. 1 using an example in $\mathbb{R}^2$ with two subsystems: at each time step, the state space is decomposed into several homogeneous *decision regions*, each of which corresponds to a pair of optimal mode and optimal-feedback gain. In addition, all the gray homogeneous regions have the same optimal mode, say mode 2. It is worth mentioning that in a higher dimensional state space, the homogeneous decision regions may become nonconvex and rather complicated. A salient feature of the DSLQR problem is that all these complex decision regions are completely encoded in a finite number of matrices in the SRSs.

IV. Properties of the Value Functions

In this section, we will derive various important properties for the family of the finite-horizon value functions $\{V_N(z)\}_{N\geq 0}^{\infty}$ and the infinite-horizon value function $V^*(z)$. These properties are crucial in the design and analysis of efficient algorithms for solving the DSLQR problems [9].

We first introduce some notations. Define $\lambda_Q^- = \min_{i\in\mathbb{M}}\{\lambda_{\min}(Q_i)\}$, and $\lambda_f^+ = \lambda_{\max}(Q_f)$. Denote by $x_{z,N}^*(t)$ for $0 \leq t \leq N$ an optimal trajectory originating from z at time 0 and denote by $\left(u_{z,N}^*(t), v_{z,N}^*(t)\right)$ the corresponding optimal hybrid-control sequence.

A. Homogeneity and Boundedness

Lemma 2 (Homogeneity): For any $z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ and $N \in \mathbb{Z}^+$, we have $V^*(\lambda z) = \lambda^2 V^*(z)$ and $V_N(\lambda z) = \lambda^2 V_N(z)$.

The homogeneity of V_N is clear from Theorem 1. The homogeneity of V^* follows from the linearity of the subsystems and the quadratic nature of the objective function (3).

The properties of the value functions presented in the rest of this section are based on the following stabilizability assumption:

$$\text{(A1)} \quad \text{At least one subsystem is stabilizable.}$$

Lemma 3 (Boundedness): Under assumption (A1), there exists a finite constant β 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 β is given by

$$\beta = \left(\|Q_f\| + \left\|Q_i + F^T R_i F\right\|\right)\cdot\left(\sum_{j=0}^{\infty}\left\|\bar{A}_i^j\right\|^2\right) < \infty. \tag{11}$$

Proof: Suppose subsystem (A_i, B_i) is stabilizable. Let $\left\{P_k^{(i)}\right\}_{k=0}^{\infty}$ be the sequence of matrices generated by the Riccati mapping using only subsystem i, i.e., $P_{k+1}^{(i)} = \rho_i\left(P_k^{(i)}\right)$ with $P_0^{(i)} = Q_f$. Since the switched system (1) can stay in subsystem (A_i, B_i) all the time, the value function of the DSLQR problem must be no greater than the value function of the LQR problem for 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 β given in (11) is an upper bound for the Euclidean norm of all the matrices in $\left\{P_k^{(i)}\right\}_{k=0}^{\infty}$. Let F be a feedback gain for which $\bar{A}_i = A_i - B_i F$ is stable. Define $\left\{\tilde{P}_k^{(i)}\right\}_{k=0}^{\infty}$ iteratively by

$$\tilde{P}_{k+1}^{(i)} = Q_i + \bar{A}_i^T \tilde{P}_k^{(i)} \bar{A}_i + F^T R_i F, \text{ with } \tilde{P}_0^{(i)} = Q_f. \tag{12}$$

In the above equation, if $F = K_i\left(\tilde{P}_k^{(i)}\right)$ for each k, where $K_i(\cdot)$ is defined in (8), then $\tilde{P}_k^{(i)}$ would coincide with $P_k^{(i)}$. In other words, $\tilde{P}_k^{(i)}$ defines the quadratic energy cost of using the stabilizing feedback gain F instead of the time-dependent optimal-feedback gain of

$$\begin{aligned} V_{k+1}(z) &= \inf_{i\in\mathbb{M}, u\in\mathbb{R}^p}\left[z^T Q_i z + u^T R_i u + V_k(A_i z + B_i u)\right] \\ &= \inf_{i\in\mathbb{M}, P\in\mathcal{H}_k, u\in\mathbb{R}^p}\left[z^T(Q_i + A_i^T P A_i)z + u^T\left(R_i + B_i^T P B_i\right)u + 2z^T A_i^T P B_i u\right] \end{aligned} \tag{9}$$

the k-horizon LQR problem. By a standard result of the Riccati equation theory ([13, Theor. 2.1]), we have $P_k^{(i)} \preceq \tilde{P}_k^{(i)}$ for all $k \geq 0$. Thus, it suffices to show $\left\|\tilde{P}_k^{(i)}\right\| \leq \beta$ for each $k \geq 0$. By (12), we have

$$\begin{aligned}\tilde{P}_k^{(i)} &= \tilde{P}_0^{(i)} + \sum_{j=1}^{k}\left(\tilde{P}_j^{(i)} - \tilde{P}_{j-1}^{(i)}\right)\\ &= \tilde{P}_0^{(i)} + \sum_{j=0}^{k-1}\left(\bar{A}_i^T\right)^j\left(\tilde{P}_1^{(i)} - \tilde{P}_0^{(i)}\right)(\bar{A}_i)^j\\ &= Q_f + \sum_{j=0}^{k-1}\left(\bar{A}_i^T\right)^{j+1} Q_f(\bar{A}_i)^{j+1}\\ &\quad + \sum_{j=0}^{k-1}\left(\bar{A}_i^T\right)^j\left(Q_i - Q_f + F^T R_i F\right)(\bar{A}_i)^j\\ &\preceq \left(\bar{A}_i^T\right)^k Q_f(\bar{A}_i)^k + \sum_{j=0}^{\infty}\left(\bar{A}_i^T\right)^j\left(Q_i + F^T R_i F\right)(\bar{A}_i)^j.\end{aligned}$$

Thus

$$\begin{aligned}\left\|P_k^{(i)}\right\| &\leq \left\|\tilde{P}_k^{(i)}\right\|\\ &\leq \left(\|Q_f\| + \left\|Q_i + F^T R_i F\right\|\right)\left(\sum_{j=0}^{\infty}\left\|\bar{A}_i^j\right\|^2\right).\end{aligned}$$

Note that the formula of the geometric series does not directly apply here, as the 2-norm of a stable matrix may not be strictly less than 1 in general. However, it is shown in [16, ch. 5] that $\lim_{k\to\infty}\left\|\bar{A}_i^k\right\|^{1/k} = \rho(\bar{A}_i) < 1$, where $\rho(\cdot)$ denotes the spectral radius of a given matrix. This guarantees that $\left\|\bar{A}_i^j\right\| < (1-\epsilon)^j$ for some small $\epsilon > 0$ and all large j. Therefore, $\sum_{j=0}^{\infty}\left\|\bar{A}_i^j\right\|^2 < \infty$ and the proposition is proven. ■

B. Exponential Stability of the Optimal Trajectory

In view of part 3) of Lemma 1, to ensure the stability of the optimal trajectory, it is natural to assume that each subsystem is stabilizable and detectable. Unfortunately, such a natural extension does not hold in the DSLQR case. As an example, consider the following DSLQR problem:

$$\begin{aligned}A_1 &= \begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}, \quad A_2 = \begin{bmatrix}0 & 0.5\\ 0.5 & 0\end{bmatrix}\\ Q_1 &= \begin{bmatrix}100 & 0\\ 0 & 0\end{bmatrix}, \quad Q_2 = \begin{bmatrix}0 & 0\\ 0 & 100\end{bmatrix}\\ x_0 &= \begin{bmatrix}0\\ 1\end{bmatrix}, \quad Q_f = 0, \text{ and } B_i = 0\\ R_i &= 0, \quad i = 1, 2.\end{aligned} \tag{13}$$

Let the horizon N be arbitrary (possibly infinite) and let $x^*(\cdot)$ be the optimal trajectory of this DSLQR problem with initial condition $x^*(0) = x_0$. Notice that each subsystem is stabilizable and detectable. However, it can be easily verified that $x^*(t) = [0, 1]^T$ if t is even and $x^*(t) = [2, 0]^T$ otherwise. Thus, to ensure the stability of the optimal trajectory, we introduce the following assumption:

$$(A2) \quad Q_i \succ 0, \quad \forall\, i \in \mathbb{M}.$$

Theorem 2: Under assumptions (A1) and (A2), the N-horizon optimal trajectory originating from z at time $t = 0$, namely, $x_{z,N}^*(\cdot)$, satisfies the following inequalities:

$$\left\|x_{z,N}^*(t)\right\|^2 \leq \frac{\beta}{\lambda_Q^-}\gamma^t\|z\|^2, \quad \text{for } t = 1, \ldots, N-1$$

$$\text{and } \left\|x_{z,N}^*(N)\right\|^2 \leq \frac{\beta\zeta^2}{\lambda_Q^-}\gamma^{N-1}\|z\|^2 \tag{14}$$

where β is defined in Lemma 3

$$\gamma = \frac{1}{1 + \frac{\lambda_Q^-}{\beta}} < 1 \text{ and } \zeta = \max_{i\in\mathbb{M}}\|A_i - B_i K_i(Q_f)\|. \tag{15}$$

In other words, the optimal trajectory is exponentially stable with a decay rate γ.

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

$$\begin{aligned}\tilde{V}_{N-(t-1)} - \tilde{V}_{N-t} &= L(\tilde{x}(t-1), \tilde{u}(t-1), \tilde{v}(t-1))\\ &\geq \tilde{x}(t-1)^T Q_{\tilde{v}(t-1)}\tilde{x}(t-1)\\ &\geq \lambda_Q^-\|\tilde{x}(t-1)\|^2\\ &\geq \frac{\lambda_Q^-}{\beta}\tilde{V}_{N-(t-1)} \geq \frac{\lambda_Q^-}{\beta}\tilde{V}_{N-t}.\end{aligned}$$

Hence, we have $\tilde{V}_{N-t} \leq 1/1 + \lambda_Q^-/\beta\tilde{V}_{N-(t-1)}$ for $t = 1, \ldots, N$. Therefore, $\tilde{V}_{N-t} \leq \left(1/1 + \lambda_Q^-/\beta\right)^t \tilde{V}_N$. Obviously, for $t \leq N-1$, $\tilde{V}_{N-t} \geq \tilde{x}(t)^T Q_{\tilde{v}(t)}\tilde{x}(t) \geq \lambda_Q^-\|\tilde{x}(t)\|^2$. Thus

$$\begin{aligned}\|\tilde{x}(t)\|^2 &\leq \frac{1}{\lambda_Q^-}\tilde{V}_{N-t} \leq \frac{1}{\lambda_Q^-}\left(\frac{1}{1 + \frac{\lambda_Q^-}{\beta}}\right)^t \tilde{V}_N\\ &\leq \frac{\beta}{\lambda_Q^-}\left(\frac{1}{1 + \frac{\lambda_Q^-}{\beta}}\right)^t \|z\|^2 = \frac{\beta}{\lambda_Q^-}\gamma^t\|z\|^2.\end{aligned} \tag{16}$$

For $t = N$, by Theorem 1, we have that $\tilde{x}(N) = (A_i - B_i K_i(Q_f)) \cdot \tilde{x}(N-1)$ for some $i \in \mathbb{M}$. Therefore, $\|\tilde{x}(N)\|^2 \leq \zeta^2\|\tilde{x}(N-1)\|^2$, where ζ is defined in (15), and then the desired result follows from (16). ■

Remark 3: It is worth pointing out that the decay rate γ given in (15) could be conservative.

C. Exponential Convergence of Value Iteration

Some classical results on the convergence of the value iteration of general DP problem can be found in [11]. Most of these results require either a discount factor with magnitude strictly less than 1 or that $\psi(z) \leq V^*(z)$ for all $z \in \mathbb{R}^n$. Neither is true for the general DSLQR problem with a nontrivial terminal cost. A more recent convergence result is given in [3] and [6], where the aforementioned assumptions are replaced with some other conditions on $V^*(z)$. Since the infinite-horizon value function $V^*(z)$ of the DSLQR problem is usually unknown, the conditions in [3], [6] are not easy to check. In view of these limitations, a further study on the convergence of the value iteration of the DSLQR problem is necessary.

By part 2) of Lemma 1, for the classical LQR problem, if the system is stabilizable, then the value iteration converges to the infinite-horizon value function. For the DSLQR problem, however, Assumption (A1) alone is not enough to ensure the convergence of the value functions. For example, consider the DSLQR problem with matrices defined by (13) except that $Q_f = I_2$. Although each subsystem is stable, it can be easily seen that $V_N(x_0)$ is 2 if N is an odd number and is 1 otherwise. Thus, the limit of $V_N(x_0)$ as $N \to \infty$ does not exist.

In the following we shall show that the value iteration will converge exponentially fast if both (A1) and (A2) are satisfied. The following lemma provides a bound for the difference between two value functions with different horizons and is the key in proving the convergence result.

Lemma 4: Let N_1 and N_2 be positive integers such that $N_1 > N_2$. For any $z \in \mathbb{R}^n$, the difference between the N_1-horizon value function and the N_2-horizon value function can be bounded as follows:

$$\begin{aligned} &V_{N_1-N_2}\left(x^*_{z,N_1}(N_2)\right) - \psi\left(x^*_{z,N_1}(N_2)\right) \\ &\quad \le V_{N_1}(z) - V_{N_2}(z) \\ &\quad \le V_{N_1-N_2}\left(x^*_{z,N_2}(N_2)\right) - \psi\left(x^*_{z,N_2}(N_2)\right). \end{aligned} \tag{17}$$

Proof: Let $z_2 = x^*_{z,N_2}(N_2)$. Define a new N_1-horizon trajectory $\tilde{x}(\cdot)$ as

$$\tilde{x}(t) = \begin{cases} x^*_{z,N_2}(t), & t \le N_2 \\ x^*_{z_2,N_1-N_2}(t-N_2), & N_2 < t \le N_1. \end{cases} \tag{18}$$

As shown in Fig. 2 (the dashdot line), $\tilde{x}(\cdot)$ is obtained by first following the N_2-horizon optimal trajectory and then the $(N_1 - N_2)$-horizon optimal trajectory. Let $(\tilde{u}(\cdot), \tilde{v}(\cdot))$ be the hybrid controls corresponding to $\tilde{x}$. Then by the definition of the value function, we have

$$\begin{aligned} V_{N_1}(z) &\le \sum_{t=0}^{N_1-1} L(\tilde{x}(t), \tilde{u}(t), \tilde{v}(t)) + \psi(\tilde{x}(N_1)) \\ &= \sum_{t=0}^{N_2-1} L\left(x^*_{z,N_2}(t), u^*_{z,N_2}(t), v^*_{z,N_2}(t)\right) \\ &\quad + \sum_{t=0}^{N_1-N_2-1} L\Bigg(x^*_{z_2,N_1-N_2}(t), u^*_{z_2,N_1-N_2}(t) \\ &\qquad\qquad v^*_{z_2,N_1-N_2}(t)\Bigg) \\ &\quad + \psi\left(x^*_{z_2,N_1-N_2}(N_1-N_2)\right) \\ &= V_{N_2}(z) - \psi\left(x^*_{z,N_2}(N_2)\right) \\ &\quad + V_{N_1-N_2}\left(x^*_{z,N_2}(N_2)\right). \end{aligned} \tag{19}$$

Equation (19) describes exactly the second inequality in (17). To prove the first one, define an N_2-horizon trajectory $\hat{x}(\cdot)$ as the solid line in Fig. 2 by taking the first N_2 steps of x^*_{z,N_1}, i.e., $\hat{x}(t) = x^*_{z,N_1}(t)$ for $0 \le t \le N_2$ and let $(\hat{u}(\cdot), \hat{v}(\cdot))$ be the corresponding hybrid control sequence. Then

$$\begin{aligned} V_{N_2}(z) &\le \sum_{t=0}^{N_2-1} L(\hat{x}(t), \hat{u}(t), \hat{v}(t)) + \psi(\hat{x}(N_2)) \\ &= \sum_{t=0}^{N_2-1} L(x^*_{z,N_1}(t), u^*_{z,N_1}(t), v^*_{z,N_1}(t)) \\ &\quad + \psi\left(x^*_{z,N_1}(N_2)\right) \\ &= V_{N_1}(z) - V_{N_1-N_2}\left(x^*_{z,N_1}(N_2)\right) \\ &\quad + \psi\left(x^*_{z,N_1}(N_2)\right) \end{aligned} \tag{20}$$

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 (19) and (20). ■

With a nontrivial terminal cost, the N-horizon value function $V_N(z)$ may not be monotone as N increases. Nevertheless, by Lemma 4, the difference between $V_{N_1}(z)$ and $V_{N_2}(z)$ can be bounded by the quadratic functions of $x^*_{z,N_1}(N_2)$ and $x^*_{z,N_2}(N_2)$. By Theorem 2, we know both quantities converge to zero as N_1 and N_2 grow to infinity. This will guarantee that by choosing N_1 and N_2 large enough, the upper and lower bounds in (17) can be made arbitrarily small. The convergence of the value iteration can thus be established.

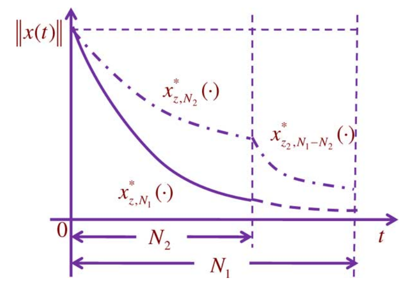

Fig. 2. Illustrating the proof of Lemma 4, where the dashdot line represents the trajectory $\tilde{x}(\cdot)$, the solid line represents the trajectory $\hat{x}(\cdot)$ and the solid line together with the dashed line represents the trajectory $x^*_{z,N_1}(\cdot)$.

Theorem 3: Under assumptions (A1) and (A2), for any $N_1 > N_2$, we have

$$|V_{N_1}(z) - V_{N_2}(z)| \le \alpha\gamma^{N_2}\|z\|^2 \tag{21}$$

where $\alpha = \max\{1, \zeta^2/\gamma\} \cdot (\beta + \lambda_f^+)\beta/\lambda_Q^-$, with β, $\gamma < 1$ and ζ defined in (11) and (15).

Proof: By Theorem 2, for any $z \in \mathbb{R}^n$, we have $\left\|x^*_{z,N_2}(N_2)\right\|^2 \le \zeta^2\beta/\lambda_Q^-\gamma\gamma^{N_2}\|z\|^2$ and $\left\|x^*_{z,N_1}(N_2)\right\|^2 \le \beta/\lambda_Q^-\gamma^{N_2}\|z\|^2$. Hence

$$\begin{aligned} V_{N_1-N_2}(x^*_{z,N_2}(N_2)) &\le \beta\left\|x^*_{z,N_2}(N_2)\right\|^2 \le \frac{\zeta^2\beta^2}{\lambda_Q^-\gamma}\gamma^{N_2}\|z\|^2 \\ \psi(x^*_{z,N_2}(N_2)) &\le \lambda_f^+\left\|x^*_{z,N_2}(N_2)\right\|^2 \\ &\le \frac{\lambda_f^+\zeta^2\beta}{\lambda_Q^-\gamma}\gamma^{N_2}\|z\|^2 \\ V_{N_1-N_2}\left(x^*_{z,N_1}(N_2)\right) &\le \beta\left\|x^*_{z,N_1}(N_2)\right\|^2 \le \frac{\beta^2}{\lambda_Q^-}\gamma^{N_2}\|z\|^2 \\ \psi\left(x^*_{z,N_1}(N_2)\right) &\le \lambda_f^+\left\|x^*_{z,N_1}(N_2)\right\|^2 \le \frac{\lambda_f^+\beta}{\lambda_Q^-}\gamma^{N_2}\|z\|^2. \end{aligned}$$

Thus, by Lemma 4 we have

$$|V_{N_1}(z) - V_{N_2}(z)| \le \max\left\{1, \frac{\zeta^2}{\gamma}\right\} \cdot \frac{(\beta + \lambda_f^+)\beta}{\lambda_Q^-}\gamma^{N_2}\|z\|^2.$$

■

By Theorem 3, assumptions (A1) and (A2) together imply the exponential convergence of the value iteration. In general, the limiting function $V_\infty(z)$ may not coincide with the infinite-horizon value function $V^*(z)$. The following Theorem shows that the two functions agree for the DSLQR problem.

Theorem 4: Under assumptions (A1) and (A2), $V_\infty(z) = V^*(z)$ for each $z \in \mathbb{R}^n$.

Proof: For any finite N, we know that

$$V_N(z) = \sum_{t=0}^{N-1} L(x^*_{z,N}(t), u^*_{z,N}(t), v^*_{z,N}(t)) + \psi(x^*_{z,N}(N)).$$

By the optimality of $V^*(z)$, we have

$$\begin{aligned} V^*(z) &\le \sum_{t=0}^{N-1} L(x^*_{z,N}(t), u^*_{z,N}(t), v^*_{z,N}(t)) + V^*(x^*_{z,N}(N)) \\ &= V_N(z) - \psi(x^*_{z,N}(N)) + V^*(x^*_{z,N}(N)). \end{aligned}$$

By Theorem 3 and Theorem 2, as $N \to \infty$, $V_N(z) \to V_\infty(z)$, $\psi(x^*_{z,N}(N)) \to 0$ and $V^*(x^*_{z,N}(N)) \to 0$. Therefore, $V^*(z) \le V_\infty(z)$. We now prove the other direction. Notice that by (A2) we must have $V^*(z) = \inf_{\pi_\infty \in \Pi^s_\infty} J_{\pi_\infty}(z)$, where Π^s_∞ denotes the set of all the infinite-horizon stabilizing policies. Let π_∞ be an arbitrary policy in Π^s_∞ and let $\hat{x}(\cdot)$ and $(\hat{u}(\cdot), \hat{v}(\cdot))$ be the corresponding trajectory and the hybrid control sequence, respectively. Since $\hat{x}(t) \to 0$ as $t \to \infty$, for any $\epsilon > 0$, there always exists an N_1 such that $\psi(\hat{x}(t)) \leq \epsilon$ for all $t \geq N_1$. Hence, for all $N \geq N_1$

$$\begin{aligned} V_N(z) &\leq \sum_{t=0}^{N-1} L(\hat{x}(t), \hat{u}(t), \hat{v}(t)) + \psi(\hat{x}(N)) \\ &\leq \sum_{t=0}^{N-1} L(\hat{x}(t), \hat{u}(t), \hat{v}(t)) + \epsilon \leq J_{\pi_\infty}(z) + \epsilon. \end{aligned}$$

Let $N \to \infty$, we have $V_\infty(z) \leq J_{\pi_\infty}(z) + \epsilon$, $\forall\, \pi_\infty \in \Pi^s_\infty$. Thus, $V_\infty(z) \leq V^*(z) + \epsilon$ and the theorem is proved as ϵ is arbitrary.

Remark 4: Compared with the previous work [3], [11], our convergence result derived specially for the DSLQR problem has several distinctions. First, it allows general terminal cost, which is especially important for the finite-horizon DSLQR problems. In addition, the convergence conditions are expressed in terms of the subsystem matrices rather than the infinite-horizon value function [3], and thus become much easier to verify. Finally, by Theorem 3, for a given tolerance on the optimal cost, the required number of iterations can be computed before the actual computation starts. This provides an efficient means to stop the value iteration with guaranteed suboptimal performance.

V. Conclusion

A number of important properties of the value functions of the DSLQR problem are derived. In particular, we have proved that any finite-horizon value function is the pointwise minimum of a finite number of quadratic functions that can be obtained recursively using the SRM. It has also been shown that under some mild conditions, the family of the finite-horizon value functions is homogeneous of degree 2, is uniformly bounded over the unit ball and converges exponentially fast to the corresponding infinite-horizon value function. Future research will focus on employing these properties to efficiently solve the DLQRS problem with guaranteed suboptimal performance.

References

[1] X. Xu and P. J. Antsaklis, “Optimal control of switched systems based on parameterization of the switching instants,” *IEEE Trans. Autom. Control*, vol. 49, no. 1, pp. 2–16, Jan. 2004.

[2] M. Egerstedt, Y. Wardi, and F. Delmotte, “Optimal control of switching times in switched dynamical systems,” *IEEE Trans. Autom. Control*, vol. 51, no. 1, pp. 110–115, Jan. 2006.

[3] B. Lincoln and A. Rantzer, “Relaxing dynamic programming,” *IEEE Trans. Autom. Control*, vol. 51, no. 8, pp. 1249–1260, Aug. 2006.

[4] F. Borrelli, M. Baotic, A. Bemporad, and M. Morari, “Dynamic programming for constrained optimal control of discrete-time linear hybrid systems,” *Automatica*, vol. 41, pp. 1709–1721, Oct. 2005.

[5] F. L. Lewis and V. L. Syrmos, *Optimal Control* , 2 ed. New York: Wiley, 1995.

[6] A. Rantzer, “Relaxed dynamic programming in switching systems,” *IEE Proc.—Control Theory & Applic.*, vol. 153, no. 5, pp. 567–574, Sep. 2006.

[7] A. Bemporad, A. Giua, and C. Seatzu, “Synthesis of state-feedback optimal controllers for continuous-time switched linear systems,” in *Proc. IEEE Conf. Decision and Control*, Las Vegas, NV, Dec. 2002, pp. 3182–3187.

[8] C. Seatzu, D. Corona, A. Giua, and A. Bemporad, “Optimal control of continuous-time switched affine systems,” *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 726–741, May 2006.

[9] W. Zhang, A. Abate, and J. Hu, “Efficient suboptimal solutions of switched LQR problems,” in *Proc. American Control Conf.*, St. Louis, MO, Jun. 2009.

[10] W. Zhang, J. Hu, and A. Abate, A Study of the Discrete-Time Switched LQR Problem Dept. Elect. and Comput. Eng., Purdue Univ., West Lafayette, IN, 2009, Tech. Rep. TR ECE 09-03.

[11] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Nashua, NH: Athena Scientific, 2001, vol. 2.

[12] D. Kleinman, “On an iterative technique for Riccati equation computations,” *IEEE Trans. Autom. Control*, vol. AC-13, no. 1, pp. 114–115, Feb. 1968.

[13] P. E. Caines and D. Q. Mayne, “On the discrete time matrix Riccati equation of optimal control,” *Int. J. Control*, vol. 12, no. 5, pp. 785–794, 1970.

[14] S. Chan, G. Goodwin, and K. Sin, “Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems,” *IEEE Trans. Autom. Control*, vol. AC-29, no. 2, pp. 110–118, Feb. 1984.

[15] A. Rantzer and M. Johansson, “Piecewise linear quadratic optimal control,” *IEEE Trans. Autom. Control*, vol. 45, no. 4, pp. 629–637, Apr. 2000.

[16] R. Horn and C. Johnson, *Matrix Analysis* . Cambridge, U.K.: Cambridge University Press, 1985.