International Journal of Innovative Computing, Information and Control
Volume 10, Number 4, August 2014
 ISSN 1349-4198
pp. 1555–1566

OPTIMAL CONTROL FOR SWITCHED DELAY SYSTEMS

Xiaomei Liu[1], Kanjian Zhang[1], Shengtao Li[2] and Haikun Wei[1]

[1]Key Laboratory of Measurement and Control of CSE, Ministry of Education
School of Automation
Southeast University
No. 2, Sipailou, Xuanwu District, Nanjing 210096, P. R. China
{ xiaomeiliu; kjzhang; hkwei }@seu.edu.cn

[2]School of Information Science and Engineering
Shandong Normal University
No. 88, East Wenhua Road, Jinan 250014, P. R. China
shengtaoli1985@163.com

Received September 2013; revised January 2014

Abstract. *Based on the calculus of variation, an optimal switching control problem for switched autonomous systems with time-delay is studied. The switching times in a sequence of fixed vector fields are the sole control variables and the subsystems do not require a refractory period, which can bring more generality. The analytic expressions for the partial derivatives of the given performance cost with respect to the control variables are derived. Furthermore, necessary conditions for a stationary solution are obtained for the switched delay systems with separable modes. Finally, the efficiency and effectiveness of the proposed method are demonstrated by two examples.*
Keywords: Optimal control, Switched systems, Time-delay, Calculus of variation

1. **Introduction.** A switched dynamic system is a particular kind of hybrid system that consists of several subsystems and a switching law specifying the active subsystem at each switching instant. Such systems arise in a variety of applications, including power systems, industrial process control, automotive systems, and networked control systems [1-7].

As the wide applications of switched systems, many results for the stability analysis and controller design have appeared in the literature [8-11]. In some practical problems, some performance should also be considered. Due to the problems' significance in theory and application, recently, considerable research efforts have been made for the optimal control of switched systems. Many of them concern the problems whose control variables consist of a proper switching law and the input $u(t)$ [3,8]. However, a more special class considers autonomous systems where the term $u(t)$ is absent, the modes sequence is fixed and the switching times are the only control variables [13-15].

Most of the approaches in the literature are based on discretization method, which may lead to computational combinatoric explosions and the obtained solutions may not be accurate enough. Without using the discretization method, [14] formulates an optimization problem in terms of minimizing a cost functional defined on the states dependent on the switching times. In that work, the gradient formula of the cost with respect to the switching times is derived and can be applied in various nonlinear programming algorithms. Further, [15] considers a similar problem, and develops an especially simpler formula, which leads itself to be directly used in conjunction with various gradient descent algorithms.

The above existing results are focused on switched systems without time-delay, but time-delay is often encountered in a variety of practical systems and often sources of instability and poor performance in systems. Therefore, the study of the switched systems with time-delay has penetrated into various branches, such as passivity and passification, stability analysis and controller design, H_∞ filtering problem, optimal control and the references therein [16-21]. It is worth noting that the switching times are not regarded as control variables among these results.

In addition, it should be noted that the optimal switching problem bears some relation to the optimal impulsive control problem [22,23]. In [22], the optimal impulsive control problem for a single delay system is studied. It is assumed that the considered systems all have a refractory period, in the sense that once an action is taken, it takes a non-infinitesimal amount of time before a subsequent action can be taken. Consequently, the optimal sequence no longer follows from the solution of the fixed sequence problem. [23] considers a similar problem, it can overcome the assumptions that the systems all require a refractory period, and hence brings more generality.

In this paper, we extend the results of [15] to a class of time-delay systems. Here, the systems do not require a refractory period. This problem has been rarely studied since the presence of delay makes the problem much more complicated. Meanwhile, its presence adds a nontrivial twist to the original problem posed in [14].

The main contribution is that for the switched delay systems with separable modes, using calculus of variations [24], we derive the analytic expressions for the partial derivatives of performance cost with respect to switching times. The paper is organized as follows. Section 2 formulates the problem and presents the notations. For switched delay systems with separable modes, Section 3 derives the necessary conditions for a stationary solution via classical variational methods. Two examples prove the feasibility of the proposed method in Section 4. Finally, Section 5 states the conclusions and discussions.

2. **Problem Formulation.**

Consider a system with a single time-delay τ. Following the standard notation in [25], denote $x_t = \{x(t+\theta)\,|-\tau \le \theta \le 0\}$. In this paper, the discussed dynamical system is a switched autonomous delay system with a fixed sequence of vector fields: $f_i(x_t)$, $i = 1, \cdots, N$. We assume that the state is continuous at switching times. The switching instants are the sole control variables.

The problem is to determine these switching times control variables such that the performance cost

$$J = \int_{T_0}^{T_N} L(x, \xi)\, dt + \sum_{i=1}^{N} \Phi_i \left(x(T_i), \{T_j\}_{j=1}^{N} \right) \tag{1}$$

is minimized for a fixed initial time ($T_0 = 0$) and terminal time ($T_N = T$). Here, $\xi(t)$ is a discrete state, taking values in the finite set $\Xi = \{1, \cdots, N\}$ and denoting the operating mode at time t. Φ_i, $i = 1, \cdots, N$ are switching costs associated with the control.

3. **Variational Approach to Optimal Switching.**

Note that the above problem is actually a multivariable parameter optimization problem. However, solving it requires the explicit solution of the state equations, which are dependent on the switching times T_i, $i = 1, \cdots, N$. We, therefore, solve the problem by classical variational methods.

For simplicity, we consider the separable modes systems with one time-delay:

$$\dot{x}(t) = f_i(x(t)) + g_i(x(t-\tau)), \quad T_{i-1} \le t \le T_i \tag{2}$$

Setting $L(x, \xi(t)) = L_i(x)$ in the interval (T_{i-1}, T_i), then the performance cost (1) expands to:

$$J = \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} L_i(x(t))\, dt + \sum_{i=1}^{N} \Phi_i\left(x(T_i), \{T_j\}_{j=1}^{N}\right) \tag{3}$$

For this, we will firstly analyze the cost variation between two systems: an "unperturbed" system x with nominal switching times T_i and a "perturbed" system $\tilde{x}$ for which arbitrary, independent perturbations ε are added to the control variables, i.e., $T_i \to T_i + \varepsilon\theta_i$ with $\varepsilon \to 0$.

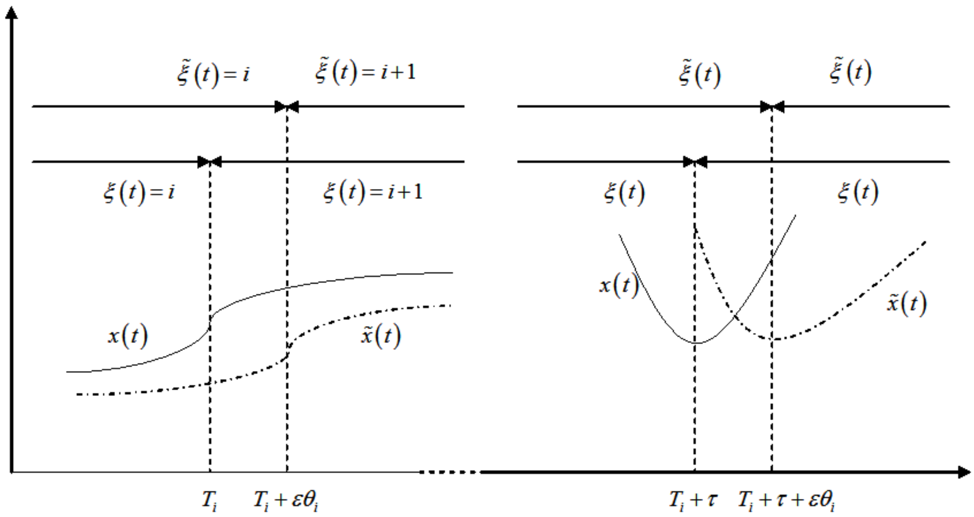

FIGURE 1. Compared trajectories

Figure 1 shows how x and $\tilde{x}$ differ, with a focus on intervals $(T_i, T_i + \varepsilon\theta_i)$ and $(T_i + \tau, T_i + \tau + \varepsilon\theta_i)$. Outside these intervals, the variation is continuous and ε-small, we write $\tilde{x}(t) = x(t) + \varepsilon\eta(t)$. The following equations describe the differences between the two systems:

1) $t \in (T_i, T_i + \varepsilon\theta_i)$,

$$\begin{aligned}
&\xi(t) = i+1 \\
&\tilde{\xi}(t) = i \\
&x(t) = x\left(T_i^{+}\right) + O(\varepsilon) = x(T_i) + O(\varepsilon) \\
&\tilde{x}(t) = x\left(T_i^{-}\right) + O(\varepsilon) = x(T_i) + O(\varepsilon) \\
&\tilde{x}_\tau(t) = x_\tau(t) + O(\varepsilon) = x_\tau(T_i) + O(\varepsilon) \\
&\dot{\tilde{x}}(t) = x\left(T_i^{-}\right) + O(\varepsilon)
\end{aligned} \tag{4}$$

2) $t \in (T_i + \tau, T_i + \tau + \varepsilon\theta_i)$,

$$\begin{aligned}
&\xi(t) = \tilde{\xi}(t) = \xi(T_i + \tau) \\
&\tilde{x}(t) = x(t) + O(\varepsilon) = x(T_i + \tau) + O(\varepsilon) \\
&x_\tau(t) = x\left(T_i^{+}\right) + O(\varepsilon) = x(T_i) + O(\varepsilon) \\
&\tilde{x}_\tau(t) = x\left(T_i^{-}\right) + O(\varepsilon) = x(T_i) + O(\varepsilon) \\
&\tilde{x}_\tau(t) = x_\tau(t) + O(\varepsilon) = x_\tau(T_i) + O(\varepsilon) \\
&\dot{\tilde{x}}(t) = \dot{x}\left((T_i + \tau)^{-}\right) + O(\varepsilon) = \dot{x}(T_i + \tau) + O(\varepsilon)
\end{aligned} \tag{5}$$

3) otherwise,

$$\begin{aligned} \tilde{\xi}(t) &= \xi(t) \\ \tilde{x}(t) &= x(t) + \varepsilon\eta(t) \\ \tilde{x}_\tau(t) &= x_\tau(t) + \varepsilon\eta_\tau(t) \\ \dot{\tilde{x}}(t) &= \dot{x}(t) + \varepsilon\dot{\eta}(t) \end{aligned} \tag{6}$$

Since $T_0 < \cdots < T_{N-1} < T_N$, there is no overlap between any two sets $(T_i, T_i + \varepsilon\theta_i)$ and $(T_j, T_j + \varepsilon\theta_j)$ as $\varepsilon \to 0$. Similar to [25], we assume no overlap between any two sets $(T_i + \tau, T_i + \tau + \varepsilon\theta_i)$ and $(T_j, T_j + \varepsilon\theta_j)$. Thus, the first equation in (5) does not require that on $(T_i + \tau, T_i + \tau + \varepsilon\theta_i)$, $\tilde{\xi}(t) = \xi(t) = i + 1$, which means that subsequent switch happening before $T_i + \tau$ (i.e., $T_{i+1} < T_i + \tau$) is allowable. In other words, a refractory period of τ seconds after each switch is not required, as was the case in [22,26,27].

Then, we analyze the induced variation in performance index. Because of the switch, we adjoin dynamical constraints with a Lagrange multiplier $\lambda(t)$, defined in the subintervals between state switches. Moreover, due to the continuity of state a change in T_j will have an effect on all the modes $i > j$. Considering the accumulation of effects, keeping track of all these effects will complicate the derivation [26]. As in [26,27], we adjoin the constraints $x\left(T_i^-\right) = x\left(T_i^+\right)$ at the switching times with a sequence of Lagrange multipliers μ_i.

Hence, assuming optimal switching control variables T_i, $i = 1, \cdots, N$ exist, and considering both the dynamical constraints and the state continuity constraints, the optimal nominal performance index $\bar{J}_0$ is:

$$\begin{aligned} \bar{J}_0 = &\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[L_{\xi(t)}(x(t)) + \lambda^T(t)\left(f_{\xi(t)}(x(t)) + g_{\xi(t)}(x(t-\tau)) - \dot{x}(t)\right)\right] dt \\ &+ \sum_{i=1}^{N} \left[\Phi_i\left(x(T_i), \{T_i\}_{j=1}^N\right) + \mu_i^T\left(x\left(T_i^+\right) - x\left(T_i^-\right)\right)\right] \end{aligned} \tag{7}$$

For simplicity, define the Hamiltonian functionals:

$$H_\xi(x, x_\tau, \lambda) \overset{def}{=} L_\xi(x) + \lambda^T\left[f_\xi(x) + g_\xi(x_\tau)\right] \tag{8}$$

Then, (7) becomes:

$$\begin{aligned} \bar{J}_0 &= \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[H_\xi(x, x_t, \lambda) - \lambda^T\dot{x}\right] dt \\ &\quad + \sum_{i=1}^{N} \left[\Phi_i\left(x(T_i), \{T_i\}_{j=1}^N\right) + \mu_i^T\left(x\left(T_i^+\right) - x\left(T_i^-\right)\right)\right] \\ &= \bar{J}_0^{(1)} + \bar{J}_0^{(2)} \end{aligned} \tag{9}$$

Remark 3.1. *The expression of the second term $\bar{J}_0^{(2)}$ of (9), i.e., switching cost adjoined with the state continuity constraints obviates the need for computing the perturbations at switching times.*

Similarly, for the perturbed systems, we have:

$$\begin{aligned} \bar{J}_\varepsilon = &\sum_{i=1}^{N} \int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_i+\varepsilon\theta_i} \left[H_{\tilde{\xi}}(\tilde{x}, \tilde{x}_t, \lambda) - \lambda^T\dot{\tilde{x}}\right] dt + \sum_{i=1}^{N} \left[\Phi_i\left(\tilde{x}(T_i + \varepsilon\theta_i), \{T_j + \varepsilon\theta_j\}_{j=1}^N\right)\right. \\ &\left. + \mu_i^T\left(\tilde{x}\left((T_i + \varepsilon\theta_i)^+\right) - \tilde{x}\left((T_i + \varepsilon\theta_i)^-\right)\right)\right] \\ = &\bar{J}_\varepsilon^{(1)} + \bar{J}_\varepsilon^{(2)} \end{aligned} \tag{10}$$

Noting that on ε-small intervals $(T_i, T_i + \varepsilon\theta_i)$ and $(T_i + \tau, T_i + \tau + \varepsilon\theta_i)$, the discrepancies between ξ and $\tilde{\xi}$, x and $\tilde{x}$, or x_τ and $\tilde{x}_\tau$ yield a discrepancy in the Hamiltonian.

Therefore, we split the integral terms $\bar{J}_0^{(1)}$ and $\bar{J}_\varepsilon^{(1)}$ of (9) and (10), and then we have:

$$\begin{aligned}
\bar{J}_0^{(1)} &= \sum_{i=1}^{N}\int_{T_{i-1}}^{T_{i-1}+\varepsilon\theta_{i-1}}\left[H_\xi\left(x,x_\tau,\lambda\right)-\lambda^T\dot{x}\right]dt+\sum_{i=1}^{N}\int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau}\left[H_\xi\left(x,x_\tau,\lambda\right)-\lambda^T\dot{x}\right]dt\\
&\quad+\sum_{i=1}^{N}\int_{T_{i-1}+\tau}^{T_{i-1}+\tau+\varepsilon\theta_{i-1}}\left[H_\xi\left(x,x_\tau,\lambda\right)-\lambda^T\dot{x}\right]dt\\
&\quad+\sum_{i=1}^{N}\int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i}\left[H_\xi\left(x,x_\tau,\lambda\right)-\lambda^T\dot{x}\right]dt\\
&=\sum_{i=1}^{N}\varepsilon\theta_{i-1}\left[H_i\left(x\left(T_{i-1}\right),x_\tau\left(T_{i-1}\right),\lambda\left(T_{i-1}^{+}\right)\right)-\lambda^T\left(T_{i-1}^{+}\right)\dot{x}\left(T_{i-1}^{+}\right)\right]\\
&\quad+\sum_{i=1}^{N}\int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau}\left[H_\xi\left(x,x_\tau,\lambda\right)-\lambda^T\dot{x}\right]dt+\sum_{i=1}^{N}\int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i}\left[H_\xi\left(x,x_\tau,\lambda\right)-\lambda^T\dot{x}\right]dt\\
&\quad+o\left(\varepsilon\right)+\sum_{i=1}^{N}\varepsilon\theta_{i-1}\left[H_{\xi(T_{i-1}+\tau)}\left(x\left(T_{i-1}+\tau\right),x\left(T_{i-1}\right),\lambda\left(\left(T_{i-1}+\tau\right)^{+}\right)\right)\right.\\
&\quad\left.-\lambda^T\left(\left(T_{i-1}+\tau\right)^{+}\right)\dot{x}\left(\left(T_{i-1}+\tau\right)^{+}\right)\right]
\end{aligned}$$

and

$$\begin{aligned}
\bar{J}_\varepsilon^{(1)} &= \sum_{i=1}^{N}\int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau}\left[H_{\tilde{\xi}}\left(\tilde{x},\tilde{x}_\tau,\lambda\right)-\lambda^T\dot{\tilde{x}}\right]dt+\sum_{i=1}^{N}\int_{T_{i-1}+\tau}^{T_{i-1}+\tau+\varepsilon\theta_{i-1}}\left[H_{\tilde{\xi}}\left(\tilde{x},\tilde{x}_\tau,\lambda\right)-\lambda^T\dot{\tilde{x}}\right]dt\\
&\quad+\sum_{i=1}^{N}\int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i}\left[H_{\tilde{\xi}}\left(\tilde{x},\tilde{x}_\tau,\lambda\right)-\lambda^T\dot{\tilde{x}}\right]dt+\sum_{i=1}^{N}\int_{T_i}^{T_i+\varepsilon\theta_i}\left[H_{\tilde{\xi}}\left(\tilde{x},\tilde{x}_\tau,\lambda\right)-\lambda^T\dot{\tilde{x}}\right]dt\\
&=\sum_{i=1}^{N}\int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau}\left[H_\xi\left(x+\varepsilon\eta,x_\tau+\varepsilon\eta_\tau,\lambda\right)-\lambda^T\left(\dot{x}+\varepsilon\dot{\eta}\right)\right]dt\\
&\quad+\sum_{i=1}^{N}\varepsilon\theta_{i-1}\left[H_{\xi(T_{i-1}+\tau)}\left(x\left(T_{i-1}+\tau\right),x\left(T_{i-1}\right),\lambda\left(\left(T_{i-1}+\tau\right)^{+}\right)\right)\right.\\
&\quad\left.-\lambda^T\left(\left(T_{i-1}+\tau\right)^{+}\right)\dot{x}\left(\left(T_{i-1}+\tau\right)^{-}\right)\right]\\
&\quad+\sum_{i=1}^{N}\int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i}\left[H_\xi\left(x+\varepsilon\eta,x_\tau+\varepsilon\eta_\tau,\lambda\right)-\lambda^T\left(\dot{x}+\varepsilon\dot{\eta}\right)\right]dt\\
&\quad+\sum_{i=1}^{N}\varepsilon\theta_i\left[H_i\left(x\left(T_i\right),x_\tau\left(T_i\right),\lambda\left(T_i^{+}\right)\right)-\lambda^T\left(T_i^{+}\right)\dot{x}\left(T_i^{-}\right)\right]+o\left(\varepsilon\right)
\end{aligned}$$

Now, by taking the first order Taylor expansion at ε, a change in variable, and noting that $\theta_0=\theta_N=0$, we get the first part of the derivative:

$$\begin{aligned}
\frac{\bar{J}_\varepsilon^{(1)}-\bar{J}_0^{(1)}}{\varepsilon} &= \sum_{i=1}^{N}\int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau}\left[D_xH_\xi-\lambda^T\dot{\eta}\right]dt+\sum_{i=1}^{N}\int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i}\left[D_xH_\xi-\lambda^T\dot{\eta}\right]dt\\
&\quad+\sum_{i=1}^{N-1}\theta_i\lambda^T\left(\left(T_i+\tau\right)^{+}\right)\left[\left(\dot{x}\left(\left(T_i+\tau\right)^{+}\right)-\dot{x}\left(\left(T_i+\tau\right)^{-}\right)\right)\right]\\
&\quad+\sum_{i=1}^{N-1}\theta_i\left[H_i\left(x\left(T_i\right),x_\tau\left(T_i\right),\lambda\left(T_i^{+}\right)\right)-\lambda^T\left(T_i^{+}\right)\dot{x}\left(T_i^{-}\right)\right]\\
&\quad-\sum_{i=1}^{N-1}\theta_i\left[H_{i+1}\left(x\left(T_i\right),x_\tau\left(T_i\right),\lambda\left(T_i^{+}\right)\right)-\lambda^T\left(T_i^{+}\right)\dot{x}\left(T_i^{+}\right)\right]+o\left(1\right)
\end{aligned}\tag{11}$$

where D_xH_ξ is the functional derivative of H_ξ:

$$D_xH_\xi=\lim_{\varepsilon\to0}\frac{H_\xi\left(x+\varepsilon\eta,x_\tau+\varepsilon\eta_\tau,\lambda\right)-H_\xi\left(x,x_\tau,\lambda\right)}{\varepsilon}$$

In (11), we replace the Hamiltonian with its expression from (8), take the first order Taylor approximation to $D_x H_\xi$, and integrate by parts the $\lambda^T \dot{\eta}$ terms:

$$\begin{aligned}
\tfrac{\bar{J}_\varepsilon^{(1)}-\bar{J}_0^{(1)}}{\varepsilon} = & \sum_{i=1}^{N} \int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau} \left[\tfrac{\partial L_\xi(x)}{\partial x}\eta + \lambda^T\left[\tfrac{\partial f_\xi(x)}{\partial x}\eta + \tfrac{\partial g_\xi(x_\tau)}{\partial x_\tau}\eta_\tau\right] + \dot{\lambda}^T\eta\right] dt - \lambda^T\eta\Big|_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau} \\
& + \sum_{i=1}^{N} \int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i} \left[\tfrac{\partial L_\xi(x)}{\partial x}\eta + \lambda^T\left[\tfrac{\partial f_\xi(x)}{\partial x}\eta + \tfrac{\partial g_\xi(x_\tau)}{\partial x_\tau}\eta_\tau\right] + \dot{\lambda}^T\eta\right] dt \\
& - \lambda^T\eta\Big|_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i} + \sum_{i=1}^{N-1} \theta_i \left[\, \left[L_i\left(x\left(T_i\right)\right) - L_{i+1}\left(x\left(T_i\right)\right)\right]\right. \\
& +\lambda^T\left(T_i^+\right)\left[\dot{x}\left(T_i^+\right) - \dot{x}\left(T_i^-\right)\right] + \lambda^T\left(T_i^+\right)\left[f_i\left(x\left(T_i\right)\right) + g_i\left(x_\tau\left(T_i\right)\right)\right] \\
& -\lambda^T\left(T_i^+\right)\left[f_{i+1}\left(x\left(T_i\right)\right) + g_{i+1}\left(x_\tau\left(T_i\right)\right)\right] \\
& \left.+\lambda^T\left(\left(T_i+\tau\right)^+\right)\left[\left(\dot{x}\left(\left(T_i+\tau\right)^+\right) - \dot{x}\left(\left(T_i+\tau\right)^-\right)\right)\right]\right]
\end{aligned} \tag{12}$$

Before concerning on the second part of the derivative, we derive useful equations:

$$\begin{aligned}
\tilde{x}\left(T_i+\varepsilon\theta_i\right) &= \tilde{x}\left(T_i\right) + \varepsilon\theta_i\dot{\tilde{x}}\left(T_i\right) + o\left(\varepsilon\right) \\
&= x\left(T_i\right) + \varepsilon\eta\left(T_i\right) + \varepsilon\theta_i\dot{x}\left(T_i\right) + o\left(\varepsilon\right)
\end{aligned}$$

$$\begin{aligned}
& \tilde{x}\left(\left(T_i+\varepsilon\theta_i\right)^+\right) - \tilde{x}\left(\left(T_i+\varepsilon\theta_i\right)^-\right) \\
&= x\left(\left(T_i+\varepsilon\theta_i\right)^+\right) + \varepsilon\eta\left(\left(T_i+\varepsilon\theta_i\right)^+\right) - \tilde{x}\left(\left(T_i+\varepsilon\theta_i\right)^-\right) \\
&= x\left(T_i^+\right) + \varepsilon\theta_i\dot{x}\left(T_i^+\right) + \varepsilon\eta\left(T_i^+\right) - x\left(T_i^-\right) - \varepsilon\eta\left(T_i^-\right) - \varepsilon\theta_i\dot{x}\left(T_i^-\right) + o\left(\varepsilon\right) \\
&= \left[x\left(T_i^+\right) - x\left(T_i^-\right)\right] + \varepsilon\left[\eta\left(T_i^+\right) - \eta\left(T_i^-\right)\right] + \varepsilon\theta_i\left[\dot{x}\left(T_i^+\right) - \dot{x}\left(T_i^-\right)\right] + o\left(\varepsilon\right)
\end{aligned}$$

Similarly, at $t = T_i + \tau + \varepsilon\theta_i$, we get:

$$\begin{aligned}
& \tilde{x}\left(\left(T_i+\tau+\varepsilon\theta_i\right)^+\right) - \tilde{x}\left(\left(T_i+\tau+\varepsilon\theta_i\right)^-\right) \\
&= \left[x\left(\left(T_i+\tau\right)^+\right) - x\left(\left(T_i+\tau\right)^-\right)\right] + \varepsilon\left[\eta\left(\left(T_i+\tau\right)^+\right) - \eta\left(\left(T_i+\tau\right)^-\right)\right] \\
& \quad +\varepsilon\theta_i\left[\dot{x}\left(\left(T_i+\tau\right)^+\right) - \dot{x}\left(\left(T_i+\tau\right)^-\right)\right] + o\left(\varepsilon\right)
\end{aligned}$$

Because both $x\left(t\right)$ and $\tilde{x}\left(t\right)$ are continuous at $T_i + \tau$, we have:

$$\left[\eta\left(\left(T_i+\tau\right)^+\right) - \eta\left(\left(T_i+\tau\right)^-\right)\right] + \theta_i\left[\dot{x}\left(\left(T_i+\tau\right)^+\right) - \dot{x}\left(\left(T_i+\tau\right)^-\right)\right] + o\left(1\right) = 0$$

Then, the non-integral term in (10) expands to

$$\begin{aligned}
\bar{J}_\varepsilon^{(2)} &= \sum_{i=1}^{N}\left[\Phi_i\left(\tilde{x}\left(T_i+\varepsilon\theta_i\right), \left\{T_j+\varepsilon\theta_j\right\}_{j=1}^{N}\right) + \mu_i^T\left(\tilde{x}\left(\left(T_i+\varepsilon\theta_i\right)^+\right) - \tilde{x}\left(\left(T_i+\varepsilon\theta_i\right)^-\right)\right)\right] \\
&= \sum_{i=1}^{N}\left[\Phi_i\left(x\left(T_i\right) + \varepsilon\eta\left(T_i\right) + \varepsilon\theta_i\dot{x}\left(T_i\right), \left\{T_i+\varepsilon\theta_j\right\}_{j=1}^{N}\right)\right. \\
& \quad \left.+\mu_i^T\left(\left[x\left(T_i^+\right) - x\left(T_i^-\right)\right] + \varepsilon\left[\eta\left(T_i^+\right) - \eta\left(T_i^-\right)\right] + \varepsilon\theta_i\left[\dot{x}\left(T_i^+\right) - \dot{x}\left(T_i^-\right)\right]\right) + o\left(\varepsilon\right)\right] \\
&= \sum_{i=1}^{N}\left[\Phi_i\left(x\left(T_i\right), \left\{T_j\right\}_{j=1}^{N}\right) + \mu_i^T\left[x\left(T_i^+\right) - x\left(T_i^-\right)\right]\right] \\
& \quad +\varepsilon\sum_{i=1}^{N}\left[\tfrac{\partial\Phi_i}{\partial x}\left(\eta\left(T_i\right) + \varepsilon\theta_i\dot{x}\left(T_i\right)\right) + \left\{\tfrac{\partial\Phi_i}{\partial T}\theta_j\right\}_{j=1}^{N}\right. \\
& \quad \left.+\mu_i^T\left(\left[\eta\left(T_i^+\right) - \eta\left(T_i^-\right)\right] + \theta_i\left[\dot{x}\left(T_i^+\right) - \dot{x}\left(T_i^-\right)\right]\right) + o\left(\varepsilon\right)\right]
\end{aligned}$$

where $\frac{\partial\Phi_i}{\partial x}$ is the partial derivative of $\Phi_i\left(x, T, \mu\right)$ taken at $\left(x\left(T_i\right), \left\{T_j\right\}_{j=1}^{N}, \mu\right)$. We get the second part of the derivative:

$$\begin{aligned}
\tfrac{\bar{J}_\varepsilon^{(2)}-\bar{J}_0^{(2)}}{\varepsilon} = & \sum_{i=1}^{N}\left[\tfrac{\partial\Phi_i}{\partial x}\left(\eta\left(T_i\right) + \varepsilon\theta_i\dot{x}\left(T_i\right)\right) + \left\{\tfrac{\partial\Phi_i}{\partial T}\theta_j\right\}_{j=1}^{N}\right. \\
& \left.+\mu_i^T\left(\left[\eta\left(T_i^+\right) - \eta\left(T_i^-\right)\right] + \theta_i\left[\dot{x}\left(T_i^+\right) - \dot{x}\left(T_i^-\right)\right]\right)\right]
\end{aligned} \tag{13}$$

By adding (12) and (13), and some rearrangements, we get an expression of the total derivative of J:

$$\begin{aligned}\delta J=\lim_{\varepsilon\to 0}\tfrac{\bar{J}_\varepsilon-\bar{J}_0}{\varepsilon}=\ &\sum_{i=1}^{N-1}\Big[\tfrac{\partial\Phi_i}{\partial x}\dot{x}\left(T_i\right)+\left\{\tfrac{\partial\Phi_i}{\partial T}\theta_j\right\}_{j=1}^{N}+\left(\lambda^T\left(T_i^+\right)+\mu_i^T\right)\left[\left(x\left(T_i^+\right)-x\left(T_i^-\right)\right)\right]\\ &+\left[L_i\left(x\left(T_i\right)\right)-L_{i+1}\left(x\left(T_i\right)\right)\right]+\lambda^T\left(T_i^+\right)\left[f_i\left(x\left(T_i\right)\right)+g_i\left(x\left(T_i-\tau\right)\right)\right]\\ &-\lambda^T\left(T_i^+\right)\left[f_{i+1}\left(x\left(T_i\right)\right)+g_{i+1}\left(x\left(T_i-\tau\right)\right)\right]\Big]\theta_i\\ &+\sum_{i=1}^{N}\int_{T_{i-1}+\varepsilon\theta_{i-1}}^{T_{i-1}+\tau}\left[\tfrac{\partial L_\xi(x)}{\partial x}\eta+\lambda^T\left[\tfrac{\partial f_\xi(x)}{\partial x}\eta+\tfrac{\partial g_\xi(x_\tau)}{\partial x_\tau}\eta_\tau\right]+\dot{\lambda}^T\eta\right]dt\\ &+\sum_{i=1}^{N}\int_{T_{i-1}+\tau+\varepsilon\theta_{i-1}}^{T_i}\left[\tfrac{\partial L_\xi(x)}{\partial x}\eta+\lambda^T\left[\tfrac{\partial f_\xi(x)}{\partial x}\eta+\tfrac{\partial g_\xi(x_\tau)}{\partial x_\tau}\eta_\tau\right]+\dot{\lambda}^T\eta\right]dt\\ &+\sum_{i=1}^{N}\left[\tfrac{\partial\Phi_i}{\partial x}-\mu_i^T-\lambda^T\left(T_i^-\right)\right]\eta\left(T_i^-\right)\\ &+\sum_{i=1}^{N}\left[\mu_i^T+\lambda^T\left(T_i^+\right)\right]\eta\left(T_i^+\right)\\ &+\sum_{i=1}^{N}\left[\lambda^T\left(\left(T_i+\tau\right)^+\right)-\lambda^T\left(\left(T_i+\tau\right)^-\right)\right]\eta\left(\left(T_i+\tau\right)^-\right)\end{aligned}\tag{14}$$

In the following, we choose λ and μ_i to make all η terms disappear. By some mathematical transformations, we can obtain the following results.

To avoid computation of η on these intervals, we choose λ so that on the intervals $(T_{i-1}, T_{i-1}+\tau)$ and $(T_{i-1}+\tau, T_i)$, $i=1,\cdots,N$:

$$\begin{aligned}&\dot{\lambda}^T\left(t\right)=-\tfrac{\partial L_{\xi(t)}(x(t))}{\partial x}-\lambda^T\left(t\right)\tfrac{\partial f_{\xi(t)}(x(t))}{\partial x}-\lambda^T\left(t+\tau\right)\tfrac{\partial g_{\xi(t+\tau)}(x(t))}{\partial x}\\ &\lambda\left(t\right)=0,\quad\left(T,T+\tau\right)\end{aligned}\tag{15}$$

To avoid computation of η at T_i^+, we choose the Lagrange multipliers μ_i so that

$$\mu_i^T=-\lambda^T\left(T_i^+\right),\quad i=1,\cdots,N\tag{16}$$

To avoid computation of η at T_i^-, we choose λ to be discontinuous at T_i with

$$\lambda^T\left(T_i^-\right)=\frac{\partial\Phi_i}{\partial x}+\lambda^T\left(T_i^+\right),\quad i=1,\cdots,N\tag{17}$$

To avoid computation of η at $T_i+\tau^-$, we choose λ to be continuous at $T_i+\tau$

$$\lambda^T\left(\left(T_i+\tau\right)^+\right)=\lambda^T\left(\left(T_i+\tau\right)^-\right),\quad i=1,\cdots,N\tag{18}$$

Now that all η terms have disappeared, we obtain:

$$\delta J=\sum_{i=1}^{N-1}\frac{\partial J}{\partial T_i}\theta_i\tag{19}$$

Therefore,

$$\begin{aligned}\tfrac{\partial J}{\partial T_i}&=\tfrac{\partial\Phi_i}{\partial x}\dot{x}\left(T_i\right)+\left\{\tfrac{\partial\Phi_i}{\partial T}\theta_j\right\}_{j=1}^{N}+\left[L_i\left(x\left(T_i\right)\right)-L_{i+1}\left(x\left(T_i\right)\right)\right]\\ &\quad+\lambda^T\left(T_i^+\right)\left[f_i\left(x\left(T_i\right)\right)+g_i\left(x\left(T_i-\tau\right)\right)\right]-\lambda^T\left(T_i^+\right)\left[f_{i+1}\left(x\left(T_i\right)\right)+g_{i+1}\left(x\left(T_i-\tau\right)\right)\right]\Big]\\ &=\left\{\tfrac{\partial\Phi_i}{\partial T}\theta_j\right\}_{j=1}^{N}+\left[L_i\left(x\left(T_i\right)\right)-L_{i+1}\left(x\left(T_i\right)\right)\right]+\lambda^T\left(T_i^-\right)\left[f_i\left(x\left(T_i\right)\right)+g_i\left(x\left(T_i-\tau\right)\right)\right]\\ &\quad-\lambda^T\left(T_i^+\right)\left[f_{i+1}\left(x\left(T_i\right)\right)+g_{i+1}\left(x\left(T_i-\tau\right)\right)\right]\Big]\end{aligned}\tag{20}$$

Finally, we summarize these results in the following theorem.

Theorem 3.1. *Considering the switched delay system with separable mode in (2) and the performance index J in (3) with fixed initial time ($T_0=0$) and terminal time ($T_N=T$),*

a necessary condition for system (2) to minimize the cost (3) is that the switching times T_i, $i = 1, \cdots, N-1$ *satisfy:*
Euler-Lagrange equations:

$$\dot{\lambda}^T(t) = -\frac{\partial L_{\xi(t)}(x(t))}{\partial x} - \lambda^T(t)\frac{\partial f_{\xi(t)}(x(t))}{\partial x} - \lambda^T(t+\tau)\frac{\partial g_{\xi(t+\tau)}(x(t))}{\partial x} \tag{21}$$

Boundary conditions:

$$\begin{array}{l} \lambda(t) = 0, \quad (T, T+\tau) \\ \lambda^T\left(T_i^-\right) = \frac{\partial \Phi_i}{\partial x} + \lambda^T\left(T_i^+\right), \quad i = 1, \cdots, N \\ \lambda^T\left((T_i+\tau)^+\right) = \lambda^T\left((T_i+\tau)^-\right), \quad i = 1, \cdots, N \end{array} \tag{22}$$

Multipliers:

$$\mu_i^T = -\lambda^T\left(T_i^+\right), \quad i = 1, \cdots, N \tag{23}$$

Optimality conditions:

$$\begin{aligned} \frac{\partial J}{\partial T_i} = {} & \frac{\partial \Phi_i}{\partial T} + \left[L_i(x(T_i)) - L_{i+1}(x(T_i))\right] + \lambda^T\left(T_i^-\right)\left[f_i(x(T_i)) + g_i(x(T_i-\tau))\right] \\ & -\lambda^T\left(T_i^+\right)\left[f_{i+1}(x(T_i)) + g_{i+1}(x(T_i-\tau))\right] \\ = {} & 0 \end{aligned} \tag{24}$$

Remark 3.2. *Although analytic solution of (24) may be quite hard to achieve, the partial derivatives of J can be used in a numerical gradient descent algorithm. See examples in next section in detail. Moreover, as (24) corresponds to the sensitivity evaluation of J with respect to* T_i, *it can also be used for providing descent direction in optimization algorithm.*

Remark 3.3. *For the delay-free case, the necessary conditions can be obtained by setting* $g_i = 0$ *of Theorem 3.1.*

4. **Example Simulation.**

In this section, we use the results of Theorem 3.1 in a descent algorithm that is applied to two examples to illustrate the effectiveness and merit of our results.

The gradient descent-based algorithm is shown as follows:

Algorithm 1:

$T_i = T_{i0}$, $i = 1, \cdots, N-1$ (initial guess)
repeat
  solve for $x(t)$, $t \in [t_0, T_N]$ and $\xi(t)$ forward in time
  solve for $\lambda(t)$, $t \in [t_0, T_N]$ backward in time
  compute the partial derivatives $\frac{\partial J}{\partial T_i}$ with

$$\begin{aligned} \frac{\partial J}{\partial T_i} = {} & \frac{\partial \Phi_i}{\partial T} + \left[L_i(x(T_i)) - L_{i+1}(x(T_i))\right] + \lambda^T\left(T_i^-\right)\left[f_i(x(T_i)) + g_i(x(T_i-\tau))\right] \\ & -\lambda^T\left(T_i^+\right)\left[f_{i+1}(x(T_i)) + g_{i+1}(x(T_i-\tau))\right] \end{aligned}$$

  updates: $T_i := T_i - \gamma_i \frac{\partial J}{\partial T_i}$
until $\left|\frac{\partial J}{\partial T_i}\right| \le \varepsilon$

Example 4.1. *Consider the following scalar time-delay system with one switching time which is taken from [22]:*

$$\dot{x}(t) = \begin{cases} x(t), & t \in (0, T] \\ x(t-1), & t \in (T, 2] \end{cases}$$

with performance cost:

$$J = \frac{1}{2}\int_0^2 (x(t) - 1)^2 \, dt + \frac{1}{T}$$

Algorithm 1 using Theorem 3.1 is applied to minimize the cost J. We choose the nominal control variable $T_1 = 0.8$ and initial condition $x(t) = 0$, $t \le 0$. After 12 iterations we find that the optimal switching time $T_1 = 1.9786$ and the corresponding optimal cost $J = 1.5025$. The numerical simulation is shown in Figure 2.

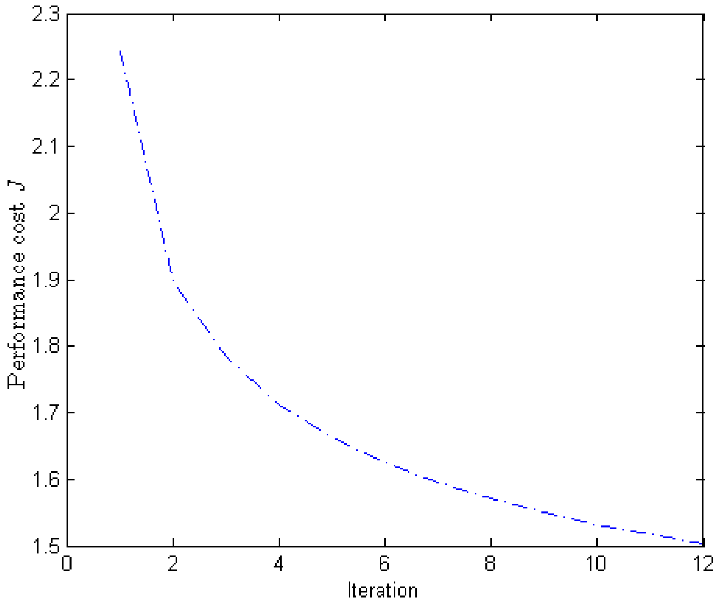

FIGURE 2. Performance cost J

Example 4.2. *We consider the scalar delay system with two switching times which is taken from [23]:*

$$\dot{x}(t) = \begin{cases} \frac{1}{2}x(t), & t \in (0, T_1] \\ \frac{1}{2}x(t-1), & t \in (T_1, T_2] \\ \frac{1}{2}x(t), & t \in (T_2, 3] \end{cases}$$

and performance cost:

$$J = \frac{1}{2}\int_0^3 (x(t) - 1)^2 \, dt + \frac{1}{T_1} + \frac{1}{T_2}$$

Find optimal switching instants T_1, T_2 to minimize the above cost criterion J. For this problem, we choose initial condition $x(t) = 0.4$, $t \le 0$, the initial nominal control variables $T_1 = 0.8$, $T_2 = 1.8$ and the parameters $\gamma_1 = 0.2$, $\gamma_2 = 0.3$. Based on Theorem 3.1, after 20 iterations the optimal switching times are found to be $T_1 = 2.4029$, $T_2 = 2.8041$ and the corresponding optimal cost is $J = 1.0193$. The result is shown in Figure 3.

As seen from Figure 3, the performance cost J can quickly converge to a minimum value with the proposed method.

Remark 4.1. *It is worth noting that, in this particular example, the optimal switching times T_1 and T_2 are within τ seconds, i.e., $T_2 < T_1 + \tau$. The previous work in [17], which requires a refractory period of τ seconds between each switching times, can only provide a suboptimal solution [23].*

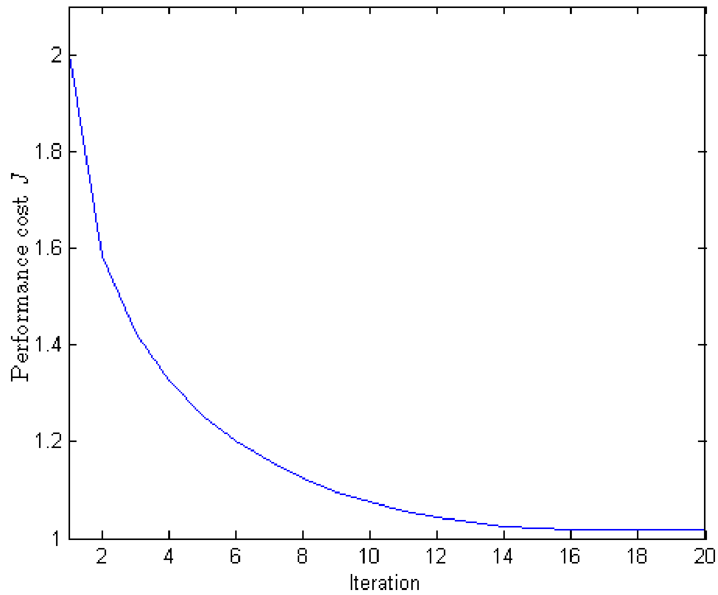

Figure 3. Performance cost J

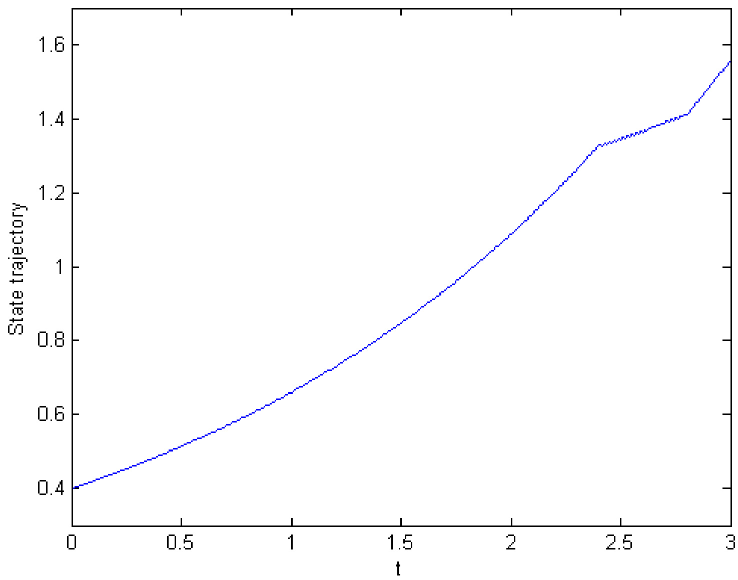

Figure 4. State trajectory

Moreover, the state trajectory associated with the optimal switching times $T_1 = 2.4029$ and $T_2 = 2.8041$ is shown in Figure 4.

5. **Conclusions.**

We have proposed an approach for solving optimal control problem of switched delay systems with prespecified sequences of active subsystems. In particular, the necessary conditions for the stationarity have been derived. This is a first step in the complete optimal control of such switched systems, where the optimal sequence of the modes needs finding.

Acknowledgment. This work is partially supported by Major Program of National Natural Science Foundation of China under Grant 11190015, and National Natural Science Foundation of China under Grant (60874030, 61374006). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] R. Brockett, Stabilization of motor networks, *Proc. of the 34th IEEE Conf. on Decision & Control*, New Orleans, LA, USA, pp.1486-1488, 1995.

[2] L. Y. Wang, A. Beydoun, J. Cook, J. Sun and I. Kolmanovsky, Optimal hybrid control with applications to automotive powertrain systems, *Control Using Logic-Based Switching, Lecture Notes in Control and Information Sciences*, vol.222, pp.190-200, 1997.

[3] M. S. Branicky, V. S. Borkar and S. K. Miner, A unified framework for hybrid control: Model and optimal control theory, *IEEE Trans. on Automat. Control*, vol.43, no.1, pp.31-45, 1998.

[4] G. Walsh, H. Ye and L. Bushnell, Stability analysis of networked control systems, *Proc. of the American Control Conference*, San Diego, CA, USA, pp.2876-2880, 1999.

[5] H. Rehbinder and M. Sanfridson, Scheduling of a limited communication channel for optimal control, *Proc. of the 39th IEEE Conf. on Decision & Control*, Sidney, Australia, pp.1011-1016, 2000.

[6] M. Egerstedt and Y. Wardi, Multi-process control using queuing theory, *Proc. of the 41th IEEE Conf. on Decision & Control*, Las Vegas, NV, USA, pp.1991-1996, 2002.

[7] P. Martin and M. Egerstedt, Optimal timing control of interconnected, switched systems with applications to robotics marionettes, *Journal of Discrete Event Dynamic Systems*, vol.20, no.2, pp.233-248, 2010.

[8] X. Zhao, L. Zhang, P. Shi and M. Liu, Stability and stabilization of switched linear systems with mode-dependent average dwell time, *IEEE Trans. on Automat. Control*, vol.57, no.7, pp.1809-1815, 2012.

[9] X. Zhao, L. Zhang, P. Shi and M. Liu, Stability of switched positive linear systems with average dwell time switching, *Automatica*, vol.48, no.6, pp.1132-1137, 2012.

[10] S. B. Attia, S. Salhi and M. Ksouri, Static switched output feedback stabilization for linear discrete-time switched systems, *International Journal of Innovative Computing, Information and Control*, vol.8, no.5, pp.3203-3213, 2012.

[11] X. Zhao, P. Shi and L. Zhang, Asynchronously switched control of a class of slowly switched linear systems, *Systems & Control Letters*, vol.61, no.12, pp.1151-1156, 2012.

[12] X. Xu and P. J. Antsaklis, An approach for solving general switched linear quadratic optimal control problems, *Proc. of the 40th IEEE Conf. on Decision & Control*, Orlando, FL, USA, pp.2478-2483, 2001.

[13] D. Hristu-Varsakelis, Feedback control systems as users of shared network communication sequences that guarantee stability, *Proc. of the 40th IEEE Conf. on Decision & Control*, Orlando, FL, USA, p.3631, 2001.

[14] X. Xu and P. Antsaklis, Optimal control of switched autonomous systems, *Proc. of the 41th IEEE Conf. on Decision & Control*, Las Vegas, NV, USA, pp.4401-4406, 2002.

[15] M. Egerstedt, Y. Wardi and F. Delmotte, Optimal control of switching times in switched dynamical systems, *Proc. of the 42th IEEE Conf. on Decision & Control*, Maui, HI, USA, pp.2138-2143, 2003.

[16] J. Lian, P. Shi and Z. Feng, Passivity and passification for a class of uncertain switched stochastic time-delay systems, *IEEE Trans. on Cybernetics*, vol.43, no.1, pp.3-13, 2013.

[17] R. Wang, P. Shi, Z. Wu and Y. Sun, Stabilization of switched delay systems with polytopic uncertainties under asynchronous switching, *J. of Franklin Institute*, vol.350, no.8, pp.2028-2043, 2013.

[18] X. Zhao, L. Zhang and P. Shi, Stability of a class of switched positive linear time-delay systems, *Int. J. of Robust and Nonlinear Control*, vol.23, no.5, pp.578-589, 2013.

[19] R. Wang, Z. Wu and P. Shi, Dynamic output feedback control for a class of switched delay systems under asynchronous switching, *Information Sciences*, vol.225, pp.72-80, 2013.

[20] J. Lian, C. Mu and P. Shi, Asynchronous H_∞ filtering for switched stochastic systems with time-varying delay, *Information Sciences*, vol.224, pp.200-212, 2013.

[21] C. Wu, K. L. Teo, R. Li and Y. Zhao, Optimal control of switched systems with time delay, *Applied Mathematics Letters*, vol.19, no.10, pp.1062-1067, 2006.

[22] E. I. Verriest, F. Delmotte and M. Egerstedt, Optimal impulsive control for point delay systems with refractory period, *Proc. of the 5th IFAC Workshop on Time-Delay Systems*, Leuven, Belgium, 2004.

[23] F. Delmotte, E. I. Verriest and M. Egerstedt, Optimal impulsive control of delay systems, *ESAIM: Control, Optimisation and Calculus of Variations*, vol.14, no.4, pp.767-779, 2008.
[24] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*, Hemisphere Publishing, New York, 1975.
[25] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Applied Math. Sciences Series, Springer-Verlag, New York, 1993.
[26] E. I. Verriest and F. Delmotte, Optimal control for switched point delay systems with refractory period, *Proc. of the 16th IFAC World Congress*, Prague, Czeck Republic, 2005.
[27] E. I. Verriest, Optimal control for switched distributed delay systems with refractory period, *Proc. of the 44th IEEE Conf. on Decision* & *Control and 2005 European Control Conference*, Seville, Spain, pp.1421-1426, 2005.