https://doi.org/10.1016/j.automatica.2019.108696

Optimal switching of switched systems with time delay in discrete time ⋆

Wei Xu [a], Zhi Guo Feng [b], Gui-Hua Lin [c], Ka Fai Cedric Yiu [d], Liying Yu [c]

[a] *School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, PR China*

[b] *Faculty of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang, PR China*

[c] *School of Management, Shanghai University, Shanghai, PR China*

[d] *Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, PR China*

Abstract

This paper addresses a kind of optimal switching problem to minimize a quadratic cost functional for the discrete-time switched linear system with time delay. Since the dynamics is influenced by the switching sequence and the time delay, most existing gradient-based methods and relaxation techniques can not be applied. In order to find the optimal solution, we first formulate the switched time-delay system into an equivalent switched system to separate the cross term of coefficient matrices. Based on the positive semi-definiteness of the system, we derive a series of lower bounds of the cost functional. By comparing them with the current optimal value, a depth-first branch and bound technique is proposed and the global optimal solution can be exactly obtained. Some numerical examples are demonstrated to verify the high efficiency of the method.

Key words: Switched systems; Time delay; Optimal switching problem; Branch and bound technique.

1 Introduction

Switched system belongs to a particular class of hybrid system, which is composed of some subsystems and a switching law. For each subsystem, its dynamics is generally described by a differential equation in continuous time or a difference equation in discrete time. The switching law is to govern these active subsystems in a given time horizon. Recently, the switched system has been widely found in many practical applications, including the choice of gears on transports [1], sensor scheduling [2], power converter system [3], dynamic supply chain network [4], and some other fields.

Due to the challenge both in theory and applications, the switched system and its related problems have been a hot topic. In the past years, many theoretical results can be found in the literature, such as the stability analysis in [5,6] and the controller design in [7,8]. To improve the performance of the system, the switching times and the switching sequence have also been considered in optimal control problems. For autonomous switched systems in continuous time, some classical nonlinear programming methods based on the gradient formula of the cost functional were extended to find the optimal switching times in [9–11]. When the control function is introduced, Xu proposed a two-stage optimization strategy in [12]. The control parameterization enhancing transform was also developed to handle the same problem in [13–15]. However, when the system involves the discrete event, such as the switching sequence, the optimal control problem becomes very complicated. In most cases, only the local optimal solution can be obtained. In [16–18], the embedding transformation with some relaxation methods was explored. But it may produce an infinite switching frequency, which is not possible to be implemented in practice. Another widely used technique is the mode-insertion method referred in [19–21], where the given finite switching sequence was updated by inserting some new subsystems to reduce the cost functional value. For the discrete-time switched system applied in [22], any

⋆ This paper was not presented at any IFAC meeting. Corresponding author Z. G. Feng. Fax +86 759 2383064.

Email addresses: xuwei9951@qq.com (Wei Xu), 18281102@qq.com (Zhi Guo Feng), guihualin@shu.edu.cn (Gui-Hua Lin), macyiu@polyu.edu.hk (Ka Fai Cedric Yiu), yly_shu@126.com (Liying Yu).

gradient information disappears and the optimal switching problem is rarely discussed. In general, the optimal switching sequence can only be searched by the enumeration or the dynamic programming. The computational complexity shows the exponential growth. Then, some other discrete optimization methods were explored, such as the discrete filled function algorithm in [23] and the branch and bound method in [24,25].

In this paper, we consider a class of time-delay switched systems in discrete time, where the evolution of the system is influenced by the current sate, the time-delay state and the switching sequence. Our purpose is to find the global optimal switching sequence such that the quadratic cost functional achieves the minimum value. As far as we know, the global optimization method has not been discussed to handle the optimal switching problem with time delay. The difficulties come from two aspects: One is that the available optimization tools are limited to obtain the global optimal solution. Furthermore, under the joint effect of the switching sequence and the time delay, the performance of the cost functional becomes more complicated. In our work, we first transform the original optimal switching problem into an equivalent augmenting form, then these parameter matrices involved different switching signals are separated in the cost functional. By constructing a set of dynamic lower bounds of the cost functional, a depth-first branch and bound algorithm adapted to the discrete-time switched system with time delay is proposed such that the global optimal solution can be obtained efficiently.

The rest of the paper is organized as follows. In Section 2, we state the switched system with time delay in discrete time and consider the optimal switching problem with a quadratic cost functional. In Section 3, we first reformulate this problem as a tractable structure, then the dynamic lower bound of the cost functional is constructed via some matrix transformations and extension techniques. Finally, a depth-first branch and bound algorithm is proposed to find the global optimal solution. In Section 4, some examples are tested to demonstrate the high efficiency of the method.

Notation We use lower-case and upper-case boldface letters to denote vectors and matrices, respectively. For any two matrices $\mathbf{Y}, \mathbf{Z} \in \mathbb{R}^{n\times n}$, the relation $\mathbf{Y} \preceq \mathbf{Z}$ means that $\mathbf{Z} - \mathbf{Y}$ is positive semi-definite. In addition, for the matrix $\mathbf{Y}$, the entry in its i-th row and j-th column is denoted by Y_{ij} and the n-dimensional identity matrix is denoted by $\mathbf{I}_n$. Let $\prod_{k=1}^{K} \mathbf{Y}_k = \mathbf{Y}_1\mathbf{Y}_2\cdots\mathbf{Y}_K$.

2 Problem formulation

We consider the following switched system with time delay in discrete time, which can switch the mode from M linear subsystems at each time.

$$\mathbf{x}(t+1) = \mathbf{A}_{\sigma(t)}(t)\mathbf{x}(t) + \mathbf{B}_{\sigma(t)}(t)\mathbf{x}(t-\tau), \quad t = 0, 1, \ldots, T-1, \tag{1}$$

with initial condition

$$\mathbf{x}(-\xi) = \mathbf{y}_\xi, \quad \xi = 0, 1, \ldots, \tau, \tag{2}$$

where $T > 0$ is the terminal time and $\tau > 0$ is the time delay. Without loss of generality, we assume that $T > \tau$. In fact, the case of $0 < T \leq \tau$ can be viewed as a special case contained in the following discussion. Since only one subsystem is active at each time, we denote the index set of these subsystems by $\mathcal{M} = \{1, 2, \ldots, M\}$ and let $\sigma(t) \in \mathcal{M}$ be the switching signal to activate the corresponding subsystem. Then, a complete switching sequence can be denoted by

$$\sigma = (\sigma(0), \sigma(1), \ldots, \sigma(T-1)) \in \mathcal{M}^T.$$

In the time-delay switched system, $\mathbf{x}(t) \in \mathbb{R}^n$ is the state, $\mathbf{A}_{\sigma(t)}(t) \in \mathbb{R}^{n\times n}$ and $\mathbf{B}_{\sigma(t)}(t) \in \mathbb{R}^{n\times n}$ are both time-varying coefficient matrices. For simplicity, we directly write them by $\mathbf{A}_\sigma(t)$ and $\mathbf{B}_\sigma(t)$ respectively and denote the m-th subsystem by the pair $(\mathbf{A}_m(t), \mathbf{B}_m(t))$ for the rest of the paper. For the initial condition, $\mathbf{y}_\xi \in \mathbb{R}^n$, $\xi = 0, 1, \ldots, \tau$, are given $\tau + 1$ column vectors.

It is worth noting that only the discrete switching sequence $\sigma \in \mathcal{M}^T$ is the decision variable, we state the optimal switching problem as follows.

Problem 1 *For the dynamics* (1) *with initial condition* (2)*, find an optimal switching sequence $\sigma \in \mathcal{M}^T$, such that the quadratic performance index described by*

$$J(\sigma) = \sum_{t=1}^{T} \mathbf{x}^\mathsf{T}(t)\mathbf{Q}(t)\mathbf{x}(t) \tag{3}$$

is minimized, where for any time $t \in \mathcal{T}_1 \triangleq \{1, 2, \ldots, T\}$, $\mathbf{Q}(t) \in \mathbb{R}^{n\times n}$ is a positive semi-definite matrix.

Since the number of all feasible switching sequences is finite, the optimal solution of Problem 1 must be existent. But the evolution of the time-delay switched system depends not only on the current state but also on the time-delay state, which are both governed by the switching sequence σ in the dynamics (1). Thus, it is difficult to analyze the performance of the cost functional (3). Moreover, the optimal switching problem is very hard to be solved all the time, especially for the global optimal solution. Therefore, some efficient optimization method to solve the optimal time-delay switching problem should be explored.

3 Method analysis

3.1 Reformulation of the optimal switching problem

Instead of $\tau+1$ individual initial states, we introduce an integral initial state vector as $\mathbf{y}\in\mathbb{R}^{(\tau+1)n}$, such that $\mathbf{y}=[\mathbf{y}_0^\intercal\ \ \mathbf{y}_1^\intercal\ \ \ldots\ \ \mathbf{y}_\tau^\intercal]^\intercal$. Then all states can be expressed by this new state $\mathbf{y}$ according to the dynamics (1) with the switching sequence σ.

Lemma 1 *Given a set of initial states $\{\mathbf{y}_\xi\in\mathbb{R}^n:\xi=0,1,\ldots,\tau\}$, the state of the time-delay switched system* (1) *with initial condition* (2) *can always be expressed by the formula*

$$\mathbf{x}(t)=\sum_{\xi=0}^{\tau}\mathbf{C}_\xi(t)\mathbf{y}_\xi,\quad t\in\{-\tau,\ldots,T\},\tag{4}$$

where for each $\xi\in\{0,1,\ldots,\tau\}$, the matrix $\mathbf{C}_\xi(t)\in\mathbb{R}^{n\times n}$ satisfies the following equation,

$$\mathbf{C}_\xi(t+1)=\mathbf{A}_\sigma(t)\mathbf{C}_\xi(t)+\mathbf{B}_\sigma(t)\mathbf{C}_\xi(t-\tau),\quad t\geq 0,\tag{5}$$

with initial condition

$$\mathbf{C}_\xi(t)=\begin{cases}\mathbf{0}, & \xi\neq -t,\\ \mathbf{I}_n, & \xi=-t,\end{cases}\qquad t=-\tau,\ldots,0.\tag{6}$$

PROOF. We use the induction method. For the initial states, without loss of generality, we set $t=-\eta$, $\forall\,\eta\in\{0,1,\ldots,\tau\}$. By the initial condition (6), we have $\mathbf{C}_\eta(-\eta)=\mathbf{I}_n$ and $\mathbf{C}_\xi(-\eta)=\mathbf{0}$ when $\xi\neq\eta$. It follows from Eq. (2), we have

$$\mathbf{x}(-\eta)=\mathbf{y}_\eta=\mathbf{C}_\eta(-\eta)\mathbf{y}_\eta=\sum_{\xi=0}^{\tau}\mathbf{C}_\xi(-\eta)\mathbf{y}_\xi.$$

Then, Eq. (4) is true in the case of $t\leq 0$.

Next, we assume that Eq. (4) holds in the case of $t=0,1,\ldots,s$, $s\geq 0$. Then let us continue to consider the time $t=s+1$. Based on the dynamics (1), we have

$$\begin{aligned}&\mathbf{x}(s+1)\\ =&\mathbf{A}_\sigma(s)\sum_{\xi=0}^{\tau}\mathbf{C}_\xi(s)\mathbf{y}_\xi+\mathbf{B}_\sigma(s)\sum_{\xi=0}^{\tau}\mathbf{C}_\xi(s-\tau)\mathbf{y}_\xi\\ =&\sum_{\xi=0}^{\tau}\left(\mathbf{A}_\sigma(s)\mathbf{C}_\xi(s)+\mathbf{B}_\sigma(s)\mathbf{C}_\xi(s-\tau)\right)\mathbf{y}_\xi.\end{aligned}$$

By Eq. (5), we obtain

$$\mathbf{x}(s+1)=\sum_{\xi=0}^{\tau}\mathbf{C}_\xi(s+1)\mathbf{y}_\xi.$$

It means that Eq. (4) also holds in the case of $t=s+1$. Thus, Eq. (4) is always true for any time $t\in\{-\tau,\ldots,T\}$.

This completes the proof. □

It follows from Lemma 1, we denote

$$\mathbf{D}(t)=[\,\mathbf{C}_0(t)\ \ \mathbf{C}_1(t)\ \ \ldots\ \ \mathbf{C}_\tau(t)\,]\in\mathbb{R}^{n\times(\tau+1)n}.$$

Eq. (4) can be equivalently written by $\mathbf{x}(t)=\mathbf{D}(t)\mathbf{y}$, where the time-varying matrix $\mathbf{D}(t)$ satisfies

$$\mathbf{D}(t+1)=\mathbf{A}_\sigma(t)\mathbf{D}(t)+\mathbf{B}_\sigma(t)\mathbf{D}(t-\tau)\tag{7}$$

with initial condition

$$\begin{cases}\mathbf{D}(0)=[\mathbf{I}_n\ \ \mathbf{0}\ \ \ldots\ \ \mathbf{0}]\in\mathbb{R}^{n\times(\tau+1)n},\\ \mathbf{D}(-1)=[\mathbf{0}\ \ \mathbf{I}_n\ \ \ldots\ \ \mathbf{0}]\in\mathbb{R}^{n\times(\tau+1)n},\\ \qquad\vdots\\ \mathbf{D}(-\tau)=[\mathbf{0}\ \ \mathbf{0}\ \ \ldots\ \ \mathbf{I}_n]\in\mathbb{R}^{n\times(\tau+1)n}.\end{cases}\tag{8}$$

Hence, Problem 1 can be equivalently reformulated by the following problem.

Problem 2 *Find a switching sequence $\sigma=(\sigma(0),\sigma(1),\ldots,\sigma(T-1))\in\mathcal{M}^T$, such that the cost functional*

$$J(\sigma)=\sum_{t=1}^{T}\mathbf{y}^\intercal\mathbf{D}^\intercal(t)\mathbf{Q}(t)\mathbf{D}(t)\mathbf{y}\tag{9}$$

is minimized, subject to the dynamics (7) *with initial condition* (8).

Compared with Problem 1, the advantage of Problem 2 is that we need not to consider the dynamic characteristic of the state in the time-delay switched system and just need to handle the positive semi-definite matrix $\mathbf{D}^\intercal(t)\mathbf{Q}(t)\mathbf{D}(t)\in\mathbb{R}^{(\tau+1)n\times(\tau+1)n}$ with respect to the switching sequence σ in the cost functional (9).

3.2 Bounds of the cost functional

Assume that the switching subsequence made up of the first t_j values is given, that is, these switching signals $\sigma(0),\sigma(1),\ldots,\sigma(t_j-1)$, $1\leq t_j\leq T$, are fixed. Then, the matrices $\mathbf{D}(1),\mathbf{D}(2),\ldots,\mathbf{D}(t_j)$ are determined by Eqs. (7) and (8). So we denote

$$L_{t_j}=\sum_{t=1}^{t_j}\mathbf{y}^\intercal\mathbf{D}^\intercal(t)\mathbf{Q}(t)\mathbf{D}(t)\mathbf{y},$$

If $t_j = T$, a complete switching sequence $\sigma = (\sigma(0), \sigma(1), \ldots, \sigma(T-1))$ is obtained. We can directly calculate the cost functional value. Otherwise, the cost functional (9) is divided into two parts as follows

$$\begin{aligned} & J(\sigma \mid \sigma(0), \ldots, \sigma(t_j - 1)) \\ =& L_{t_j} + \sum_{t=t_j+1}^{T} \mathbf{y}^\intercal \mathbf{D}^\intercal(t)\mathbf{Q}(t)\mathbf{D}(t)\mathbf{y}, \end{aligned} \tag{10}$$

where the second term of the right hand side of Eq. (10) is unknown because $\mathbf{D}(t_j+1), \ldots, \mathbf{D}(T)$ are not given. Thus, we should construct a lower bound of Eq. (10) as a selection rule to determine the rest of unknown switching signals. However, the matrix $\mathbf{D}(t+1)$ depends on the joint effect of the current matrix $\mathbf{D}(t)$ and the time-delay matrix $\mathbf{D}(t-\tau)$. There are some cross terms in the expansion of the cost functional, such as $\mathbf{D}^\intercal(t-\tau)\mathbf{B}_\sigma^\intercal(t)\mathbf{Q}(t+1)\mathbf{A}_\sigma(t)\mathbf{D}(t)$. Clearly, its positive semi-definiteness can not be guaranteed. The analysis of the lower bound becomes very difficult.

In order to handle these difficulties to obtain a unified lower bound for all scenarios of $\sigma(t_j), \sigma(t_j+1), \ldots, \sigma(T-1) \in \mathcal{M}$, we have the theorem below.

Theorem 1 *Assume that the first t_j switching signals $\sigma(0), \ldots, \sigma(t_j-1)$ are given in $\sigma \in \mathcal{M}^T$. For each $k = 0, 1, \ldots, T-t_j-1$, $l \in \{0, 1, \ldots, k\}$, we denote*

$$\mathbf{P}_{t_j}^k = \begin{bmatrix}\mathbf{D}^\intercal(t_j) & \mathbf{D}^\intercal(t_j-\tau) & \ldots & \mathbf{D}^\intercal(t_j-\tau+k)\end{bmatrix}^\intercal,$$

$$\mathbf{\Phi}_{\sigma(t_j+l)}^k = \begin{bmatrix}\mathbf{A}_\sigma(t_j+l) & \mathbf{B}_\sigma(t_j+l) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{(k-l)n},\end{bmatrix},$$

and for any $h \in \{0, 1, \ldots, k-\tau\}$ when $k \geq \tau$, let

$$\mathbf{\Psi}_{\sigma(t_j+h)}^k = \begin{bmatrix}\mathbf{A}_\sigma(t_j+h) & \mathbf{B}_\sigma(t_j+h) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{(\tau-1)n} \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0}\end{bmatrix}.$$

Then, it holds that

$$\mathbf{D}(t_j+k+1) = \begin{cases} (\prod_{l=0}^{k} \mathbf{\Phi}_{\sigma(t_j+k-l)}^k)\mathbf{P}_{t_j}^k, & 0 \leq k < \tau, \\ (\prod_{l=0}^{\tau-1} \mathbf{\Phi}_{\sigma(t_j+k-l)}^k)(\prod_{h=0}^{k-\tau} \mathbf{\Psi}_{\sigma(t_j+k-\tau-h)}^k)\mathbf{P}_{t_j}^{\tau-1}, & k \geq \tau, \end{cases} \tag{11}$$

where the dimensions of the block matrices $\mathbf{P}_{t_j}^k \in \mathbb{R}^{(k+2)n \times (\tau+1)n}$ and $\mathbf{\Phi}_{\sigma(t_j+l)}^k \in \mathbb{R}^{(k-l+1)n \times (k-l+2)n}$ are extensible with respect to k and l.

PROOF. First, let us consider the first case of Eq. (11) when $0 \leq k < \tau$ and the induction method is used.

When $k = 0$, we obtain $l = 0$, then

$$\begin{aligned} & \mathbf{D}(t_j+1) \\ =& \begin{bmatrix}\mathbf{A}_\sigma(t_j) & \mathbf{B}_\sigma(t_j)\end{bmatrix} \begin{bmatrix}\mathbf{D}(t_j) \\ \mathbf{D}(t_j-\tau)\end{bmatrix} = \mathbf{\Phi}_{\sigma(t_j)}^0 \mathbf{P}_{t_j}^0. \end{aligned}$$

It means that Eq. (11) is true in the case of $k = 0$.

Assume that Eq. (11) also holds when $k = r-1 \geq 0$, where $1 \leq r \leq \tau$, let us consider the case of $k = r$. From the definitions of $\mathbf{P}_{t_j}^r$ and $\mathbf{\Phi}_{\sigma(t_j+r-l)}^r$, we obtain the relations

$$\mathbf{P}_{t_j}^r = \begin{bmatrix}\mathbf{P}_{t_j}^{r-1} \\ \mathbf{D}(t_j+r-\tau)\end{bmatrix}, \mathbf{\Phi}_{\sigma(t_j+r-l)}^r = \begin{bmatrix}\mathbf{\Phi}_{\sigma(t_j+r-l)}^{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n\end{bmatrix}.$$

Then,

$$\begin{aligned} & \mathbf{D}(t_j+r+1) \\ =& \begin{bmatrix}\mathbf{A}_\sigma(t_j+r) & \mathbf{B}_\sigma(t_j+r)\end{bmatrix} \begin{bmatrix}\mathbf{D}(t_j+r) \\ \mathbf{D}(t_j+r-\tau)\end{bmatrix} \\ =& \mathbf{\Phi}_{\sigma(t_j+r)}^r \begin{bmatrix}\prod_{l=0}^{r-1} \mathbf{\Phi}_{\sigma(t_j+r-1-l)}^{r-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n\end{bmatrix} \begin{bmatrix}\mathbf{P}_{t_j}^{r-1} \\ \mathbf{D}(t_j+r-\tau)\end{bmatrix} \\ =& (\prod_{l=0}^{r} \mathbf{\Phi}_{\sigma(t_j+r-l)}^r)\mathbf{P}_{t_j}^r. \end{aligned}$$

Hence, the first case of Eq. (11) is always true. In fact, it also holds when $k \geq \tau$, then we rewrite $\mathbf{D}(t_j+k+1)$ as

$$\mathbf{D}(t_j+k+1) = (\prod_{l_1=0}^{\tau-1} \mathbf{\Phi}_{\sigma(t_j+k-l_1)}^k)(\prod_{l_2=\tau}^{k} \mathbf{\Phi}_{\sigma(t_j+k-l_2)}^k)\mathbf{P}_{t_j}^k.$$

Now we transform it into the second case of Eq. (11).

On the one hand, for any $0 \leq q \leq k-\tau$,

$$\begin{aligned} & \mathbf{\Phi}_{\sigma(t_j+q)}^k \mathbf{P}_{t_j+q}^{k-q} \\ =& \begin{bmatrix}\mathbf{A}_\sigma(t_j+q) & \mathbf{B}_\sigma(t_j+q) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{(k-q)n}\end{bmatrix} \begin{bmatrix}\mathbf{D}(t_j+q) \\ \mathbf{D}(t_j+q-\tau) \\ \vdots \\ \mathbf{D}(t_j-\tau+k)\end{bmatrix} \\ =& \begin{bmatrix}\mathbf{D}(t_j+q+1) \\ \mathbf{D}(t_j+q-\tau+1) \\ \vdots \\ \mathbf{D}(t_j-\tau+k)\end{bmatrix} = \mathbf{P}_{t_j+q+1}^{k-q-1}. \end{aligned}$$

By the recursion, we have

$$(\prod_{l_2=\tau}^{k} \mathbf{\Phi}^k_{\sigma(t_j+k-l_2)})\mathbf{P}^k_{t_j} = \mathbf{P}^{\tau-1}_{t_j+k-\tau+1}.$$

On the other hand,

$$\begin{aligned}\mathbf{\Psi}^k_{\sigma(t_j)}\mathbf{P}^{\tau-1}_{t_j} &= \begin{bmatrix} \mathbf{A}_\sigma(t_j) & \mathbf{B}_\sigma(t_j) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{(\tau-1)n} \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{D}(t_j) \\ \mathbf{D}(t_j-\tau) \\ \vdots \\ \mathbf{D}(t_j-1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{D}(t_j+1) \\ \mathbf{D}(t_j-\tau+1) \\ \vdots \\ \mathbf{D}(t_j) \end{bmatrix} = \mathbf{P}^{\tau-1}_{t_j+1},\end{aligned}$$

then we obtain

$$(\prod_{h=0}^{k-\tau} \mathbf{\Psi}^k_{\sigma(t_j+k-\tau-h)})\mathbf{P}^{\tau-1}_{t_j} = \mathbf{P}^{\tau-1}_{t_j+k-\tau+1}.$$

It shows that, for any $k \geq \tau$,

$$(\prod_{l_2=\tau}^{k} \mathbf{\Phi}^k_{\sigma(t_j+k-l_2)})\mathbf{P}^k_{t_j} =(\prod_{h=0}^{k-\tau} \mathbf{\Psi}^k_{\sigma(t_j+k-\tau-h)})\mathbf{P}^{\tau-1}_{t_j}.$$

Thus, we obtain the second case of Eq. (11) when $k \geq \tau$.

This completes the proof. □

For simplicity, we denote

$$\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k-l)} = \begin{cases} \mathbf{\Phi}^k_{\sigma(t_j+k-l)}, & 0 \leq l < \tau, \\ \mathbf{\Psi}^k_{\sigma(t_j+k-l)}, & \tau \leq l, \end{cases}$$

$$\hat{\mathbf{P}}^k_{t_j} = \begin{cases} \mathbf{P}^k_{t_j}, & 0 \leq k < \tau, \\ \mathbf{P}^{\tau-1}_{t_j}, & \tau \leq k, \end{cases}$$

then Eq. (11) can be rewritten by

$$\mathbf{D}(t_j+k+1) = (\prod_{l=0}^{k} \hat{\mathbf{\Phi}}^k_{\sigma(t_j+k-l)})\hat{\mathbf{P}}^k_{t_j}. \tag{12}$$

Hence, the cost functional (10) is equal to

$$\begin{aligned}&J(\sigma \mid \sigma(0),\dots,\sigma(t_j-1)) \\ =&L_{t_j} + \sum_{k=0}^{T-t_j-1} \mathbf{y}^\intercal(\hat{\mathbf{P}}^k_{t_j})^\intercal(\hat{\mathbf{\Phi}}^k_{\sigma(t_j)})^\intercal \cdots (\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)})^\intercal \cdot \\ &\cdot \mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)} \cdots \hat{\mathbf{\Phi}}^k_{\sigma(t_j)}\hat{\mathbf{P}}^k_{t_j}\mathbf{y}.\end{aligned} \tag{13}$$

Based on the positive semi-definiteness of matrix, we introduce the following theorem referred in [25].

Theorem 2 *For a positive semi-definite matrix $\mathbf{\Upsilon} \in \mathbb{R}^{n\times n}$, if a diagonal matrix $\mathbf{\Lambda} \in \mathbb{R}^{n\times n}$ is given by*

$$\mathbf{\Lambda} = \begin{bmatrix} \sum_{j=1}^{n} |\Upsilon_{1j}| & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_{j=1}^{n} |\Upsilon_{nj}| \end{bmatrix},$$

then $\mathbf{\Upsilon} \preceq \mathbf{\Lambda}$.

Corollary 1 *For a positive semi-definite matrix $\mathbf{\Upsilon} \in \mathbb{R}^{n\times n}$, let $\mathbf{\Upsilon}_\mu = \mathbf{\Upsilon} + \mu\mathbf{I}_n$, where $\mu > 0$ is a regularization parameter. Then, we can always obtain a diagonal matrix $\mathbf{\Gamma} \in \mathbb{R}^{n\times n}$, such that $\mathbf{\Gamma} \preceq \mathbf{\Upsilon}_\mu$.*

PROOF. For the positive semi-definite matrix $\mathbf{\Upsilon} \in \mathbb{R}^{n\times n}$ and $\mu > 0$, we know that $\mathbf{\Upsilon}_\mu = \mathbf{\Upsilon} + \mu\mathbf{I}_n$ is an invertible positive definite matrix. Then, its inverse matrix denoted by $\mathbf{\Upsilon}_\mu^{-1}$ is also positive definite. Based on Theorem 2, we can obtain a diagonal positive definite matrix $\tilde{\mathbf{\Lambda}}$, such that $\mathbf{\Upsilon}_\mu^{-1} \preceq \tilde{\mathbf{\Lambda}}$. Let $\mathbf{\Gamma} = \tilde{\mathbf{\Lambda}}^{-1}$, then it holds that $\mathbf{\Gamma} \preceq \mathbf{\Upsilon}_\mu$, where $\mathbf{\Gamma} \in \mathbb{R}^{n\times n}$ is a diagonal positive definite matrix. This completes the proof. □

Theorem 3 *Given the first t_j switching signals $\sigma(0)$, $\sigma(1),\dots,\sigma(t_j-1)$, $1 \leq t_j \leq T-1$, $k \in \{0,\dots,T-t_j-1\}$ and $\mu > 0$, we can always find a number $\Xi^k_\mu \in \mathbb{R}$, such that, for any $\sigma(t_j),\dots,\sigma(t_j+k) \in \mathcal{M}$,*

$$\begin{aligned}\Xi^k_\mu \leq& \mathbf{y}^\intercal(\hat{\mathbf{P}}^k_{t_j})^\intercal(\hat{\mathbf{\Phi}}^k_{\sigma(t_j)})^\intercal \cdots (\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)})^\intercal \cdot \\ &\cdot \mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)} \cdots \hat{\mathbf{\Phi}}^k_{\sigma(t_j)}\hat{\mathbf{P}}^k_{t_j}\mathbf{y}.\end{aligned} \tag{14}$$

PROOF. For the given first t_j switching signals $\sigma(0)$, $\dots,\sigma(t_j-1)$ and $k \in \{0,\dots,T-t_j-1\}$, $\hat{\mathbf{P}}^k_{t_j}$ can be determined. We just need to consider the formula

$$(\hat{\mathbf{\Phi}}^k_{\sigma(t_j)})^\intercal \cdots (\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)})^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)} \cdots \hat{\mathbf{\Phi}}^k_{\sigma(t_j)}.$$

Since $\mathbf{Q}(t_j+k+1) \in \mathbb{R}^{n\times n}$ is a positive semi-definite matrix, then for any $\sigma(t_j+k) \in \mathcal{M}$, the matrix

$$(\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)})^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)}$$

is also positive semi-definite. Based on Corollary 1, we obtain a diagonal matrix $\mathbf{\Gamma}^{k,1}_{\sigma(t_j+k)} \in \mathbb{R}^{2n\times 2n}$, such that

$$\mathbf{\Gamma}^{k,1}_{\sigma(t_j+k)} \preceq (\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)})^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}^k_{\sigma(t_j+k)} + \mu\mathbf{I}_{2n}.$$

Let $\sigma(t_j+k)=1,\dots,M$, a set of diagonal matrices $\mathbf{\Gamma}_1^{k,1}$, $\dots,\mathbf{\Gamma}_M^{k,1}$ are computed. We define a minimal diagonal matrix $\mathbf{\Gamma}_{\min}^{k,1}$, such that its i-th diagonal element satisfies

$$(\Gamma_{\min}^{k,1})_{ii}=\min\{(\Gamma_1^{k,1})_{ii},(\Gamma_2^{k,1})_{ii},\dots,(\Gamma_M^{k,1})_{ii}\},$$

where $i=1,2,\dots,2n$. Then, for any $\sigma(t_j+k)\in\mathcal{M}$,

$$\mathbf{\Gamma}_{\min}^{k,1}\preceq(\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k)^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k+\mu\mathbf{I}_{2n}.$$

Furthermore, for any $\sigma(t_j+k-1)\in\mathcal{M}$, we have

$$\begin{aligned}&(\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k)^\intercal\mathbf{\Gamma}_{\min}^{k,1}\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k+\mu\mathbf{I}_{3n}\preceq\\&(\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k)^\intercal(\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k)^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k\cdot\\&\cdot\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k+\mu((\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k)^\intercal\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k+\mathbf{I}_{3n}).\end{aligned}$$

Denote $\mathbf{\Lambda}_{\max}^{k,0}=\mathbf{I}_{2n}$, we construct two diagonal matrices $\mathbf{\Gamma}_{\sigma(t_j+k-1)}^{k,2},\mathbf{\Lambda}_{\sigma(t_j+k-1)}^{k,1}\in\mathbb{R}^{3n\times3n}$ by Corollary 1 and Theorem 2, respectively, such that

$$\begin{aligned}&\mathbf{\Gamma}_{\sigma(t_j+k-1)}^{k,2}\preceq(\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k)^\intercal\mathbf{\Gamma}_{\min}^{k,1}\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k+\mu\mathbf{I}_{3n}\\&\mathbf{\Lambda}_{\sigma(t_j+k-1)}^{k,1}\succeq(\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k)^\intercal\mathbf{\Lambda}_{\max}^{k,0}\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k+\mathbf{I}_{3n}.\end{aligned}$$

Let $\sigma(t_j+k-1)=1,\dots,M$, then we obtain a minimal diagonal matrix $\mathbf{\Gamma}_{\min}^{k,2}\in\mathbb{R}^{3n\times3n}$ and a maximal diagonal matrix $\mathbf{\Lambda}_{\max}^{k,1}\in\mathbb{R}^{3n\times3n}$, such that for each $i=1,\dots,3n$,

$$\begin{aligned}&(\Gamma_{\min}^{k,2})_{ii}=\min\{(\Gamma_1^{k,2})_{ii},(\Gamma_2^{k,2})_{ii},\dots,(\Gamma_M^{k,2})_{ii}\},\\&(\Lambda_{\max}^{k,1})_{ii}=\max\{(\Lambda_1^{k,1})_{ii},(\Lambda_2^{k,1})_{ii},\dots,(\Lambda_M^{k,1})_{ii}\}.\end{aligned}$$

Thus, for any $\sigma(t_j+k-1),\sigma(t_j+k)\in\mathcal{M}$, it holds that

$$\begin{aligned}\mathbf{\Gamma}_{\min}^{k,2}\preceq&(\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k)^\intercal(\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k)^\intercal\mathbf{Q}(t_j+k+1)\cdot\\&\cdot\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k\hat{\mathbf{\Phi}}_{\sigma(t_j+k-1)}^k+\mu\mathbf{\Lambda}_{\max}^{k,1}.\end{aligned}$$

By the recursion, we obtain a series of minimal diagonal matrices $\mathbf{\Gamma}_{\min}^{k,3},\dots,\mathbf{\Gamma}_{\min}^{k,k+1}$ and maximal diagonal matrices $\mathbf{\Lambda}_{\max}^{k,2},\dots,\mathbf{\Lambda}_{\max}^{k,k}$. Then, for any $\sigma(t_j),\dots,\sigma(t_j+k)\in\mathcal{M}$, we have

$$\begin{aligned}&\mathbf{\Gamma}_{\min}^{k,k+1}\preceq(\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k)^\intercal\mathbf{\Gamma}_{\min}^{k,k}\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k+\mu\mathbf{I}_{\lambda n}\\&\preceq(\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k)^\intercal\cdots(\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k)^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k\cdots\\&\quad\cdot\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k+\mu((\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k)^\intercal\mathbf{\Lambda}_{\max}^{k,k-1}\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k)+\mathbf{I}_{\lambda n})\\&\preceq(\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k)^\intercal\cdots(\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k)^\intercal\mathbf{Q}(t_j+k+1)\hat{\mathbf{\Phi}}_{\sigma(t_j+k)}^k\cdots\\&\quad\cdot\hat{\mathbf{\Phi}}_{\sigma(t_j)}^k+\mu\mathbf{\Lambda}_{\max}^{k,k}.\end{aligned}$$

where

$$\lambda=\begin{cases}k+2, & 0\le k<\tau,\\ \tau+1, & \tau\le k.\end{cases}$$

Hence, we just let

$$\Xi_\mu^k=\mathbf{y}^\intercal(\hat{\mathbf{P}}_{t_j}^k)^\intercal(\mathbf{\Gamma}_{\min}^{k,k+1}-\mu\mathbf{\Lambda}_{\max}^{k,k})\hat{\mathbf{P}}_{t_j}^k\mathbf{y},$$

then for any $\sigma(t_j),\dots,\sigma(t_j+k)\in\mathcal{M}$, Eq. (14) holds.

This completes the proof. □

It follows from Theorem 3, the lower bound of Eq. (13) with the regularization parameter $\mu>0$ is expressed by

$$LB(\sigma(0),\dots,\sigma(t_j-1))=L_{t_j}+\sum_{k=0}^{T-t_j-1}\Xi_\mu^k. \tag{15}$$

3.3 Depth-first branch and bound method

The depth-first branch and bound method is an improved version of tree search method, which can search all global optimal solutions efficiently. But its searching principle depends on the comparison between the lower bound and the current optimal value. In the time-delay switched system, we use the dynamical lower bound expression (15) to implement the branch and bound method. The searching process is summarized as follows.

Algorithm 1 (Branch and bound method)

Step 0: (Initialization) *Given the parameters T, M, τ, $\mathbf{Q}(t)$ and $\mathbf{y}=[\mathbf{y}_0^\intercal\ \ \mathbf{y}_1^\intercal\ \ \dots\ \ \mathbf{y}_\tau^\intercal]^\intercal$. Set $J_{\min}=+\infty$, $\mu>0$ and $t:=0$.*

Step 1: (Sorting lower bounds) *Compute the lower bounds $LB(\sigma(0),\dots,\sigma(t))$ with $\sigma(t)=1$ to M and sort them by ascending rule in $\gamma_1(t),\dots,\gamma_M(t)$. Set $k_t:=1$, go to Step 2.*

Step 2: (Bounding) *Determine $\sigma(t):=\gamma_{k_t}(t)$. If $LB(\sigma(0),\dots,\sigma(t))>J_{\min}$ or $k_t=M$, set $t:=t-1$ and go to Step 5. If $LB(\sigma(0),\dots,\sigma(t))\le J_{\min}$, go to Step 3.*

Step 3: (Branching) *Set $t:=t+1$. If $t=T-1$, go to Step 4. Otherwise, go to Step 1.*

Step 4: (Update the current optimal value) *Choose $\sigma(t)$ from 1 to M and compute $J(\sigma)$. Update the current optimal value by $J_{\min}:=J(\sigma)$ and store the solution σ if $J(\sigma)\le J_{\min}$. Set $t:=t-1$, go to Step 5.*

Step 5: (Terminate) *If $t<0$, terminate and output the optimal solution σ. Otherwise, set $k_t:=k_t+1$, go to Step 2.*

The following theorem states that the global optimal solution can be searched by Algorithm 1.

Theorem 4 *The solution obtained by Algorithm 1 is the global optimal switching sequence.*

PROOF. We prove it using proof by contradiction. Let the solution obtained by Algorithm 1 be $\sigma^* = (\sigma^*(0), \sigma^*(1), \ldots, \sigma^*(T-1))$ and assume that it is not the global optimal solution. It means that there is a better solution $\tilde{\sigma} = (\tilde{\sigma}(0), \tilde{\sigma}(1), \ldots, \tilde{\sigma}(T-1))$, such that $J(\tilde{\sigma}) < J(\sigma^*)$. Then, there is at least one switching signal $\tilde{\sigma}(t) \neq \sigma^*(t)$ at some time t. Without loss of generality, we assume that the first t_j $(t = 0, 1, \ldots, t_j - 1)$ switching signals of σ^* and $\tilde{\sigma}$ are the same. Suppose that $\tilde{\sigma}(t_j) = \gamma_m(t_j)$ and $\sigma^*(t_j) = \gamma_l(t_j)$, where $m, l \in \mathcal{M}$. If $m < l$, σ^* is searched later than $\tilde{\sigma}$ by the ascending rule of lower bounds in Step 1 of Algorithm 1. Then, the solution σ^* is better than or equals to $\tilde{\sigma}$, that is, $J(\sigma^*) \leq J(\tilde{\sigma})$. It is a contradiction. Otherwise, when $m > l$, there are the following two cases,

$$\begin{aligned} &LB(\tilde{\sigma}(0), \ldots, \tilde{\sigma}(t_j - 1), \gamma_m(t_j)) \\ \geq &J(\sigma^*) \geq LB(\sigma^*(0), \ldots, \sigma^*(t_j - 1), \sigma^*(t_j)) \end{aligned}$$

or

$$J_{\min} = J(\sigma^*) > LB(\tilde{\sigma}(0), \ldots, \tilde{\sigma}(t_j - 1), \gamma_m(t_j)).$$

For the first case, we have $J(\tilde{\sigma}) \geq J(\sigma^*)$. This is also a contradiction. For the second case, the solution can be further branched. Similarly, we can obtain that

$$\begin{aligned} J_{\min} = J(\sigma^*) > LB(\tilde{\sigma}(0), \ldots, \tilde{\sigma}(t-1), \gamma_m(t)), \\ \forall t_j < t \leq T - 1. \end{aligned}$$

It means that the solution $\tilde{\sigma}$ will be searched in Algorithm 1. Then, the optimal solution can be updated at least by $J_{\min} \leq J_{\tilde{\sigma}} < J_{\sigma^*}$. It contradicts the optimality of σ^*. Hence, the assumption does not hold and the global optimal solution can be obtained by Algorithm 1.

This completes the proof. □

4 Numerical examples

In this section, we first consider two optimal switching problems of switched systems with time delay to illustrate the efficiency of the proposed method. Then an application of this system is given. All codes are compiled and executed in Matlab 2012a on a Windows 7 laptop with 2.5 GHz CPU and 4G RAM.

Example 1 *Consider the time-invariant switched system with the time delay $\tau = 3$ and the terminal time $T = 6$, where there are 10 subsystems with the parameters $(\mathbf{A}_i, \mathbf{B}_i)$, $i = 1, 2, \ldots, 10$ given as follows:*

$$\left(\begin{bmatrix}1&0&1\\0&1&0\\1&0&1\end{bmatrix}, \begin{bmatrix}-1&1&1\\1&-1&1\\0&0&0\end{bmatrix}\right), \left(\begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix}, \begin{bmatrix}1&1&1\\0&0&0\\-1&-1&1\end{bmatrix}\right),$$

$$\left(\begin{bmatrix}1&0&1\\1&0&1\\0&1&0\end{bmatrix}, \begin{bmatrix}0&0&0\\1&-1&1\\-1&1&1\end{bmatrix}\right), \left(\begin{bmatrix}0&1&0\\0&1&0\\1&0&1\end{bmatrix}, \begin{bmatrix}1&1&0\\-1&1&0\\1&-1&0\end{bmatrix}\right),$$

$$\left(\begin{bmatrix}0&1&0\\1&0&1\\1&0&1\end{bmatrix}, \begin{bmatrix}-1&0&1\\1&0&1\\1&0&-1\end{bmatrix}\right), \left(\begin{bmatrix}1&0&1\\0&1&0\\0&1&0\end{bmatrix}, \begin{bmatrix}0&1&-1\\0&-1&1\\0&1&1\end{bmatrix}\right),$$

$$\left(\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, \begin{bmatrix}1&1&-1\\1&0&-1\\0&1&0\end{bmatrix}\right), \left(\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}, \begin{bmatrix}0&1&0\\-1&0&1\\1&-1&1\end{bmatrix}\right),$$

$$\left(\begin{bmatrix}0&1&0\\0&1&0\\0&1&0\end{bmatrix}, \begin{bmatrix}-1&1&0\\1&0&1\\1&-1&0\end{bmatrix}\right), \left(\begin{bmatrix}0&0&0\\1&1&1\\0&0&0\end{bmatrix}, \begin{bmatrix}0&1&-1\\1&0&1\\0&1&-1\end{bmatrix}\right).$$

The initial states $\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ and $\mathbf{Q}$ are given by

$$\begin{bmatrix}1\\2\\2\end{bmatrix}, \begin{bmatrix}2\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}5&3&2\\3&2&1\\2&1&3\end{bmatrix}.$$

In this problem, there are totally 10^6 switching sequences. Since the enumeration is the only method for the optimal switching problem of the time-delay switched system, we use it to determine the global optimal solution, which takes about 354.16 seconds and obtains the optimal solution $\sigma = (10, 3, 9, 10, 1, 9)$. The corresponding optimal value is $J(\sigma) = 704$. We list the distribution of the cost functional values in Table 1. It can be seen that most of the cost functional values are greater than 3000.

Table 1
The distribution of cost functional values in Example 1.

Range of cost values	Number of solutions
704	1
$(704, 10^3]$	33
$(10^3, 3 \times 10^3]$	5701
$(3 \times 10^3, 8 \times 10^3]$	104996
$(8 \times 10^3, 2 \times 10^4]$	380701
$(2 \times 10^4, 5 \times 10^4]$	404862
$(5 \times 10^4, 10^5]$	95368
$(10^5, 354698]$	8338

Next, we apply the branch and bound method to solve this problem. Table 2 shows the searching efficiency of the method with different regularization parameters. When μ approaches zero, the algorithm is more efficient. For the case of $\mu = 10^{-4}$, the method only searches 230 feasible solutions, which takes about 6.046 seconds. That is, only 0.23% switching sequences are searched. The optimal solution and the optimal value are the same given by the enumeration.

Table 2
Branch and bound method with different μ in Example 1.

μ	Searching times	Time	Optimal value
1	1110	71.253s	704
0.5	930	44.257s	704
10^{-1}	800	26.895s	704
10^{-4}	230	6.046s	704

The searching process of our proposed method is depicted in Figure 1. It can be seen that the current optimal value decreases gradually and the range of cost functional values is only limited to $[704, 7678]$. That is, most of insignificant points are pruned. So our method is effective to search the global optimal solution.

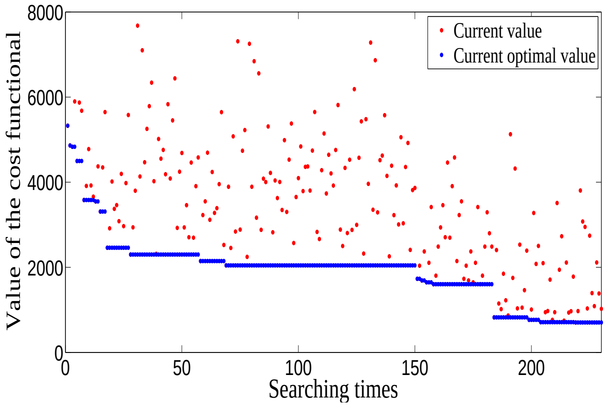

Fig. 1. The searching process of our proposed method in Example 1.

Example 2 *Consider the time-varying switched system with the time delay $\tau = 2$ and the terminal time $T = 12$, where there are 4 subsystems with the parameters $(\mathbf{A}_i(t), \mathbf{B}_i(t))$, $i = 1, 2, 3, 4$, given as follows:*

$$\left(\begin{bmatrix} \sin t & 0 & 0 \\ 0 & 2 & \cos t \\ 1+0.5t & 0 & \cos t \end{bmatrix}, \begin{bmatrix} \sin t & 0 & 0 \\ 0 & -\sin t & 0 \\ 0 & 0 & \cos t \end{bmatrix}\right),$$

$$\left(\begin{bmatrix} 1+\cos t & 0 & 0 \\ 0.3t & \sin t & 0 \\ 0 & \cos t & 2+\sin t \end{bmatrix}, \begin{bmatrix} \cos t & 0 & 0 \\ 0 & \sin t & 0 \\ 0 & -\cos t & 1 \end{bmatrix}\right),$$

$$\left(\begin{bmatrix} 0.6 & 0 & 1-\sin t \\ 0 & t & \cos t \\ \cos t & 0 & t \end{bmatrix}, \begin{bmatrix} \sin t & 0 & 0 \\ 0 & \cos t & 0 \\ 0 & 0 & \sin t \end{bmatrix}\right),$$

$$\left(\begin{bmatrix} 0.4t+1 & 0 & 3 \\ 1 & \cos t - 1 & 0 \\ \sin t & 0 & \sin t \end{bmatrix}, \begin{bmatrix} -\cos t & 0 & 0 \\ 0 & \cos t & 0 \\ 0 & 0 & \sin t \end{bmatrix}\right).$$

The initial states $\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2$ and $\mathbf{Q}(t)$ are given by

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1+\cos t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1+\sin t \end{bmatrix}.$$

In this problem, the number of feasible switching sequences is $4^{12} \approx 1.6777 \times 10^7$. First, we use the enumeration to find the global optimal solution. It takes about 5512.57 seconds to obtain the optimal switching sequence as $\sigma = (3, 3, 1, 1, 2, 1, 1, 3, 1, 2, 2, 4)$ and the optimal value as $J(\sigma) = 146.3816$. The distribution of the cost functional values is listed in Table 3. In fact, the performance of the cost functional is in the range $[146.3816, 1.7401 \times 10^{18}]$. Note that there exists only one global optimal solution.

Table 3
The distribution of cost functional values in Example 2.

Range of cost values	Number of solutions
146.3816	1
$(146.3816, 10^3]$	254
$(10^3, 10^4]$	6488
$(10^4, 10^5]$	55957
$(10^5, 10^6]$	257778
$(10^6, 10^7]$	778576
$(10^7, 10^8]$	1697970
$(10^8, +\infty)$	13980192

For comparison, we apply the branch and bound method to solve this problem and set $\mu = 10^{-10}$. It only searches 96 feasible solutions to determine the same global optimal solution, which takes only about 6.09 seconds. The searching process of the method is depicted in Figure 2. It can be seen that most of feasible solutions are ignored. Thus, the proposed method is also efficient in the time-varying switched system.

Example 3 *Consider the dynamical model of the supply chain with time delay, which is proposed in [26]. More details can be referred to [4,27,28]. We suppose that a firm produces two kinds of products and denote the amounts of the production, sale and inventory of product j by p_j, $s_j(t)$ and $i_j(t)$ at period t, respectively, where $j = 1, 2$. Their advertisement costs are $a_1 = 12$ and $a_2 = 17$,*

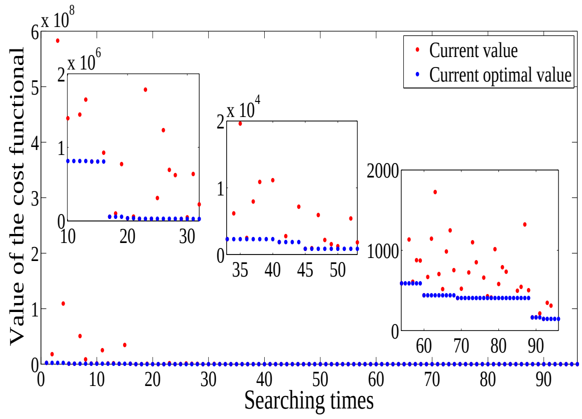

Fig. 2. The searching process of our proposed method in Example 2.

respectively. In order to reduce the cost and the inventory level, the firm have the following four production modes.

$$\begin{cases} s_1(t+1) = 0.7s_1(t) + 0.2s_2(t-2) + a_1, \\ s_2(t+1) = 0.1s_1(t-2) + 0.5s_2(t) + a_2, \\ i_j(t+1) = i_j(t) + p_j - s_j(t+1), \quad j = 1, 2, \end{cases}$$

where $p_1 = 0$ *or* 40, $p_2 = 0$ *or* 80. *That means each product has two cases: no production and normal production.*

Let

$$\begin{aligned} \mathbf{x}(t) &= [s_1(t) \;\; s_2(t) \;\; i_1(t) \;\; i_2(t) \;\; 1]^\intercal, \\ \mathbf{b} &= [a_1 \;\; a_2 \;\; p_1 - a_1 \;\; p_2 - a_2]^\intercal, \end{aligned}$$

then the supply chain system can be formulated by

$$\mathbf{x}(t+1) = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \mathbf{x}(t-2),$$

where $\mathbf{b}$ *takes value from the following four modes:*

$$\begin{aligned} \mathbf{b}_1 &= [12 \;\; 17 \;\; 28 \;\; 63]^\intercal, \\ \mathbf{b}_2 &= [12 \;\; 17 \;\; -12 \;\; 63]^\intercal, \\ \mathbf{b}_3 &= [12 \;\; 17 \;\; 28 \;\; -17]^\intercal, \\ \mathbf{b}_4 &= [12 \;\; 17 \;\; -12 \;\; -17]^\intercal, \end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ -0.7 & 0 & 1 & 0 \\ 0 & -0.5 & 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 0.2 & 0 & 0 \\ 0.1 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 \\ -0.1 & 0 & 0 & 0 \end{bmatrix}.$$

The initial states are given by

$$\begin{aligned} \mathbf{x}(-2) &= [0 \;\; 0 \;\; 20 \;\; 0 \;\; 1]^\intercal, \\ \mathbf{x}(-1) &= [0 \;\; 0 \;\; 20 \;\; 0 \;\; 1]^\intercal, \\ \mathbf{x}(0) &= [-10 \;\; 0 \;\; 20 \;\; 0 \;\; 1]^\intercal, \end{aligned}$$

and the maximal and minimal values of the state $\mathbf{x}(t)$ *at period* t *are, respectively, given by*

$$\begin{aligned} \mathbf{x}_{\max}(t) &= [+\infty \;\; +\infty \;\; 50 \;\; 500 \;\; 1]^\intercal, \\ \mathbf{x}_{\min}(t) &= [-20 \;\; -200 \;\; -50 \;\; -500 \;\; 1]^\intercal. \end{aligned}$$

The purpose is to minimize the following cost functional

$$J = \sum_{t=1}^{9} \mathbf{x}^\intercal(t)\mathbf{Q}\mathbf{x}(t),$$

where $T = 9$, *and*

$$\mathbf{Q} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In the system, the number of scheduling sequences is $4^9 = 262144$. We set $\mu = 10^{-6}$ and apply the branch and bound method to solve this problem. It takes about 13.374 seconds to search 862 feasible solutions and obtains the global optimal solution as $\sigma = (4, 2, 3, 3, 1, 3, 1, 3, 1)$. That is, only 0.329% scheduling sequences are searched. The optimal value is $J(\sigma) = 5.3353 \times 10^4$. Thus, the proposed method is very efficient in the application.

5 Conclusion

The optimal switching problem for linear switched systems in discrete time with time delay is considered in this paper. The purpose is to design an optimal switching sequence, such that a quadratic cost functional is minimized. As a particular class of optimal control problem, it is NP-complete and the gradient information disappears. To find the global optimal solution, we reformulate the original problem with the matrix transformation and extension techniques to reduce the complexity caused by the time delay. A set of lower bounds of the optimal value are constructed for the cost functional. Then, we develop a depth-first branch and bound method to solve this problem, which is verified in the optimal switching problems of time-invariant and time-varying switched time-delay systems. Numerical results have been illustrated to show that the proposed method is efficient and effective.

Acknowledgements

This paper is supported by the grant of National Natural Science Foundation of China (Nos. 11431004, 11671250), the grant of Humanity and Social Science

Foundation of Ministry of Education of China (No. 15YJA630034), the grant of the Startup Foundation for Introducing Talent of NUIST (No. 1091181901002), the grant of Chongqing Science and Technology Commission (No. cstc2017jcyjAX0161), the grant of Chongqing Normal University (No. 17XLB010), and program for scientific research start-up funds of Guangdong Ocean University. The fourth author is supported by PolyU grant ZZGS.

References

[1] Kirches, C., Sager, S., Bock, H. G., & Schlöder, J. P. (2010). Time-optimal control of automobile test drives with gear shifts. *Optimal Control Applications and Methods*, 31, 137-153.

[2] Woon, S. F., Rehbock, V., & Loxton, R. C. (2010). Global optimization method for continuous-time sensor scheduling. *Nonlinear Dynamics and Systems Theory*, 10(2), 175-188.

[3] Vazquez, S., Marquez, A., Aguilera, R., Quevedo, D., Leon, J. I., & Franquelo, L. G. (2015). Predictive optimal switching sequence direct power control for grid-connected power converters. *IEEE Transactions on Industrial Electronics*, 62(4), 2010-2020.

[4] Zhang, S., Zhang, C., Zhang, S., & Zhang, M. (2018). Discrete switched model and fuzzy robust control of dynamic supply chain network. *Complexity*, 9, 1-11.

[5] Zhao, X., Zhang, L., & Shi, P. (2013). Stability of a class of switched positive linear time-delay systems. *International Journal of Robust and Nonlinear Control*, 23(5), 578-589.

[6] Zhang, L., Zhuang, S., & Braatz, R. D. (2016). Switched model predictive control of switched linear systems: Feasibility, stability and robustness. *Automatica*, 67, 8-21.

[7] Girard, A., Gössler, G., & Mouelhi, S. (2016). Safety controller synthesis for incrementally stable switched systems using multiscale symbolic models. *IEEE Transactions on Automatic Control*, 61(6), 1537-1549.

[8] Liu, J., Zhang, K., Pang, G., & Wei, H. (2016). Controller synthesis for constrained discrete-time switched positive linear systems. *Nonlinear Analysis Hybrid Systems*, 19, 1-12.

[9] Xu, X., & Antsaklis, P. J. (2002). Optimal control of switched autonomous systems. In *Proc. of the 41st IEEE conf. on dec. and cont.*, Vol. 4 (pp. 4401-4406).

[10] Egerstedt, M., Wardi, Y., & Delmotte, F. (2003). Optimal control of switching times in switched dynamical systems. In *Proc. of the 42nd IEEE conf. on dec. and cont.*, Vol. 3 (pp. 2138-2143).

[11] Egerstedt, M., Wardi, Y., & Axelsson, H. (2006). Transition-time optimization for switched-mode dynamical systems. *IEEE Transactions on Automatic Control*, 51(1), 110-115.

[12] Xu, X., & Antsaklis P. J. (2004). Optimal control of switched systems based on parameterization of the switching instants. *IEEE Transactions on Automatic Control*, 49(1), 2-16.

[13] Loxton, R., Teo, K. L., & Rehbock, V. (2008). Optimal control problems with multiple characteristic time points in the objective and constraints. *Automatica*, 44(11), 2923-2929.

[14] Loxton, R., Lin, Q., & Teo, K. L. (2013). Minimizing control variation in nonlinear optimal control. *Automatica*, 49(9), 2652-2664.

[15] Lin, Q., Loxton, R., & Teo, K. L. (2014). The control parameterization method for nonlinear optimal control: A survey. *Journal of Industrial and Management Optimization*, 10(1), 275-309.

[16] Bengea, S. C., & DeCarlo, R. A. (2005). Optimal control of switching systems. *Automatica*, 41, 11-27.

[17] Das, T., & Mukherjee, R. (2008). Optimally switched linear systems. *Automatica*, 44, 1437-1441.

[18] Zhai, J., Niu, T., Ye, J., & Feng, E. (2017). Optimal control of nonlinear switched system with mixed constraints and its parallel optimization algorithm. *Nonlinear Analysis Hybrid Systems*, 25, 21-40.

[19] Axelsson, H., Wardi, Y., Egerstedt, M., & Verriest, E. I. (2008). Gradient descent approach to optimal mode scheduling in hybrid dynamical systems. *Journal of Optimization Theory and Applications*, 136(2), 167-186.

[20] Wardi, Y., Egerstedt, M., & Hale, M. (2015). Switched-mode systems: Gradient-descent algorithms with Armijo step sizes. *Discrete Event Dynamic Systems*, 25(4), 571-599.

[21] Caldwell, T. M., & Murphey, T. D. (2016). Projection-based iterative mode scheduling for switched systems. *Nonlinear Analysis Hybrid Systems*, 21, 59-83.

[22] Esteban, H. V., Patrizio, C., Richard, M., & Franco, B. (2011). Discrete-time control for switched positive systems with application to mitigating viral escape. *International Journal of Robust and Nonlinear Control*, 21, 1093-1111.

[23] Feng, Z. G., Teo, K. L., & Rehbock, V. (2009). A discrete filled function method for the optimal control of switched systems in discrete time. *Optimal Control Applications and Methods* , 30(6), 585-593.

[24] Gao, J., & Li, D. (2012). Linear-quadratic switching control with switching cost. *Automatica*, 48, 1138-1143.

[25] Xu, W., Feng, Z. G., Peng, J. W., & Yiu, K. F. C. (2017). Optimal switching for linear quadratic problem of switched systems in discrete time. *Automatica*, 78, 185-193.

[26] Tamura, H. (1975). Decentralized optimization for distributed-lag models of discrete systems. *Automatica*, 11(6), 593-602.

[27] Ignaciuk, P., & Bartoszewicz, A. (2012). Linear-quadratic optimal control of periodic-review perishable inventory systems. *IEEE Transactions on Control Systems Technology*, 20(5), 1400-1407.

[28] Dashkovskiy, S., Karimi, H. R., & Kosmykov, M. (2013). Stability analysis of logistics networks with time-delays. *Production Planning and Control*, 24(7), 567-574.