Optimal Control of Switched Systems Based on Parameterization of the Switching Instants

Xuping Xu, *Member, IEEE,* and Panos J. Antsaklis, *Fellow, IEEE*

***Abstract*—This paper presents a new approach for solving optimal control problems for switched systems. We focus on problems in which a prespecified sequence of active subsystems is given. For such problems, we need to seek both the optimal switching instants and the optimal continuous inputs. In order to search for the optimal switching instants, the derivatives of the optimal cost with respect to the switching instants need to be known. The most important contribution of the paper is a method which first transcribes an optimal control problem into an equivalent problem parameterized by the switching instants and then obtains the values of the derivatives based on the solution of a two point boundary value differential algebraic equation formed by the state, costate, stationarity equations, the boundary and continuity conditions, along with their differentiations. This method is applied to general switched linear quadratic problems and an efficient method based on the solution of an initial value ordinary differential equation is developed. An extension of the method is also applied to problems with internally forced switching. Examples are shown to illustrate the results in the paper.**

***Index Terms*—Hybrid systems, linear quadratic problems, optimal control, switched systems.**

I. Introduction

SWITCHED systems are a particular class of hybrid systems that consist of several subsystems and switching laws orchestrating the active subsystem at each time instant. Many real-world processes such as chemical processes, automotive systems, and manufacturing processes, etc., can be modeled as such systems.

Optimal control problems for switched systems, which require the solutions of both the optimal switching sequences and the optimal continuous inputs, have attracted many researchers recently. This phenomenon is due to the problems' significance in theory and application. Many results, which report progresses regarding theoretical or practical issues for continuous-time or discrete-time versions of such problems, have appeared in the literature (see, e.g., [4], [5], [7]–[13], [15], [16], and [18]–[32]). However, there are many issues not yet addressed. For example, even for problems with linear subsystems and quadratic costs, how to obtain a closed form solution of the optimal switching instants is still a largely open problem. For more discussions on various literature results, the reader is referred to [26], [30], [31], and the references therein.

In this paper, we explore numerical solutions to such optimal control problems. Since many practical problems only involve optimizations in which a prespecified sequence of active subsystems is given (e.g., the speeding up of an automobile power train only requires switchings from gear 1–4), we concentrate on such problems. For discussions on possible solution methodologies for general optimal control problems, the reader is referred to [30, Sec. 3] and [31, Sec. 4]. Given a prespecified sequence of active subsystems, one needs to seek the solutions of both the optimal switching instants and the optimal continuous input. In [26], [30], and [31], we proposed an idea of decomposing the problem into stage (a), which is a conventional optimal control problem that finds the optimal cost given the sequence of active subsystems and the switching instants, and stage (b), which is a nonlinear optimization problem that finds the local optimal switching instants. It is worth mentioning that in [10], [11], Cassandras *et al.* proposed a similar two-stage hierarchical decomposition idea through their independent studies of similar problems for hybrid systems. In [11], Cassandras *et al.* studied a problem motivated by manufacturing systems and a quadratic optimal control problem with linear subsystems. The two problems were solved by an iterative methodology which first finds the analytical solutions to the lower level (i.e., finds the continuous input and the optimal cost) and then substitute the results into the high level and seek the optimal switching instants using nonlinear optimization methods. However it should be pointed out that it is not always possible to derive analytical solutions to the lower level optimal control problems. This is evident from the fact that only few classes of conventional optimal control problems possess closed form solutions. Even for the case of linear quadratic (LQ) problems, it is well known that the optimal costs are quadratic forms in which the coefficients can be obtained numerically by solving Riccati equations backward in time. Therefore, we do not even have a closed form solution for an LQ problem. Being presented with such difficulties, we propose a different solution approach in this paper. Our approach is motivated by the observation that, in order to solve stage (b), it is not necessary to find closed form solutions to stage (a). As long as we know the derivatives of the optimal cost with respect to the switching instants, the nonlinear optimization in stage (b) can be carried out using constrained nonlinear optimization techniques.

In general, it is hard to obtain the values of the derivatives of the stage (a) optimal cost with respect to the switching instants. To address these difficulties, in a previous paper [30], we proposed an approach which approximates such derivatives

Manuscript received August 21, 2001; revised April 18, 2002 and October 5, 2003. Recommended by Associate Editor A. Bemporad. This work was supported by the National Science Foundation under Grant NSF ECS99-12458 and Grant CCR01-13131, and by the DARPA/ITO-NEST Program under Grant AF-F30602-01-2-0526.

X. Xu is with the Department of Electrical and Computer Engineering, Penn State Erie, Erie, PA 16563 USA (e-mail: Xuping-Xu@psu.edu).

P. J. Antsaklis is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: antsaklis.1@nd.edu).

Digital Object Identifier 10.1109/TAC.2003.821417

by direct differentiations of value functions. In this paper, a method is proposed which can provide us with accurate numerical values of the derivatives instead of approximations. The method is faster and more straightforward than the approximation method when implemented. The method is based on the solution of a two point boundary value differential algebraic equation (DAE) formed by the state, costate, stationarity equations, the boundary and continuity conditions, along with their differentiations. We also apply the method to general switched linear quadratic (GSLQ) problems and show that the burden of solving a DAE can be reduced to solving an initial value ordinary differential equation (ODE). The method is much easier to implement and much faster than the approximation method for GSLQ problems. Finally, an extension of the method is applied to problems with internally forced switching (IFS). Overall, we believe the method is new and is the first one that can obtain accurate derivative values of the optimal costs.

The structure of the paper is as follows. In Section II, we formulate the optimal control problem studied in this paper. In Section III, we outline the two stage decomposition idea and discuss each stage. In Section IV, we transcribe a problem into an equivalent problem parameterized by the switching instants and develop a method to obtain the derivative value based on the solution of a two point boundary value DAE. Similar ideas are applied to GSLQ problems in Section V and a more efficient method based on the solution of an initial value ODE is developed. Section VI reports results for problems with IFS. Examples are given in Section VII to illustrate the effectiveness of the method. Section VIII concludes the paper.

II. Problem Formulation

A. Switched Systems

In this paper, we consider *switched systems* consisting of the subsystems

$$\dot{x} = f_i(x,u) \quad f_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$$
$$i \in I \triangleq \{1, 2, \ldots, M\}. \tag{1}$$

In order to control a switched system, one needs to choose not only a continuous input but also a switching sequence. A *switching sequence* in $t \in [t_0, t_f]$ regulates the sequence of active subsystems and is defined as

$$\sigma = ((t_0, i_0), (t_1, i_1), \ldots, (t_K, i_K)) \tag{2}$$

where $0 \leq K < \infty$, $t_0 \leq t_1 \leq \cdots \leq t_K \leq t_f$, and $i_k \in I$ for $k = 0, 1, \ldots, K$. Note here (t_k, i_k) indicates that at instant t_k, the system switches from subsystem i_{k-1} to subsystem i_k; during the time interval $[t_k, t_{k+1})$ ($[t_K, t_f]$ if $k = K$), subsystem i_k is active. For a switched system to be well-behaved, we only consider nonZeno sequences which switch at most a finite number of times in $[t_0, t_f]$, though different sequences may have different numbers of switchings. If we regard σ as a discrete input, then the overall control input to the system is a pair (σ, u). Finally, we note that the feature distinguishing a switched system from a general hybrid system is that its continuous state does not exhibit jumps at the switching instants. Such a feature makes the computation of continuous inputs amenable via the usage of conventional optimal control methods.

B. Optimal Control Problem

In the sequel, we define $\mathcal{U}_{[t_0,t_f]} \triangleq \{u | u \in C_p[t_0, t_f], u(t) \in \mathbb{R}^m\}$; in other words, $\mathcal{U}_{[t_0,t_f]}$ is the set of piecewise continuous functions for $t \in [t_0, t_f]$ that take values in $\mathbb{R}^m$. Since many practical problems only involve optimizations in which a prespecified sequence of active subsystems (i.e., the untimed sequence $(i_0, i_1, \ldots, i_K)$) is given, we concentrate on such problems. (Such problems appear, e.g., in the speeding up of an automobile power train which only requires switchings from gear 1–4.)

Problem 1: Consider a switched system consisting of subsystems $\dot{x} = f_i(x,u)$, $i \in I$. Given a fixed time interval $[t_0, t_f]$ and a prespecified sequence of active subsystems $(i_0, i_1, \ldots, i_K)$, find a continuous input $u \in \mathcal{U}_{[t_0,t_f]}$ and switching instants $t_1, \ldots, t_K$ such that the corresponding continuous state trajectory x departs from a given initial state $x(t_0) = x_0$ and meets an $(n - l_f)$-dimensional smooth manifold $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \to \mathbb{R}^{l_f}\}$ at t_f and the cost functional

$$J = \psi(x(t_f)) + \int_{t_0}^{t_f} L(x(t), u(t))dt \tag{3}$$

is minimized. □

Problem 1 is a basic optimal control problem in Bolza form. As in the usual practice of formulating optimal control problems (see [1]), in the sequel, we assume that f, L are continuous and have continuous partial derivatives with respect to x; ϕ_f is continuously differentiable; ψ has twice continuous derivatives. Besides these assumptions, whenever necessary, we will further assume that they possess enough smoothness properties we need in our derivations.

The way we formulate Problem 1 with a fixed final time is mainly for the convenience of subsequent studies. For a problem with free final time t_f, we can introduce an additional state variable and transcribe it to a problem with fixed final time. Analytical tools such as the maximum principle and the Hamilton–Jacobi–Bellman (HJB) equation for hybrid and switched systems have been derived in the literature (see [18], [23], [25], [26], [31], and [32]). However, it is difficult to directly use these results to find optimal controls even for switched systems with linear subsystems. For details and comments on the difficulties of using them to obtain optimal solutions, see [31] and [27, Ch. 5].

III. Two Stage Decomposition

In [26], [30], and [31], we proposed an idea which decomposes Problem 1 into two stages. Stage (a) is a conventional optimal control problem which seeks the minimum value of J with respect to u under a given switching sequence $\sigma = ((t_0, i_0), (t_1, i_1), \ldots, (t_K, i_K))$. In the sequel, we denote the corresponding optimal cost as a function $J_1(\hat{t})$, where $\hat{t} \triangleq (t_1, \ldots, t_K)^T$. Stage (b) is a constrained nonlinear optimization problem

$$\begin{aligned} \min_{\hat{t}} \quad & J_1(\hat{t}) \\ \text{subject to} \quad & \hat{t} \in T \end{aligned} \tag{4}$$

where $T \triangleq \{\hat{t} = (t_1, \ldots, t_K)^T | t_0 \leq t_1 \leq \cdots \leq t_K \leq t_f\}$.

In order to solve Problem 1, one needs to resort to not only optimal control methods, but also nonlinear optimization techniques. Except for very few classes of problems (e.g., minimum energy problems in [27] and the two examples in [11]), analytical expressions of $J_1(\hat{t})$ are almost impossible to obtain. This is evident from the fact that very few classes of conventional optimal control problems possess analytical solutions. The unavailability of analytical expressions of $J_1(\hat{t})$ henceforth makes stage (b) optimization difficult to carry out. However, even without the expressions of $J_1(\hat{t})$, if we can find the values of the derivatives $\partial J_1/\partial \hat{t}$ and $\partial^2 J_1/\partial \hat{t}^2$, we can still solve stage (b) by employing some nonlinear optimization techniques. Let us elaborate more on stages (a) and (b) in the following.

Stage (a): In stage (a), we need to find an optimal continuous input u and the corresponding minimum J. Although different subsystems are active in different time intervals, stage (a) which seeks $J_1(\hat{t})$ for the corresponding $\hat{t} = (t_1, \ldots, t_K)^T$ is conventional since these intervals are fixed. The only difference between stage (a) and most of the conventional optimal control problems is that in stage (a), the system dynamics changes with respect to different time intervals. However, it is not difficult to use the calculus of variations techniques (see, e.g., [14]) to prove the following necessary conditions (in fact, it is a form of the maximum principle). For simplicity of notation, in the following theorem, we assume that subsystem k is active in the time interval $[t_{k-1}, t_k)$ for $k = 1, \ldots, K$ and subsystem $K+1$ is active in $[t_K, t_{K+1}]$ where $t_{K+1} = t_f$.

Theorem 1—Necessary Conditions for Stage (a): Consider the stage (a) problem for Problem 1. Assume that subsystem k is active in $[t_{k-1}, t_k)$ for $1 \le k \le K$ and subsystem $K+1$ in $[t_K, t_{K+1}]$ where $t_{K+1} = t_f$. Let $u \in \mathcal{U}_{[t_0, t_f]}$ be a continuous input such that the corresponding continuous state trajectory x departs from a given initial state $x(t_0) = x_0$ and meets $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \to \mathbb{R}^{l_f}\}$ at t_f. In order for u to be optimal, it is necessary that there exists a vector function $p(t) = [p_1(t), \ldots, p_n(t)]^T$, $t \in [t_0, t_f]$, such that the following conditions hold.

a) For almost any $t \in [t_0, t_f]$ the following state and costate equations hold:

$$\text{state eqn: } \frac{dx(t)}{dt} = \left(\frac{\partial H}{\partial p}(x(t), p(t), u(t))\right)^T \tag{5}$$

$$\text{costate eqn: } \frac{dp(t)}{dt} = -\left(\frac{\partial H}{\partial x}(x(t), p(t), u(t))\right)^T. \tag{6}$$

Here, $H(x, p, u) \triangleq L(x, u) + p^T f_k(x, u)$, if $t \in [t_{k-1}, t_k)$ ($k = K+1$ if $t \in [t_K, t_f]$).

b) For almost any $t \in [t_0, t_f]$, the stationarity condition holds

$$0 = \left(\frac{\partial H}{\partial u}(x(t), p(t), u(t))\right)^T. \tag{7}$$

c) At t_f, the function p satisfies

$$p(t_f) = \left(\frac{\partial \psi}{\partial x}(x(t_f))\right)^T + \left(\frac{\partial \phi_f}{\partial x}(x(t_f))\right)^T \lambda \tag{8}$$

where λ is an l_f-dimensional vector.

d) At any t_k, $k = 1, 2, \ldots, K$, we have

$$p(t_k-) = p(t_k+). \tag{9}$$

Proof: See Appendix A. □

These necessary conditions will be used in Section IV in the development of a method for finding $\partial J_1/\partial \hat{t}$ and $\partial^2 J_1/\partial \hat{t}^2$. In general, it is difficult or even impossible to find an analytical expression of $J_1(\hat{t})$ using the above conditions. The reason is that conditions (a)–(d) present a two point boundary value differential algebraic equation (DAE) which, in most cases, cannot be solved analytically. However, the above DAE can be solved efficiently using many numerical methods (e.g., shooting methods).

Stage (b): In stage (b), we need to solve the constrained nonlinear optimization problem (4) with simple constraints. Computational methods for finding local optimal solutions of such problems are abundant in the nonlinear optimization literature. For example, feasible direction methods and penalty function methods are two commonly used classes of methods. These methods use first-order derivative $\partial J_1/\partial \hat{t}$ and second-order derivative $\partial^2 J_1/\partial \hat{t}^2$. In the computation of the examples in this paper, we use the gradient projection method (using $\partial J_1/\partial \hat{t}$) and its variations (see [6, Sec. 2.3] for details). For more information on various methods for solving constrained nonlinear optimization problems, see [3] and [17].

Remark 1: In this paper, we use methods using gradient information as opposed to nongradient ones for stage (b). The reasons are as follows. First, we note that, albeit conceptually applicable, nongradient methods that are based on brute force perturbation of $\hat{t}$ usually incur heavy computations (note for each perturbation of $\hat{t}$, an optimal control problem needs to be solved, which incurs nontrivial computational effort) and converge quite slowly; while gradient information provides a better direction for searching and hence reduces computational burden and help the methods converge faster. Second, in the case of more than two switchings, stage (b) poses a problem in higher dimensional spaces, which will create a huge number of possible perturbation directions for the nongradient method; however, the effectiveness of gradient based methods will not be hindered by higher dimensionalities. □

A Conceptual Algorithm: The following conceptual algorithm provides a framework for the optimization methodologies in Sections IV–VII.

Algorithm 1—A Conceptual Algorithm for Stage 1 Optimization

1) Set the iteration index $j = 0$. Choose an initial $\hat{t}^j$.
2) By solving an optimal control problem (i.e., stage (a)), find $J_1(\hat{t}^j)$.
3) Find $(\partial J_1/\partial \hat{t})(\hat{t}^j)$ (and $(\partial^2 J_1/\partial \hat{t}^2)(\hat{t}^j)$ if second-order method is to be used).
4) Use some feasible direction method to update $\hat{t}^j$ to be $\hat{t}^{j+1} = \hat{t}^j + \alpha^j d\hat{t}^j$ (here $d\hat{t}^j$ is formed by using the gradient information of J_1, the stepsize α^j can be chosen using some stepsize rule, e.g.,

Armijo's rule [6]). Set the iteration index $j=j+1$.

5) Repeat Steps (2), (3), (4) and (5), until a prespecified termination condition is satisfied (e.g., the norm of the projection of $(\partial J_1/\partial \hat{t})(\hat{t}^j)$ on any feasible direction is smaller than a given small number ϵ). □

It should be pointed out that the key elements of the above algorithm are

a) an optimal control algorithm for step 2);
b) the derivations of $\partial J_1/\partial \hat{t}$ and $\partial^2 J_1/\partial \hat{t}^2$ for step 3);
c) a nonlinear optimization algorithm for step 4).

In these discussions, we have already addressed elements a) and c). Element b) poses an obstacle for the usage of Algorithm 1 because $\partial J_1/\partial \hat{t}$ and $\partial^2 J_1/\partial \hat{t}^2$ are not readily available. It is the task of Sections IV–VII to address b) and devise a method for deriving the values of $\partial J_1/\partial \hat{t}$ and $\partial^2 J_1/\partial \hat{t}^2$. Finally, it should be pointed out that hidden in step 4), when we are searching for α^j, optimal control algorithm for stage a) will also be used in order to obtain the value of J_1 at the trial $\hat{t}$'s.

IV. Approach Based on Parameterization of the Switching Instants

In this section, an approach to Problem 1 based on parameterization of the switching instants is presented. The first step is to transcribe an optimal control problem into an equivalent conventional optimal control problem parameterized by the switching instants. Based on the equivalent problem formulation, a method based on the solution of a two point boundary value DAE is then developed for deriving accurate values of $\partial J_1/\partial \hat{t}$ and $\partial^2 J_1/\partial \hat{t}^2$.

A. Equivalent Problem Formulation

Now, we describe the transcription of Problem 1 into an equivalent problem parameterized by the unknown switching instants. The equivalent problem has the property that the switching instants are fixed with respect to the new independent time variable.

For simplicity of notation, we concentrate on the case of two subsystems where subsystem 1 is active in the interval $t \in [t_0, t_1)$ and subsystem 2 is active in the interval $t \in [t_1, t_f]$ (t_1 is the switching instant to be determined). We also assume that $S_f = \mathbb{R}^n$ (for general S_f, we can introduce Lagrange multipliers and develop a similar method). It is straightforward to apply the methods developed in this section to problems with several subsystems and more than one switchings; we will remark on this at the end of Section V-A.I. We consider the following problem.

Problem 2: For a switched system

$$\dot{x} = f_1(x,u), \qquad t_0 \le t < t_1 \tag{10}$$

$$\dot{x} = f_2(x,u), \qquad t_1 \le t \le t_f \tag{11}$$

find a switching instant t_1 and a continuous input $u(t)$, $t \in [t_0, t_f]$ such that the cost functional

$$J = \psi(x(t_f)) + \int_{t_0}^{t_f} L(x,u)dt \tag{12}$$

is minimized. Here t_0, t_f and $x(t_0)=x_0$ are given. □

Problem 2 can be transcribed into an equivalent problem as follows. We introduce a state variable x_{n+1} corresponding to the switching instant t_1. Let x_{n+1} satisfy

$$\frac{dx_{n+1}}{dt} = 0 \tag{13}$$

$$x_{n+1}(0) = t_1. \tag{14}$$

Next, a new independent time variable τ is introduced. A piecewise linear relationship between t and τ is established as

$$t = \begin{cases} t_0 + (x_{n+1} - t_0)\tau, & 0 \le \tau \le 1 \\ x_{n+1} + (t_f - x_{n+1})(\tau - 1), & 1 \le \tau \le 2. \end{cases} \tag{15}$$

Clearly, $\tau = 0$ corresponds to $t = t_0$, $\tau = 1$ to $t = t_1$, and $\tau = 2$ to $t = t_f$. By introducing x_{n+1} and τ, Problem 2 is transcribed into the following equivalent problem.

Problem 3—(An Equivalent Problem): For a system with dynamics

$$\frac{dx(\tau)}{d\tau} = (x_{n+1} - t_0) f_1(x,u) \tag{16}$$

$$\frac{dx_{n+1}}{d\tau} = 0 \tag{17}$$

in the interval $\tau \in [0,1)$ and

$$\frac{dx(\tau)}{d\tau} = (t_f - x_{n+1}) f_2(x,u) \tag{18}$$

$$\frac{dx_{n+1}}{d\tau} = 0 \tag{19}$$

in the interval $\tau \in [1,2]$, find an x_{n+1} and a $u(\tau)$, $\tau \in [0,2]$ such that the cost functional

$$J = \psi(x(2)) + \int_0^1 (x_{n+1} - t_0) L(x,u) d\tau + \int_1^2 (t_f - x_{n+1}) L(x,u) d\tau \tag{20}$$

is minimized. Here, t_0, t_f and $x(0) = x_0$ are given. □

Remark 2: Problem 3 and Problem 2 are equivalent in the sense that an optimal solution for Problem 3 is an optimal solution for Problem 2 by a proper change of independent variable as in (15) and by regarding $x_{n+1} = t_1$, and *vice versa*. □

Remark 3: The equivalent Problem 3 provides us with some advantage, namely that it no longer has a varying switching instant and therefore is conventional. Because x_{n+1} is actually an unknown constant throughout $\tau \in [0,2]$, in the subsequent discussion, we regard x_{n+1} as an unknown parameter for optimal control problem with cost (20) and subsystems (16) and (18), i.e., we can regard Problem 3 as an optimal control problem parameterized by the switching instant x_{n+1}. It is also worth noting that by regarding x_{n+1} as a parameter, the dimensionality of Problem 3 is the same as that of Problem 2. This is because, given the value of x_{n+1}, we only need to consider (16) and (18) to solve for the state trajectory. In fact, in the case of more than one switchings, if we apply similar transcriptions, the dimensionality of the equivalent problem is still the same as the original problem. □

Remark 4: Problem 2 and 3 allows for the special cases $t_1 = t_0$ and $t_1 = t_f$. In fact, $t_1 = t_0$ (respectively, $t_1 = t_f$) corresponds to the case when only subsystem 2 (respectively, 1) is active for $t \in [t_0, t_f]$. Algorithm 1 and the method we will develop also allow for such special solutions. □

B. Method Based on Solving a Boundary Value Differential Algebraic Equation

Based on the equivalent Problem 3, we now develop a method for deriving accurate numerical value of dJ_1/dt_1. The method is based on the solution of a two point boundary value DAE formed by the state, costate, stationarity equations, the boundary and continuity conditions for Problem 3, along with their derivatives with respect to the parameter x_{n+1}. In the following, we denote $\partial L/\partial x$, $\partial L/\partial u$ as row vectors and $\partial f/\partial x$ as an $n \times n$ matrix whose (i_1, i_2)-th element is $\partial f_{i_1}/\partial x_{i_2}$. Similar notations apply to $\partial H/\partial x$, $\partial H/\partial u$, $\partial f/\partial u$, etc.

Consider the equivalent Problem 3, define

$$\tilde{f}_1(x,u,x_{n+1}) \triangleq (x_{n+1} - t_0) f_1(x,u) \tag{21}$$

$$\tilde{f}_2(x,u,x_{n+1}) \triangleq (t_f - x_{n+1}) f_2(x,u) \tag{22}$$

$$\tilde{L}_1(x,u,x_{n+1}) \triangleq (x_{n+1} - t_0) L(x,u) \tag{23}$$

$$\tilde{L}_2(x,u,x_{n+1}) \triangleq (t_f - x_{n+1}) L(x,u). \tag{24}$$

Regarding x_{n+1} as a parameter, it is not difficult to see that the optimal state trajectory $x(\tau)$ for stage (a) is actually a function parameterized by x_{n+1}. Consequently, we denote it as $x(\tau, x_{n+1})$. We define the parameterized Hamiltonian as

$$H(x,p,u,x_{n+1}) \triangleq \begin{cases} \tilde{L}_1(x,u,x_{n+1}) + p^T \tilde{f}_1(x,u,x_{n+1}), & \tau \in [0,1) \\ \tilde{L}_2(x,u,x_{n+1}) + p^T \tilde{f}_2(x,u,x_{n+1}), & \tau \in [1,2] \end{cases}. \tag{25}$$

Assume that a parameter x_{n+1} is given, then we can apply Theorem 1 to Problem 3. The necessary conditions a) and b) provide us with the following state, costate, and stationarity equations:

$$\frac{\partial x}{\partial \tau} = \left(\frac{\partial H}{\partial p}\right)^T = \tilde{f}_k(x,u,x_{n+1}) \tag{26}$$

$$\frac{\partial p}{\partial \tau} = -\left(\frac{\partial H}{\partial x}\right)^T = -\left(\frac{\partial \tilde{f}_k}{\partial x}\right)^T p - \left(\frac{\partial \tilde{L}_k}{\partial x}\right)^T \tag{27}$$

$$0 = \left(\frac{\partial H}{\partial u}\right)^T = \left(\frac{\partial \tilde{f}_k}{\partial u}\right)^T p + \left(\frac{\partial \tilde{L}_k}{\partial u}\right)^T. \tag{28}$$

In (26)–(28), the subscript $k = 1$ for $\tau \in [0,1)$ and $k = 2$ for $\tau \in [1,2]$. Note that the p and u corresponding to the optimal solution are also functions of τ and x_{n+1}, i.e., $p = p(\tau, x_{n+1})$ and $u = u(\tau, x_{n+1})$.

From the necessary condition c) of Theorem 1, we obtain the boundary conditions

$$x(0, x_{n+1}) = x_0 \tag{29}$$

$$p(2, x_{n+1}) = \left(\frac{\partial \psi}{\partial x}(x(2, x_{n+1}))\right)^T. \tag{30}$$

The necessary condition d) tells us that $p(\tau, x_{n+1})$ is continuous at $\tau = 1$ for fixed x_{n+1}, i.e.,

$$p(1-, x_{n+1}) = p(1+, x_{n+1}). \tag{31}$$

Equation (26)–(28) along with boundary conditions (29)–(30) and the continuity condition (31) form a two point boundary value DAE parameterized by x_{n+1}. For each given x_{n+1}, the DAE can be solved using numerical methods. Now, assume that we have solved the above DAE and obtained the optimal $x(\tau, x_{n+1})$, $p(\tau, x_{n+1})$ and $u(\tau, x_{n+1})$, we then have the optimal value of J which is a function of the parameter x_{n+1}

$$J_1(x_{n+1}) = \psi(x(2, x_{n+1})) + \int_0^1 \tilde{L}_1(x,u,x_{n+1}) d\tau + \int_1^2 \tilde{L}_2(x,u,x_{n+1}) d\tau. \tag{32}$$

Differentiating J_1 with respect to x_{n+1} provides us with

$$\begin{aligned} \frac{dJ_1}{dx_{n+1}} = {} & \frac{\partial \psi(x(2, x_{n+1}))}{\partial x} \frac{\partial x(2, x_{n+1})}{\partial x_{n+1}} \\ & + \int_0^1 \Bigg(L(x,u) + (x_{n+1} - t_0) \\ & \quad \times \left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial L}{\partial u} \frac{\partial u}{\partial x_{n+1}} \right) \Bigg) d\tau \\ & + \int_1^2 \Bigg(-L(x,u) + (t_f - x_{n+1}) \\ & \quad \times \left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial L}{\partial u} \frac{\partial u}{\partial x_{n+1}} \right) \Bigg) d\tau. \end{aligned} \tag{33}$$

So, we need to obtain the function $(\partial x(\tau, x_{n+1}))/\partial x_{n+1}$ and $(\partial u(\tau, x_{n+1}))/\partial x_{n+1}$ (here we assume that x_{n+1} is fixed) in order to obtain the value dJ_1/dx_{n+1}. By differentiating the above state, costate and stationarity (26)–(28) with respect to x_{n+1} (note that $(\partial/\partial x_{n+1})(\partial/\partial \tau) = (\partial/\partial \tau)(\partial/\partial x_{n+1})$), we obtain

$$\begin{aligned} \frac{\partial}{\partial \tau}\left(\frac{\partial x}{\partial x_{n+1}}\right) & = \frac{\partial}{\partial x_{n+1}}\left(\frac{\partial x}{\partial \tau}\right) \\ & = f_1 + (x_{n+1} - t_0) \\ & \quad \times \left(\frac{\partial f_1}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial f_1}{\partial u} \frac{\partial u}{\partial x_{n+1}} \right) \end{aligned} \tag{34}$$

$$\begin{aligned} \frac{\partial}{\partial \tau}\left(\frac{\partial p}{\partial x_{n+1}}\right) & = -\frac{\partial}{\partial x_{n+1}}\left(\frac{\partial p}{\partial \tau}\right) \\ & = -\left(\frac{\partial f_1}{\partial x}\right)^T p - \left(\frac{\partial L}{\partial x}\right)^T \\ & \quad - (x_{n+1} - t_0) \\ & \quad \times \Bigg(\left(\frac{\partial f_1}{\partial x}\right)^T \frac{\partial p}{\partial x_{n+1}} + \left(p^T \frac{\partial^2 f_1}{\partial x^2} \frac{\partial x}{\partial x_{n+1}}\right)^T \\ & \qquad + \left(p^T \frac{\partial^2 f_1}{\partial x \partial u} \frac{\partial u}{\partial x_{n+1}}\right)^T + \frac{\partial^2 L}{\partial x^2} \frac{\partial x}{\partial x_{n+1}} \\ & \qquad + \frac{\partial^2 L}{\partial x \partial u} \frac{\partial u}{\partial x_{n+1}} \Bigg) \end{aligned} \tag{35}$$

$$0 = \left(\frac{\partial f_1}{\partial u}\right)^T p + \left(\frac{\partial L}{\partial u}\right)^T + (x_{n+1} - t_0) \times \left(\left(\frac{\partial f_1}{\partial u}\right)^T \frac{\partial p}{\partial x_{n+1}} + \left(p^T \frac{\partial^2 f_1}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}}\right)^T + \left(p^T \frac{\partial^2 f_1}{\partial u^2}\frac{\partial u}{\partial x_{n+1}}\right)^T + \frac{\partial^2 L}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial u^2}\frac{\partial u}{\partial x_{n+1}}\right) \quad (36)$$

for $\tau \in [0, 1)$ and

$$\frac{\partial}{\partial \tau}\left(\frac{\partial x}{\partial x_{n+1}}\right) = \frac{\partial}{\partial x_{n+1}}\left(\frac{\partial x}{\partial \tau}\right) = -f_2 + (t_f - x_{n+1}) \times \left(\frac{\partial f_2}{\partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial f_2}{\partial u}\frac{\partial u}{\partial x_{n+1}}\right) \quad (37)$$

$$\frac{\partial}{\partial \tau}\left(\frac{\partial p}{\partial x_{n+1}}\right) = -\frac{\partial}{\partial x_{n+1}}\left(\frac{\partial p}{\partial \tau}\right) = \left(\frac{\partial f_2}{\partial x}\right)^T p + \left(\frac{\partial L}{\partial x}\right)^T - (t_f - x_{n+1}) \times \left(\left(\frac{\partial f_2}{\partial x}\right)^T \frac{\partial p}{\partial x_{n+1}} \left(p^T \frac{\partial^2 f_2}{\partial x^2}\frac{\partial x}{\partial x_{n+1}}\right)^T + \left(p^T \frac{\partial^2 f_2}{\partial x \partial u}\frac{\partial u}{\partial x_{n+1}}\right)^T + \frac{\partial^2 L}{\partial x^2}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial x \partial u}\frac{\partial u}{\partial x_{n+1}}\right) \quad (38)$$

$$0 = -\left(\frac{\partial f_2}{\partial u}\right)^T p - \left(\frac{\partial L}{\partial u}\right)^T + (t_f - x_{n+1}) \times \left(\left(\frac{\partial f_2}{\partial u}\right)^T \frac{\partial p}{\partial x_{n+1}} + \left(p^T \frac{\partial^2 f_2}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}}\right)^T + \left(p^T \frac{\partial^2 f_2}{\partial u^2}\frac{\partial u}{\partial x_{n+1}}\right)^T + \frac{\partial^2 L}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial u^2}\frac{\partial u}{\partial x_{n+1}}\right) \quad (39)$$

for $\tau \in [1, 2]$.

In the previous equations, $\partial^2 f_1/\partial x^2$ is an $n \times n \times n$ array whose (j_1, j_2, j_3) element is $\partial^2 f_{1,j_1}/\partial x_{j_2}\partial x_{j_3}$ and the notation $p^T(\partial^2 f_1/\partial x^2)(\partial x/\partial x_{n+1})$ denotes an $1 \times n$ row vector which has its j_2th element as $\sum_{j_1=1}^n \sum_{j_3=1}^n p_{j_1}(\partial^2 f_{1,j_1}/\partial x_{j_2}\partial x_{j_3})(\partial x_{j_3}/\partial x_{n+1})$ where f_{1,j_1} is the j_1-th element of f_1, p_{j_1} is the j_1-th element of p and x_{j_2} is the j_2-th element of x. Similarly, $p^T(\partial^2 f_1/\partial x \partial u)(\partial u/\partial x_{n+1})$ denotes an $1 \times n$ row vector which has its j_2th element as $\sum_{j_1=1}^n \sum_{j_3=1}^m p_{j_1}(\partial^2 f_{1,j_1}/\partial x_{j_2}\partial u_{j_3})(\partial u_{j_3}/\partial x_{n+1})$; $p^T(\partial^2 f_1/\partial u^2)(\partial u/\partial x_{n+1})$ denotes an $1 \times m$ row vector which has its j_2th element as $\sum_{j_1=1}^n \sum_{j_3=1}^m p_{j_1}(\partial^2 f_{1,j_1}/\partial u_{j_2}\partial u_{j_3})(\partial u_{j_3}/\partial x_{n+1})$; $p^T(\partial^2 f_1/\partial u \partial x)(\partial x/\partial x_{n+1})$ denotes an $1 \times m$ row vector which has its j_2th element as $\sum_{j_1=1}^n \sum_{j_3=1}^n p_{j_1}(\partial^2 f_{1,j_1}/\partial u_{j_2}\partial x_{j_3})(\partial x_{j_3}/\partial x_{n+1})$. The expressions of $p^T(\partial^2 f_2/\partial x^2)(\partial x/\partial x_{n+1})$, $p^T(\partial^2 f_2/\partial x \partial u)(\partial u/\partial x_{n+1})$, $p^T(\partial^2 f_2/\partial u^2)(\partial u/\partial x_{n+1})$ and $p^T(\partial^2 f_2/\partial u \partial x)(\partial x/\partial x_{n+1})$ are understood similarly.

Differentiating the boundary conditions (29) and (30) and the continuity condition (31) with respect to x_{n+1}, we obtain

$$\frac{\partial x(0, x_{n+1})}{\partial x_{n+1}} = 0 \quad (40)$$

$$\frac{\partial p(2, x_{n+1})}{\partial x_{n+1}} = \frac{\partial^2 \psi(x(2, x_{n+1}))}{\partial x^2}\frac{\partial x(2, x_{n+1})}{\partial x_{n+1}} \quad (41)$$

$$\frac{\partial p(1-, x_{n+1})}{\partial x_{n+1}} = \frac{\partial p(1+, x_{n+1})}{\partial x_{n+1}}. \quad (42)$$

It can now be observed that (26)–(28), the boundary conditions (29) and (30) and the continuity condition (31), along with their differentiations (34)–(39), (40)–(42), form a two point boundary value DAE for $x(\tau, x_{n+1})$, $p(\tau, x_{n+1})$, $u(\tau, x_{n+1})$ and $\partial x(\tau, x_{n+1})/\partial x_{n+1}$, $\partial p(\tau, x_{n+1})/\partial x_{n+1}$, $\partial u(\tau, x_{n+1})/\partial x_{n+1}$. By solving them and substituting the result into (33), we can obtain dJ_1/dx_{n+1}.

Remark 5: If all subsystems are linear in control and the cost function L is quadratic in control, then u can be solved from the stationarity equation as a function of x and p. By differentiation with respect to x_{n+1}, $\partial u/\partial x_{n+1}$ can also be expressed as a function of x, p, $\partial x/\partial x_{n+1}$ and $\partial p/\partial x_{n+1}$. If we substitute these functions for u and $\partial u/\partial x_{n+1}$ into the state, costate equations and their differentiations, the two point boundary value DAE can, hence, be reduced to a two point boundary value differential equation in x, p, $\partial x/\partial x_{n+1}$ and $\partial p/\partial x_{n+1}$, which can be solved more easily than the DAE (e.g., using shooting methods for two point boundary value differential equations, or directly using the function `bvp4c` in Matlab). □

Remark 6—Several Subsystems and More Than One Switchings: There is no difficulty in applying the previous method to problems with several subsystems and more than one switchings. Assuming that there are K switchings, we can transcribe the problem into an equivalent problem by introducing K new state variables x_{n+k}'s, $k = 1, \ldots, K$ which correspond to the switching instants t_k's and satisfy

$$\frac{dx_{n+k}}{dt} = 0 \quad (43)$$

$$x_{n+k}(0) = t_k. \quad (44)$$

The new independent time variable τ has a piecewise linear relationship with t where $\tau = 0$ corresponds to $t = t_0$, $\tau = 1$ corresponds to $t = t_1, \ldots,$ and $\tau = K + 1$ corresponds to $t = t_f$. It is then straightforward to apply the necessary conditions in Theorem 1 to the equivalent problem to come up with the state, costate, stationarity equations, the boundary, and continuity conditions. Similar to the case of a single switching, we can then obtain x, p, u, $\partial x/\partial x_{n+k}$'s, $\partial p/\partial x_{n+k}$'s and $\partial u/\partial x_{n+k}$'s by solving the two point boundary value DAE formed by the state, costate, stationarity equations, the boundary and continuity conditions, along with their derivatives with respect to x_{n+k}'s. By substituting them into the expressions of $\partial J_1/\partial x_{n+k}$'s which can be derived similarly to (33), we can then find the accurate values of $\partial J_1/\partial x_{n+k}$'s. □

Remark 7—Second Order Derivatives: If we take second-order partial derivatives on (32), we can obtain the expression for $d^2 J_1(t_1)/dt_1^2$ which depends on the values of $\partial x/\partial x_{n+1}$, $\partial u/\partial x_{n+1}$, $\partial^2 x(\tau, x_{n+1})/\partial x_{n+1}^2$ and $\partial^2 u(\tau, x_{n+1})/\partial x_{n+1}^2$.

Similarly to the above procedure, taking first and second order derivatives of the state, costate, and stationarity (26)–(28), the boundary conditions (29) and (30) and the continuity condition (31) with respect to x_{n+1} will result in a two point boundary value DAE in $x, p, u, \partial x/\partial x_{n+1}, \partial p/\partial x_{n+1}, \partial u/\partial x_{n+1}, \partial^2 x/\partial x_{n+1}^2, \partial^2 p/\partial x_{n+1}^2, \partial^2 u/\partial x_{n+1}^2$. There is no difficulty in obtaining the values of $dJ_1(t_1)/dt_1$ and $d^2 J_1(t_1)/dt_1^2$ by solving the DAE and substituting the results to the expressions of them. This procedure can similarly be applied to the case of several subsystems and more than one switchings. □

Remark 8—Comments on Computation: In general, In order to find $\partial J_1/\partial x_{n+1}$, we need to resort to numerical methods to solve a two point boundary value DAE in x, p, u, $\partial x/\partial x_{n+1}$, $\partial p/\partial x_{n+1}$ and $\partial u/\partial x_{n+1}$. Such a DAE might seem to be twice the size of the equivalent optimal control problem where only x, p and u need to be found. However, due to the structure of such a DAE, we can first solve x, p and u from a DAE formed by (26)–(31) and then solve $\partial x/\partial x_{n+1}$, $\partial p/\partial x_{n+1}$ and $\partial u/\partial x_{n+1}$ from another DAE of the same size formed by (34)–(42). Such a method of repeatedly solving small size DAEs relieves us from the burden of solving a DAE of large size. Moreover, in the case of K switchings, we do not have to actually solve a large DAE in $x, p, u, \partial x/\partial x_{n+1}, \partial p/\partial x_{n+1}, \partial u/\partial x_{n+1}, \ldots, \partial x/\partial x_{n+K}, \partial p/\partial x_{n+K}$ and $\partial u/\partial x_{n+K}$. Instead, we only need to first solve a DAE for x, p and u and then solve K DAEs of the same size in $\partial x/\partial x_{n+k}$, $\partial p/\partial x_{n+K}$, $\partial u/\partial x_{n+K}$ for $k = 1, \ldots, K$. This explains why our method can deal with multiple switchings without enlarging the size of DAEs. □

V. General Switched Linear Quadratic Problems

As remarked in Section I, even for conventional optimal control problems with linear subsystems and quadratic costs, there is no closed form solution for stage (a). For such problems, we may only conclude that the optimal costs are quadratic forms in which the coefficients can be obtained numerically by solving Riccati equations backward in time. Therefore, to derive accurate numerical value of dJ_1/dx_{n+1}, one still needs to resort to numerical methods. In this section, we apply the idea in Section IV to general switched linear quadratic (GSLQ) problems and develop a more efficient method for deriving accurate numerical values of dJ_1/dx_{n+1}. Due to the problem's special structure, the method has the advantage that it only needs to solve an initial value ODE formed by the parameterized Riccati equation and its differentiation with respect to the switching instant in order to compute the value of dJ_1/dx_{n+1}. For simplicity of notation, we consider the following GSLQ problem with two subsystems and one switching.

Problem 4—GSLQ Problem: For a switched system

$$\dot{x} = A_1 x + B_1 u,\ t_0 \le t < t_1 \tag{45}$$

$$\dot{x} = A_2 x + B_2 u,\ t_1 \le t \le t_f \tag{46}$$

find a switching instant t_1 and a continuous input $u(t)$ such that the cost functional in general quadratic form

$$J = \frac{1}{2}x(t_f)^T Q_f x(t_f) + M_f x(t_f) + W_f + \int_{t_0}^{t_f} \times \left(\frac{1}{2}x^T Q x + x^T V u + \frac{1}{2}u^T R u + M x + N u + W\right) dt \tag{47}$$

is minimized. Here, t_0, t_f and $x(t_0) = x_0$ are given; $Q_f, M_f, W_f, Q, V, R, M, N, W$ are matrices of appropriate dimensions with $Q_f \ge 0$, $Q \ge 0$ and $R > 0$. □

In view of the method in Section IV, we transcribe Problem 4 into its equivalent problem.

Problem 5—Equivalent GSLQ Problem: For a system with dynamics

$$\frac{dx(\tau)}{d\tau} = (x_{n+1} - t_0)(A_1 x + B_1 u) \tag{48}$$

$$\frac{dx_{n+1}}{d\tau} = 0 \tag{49}$$

in the interval $\tau \in [0, 1)$ and

$$\frac{dx(\tau)}{d\tau} = (t_f - x_{n+1})(A_2 x + B_2 u) \tag{50}$$

$$\frac{dx_{n+1}}{d\tau} = 0 \tag{51}$$

in the interval $\tau \in [1, 2]$, find an x_{n+1} and a $u(\tau)$ such that the cost functional

$$J = \frac{1}{2}x(2)^T Q_f x(2) + M_f x(2) + W_f + \int_0^1 (x_{n+1} - t_0) \times L(x, u) d\tau + \int_1^2 (t_f - x_{n+1}) L(x, u) d\tau \tag{52}$$

where

$$L(x, u) = \frac{1}{2}x^T Q x + x^T V u + \frac{1}{2}u^T R u + M x + N u + W \tag{53}$$

is minimized. Here, t_0, t_f and $x(0) = x_0$ are given. □

Similar to Remark 3, Problem 5 can be regarded as a GSLQ problem parameterized by the switching instant x_{n+1}. Assume that we are given a fixed x_{n+1}, we can apply the principle of optimality to Problem 5 as follows. We assume that the optimal value function is

$$V^*(x, \tau, x_{n+1}) = \frac{1}{2}x^T P(\tau, x_{n+1}) x + S(\tau, x_{n+1}) x + T(\tau, x_{n+1}) \tag{54}$$

where $P^T(\tau, x_{n+1}) = P(\tau, x_{n+1})$. The HJB equation is

$$-\frac{\partial V^*}{\partial \tau}(x, \tau, x_{n+1}) = \min_{u(\tau)} \Big\{ (x_{n+1} - t_0)\big(L(x, u) + \frac{\partial V^*}{\partial x}(x, \tau, x_{n+1}) f_1(x, u)\big) \Big\} \tag{55}$$

in the interval $\tau \in [0, 1)$ and

$$-\frac{\partial V^*}{\partial \tau}(x, \tau, x_{n+1}) = \min_{u(\tau)} \Big\{ (t_f - x_{n+1})\big(L(x, u) + \frac{\partial V^*}{\partial x}(x, \tau, x_{n+1}) f_2(x, u)\big) \Big\} \tag{56}$$

in the interval $\tau \in [1, 2]$.

Using a method similar to the method for solving conventional linear quadratic regulator problems (see, e.g., [2]), it can be obtained that the solution to the above HJB equation is

$$u(x, \tau, x_{n+1}) = -K(\tau, x_{n+1}) x(\tau, x_{n+1}) - E(\tau, x_{n+1}) \tag{57}$$

where

$$K(\tau, x_{n+1}) = R^{-1}\big(B_k^T P(\tau, x_{n+1}) + V^T\big) \tag{58}$$

$$E(\tau, x_{n+1}) = R^{-1}\big(B_k^T S^T(\tau, x_{n+1}) + N^T\big) \tag{59}$$

(here, the subscript $k = 1$ for $\tau \in [0, 1)$ and $k = 2$ for $\tau \in [1, 2]$) and $P(\tau, x_{n+1})$, $S(\tau, x_{n+1})$ and $T(\tau, x_{n+1})$ (in the following abbreviated as P, S and T) satisfy the following parameterized general Riccati equation (parameterized by x_{n+1})

$$-\frac{\partial P}{\partial \tau} = (x_{n+1} - t_0) \times \left(Q + PA_1 + A_1^T P - (PB_1 + V)R^{-1}(B_1^T P + V^T)\right) \quad (60)$$

$$-\frac{\partial S}{\partial \tau} = (x_{n+1} - t_0) \times \left(M + SA_1 - (N + SB_1)R^{-1}(B_1^T P + V^T)\right) \quad (61)$$

$$-\frac{\partial T}{\partial \tau} = (x_{n+1} - t_0) \times \left(W - \frac{1}{2}(N + SB_1)R^{-1}(B_1^T S^T + N^T)\right) \quad (62)$$

for $\tau \in [0, 1)$ and

$$-\frac{\partial P}{\partial \tau} = (t_f - x_{n+1}) \times \left(Q + PA_2 + A_2^T P - (PB_2 + V)R^{-1}(B_2^T P + V^T)\right) \quad (63)$$

$$-\frac{\partial S}{\partial \tau} = (t_f - x_{n+1}) \times \left(M + SA_2 - (N + SB_2)R^{-1}(B_2^T P + V^T)\right) \quad (64)$$

$$-\frac{\partial T}{\partial \tau} = (t_f - x_{n+1}) \times \left(W - \frac{1}{2}(N + SB_2)R^{-1}(B_2^T S^T + N^T)\right) \quad (65)$$

for $\tau \in [1, 2]$.

Along with the boundary conditions $P(2, x_{n+1}) = Q_f$, $S(2, x_{n+1}) = M_f$, and $T(2, x_{n+1}) = W_f$, we can solve (60)–(65) (for a fixed x_{n+1}) backward in τ and obtain the parameterized optimal cost at $\tau = 0$ (i.e., the optimal J_1 under fixed x_{n+1}) as

$$J_1(t_1) = J_1(x_{n+1}) = V^*(x_0, 0, x_{n+1}) = \frac{1}{2}x_0^T P(0, x_{n+1})x_0 + S(0, x_{n+1})x_0 + T(0, x_{n+1}). \quad (66)$$

From (66), we have

$$\frac{dJ_1}{dx_{n+1}}(x_{n+1}) = \frac{\partial V^*}{\partial x_{n+1}}(x_0, 0, x_{n+1}) = \frac{1}{2}x_0^T \frac{\partial P}{\partial x_{n+1}}(0, x_{n+1})x_0 + \frac{\partial S}{\partial x_{n+1}}(0, x_{n+1})x_0 + \frac{\partial T}{\partial x_{n+1}}(0, x_{n+1}). \quad (67)$$

In order to obtain the value of dJ_1/dx_{n+1} by using (67), we need to know $\partial P/\partial x_{n+1}$, $\partial S/\partial x_{n+1}$ and $\partial T/\partial x_{n+1}$ at $(0, x_{n+1})$. To obtain these values, we differentiate (60)–(65) with respect to x_{n+1} to obtain the following equations (note that $(\partial/\partial x_{n+1})(\partial/\partial \tau) = (\partial/\partial \tau)(\partial/\partial x_{n+1})$)

$$-\frac{\partial}{\partial \tau}\left(\frac{\partial P}{\partial x_{n+1}}\right) = \left(Q + PA_1 + A_1^T P - (PB_1 + V)R^{-1} \times \left(B_1^T P + V^T\right)\right) + (x_{n+1} - t_0) \times \left(\frac{\partial P}{\partial x_{n+1}}A_1 + A_1^T \frac{\partial P}{\partial x_{n+1}} - \frac{\partial P}{\partial x_{n+1}}B_1 \times R^{-1}(B_1^T P + V^T) - (PB_1 + V)R^{-1} \times B_1^T \frac{\partial P}{\partial x_{n+1}}\right) \quad (68)$$

$$-\frac{\partial}{\partial \tau}\left(\frac{\partial S}{\partial x_{n+1}}\right) = \left(M + SA_1 - (N + SB_1) \times R^{-1}\left(B_1^T P + V^T\right)\right) + (x_{n+1} - t_0) \times \left(\frac{\partial S}{\partial x_{n+1}}A_1 - \frac{\partial S}{\partial x_{n+1}} \times B_1 R^{-1}(B_1^T P + V^T) - (N + SB_1)R^{-1}B_1^T \frac{\partial P}{\partial x_{n+1}}\right) \quad (69)$$

$$-\frac{\partial}{\partial \tau}\left(\frac{\partial T}{\partial x_{n+1}}\right) = \left(W - \frac{1}{2}(N + SB_1) \times R^{-1}(B_1^T S^T + N^T)\right) + (x_{n+1} - t_0) \times \left(-\frac{1}{2}\frac{\partial S}{\partial x_{n+1}} \times B_1 R^{-1}(B_1^T S^T + N^T) - \frac{1}{2}(N + SB_1)R^{-1}B_1^T \times \left(\frac{\partial S}{\partial x_{n+1}}\right)^T\right) \quad (70)$$

in the interval $\tau \in [0, 1)$ and

$$-\frac{\partial}{\partial \tau}\left(\frac{\partial P}{\partial x_{n+1}}\right) = -\left(Q + PA_2 + A_2^T P - (PB_2 + V) \times R^{-1}(B_2^T P + V^T)\right) + (t_f - x_{n+1}) \times \left(\frac{\partial P}{\partial x_{n+1}}A_2 + A_2^T \frac{\partial P}{\partial x_{n+1}} - \frac{\partial P}{\partial x_{n+1}}B_2 \times R^{-1}(B_2^T P + V^T) - (PB_2 + V) \times R^{-1}B_2^T \frac{\partial P}{\partial x_{n+1}}\right) \quad (71)$$

$$
\begin{aligned}
-\frac{\partial}{\partial \tau}\left(\frac{\partial S}{\partial x_{n+1}}\right) =& -\Big(M + SA_2 - (N+SB_2) \\
&\quad \times R^{-1}(B_2^T P + V^T)\Big) \\
&+ (t_f - x_{n+1}) \\
&\times \Big(\frac{\partial S}{\partial x_{n+1}} A_2 - \frac{\partial S}{\partial x_{n+1}} B_2 \\
&\quad \times R^{-1}(B_2^T P + V^T) \\
&\quad - (N+SB_2)R^{-1}B_2^T \frac{\partial P}{\partial x_{n+1}}\Big) \quad (72)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial}{\partial \tau}\left(\frac{\partial T}{\partial x_{n+1}}\right) =& -\Big(W - \frac{1}{2}(N+SB_2) \\
&\quad \times R^{-1}(B_2^T S^T + N^T)\Big) \\
&+ (t_f - x_{n+1}) \\
&\times \Big(-\frac{1}{2}\frac{\partial S}{\partial x_{n+1}} B_2 R^{-1}(B_2^T S^T + N^T) \\
&\quad - \frac{1}{2}(N+SB_2)R^{-1}B_2^T \\
&\quad \times \left(\frac{\partial S}{\partial x_{n+1}}\right)^T\Big) \quad (73)
\end{aligned}
$$

in the interval $\tau \in [1,2]$.

Equations (60)–(62) (when $k = 1$) and (68)–(70) for $\tau \in [0,1)$, and the equations (60)–(62) (when $k = 2$) and (71)–(73) for $\tau \in [1,2]$ together with the following boundary conditions at $\tau = 2$:

$$P(2, x_{n+1}) = Q_f \quad (74)$$

$$S(2, x_{n+1}) = M_f \quad (75)$$

$$T(2, x_{n+1}) = W_f \quad (76)$$

$$\frac{\partial P}{\partial x_{n+1}}(2, x_{n+1}) = 0 \quad (77)$$

$$\frac{\partial S}{\partial x_{n+1}}(2, x_{n+1}) = 0 \quad (78)$$

$$\frac{\partial T}{\partial x_{n+1}}(2, x_{n+1}) = 0 \quad (79)$$

form an initial value ODE for $P, S, T, \partial P/\partial x_{n+1}, \partial S/\partial x_{n+1}$ and $\partial T/\partial x_{n+1}$ which can be efficiently solved backward in τ using the Runge–Kutta method. From the solution of this ODE, values of $(\partial P/\partial x_{n+1})(0, x_{n+1})$, $(\partial S/\partial x_{n+1})(0, x_{n+1})$ and $(\partial T/\partial x_{n+1})(0, x_{n+1})$ can be obtained and substituted into (67) to obtain the value of dJ_1/dt_1. The conceptual Algorithm 1 can then be applied.

Remark 9—Several Subsystems and More Than One Switchings: It can be seen that there is no difficulty in applying the previous method to GSLQ problems with several subsystems and more than one switchings. First of all, we can transcribe the problem to an equivalent problem in $\tau \in [0, K+1]$ if there are K switchings as mentioned in Remark 6. It is then straightforward to differentiate the Riccati equation parameterized by $x_{n+1}, \ldots, x_{n+K}$ (i.e., $t_1, \ldots, t_K$) to obtain additional differential equations for $\partial p/\partial x_{n+K}$'s, $\partial S/\partial x_{n+k}$'s and $\partial T/\partial x_{n+k}$'s. Along with the boundary conditions $P(K+1, x_{n+1}, \ldots, x_{n+K}) = Q_f$, $S(K+1, x_{n+1}, \ldots, x_{n+K}) = M_f$, $T(K+1, x_{n+1}, \ldots, x_{n+K}) = W_f$, $\partial P/\partial x_{n+k}(K+1, x_{n+1}, \ldots, x_{n+K}) = 0$, $\partial S/\partial x_{n+k}(K+1, x_{n+1}, \ldots, x_{n+K}) = 0$ and $\partial T/\partial x_{n+k}(K+1, x_{n+1}, \ldots, x_{n+K}) = 0$ for all $1 \le k \le K$, we can solve the resultant initial value ODE backward in τ to find the values of P, S, T, and their derivatives with respect to x_{n+k} at $\tau = 0$. Once we have their values at $\tau = 0$, we can substitute the values into

$$
\begin{aligned}
\frac{\partial J_1}{\partial x_{n+k}} &= \frac{\partial V^*}{\partial x_{n+k}}(x_0, 0, x_{n+1}, \ldots, x_{n+K}) \\
&= \frac{1}{2} x_0^T \frac{\partial P}{\partial x_{n+k}}(0, x_{n+1}, \ldots, x_{n+k}) x_0 \\
&\quad + \frac{\partial S}{\partial x_{n+k}}(0, x_{n+1}, \ldots, x_{n+k}) x_0 \\
&\quad + \frac{\partial T}{\partial x_{n+k}}(0, x_{n+1}, \ldots, x_{n+k}) \quad (80)
\end{aligned}
$$

to obtain the accurate values of $\partial J_1/\partial t_k$'s. □

Remark 10—Second-Order Derivatives: It is not difficult to see that if we take second order partial derivatives of (66), we obtain

$$
\begin{aligned}
\frac{d^2 J_1}{dx_{n+1}^2}(t_1) &= \frac{\partial^2 V^*}{\partial x_{n+1}^2}(x_0, x_{n+1}, 0) \\
&= \frac{1}{2} x_0^T \frac{\partial^2 P}{\partial x_{n+1}^2}(0, x_{n+1}) x_0 \\
&\quad + \frac{\partial^2 S}{\partial x_{n+1}^2}(0, x_{n+1}) x_0 \\
&\quad + \frac{\partial^2 T}{\partial x_{n+1}^2}(0, x_{n+1}). \quad (81)
\end{aligned}
$$

While following similar ideas of differentiation of the parameterized Riccati equation, we can take first and second-order differentiations of (60)–(62) and (63)–(65) with respect to x_{n+1} and form a set of ordinary differential equations. Along with the initial conditions (74)–(79) and 0's at $\tau = 2$ for $\partial^2 P/\partial x_{n+1}^2$, $\partial^2 S/\partial x_{n+1}^2$ and $\partial^2 T/\partial x_{n+1}^2$, the resultant initial value ODE for $P, S, T, \partial P/\partial x_{n+1}, \partial S/\partial x_{n+1}, \partial T/\partial x_{n+1}, \partial^2 P/\partial x_{n+1}^2$, $\partial^2 S/\partial x_{n+1}^2$ and $\partial^2 T/\partial x_{n+1}^2$ can be readily solved and hence the accurate value of $d^2 J_1/dx_{n+1}^2$ can be obtained. □

Remark 11: Note that the method for GSLQ problems is new and it can be easily implemented using any ODE solver (e.g., `ode45` in Matlab) to address continuous-time linear quadratic problems. It is much easier to implement and much faster than the approximation method (see [30]) for GSLQ problems. Since we do not resort to the discretizations of the time and state spaces, accurate values of the switching instants can be obtained. Moreover, because we focus on the continuous-time case, our method is quite different than methods for discrete-time problems which usually resort to backward searching (e.g., via pruning of search trees as in [15]) or multiparametric programming (e.g.,[4]). □

VI. Problems With Internally Forced Switching

For all the switched systems we study in Sections II–V, we have direct control over the switchings (i.e., the switchings are generated externally). We call such systems switched systems with externally forced switching (EFS). It is worth noting that there is another important class of switched systems in which the switchings are generated implicitly when the state trajectory intersects some switching sets. Such systems are said to be switched systems with internally forced switching (IFS). In this section, we extend the result in Section IV to optimal control problems for such systems.

A. Optimal Control Problems for Switched Systems With IFS

The specifications of a *switched system with IFS* include not only the subsystems

$$\dot{x} = f_i(x,u) \quad f_i : X_i \times \mathbb{R}^m \to \mathbb{R}^n$$
$$i \in I \triangleq \{1,2,\ldots,M\} \tag{82}$$

but also the switching sets

$$\Gamma_{(i_1,i_2)} \subseteq X_{i_1} \cap X_{i_2}, \qquad i_1, i_2 \in I \quad i_1 \neq i_2 \tag{83}$$

where $X_i \in \mathbb{R}^n$. In this paper, we consider $\Gamma_{(i_1,i_2)} = \{x | \gamma_{(i_1,i_2)}(x) = 0, \gamma_{(i_1,i_2)} : \mathbb{R}^n \to \mathbb{R}^{l_{(i_1,i_2)}}\}$. For such systems, if the state trajectory intersects $\Gamma_{(i_1,i_2)}$ at subsystem i_1, the system will switch from subsystem i_1 to i_2. The only control input for such systems is the continuous input. Although one can only directly control the system by the continuous input $u(t), t \in [t_0, t_f]$, a switching sequence $\big((t_0,i_0),(t_1,i_1),\ldots,(t_K,i_K)\big)$ will be generated implicitly along with the evolution of the system state trajectory.

In this section, we focus on optimal control problems for switched systems with IFS in which a prespecified sequence of active subsystems is given. Many practical problems with IFS are in fact such problems. For example, the speeding-up of an automatic transmission automobile only requires switchings from gear 1–4 (although the switchings cannot be externally forced by the driver). Formally, the problem is stated as follows.

Problem 6 (Problem With IFS): Consider a switched system with IFS. Given a fixed time interval $[t_0, t_f]$ and a prespecified sequence of active subsystems $(i_0, i_1, \ldots, i_K)$, find a continuous input $u \in \mathcal{U}_{[t_0,t_f]}$ such that the corresponding continuous state trajectory x departs from a given initial state $x(t_0) = x_0$ at initial active subsystem i_0 and meets an $(n-l_f)$-dimensional smooth manifold $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \to \mathbb{R}^{l_f}\}$ at t_f, and the prespecified switching sequence is generated, and the cost functional

$$J = \psi(x(t_f)) + \int_{t_0}^{t_f} L(x(t), u(t))\, dt \tag{84}$$

is minimized. □

B. Method for Problems With IFS

Problems with IFS are more difficult due to the additional constraint that the state x must be in the set $\Gamma_{(i_1,i_2)}$ when the system switches from subsystem i_1 to i_2. Moreover, the switching instants can depend on the continuous input in a complicated way (in contrast, the switching instants and the continuous input are independent and can be generated separately for problems with EFS). To address these difficulties, we propose the following approach which leads to a method based on an extension of the method for problems with EFS.

Approach 1—An Approach for Problems With IFS:

1) Denote in a redundant fashion that an optimal solution to an IFS problem contains both an optimal switching sequence (starting at subsystem i_0) and an optimal continuous input, i.e., regard an IFS problem as an EFS problem with additional state constraints at the switching instants. Solve the corresponding EFS problem.
2) Verify the validity of the solution for the IFS problem (i.e., if the system under the continuous input can evolve validly and generate the corresponding switching sequence). □

The decomposition of the problem into two stages and the conceptual Algorithm 1 are still applicable to step 1) in the previous approach. Step 1) can then be solved using an extension of the method in Section IV. Such an extension must address the additional requirement that the system's state be restricted to a switching hypersurface at each switching instant.

Now, we look into stage (a) for step (1). It is in essence a conventional optimal control problem which seeks the minimum value J with respect to u under a given switching sequence $\sigma = \big((t_0,i_0),(t_1,i_1),\ldots,(t_K,i_K)\big)$. The following theorem provides necessary conditions similar to those in Theorem 1.

Theorem 2: Consider the stage (a) for Problem 6. Assume that subsystem k is active in $[t_{k-1}, t_k)$ for $1 \leq k \leq K$ and subsystem $K+1$ in $[t_K, t_{K+1}]$ where $t_{K+1} = t_f$. Assume that $x \in \Gamma_k = \{x | \gamma_k(x) = 0, \gamma_k : \mathbb{R}^n \to \mathbb{R}^{l_k}\}$ at the switching instant t_k. Let $u \in \mathcal{U}_{[t_0,t_f]}$ be a continuous input such that the corresponding continuous state trajectory x departs from a given initial state $x(t_0) = x_0$ and meets $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \to \mathbb{R}^{l_f}\}$ at t_f. Also assume that $x(t) \in \operatorname{Int}(X_k)$ for $t \in (t_{k-1}, t_k), 1 \leq k \leq K$ and $x(t) \in \operatorname{Int}(X_{K+1})$ for $t \in (t_K, t_f)$. In order that u be optimal, it is necessary that there exists a vector function $p(t) = [p_1(t), \ldots, p_n(t)]^T, t \in [t_0, t_f]$, such that the conditions a)–c) as in Theorem 1 hold, and moreover, the following condition holds d). At any $t_k, k = 1, 2, \ldots, K$, we have

$$p(t_k+) - p(t_k-) + \left(\frac{\partial \gamma_k}{\partial x}(x(t_k))\right)^T \nu_k = 0 \tag{85}$$

where ν_k is an l_k-dimensional vector.

Proof: See Appendix B. □

Equipped with Theorem 2, we can now develop a method for solving Problem 6 with IFS similarly to the method in Section IV. For simplicity of notation, we only consider the case of two subsystems. Assume that subsystem 1 is active for $t \in [t_0, t_1)$ and $x(t) \in \operatorname{Int}(X_1)$ for $t \in (t_0, t_1)$, and subsystem 2 is active for $t \in [t_1, t_f]$ and $x(t) \in \operatorname{Int}(X_2)$ for $t \in (t_1, t_f)$. Also assume that $S_f = \mathbb{R}^n$ (for general S_f, we can introduce Lagrange multipliers and develop a similar method). Using similar transcription as in Section IV-A, we can obtain the corresponding equivalent problem. Now similar to the procedure in Section IV-B, we can apply Theorem 2 to the equivalent problem and obtain the state, costate and stationarity equations and the boundary conditions which are the same as (26)–(30). However, instead of (31), the necessary condition d) in Theorem 2 leads to

$$p(1+, x_{n+1}) = p(1-, x_{n+1}) - \left(\frac{\partial \gamma_1}{\partial x}(x(1, x_{n+1}))\right)^T \nu_1(x_{n+1}) \tag{86}$$

where the multiplier ν_1 is also an unknown function of x_{n+1}. Moreover, besides the boundary condition and condition d), we require that

$$\gamma_1\big(x(1, x_{n+1})\big) = 0. \tag{87}$$

If we differentiate the cost functional J and the aforementioned equations with respect to x_{n+1}, we can obtain equations same as (33)–(41). Equation (42) will be replaced by

$$\begin{aligned}\frac{\partial p}{\partial x_{n+1}}(1+, x_{n+1}) = &\frac{\partial p}{\partial x_{n+1}}(1-, x_{n+1}) \\ &- \left(\nu_1^T(x_{n+1})\frac{\partial^2\gamma_1}{\partial x^2}\big(x(1, x_{n+1})\big) \times \frac{\partial x}{\partial x_{n+1}}\right)^T \\ &- \left(\frac{\partial \gamma_1}{\partial x}\big(x(1, x_{n+1})\big)\right)^T \times \frac{d\nu_1}{dx_{n+1}}(x_{n+1})\end{aligned} \tag{88}$$

where $\nu_1^T(\partial^2\gamma_1/\partial x^2)(\partial x/\partial x_{n+1})$ denotes an $1 \times n$ row vector which has its j_2th element as $\sum_{j_1=1}^{l_1}\sum_{j_3=1}^{n}\nu_{1,j_1}(\partial^2\gamma_{1,j_1}/\partial x_{j_2}\partial x_{j_3})(\partial x_{j_3}/\partial x_{n+1})$. The differentiation of (87) is

$$\frac{\partial\gamma_1}{\partial x}\big(x(1, x_{n+1})\big)\frac{\partial x}{\partial x_{n+1}}(1, x_{n+1}) = 0. \tag{89}$$

The aforementioned state, costate and stationarity equations, boundary conditions and their differentiations, along with (86)–(89) form a two point boundary value DAE (with jumps for p and $\partial p/\partial x_{n+1}$ at $\tau = 1$) for $x(\tau, x_{n+1})$, $p(\tau, x_{n+1})$, $u(\tau, x_{n+1})$, $(\partial x/\partial x_{n+1})(\tau, x_{n+1})$, $(\partial p/\partial x_{n+1})(\tau, x_{n+1})$, $(\partial p/\partial x_{n+1})(\tau, x_{n+1})$, $\nu_1(x_{n+1})$ and $(d\nu_1/dx_{n+1})(x_{n+1})$. By solving them and substituting the result into (33), we can obtain dJ_1/dx_{n+1}.

Remark 12: The approach developed in this section can be extended in a straightforward manner to the case of several subsystems and more than one switchings. The value of $d^2J_1(t_1)/dt_1^2$ can also be similarly obtained. □

Remark 13: Note that in the solution process of the two point boundary value DAE with jumps, we have not enforced the requirements that $x(\tau, x_{n+1}) \in \mathrm{Int}(X_1)$ and $x(\tau, x_{n+1}) \cap \Gamma_1 = \emptyset$ for $\tau \in (0, 1)$, $x(\tau, x_{n+1}) \in \mathrm{Int}(X_2)$ for $\tau \in (1, 2)$. However, after a solution has been found, we need to verify these conditions for the result. This is step (2) of Approach 1 which verifies the validity of the solution. Note that there is no guarantee that the verification will always be successful. How to modify the method so that the requirements in the verification process can be enforced poses a future research topic. □

VII. Some Examples

In this section, we illustrate the effectiveness of the methods developed in Sections IV–VI using several examples.

Example 1: Consider a switched system consisting of nonlinear subsystems

$$\text{subsystem 1:} \begin{cases}\dot{x}_1 = x_1 + u\sin x_1\\ \dot{x}_2 = -x_2 - u\cos x_2\end{cases}$$

$$\text{subsystem 2:} \begin{cases}\dot{x}_1 = x_2 + u\sin x_2\\ \dot{x}_2 = -x_1 - u\cos x_1\end{cases}$$

$$\text{subsystem 3:} \begin{cases}\dot{x}_1 = -x_1 - u\sin x_1\\ \dot{x}_2 = x_2 + u\cos x_2\end{cases}.$$

Assume that $t_0 = 0$, $t_f = 3$ and the system switches at $t = t_1$ from subsystem 1 to 2 and at $t = t_2$ from subsystem 2 to 3 ($0 \le t_1 \le t_2 \le 3$). We want to find optimal switching instants t_1, t_2 and an optimal input u such that the cost functional $J = (1/2)(x_1(3)-1)^2 + (1/2)(x_2(3)+1)^2 + (1/2)\int_0^3((x_1-1)^2 + (x_2+1)^2 + u^2(t))dt$ is minimized. Here $x_1(0) = 2$ and $x_2(0) = 3$.

For this problem, we use the method in Section IV to obtain the values of $\partial J_1/\partial t_1$ and $\partial J_1/\partial t_2$. Since the problem is linear in control, we need only to solve a two point boundary value differential equation. The resultant differential equation for x_1, x_2, p_1, p_2, $\partial x_1/\partial x_{n+1}$, $\partial x_2/\partial x_{n+1}$, $\partial p_1/\partial x_{n+1}$, $\partial p_2/\partial x_{n+1}$, $\partial x_1/\partial x_{n+2}$, $\partial x_2/\partial x_{n+2}$, $\partial p_1/\partial x_{n+2}$ and $\partial p_2/\partial x_{n+2}$ is formed by the following differential equations and their derivatives with respect to x_{n+1} and x_{n+2}.

For $\tau \in [0, 1)$, the differential equations are

$$\begin{cases}\frac{\partial x_1}{\partial\tau} = x_{n+1}(x_1 + u\sin x_1)\\ \frac{\partial x_2}{\partial\tau} = x_{n+1}(-x_2 - u\cos x_2)\\ \frac{\partial p_1}{\partial\tau} = -x_{n+1}((x_1-1) + (p_1 + up_1\cos x_1))\\ \frac{\partial p_2}{\partial\tau} = -x_{n+1}((x_2+1) + (-p_2 + up_2\sin x_2))\end{cases}$$

where $u = -p_1\sin x_1 + p_2\cos x_2$.

For $\tau \in [1, 2)$, the differential equations are shown in the equation at the bottom of the page, where $u = -p_1\sin x_2 + p_2\cos x_1$.

For $\tau \in [2, 3]$, the differential equations are

$$\begin{cases}\frac{\partial x_1}{\partial\tau} = (3 - x_{n+2})(-x_1 - u\sin x_1)\\ \frac{\partial x_2}{\partial\tau} = (3 - x_{n+2})(x_2 + u\cos x_2)\\ \frac{\partial p_1}{\partial\tau} = -(3 - x_{n+2})\left((x_1-1) + (-p_1 - up_1\cos x_1)\right)\\ \frac{\partial p_2}{\partial\tau} = -(3 - x_{n+2})\left((x_2+1) + (p_2 - up_2\sin x_2)\right)\end{cases}$$

where $u = p_1\sin x_1 - p_2\cos x_2$.

$$\begin{cases}\frac{\partial x_1}{\partial\tau} = (x_{n+2} - x_{n+1})(x_2 + u\sin x_2)\\ \frac{\partial x_2}{\partial\tau} = (x_{n+2} - x_{n+1})(-x_1 - u\cos x_1)\\ \frac{\partial p_1}{\partial\tau} = -(x_{n+2} - x_{n+1})\left((x_1-1) + (-p_2 + up_2\sin x_1)\right)\\ \frac{\partial p_2}{\partial\tau} = -(x_{n+2} - x_{n+1})\left((x_2+1) + (p_1 + up_1\cos x_2)\right)\end{cases}$$

The boundary conditions are

$$
\begin{aligned}
&x_1(0) = 2 \quad x_2(0) = 3\\
&p_1(3) = x_1(3) - 1 \quad p_2(3) = x_2(3) + 1\\
&\frac{\partial x_1}{\partial x_{n+1}}(0) = 0 \quad \frac{\partial x_2}{\partial x_{n+1}}(0) = 0\\
&\frac{\partial p_1}{\partial x_{n+1}}(3) = \frac{\partial x_1}{\partial x_{n+1}}(3) \quad \frac{\partial p_2}{\partial x_{n+1}}(3) = \frac{\partial x_2}{\partial x_{n+1}}(3)\\
&\frac{\partial x_1}{\partial x_{n+2}}(0) = 0 \quad \frac{\partial x_2}{\partial x_{n+2}}(0) = 0\\
&\frac{\partial p_1}{\partial x_{n+2}}(3) = \frac{\partial x_1}{\partial x_{n+2}}(3) \quad \frac{\partial p_2}{\partial x_{n+2}}(3) = \frac{\partial x_2}{\partial x_{n+2}}(3)
\end{aligned}
$$

In the expressions of the previous boundary conditions, in order to keep the notation simple, we omit the arguments x_{n+1}, x_{n+2} for the functions. The values of $\partial J_1/\partial t_1$ and $\partial J_1/\partial t_2$ can be obtained from the evaluations of $\partial J_1/\partial x_{n+1}$ and $\partial J_1/\partial x_{n+2}$ similar to (33).

Choose initial nominal values $t_1 = 1$ and $t_2 = 2$. By applying Algorithm 1 with the gradient projection method, after 15 iterations we find that the optimal switching instants are $t_1 = 0.2262$, $t_2 = 1.0176$ and the corresponding optimal cost is 5.4399. The computation takes 258.31 seconds of CPU time when it is performed using Matlab 6.1 on an AMD Athlon 4 900–MHz PC with 256 MB of RAM, as opposed to the less accurate results and much longer CPU time (more than 1 hour) when the approach in [30] is applied. The corresponding continuous control and state trajectory are shown in Fig. 1(a) and (b). Fig. 2 shows the optimal cost J_1 for different (t_1, t_2)'s. It can be observed that J_1 is nonconvex, and therefore in general our algorithm ends up with a local minimum (however here we have actually obtained the global minimum for this problem). □

Example 2: Consider a switched system consisting of

$$
\begin{aligned}
\text{subsystem } 1: \dot{x} &= \begin{bmatrix} 0.6 & 1.2 \\ -0.8 & 3.4 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u\\
\text{subsystem } 2: \dot{x} &= \begin{bmatrix} 4 & 3 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 2 \\ -1 \end{bmatrix} u.
\end{aligned}
$$

Assume that $t_0 = 0$, $t_f = 2$ and the system switches once at $t = t_1$ $(0 \le t_1 \le 2)$ from subsystem 1 to 2. We want to find an optimal switching instant t_1 and an optimal input u such that the cost functional $J = (1/2)\left(x_1(2) - 4\right)^2 + (1/2)\left(x_2(2) - 2\right)^2 + (1/2)\int_0^2 \left(x_2(t) - 2\right)^2 + u^2(t)dt$ is minimized. Here, $x(0) = [0,\ 2]^T$.

We use the method in Section V to obtain the value of dJ_1/dt_1. From an initial nominal $t_1 = 1.0$, by using Algorithm 1 with the gradient projection method, after 17 iterations we find that the optimal switching instant is $t_1 = 0.1897$ and the corresponding optimal cost is 9.7667. The computation takes 30.75 seconds of CPU time when it is performed using Matlab 6.1 on an AMD Athlon 4900–MHz PC with 256 MB of RAM, as opposed to 323.18 s of CPU time when the approach in [30] is applied to achieve the same accuracy of the result. The corresponding continuous control and state trajectory are shown in Fig. 3(a) and (b). Fig. 4 shows the optimal cost for different t_1's. □

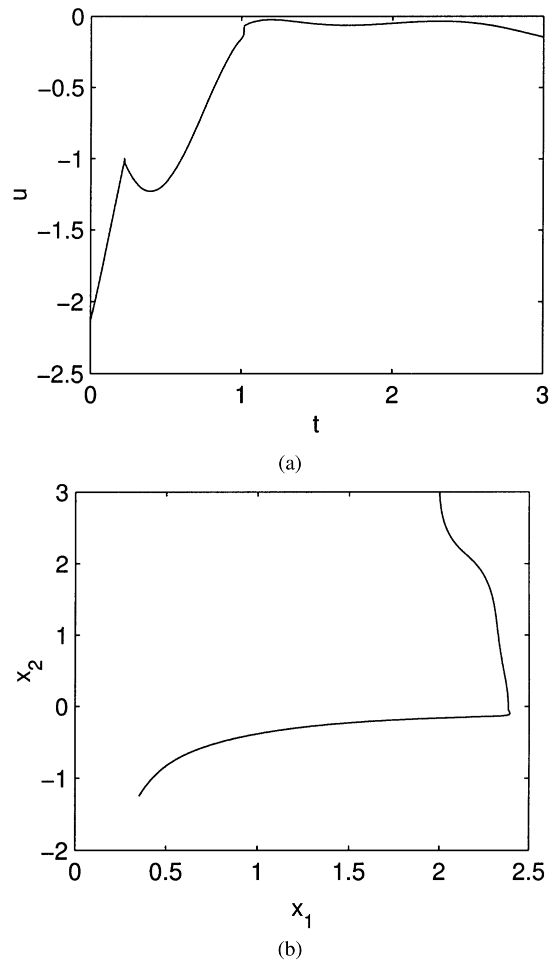

Fig. 1. Example 1. (a) Control input. (b) State trajectory.

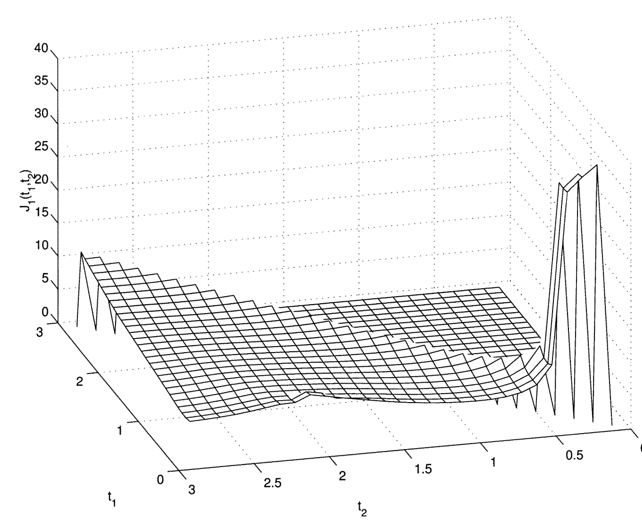

Fig. 2. Optimal cost for Example 1 for different (t_1, t_2)'s.

Example 3: Consider a switched system with IFS only consisting of

$$
\text{subsystem } 1: \dot{x} = \begin{bmatrix} 1.5 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \tag{90}
$$

$$
\text{subsystem } 2: \dot{x} = \begin{bmatrix} 0.5 & 0.866 \\ 0.866 & -0.5 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u. \tag{91}
$$

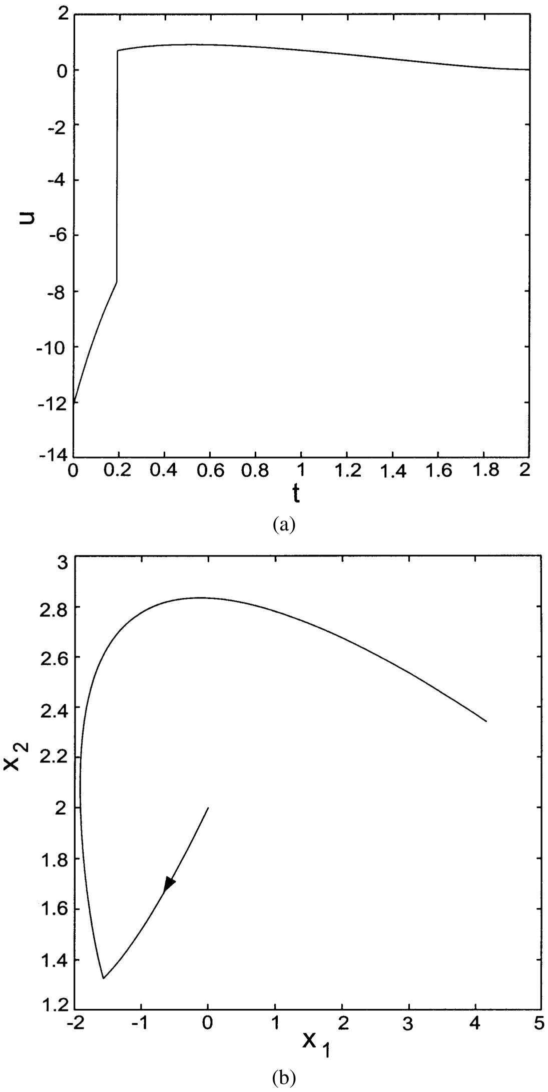

Fig. 3. Example 2. (a) Control input. (b) State trajectory.

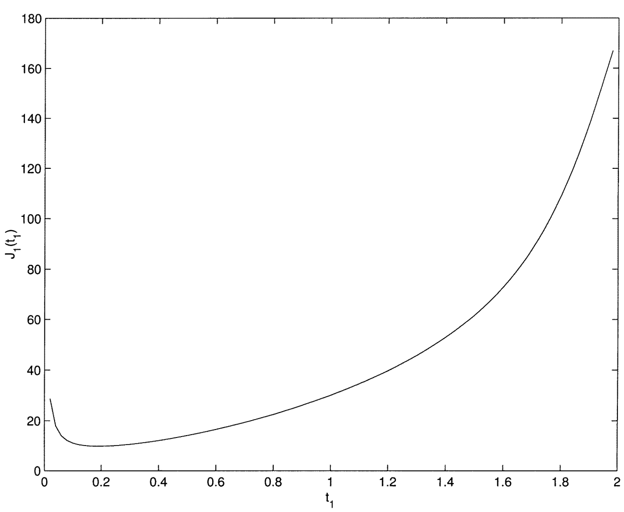

Fig. 4. Optimal cost for Example 2 for different t_1's.

Assume that $t_0 = 0$, $t_f = 2$ and the system state starts at $x(0) = [1,\ 1]^T$ following subsystem 1 (subsystem 1 is active for $\gamma_1(x_1, x_2) = x_1 + x_2 - 7 \leq 0$ and subsystem 2 is active for $\gamma_1(x_1, x_2) \geq 0$). Assume that upon intersecting the hypersurface $\gamma_1(x_1, x_2) = 0$, the system switches from subsystem 1 to 2. Also, assume there is only one switching which takes place at time t_1 $(0 \leq t_1 \leq 2)$. We want to find an optimal input u such that the cost functional $J = (1/2)\big(x_1(2) - 10\big)^2 + (1/2)\big(x_2(2) - 6\big)^2 + (1/2)\int_0^2 u^2(t)dt$ is minimized.

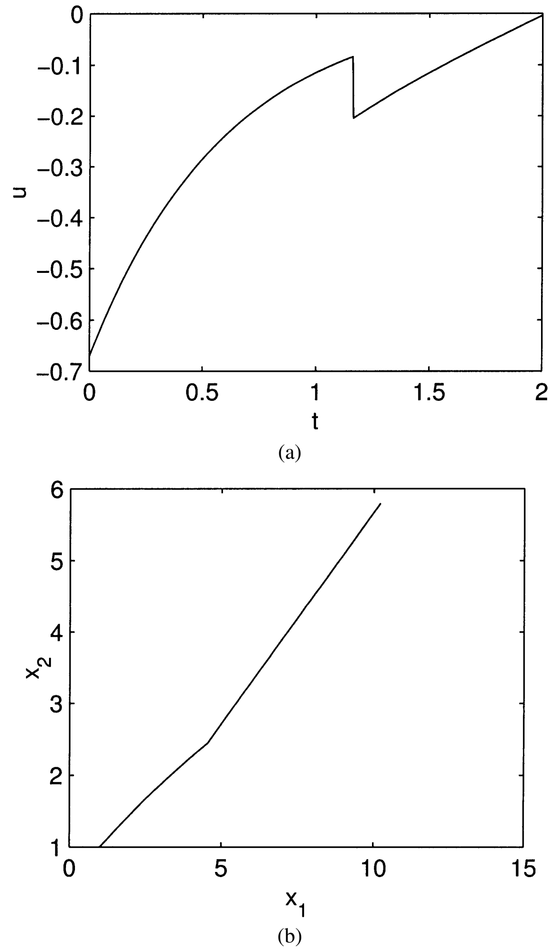

Fig. 5. Example 3. (a) Control input. (b) State trajectory.

We apply the approach developed in Section IV to this problem. We choose an initial nominal $t_1 = 1.5$. After 12 iterations we find that the optimal switching instant is $t_1 = 1.1624$ and the corresponding optimal cost is 0.1130. The corresponding $\nu_1 = -0.0605$. The computation takes about 34 minutes of CPU time when it is performed using Matlab 6.1 on an AMD Athlon 4900–MHz PC with 256 MB of RAM (it takes more time than the previous two examples due to the unavailability of an efficient Matlab subroutine for DAE with jumps; we write our own solver which is not efficient enough). The corresponding continuous control and state trajectory are shown in Fig. 5(a) and (b). The results are verified to be valid. Fig. 6 shows the optimal cost for different t_1's. □

VIII. Conclusion

In this paper, we studied optimal control problems for switched systems in which a prespecified sequence of active subsystems is given. Based on the two stage optimization idea, we proposed a method to obtain the accurate values of the derivatives that is necessary for stage (b). The method first transcribes an optimal control problem into an equivalent problem parameterized by

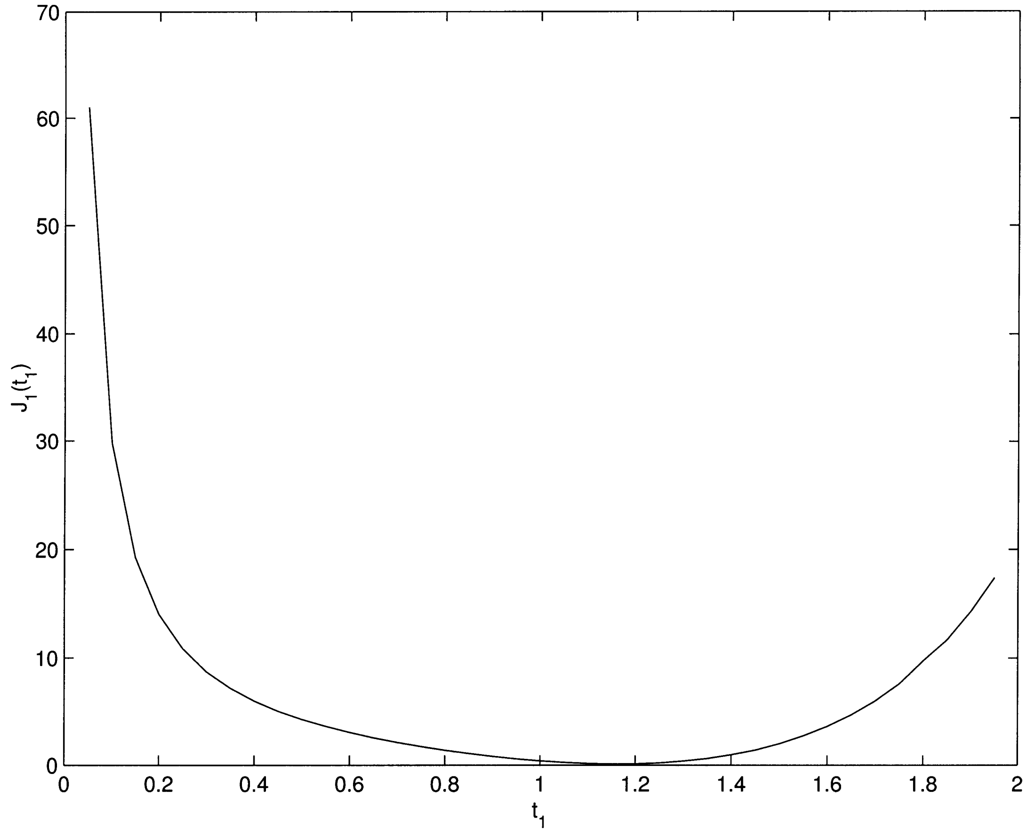

Fig. 6. Optimal cost for Example 3 for different t_1's.

the switching instants and then derives the derivatives based on the solution of a two point boundary value DAE formed by the state, costate, stationarity equations, the boundary and continuity conditions, along with their differentiations. The method was also applied to GSLQ problems and a method based on the solution of an initial value ODE was developed. An extension of the method was applied to problems with IFS. Note that earlier results of Sections IV and V have appeared in [28] and [29]. Another earlier result by the authors, which obtains approximations of the derivatives, is reported in [30]. However, note that the approach in this paper is more accurate and straightforward than that in [30]. We believe that the method described here has advantages over existing methods in that it combines good numerical characteristics and it is based on concrete theoretical results. It is particularly effective in the case of general switched linear quadratic problems and it may be used to address practical problems.

APPENDIX A

Proof of Theorem 1

Proof: We use Lagrange multipliers to adjoin the constraints $\dot{x} = f_k(x,u), k = 1,\ldots,K+1$ and $\phi_f(x(t_f)) = 0$ to J. The augmented performance index is thus

$$J' = \psi\big(x(t_f)\big) + \lambda^T \phi_f\big(x(t_f)\big) + \sum_{k=1}^{K+1} \int_{t_{k-1}}^{t_k} \Big(L(x,u) + p^T(t)\big(f_k(x,u) - \dot{x}\big)\Big) dt. \quad (92)$$

By defining $H(x,p,u) \triangleq L(x,u) + p^T f_k(x,u)$, for $t \in [t_{k-1}, t_k)$, $1 \le k \le K$ and $t \in [t_K, t_{K+1}]$ with $t_{K+1} = t_f$ if $k = K+1$, we have

$$J' = \psi\big(x(t_f)\big) + \lambda^T \phi_f\big(x(t_f)\big) + \sum_{k=1}^{K+1} \int_{t_{k-1}}^{t_k} \big(H(x,p,u) - p^T \dot{x}\big)\, dt. \quad (93)$$

From the calculus of variations, we can obtain the first variation of J' as

$$\begin{aligned}\delta J' = &\left(\frac{\partial \psi}{\partial x}\big(x(t_f)\big) + \lambda^T \frac{\partial \phi_f}{\partial x}\big(x(t_f)\big) - p^T(t_f)\right)\delta x(t_f) \\ &+ \sum_{k=1}^{K}\Big(p^T(t_k+) - p^T(t_k-)\Big)\delta x(t_k) \\ &+ \sum_{k=1}^{K+1}\int_{t_{k-1}}^{t_k}\left(\left(\frac{\partial H}{\partial x} + \dot{p}^T\right)\delta x + \frac{\partial H}{\partial u}\delta u \right. \\ &\left. + \left(\frac{\partial H}{\partial p} - \dot{x}^T\right)\delta p\right) dt. \quad (94)\end{aligned}$$

According to the Lagrange theory, a necessary condition for a solution to be optimal is $\delta J' = 0$. Setting to zero the coefficients of the independent increments $\delta x(t_f)$, $\delta x(t_k)$'s, δx, δu and δp yields the necessary conditions a)–d). □

APPENDIX B

Proof of Theorem 2

Proof: The proof is similar to that of Theorem 1, except that here in J' we introduce a term $\nu_k^T \gamma_k(x(t_k))$ and in the expansion of $\delta J'$, we have the coefficients of $\delta x(t_k)$ as $\left(p(t_k+) - p(t_k-) + \left((\partial\gamma_k/\partial x)(x(t_k))\right)^T \nu_k\right)^T$. Setting to zero the coefficients of the independent increments $\delta x(t_f)$, $\delta x(t_k)$'s, δx, δu and δp therefore yields the necessary conditions a)–d). □

REFERENCES

[1] M. Athans and P. Falb, *Optimal Control*. New York: McGraw-Hill, 1966.

[2] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. Philadelphia, PA: SIAM, 1999.

[3] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming Theory and Algorithms*, 2nd ed. New York: Wiley, 1993.

[4] A. Bemporad, F. Borrelli, and M. Morari, "Optimal controllers for hybrid systems: stability and piecewise linear explicit form," in *Proc. 39th IEEE Conf. Decision Control*, 2000, pp. 1810–1815.

[5] ——, "On the optimal control law for linear discrete time hybrid systems," in *Proc. Hybrid Systems: Computation and Control (HSCC 2002)*, vol. 2289, 2002, pp. 105–119.

[6] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Singapore: Athena Scientific, 1999.

[7] M. S. Branicky, V. S. Borkar, and S. K. Mitter, "A unified framework for hybrid control: model and optimal control theory," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 31–45, Jan. 1998.

[8] A. Giua, C. Seatzu, and C. V. D. Mee, "Optimal control of autonomous linear systems switched with a pre-assigned finite sequence," in *Proc. 2001 IEEE Int. Symp. Intelligent Control*, 2001, pp. 144–149.

[9] ——, "Optimal control of switched autonomous linear systems," in *Proc. 40th IEEE Conf. Decision Control*, 2001, pp. 2472–2477.

[10] K. Gokbayrak and C. G. Cassandras, "Hybrid controllers for hierarchically decomposed systems," in *Proc. Hybrid Systems: Computation Control (HSCC 2000)*, vol. 1790, 2000, pp. 117–129.

[11] ——, "A hierarchical decomposition method for optimal control of hybrid systems," in *Proc. 39th IEEE Conf. Decision Control*, 2000, pp. 2472–2477.

[12] S. Hedlund and A. Rantzer, "Optimal control of hybrid system," in *Proc. 38th IEEE Conf. Decision Control*, 1999, pp. 3972–3977.

[13] Johansson, "Piecewise linear control systems," Ph.D. dissertation, Lund Inst. Technol., Lund, Sweden, 1999.

[14] F. L. Lewis, *Optimal Control*. New York: Wiley, 1986.

[15] B. Lincoln and B. M. Bernhardsson, "Efficient pruning of search trees in LQR control of switched linear systems," in *Proc. 39th IEEE Conf. Decision Control*, 2000, pp. 1828–1833.

[16] J. Lu, L. Liao, A. Nerode, and J. H. Taylor, "Optimal control of systems with continuous and discrete states," in *Proc. 32nd IEEE Conf. Decision Control*, 1993, pp. 2292–2297.

[17] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*. New York: McGraw-Hill, 1996.

[18] B. Piccoli, "Hybrid systems and optimal control," in *Proc. 37th IEEE Conf. Decision Control*, 1998, pp. 13–18.

[19] A. Rantzer and M. Johansson, "Piecewise linear quadratic optimal control," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 629–637, Ap. 2000.

[20] P. Riedinger, C. Zanne, and F. Kratz, "Time optimal control of hybrid systems," in *Proc. 1999 Amer. Control Conf.*, 1999, pp. 2466–2470.

[21] P. Riedinger, F. Kratz, C. Iung, and C. Zanne, "Linear quadratic optimization for hybrid systems," in *Proc. 38th IEEE Conf. Decision Control*, 1999, pp. 3059–3064.

[22] T. I. Seidman, "Optimal control for switching systems," in *Proc. 21st Annu. Conf. Information Sciences Systems*, 1987, pp. 485–489.

[23] H. J. Sussmann, "A maximum principle for hybrid optimal control problems," in *Proc. 38th IEEE Conf. Decision Control*, 1999, pp. 425–430.

[24] L. Y. Wang, A. Beydoun, J. Cook, J. Sun, and I. Kolmanovsky, "Optimal hybrid control with applications to automotive powertrain systems," in *Control Using Logic-Based Switching*: Springer-Verlag, 1997, vol. LNCIS 222, pp. 190–200.

[25] H. S. Witsenhausen, "A class of hybrid-state continuous-time dynamic systems," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 161–167, Apr. 1966.

[26] X. Xu and P. J. Antsaklis, "Optimal control of switched systems: new results and open problems," in *Proc. 2000 Amer. Control Conf.*, 2000, pp. 2683–2687.

[27] X. Xu, "Analysis and design of switched systems," Ph.D. dissertation, Univ. Notre Dame, Notre Dame, IN, 2001.

[28] X. Xu and P. J. Antsaklis, "An approach for solving general switched linear quadratic optimal control problems," in *Proc. 40th IEEE Conf. Decision Control*, 2001, pp. 2478–2483.

[29] ——, "An approach to optimal control of switched systems with internally forced switchings," in *Proc. 2002 Amer. Control Conf.*, 2002, pp. 148–153.

[30] ——, "Optimal control of switched systems via nonlinear optimization based on direct differentiations of value functions," *Int. J. Control*, vol. 75, no. 16/17, pp. 1406–1426, 2002.

[31] ——, "Results and perspectives on computational methods for optimal control of switched systems," in *Hybrid Systems: Computation and Control (HSCC 2003)*, vol. 2623, 2003, pp. 540–555.

[32] J. Yong, "Systems governed by ordinary differential equations with continuous, switching and impulse controls," *Appl. Math. Optim.*, vol. 20, pp. 223–235, 1989.

Xuping Xu (S'98–M'01) was born in Shanghai, China, in 1974. He received the B.S. degrees in electrical engineering and applied mathematics from Shanghai Jiao Tong University, Shanghai, China, the M.S. degrees in electrical engineering and applied mathematics, and the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1995, 1998, 1999, and 2001, respectively.

Since 2001, he has been an Assistant Professor in the Department of Electrical and Computer Engineering at Penn State Erie, Erie, PA. His current research interests include stability and optimal control of hybrid and switched systems, discrete event systems, and autonomous intelligent systems.

Dr. Xu serves on the Conference Editorial Board of the IEEE Control Systems Society.

Panos J. Antsaklis (S'74–M'76–SM'86–F'91) received the Undergraduate degree from the National Technical University of Athens (NTUA), Athens, Greece, and the M.S. and Ph.D. degrees from Brown University, Providence, RI.

He is the H.C. and E.A. Brosey Professor of Electrical Engineering and Director of the Center for Applied Mathematics at the University of Notre Dame, Notre Dame, IN. His work includes analysis of behavior and design of control strategies for complex autonomous, intelligent systems. His recent research focuses on networked embedded systems and addresses problems in the interdisciplinary research area of control, computing and communication networks, and on hybrid and discrete-event dynamical systems.

He has authored a number of publications in journals, conference proceedings and books, and he has edited four books on Intelligent Autonomous Control and Hybrid Systems. In addition, he coauthored (with J. Moody) the research monograph *Supervisory Control of Discrete Event Systems Using Petri Nets* (Norwell, MA: Kluwer, 1998) and (with A. N. Michel) the graduate textbook *Linear Systems* (New York: McGraw-Hill, 1997).

Dr. Antsaklis serves on the Editorial Boards of several journals, and he currently serves as Associate Editor at Large of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He has been Guest Editor of Special Issues on Hybrid Systems in the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and the PROCEEDINGS OF THE IEEE. He has served as Program Chair and General Chair of major systems and control conferences and he was the 1997 President of the IEEE Control Systems Society (CSS). He is a Distinguished Lecturer of the IEEE Control Systems Society, a recipient of the IEEE Distinguished Member Award of the Control Systems Society, and an IEEE Third Millennium Medal recipient.