

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

Xuping Xu[†] Panos J. Antsaklis[‡]

Abstract

This paper presents an approach for solving optimal control problems for switched systems. In general, in such problems one needs to find both optimal continuous inputs and optimal switching sequences, since the system dynamics vary before and after every switching instant. After formulating the optimal control problem, we propose a two stage optimization methodology. Since many practical problems only involve optimization where the number of switchings and the sequence of active subsystems are given, we focus on Stage 1 optimization problems and propose a conceptual algorithm for solving them. In order to implement the algorithm, the derivatives of the optimal cost with respect to the switching instants need to be known. We propose a method which first transcribes a Stage 1 problem into an equivalent problem parameterized by the switching instants and then the values of the derivatives are obtained 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 and their differentiations. This approach is applied to general switched linear quadratic (GSLQ) problems and an efficient method based on the solution of an initial value ordinary differential equation (ODE) is developed. Examples are shown to illustrate the results in the paper.

1 Introduction

A switched system is a particular kind of hybrid system that consists of several subsystems and a switching law specifying 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.

Recently, optimal control problems of hybrid and switched systems have been attracting researchers from various fields in science and engineering, due to the problems' significance in theory and applications. The available results in the literature for such problems can be classified into two categories, i.e., theoretical and practical. [4, 8, 16, 19, 20, 22, 29] are some primarily theoretical results. These results extended the classical maximum principle or the HJB equation to such problems. However, because there is no efficient constructive methodology suggested in these papers for

*The research in this paper is supported by the Army Research Office (DAAG 55-98-1-0199) and National Science Foundation (NSF ECS-9912458).

[†]Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA, E-mail: Xuping.Xu.15@nd.edu.

[‡]Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA. Tel: 1-219-631-5792; Fax: 1-219-631-4393; E-mail: antsaklis.1@nd.edu.

finding optimal solutions, there is a significant gap between theoretical results and their applications to real-world examples. As to the second category of practical results, researchers take advantage of the availability of high speed computers and efficient nonlinear optimization techniques to develop approaches for solving such problems (see e.g., [4, 5, 6, 9, 10, 12, 13, 18, 17, 21]). It is worth noting that because there are many different models and optimal control objectives for hybrid systems, the above papers often differ greatly in their problem formulations and approaches. Switched systems, on the other hand, tend to be described by similar models, and similar optimal control problem formulations have appeared in the literature (e.g., [9, 10, 13, 18, 21, 26]). For an optimal control problem for a switched system, one needs to find both an optimal continuous input and an optimal switching sequence since the system dynamics may vary after every switching instant. Due to the involvement of switching sequences, such a problem is in general difficult to solve. Interested readers may refer to [26] for an overview of the problem and its difficulties. Most of the methods in the literature that we are aware of are based on some discretization of continuous-time space and/or discretization of state space into grids and use search methods for the resultant discrete model to find optimal/suboptimal solutions. But the discretization approaches may lead to computational combinatoric explosion and the solutions obtained may not be accurate enough (see [23]). In view of this, in our research, we explore approaches that are not based on the discretization of the continuous time space.

This paper presents an approach for solving optimal control problems of switched systems which is not based on the discretization of the time space and emphasizes accurate optimization of switching instants. Optimal control problems for switched systems are first carefully formulated. We then propose a two stage optimization methodology. Since the two stage optimization methodology is still difficult to implement, we then concentrate on Stage 1 optimization where the number of switchings and the sequence of active subsystems are given. Focusing on Stage 1 problems is appropriate because in many practical situations, we only need to study problems with a fixed number of switchings and a fixed order of active subsystems (e.g., the speeding up of an automobile power train only requires switchings from gear 1 to 2 to 3 to 4) and in such cases the solution to Stage 1 is indeed optimal for the problem. On the other hand, Stage 1 optimization itself is already challenging enough and solving it is a first step toward solving the general problem which does not possess a good solution up to now. A Stage 1 problem can further be decomposed into Stage 1(a), which is a conventional optimal control problem that finds the optimal cost given the order of active subsystems and the switching instants, and Stage 1(b), which is a nonlinear optimization problem that finds the optimal switching instants. Stage 1(b) poses difficulties because in general it is hard to obtain the values of the derivatives of the Stage 1(a) optimal cost with respect to the switching instants. To address these difficulties, in our previous papers [28, 24], we proposed an approach which approximates such derivatives 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 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 and their differentiations. Our approach is also applied to general switched linear quadratic (GSLQ) problems and an efficient method is derived.

The structure of the paper is as follows. In Section 2, we formulate the optimal control problem studied in this paper. In Section 3, we show that such a problem can be posed as a two stage

optimization problem under some additional assumptions. From Section 4 on, we concentrate on Stage 1 optimization problems. In Section 4, we discuss Stage 1(a) and 1(b) and propose a conceptual algorithm. In Section 5, we transcribe a Stage 1 problem into an equivalent problem parameterized by the switching instants. In Section 6, we develop a method to obtain the derivative information based on the solution of a two point boundary value DAE. Moreover, similar ideas are applied to general switched linear quadratic (GSLQ) problems in Section 7 and a more efficient method based on the solution of an initial value ordinary differential equation (ODE) is developed. Examples are given in Section 8 to illustrate the effectiveness of the method. Section 9 concludes the paper.

It should be pointed out that parts of Sections 2, 3 and 4, which introduce the problem formulation and the two stage optimization approach, are similar to corresponding parts in [28, 24]. However, the major contributions of this paper, which are the methods for deriving accurate values of the derivatives of optimal cost with respect to the switching instants reported in Sections 5, 6 and 7, are very different from the methods in [28, 24]; they are more accurate and straightforward than the approximation methods in [28, 24].

2 Problem Formulation

2.1 Switched Systems

Switched Systems

A switched system is a particular kind of hybrid system that consists of several subsystems and a switching logic among them. The feature that distinguishes a switched system from a general hybrid system is that its continuous state does not exhibit jumps at the switching instants. The switched systems we shall consider in this paper are defined as follows.

Definition 2.1 (Switched System) *A switched system is a tuple $\mathcal{S} = (\mathcal{D}, \mathcal{F}, \mathcal{L})$ where*

- $\mathcal{D} = (I, E)$ is a directed graph indicating the discrete structure of the system. The node set $I = \{1, 2, \dots, M\}$ is the set of indices for subsystems. The directed edge set E is a subset of $I \times I - \{(i, i) | i \in I\}$ which contains all valid events. If an event $e = (i_1, i_2)$ takes place, the system switches from subsystem i_1 to i_2 .
- $\mathcal{F} = \{f_i : X_i \times U_i \times \mathbb{R} \rightarrow \mathbb{R}^n | X_i \subseteq \mathbb{R}^n, U_i \subseteq \mathbb{R}^m, i \in I\}$ with f_i describing the vector field for the i -th subsystem $\dot{x} = f_i(x, u, t)$.
- $\mathcal{L} = \{\Lambda_e | \Lambda_e \subseteq \mathbb{R}^n, e \in E\}$ provides us with a logic constraint that relates the continuous state and mode switchings. Note for any $e \in E$, $\Lambda_e \neq \emptyset$. Only when $x \in \Lambda_e$, $e = (i_1, i_2)$, a switching from i_1 to i_2 is possible. \square

In view of Definition 2.1, a switched system is a collection of subsystems related by a switching logic described by \mathcal{D} and \mathcal{L} . Note that one distinct feature of a switched system is that it has no discontinuities of the state x at the switching instants. If a particular switching law has been

specified (the law may be specified by state space partitions or by time involvements), then the switched system can be described as

$$\dot{x}(t) = f_{i(t)}(x(t), u(t), t) \quad (2.1)$$

$$i(t) = \varphi(x(t), i(t^-), t), \quad (2.2)$$

where $\varphi : \mathbb{R}^n \times I \times \mathbb{R} \rightarrow I$ determines the active subsystem at instant t . Note that (2.1)-(2.2) are used as the definition of switched systems in some of the literature (e.g., [9]). We adopt Definition 2.1 rather than (2.1)-(2.2) because in design problems, in general, φ is not defined a priori and it is a designer's task to find a switching law.

Switching Sequences

For a switched system \mathcal{S} , the control input of the system consists of both a continuous input $u(t)$, $t \in [t_0, t_f]$ and a switching sequence. We define a switching sequence as follows.

Definition 2.2 (Switching Sequence) For a switched system \mathcal{S} , a switching sequence σ in $[t_0, t_f]$ is defined to be

$$\sigma = ((t_0, i_0), (t_1, e_1), (t_2, e_2), \dots, (t_K, e_K)), \quad (2.3)$$

with $0 \leq K < \infty$, $t_0 \leq t_1 \leq t_2 \leq \dots \leq t_K \leq t_f$, and $i_0 \in I$, $e_k = (i_{k-1}, i_k) \in E$ for $k = 1, 2, \dots, K$. We define $\Sigma_{[t_0, t_f]} \triangleq \{\sigma \text{'s in } [t_0, t_f]\}$. \square

A switching sequence σ as defined above indicates that, if $t_k < t_{k+1}$, then subsystem i_k is active in $[t_k, t_{k+1})$ ($[t_K, t_f]$ if $k = K$); if $t_k = t_{k+1}$, then i_k is switched through at instant t_k ('switched through' means that the system switches from subsystem i_{k-1} to i_k and then to i_{k+1} all at instant t_k). For a switched system to be well-behaved, we generally exclude the undesirable *Zeno* phenomenon, i.e., infinitely many switchings in finite amount of time. Hence in Definition 2.2, we only allow 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. We specify $\sigma \in \Sigma_{[t_0, t_f]}$ as a discrete input to a switched system. The overall control input to the system is therefore a pair (σ, u) .

Example 2.1 (An Automotive Control System) A manual transmission car with four gears is a good example of a switched system. If we denote the lateral position as x_1 and the velocity as x_2 , the system dynamics at gear i can be described by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \alpha_i(x_2)u, \end{aligned}$$

where the nonlinear gear efficiency function $\alpha_i(x_2)$ is depicted in figure 1.

For this system, $I = \{1, 2, 3, 4\}$, all $X_i = \mathbb{R}^n$ and all $U_i = [0, u_{max}]$, where u_{max} is given. If for each switching, the car can only be shifted one gear up or down, we have $E = \{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3)\}$; moreover, $\Lambda_{(1,2)} = \Lambda_{(2,1)} = \{x|x_2 \in [10, 20]\}$, $\Lambda_{(2,3)} = \Lambda_{(3,2)} = \{x|x_2 \in [20, 40]\}$, $\Lambda_{(3,4)} = \Lambda_{(4,3)} = \{x|x_2 \in [40, 60]\}$. The control input of this system consists of the continuous input u (the throttle position) and the external switching sequence (gear shifting). \square

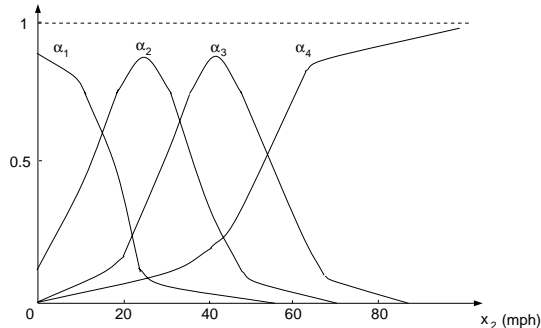


Figure 1: The nonlinear gear efficiency functions α_i .

2.2 An Optimal Control Problem

Note that in the sequel of this paper, we assume that $X_i = \mathbb{R}^n$, $U_i = \mathbb{R}^m$ and $\Lambda_e = \mathbb{R}^n$ for any $i \in I$, $e \in E$ and report results under these assumptions. We assume these because on one hand optimal control problems under these assumptions are already challenging and interesting enough and well deserve our attention; on the other hand problems under more general constraints are still under extensive researches. We also 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 all piecewise continuous functions for $t \in [t_0, t_f]$ that take values in \mathbb{R}^m .

Problem 2.1 Consider a switched system $\mathcal{S} = (\mathcal{D}, \mathcal{F}, \mathcal{L})$. Given a fixed time interval $[t_0, t_f]$, find a continuous input $u \in \mathcal{U}_{[t_0, t_f]}$ and a switching sequence $\sigma \in \Sigma_{[t_0, t_f]}$ 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 $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \rightarrow \mathbb{R}^{l_f}\}$ at t_f and the cost functional

$$J = \psi(x(t_f)) + \int_{t_0}^{t_f} L(x(t), u(t), t) dt \quad (2.4)$$

is minimized. □

Problem 2.1 is a basic optimal control problem with a fixed end-time where the final state is on a smooth manifold. 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 and t ; ϕ_f is assumed to be continuously differentiable; ψ has twice continuous derivatives. Besides these assumptions, in the following, whenever necessary, we will further assume that they possess enough smoothness properties we need in our derivations.

The way we formulate Problem 2.1 with a fixed final time is mainly for the convenience of subsequent studies in this paper. Note that for a problem with free end-time t_f , we can introduce an additional state variable and transcribe it to a problem with fixed end-time (for more details, see [23]).

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 [16, 20, 22, 26, 29]). However, it is difficult to directly use these tools 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 [23].

3 Two Stage Optimization

In general, we need to find an optimal control solution (σ^*, u^*) for Problem 2.1 such that

$$J(\sigma^*, u^*) = \min_{\sigma \in \Sigma_{[t_0, t_f]}, u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u). \quad (3.1)$$

Notice that for any fixed switching sequence σ , Problem 2.1 reduces to a conventional optimal control problem for which we only need to find an optimal continuous input u that minimizes $J_\sigma(u) = J(\sigma, u)$. This idea naturally leads us toward considering Problem 2.1 as a two stage optimization problem. Under some additional assumptions, we can prove the following lemma that provides a way to do so.

Lemma 3.1 *For Problem 2.1, if*

(a). *an optimal solution (σ^*, u^*) exists and*

(b). *for any given switching sequence σ , there exists a corresponding $u^* = u_\sigma^*$ such that $J_\sigma(u) = J(\sigma, u)$ is minimized,*

then the following equation holds

$$\min_{\sigma \in \Sigma_{[t_0, t_f]}, u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u) = \min_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u). \quad (3.2)$$

Proof: First we claim that

$$\min_{\sigma \in \Sigma_{[t_0, t_f]}, u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u) \leq \inf_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u). \quad (3.3)$$

This is because for any fixed σ , there exists a u_σ^* such that $J(\sigma, u_\sigma^*) = \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u)$. But for every pair (σ, u_σ^*) , we must have $J(\sigma^*, u^*) \leq J(\sigma, u_\sigma^*)$, therefore from (3.3) we must have

$$J(\sigma^*, u^*) \leq \inf_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u_\sigma^*) = \inf_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u). \quad (3.4)$$

While we also have the inequality

$$\inf_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u) \leq \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma^*, u) = J(\sigma^*, u_{\sigma^*}^*). \quad (3.5)$$

In (3.5) we can choose $u_{\sigma^*}^* = u^*$, since for any other u , we must have $J(\sigma^*, u^*) \leq J(\sigma^*, u)$ due to the optimality of (σ^*, u^*) . Hence combining (3.4) and (3.5) we have

$$J(\sigma^*, u^*) \leq \inf_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u) \leq J(\sigma^*, u_{\sigma^*}^*) = J(\sigma^*, u^*). \quad (3.6)$$

Hence all inequalities in (3.6) must be equalities and the $\inf_{\sigma \in \Sigma_{[t_0, t_f]}}$ can be replaced by $\min_{\sigma \in \Sigma_{[t_0, t_f]}}$ so we obtain

$$J(\sigma^*, u^*) = \min_{\sigma \in \Sigma_{[t_0, t_f]}, u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u) = \min_{\sigma \in \Sigma_{[t_0, t_f]}} \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u). \quad (3.7)$$

□

The right hand side of (3.2) needs twice the minimization process. This supports the validity of the following two stage optimization methodology.

A Two Stage Optimization Methodology

Stage 1. Fixing σ , solve the inner minimization problem.

Stage 2. Regarding the optimal cost for each σ as a function

$$J_1 = J_1(\sigma) = \min_{u \in \mathcal{U}_{[t_0, t_f]}} J(\sigma, u), \quad (3.8)$$

minimize J_1 with respect to $\sigma \in \Sigma_{[t_0, t_f]}$. □

In more detail, we can implement the above methodology by the following algorithm.

Algorithm 3.1 (A Two Stage Algorithm)

Stage 1. (a). Fix the total number of switchings to be K and the sequence of active subsystems and let the minimum value of J with respect to u be a function of the K switching instants, i.e., $J_1 = J_1(t_1, t_2, \dots, t_K)$ for $K \geq 0$ ($t_0 \leq t_1 \leq t_2 \leq \dots \leq t_K \leq t_f$). Find J_1 .

(b). Minimize J_1 with respect to t_1, t_2, \dots, t_K .

Stage 2. (a). Vary the order of active subsystems to find an optimal solution under K switchings.

(b). Vary the number of switchings K to find an optimal solution for Problem 2.1. □

Note that the Stage 1 in Algorithm 3.1 is slightly different from the Stage 1 in the two stage methodology. In fact, Stage 1(a) in Algorithm 3.1 corresponds to Stage 1 in the two stage methodology. Stage 1(b) and Stage 2 together correspond to Stage 2 in the two stage methodology. In the following, when we mention Stage 1 optimization, we actually refer to Stage 1 in Algorithm 3.1. Algorithm 3.1 needs further implementations. In practice, many problems only require the solutions of optimal continuous inputs and optimal switching instants for Stage 1 optimization where a fixed number of switchings and a fixed sequence of active subsystems are given. In general, explicit expressions of J_1 are difficult to obtain or quite complicated even for very simple problems. Therefore it is necessary to devise optimization methods that do not require the explicit expression of J_1 as a function of t_k 's. In the next section, we shall discuss Stage 1 optimization in detail.

4 More on Stage 1 Optimization

Now we concentrate on Stage 1 optimization. On one hand, Stage 1 optimization has already presented enough challenge to us. On the other hand, since many real world problems are in fact stage 1 optimization problems, Stage 1 does deserve our attention. For example, the speeding-up of a power train only requires switchings from gear 1 to 2 to 3 to 4. As can be seen from Algorithm 3.1 in Section 3, Stage 1 can further be decomposed into two sub-steps (a) and (b). Stage 1(a) is in essence a conventional optimal control problem which seeks the minimum value of J with respect to u under a given switching sequence $\sigma = ((t_0, i_0), (t_1, e_1), \dots, (t_K, e_K))$. We denote the corresponding optimal cost as a function $J_1(\hat{t})$, where $\hat{t} \triangleq (t_1, t_2, \dots, t_K)^T$. Stage 1(b) is in essence a constrained nonlinear optimization problem

$$\begin{aligned} & \min_{\hat{t}} J_1(\hat{t}) \\ & \text{subject to } \hat{t} \in T \end{aligned} \quad (4.1)$$

where $T \triangleq \{\hat{t} = (t_1, t_2, \dots, t_K)^T | t_0 \leq t_1 \leq t_2 \leq \dots \leq t_K \leq t_f\}$.

In order to solve a Stage 1 problem, 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 [23]), 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 1(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 $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$, we can still solve Stage 1(b) by employing some nonlinear optimization algorithms. Let us elaborate more on Stage 1(a) and 1(b) in the followings.

Stage 1(a)

For Stage 1(a) where a switching sequence $\sigma = ((t_0, i_0), (t_1, e_1), \dots, (t_K, e_K))$ is given, finding $J_1(\hat{t})$ for the corresponding $\hat{t} = (t_1, \dots, t_K)^T$ is a conventional optimal control problem. Note that although different subsystems are active in different time intervals, the problem is conventional since these intervals are fixed. In Stage 1(a), we need to find an optimal continuous input u and the corresponding minimum J . The only difference between Stage 1(a) and most of the problems in many optimal control texts is that in Stage 1(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. [11]) to prove the following necessary conditions. For simplicity of notations, in the following theorem, we assume that subsystem k is active in the time interval $[t_{k-1}, t_k)$ for $k = 1, \dots, K$ and subsystem $K + 1$ is active in $[t_K, t_{K+1}]$ where $t_{K+1} = t_f$.

Theorem 4.1 (Necessary Conditions for Stage 1(a)) *Consider the Stage 1(a) problem for Problem 2.1. 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 $t_{K+1} = t_f$. 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 $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \rightarrow \mathbb{R}^{l_f}\}$ at t_f . In order that u be optimal, it is necessary that there exists a vector function $p(t) = [p_1(t), \dots, p_n(t)]^T$, $t \in [t_0, t_f]$, such that the following conditions hold*

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

$$\text{State equation: } \frac{dx(t)}{dt} = \left[\frac{\partial H}{\partial p}(x(t), p(t), u(t), t) \right]^T \quad (4.2)$$

$$\text{Costate equation: } \frac{dp(t)}{dt} = - \left[\frac{\partial H}{\partial x}(x(t), p(t), u(t), t) \right]^T, \quad (4.3)$$

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

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

$$0 = \left[\frac{\partial H}{\partial u}(x(t), p(t), u(t), t) \right]^T. \quad (4.4)$$

(c). *At t_f , the function p satisfies*

$$p(t_f) = \left[\frac{\partial \psi}{\partial x}(x(t_f)) \right]^T + \left[\frac{\partial \phi_f}{\partial x}(x(t_f)) \right]^T \lambda \quad (4.5)$$

where λ is an l_f -dimensional vector.

(d). At any t_k , $k = 1, 2, \dots, K$, we have

$$p(t_k-) = p(t_k+). \quad (4.6)$$

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

$$J' = \psi(x(t_f)) + \lambda^T \phi_f(x(t_f)) + \sum_{k=1}^{K+1} \int_{t_{k-1}}^{t_k} [L(x, u, t) + p^T(t)(f_k(x, u, t) - \dot{x})] dt. \quad (4.7)$$

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

$$J' = \psi(x(t_f)) + \lambda^T \phi_f(x(t_f)) + \sum_{k=1}^{K+1} \int_{t_{k-1}}^{t_k} [H(x, p, u, t) - p^T \dot{x}] dt. \quad (4.8)$$

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

$$\begin{aligned} \delta J' &= \left[\frac{\partial \psi}{\partial x}(x(t_f)) + \lambda^T \frac{\partial \phi_f}{\partial x}(x(t_f)) - p^T(t_f) \right] \delta x(t_f) + \sum_{k=1}^K [p^T(t_k+) - p^T(t_k-)] \delta x(t_k) \\ &+ \sum_{k=1}^{K+1} \int_{t_{k-1}}^{t_k} \left[\left(\frac{\partial H}{\partial x} + \dot{p}^T \right) \delta x + \frac{\partial H}{\partial u} \delta u + \left(\frac{\partial H}{\partial p} - \dot{x}^T \right) \delta p \right] dt. \end{aligned} \quad (4.9)$$

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). \square

The above necessary conditions will be used in Section 6 in the development of a method for finding $\frac{\partial J_1}{\partial t}$ and $\frac{\partial^2 J_1}{\partial t^2}$. 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 1(b)

In Stage 1(b), we need to solve the constrained nonlinear optimization problem (4.1) with simple constraints. Computational methods for the solution 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 $\frac{\partial J_1}{\partial t}$ and second-order derivative $\frac{\partial^2 J_1}{\partial t^2}$. In the computation of the examples in this paper, we use the gradient projection method (using $\frac{\partial J_1}{\partial t}$) and the constrained Newton's method (using $\frac{\partial J_1}{\partial t}$ and $\frac{\partial^2 J_1}{\partial t^2}$)

and their variations (see Section 2.3 in Bertsekas [3] for details). For more discussions on various methods for solving constrained nonlinear optimization problems, please also see [2, 15].

A Conceptual Algorithm

The following conceptual algorithm provides a framework for the optimization methodologies in the subsequent sections.

Algorithm 4.1 (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 (Stage 1(a)), find $J_1(\hat{t}^j)$.
- (3). Find $\frac{\partial J_1}{\partial t}(\hat{t}^j)$ and $\frac{\partial^2 J_1}{\partial t^2}(\hat{t}^j)$.
- (4). Use the gradient projection method or the constrained Newton's method (if $\frac{\partial^2 J_1}{\partial t^2}(\hat{t}^j)$ is known) to update \hat{t}^j to be $\hat{t}^{j+1} = \hat{t}^j + \alpha^j d\hat{t}^j$ (here the stepsize α^j is chosen using the Armijo's rule [3]). Set the iteration index $j = j + 1$.
- (5). Repeat Steps (2), (3), (4) and (5), until a prespecified termination condition is satisfied. \square

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 $\frac{\partial J_1}{\partial t}$ and $\frac{\partial^2 J_1}{\partial t^2}$ for Step (3).
- (c). A nonlinear optimization algorithm for Step (4).

In the above discussions, we have already addressed elements (a) and (c). (b) poses an obstacle for the use of Algorithm 4.1 because $\frac{\partial J_1}{\partial t}$ and $\frac{\partial^2 J_1}{\partial t^2}$ are not readily available. It is the task of the subsequent sections to address (b) and devise a method for deriving the values of $\frac{\partial J_1}{\partial t}$ and $\frac{\partial^2 J_1}{\partial t^2}$. Lastly, it should be pointed out that in Step (4), when we are searching for α^j , the optimal control algorithm for Stage 1(a) will also be used in order to obtain the value of J_1 at the intermediate trial \hat{t} 's.

5 An Equivalent Problem Formulation Based on Parameterization of the Switching Instants

In the following three sections, an approach to Stage 1 optimization based on equivalent transcription is presented. The first step is to transcribe a Stage 1 problem into an equivalent conventional optimal control problem parameterized by the switching instants. Based on the equivalent problem formulation, a method is then developed for searching for accurate values of $\frac{\partial J_1}{\partial t}$ and $\frac{\partial^2 J_1}{\partial t^2}$. The method is based on the solution of a two point boundary value differential algebraic equation (DAE). Furthermore, an efficient method for deriving the derivatives based on the solution of an initial value ordinary differential equation (ODE) is proposed for GSLQ problems.

Now let us describe the transcription of a Stage 1 problem into an equivalent optimal control 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 and clarity of presentation, in the followings, we concentrate on the case of two subsystems where subsystem 1 is active in the interval $t \in [0, t_1)$ 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 similar methods). It is straightforward to apply the methods developed in this paper to problems with several subsystems and more than one switchings; we will remark on this in the subsequent sections.

We consider the following Stage 1 problem.

Problem 5.1 *For a switched system*

$$\dot{x} = f_1(x, u, t), \quad 0 \leq t < t_1, \quad (5.1)$$

$$\dot{x} = f_2(x, u, t), \quad t_1 \leq t \leq t_f, \quad (5.2)$$

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

$$J = \psi(x(t_f)) + \int_{t_0}^{t_f} L(x, u, t) dt \quad (5.3)$$

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

We transcribe Problem 5.1 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

$$\frac{dx_{n+1}}{dt} = 0 \quad (5.4)$$

$$x_{n+1}(0) = t_1 \quad (5.5)$$

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

$$t = \begin{cases} t_0 + (x_{n+1} - t_0)\tau, & 0 \leq \tau \leq 1 \\ x_{n+1} + (t_f - x_{n+1})(\tau - 1), & 1 \leq \tau \leq 2. \end{cases} \quad (5.6)$$

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 5.1 can be transcribed into the following equivalent problem.

Problem 5.2 (An Equivalent Problem) *For a system with dynamics*

$$\frac{dx(\tau)}{d\tau} = (x_{n+1} - t_0)f_1(x, u, t_0 + (x_{n+1} - t_0)\tau) \quad (5.7)$$

$$\frac{dx_{n+1}}{d\tau} = 0 \quad (5.8)$$

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

$$\frac{dx(\tau)}{d\tau} = (t_f - x_{n+1})f_2(x, u, x_{n+1} + (t_f - x_{n+1})(\tau - 1)) \quad (5.9)$$

$$\frac{dx_{n+1}}{d\tau} = 0 \quad (5.10)$$

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

$$\begin{aligned} J = & \psi(x(2)) + \int_0^1 (x_{n+1} - t_0)L(x, u, t_0 + (x_{n+1} - t_0)\tau) d\tau \\ & + \int_1^2 (t_f - x_{n+1})L(x, u, x_{n+1} + (t_f - x_{n+1})(\tau - 1)) d\tau \end{aligned} \quad (5.11)$$

is minimized. Here t_f , $x(0) = x_0$ are given. \square

Remark 5.1 Problem 5.2 and Problem 5.1 are equivalent in the sense that an optimal solution for Problem 5.2 is an optimal solution for Problem 5.1 by a proper change of independent variables as in (5.6) and by regarding $x_{n+1} = t_1$, and vice versa. \square

Remark 5.2 The equivalent Problem 5.2 provides us with the advantage, namely that it no longer has a varying switching instant. In fact, because x_{n+1} is actually an unknown constant throughout $\tau \in [0, 2]$, Problem 5.2 can be regarded as a conventional optimal control problem with an unknown parameter x_{n+1} . The problem is conventional because it has fixed time instant when the system dynamics change. In the subsequent discussion, we adopt the viewpoint of regarding x_{n+1} as an unknown parameter for optimal control problem with cost (5.11) and subsystems (5.7) and (5.9), i.e., we can regard Problem 5.2 as an optimal control problem parameterized by the switching instant x_{n+1} . \square

6 A Method Based on Solving a Boundary Value Differential Algebraic Equation

In this section, based on the equivalent Problem 5.2, we develop a method which can give us accurate numerical value of $\frac{dJ}{dt_1}$. The method is based on the solution of a two point boundary value differential algebraic equation (DAE) which is formed by the state, costate, stationarity equations, the boundary and continuity conditions for Problem 5.2 and their differentiations with respect to the parameter x_{n+1} . In the followings, we denote $\frac{\partial L}{\partial x}$, $\frac{\partial L}{\partial u}$ as row vectors and we denote $\frac{\partial f}{\partial x}$ as an $n \times n$ matrix whose (i_1, i_2) -th element is $\frac{\partial f_{i_1}}{\partial x_{i_2}}$. Similar notations apply to $\frac{\partial H}{\partial x}$, $\frac{\partial H}{\partial u}$, $\frac{\partial f}{\partial u}$, etc.

Consider the equivalent Problem 5.2, define

$$\tilde{f}_1(x, u, \tau, x_{n+1}) \triangleq (x_{n+1} - t_0)f_1(x, u, t_0 + (x_{n+1} - t_0)\tau), \quad (6.1)$$

$$\tilde{f}_2(x, u, \tau, x_{n+1}) \triangleq (t_f - x_{n+1})f_2(x, u, x_{n+1} + (t_f - x_{n+1})(\tau - 1)), \quad (6.2)$$

$$\tilde{L}_1(x, u, \tau, x_{n+1}) \triangleq (x_{n+1} - t_0)L(x, u, t_0 + (x_{n+1} - t_0)\tau), \quad (6.3)$$

$$\tilde{L}_2(x, u, \tau, x_{n+1}) \triangleq (t_f - x_{n+1})L(x, u, x_{n+1} + (t_f - x_{n+1})(\tau - 1)). \quad (6.4)$$

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

$$H(x, p, u, \tau, x_{n+1}) \triangleq \begin{cases} \tilde{L}_1(x, u, \tau, x_{n+1}) + p^T \tilde{f}_1(x, u, \tau, x_{n+1}), \\ \quad \text{for } \tau \in [0, 1), \\ \tilde{L}_2(x, u, \tau, x_{n+1}) + p^T \tilde{f}_2(x, u, \tau, x_{n+1}), \\ \quad \text{for } \tau \in [1, 2]. \end{cases} \quad (6.5)$$

Assume that a parameter x_{n+1} is given, then we can apply Theorem 4.1 to the equivalent Problem 5.2. The necessary conditions (a) and (b) provides us with the following state, costate and stationarity equations

$$\text{State equation: } \frac{\partial x}{\partial \tau} = \left(\frac{\partial H}{\partial p}\right)^T = \tilde{f}_1(x, u, \tau, x_{n+1}), \quad (6.6)$$

$$\text{Costate equation: } \frac{\partial p}{\partial \tau} = -\left(\frac{\partial H}{\partial x}\right)^T = -\left(\frac{\partial \tilde{f}_1}{\partial x}\right)^T p - \left(\frac{\partial \tilde{L}_1}{\partial x}\right)^T, \quad (6.7)$$

$$\text{Stationarity equation: } 0 = \left(\frac{\partial H}{\partial u}\right)^T = \left(\frac{\partial \tilde{f}_1}{\partial u}\right)^T p + \left(\frac{\partial \tilde{L}_1}{\partial u}\right)^T. \quad (6.8)$$

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

$$\text{State equation: } \frac{\partial x}{\partial \tau} = \left(\frac{\partial H}{\partial p}\right)^T = \tilde{f}_2(x, u, \tau, x_{n+1}), \quad (6.9)$$

$$\text{Costate equation: } \frac{\partial p}{\partial \tau} = -\left(\frac{\partial H}{\partial x}\right)^T = -\left(\frac{\partial \tilde{f}_2}{\partial x}\right)^T p - \left(\frac{\partial \tilde{L}_2}{\partial x}\right)^T, \quad (6.10)$$

$$\text{Stationarity equation: } 0 = \left(\frac{\partial H}{\partial u}\right)^T = \left(\frac{\partial \tilde{f}_2}{\partial u}\right)^T p + \left(\frac{\partial \tilde{L}_2}{\partial u}\right)^T. \quad (6.11)$$

in $\tau \in [1, 2]$. Note that the p and u corresponding to the optimal solution are also functions of τ and x_{n+1} . Therefore, in the following, we denote them as $p = p(\tau, x_{n+1})$ and $u = u(\tau, x_{n+1})$.

From the necessary condition (c) of Theorem 4.1, we obtain the boundary conditions

$$x(0, x_{n+1}) = x_0, \quad (6.12)$$

$$p(2, x_{n+1}) = \left(\frac{\partial \psi}{\partial x}(x(2, x_{n+1}))\right)^T. \quad (6.13)$$

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

$$p(1-, x_{n+1}) = p(1+, x_{n+1}). \quad (6.14)$$

(6.6)-(6.8), (6.9)-(6.11) along with boundary conditions (6.12) and (6.13) form a two point boundary value differential algebraic equation (DAE) which is 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 obtain 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}

$$J_1(x_{n+1}) = \psi(x(2, x_{n+1})) + \int_0^1 \tilde{L}_1(x, u, \tau, x_{n+1}) d\tau + \int_1^2 \tilde{L}_2(x, u, \tau, x_{n+1}) d\tau. \quad (6.15)$$

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

$$\begin{aligned} \frac{dJ_1}{dx_{n+1}} &= \frac{\partial \psi(x(2, x_{n+1}))}{\partial x} \frac{\partial x(2, x_{n+1})}{\partial x_{n+1}} + \int_0^1 [L(x, u, t_0 + (x_{n+1} - t_0)\tau) \\ &\quad + (x_{n+1} - t_0) \left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial L}{\partial u} \frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial L}{\partial t} \right)] d\tau \\ &\quad + \int_1^2 [-L(x, u, x_{n+1} + (t_f - x_{n+1})(\tau - 1)) \\ &\quad + (t_f - x_{n+1}) \left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial L}{\partial u} \frac{\partial u}{\partial x_{n+1}} + (2 - \tau) \frac{\partial L}{\partial t} \right)] d\tau. \end{aligned} \quad (6.16)$$

So we need to obtain the function $\frac{\partial x(\tau, x_{n+1})}{\partial x_{n+1}}$ and $\frac{\partial u(\tau, x_{n+1})}{\partial x_{n+1}}$ (here we assume that x_{n+1} is fixed) in order to obtain the value $\frac{dJ_1}{dx_{n+1}}$. By differentiating the above state, costate and stationarity equations (6.6)-(6.8) and (6.9)-(6.11) with respect to x_{n+1} , we obtain

$$\begin{aligned} \frac{\partial}{\partial \tau} \left(\frac{\partial x}{\partial x_{n+1}} \right) &= \frac{\partial}{\partial x_{n+1}} \left(\frac{\partial x}{\partial \tau} \right) \\ &= f_1 + (x_{n+1} - t_0) \left(\frac{\partial f_1}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial f_1}{\partial u} \frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial f_1}{\partial t} \right), \end{aligned} \quad (6.17)$$

$$\begin{aligned} \frac{\partial}{\partial \tau} \left(\frac{\partial p}{\partial x_{n+1}} \right) &= -\frac{\partial}{\partial x_{n+1}} \left(\frac{\partial p}{\partial \tau} \right) = -\left(\frac{\partial f_1}{\partial x} \right)^T p - \left(\frac{\partial L}{\partial x} \right)^T - (x_{n+1} - t_0) \left[\left(\frac{\partial f_1}{\partial x} \right)^T \frac{\partial p}{\partial x_{n+1}} \right. \\ &\quad \left. + (p^T \frac{\partial^2 f_1}{\partial x^2} \frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_1}{\partial x \partial u} \frac{\partial u}{\partial x_{n+1}})^T + \tau (p^T \frac{\partial^2 f_1}{\partial x \partial t})^T \right. \\ &\quad \left. + \frac{\partial^2 L}{\partial x^2} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial x \partial u} \frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial^2 L}{\partial x \partial t} \right], \end{aligned} \quad (6.18)$$

$$\begin{aligned} 0 &= \left(\frac{\partial f_1}{\partial u} \right)^T p + \left(\frac{\partial L}{\partial u} \right)^T + (x_{n+1} - t_0) \left[\left(\frac{\partial f_1}{\partial u} \right)^T \frac{\partial p}{\partial x_{n+1}} \right. \\ &\quad \left. + (p^T \frac{\partial^2 f_1}{\partial u \partial x} \frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_1}{\partial u^2} \frac{\partial u}{\partial x_{n+1}})^T + \tau (p^T \frac{\partial^2 f_1}{\partial u \partial t})^T \right. \\ &\quad \left. + \frac{\partial^2 L}{\partial u \partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial u^2} \frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial^2 L}{\partial u \partial t} \right], \end{aligned} \quad (6.19)$$

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

$$\begin{aligned} \frac{\partial}{\partial \tau} \left(\frac{\partial x}{\partial x_{n+1}} \right) &= \frac{\partial}{\partial x_{n+1}} \left(\frac{\partial x}{\partial \tau} \right) \\ &= -f_2 + (t_f - x_{n+1}) \left(\frac{\partial f_2}{\partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial f_2}{\partial u} \frac{\partial u}{\partial x_{n+1}} + (2 - \tau) \frac{\partial f_2}{\partial t} \right), \end{aligned} \quad (6.20)$$

$$\begin{aligned} \frac{\partial}{\partial \tau} \left(\frac{\partial p}{\partial x_{n+1}} \right) &= -\frac{\partial}{\partial x_{n+1}} \left(\frac{\partial p}{\partial \tau} \right) = \left(\frac{\partial f_2}{\partial x} \right)^T p + \left(\frac{\partial L}{\partial x} \right)^T - (t_f - x_{n+1}) \left[\left(\frac{\partial f_2}{\partial x} \right)^T \frac{\partial p}{\partial x_{n+1}} \right. \\ &\quad \left. + (p^T \frac{\partial^2 f_2}{\partial x^2} \frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_2}{\partial x \partial u} \frac{\partial u}{\partial x_{n+1}})^T + (2 - \tau) (p^T \frac{\partial^2 f_2}{\partial x \partial t})^T \right. \\ &\quad \left. + \frac{\partial^2 L}{\partial x^2} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial x \partial u} \frac{\partial u}{\partial x_{n+1}} + (2 - \tau) \frac{\partial^2 L}{\partial x \partial t} \right], \end{aligned} \quad (6.21)$$

$$\begin{aligned} 0 &= -\left(\frac{\partial f_2}{\partial u} \right)^T p - \left(\frac{\partial L}{\partial u} \right)^T + (x_{n+1} - t_0) \left[\left(\frac{\partial f_2}{\partial u} \right)^T \frac{\partial p}{\partial x_{n+1}} \right. \\ &\quad \left. + (p^T \frac{\partial^2 f_2}{\partial u \partial x} \frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_2}{\partial u^2} \frac{\partial u}{\partial x_{n+1}})^T + (2 - \tau) (p^T \frac{\partial^2 f_2}{\partial u \partial t})^T \right. \\ &\quad \left. + \frac{\partial^2 L}{\partial u \partial x} \frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial u^2} \frac{\partial u}{\partial x_{n+1}} + (2 - \tau) \frac{\partial^2 L}{\partial u \partial t} \right], \end{aligned} \quad (6.22)$$

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

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

Differentiating the boundary conditions (6.12), (6.13) and the continuity condition (6.14) with respect to x_{n+1} , we obtain

$$\frac{\partial x(0, x_{n+1})}{\partial x_{n+1}} = 0, \quad (6.23)$$

$$\frac{\partial p(2, x_{n+1})}{\partial x_{n+1}} = \frac{\partial^2 \psi(x(2, x_{n+1}))}{\partial x^2} \frac{\partial x(2, x_{n+1})}{\partial x_{n+1}}, \quad (6.24)$$

$$\frac{\partial p(1-, x_{n+1})}{\partial x_{n+1}} = \frac{\partial p(1+, x_{n+1})}{\partial x_{n+1}}. \quad (6.25)$$

It can now be observed that (6.6)-(6.8), (6.9)-(6.11) and (6.17)-(6.19), (6.20)-(6.22) along with the boundary conditions (6.12), (6.13) and (6.23), (6.24) and with the continuity conditions (6.14), (6.25) form a two point boundary value DAE for $x(\tau, x_{n+1})$, $p(\tau, x_{n+1})$, $u(\tau, x_{n+1})$ and $\frac{\partial x(\tau, x_{n+1})}{\partial x_{n+1}}$, $\frac{\partial p(\tau, x_{n+1})}{\partial x_{n+1}}$, $\frac{\partial u(\tau, x_{n+1})}{\partial x_{n+1}}$. By solving them and substitute the result into (6.16), we can obtain $\frac{dJ_1}{dx_{n+1}}$.

Remark 6.1 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} , $\frac{\partial u}{\partial x_{n+1}}$ can also be expressed as a function of x , p , $\frac{\partial x}{\partial x_{n+1}}$ and $\frac{\partial p}{\partial x_{n+1}}$. If we substitute these functions for u and $\frac{\partial u}{\partial x_{n+1}}$ 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 , $\frac{\partial x}{\partial x_{n+1}}$ and $\frac{\partial p}{\partial x_{n+1}}$, which can be solved more easily than the DAE (e.g., using shooting methods for two point boundary value differential equations). \square

Remark 6.2 In general, we need to resort to numerical methods to find the solution to the two point boundary value DAE. Methods similar to the shooting methods for solving boundary value differential equations can be adopted. We can guess the unknown boundary conditions at $\tau = 0$ and then discretize the DAE and numerically solve it forward in τ . At the end time $\tau = 2$, we regard the minimum square error between the required end condition and the corresponding values for the current solution as a function of the unknown initial conditions. Numerical optimization methods are then used to minimize the minimum square error. Ideally, it will converge to 0. Once we have successfully located the correct initial conditions, the trajectory of $\frac{\partial x}{\partial x_{n+1}}$ and $\frac{\partial u}{\partial x_{n+1}}$ as a function of τ can be solved from the DAE. However, since we are using numerical methods to solve it, we can only obtain their values at some discrete instants. In this case, if we regard the integration in (6.16) as a sum, it may not be accurate enough. One way to address this issue is to

introduce an auxiliary state variable y that satisfies

$$\frac{dy}{d\tau} = \begin{cases} L(x, u, t_0 + (x_{n+1} - t_0)\tau) + (x_{n+1} - t_0)\left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial x_{n+1}}\right. \\ \left. + \frac{\partial L}{\partial u} \frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial L}{\partial t}\right), \text{ for } \tau \in [0, 1), \\ -L(x, u, x_{n+1} + (t_f - x_{n+1})(\tau - 1)) + (t_f - x_{n+1})\left(\frac{\partial L}{\partial x} \frac{\partial x}{\partial x_{n+1}}\right. \\ \left. + \frac{\partial L}{\partial u} \frac{\partial u}{\partial x_{n+1}} + (2 - \tau) \frac{\partial L}{\partial t}\right), \text{ for } \tau \in [1, 2], \end{cases} \quad (6.26)$$

$$y(0) = 0. \quad (6.27)$$

Solving the differential equation (6.26) along with the DAE, we can obtain the value

$$\frac{dJ_1}{dx_{n+1}} = \frac{\partial \psi(x(2), x_{n+1})}{\partial x} \frac{\partial x(2), x_{n+1}}{\partial x_{n+1}} + y(2). \quad (6.28)$$

In this way, the resultant $\frac{dJ_1}{dx_{n+1}}$ would be quite accurate. \square

Remark 6.3 (Several Subsystems and More Than One Switchings) There is no difficulty in applying the above method to problems with several subsystems and more than one switchings. Assuming that there are K switchings, we can transcribe the Stage 1 problem into an equivalent problem by introducing K new state variables x_{n+k} 's, $k = 1, \dots, K$ which correspond to the switching instants t_k 's and satisfies

$$\frac{dx_{n+k}}{dt} = 0, \quad (6.29)$$

$$x_{n+k}(0) = t_k. \quad (6.30)$$

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 $t = t_1$, \dots , $\tau = K + 1$ corresponds to $t = t_f$. It is then straightforward to apply the necessary conditions in Theorem 4.1 to the equivalent problem to come up with the state, costate, stationarity equations, the boundary and continuity conditions. Similarly to the case of a single switching, we can then obtain x , p , u , $\frac{\partial x}{\partial x_{n+k}}$'s, $\frac{\partial p}{\partial x_{n+k}}$'s and $\frac{\partial u}{\partial x_{n+k}}$'s by solving the two point boundary value DAE formed by the state, costate, stationarity equations, the boundary and continuity conditions along with their differentiations with respect to x_{n+k} 's. By substituting them into the expressions of $\frac{\partial J_1}{\partial x_{n+k}}$'s which can be derived similarly to (6.16), we can then find the accurate values of $\frac{\partial J_1}{\partial x_{n+k}}$'s. \square

Remark 6.4 (Second Order Derivatives) If we take second order partial derivatives on equation (6.15), we can obtain the expression for $\frac{d^2 J_1(t_1)}{dt_1^2}$ which depends on the values of $\frac{\partial x}{\partial x_{n+1}}$, $\frac{\partial u}{\partial x_{n+1}}$, $\frac{\partial^2 x(\tau, x_{n+1})}{\partial x_{n+1}^2}$ and $\frac{\partial^2 u(\tau, x_{n+1})}{\partial x_{n+1}^2}$. Similarly to the above procedure, taking first and second order derivatives of the state, costate and stationarity equations (6.6)-(6.8) and (6.9)-(6.11), the boundary conditions (6.12)-(6.13) and the continuity condition (6.14) with respect to x_{n+1} will result in a two point boundary value DAE in x , p , u , $\frac{\partial x}{\partial x_{n+1}}$, $\frac{\partial p}{\partial x_{n+1}}$, $\frac{\partial u}{\partial x_{n+1}}$, $\frac{\partial^2 x}{\partial x_{n+1}^2}$, $\frac{\partial^2 p}{\partial x_{n+1}^2}$, $\frac{\partial^2 u}{\partial x_{n+1}^2}$. There is no difficulty in obtaining the values of $\frac{dJ_1(t_1)}{dt_1}$ and $\frac{d^2 J_1(t_1)}{dt_1^2}$ 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. \square

7 General Switched Linear Quadratic Problems

In this section, the idea in Section 6 is applied to general switched linear quadratic (GSLQ) problems and a more efficient method for deriving accurate numerical values of $\frac{dJ_1}{dx_{n+1}}$ is developed. The method only needs to solve an initial value ODE which is formed by the parameterized general Riccati equation and its differentiation with respect to the switching instant in order to compute the values of $\frac{dJ_1}{dx_{n+1}}$. For simplicity of notation, we consider the following Stage 1 GSLQ problem with two subsystems and one switching.

Problem 7.1 (Stage 1 of a GSLQ problem) *For a switched system*

$$\dot{x} = A_1x + B_1u, \quad t_0 \leq t < t_1 \quad (7.1)$$

$$\dot{x} = A_2x + B_2u, \quad t_1 \leq t \leq t_f, \quad (7.2)$$

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

$$\begin{aligned} J = & \frac{1}{2}x(t_f)^T Q_f x(t_f) + M_f x(t_f) + W_f + \int_{t_0}^{t_f} \left(\frac{1}{2}x^T Q x \right. \\ & \left. + x^T V u + \frac{1}{2}u^T R u + M x + N u + W \right) dt \end{aligned} \quad (7.3)$$

is minimized. Here 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 $Q_f \geq 0$, $Q \geq 0$ and $R > 0$. \square

In view of the method in Section 5, we transcribe Problem 7.1 into its equivalent problem.

Problem 7.2 (Equivalent to Problem 7.1) *For a system with dynamics*

$$\frac{dx(\tau)}{d\tau} = (x_{n+1} - t_0)(A_1x + B_1u), \quad (7.4)$$

$$\frac{dx_{n+1}}{d\tau} = 0, \quad (7.5)$$

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

$$\frac{dx(\tau)}{d\tau} = (t_f - x_{n+1})(A_2x + B_2u), \quad (7.6)$$

$$\frac{dx_{n+1}}{d\tau} = 0, \quad (7.7)$$

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

$$\begin{aligned} J = & \frac{1}{2}x(2)^T Q_f x(2) + M_f x(2) + W_f + \int_0^1 (x_{n+1} - t_0)L(x, u) \, d\tau \\ & + \int_1^2 (t_f - x_{n+1})L(x, u) \, d\tau \end{aligned} \quad (7.8)$$

where

$$L(x, u) = \frac{1}{2}x^T Q x + x^T V u + \frac{1}{2}u^T R u + M x + N u + W \quad (7.9)$$

is minimized. Here t_f , $x(0) = x_0$ are given. \square

Similar to Remark 5.2, Problem 7.2 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 7.2 as follows. We assume that the optimal value function is

$$V^*(x, \tau, x_{n+1}) = \frac{1}{2}x^T P(\tau, x_{n+1})x + S(\tau, x_{n+1})x + T(\tau, x_{n+1}) \quad (7.10)$$

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

$$-\frac{\partial V^*}{\partial \tau}(x, \tau, x_{n+1}) = \min_{u(\tau)} \{ (x_{n+1} - t_0)[L(x, u) + \frac{\partial V^*}{\partial x}(x, \tau, x_{n+1})f_1(x, u)] \} \quad (7.11)$$

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

$$-\frac{\partial V^*}{\partial \tau}(x, \tau, x_{n+1}) = \min_{u(\tau)} \{ (t_f - x_{n+1})[L(x, u) + \frac{\partial V^*}{\partial x}(x, \tau, x_{n+1})f_2(x, u)] \} \quad (7.12)$$

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

Using a method similar to the method for solving conventional linear quadratic regulator problems (see e.g., [7]), it can be obtained that the solution to (7.11) in the interval $\tau \in [0, 1)$ is

$$u(x, \tau, x_{n+1}) = -K(\tau, x_{n+1})x(\tau, x_{n+1}) - E(\tau, x_{n+1}) \quad (7.13)$$

where

$$K(\tau, x_{n+1}) = R^{-1}(B_1^T P(\tau, x_{n+1}) + V^T), \quad (7.14)$$

$$E(\tau, x_{n+1}) = R^{-1}(B_1^T S^T(\tau, x_{n+1}) + N^T), \quad (7.15)$$

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})

$$-\frac{\partial P}{\partial \tau} = (x_{n+1} - t_0)(Q + PA_1 + A_1^T P - (PB_1 + V)R^{-1}(B_1^T P + V^T)), \quad (7.16)$$

$$-\frac{\partial S}{\partial \tau} = (x_{n+1} - t_0)(M + SA_1 - (N + SB_1)R^{-1}(B_1^T P + V^T)), \quad (7.17)$$

$$-\frac{\partial T}{\partial \tau} = (x_{n+1} - t_0)(W - \frac{1}{2}(N + SB_1)R^{-1}(B_1^T S^T + N^T)). \quad (7.18)$$

The solution to (7.12) in the interval $\tau \in [1, 2]$ is

$$u(x, \tau, x_{n+1}) = -K(\tau, x_{n+1})x(\tau, x_{n+1}) - E(\tau, x_{n+1}) \quad (7.19)$$

where

$$K(\tau, x_{n+1}) = R^{-1}(B_2^T P(\tau, x_{n+1}) + V^T), \quad (7.20)$$

$$E(\tau, x_{n+1}) = R^{-1}(B_2^T S^T(\tau, x_{n+1}) + N^T), \quad (7.21)$$

and P , S and T satisfies the following parameterized general Riccati equation

$$-\frac{\partial P}{\partial \tau} = (t_f - x_{n+1})(Q + PA_2 + A_2^T P - (PB_2 + V)R^{-1}(B_2^T P + V^T)), \quad (7.22)$$

$$-\frac{\partial S}{\partial \tau} = (t_f - x_{n+1})(M + SA_2 - (N + SB_2)R^{-1}(B_2^T P + V^T)), \quad (7.23)$$

$$-\frac{\partial T}{\partial \tau} = (t_f - x_{n+1})(W - \frac{1}{2}(N + SB_2)R^{-1}(B_2^T S^T + N^T)). \quad (7.24)$$

Once we have solved (7.16-7.18) and (7.22-7.24) (for a fixed x_{n+1}), we can obtain the parameterized optimal cost at $\tau = 0$, i.e., the optimal J_1 under fixed x_{n+1} as

$$\begin{aligned} J_1(t_1) &= J_1(x_{n+1}) = V^*(x_0, 0, x_{n+1}) \\ &= \frac{1}{2}x_0^T P(0, x_{n+1})x_0 + S(0, x_{n+1})x_0 + T(0, x_{n+1}). \end{aligned} \quad (7.25)$$

From (7.25), we have

$$\begin{aligned} \frac{dJ_1}{dx_{n+1}}(x_{n+1}) &= \frac{\partial V^*}{\partial x_{n+1}}(x_0, 0, x_{n+1}) \\ &= \frac{1}{2}x_0^T \frac{\partial P}{\partial x_{n+1}}(0, x_{n+1})x_0 + \frac{\partial S}{\partial x_{n+1}}(0, x_{n+1})x_0 + \frac{\partial T}{\partial x_{n+1}}(0, x_{n+1}). \end{aligned} \quad (7.26)$$

In order to obtain the value of $\frac{dJ_1}{dx_{n+1}}$ from (7.26), we need to know $\frac{\partial P}{\partial x_{n+1}}$, $\frac{\partial S}{\partial x_{n+1}}$ and $\frac{\partial T}{\partial x_{n+1}}$ at $(0, x_{n+1})$. To obtain these values, we differentiate (7.16-7.18) and (7.22-7.24) with respect to x_{n+1} to obtain the following equations

$$\begin{aligned} -\frac{\partial}{\partial \tau} \left(\frac{\partial P}{\partial x_{n+1}} \right) &= (Q + PA_1 + A_1^T P - (PB_1 + V)R^{-1}(B_1^T P + V^T)) \\ &\quad + (x_{n+1} - t_0) \left(\frac{\partial P}{\partial x_{n+1}} A_1 + A_1^T \frac{\partial P}{\partial x_{n+1}} - \left(\frac{\partial P}{\partial x_{n+1}} B_1 \right) R^{-1} (B_1^T P + V^T) \right) \\ &\quad - (PB_1 + V)R^{-1} \left(B_1^T \frac{\partial P}{\partial x_{n+1}} \right) \end{aligned} \quad (7.27)$$

$$\begin{aligned} -\frac{\partial}{\partial \tau} \left(\frac{\partial S}{\partial x_{n+1}} \right) &= (M + SA_1 - (N + SB_1)R^{-1}(B_1^T P + V^T)) + (x_{n+1} - t_0) \left(\frac{\partial S}{\partial x_{n+1}} A_1 \right. \\ &\quad \left. - \left(\frac{\partial S}{\partial x_{n+1}} B_1 \right) R^{-1} (B_1^T P + V^T) - (N + SB_1)R^{-1} \left(B_1^T \frac{\partial P}{\partial x_{n+1}} \right) \right) \end{aligned} \quad (7.28)$$

$$\begin{aligned} -\frac{\partial}{\partial \tau} \left(\frac{\partial T}{\partial x_{n+1}} \right) &= \left(W - \frac{1}{2}(N + SB_1)R^{-1}(B_1^T S^T + N^T) \right) \\ &\quad + (x_{n+1} - t_0) \left(-\frac{1}{2} \left(\frac{\partial S}{\partial x_{n+1}} B_1 \right) R^{-1} (B_1^T S^T + N^T) \right. \\ &\quad \left. - \frac{1}{2} (N + SB_1)R^{-1} \left(B_1^T \left(\frac{\partial S}{\partial x_{n+1}} \right)^T \right) \right) \end{aligned} \quad (7.29)$$

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

$$\begin{aligned} -\frac{\partial}{\partial \tau} \left(\frac{\partial P}{\partial x_{n+1}} \right) &= -(Q + PA_2 + A_2^T P - (PB_2 + V)R^{-1}(B_2^T P + V^T)) \\ &\quad + (t_f - x_{n+1}) \left(\frac{\partial P}{\partial x_{n+1}} A_2 + A_2^T \frac{\partial P}{\partial x_{n+1}} - \left(\frac{\partial P}{\partial x_{n+1}} B_2 \right) R^{-1} (B_2^T P + V^T) \right) \\ &\quad - (PB_2 + V)R^{-1} \left(B_2^T \frac{\partial P}{\partial x_{n+1}} \right) \end{aligned} \quad (7.30)$$

$$\begin{aligned} -\frac{\partial}{\partial \tau} \left(\frac{\partial S}{\partial x_{n+1}} \right) &= -(M + SA_2 - (N + SB_2)R^{-1}(B_2^T P + V^T)) + (t_f - x_{n+1}) \left(\frac{\partial S}{\partial x_{n+1}} A_2 \right. \\ &\quad \left. - \left(\frac{\partial S}{\partial x_{n+1}} B_2 \right) R^{-1} (B_2^T P + V^T) - (N + SB_2)R^{-1} \left(B_2^T \frac{\partial P}{\partial x_{n+1}} \right) \right) \end{aligned} \quad (7.31)$$

$$\begin{aligned} -\frac{\partial}{\partial \tau} \left(\frac{\partial T}{\partial x_{n+1}} \right) &= -(W - \frac{1}{2}(N + SB_2)R^{-1}(B_2^T S^T + N^T)) \\ &\quad + (t_f - x_{n+1}) \left(-\frac{1}{2} \left(\frac{\partial S}{\partial x_{n+1}} B_2 \right) R^{-1} (B_2^T S^T + N^T) \right. \\ &\quad \left. - \frac{1}{2} (N + SB_2)R^{-1} \left(B_2^T \left(\frac{\partial S}{\partial x_{n+1}} \right)^T \right) \right) \end{aligned} \quad (7.32)$$

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

The equations (7.16-7.18) and (7.27-7.29) for $\tau \in [0, 1)$ and the equations (7.22-7.24) and (7.30-7.32) for $\tau \in [1, 2]$ together with the following boundary conditions at $\tau = 2$

$$P(2, x_{n+1}) = Q_f \quad (7.33)$$

$$S(2, x_{n+1}) = M_f \quad (7.34)$$

$$T(2, x_{n+1}) = W_f \quad (7.35)$$

$$\frac{\partial P}{\partial x_{n+1}}(2, x_{n+1}) = 0 \quad (7.36)$$

$$\frac{\partial S}{\partial x_{n+1}}(2, x_{n+1}) = 0 \quad (7.37)$$

$$\frac{\partial T}{\partial x_{n+1}}(2, x_{n+1}) = 0 \quad (7.38)$$

form an initial value ordinary differential equation (ODE) for P , S , T , $\frac{\partial P}{\partial x_{n+1}}$, $\frac{\partial S}{\partial x_{n+1}}$ and $\frac{\partial T}{\partial x_{n+1}}$ which can be solved efficiently using the function `ode45` in MATLAB. From the solution of this ODE, values of $\frac{\partial P}{\partial x_{n+1}}(0, x_{n+1})$, $\frac{\partial S}{\partial x_{n+1}}(0, x_{n+1})$ and $\frac{\partial T}{\partial x_{n+1}}(0, x_{n+1})$ can be obtained and substituted into (7.26) to obtain the value of $\frac{dJ_1}{dt_1}$. The conceptual Algorithm 4.1 can then be applied.

Remark 7.1 (Several Subsystems and More Than One Switchings) It can be seen that there is no difficulty in applying the above method to GSLQ problems with several subsystems and more than one switchings. First of all, we can translate the problem to an equivalent problem in $\tau \in [0, K + 1]$ if there are K switchings as mentioned in Remark 6.3. It is then straightforward to differentiate the Riccati equation which are parameterized by x_{n+1}, \dots, x_{n+K} (i.e., t_1, \dots, t_K) to obtain additional differential equations for $\frac{\partial P}{\partial x_{n+k}}$'s, $\frac{\partial S}{\partial x_{n+k}}$'s and $\frac{\partial T}{\partial x_{n+k}}$'s. Along with the boundary conditions $P(K + 1, x_{n+1}, \dots, x_{n+K}) = Q_f$, $S(K + 1, x_{n+1}, \dots, x_{n+K}) = M_f$, $T(K + 1, x_{n+1}, \dots, x_{n+K}) = W_f$, $\frac{\partial P}{\partial x_{n+k}}(K + 1, x_{n+1}, \dots, x_{n+K}) = 0$, $\frac{\partial S}{\partial x_{n+k}}(K + 1, x_{n+1}, \dots, x_{n+K}) = 0$

and $\frac{\partial T}{\partial x_{n+k}}(K+1, x_{n+1}, \dots, x_{n+K}) = 0$ for all $1 \leq k \leq K$, we can solve the resultant initial value ODE backwards in τ to find their values at $\tau = 0$. Once we have their values at $\tau = 0$, we can substitute the values into

$$\begin{aligned} \frac{\partial J_1}{\partial x_{n+k}} &= \frac{\partial V^*}{\partial x_{n+k}}(x_0, 0, x_{n+1}, \dots, x_{n+K}) = \frac{1}{2}x_0^T \frac{\partial P}{\partial x_{n+k}}(0, x_{n+1}, \dots, x_{n+k})x_0 \\ &+ \frac{\partial S}{\partial x_{n+k}}(0, x_{n+1}, \dots, x_{n+k})x_0 + \frac{\partial T}{\partial x_{n+k}}(0, x_{n+1}, \dots, x_{n+k}) \end{aligned} \quad (7.39)$$

to derive the accurate values of $\frac{\partial J_1}{\partial t_k}$'s. \square

Remark 7.2 (Second Order Derivatives) It is not difficult to see that if we take second order partial derivatives of equation (7.25), we obtain

$$\begin{aligned} \frac{d^2 J_1}{dx_{n+1}^2}(t_1) &= \frac{\partial^2 V^*}{\partial x_{n+1}^2}(x_0, x_{n+1}, 0) \\ &= \frac{1}{2}x_0^T \frac{\partial^2 P}{\partial x_{n+1}^2}(0, x_{n+1})x_0 + \frac{\partial^2 S}{\partial x_{n+1}^2}(0, x_{n+1})x_0 + \frac{\partial^2 T}{\partial x_{n+1}^2}(0, x_{n+1}). \end{aligned} \quad (7.40)$$

While following similar ideas of differentiation of the parameterized general Riccati equation, we can take first and second-order differentiations of (7.16)-(7.18) and (7.22)-(7.24) with respect to x_{n+1} and form a set of ordinary differential equations. Along with the initial conditions (7.33)-(7.38) and 0's at $\tau = 2$ for $\frac{\partial^2 P}{\partial x_{n+1}^2}$, $\frac{\partial^2 S}{\partial x_{n+1}^2}$ and $\frac{\partial^2 T}{\partial x_{n+1}^2}$, the resultant initial value ODE for P , S , T , $\frac{\partial P}{\partial x_{n+1}}$, $\frac{\partial S}{\partial x_{n+1}}$, $\frac{\partial T}{\partial x_{n+1}}$, $\frac{\partial^2 P}{\partial x_{n+1}^2}$, $\frac{\partial^2 S}{\partial x_{n+1}^2}$ and $\frac{\partial^2 T}{\partial x_{n+1}^2}$ can be readily solved and hence the accurate value of $\frac{d^2 J_1}{dx_{n+1}^2}$ can be obtained. With this information, the conceptual Algorithm 4.1 can then be applied. \square

8 Some Examples

In this section, we illustrate the effectiveness of the methods developed in Sections 6 and 7 using several examples.

Example 8.1 Consider a switched system consisting of

$$\text{subsystem 1: } \begin{cases} \dot{x}_1 = -x_1 + 2x_1u \\ \dot{x}_2 = x_2 + x_2u \end{cases} \quad (8.1)$$

$$\text{subsystem 2: } \begin{cases} \dot{x}_1 = x_1 - 3x_1u \\ \dot{x}_2 = 2x_2 - 2x_2u \end{cases} \quad (8.2)$$

$$\text{subsystem 3: } \begin{cases} \dot{x}_1 = 2x_1 + x_1u \\ \dot{x}_2 = -x_2 + 3x_2u \end{cases} \quad (8.3)$$

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 ($0 \leq t_1 \leq t_2 \leq 3$). We want to find optimal switching instants t_1 , t_2 and an optimal input u such that the cost functional $J = \frac{1}{2}(x_1(3) - e^2)^2 + \frac{1}{2}(x_2(3) - e^2)^2 + \frac{1}{2} \int_0^3 u^2(t) dt$ is minimized. Here $x_1(0) = 1$ and $x_2(0) = 1$.

For this problem, we use the method in Section 6 to obtain the values of $\frac{\partial J_1}{\partial t_1}$ and $\frac{\partial J_1}{\partial t_2}$. Since the problem is linear in control, we need only to solve a two point boundary value differential equation. The resultant differential equation for $x_1, x_2, p_1, p_2, \frac{\partial x_1}{\partial x_{n+1}}, \frac{\partial x_2}{\partial x_{n+1}}, \frac{\partial p_1}{\partial x_{n+1}}, \frac{\partial p_2}{\partial x_{n+1}}, \frac{\partial x_1}{\partial x_{n+2}}, \frac{\partial x_2}{\partial x_{n+2}}, \frac{\partial p_1}{\partial x_{n+2}}$ and $\frac{\partial p_2}{\partial x_{n+2}}$ is formed by the following differential equations and their differentiations with respect to x_{n+1} and x_{n+2} .

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

$$\begin{cases} \frac{\partial x_1}{\partial \tau} = x_{n+1}(-x_1 + 2x_1 u) \\ \frac{\partial x_2}{\partial \tau} = x_{n+1}(x_2 + x_2 u) \\ \frac{\partial p_1}{\partial \tau} = -x_{n+1}(-p_1 + 2p_1 u) \\ \frac{\partial p_2}{\partial \tau} = -x_{n+1}(p_2 + p_2 u) \end{cases} \quad (8.4)$$

where $u = -2x_1 p_1 - x_2 p_2$.

For $\tau \in [1, 2)$, the differential equations are

$$\begin{cases} \frac{\partial x_1}{\partial \tau} = (x_{n+2} - x_{n+1})(x_1 - 3x_1 u) \\ \frac{\partial x_2}{\partial \tau} = (x_{n+2} - x_{n+1})(2x_2 - 2x_2 u) \\ \frac{\partial p_1}{\partial \tau} = -(x_{n+2} - x_{n+1})(p_1 - 3p_1 u) \\ \frac{\partial p_2}{\partial \tau} = -(x_{n+2} - x_{n+1})(2p_2 - 2p_2 u) \end{cases} \quad (8.5)$$

where $u = 3x_1 p_1 + 2x_2 p_2$.

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

$$\begin{cases} \frac{\partial x_1}{\partial \tau} = (3 - x_{n+2})(2x_1 + x_1 u) \\ \frac{\partial x_2}{\partial \tau} = (3 - x_{n+2})(-x_2 + 3x_2 u) \\ \frac{\partial p_1}{\partial \tau} = -(3 - x_{n+2})(2p_1 + p_1 u) \\ \frac{\partial p_2}{\partial \tau} = -(3 - x_{n+2})(-p_2 + 3p_2 u) \end{cases} \quad (8.6)$$

where $u = -x_1 p_1 - 3x_2 p_2$.

The boundary conditions are

$$\begin{aligned} x_1(0) = 1, & \quad x_2(0) = 1, & \quad p_1(3) = x_1(3) - e^2, \\ p_2(3) = x_2(3) - e^2, & \quad \frac{\partial x_1}{\partial x_{n+1}}(0) = 0, & \quad \frac{\partial x_2}{\partial x_{n+1}}(0) = 0, \\ \frac{\partial p_1}{\partial x_{n+1}}(3) = \frac{\partial x_1}{\partial x_{n+1}}(3), & \quad \frac{\partial p_2}{\partial x_{n+1}}(3) = \frac{\partial x_2}{\partial x_{n+1}}(3), & \quad \frac{\partial x_1}{\partial x_{n+2}}(0) = 0, \\ \frac{\partial x_2}{\partial x_{n+2}}(0) = 0, & \quad \frac{\partial p_1}{\partial x_{n+2}}(3) = \frac{\partial x_1}{\partial x_{n+2}}(3), & \quad \frac{\partial p_2}{\partial x_{n+2}}(3) = \frac{\partial x_2}{\partial x_{n+2}}(3). \end{aligned} \quad (8.7)$$

In the expressions of the above boundary conditions, in order to keep the notation simple, we omit the arguments x_{n+1}, x_{n+2} for the functions. For example, we write $x_1(0, x_{n+1}, x_{n+2})$ simply as

$x_1(0)$. The values of $\frac{\partial J_1}{\partial t_1}$ and $\frac{\partial J_1}{\partial t_2}$ can be obtained from

$$\begin{aligned} \frac{\partial J_1}{\partial x_{n+1}} &= (x_1(3) - e^2) \frac{\partial x_1(3)}{\partial x_{n+1}} + (x_2(3) - e^2) \frac{\partial x_2(3)}{\partial x_{n+1}} \\ &+ \int_0^1 \left(\frac{1}{2} u^2 + x_{n+1} u \frac{\partial u}{\partial x_{n+1}} \right) d\tau + \int_1^2 \left[-\frac{1}{2} u^2 + (x_{n+2} - x_{n+1}) u \frac{\partial u}{\partial x_{n+1}} \right] d\tau \\ &+ \int_2^3 \left[(3 - x_{n+1}) u \frac{\partial u}{\partial x_{n+1}} \right] d\tau, \end{aligned} \quad (8.8)$$

$$\begin{aligned} \frac{\partial J_1}{\partial x_{n+2}} &= (x_1(3) - e^2) \frac{\partial x_1(3)}{\partial x_{n+2}} + (x_1(3) - e^2) \frac{\partial x_1(3)}{\partial x_{n+2}} \\ &+ \int_0^1 \left(x_{n+1} u \frac{\partial u}{\partial x_{n+2}} \right) d\tau + \int_1^2 \left[\frac{1}{2} u^2 + (x_{n+2} - x_{n+1}) u \frac{\partial u}{\partial x_{n+2}} \right] d\tau \\ &+ \int_2^3 \left[-\frac{1}{2} u^2 + (3 - x_{n+1}) u \frac{\partial u}{\partial x_{n+2}} \right] d\tau. \end{aligned} \quad (8.9)$$

Choose initial nominal values $t_1 = 1.1$ and $t_2 = 2.1$. By applying the conceptual Algorithm 4.1 with the gradient projection method, after 18 iterations we find that the optimal switching instants are $t_1 = 1.0050$, $t_2 = 1.9993$ and the corresponding optimal cost is 2.7599×10^{-6} . After translating the result into the form suitable for the original problem, the corresponding continuous control and state trajectory are shown in Figure 2 (a) and (b). Note that the theoretical optimal solutions for this problem are $t_1^{opt} = 1$, $t_2^{opt} = 2$, $u^{opt} \equiv 0$ and $J^{opt} = 0$, so the result we obtained is quite accurate. Also note that the result for this problem obtained using the method in this paper is more accurate than that reported in [28, 24]. \square

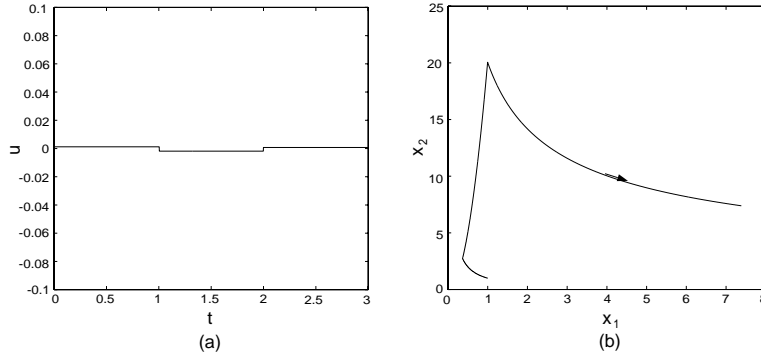


Figure 2: Example 8.1: (a) The control input. (b) The state trajectory.

Example 8.2 Consider a switched system consisting of

$$\text{subsystem 1: } \dot{x} = \begin{bmatrix} 0.6 & 1.2 \\ -0.8 & 3.4 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u, \quad (8.10)$$

$$\text{subsystem 2: } \dot{x} = \begin{bmatrix} 4 & 3 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 2 \\ -1 \end{bmatrix} u. \quad (8.11)$$

Assume that $t_0 = 0$, $t_f = 2$ and the system switches once at $t = t_1$ ($0 \leq t_1 \leq 2$) 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

$$J = \frac{1}{2}(x_1(2) - 4)^2 + \frac{1}{2}(x_2(2) - 2)^2 + \frac{1}{2} \int_0^2 (x_2(t) - 2)^2 + u^2(t) dt \quad (8.12)$$

is minimized. Here $x(0) = [0, 2]^T$.

We use the method in Section 7 to obtain the value of $\frac{dJ}{dt_1}$. From an initial nominal $t_1 = 1.0$, by using the the conceptual Algorithm 4.1 with the gradient projection method, after 12 iterations we find that the optimal switching instant is $t_1 = 0.1897$ and the corresponding optimal cost is 9.7667. The optimal continuous control input is in the state feedback form $u(\tau) = -K(\tau, 0.1897)x(\tau) - E(\tau, 0.1897)$ for the equivalent problem. We can then translate the result into the form suitable for the original problem. The corresponding continuous control and state trajectory are shown in Figure 3 (a) and (b). Figure 4 shows the optimal cost for different t_1 's. \square

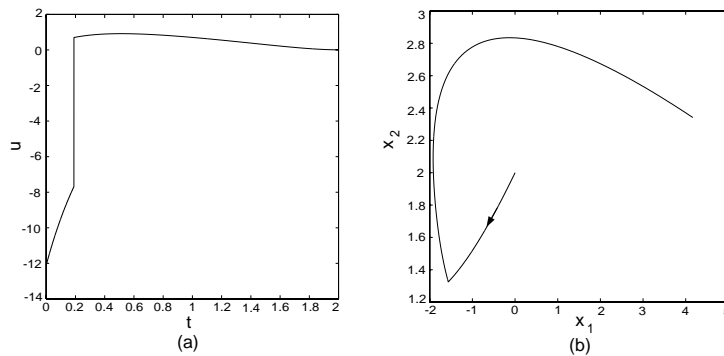


Figure 3: Example 8.2: (a) The control input. (b) The state trajectory.

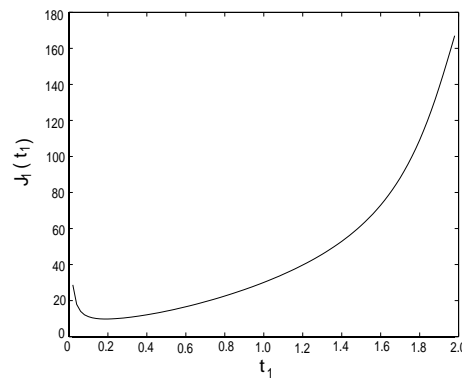


Figure 4: The optimal cost for Example 8.2 for different t_1 's.

As remarked in Remark 7.1, the approach we develop in Section 7 is also applicable to systems with several subsystems and more than one switchings following the same idea except for more notations. The following example illustrates such an application.

Example 8.3 Consider a switched system consisting of

$$\text{subsystem 1: } \dot{x} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \quad (8.13)$$

$$\text{subsystem 2: } \dot{x} = \begin{bmatrix} 0.5 & 5.3 \\ -5.3 & 0.5 \end{bmatrix} x + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u, \quad (8.14)$$

$$\text{subsystem 3: } \dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1.5 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u. \quad (8.15)$$

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 ($0 \leq t_1 \leq t_2 \leq 3$). We want to find optimal switching instants t_1, t_2 and an optimal input u such that the cost functional $J = \frac{1}{2} \int_0^3 u^2(t) dt$ is minimized. Here $x(0) = [4, 4]^T$ and $x(3)$ is required to be close to $[-4.1437, 9.3569]^T$.

For this problem, we adjoin a penalty term $[(x_1(3) + 4.1437)^2 + (x_2(3) - 9.3569)^2]$ to J and then consider the expanded cost functional J_{exp} . We use the method in Section 7 to obtain the values of $\frac{\partial J_1}{\partial t_1}$ and $\frac{\partial J_1}{\partial t_2}$. From initial nominal values $t_1 = 0.8$, $t_2 = 1.8$, by using the the conceptual Algorithm 4.1 with the gradient projection method, after 20 iterations we find that the optimal switching instant is $t_1 = 0.9982$, $t_2 = 1.9983$ and the corresponding optimal cost is 4.4087×10^{-5} . The optimal continuous control input is in the state feedback form $u(\tau) = -K(\tau, 0.9982, 1.9983)x(\tau) - E(\tau, 0.9982, 1.9983)$ for the equivalent problem. We can then translate the result into the form suitable for the original problem. The corresponding continuous control and state trajectory are shown in Figure 5 (a) and (b). Note that the theoretical optimal solutions for this problem are $t_1^{opt} = 1$, $t_2^{opt} = 2$, $u^{opt} \equiv 0$ and $J_{exp}^{opt} = 0$, so the result we obtained is quite accurate. Figure 6 shows the optimal cost for different $t_1 < t_2$. \square

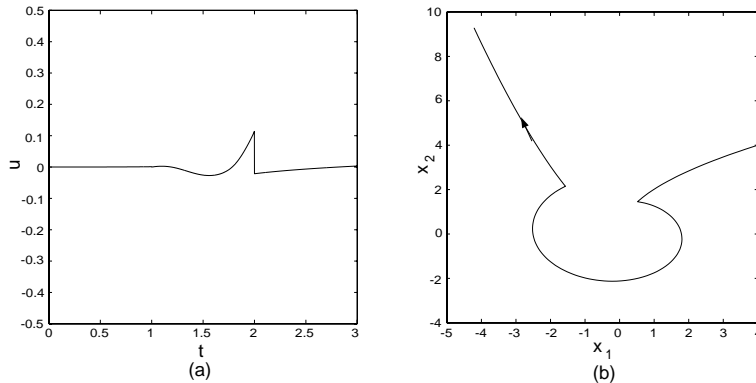


Figure 5: Example 8.3: (a) The control input. (b) The state trajectory.

It can be observed from Figure 6 that the function $J_1(t_1, t_2)$ has several ripples. Hence it is not convex even for this simple GSLQ problem; that is why such problems pose significant difficulties.

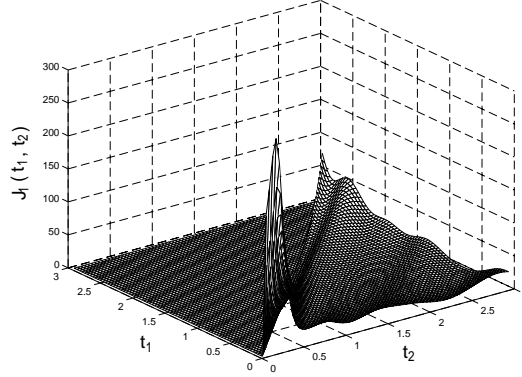


Figure 6: The optimal cost for Example 8.3 for different (t_1, t_2) 's.

9 Conclusion

In this paper, we first formulated an optimal control problem of switched systems and proposed a two stage optimization methodology for it. Then we focused on Stage 1 optimization problems which can further be decomposed into Stage 1(a) and Stage 1(b). We proposed an approach to obtain the accurate values of the derivatives that is necessary for Stage 1(b) optimizations. The method first transcribes a Stage 1 problem into an equivalent problem parameterized by 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 and their differentiations. In particular, a modified version of the method was proposed for GSLQ problems which only need to solve an initial value ODE. Note that earlier results of this paper have appeared in [26, 25, 27] and a more complete version can be found in [23]. Another earlier result by the authors, which obtains approximations of the derivatives, is reported in [28, 24]. However, note that the result in this paper is more accurate and straightforward than the one in [28, 24]. 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.

References

- [1] M. Athans and P. Falb. *Optimal Control*. McGraw-Hill, 1966.
- [2] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming Theory and Algorithms, Second Edition*. John Wiley & Sons, Inc., 1993.
- [3] D.P. Bertsekas. *Nonlinear Programming, Second Edition*. Athena Scientific, 1999.
- [4] M.S. Branicky, V.S. Borkar, and S.K. Mitter. A unified framework for hybrid control: model and optimal control theory. *IEEE Transactions on Automatic Control*, 43(1):31–45, January 1998.

- [5] M.S. Branicky, R. Hebbbar, and G. Zhang. A fast marching algorithm for hybrid systems. In *Proceedings of the 38th IEEE Conference on Decision and Control*, pages 4897–4902, 1999.
- [6] M.S. Branicky, T.A. Johansen, I. Petersen, and E. Frazzoli. On-line techniques for behavioral programming. In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 1840–1845, Sydney, Australia, December 2000.
- [7] A.E. Bryson and Y.-C. Ho. *Applied Optimal Control: Optimization, Estimation, and Control*. Hemisphere Washington, DC, 1975.
- [8] I. Capuzzo Dolcetta and L.C. Evans. Optimal switching for ordinary differential equations. *SIAM Journal of Control and Optimization*, 22(1):143–161, January 1984.
- [9] S. Hedlund and A. Rantzer. Optimal control of hybrid system. In *Proceedings of the 38th IEEE Conference on Decision and Control*, pages 3972–3977, Phoenix, AZ, December 1999.
- [10] M. Johansson. *Piecewise linear control systems*. PhD thesis, Lund Institute of Technology, Sweden, 1999.
- [11] F.L Lewis. *Optimal Control*. Wiley Interscience, 1986.
- [12] B. Lincoln and B.M. Bernhardsson. Efficient pruning of search trees in lqr control of switched linear systems. In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 1828–1833, Sydney, Australia, December 2000.
- [13] J. Lu, L. Liao, A. Nerode, and J.H. Taylor. Optimal control of systems with continuous and discrete states. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 2292–2297, San Antonio, TX, December 1993.
- [14] A.S. Morse, editor. *Control Using Logic-Based Switching*, volume 222 of *Lecture Notes in Control and Information Sciences*. Springer, 1997.
- [15] S.G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, 1996.
- [16] B. Piccoli. Hybrid systems and optimal control. In *Proceedings of the 37th IEEE Conference on Decision and Control*, pages 13–18, Tempa, FL, December 1998.
- [17] P. Riedinger, F. Kratz, C. Iung, and C. Zanne. Linear quadratic optimization for hybrid systems. In *Proceedings of the 38th IEEE Conference on Decision and Control*, pages 3059–3064, Phoenix, AZ, December 1999.
- [18] P. Riedinger, C. Zanne, and F. Kratz. Time optimal control of hybrid systems. In *Proceedings of the 1999 American Control Conference*, pages 2466–2470, San Diego, CA, June 1999.
- [19] T.I. Seidman. Optimal control for switching systems. In *Proceedings of the 21st Annual Conference on Information Sciences and Systems*, pages 485–489, The Johns Hopkins University, Baltimore, Maryland, December 1987.
- [20] H.J. Sussmann. A maximum principle for hybrid optimal control problems. In *Proceedings of the 38th IEEE Conference on Decision and Control*, Phoenix, AZ, December 1999.

- [21] L.Y. Wang, A. Beydoun, J. Cook, J. Sun, and I. Kolmanovsky. Optimal hybrid control with applications to automotive powertrain systems. In [14], pages 190–200, 1997.
- [22] H.S. Witsenhausen. A class of hybrid-state continuous-time dynamic systems. *IEEE Transactions on Automatic Control*, 11(2):161–167, April 1966.
- [23] X. Xu. *Analysis and design of switched systems*. PhD thesis, University of Notre Dame, July 2001.
- [24] X. Xu and P. J. Antsaklis. Optimal control of switched systems via nonlinear optimization based on direct differentiations of value functions. ISIS Technical Report ISIS-2001-005, University of Notre Dame, August 2001. (<http://www.nd.edu/~isis/tech.html>).
- [25] X. Xu and P.J. Antsaklis. A dynamic programming approach for optimal control of switched systems. In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 1822–1827, Sydney, Australia, December 2000.
- [26] X. Xu and P.J. Antsaklis. Optimal control of switched systems: new results and open problems. In *Proceedings of the 2000 American Control Conference*, pages 2683–2687, 2000.
- [27] X. Xu and P.J. Antsaklis. An approach for solving general switched linear quadratic optimal control problems. In *Proceedings the 40th IEEE Conference on Decision and Control*, 2001. to appear.
- [28] X. Xu and P.J. Antsaklis. Optimal control of switched systems via nonlinear optimization based on direct differentiations of value functions. *International Journal of Control*, 2001. submitted.
- [29] J. Yong. Systems governed by ordinary differential equations with continuous, switching and impulse controls. *Appl. Math. Optim.*, 20:223–235, 1989.