OPTIMAL CONTROL OF SWITCHED SYSTEMS WITH MULTIPLE TIME-DELAYS AND A COST ON CHANGING CONTROL

ZHAOHUA GONG AND CHONGYANG LIU*
School of Mathematics and Information Science
Shandong Institute of Business and Technology
Yantai 264005, China

YIJING WANG
Department of Mathematics and Statistics
Curtin University
Perth 6845, Australia

(Communicated by Kok Lay Teo)

Abstract. In this paper, we consider a class of optimal switching control problems with multiple time-delays and a cost on changing control and subject to terminal state constraints. A computational method involving three stages is developed to solve this class of optimal control problems. First, by parameterizing the control function with piecewise-constant functions, the optimal switching control problem is approximated by a sequence of finite-dimensional optimization problems, where the original switching times, the control heights and the control switching times are decision variables. Second, by introducing new variables, the total variation of the control variables is transformed into an equivalently smooth function. Third, we convert the constrained optimization problem into one only with box constraints by an exact penalty function method. The gradients of the cost functional are then derived, which can be combined with any gradient-based optimization method to determine the optimal solution. Finally, a numerical example is given to illustrate the effectiveness of the proposed algorithm.

1. Introduction. Switched systems are encountered in many real life applications, such as locomotives [5], bio-chemical reactors [9], and hybrid power systems [24]. On the other hands, the existence of time delays must not be ignored in many practical engineering problems [17]. As a consequence, optimal control of switched systems with or without delays has become an important and challenging research topic for many applied mathematicians and engineers [2, 18, 26]. For conventionally optimal control of switched system, its aim is to determine an optimal control function and an optimal switching sequence of the switched systems involved such that a cost functional is minimized subject to constraints on the state and/or the control. However, the presence of delays in a switched system complicates the search for an optimal operation policy. In particular, the time-scaling transformation [8, 12, 26],

2010 Mathematics Subject Classification. Primary: 49J21, 49M37; Secondary: 34K34.
Key words and phrases. Switched system, time-delay system, optimal control, nonlinear optimization, total variation.
* Corresponding author: Chongyang Liu.
a powerful tool for solving optimal switching control problems, cannot be directly applied to switched systems with time-delays [7]. In fact, optimization techniques for switched systems with time-delays are scarce in the literature. Necessary conditions for determining optimal switching times and/or optimal impulse magnitudes for such systems are derived in [4, 21, 22] via classical variational techniques. However, these analytical results are only applicable to separable systems with a single delay. Reference [25] presents an algorithm for optimization of switched systems with time-delays, which is based on a parameterization scheme in which the switching times are expressed in terms of the subsystem durations. However, this algorithm is only applicable to switched systems with a single delay and no control input. Recently, an effective computational method is developed to find the optimal switching times and system parameters for a general switched system with multiple time-delays in [11]. However, the control input is not considered as a decision variable in [11].

In this paper, we will consider a new class of optimal control problems, where the dynamical system is governed by a switched system with multiple time-delays and there is a cost on changing control. In a standard optimal control problem, the cost functional is usually expressed in terms of the final state and/or an integral term involving the state and control values at each time. The cost of changing the control input is rarely considered. Thus, two different control laws that result in the same cost value are deemed to perform equally even if one of them is constant while the other fluctuates widely. However, in reality, there will be a cost when the control input changes from one value to another. The cost may be just wear and tear on the system actuators [14]. Some papers are now available in the literature that consider delay-free optimal control problems with costs on changing control input. For example, necessary optimality conditions are derived in [3] for an optimal control problem in which the control input can only assume two possible values and there is a cost associated with changing from one value to another. In [15], further optimality conditions for similar problems (but more general), in which the total variation of the control input is incorporated as a penalty in the cost functional, is investigated. In [20], a computational algorithm is developed to solve a class of optimal control problems in which the cost functional is the sum of the terminal cost, the integral cost and the full variation of control. Recently, in [14], a new computational method based on the control parameterization [19], time-scaling transformation [12] and a smoothing technique is proposed to solve similar problems as in [20]. However, these computational methods are not applicable to optimal switching control problems with multiple time-delays.

The aim of this paper is to propose a computational method, which is applicable to a class of optimal control problems with the following characteristics:

(i). The dynamical system is a switched system with variable switching times;
(ii). The dynamical system is with multiple time-delays;
(iii). The cost functional contains a total variation measuring the changes of control inputs; and
(iv). The state variables are required to satisfy the terminal equality constraints.

We first convert the optimal switching control problem into a sequence of finite-dimensional optimization problems by parameterizing the control input with piecewise-constant functions. In view of the non-smoothness of the cost functional, the smoothing technique [14] is then used to transform the resulting optimization problems into equivalently smooth constrained optimization problems. The exact
penalty method [27] is further utilized to transcribe the constrained optimization problems into optimization problems only with box constraints. For these optimization problems, a computational method is developed based on the gradients of the penalized cost functional with respect to decision variables. Finally, a numerical example is given to illustrate the effectiveness of the proposed computational method.

The rest of this paper is organized as follows. The problem formulation is described in Section 2. The transformation procedure, including the control parameterization technique, the smoothing technique and the exact penalty method, is introduced in Section 3. The gradient formulas are derived in Section 4. A numerical example is given in Section 5. Section 6 concludes the paper.

2. Problem formulation. Consider the following switched time-delay system with \( N \) subsystems, and \( m \) time-delays:

\[
\begin{align*}
\dot{x}(t) &= f_i(t, x(t), x(t-\alpha_1), \ldots, x(t-\alpha_m), u(t)), \quad t \in (\tau_{i-1}, \tau_i], \\
x(t) &= \phi_1, \quad t \leq 0,
\end{align*}
\]

where \( t \) is the time; \( x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^\top \in \mathbb{R}^n \) is the state vector; \( u(t) = (u_1(t), \ldots, u_r(t))^\top \in \mathbb{R}^r \) is the control vector; \( \tau_0 = 0 \) is the initial time; \( \tau_N = T > 0 \) is a given terminal time; \( \tau_i, \ i = 1, 2, \ldots, N - 1 \), are switching times; \( \alpha_j, \ j = 1, 2, \ldots, m \), are given time-delays; and \( f^i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{nm} \times \mathbb{R}^r \rightarrow \mathbb{R}^n, i = 1, 2, \ldots, N \), and \( \phi : \mathbb{R} \rightarrow \mathbb{R}^n \) are given functions.

System (1) is controlled by manipulating the switching times and the control. Define

\[
T := \{(\tau_1, \tau_2, \ldots, \tau_{N-1})^\top \in \mathbb{R}^{N-1} : \tau_i - \tau_{i-1} \geq \Delta_i, \ i = 1, 2, \ldots, N\},
\]

where \( \Delta_i > 0 \) is the minimum duration of the \( i \)th subsystem. Any \( \tau \in T \) is called a feasible switching time vector.

Define

\[
U := \{(u_1, u_2, \ldots, u_r)^\top \in \mathbb{R}^r : a_h \leq u_h \leq b_h, h = 1, 2, \ldots, r\},
\]

where \( a_h \) and \( b_h \) are given real numbers such that \( a_h \leq b_h, h = 1, 2, \ldots, r \). It is obvious that \( U \) is a compact and convex subset of \( \mathbb{R}^r \). The total variation of the control \( u : [0, T] \rightarrow \mathbb{R}^r \) is defined by

\[
\bigvee_0^T u = \sum_{h=1}^r \bigvee_0^T u_h,
\]

where \( u_h : [0, T] \rightarrow \mathbb{R} \) is the \( h \)th component of the control \( u \) and \( \bigvee_0^T u_h \) denotes the total variation of \( u_h \) defined by

\[
\bigvee_0^T u_h = \sup \sum_{i=1}^p |u_h(t_i) - u_h(t_{i-1})|.
\]

Here, the supremum is taken over all finite partitions \( \{t_i\}_{i=0}^p \subset [0, T] \) satisfying \( 0 = t_0 < t_1 < \cdots < t_{p-1} < t_p = T \). If the total variation of \( u \) is finite, then we say that \( u \) is of bounded variation. Let \( u : [0, T] \rightarrow U \) be a function of bounded variation over \([0, T]\), and let \( U \) be the set of all such control functions.
We assume that the following conditions are satisfied throughout this paper.

Assumption 1. The functions \( f_i, i = 1, 2, \ldots, N \), are continuously differentiable. Moreover, the function \( \varphi \) is twice continuously differentiable.

Assumption 2. For each \( i = 1, 2, \ldots, N \), there exists a positive real number \( L_1 > 0 \) such that for all \( t \in [0, T] \), \( y_j^i \in \mathbb{R}^n, j = 0, \ldots, m \), and \( v \in \mathbb{R}^r \),
\[
\| f^i(t, y_0^i, y_1^i, \ldots, y_m^i, v) \| \leq L_1 (1 + \| y_0^i \| + \| y_1^i \| + \cdots + \| y_m^i \|),
\]
where \( \| \cdot \| \) is the Euclidean norm.

Assumptions 1 and 2 ensure that the switched system (1) has a unique solution \( x(\cdot|\tau, u) \) corresponding to each \( \tau \in \mathcal{T} \) and \( u \in \mathcal{U} \).

Now, we suppose that system (1) is subjected to the following terminal state constraints:
\[
\varphi_l(x(T|\tau, u)) = 0, \quad l = 1, 2, \ldots, q,
\]
where \( \varphi_l : \mathbb{R}^n \to \mathbb{R}, l = 1, 2, \ldots, q \), are given continuously differentiable functions.

In real practical situations, large fluctuations of the control are clearly undesirable. Thus, the cost functional is expressed as a weighted sum of the function of the terminal state and the total variation of the control, that is, the cost functional can be defined as
\[
J(\tau, u) := \psi(x(T|\tau, u)) + \gamma \int_0^T u^T \, dt,
\]
where \( \psi : \mathbb{R}^n \to \mathbb{R} \) is a given continuously differentiable function and \( \gamma \in [0, 1] \) is a weighting factor. Note that we can easily transform an integral running cost into the form of (7) by introducing an additional state variable.

Our object is to choose a switching time vector \( \tau \in \mathcal{T} \) and a control \( u \in \mathcal{U} \) such that the cost functional (7) is minimized subject to the terminal constraints (6). Thus, our optimal switching control problem is stated formally as follows.

**Problem (P).** Find a switching time vector \( \tau \in \mathcal{T} \) and a control \( u \in \mathcal{U} \) such that the cost functional (7) is minimized subject to the switched time-delay system (1) and the terminal constraints (6).

Problem (P) is an optimal control problem involving nonlinear switched system with multiple time-delays and a cost on changing control. It is well known that variable switching times pose a significant challenge for conventional numerical optimization techniques [12, 14]. One of the most widely used methods for overcoming this challenge is the time-scaling transformation described in [6, 13, 14]. This transformation involves mapping the variable switching times to fixed points in a new time horizon. Unfortunately, since the time-scaling transformation is not applicable to delay systems [7], it cannot be used to solve Problem (P). Therefore, a new approach is needed to solve Problem (P).

3. Problem transformation.

3.1. Control parameterization. To solve Problem (P) numerically, we apply the control parameterization scheme [19]. That is, let each of the time subintervals \((\tau_{i-1}, \tau_i), i = 1, 2, \ldots, N\), be partitioned into \( p_i \) subintervals with \( p_i - 1 \) partition points, i.e.,
\[
\tau_1^i, \tau_2^i, \ldots, \tau_{p_i-1}^i.
\]
Then, the control is approximated as a piecewise-constant function defined by

\[
    u^p(t) = \sum_{i=1}^{N} \sum_{k=1}^{P_i} \sigma^{i,k} \chi(\tau_{i-1}^k, \tau_i^k)(t), \quad t \in (\tau_{i-1}, \tau_i], \quad i = 1, 2, \ldots, N,
\]

where \( \tau_0^i = \tau_{i-1}^i \); \( \tau_p^i = \tau_i^i \); \( \sigma^{i,k} \) is the \( k \)th control height vector of the \( i \)th switch interval, \( i = 1, 2, \ldots, N \), \( k = 1, 2, \ldots, P_i \); and for a given interval \( I \), \( \chi_I \) denotes the indicator function defined by

\[
    \chi_I(t) = \begin{cases} 
    1, & \text{if } t \in I, \\
    0, & \text{otherwise.} 
    \end{cases}
\]  

Furthermore, the partition times satisfy the following constraints:

\[
    \tau_{i-1} = \tau_0^i \leq \tau_1^i \leq \cdots \leq \tau_{p_i}^i = \tau_i \quad \text{and} \quad \tau_{p_i}^i - \tau_{p_i-1}^i \geq \Delta_i > 0, \quad i = 1, 2, \ldots, N,
\]

where \( p_0 = 0 \) and \( \Delta_i \) is as defined in (2). Thus, by the control parameterization, the system (1) becomes

\[
\begin{align*}
    \dot{x}(t) &= f^i(t, x(t), x(t-\alpha_1), \ldots, x(t-\alpha_m), \sum_{k=1}^{p_i} \sigma^{i,k} \chi(\tau_{i-1}^k, \tau_i^k)(t)), \\
    x(t) &= \varphi(t), \quad t \leq 0.
\end{align*}
\]  

Now, denote \( \sum_{i=1}^{N} P_i \) by \( \tilde{N} \), and let \( \nu := (\nu_1, \nu_2, \ldots, \nu_{N-1})^\top \in \mathbb{R}^{\tilde{N}-1} \) such that

\[
    \tau_0^i = \nu_0 \leq \tau_1^i = \nu_1 \leq \tau_2^i = \nu_2 \leq \cdots \leq \tau_{p_i}^i = \nu_{p_i} \\
    \leq \tau_{p_i+1}^i = \nu_{p_i+1} \leq \tau_{p_i+2}^i = \nu_{p_i+2} \leq \cdots \leq \tau_p^i = \nu_{p-1} \leq T,
\]

and

\[
    \nu_{p_0+\cdots+p_i} - \nu_{p_0+\cdots+p_{i-1}} \geq \Delta_i > 0, \quad i = 1, \ldots, N.
\]

Furthermore, let

\[
    \sigma := ((\sigma^1)^\top, \ldots, (\sigma^{N,1})^\top, \ldots, (\sigma^{N,1})^\top, \ldots, (\sigma^{N,1})^\top, \ldots, (\sigma^{N,N})^\top)^\top \in \mathbb{R}^{\tilde{N}}.
\]

The approximated control can be rewritten as

\[
    \tilde{u}^p(t) = \sum_{j=1}^{\tilde{N}} \sigma^j \chi(\nu_{j-1}, \nu_j)(t).
\]

Define

\[
    \ell(j) = i, \quad \text{if} \quad \sum_{m=0}^{i-1} p_m < j \leq \sum_{m=0}^{i} p_m, \quad i = 1, 2, \ldots, N.
\]

Then, the switched system (10) can be written as

\[
\begin{align*}
    \dot{x}(t) &= f^{\ell(j)}(t, x(t), x(t-\alpha_1), \ldots, x(t-\alpha_m), \sigma^j), \quad t \in (v_{j-1}, v_j], \\
    x(t) &= \varphi(t), \quad t \leq 0.
\end{align*}
\]  

Let \( \Gamma \) be the set of all those switching vectors \( \nu \) such that conditions (11) and (12) are satisfied, \( \Omega \) be the set of all those control parameter vectors \( \sigma = \)
Then, we can state a new optimization problem as given below:

\( \phi_l(\mathbf{x}^p(T|\mathbf{v}, \mathbf{\sigma})) = 0, \quad l = 1, 2, \ldots, q, \tag{16} \)

and the cost functional becomes

\[ J(\mathbf{v}, \mathbf{\sigma}) = \psi(\mathbf{x}^p(T|\mathbf{v}, \mathbf{\sigma})) + \gamma \sum_{h=1}^{r} \sum_{j=1}^{N-1} |\sigma^{j+1}_h - \sigma^j_h|. \tag{17} \]

As a result, Problem (P) can be approximated by the following finite-dimensional optimization problem:

Problem (P(p)). Find a feasible pair \((\mathbf{v}, \mathbf{\sigma}) \in \Gamma \times \Omega\) such that the cost functional

\[ J(\mathbf{v}, \mathbf{\sigma}) = \psi(\mathbf{x}^p(T|\mathbf{v}, \mathbf{\sigma})) + \gamma \sum_{h=1}^{r} \sum_{j=1}^{N-1} |\sigma^{j+1}_h - \sigma^j_h| \]

is minimized subject to switched time-delay system (15) and terminal constraints (16).

3.2. An equivalent problem. Since the cost functional (17) is non-smooth, the standard gradient-based optimization algorithms cannot be used directly to solve Problem (P(p)). In what follows, we shall show that Problem (P(p)) can be converted into an equivalently smooth optimization problem.

Let

\[ \mathbf{\zeta} := (\mathbf{\beta}^\top, (\mathbf{c}^1)^\top, \ldots, (\mathbf{c}^{N-1})^\top, (\mathbf{\xi}^1)^\top, \ldots, (\mathbf{\xi}^{N-1})^\top)^\top \in \mathbb{R}^{(2N-1)r}, \tag{18} \]

where \( \mathbf{\beta}, \mathbf{\xi}^j, \mathbf{\xi}^j \) are all vectors in \( \mathbb{R}^r \). Define the function

\[ \varrho^j(\mathbf{\zeta}) = \begin{cases} \mathbf{\beta} + \sum_{l=j}^{N-1} (\mathbf{c}^l - \mathbf{\xi}^l), & \text{if } j = 1, 2, \ldots, N-1, \\ \mathbf{\beta}, & \text{if } j = N. \end{cases} \]

Furthermore, let \( \mathcal{Z} \) be the set of \( \mathbf{\zeta} \in \mathbb{R}^{(2N-1)r} \) in the form of (18) such that for each \( h = 1, 2, \ldots, r \),

\[ a_h \leq \varrho^j_h(\mathbf{\zeta}) \leq b_h, \quad j = 1, 2, \ldots, N; \]

and

\[ \varsigma^j_h \geq 0, \quad \xi^j_h \geq 0, \quad j = 1, 2, \ldots, N-1. \]

Next, we consider the following dynamic system:

\[ \begin{cases} \dot{y}(t) = f^{(j)}(t, y(t), y(t - \alpha_1), \ldots, y(t - \alpha_m), \varrho^j(\mathbf{\zeta})), & t \in (v_{j-1}, v_j], \\ y(t) = \varphi(t), & t \leq 0. \end{cases} \tag{19} \]

Let \( \mathbf{y}(\cdot|\mathbf{v}, \mathbf{\zeta}) \) be the unique solution of the system (19) corresponding to \((\mathbf{v}, \mathbf{\zeta}) \in \Gamma \times \mathcal{Z}\) and let the terminal constraints be defined by

\[ \phi_l(\mathbf{y}(T|\mathbf{v}, \mathbf{\zeta})) = 0, \quad l = 1, 2, \ldots, q. \tag{20} \]

Then, we can state a new optimization problem as given below:
Problem (Q). Find a feasible pair \((\nu, \zeta) \in \Gamma \times \mathcal{Z}\) such that the cost functional
\[
\tilde{J}(\nu, \zeta) = \psi(y(T|\nu, \zeta)) + \gamma \sum_{h=1}^{\bar{N}} \sum_{j=1}^{\nu_h} (\varsigma_{j}^h + \xi_{j}^h)
\]
is minimized subject to the system (19) and the terminal constraints (20).

Note that Problem (Q) is a smooth optimization problem. For Problems (P(p)) and (Q), we have the following theorem.

**Theorem 3.1.** For each \((\nu, \zeta) \in \Gamma \times \mathcal{Z}\), it holds that
\[
y(t|\nu, \zeta) = x^p(t|\nu, \tilde{\sigma}(\zeta)), \quad t \in (-\infty, T],
\]
where
\[
\tilde{\sigma}(\zeta) = ((\varphi^1(\zeta))^\top, (\varphi^2(\zeta))^\top, \ldots, (\varphi^{\bar{N}}(\zeta))^\top)^\top.
\]

**Proof.** For notation simplicity, let
\[
\tilde{x}(\cdot) = x^p(\cdot|\nu, \tilde{\sigma}(\zeta)).
\]
By system (15), it is obvious that
\[
\tilde{x}(t) = \varphi(t) = y(t|\nu, \zeta), \quad t \in (-\infty, 0].
\]
By differentiating \(\tilde{x}(t)\) and using the chain rule, we obtain that for \(t \in (\nu_{j-1}, \nu_j)\), \(j = 1, 2, \ldots, \bar{N}\),
\[
\dot{\tilde{x}}(t) = f^{l(j)}(t, \tilde{x}(t), \tilde{x}(t - \alpha_1), \ldots, \tilde{x}(t - \alpha_m), \tilde{\sigma}(\zeta))
= f^{l(j)}(t, \tilde{x}(t), \tilde{x}(t - \alpha_1), \ldots, \tilde{x}(t - \alpha_m), \varphi^j(\zeta)).
\]
This implies that \(\tilde{x}(\cdot)\) is the solution of system (19). It follows that \(\tilde{x}(\cdot)\) is the unique solution of system (19), as required.

Based on Theorems 3.1 and 3.2 and by similar arguments as given in [14], we can obtain the following result.

**Theorem 3.3.** Problems (P(p)) and (Q) are equivalent.

By Theorem 3.3, we see that the optimal solution of Problem (P(p)) can be generated by that of Problem (Q).
3.3. A penalty problem. Problem (Q) is a smooth dynamic optimization problem with terminal state constraints. In this subsection, the exact penalty method [27] will be applied to these terminal state constraints.

Define

\[ \Upsilon := \{(v_1, v, \ldots, v_{N-1})^T \in \mathbb{R}^{N-1} : 0 \leq v_i \leq T, i = 1, 2, \ldots, N-1 \}, \]

and

\[ \mathcal{Z}_0 := \{(\beta^T, (\zeta^1)^T, \ldots, (\zeta^{N-1})^T, (\xi^1)^T, \ldots, (\xi^{N-1})^T)^T \in \mathbb{R}^{(2N-1)r} : a_h \leq \beta_h \leq b_h, \zeta_h^j, \xi_h^j \geq 0, j = 1, 2, \ldots, N-1, h = 1, 2, \ldots, r \}, \]

It is obvious that \( \Gamma \subset \Upsilon \times \mathcal{Z}_0 \). Now, we define a constraint violation function on \( \Upsilon \times \mathcal{Z}_0 \) as

\[
\Pi(v, \zeta) = \sum_{i=1}^{N} \max\{v_{p_0+p_1+\ldots+p_i} - v_{p_0+p_1+\ldots+p_{i-1}} - \Delta_i, 0\}^2 \\
+ \sum_{i=1}^{N} \sum_{r=1}^{i} \sum_{\nu=1}^{\nu} \max\{a_{r} - \beta_{r} - \sum_{\nu=1}^{N-1} (\zeta_{r}^\nu - \xi_{r}^\nu), 0\}^2 \\
+ \sum_{i=1}^{N} \sum_{r=1}^{i} \sum_{\nu=1}^{\nu} \max\{\beta_{r} + \sum_{\nu=1}^{N-1} (\zeta_{r}^\nu - \xi_{r}^\nu) - b_{r}, 0\}^2. \tag{22}
\]

Then, \( \Pi(v, \zeta) = 0 \) if and only if \( (v, \zeta) \in \Upsilon \times \mathcal{Z}_0 \) is feasible for Problem (Q). On this basis, we define a penalty function \( J_p(v, \zeta) \) as follows:

\[
J_p(v, \zeta, \epsilon) = J(v, \zeta) + \epsilon^{-\delta_1} \Pi(v, \zeta) + \rho \epsilon^{\delta_2}, \tag{23}
\]

where \( \epsilon > 0 \) is a new decision variable, \( \rho > 0 \) is a penalty parameter, and \( \delta_1 \) and \( \delta_2 \) are fixed parameters satisfying \( 1 \leq \delta_2 \leq \delta_1 \).

With the penalty cost functional (23), we define the following penalty problem.

**Problem (Q).** Find a feasible pair \( (v, \zeta, \epsilon) \in \Upsilon \times \mathcal{Z}_0 \times (0, \bar{\epsilon}] \) such that the cost functional (23) is minimized subject to the switched system (19), \( \bar{\epsilon} \) is a given small positive constant.

The following convergence results can be obtained based on the result in [27].

**Theorem 3.4.** Let \( (v^*, \zeta^*, \epsilon^*) \) be a local optimal solution of Problem (Q). Then, \( (v^*, \zeta^*) \) is a local optimal solution of Problem (Q) if and only if \( \epsilon^* = 0 \).

**Theorem 3.5.** Let \( \{\rho_k\}_{k=1}^{\infty} \) be an increasing sequence of penalty parameters such that \( \rho_k \to +\infty \) as \( k \to +\infty \), and let \( (v^{k*,}, \zeta^{k*}, \epsilon^{k*}) \) be the solution of Problem (Q) with \( \rho = \rho_k \). Then, the limit point of the sequence \( \{(v^{k*,}, \zeta^{k*})\} \) is a solution of Problem (Q).

From the above two theorems, we can solve Problem (Q) where only the box constraints are involved. Furthermore, the solution of Problem (Q) with \( \rho \) being sufficiently large is the optimal solution of Problem (Q). In the next section, we will derive the gradient formulas of \( J_p(v, \zeta, \epsilon) \) with respect to the decision variables.
4. Gradient formulas. Problem (\(\bar{Q}\)) is an optimization problem involving a non-linear switched time-delay system, where the switching time vector \(\nu\), parameter vector \(\zeta\) and variable \(\epsilon\) are decision variables to be optimized. Note that gradient-based optimization algorithms, e.g., sequential quadratic programming (SQP)[16], are effective in solving such optimization problems [19]. For this, the gradients of the cost functional (23) with respect to the decision variables are needed. To this end, define
\[
\dot{\mathbf{y}}(t) := (\mathbf{y}(t - \alpha_1)^T, \mathbf{y}(t - \alpha_2)^T, \ldots, \mathbf{y}(t - \alpha_m)^T)^T, \\
\tilde{f}^{(j)}(t|\mathbf{v}, \zeta) := f^{(j)}(t, \mathbf{y}(t|\mathbf{v}, \zeta), \dot{\mathbf{y}}(t|\mathbf{v}, \zeta), \vartheta^j(\zeta)),
\]
and for given \(i\) and \(j\),
\[
\delta_{ij} := \begin{cases} 
1, & \text{if } i = j, \\
0, & \text{otherwise.}
\end{cases}
\]
For each \((\mathbf{v}, \zeta, \epsilon) \in \mathcal{T} \times \mathcal{Z}_0 \times [0, \bar{c}]\), consider the following costate system:
\[
\dot{\lambda}(t) = -\sum_{\kappa=1}^{\bar{N}} \mathbf{y}(t)\mathbf{z}^{(\kappa)}(t|\mathbf{v}, \zeta)|_{(\nu_{\kappa-1}, \nu_{\kappa})}(t) - \sum_{\kappa=1}^{\bar{N}} \sum_{j=1}^{m} \lambda(t + \alpha_j) \frac{\partial f^{(\kappa)}(t + \alpha_j|\mathbf{v}, \zeta)}{\partial \mathbf{y}}|_{(\nu_{\kappa-1}, \nu_{\kappa} - \alpha_j)}(t), \quad t \in [0, T],
\]
with the conditions
\[
\lambda(t) = 0, \quad t > T, \\
\lambda(T) = \frac{\partial \psi(y(T))}{\partial \mathbf{y}} + 2\epsilon \delta_1 \sum_{l=1}^{q} \phi_l(y(T)) \frac{\partial \phi_l(y(T))}{\partial \mathbf{y}}.
\]
Let \(\lambda(\cdot|\mathbf{v}, \zeta)\) denote the solution of system (25)-(27) corresponding to the given pair \((\mathbf{v}, \zeta, \epsilon) \in \mathcal{T} \times \mathcal{Z}_0 \times [0, \bar{c}]\). The gradients of the cost functional (23) with respect to \(\zeta\) and \(\epsilon\) are given in the following theorem.

**Theorem 4.1.** Let \((\mathbf{v}, \zeta, \epsilon) \in \mathcal{T} \times \mathcal{Z}_0 \times [0, \bar{c}]\). Then,
\[
\frac{\partial J_\rho}{\partial \beta_h} = \int_0^T \sum_{\kappa=1}^{\bar{N}} \lambda(s)^T \frac{\partial f^{(\kappa)}(s|\mathbf{v}, \zeta)}{\partial \beta_h}|_{(\nu_{\kappa-1}, \nu_{\kappa})}(s)ds - 2\epsilon \delta_1 \sum_{i=1}^{\bar{N}-1} \max\{a_h - \beta_h - \sum_{\nu=1}^{\bar{N}} (s_h^\nu - \xi_h^\nu), 0\} + 2\epsilon \delta_1 \sum_{i=1}^{\bar{N}} \max\{\beta_h + \sum_{\nu=1}^{\bar{N}} (s_h^\nu - \xi_h^\nu) - b_h, 0\}, \\
\text{for } h = 1, 2, \ldots, r,
\]
\[
\frac{\partial J_\rho}{\partial \xi^j_h} = \int_0^{\nu_h} \lambda(s)^T \frac{\partial f^{(\kappa)}(s|\mathbf{v}, \zeta)}{\partial \xi^j_h}|_{(\nu_{\kappa-1}, \nu_{\kappa})}(s)ds + \gamma - 2\epsilon \delta_1 \sum_{i=1}^{\bar{N}-1} \max\{a_h - \beta_h - \sum_{\nu=1}^{\bar{N}} (s_h^\nu - \xi_h^\nu), 0\} + 2\epsilon \delta_1 \sum_{i=1}^{\bar{N}} \max\{\beta_h + \sum_{\nu=1}^{\bar{N}} (s_h^\nu - \xi_h^\nu) - b_h, 0\}, \\
\text{for } j = 1, 2, \ldots, \bar{N} - 1, \ h = 1, 2, \ldots, r,
\]
\[
\frac{\partial J_\rho}{\partial \xi^j_h} = \int_0^{\nu_h} \lambda(s)^T \frac{\partial f^{(\kappa)}(s|\mathbf{v}, \zeta)}{\partial \xi^j_h}|_{(\nu_{\kappa-1}, \nu_{\kappa})}(s)ds + \gamma + 2\epsilon \delta_1 \sum_{i=1}^{\bar{N}} \max\{a_h - \beta_h - \sum_{\nu=1}^{\bar{N}} (s_h^\nu - \xi_h^\nu), 0\} + 2\epsilon \delta_1 \sum_{i=1}^{\bar{N}} \max\{\beta_h + \sum_{\nu=1}^{\bar{N}} (s_h^\nu - \xi_h^\nu) - b_h, 0\}.
\]
and \( \Pi(\cdot, \zeta, \nu, \rho) \) is as defined in (22).

Proof. Let \( \omega : (0, +\infty) \to \mathbb{R}^n \) be an arbitrary function that is continuous and differentiable almost everywhere. The cost functional (23) can be rewritten as

\[
\tilde{J}_p(v, \zeta, \epsilon) = \psi(y(T)) + \gamma \sum_{h=1}^{N-1} \left( \zeta_h + \xi_h^T \right) + \epsilon^{-\delta_1} \Pi(v, \zeta, \nu, \rho) + \rho \epsilon^{\delta_2}
\]

\[
+ \sum_{k=1}^{\tilde{N}} \int_{t_{k-1}}^{t_k} \left( \omega(s)^T \tilde{f}^{(n)}(s) - \omega(s)^T \dot{y}(s) \right) ds,
\]

where the arguments \( v \) and \( \zeta \) in \( \Psi(\cdot, v, \zeta) \) and \( \tilde{f}^{(n)}(\cdot, v, \zeta) \) are omitted for simplicity; and \( \Pi(\cdot, \cdot) \) is as defined in (22). Then, by applying integration by parts, we have

\[
\tilde{J}_p(v, \zeta, \epsilon) = \psi(y(T)) + \gamma \sum_{h=1}^{N-1} \left( \zeta_h + \xi_h^T \right) + \epsilon^{-\delta_1} \Pi(v, \zeta, \nu, \rho) + \rho \epsilon^{\delta_2}
\]

\[
- \omega(T)^T y(T) + \omega(0)^T y(0)
\]

\[
+ \sum_{k=1}^{\tilde{N}} \int_{t_{k-1}}^{t_k} \left( \omega(s)^T \tilde{f}^{(n)}(s) + \dot{\omega}(s)^T y(s) \right) ds.
\]  

(32)

Differentiating (32) with respect to \( \beta_h \) gives

\[
\frac{\partial \tilde{J}_p(v, \zeta, \epsilon)}{\partial \beta_h} = \left( \frac{\partial \psi(y(T))}{\partial y} - \omega(T)^T + 2\epsilon^{-\delta_1} \sum_{l=1}^{q} \phi_l(y(T)) \frac{\partial \phi_l(y(T))}{\partial y} \right) \frac{\partial y(T)}{\partial \beta_h}
\]

\[
- 2\epsilon^{-\delta_1} \sum_{l=1}^{\tilde{N}-1} \max\{a_{\nu} - \beta_h - \sum_{\nu=1}^{\tilde{N}-1} (\zeta_h^\nu - \xi_h^\nu), 0\}
\]

\[
+ 2\epsilon^{-\delta_1} \sum_{l=1}^{\tilde{N}-1} \max\{\beta_h + \sum_{\nu=1}^{\tilde{N}-1} (\zeta_h^\nu - \xi_h^\nu) - b_h, 0\}
\]

\[
+ \sum_{k=1}^{\tilde{N}} \int_{t_{k-1}}^{t_k} \omega(s)^T \frac{\partial \tilde{f}^{(n)}(s)}{\partial \beta_h} ds
\]

\[
+ \sum_{k=1}^{\tilde{N}} \int_{t_{k-1}}^{t_k} \left( \omega(s)^T \frac{\partial \tilde{f}^{(n)}(s)}{\partial y} + \dot{\omega}(s)^T \frac{\partial y(s)}{\partial \beta_h} \right) ds
\]

\[
+ \sum_{k=1}^{\tilde{N}} \int_{t_{k-1}}^{t_k} \omega(s)^T \frac{\partial \tilde{f}^{(n)}(s)}{\partial y} \frac{\partial y(s - \alpha_k)}{\partial \beta_h} ds.
\]  

(33)

Performing a change of variable in the last term on the right-hand side of (33) gives
Since \( y(s) = \varphi(s) \) for \( s \leq 0 \), equation (34) can be rewritten as
\[
\sum_{\kappa=1}^{\bar{N}} \sum_{j=1}^{m} \int_{v_{\kappa-1}}^{v_{\kappa}} \omega(s) \partial_{\kappa} f_{(\kappa)}(s) \frac{\partial y(s - \alpha_j)}{\partial h} ds = \sum_{\kappa=1}^{\bar{N}} \sum_{j=1}^{m} \int_{v_{\kappa-1}}^{v_{\kappa} - \alpha_j} \omega(s + \alpha_j) \partial_{\kappa} f_{(\kappa)}(s + \alpha_j) \frac{\partial y(s)}{\partial h} ds, \tag{35}
\]
where \( \chi(j) \) is as defined in (8). Substituting (35) into (33) yields
\[
\frac{\partial \hat{J}_p(v, \zeta, \epsilon)}{\partial h} = \left( \frac{\partial \psi(y(T))}{\partial y} - \omega(T) \partial_{\kappa} f_{(\kappa)}(s) \frac{\partial \varphi(y(T))}{\partial y} \right) \frac{\partial y(T)}{\partial h}
- 2 \epsilon \delta_1 \sum_{i=1}^{\bar{N}-1} \max\{a_i - \beta_i - \sum_{\nu=1}^{\bar{N}-1} (s_i^\nu - \xi_i^\nu), 0\}
+ 2 \epsilon \delta_1 \sum_{i=1}^{\bar{N}-1} \max\{\beta_i + \sum_{\nu=1}^{\bar{N}-1} (s_i^\nu - \xi_i^\nu) - b_i, 0\}
+ \int_0^T \sum_{\kappa=1}^{\bar{N}} \omega(s) \partial_{\kappa} f_{(\kappa)}(s) \chi(v_{\kappa-1}, v_{\kappa}) ds
+ \int_0^T \left( \sum_{\kappa=1}^{\bar{N}} \omega(s) \partial_{\kappa} f_{(\kappa)}(s) \chi(v_{\kappa-1}, v_{\kappa}) + \dot{\omega}(s) \right) \frac{\partial y(s)}{\partial h} ds
+ \int_0^T \left( \sum_{\kappa=1}^{\bar{N}} \sum_{j=1}^{m} \omega(s + \alpha_j) \partial_{\kappa} f_{(\kappa)}(s + \alpha_j) \chi(v_{\kappa-1} - \alpha_j, v_{\kappa} - \alpha_j) \right) \frac{\partial y(s)}{\partial h} ds. \tag{36}
\]
Choosing \( \omega(\cdot) = \lambda(\cdot | v, \zeta) \) and substituting (25)-(27) into (36) gives equation (28). By similar methods, we can obtain the gradient formulas (29), (30) and (31). The proof is completed. \( \blacksquare \)

The following theorem gives the gradients of the cost functional (23) with respect to the switching times \( v \).

**Theorem 4.2.** Let \((v, \zeta, \epsilon) \in \mathcal{Y} \times \mathcal{Z}_0 \times [0, \bar{\epsilon}]\). Then,
\[
\frac{\partial \hat{J}_p(v, \zeta, \epsilon)}{\partial v_j} = \lambda(v_j)^\top \tilde{f}^{(j)}(v_j, \zeta) - \lambda(v_j)^\top \tilde{f}^{(j+1)}(v_j, \zeta)
\]
and

\[ 2\epsilon^{-\delta_i} \sum_{i=1}^{N} \delta_{(p_0+p_1+\ldots+p_i)} \max\{v_{p_0+p_1+\ldots+p_i} - v_{p_0+p_1+\ldots+p_i-1} - \Delta_i, 0\} \]

\[ -2\epsilon^{-\delta_i} \sum_{i=1}^{N} \delta_{(p_0+p_1+\ldots+p_i)} \max\{v_{p_0+p_1+\ldots+p_i+1} - v_{p_0+p_1+\ldots+p_i} - \Delta_{i+1}, 0\} \]

\[ -2\epsilon^{-\delta_i} \sum_{i=1}^{N} \delta_{ij} \max\{v_{i-1} - v_i, 0\} + 2\epsilon^{-\delta_i} \sum_{i=1}^{N-1} \delta_{ij} \max\{v_{i} - v_{i+1}, 0\}, \]

\[ j = 1, 2, \ldots, N-1, \]

where \( \lambda(\cdot) = \lambda(\cdot|v, \zeta); \) and \( \delta_{(p_0+p_1+\ldots+p_i)} \) and \( \delta_{ij} \) are as defined in (24).

**Proof.** The proof is similar to that given for Theorem 4.1.

On the basis of Theorems 4.1 and 4.2, we now propose the following algorithm for solving Problem \((\bar{Q})\).

**Step 1.** Choose the initial guess \((v^0, \zeta^0, \epsilon^0) \in \mathcal{Y} \times \mathcal{Z} \times [0, \bar{\epsilon}],\) the initial penalty parameter \(\rho^0,\) tolerance \(\epsilon_{\text{min}} > 0,\) and parameters \(\delta \geq 1, \delta_1 \geq \delta_2 > 1,\) and \(\delta_3 > 1.\) Set \(\rho := \rho^0,\) \((v, \zeta, \epsilon) := (v^0, \zeta^0, \epsilon^0).\)

**Step 2.** Solve system (19) from \(t = 0\) to \(t = T\) to obtain \(y(\cdot|v, \zeta).\)

**Step 3.** Using \(y(T|v, \zeta),\) compute \(\bar{y}_p(v, \zeta, \epsilon)\) by (23).

**Step 4.** Using \(y(\cdot|v, \zeta),\) solve the costate systems (25)-(26) from \(t = T\) to \(t = 0\) to obtain \(\lambda(\cdot|v, \zeta).\)

**Step 5.** Using \(y(\cdot|v, \zeta)\) and \(\lambda(\cdot|v, \zeta),\) compute \(\frac{\partial \bar{J}_p(v, \zeta, \epsilon)}{\partial h}, \frac{\partial \bar{J}_p(v, \zeta, \epsilon)}{\partial \zeta_h}, \frac{\partial \bar{J}_p(v, \zeta, \epsilon)}{\partial v_j}, h = 1, 2, \ldots, r, j = 1, 2, \ldots, \bar{N},\)

according to the formulas in Theorems 4.1 and 4.2.

**Step 6.** Solve Problem \((\bar{Q})\) by using a standard gradient-based nonlinear optimization method, e.g., SQP, to obtain \((v^*, \zeta^*, \epsilon^*).\)

**Step 7.** If \(\epsilon^* < \epsilon_{\text{min}},\) then stop and construct the optimal solution \((v^*, \delta(\epsilon^*))\) of Problem \((P(p))\) from \((v^*, \zeta^*, \epsilon^*)\) by (21). Otherwise, set \(\rho := \delta_3 \rho,\) \((v, \zeta, \epsilon) := (v^*, \zeta^*, \epsilon^*)\) and go to Step 2.

5. **Numerical example.** Consider the following nonlinear switched system with two time delays:

\[
\begin{align*}
\text{subsystem 1:} & \quad \begin{cases} \dot{x}_1(t) = -5x_1(t) - 4x_2(t) - 3x_1(t - 0.5) + 2x_2(t - 0.1) + u(t) + 0.1 \tanh(x_1(t)), \\ \dot{x}_2(t) = 0.1x_1(t) - 7x_2(t) + 0.5u(t) - \sin(x_2(t - 0.1)), \end{cases} \\
\text{subsystem 2:} & \quad \begin{cases} \dot{x}_1(t) = -4x_1(t) + 0.5x_2(t) + 0.2 \sin(x_2(t)) + t^2 + 8, \\ \dot{x}_2(t) = 5x_1(t) - 5x_2(t) + 0.5 \sin(x_1(t - 0.1)) - u(t), \end{cases}
\end{align*}
\]

with the initial conditions

\[ x_1(t) = 6, \quad x_2(t) = t^2 + 2, \quad t \leq 0. \]

The time horizon is \([0, 1.5].\) We assume that the system switches once during the time horizon from subsystem 1 to subsystem 2. Suppose that the switched point \(\tau_1 \in [0.1, 1.4],\) and the control \(u(t) \in [-5, 10]\) for each \(t \in [0, 1.5].\) The terminal state should satisfy the following condition:

\[ x_2(1.5) = 1. \]
Table 1. Cost, terminal constraint and total variation for different weighting coefficients

| Weight | Cost | Terminal constraint | Total variation |
|--------|------|---------------------|-----------------|
| γ      | $x_1(1.5) - 2$ | $x_2(1.5) - 1$ | $\int_0^{1.5} u$ |
| 0      | $4.3054 \times 10^{-5}$ | $3.7923 \times 10^{-8}$ | 185.4279 |
| 0.01   | 0.0024 | $4.9779 \times 10^{-8}$ | 95.1071 |
| 0.05   | 0.0173 | $2.9718 \times 10^{-7}$ | 53.5099 |
| 0.1    | 0.0716 | $1.3383 \times 10^{-7}$ | 31.3635 |
| 0.5    | 0.0234 | $3.8322 \times 10^{-5}$ | 5.4157 |

Our goal is to find an optimal switching time $\tau_1 \in [0, 1.4]$ and an optimal control $u : [0, 1.5] \rightarrow [-5, 10]$ such that the cost functional

$$J(\tau_1, u) = (x_1(1.5) - 2)^2 + \gamma \int_0^{1.5} u$$

is minimized subject to the switched time-delay system (37)-(39) and the terminal constraint (40).

We use our proposed computational method in the previous section with $\epsilon_{\text{min}} = 10^{-6}$, $\rho_0 = 1$, $\delta_1 = 1$, $\delta_2 = 2$, and $\delta_3 = 5$ to solve this example for different weighting factors $\gamma = 0$, $\gamma = 0.01$, $\gamma = 0.05$, $\gamma = 0.1$, and $\gamma = 0.5$, respectively. Our numerical results are summarized in Table 1. The optimal control and the corresponding state trajectories are shown in Figure 1 and Figure 2, respectively. From Figure 1, we can see that the optimal controls become more smooth when the weighting factor $\gamma$ is increased, that is, the total variation is emphasized when the weighting factor $\gamma$ is increased.

6. Conclusions. This paper has studied optimal control problem involving switched systems with multiple time-delays and subject to terminal state constraints. The objective functional in the problem contains a cost on changing control. A gradient-based optimization method was developed to solve the optimal control problem by the control parameterization, the smoothing technique and the exact penalty method. A numerical example indicates the effectiveness of the proposed method.

In fact, the optimal control problem considered in this paper has wide applications in the practical engineering; see, for example [10, 23]. To solve these engineering problems using our developed computational method is our undergoing studies.

Acknowledgments. This work is partially supported by the Natural Science Foundation of China (Nos. 11201267, 11001153 and 71471103) and the Shandong Province Natural Science Foundation of China (Nos. ZR2015AL010, ZR2015PG006 and ZR2013AQ022).
Figure 1. Optimal control.

REFERENCES

[1] N. U. Ahmed, Elements of Finite-Dimensional Systems and Control Theory, Longman Scientific and Technical, Essex, 1988.
[2] S. C. Bengea and A. D. Raymond, Optimal control of switching systems, Automatica, 41 (2005), 11–27.
[3] J. M. Blatt, Optimal control with a cost of switching control, Journal of the Australian Mathematical Society-Series B, 19 (1976), 316–332.
[4] F. Delmotte, E. I. Verriest and M. Egerstedt, Optimal impulsive control of delay systems, ESAIM Control Optimisation and Calculus of Variations, 14 (2008), 767–779.
[5] P. Howlett, Optimal strategies for the control of a train, Automatica, 32 (1996), 519–532.
[6] R. Li, K. L. Teo, K. H. Wong and G. R. Duan, Control parameterization enhancing transform for optimal control of switched systems, Mathematical and Computer Modelling, 43 (2006), 1393–1403.
[7] Q. Lin, R. Loxton and K. L. Teo, The control parameterization method for nonlinear optimal control: A survey, Journal of Industrial and Management Optimization, 10 (2014), 275–309.
[8] Q. Lin, R. Loxton, K. L. Teo and Y. H. Wu, A new computational method for optimizing nonlinear impulsive systems, Dynamics of Continuous, Discrete and Impulsive Systems–Series B, 18 (2011), 59–76.
[9] C. Liu, Z. Gong, B. Shen and E. Feng, Modelling and optimal control for a fed-batch fermentation process, Applied Mathematical Modelling, 37 (2013), 695–706.
[10] C. Liu, Z. Gong, K. L. Teo and E. Feng, Multi-objective optimization of nonlinear switched time-delay systems in fed-batch process, Applied Mathematical Modelling, 40 (2016), 10533–10548.
[11] C. Liu, R. Loxton and K. L. Teo, Switching time and parameter optimization in nonlinear switched systems with multiple time–delays, Journal of Optimization Theory and Applications, 163 (2014), 957–988.
[12] R. Loxton, K. L. Teo and V. Rehbock, Computational method for a class of switched system optimal control problems, IEEE Transactions on Automatic Control, 54 (2009), 2455–2460.
[13] R. Loxton, K. L. Teo, V. Rehbock and W. K. Ling, Optimal switching instants for a switched–capacitor DC/DC power converter, Automatica, 45 (2009), 973–980.
[14] R. Loxton, Q. Lin and K. L. Teo, Minimizing control variation in nonlinear optimal control, Automatica, 49 (2013), 2652–2664.
[15] J. Matula, On an extremum problem, Journal of the Australian Mathematical Society-Series B, 28 (1987), 376–392.
[16] J. Nocedal and S. J. Wright, Numerical Optimization, Springer, New York, 1999.
OPTIMAL CONTROL OF SWITCHED TIME-DELAY SYSTEMS

Figure 2. Optimal state trajectories.

[17] J. P. Richard, *Time-delay systems: An overview of some recent advances and open problems*, *Automatica*, 39 (2003), 1667–1694.
[18] T. I. Seidman, *Optimal control for switching systems*, Proceedings of the 21st Annual Conference on Information Science and Systems, 1987.
[19] K. L. Teo, C. J. Goh and K. H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Longman Scientific and Technical, Essex, 1991.
[20] K. L. Teo and L. S. Jennings, *Optimal control with a cost on changing control*, *Journal of Optimization Theory and Applications*, 68 (1991), 336–357.
[21] E. I. Verriest, *Optimal control for switched point delay systems with refractory period*, in *The 16th IFAC World Congress*, 38 (2005), 413–418.
[22] E. I. Verriest, F. Delmotte and M. Egerstedt, *Optimal impulsive control of point delay systems with refractory period*, Proceedings of the 5th IFAC Workshop on Time Delay Systems, 2004.
[23] L. Wang, Q. Lin, R. Loxton, K. L. Teo and G. Cheng, *Optimal 1,3-propanediol production: Exploring the trade-off between process yield and feeding rate variation*, *Journal of Process Control*, 32 (2015), 1–9.
[24] S. F. Woon, V. Rehbock and R. Loxton, *Towards global solutions of optimal discrete-valued control problems*, *Optimal Control Applications and Methods*, 33 (2012), 576–594.
[25] C. Wu, K. L. Teo, R. Li and Y. Zhao, Optimal control of switched systems with time delay, *Applied Mathematics Letters*, 19 (2006), 1062–1067.

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

[27] C. Yu, B. Li, R. Loxton and K. L. Teo, A new exact penalty function method for continuous inequality constrained optimization problems, *Journal of Industrial and Management Optimization*, 6 (2010), 895–910.

Received July 2016; revised October 2016.

E-mail address: zhaohuagong@163.com
E-mail address: liu_chongyang@yahoo.com
E-mail address: wyjhit@gmail.com