15th Triennial World Congress, Barcelona, Spain

AN APPROACH TO SWITCHED SYSTEMS OPTIMAL CONTROL BASED ON PARAMETERIZATION OF THE SWITCHING INSTANTS ⋆

Xuping Xu * **Panos J. Antsaklis** **

* *Corresponding author. Penn State Erie, ECE Dept., 5091 Station Road, Erie, PA 16563 USA*
Tel: 1-814-898-7169; E-mail: Xuping-Xu@psu.edu
** *Dept. of EE, Univ. of Notre Dame, Notre Dame, IN 46556 USA*
Tel: 1-574-631-5792; E-mail: antsaklis.1@nd.edu

Abstract: This paper provides a viable general approach to switched systems optimal control. Such optimal control problems require the solutions of not only optimal continuous inputs but also optimal switching sequences. Many practical problems only involve optimization where the number of switchings and the sequence of active subsystems are given. This is stage 1 of the two stage optimization methodology proposed by the authors in previous papers. In order to solve stage 1 problems, the derivatives of the optimal cost with respect to the switching instants need to be known. In (Xu and Antsaklis, 2001), we proposed an approach for solving a special class of such problems, namely, general switched linear quadratic problems. In this paper, the idea of (Xu and Antsaklis, 2001) is extended to general switched systems optimal control problems and an approach is proposed for solving them. The approach 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 formed by the state, costate, stationarity equations, the boundary and continuity conditions and their differentiations. Examples are shown to illustrate the results in the paper.

Keywords: Switching systems; Optimal control; Hybrid systems; Control synthesis

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 electrical circuit systems, etc., can be modeled as switched systems.

Optimal control problems for switched systems have attracted the attention of researchers recently. For such a problem, one needs to find both an optimal continuous input and an optimal switching sequence since the system dynamics vary before and after every switching instant. The available results in the literature on such problems can be classified as theoretical (e.g., (Branicky *et al.*, 1998; Piccoli, 1998)) and practical (see e.g., (Gokbayrak and Cassandras, 2000; Hedlund and Rantzer, 1999; Lu *et al.*, 1993; Wang *et al.*, 1997)). Most of the practical methods that we are aware of are using numerical methods and 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 problem to find optimal/suboptimal solutions. But the discretization of time space may lead to computational combinatoric explosion and the solutions obtained may not be accurate enough. In view of this, in some previous papers by the authors

⋆ The partial support of the National Science Foundation (NSF ECS99-12458 & CCR01-13131), and of the DARPA/ITO-NEST Program (AF-F30602-01-2-0526) is gratefully acknowledged.

(see (Xu and Antsaklis, 2000; Xu and Antsaklis, 2001)), approaches not based on discretization of continuous time space were explored.

In this paper, we further explore approaches not based on discretizations. Since many practical problems only involve optimization where the number of switchings and the sequence of active subsystems are given (yet the switching instants are unknown), we focus on such problems. This is stage 1 of the two stage optimization methodology proposed by the authors in previous papers (Xu and Antsaklis, 2000; Xu and Antsaklis, 2001). In order to solve stage 1 problems, the derivatives of the optimal cost with respect to the switching instants need to be known. In (Xu and Antsaklis, 2001), we proposed an approach for solving a special class of such problems, namely, general switched linear quadratic problems. In this paper, the approach is extended to general switched systems optimal control problems. The approach 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.

2. PROBLEM FORMULATION

2.1 *Switched Systems*

The switched systems we shall consider in this paper are defined as follows.

Definition 1. (Switched System). A switched system is a tuple $\mathcal{S} = (\mathcal{D}, \mathcal{F})$ where

• $\mathcal{D} = (I, E)$ is a directed graph indicating the discrete structure of the system. The node set $I = \{1, 2, \cdots, M\}$ is the set of indices for subsystems. The directed edge set E is a subset of $I \times I - \{(i,i)|i \in I\}$ containing all valid events. If an event $e = (i_1, i_2)$ takes place, the system switches from subsystem i_1 to i_2.

• $\mathcal{F} = \{f_i : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^n |\ i \in I\}$ with f_i describing the vector field for the i-th subsystem $\dot{x} = f_i(x, u, t)$. □

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

Definition 2. (Switching Sequence). For a switched system $\mathcal{S}$, a switching sequence σ in $[t_0, t_f]$ is defined as

$$\sigma = ((t_0, i_0), (t_1, e_1), (t_2, e_2), \cdots, (t_K, e_K)), \quad (1)$$

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

A switching sequence σ as defined above indicates that subsystem i_k is active in $[t_k, t_{k+1})$. We specify $\sigma \in \Sigma_{[t_0,t_f]}$ as a discrete input to the system (see (Xu and Antsaklis, 2000) for more details).

2.2 *An Optimal Control Problem*

Now we formulate the optimal control problem we will study in this paper. In the sequel, we denote $\mathcal{U}_{[t_0,t_f]} \triangleq \{u | u \in C_p[t_0, t_f], u(t) \in \mathbb{R}^m\}$; i.e., the set of all piecewise continuous functions for $t \in [t_0, t_f]$ with values in $\mathbb{R}^m$.

Problem 1. Consider a switched system $\mathcal{S}$. Given a fixed interval $[t_0, t_f]$, find $u \in \mathcal{U}_{[t_0,t_f]}$ and a switching sequence $\sigma \in \Sigma_{[t_0,t_f]}$ such that $x(t)$ departs from a given $x(t_0) = x_0$ and meets $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \to \mathbb{R}^{l_f}\}$ at t_f and the cost functional

$$J = \psi(x(t_f)) + \int_{t_0}^{t_f} L(x(t), u(t), t)\ dt \quad (2)$$

is minimized. □

In the sequel, we assume that f, L, ϕ_f and ψ 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. 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 (Xu, 2001)).

3. TWO STAGE OPTIMIZATION

Now we review the two stage algorithm (see (Xu and Antsaklis, 2001)) in the following.

Algorithm 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., $J_1 = J_1(t_1, t_2, \cdots, t_K)$ for $K \geq 0$ $(t_0 \leq t_1 \leq \cdots \leq t_K \leq t_f)$. Find J_1.
(b). Minimize J_1 with respect to $t_1, t_2, \cdots, t_K$.

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 the optimal control problem. □

Algorithm 1 has high computational costs. In the followings, we concentrate on stage 1 optimization. Note that many practical problems are in fact stage 1 problems. For example, the speeding-up of a power train in an automobile only requires switchings from gear 1 to 2 to 3 to 4. As can be seen from Algorithm 1, stage 1 can further be decomposed into two sub-steps (a) and (b) (note that a similar hierarchical decomposition method can be found in (Gokbayrak and Cassandras, 2000)).

Stage 1(a)

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 $\sigma = ((t_0, i_0), (t_1, e_1), \cdots, (t_K, e_K))$. We denote the corresponding optimal cost as a function $J_1(\hat{t})$, where $\hat{t} \triangleq (t_1, t_2, \cdots, t_K)^T$. In stage 1(a), we need to find an optimal u and the corresponding minimum J. For stage 1(a), although different subsystems are active in different time intervals, the problem is conventional since these intervals are fixed. It is not difficult to use the calculus of variations techniques to prove the following necessary conditions.

Theorem 1. (Necessary Conditions for Stage 1(a)). Consider the stage 1(a) problem for Problem 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 $x(t)$ departs from a given initial $x(t_0) = x_0$ and meets $S_f = \{x | \phi_f(x) = 0, \phi_f : \mathbb{R}^n \to \mathbb{R}^{l_f}\}$ at t_f. In order that u be optimal, it is necessary that there exists a vector function $p(t) = [p_1(t), \cdots, p_n(t)]^T$, $t \in [t_0, t_f]$, such that the following conditions hold

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

$$\text{State eq: } \frac{dx(t)}{dt} = [\frac{\partial H}{\partial p}(x(t), p(t), u(t), t)]^T \quad (1)$$

$$\text{Costate eq:} \frac{dp(t)}{dt} = -[\frac{\partial H}{\partial x}(x(t), p(t), u(t), t)]^T, (2)$$

$H(x, p, u, t) \triangleq L(x, u, t) + p^T f_k(x, u, t)$, 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

$$0 = [\frac{\partial H}{\partial u}(x(t), p(t), u(t), t)]^T. \quad (3)$$

(c). At t_f, the function p satisfies

$$p(t_f) = [\frac{\partial \psi}{\partial x}(x(t_f))]^T + [\frac{\partial \phi_f}{\partial x}(x(t_f))]^T \lambda \quad (4)$$

where λ is an l_f-dimensional vector.

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

$$p(t_k-) = p(t_k+). \quad (5)$$

Proof: See Chapter 6 of (Xu, 2001). □

The above necessary conditions will be used in Section 5 in the development of a method for finding $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$. In general, it is difficult or even impossible to find an analytical expression of $J_1(\hat{t})$ using them. However, we can find the numerical solutions by solving the two point boundary value differential algebraic equation (DAE) formed by conditions (a)-(d) using numerical methods.

Stage 1(b)

Stage 1(b) is in essence a constrained nonlinear optimization problem with simple constraints

$$\min_{\hat{t}} J_1(\hat{t}) \qquad \text{subject to } \hat{t} \in T \quad (6)$$

where $T \triangleq \{\hat{t} = (t_1, t_2, \cdots, t_K)^T | t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_K \leq t_f\}$. Feasible direction methods can be applied to such problems. These methods use $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$. In our computations, we use the gradient projection method (using $\frac{\partial J_1}{\partial \hat{t}}$) and the constrained Newton's method (using $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$) (see Section 2.3 in Bertsekas (Bertsekas, 1999) for details).

A Conceptual Algorithm

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

Algorithm 2. **(A Conceptual Algorithm for Stage 1)**

(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 $\frac{\partial J_1}{\partial \hat{t}}(\hat{t}^j)$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}(\hat{t}^j)$.

(4). Use the gradient projection method or the constrained Newton's method (if $\frac{\partial^2 J_1}{\partial \hat{t}^2}(\hat{t}^j)$ is known) to update $\hat{t}^j$ to be $\hat{t}^{j+1} = \hat{t}^j + \alpha^j d\hat{t}^j$ (here α^j is chosen using the Armijo's rule (Bertsekas, 1999)). Set $j = j+1$.

(5). Repeat Steps (2), (3), (4) and (5), until a prespecified termination condition is satisfied. □

The key elements of the above algorithm are

(a). An optimal control algorithm for Step (2).

(b). The derivations of $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$ 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 2 because $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$ are not readily available. It is the task of the subsequent sections to address (b) and devise a method for deriving the values of $\frac{\partial J_1}{\partial \hat{t}}$ and $\frac{\partial^2 J_1}{\partial \hat{t}^2}$.

4. AN EQUIVALENT PROBLEM FORMULATION

In this section we transcribe a stage 1 problem into an equivalent conventional optimal control problem parameterized by the switching instants, which will be used in next section. For simplicity of notation, in the followings, we concentrate on the case of two subsystems where subsystem 1 is active for $t \in [0, t_1)$ and subsystem 2 is active for $t \in [t_1, t_f]$ (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). We consider the following stage 1 problem.

Problem 2. For a switched system

$$\dot{x} = f_1(x,u,t),\ 0 \le t < t_1, \quad (1)$$

$$\dot{x} = f_2(x,u,t),\ t_1 \le t \le t_f, \quad (2)$$

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

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

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

We transcribe Problem 2 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

$$\frac{dx_{n+1}}{dt} = 0 \quad (4)$$

$$x_{n+1}(0) = t_1 \quad (5)$$

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

$$t(\tau) = \begin{cases} t_0 + (x_{n+1} - t_0)\tau, & 0 \le \tau \le 1 \\ x_{n+1} + (t_f - x_{n+1})(\tau - 1), & 1 \le \tau \le 2. \end{cases} \quad (6)$$

By introducing x_{n+1} and τ, Problem 2 can be transcribed into the following equivalent problem.

Problem 3. (An Equivalent Problem). For a system with dynamics

$$\frac{dx(\tau)}{d\tau} = (x_{n+1} - t_0) f_1(x,u,t(\tau)) \quad (7)$$

$$\frac{dx_{n+1}}{d\tau} = 0 \quad (8)$$

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

$$\frac{dx(\tau)}{d\tau} = (t_f - x_{n+1}) f_2(x,u,t(\tau)) \quad (9)$$

$$\frac{dx_{n+1}}{d\tau} = 0 \quad (10)$$

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

$$J = \psi(x(2)) + \int_0^1 (x_{n+1} - t_0) L(x,u,t(\tau))\ d\tau + \int_1^2 (t_f - x_{n+1}) L(x,u,t(\tau))\ d\tau \quad (11)$$

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

Remark 1. 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 variables as in (6) and by regarding $x_{n+1} = t_1$, and vice versa. □

Remark 2. Problem 3 is conventional because it has fixed time instant when the system dynamics change. In fact, because x_{n+1} is an unknown constant throughout $\tau \in [0,2]$, it can be regarded as a conventional optimal control problem with an unknown parameter x_{n+1}. In the sequel, we regard Problem 3 as an optimal control problem parameterized by the switching instant x_{n+1} with cost (11) and subsystems (7) and (9). □

5. THE DEVELOPMENT OF THE APPROACH

In this section, based on the equivalent Problem 3, we develop a method for deriving accurate numerical value of $\frac{dJ_1}{dt_1}$. 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 3 and their differentiations with respect to the parameter x_{n+1}. In the following, 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 3, define

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

$$\tilde{f}_2(x,u,\tau,x_{n+1}) \triangleq (t_f - x_{n+1}) f_2(x,u,t(\tau)), \quad (2)$$

$$\tilde{L}_1(x,u,\tau,x_{n+1}) \triangleq (x_{n+1} - t_0) L(x,u,t(\tau)), \quad (3)$$

$$\tilde{L}_2(x,u,\tau,x_{n+1}) \triangleq (t_f - x_{n+1}) L(x,u,t(\tau)). \quad (4)$$

Regarding x_{n+1} as a parameter, we denote the optimal state trajectory as $x(\tau, x_{n+1})$. We define the parameterized Hamiltonian as

$$H(x,p,u,\tau,x_{n+1}) \triangleq \begin{cases} \tilde{L}_1(x,u,\tau,x_{n+1}) + p^T \tilde{f}_1(x,u,\tau,x_{n+1}), \\ \quad \text{for } \tau \in [0,1), \\ \tilde{L}_2(x,u,\tau,x_{n+1}) + p^T \tilde{f}_2(x,u,\tau,x_{n+1}), \\ \quad \text{for } \tau \in [1,2]. \end{cases} \quad (5)$$

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 equations

$$\text{State eq: } \frac{\partial x}{\partial \tau} = (\frac{\partial H}{\partial p})^T = \tilde{f}_1(x,u,\tau,x_{n+1}) \quad (6)$$

$$\text{Costate eq: } \frac{\partial p}{\partial \tau} = -(\frac{\partial H}{\partial x})^T = -(\frac{\partial \tilde{f}_1}{\partial x})^T p - (\frac{\partial \tilde{L}_1}{\partial x})^T \quad (7)$$

$$\text{Stationarity eq: } 0 = (\frac{\partial H}{\partial u})^T = (\frac{\partial \tilde{f}_1}{\partial u})^T p + (\frac{\partial \tilde{L}_1}{\partial u})^T \quad (8)$$

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

$$\text{State eq: } \frac{\partial x}{\partial \tau} = (\frac{\partial H}{\partial p})^T = \tilde{f}_2(x,u,\tau,x_{n+1}) \quad (9)$$

$$\text{Costate eq: } \frac{\partial p}{\partial \tau} = -(\frac{\partial H}{\partial x})^T = -(\frac{\partial \tilde{f}_2}{\partial x})^T p - (\frac{\partial \tilde{L}_2}{\partial x})^T \quad (10)$$

$$\text{Stationarity eq: } 0 = (\frac{\partial H}{\partial u})^T = (\frac{\partial \tilde{f}_2}{\partial u})^T p + (\frac{\partial \tilde{L}_2}{\partial u})^T \quad (11)$$

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

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

$$x(0, x_{n+1}) = x_0, \tag{12}$$

$$p(2, x_{n+1}) = (\frac{\partial \psi}{\partial x}(x(2, x_{n+1})))^T. \tag{13}$$

The necessary condition (d) tells us that

$$p(1-, x_{n+1}) = p(1+, x_{n+1}). \tag{14}$$

(6)-(8), (9)-(11) along with (12) and (13) form a two point boundary value differential algebraic equation (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 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}

$$J_1(x_{n+1}) = \psi(x(2, x_{n+1})) + \int_0^1 \tilde{L}_1(x, u, \tau, x_{n+1})\, d\tau + \int_1^2 \tilde{L}_2(x, u, \tau, x_{n+1})\, d\tau. \tag{15}$$

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

$$\begin{aligned} \frac{dJ_1}{dx_{n+1}} &= \frac{\partial \psi(x(2, x_{n+1}))}{\partial x} \frac{\partial x(2, x_{n+1})}{\partial x_{n+1}} + \int_0^1 [L(x, u, t(\tau)) \\ &+ (x_{n+1} - t_0)(\frac{\partial L}{\partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial L}{\partial u}\frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial L}{\partial t})]\, d\tau \\ &+ \int_1^2 [-L(x, u, t(\tau)) + (t_f - x_{n+1})(\frac{\partial L}{\partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial L}{\partial u}\frac{\partial u}{\partial x_{n+1}} \\ &+ (2-\tau)\frac{\partial L}{\partial t})]\, d\tau. \end{aligned} \tag{16}$$

So we need to obtain the functions $\frac{\partial x(\tau, x_{n+1})}{\partial x_{n+1}}$ 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 $\frac{dJ_1}{dx_{n+1}}$. By differentiating (6)-(8) and (9)-(11) with respect to x_{n+1}, we obtain

$$\begin{aligned} \frac{\partial}{\partial \tau}(\frac{\partial x}{\partial x_{n+1}}) &= \frac{\partial}{\partial x_{n+1}}(\frac{\partial x}{\partial \tau}) = f_1 + (x_{n+1} - \\ & -t_0)(\frac{\partial f_1}{\partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial f_1}{\partial u}\frac{\partial u}{\partial x_{n+1}} + \tau\frac{\partial f_1}{\partial t}), \end{aligned} \tag{17}$$

$$\begin{aligned} \frac{\partial}{\partial \tau}(\frac{\partial p}{\partial x_{n+1}}) &= -\frac{\partial}{\partial x_{n+1}}(\frac{\partial p}{\partial \tau}) = -(\frac{\partial f_1}{\partial x})^T p - (\frac{\partial L}{\partial x})^T - (x_{n+1} \\ & -t_0)[(\frac{\partial f_1}{\partial x})^T \frac{\partial p}{\partial x_{n+1}} + (p^T \frac{\partial^2 f_1}{\partial x^2}\frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_1}{\partial x \partial u}\frac{\partial u}{\partial x_{n+1}})^T \\ & +\tau(p^T \frac{\partial^2 f_1}{\partial x \partial t})^T \frac{\partial^2 L}{\partial x^2}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial x \partial u}\frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial^2 L}{\partial x \partial t}], \end{aligned} \tag{18}$$

$$\begin{aligned} 0 &= (\frac{\partial f_1}{\partial u})^T p + (\frac{\partial L}{\partial u})^T + (x_{n+1} - t_0)[(\frac{\partial f_1}{\partial u})^T \frac{\partial p}{\partial x_{n+1}} \\ & +(p^T \frac{\partial^2 f_1}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_1}{\partial u^2}\frac{\partial u}{\partial x_{n+1}})^T + \tau(p^T \frac{\partial^2 f_1}{\partial u \partial t})^T \\ & + \frac{\partial^2 L}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial u^2}\frac{\partial u}{\partial x_{n+1}} + \tau \frac{\partial^2 L}{\partial u \partial t}], \end{aligned} \tag{19}$$

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

$$\begin{aligned} \frac{\partial}{\partial \tau}(\frac{\partial x}{\partial x_{n+1}}) &= \frac{\partial}{\partial x_{n+1}}(\frac{\partial x}{\partial \tau}) = -f_2 + (t_f \\ & -x_{n+1})(\frac{\partial f_2}{\partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial f_2}{\partial u}\frac{\partial u}{\partial x_{n+1}} + (2-\tau)\frac{\partial f_2}{\partial t}), \end{aligned} \tag{20}$$

$$\begin{aligned} \frac{\partial}{\partial \tau}(\frac{\partial p}{\partial x_{n+1}}) &= -\frac{\partial}{\partial x_{n+1}}(\frac{\partial p}{\partial \tau}) = (\frac{\partial f_2}{\partial x})^T p + (\frac{\partial L}{\partial x})^T - (t_f \\ & -x_{n+1})[(\frac{\partial f_2}{\partial x})^T \frac{\partial p}{\partial x_{n+1}} + (p^T \frac{\partial^2 f_2}{\partial x^2}\frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_2}{\partial x \partial u}\frac{\partial u}{\partial x_{n+1}})^T \\ & +(2-\tau)(p^T \frac{\partial^2 f_2}{\partial x \partial t})^T + \frac{\partial^2 L}{\partial x^2}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial x \partial u}\frac{\partial u}{\partial x_{n+1}} \\ & +(2-\tau)\frac{\partial^2 L}{\partial x \partial t}], \end{aligned} \tag{21}$$

$$\begin{aligned} 0 &= -(\frac{\partial f_2}{\partial u})^T p - (\frac{\partial L}{\partial u})^T + (x_{n+1} - t_0)[(\frac{\partial f_2}{\partial u})^T \frac{\partial p}{\partial x_{n+1}} \\ & +(p^T \frac{\partial^2 f_2}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}})^T + (p^T \frac{\partial^2 f_2}{\partial u^2}\frac{\partial u}{\partial x_{n+1}})^T + (2-\tau)(p^T \frac{\partial^2 f_2}{\partial u \partial t})^T \\ & + \frac{\partial^2 L}{\partial u \partial x}\frac{\partial x}{\partial x_{n+1}} + \frac{\partial^2 L}{\partial u^2}\frac{\partial u}{\partial x_{n+1}} + (2-\tau)\frac{\partial^2 L}{\partial u \partial t}], \end{aligned} \tag{22}$$

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. Other notations can be interpreted similarly (see (Xu, 2001) for details).

Differentiating (12), (13) and (14) with respect to x_{n+1}, we obtain

$$\frac{\partial x(0, x_{n+1})}{\partial x_{n+1}} = 0, \tag{23}$$

$$\frac{\partial p(2, x_{n+1})}{\partial x_{n+1}} = \frac{\partial^2 \psi(x(2, x_{n+1}))}{\partial x^2}\frac{\partial x(2, x_{n+1})}{\partial x_{n+1}}, \tag{24}$$

$$\frac{\partial p(1-, x_{n+1})}{\partial x_{n+1}} = \frac{\partial p(1+, x_{n+1})}{\partial x_{n+1}}. \tag{25}$$

It can now be observed that (6)-(8), (9)-(11) and (17)-(19), (20)-(22) along with the boundary conditions (12), (13) and (23), (24) and with the continuity conditions (14), (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 substituting the result into (16), we can obtain $\frac{dJ_1}{dx_{n+1}}$.

Remark 3. In general, we need to resort to numerical methods (e.g., shooting methods) to find the solution to the two point boundary value DAE. In particular, if all subsystems are linear in control and the cost function L is quadratic in control, then the two point boundary value DAE can hence be reduced to a two point boundary value differential equation which can be solved more easily. See Chapter 8 of (Xu, 2001) for details. □

Remark 4. 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 $\frac{d^2 J_1(t_1)}{dt_1^2}$ can also be similarly obtained. See Chapter 8 of (Xu, 2001) for details. □

6. SOME EXAMPLES

Example 1. Consider a switched system with

$$\text{subsystem 1: } \dot{x} = x + 2xu, \quad (1)$$

$$\text{subsystem 2: } \dot{x} = -x - 3xu. \quad (2)$$

Assume $t_0 = 0$, $t_f = 2$ and the system switches at $t = t_1$ ($0 \leq t_1 \leq 2$) from subsystem 1 to 2. Find an optimal switching instant t_1 and an optimal input u such that $J = \frac{1}{2}(x(2)-1)^2 + \frac{1}{2}\int_0^2 u^2(t)\, dt$ is minimized. Here $x(0) = 1$.

The method in Section 5 is used to obtain $\frac{dJ_1}{dt_1}$. Choose an initial nominal $t_1 = 1.2$. By applying Algorithm 2 with the gradient projection method, after 20 iterations we find the optimal t_1 to be $t_1 = 0.9994$ and the corresponding cost to be 1.1848×10^{-7}. Figure 1 (a) and (b) show the continuous input and the state trajectory. Note that the theoretical solutions are $t_1^{opt} = 1$, $u^{opt} \equiv 0$ and $J^{opt} = 0$. □

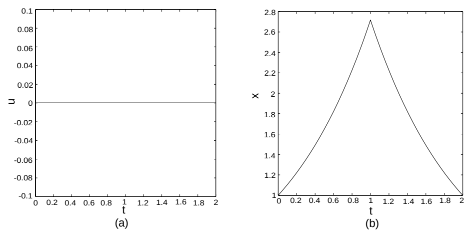

Fig. 1. Example 1: (a) The control input. (b) The state trajectory $x(t)$.

Example 2. Consider a switched system with

$$\text{subsystem 1: } \begin{cases} \dot{x}_1 = -x_1 + 2x_1 u \\ \dot{x}_2 = x_2 + x_2 u \end{cases} \quad (3)$$

$$\text{subsystem 2: } \begin{cases} \dot{x}_1 = x_1 - 3x_1 u \\ \dot{x}_2 = 2x_2 - 2x_2 u \end{cases} \quad (4)$$

$$\text{subsystem 3: } \begin{cases} \dot{x}_1 = 2x_1 + x_1 u \\ \dot{x}_2 = -x_2 + 3x_2 u \end{cases} \quad (5)$$

Assume $t_0 = 0$, $t_f = 3$ and the system switches at $t = t_1$ from subsystem 1 to 2 and at $t = t_2$ from subsystem 2 to 3 ($0 \leq t_1 \leq t_2 \leq 3$). Find optimal switching instants t_1, t_2 and an optimal input u such that $J = \frac{1}{2}(x_1(3) - e^2)^2 + \frac{1}{2}(x_2(3) - e^2)^2 + \frac{1}{2}\int_0^3 u^2(t)\, dt$ is minimized. Here $x_1(0) = 1$ and $x_2(0) = 1$.

The method in Section 5 is used to obtain $\frac{\partial J_1}{\partial t_1}$ and $\frac{\partial J_1}{\partial t_2}$. We choose initial nominal $t_1 = 1.1$ and $t_2 = 2.1$. By applying Algorithm 2 with the gradient projection method, after 18 iterations we find that the optimal t_1 and t_2 to be $t_1 = 1.0050$, $t_2 = 1.9993$ and the corresponding cost to be 2.7599×10^{-6}. The corresponding continuous control and state trajectory are shown in Figure 2 (a) and (b) show the continuous input and the state trajectory. Note that the theoretical solutions 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. □

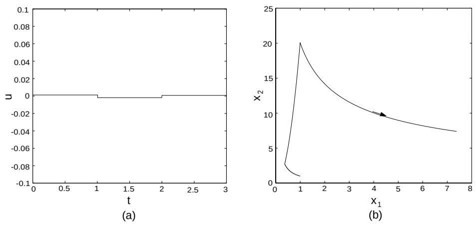

Fig. 2. Example 2: (a) The control input. (b) The state trajectory.

7. CONCLUSION

In this paper, a general approach to switched systems optimal control is developed. It is mainly developed in Sections 4 and 5 and is applicable to problems with many subsystems and more than one switchings. The approach is based on solving a two point boundary value DAE derived in Section 5. Derivatives of the optimal cost with respect to the switching instants can be obtained accurately and therefore nonlinear optimization algorithms can be used to find the optimal switching instants. Future research topics include the extension of the approach to systems with internally-forced switchings and systems with state discontinuities.

8. REFERENCES

Bertsekas, D.P. (1999). *Nonlinear Programming, Second Edition.* Athena Scientific.

Branicky, M.S., V.S. Borkar and S.K. Mitter (1998). A unified framework for hybrid control: model and optimal control theory. *IEEE Trans. on AC* **43**(1), 31–45.

Gokbayrak, K. and C.G. Cassandras (2000). A hierarchical decomposition method for optimal control of hybrid systems. In: *Proc. 39th CDC.* pp. 1816–1821.

Hedlund, S. and A. Rantzer (1999). Optimal control of hybrid system. In: *Proc. 38th CDC.* pp. 3972–3977.

Lu, J., L. Liao, A. Nerode and J.H. Taylor (1993). Optimal control of systems with continuous and discrete states. In: *Proc. 32nd CDC.* pp. 2292–2297.

Piccoli, B. (1998). Hybrid systems and optimal control. In: *Proc. 37th CDC.* pp. 13–18.

Wang, L.Y., A. Beydoun, J. Cook, J. Sun and I. Kolmanovsky (1997). Optimal hybrid control with applications to automotive powertrain systems. In: *LNCIS 222.* Springer. pp. 190–200.

Xu, X. (2001). Analysis and design of switched systems. PhD thesis. Univ. of Notre Dame. (`http://www.nd.edu/~isis/oap.html`).

Xu, X. and P.J. Antsaklis (2000). Optimal control of switched systems: new results and open problems. In: *Proc. 2000 ACC.* pp. 2683–2687.

Xu, X. and P.J. Antsaklis (2001). An approach for solving general switched linear quadratic optimal control problems. In: *Proc. 40th CDC.* pp. 2478–2483.