MINIMIZING ALMOST SMOOTH CONTROL VARIATION IN NONLINEAR OPTIMAL CONTROL PROBLEMS

YING ZHANG
Department of Mathematics
Zhejiang Normal University
Jinhua 321004, Zhejiang, China

CHANGJUN YU*
Department of Mathematics
Shanghai University
Baoshan 200444, Shanghai, China

YINGTAO XU
Xingzhi College
Zhejiang Normal University
Jinhua 321004, Zhejiang, China

YANQIN BAI
Department of Mathematics
Shanghai University
Baoshan 200444, Shanghai, China

(Communicated by Cedric Yiu)

ABSTRACT. In this paper, we consider an optimal control problem in which the control is almost smooth and the state and control are subject to terminal state constraints and continuous state and control inequality constraints. By introducing an extra set of differential equations for this almost smooth control, we transform this constrained optimal control problem into an equivalent problem involving both control function and system parameter vector as decision variables. Then, by the control parametrization technique and a time scaling transformation, the equivalent problem is approximated by a sequence of constrained optimal parameter selection problems, each of which is a finite dimensional optimization problem. For each of these constrained optimal parameter selection problems, a novel exact penalty function method is constructed by appending penalized constraint violations to the cost function. This gives rise to a sequence of unconstrained optimal parameter selection problems; and each of which can be solved by existing optimization algorithms or software packages. Finally, a practical container crane operation problem is solved, showing the effectiveness and applicability of the proposed approach.

2010 Mathematics Subject Classification. Primary: 58F15, 58F17; Secondary: 53C35.

Key words and phrases. Optimal control, control variation, almost smooth control.

This paper is supported by NSFC grant 11871039, 11771275, the Scientific Research Project of Zhejiang Provincial Department of science and technology in China (Grant No. LGN19C040001), and the Scientific Research Project of Zhejiang Provincial Department of Education in China (Grant No. Y201329106).
* Corresponding author: Changjun Yu.
1. **Introduction.** Optimal control theory has many successful applications in engineering, physical and medical sciences, such as soft landing for moon lander [15], robotics [4], zinc sulphate electrolyte purification process [9], hybrid control system [6], switched system [22] and fund management [3]. In practice, the optimal control problems are often subject to various constraints, which include boundedness constraints on the control variables, terminal state constraints, and continuous state and control inequality constraints. There are many computational methods available in the literature for solving various classes of optimal control problems. They include control parametrization method [19], direct transcription method [16], discretization method [5]. For a large number of the optimal control problems, the control functions are assumed to be essentially bounded measurable functions or L-2 functions. There is no smoothness condition imposed. In many practical optimal control problems arising in real-world scenarios, a discontinuous control signal is allowed. However, in some engineering applications, discontinuous control signals should be avoided because abrupt changes in the control signal could entail infinite jerk at the switching points, and excite a large bandwidth of vibration modes causing structural vibration. Thus, for these applications, continuous control is much preferred.

In [20], an optimal control problem with almost smooth controls is considered. This optimal control problem is subject to boundedness constraints on the control variables, terminal state constraints, and continuous state and control inequality constraints. An additional differential equation with a new control appearing in its right hand side is introduced, where the original control becomes a new state and its initial condition becomes a system parameter to be chosen optimally. Thus, an equivalent optimal control problem involving both control function and system parameter, which is referred to as a combined optimal control and optimal parameter selection problem, is obtained, where the (new) control function is not required to be almost smooth. Clearly, the original control function will be continuous and piecewise linear if the new control function is approximated by a piecewise constant function. The control parameterization method [19] is used to solve this equivalent combined optimal control and optimal parameter selection problem, where the control function is approximated by a piecewise constant function and its heights are considered as decision variables. For the continuous state and control inequality constraints, they are handled by the constraint transcription technique.

By applying the control parameterization method to an optimal control problem, the time horizon is partitioned into a number of subintervals, and the control is approximated by a piecewise constant function consistent with this partition. The discontinuities of this piecewise constant function can only occur at the partition points, also called the switching times. The heights of this piecewise constant function are regarded as decision variables. In this way, the optimal control problem is approximated by an optimal parameter selection problem, which is a finite dimensional optimization problem. Clearly, a better solution would be obtained if the switching times of this piecewise constant function are also considered as decision variables. However, some drawbacks will be encountered during the numerical computation (see [19] for details). The time scaling transformation was first introduced for time optimal control problems in [11] and generalized to more general optimal control problems in [2, 12, 13]. It is introduced to circumvent the difficulties caused by regarding switching times as decision variables. It works by mapping the variable
switching times to fixed points in a new time horizon, thus yielding a new optimization problem in which the switching times are fixed. The transformed problem can then be solved more readily by existing gradient-based optimization algorithms.

In this paper, the time scaling transformation will be applied to supplement the control parameterization technique to solve the optimal control problem with almost smooth controls. The exact penalty function method [24, 25], rather than the constraint transcription technique, will be used to handle the continuous state and control inequality constraints. A gradient-based computational method will be developed to solve the optimal control problem under consideration.

The remainder of the paper is organized as follows. Following this introduction, we first introduce an extra set of differential equations for the almost smooth control, then the original problem in Section 2 can be written equivalently as a combined optimal control and optimal parameter selection problem in Section 3, where the initial conditions for the extra set of differential equations are regarded as system parameter vector. Then, in Section 4, after the application of the control parameterization technique together with a time scaling transformation, the problem is approximated by a sequence of constrained optimal parameter selection problems, i.e. optimal control problems with piecewise constant controls and fixed switching times subject to terminal state constraints and continuous state and control inequality constraints. To solve this new problem, we propose, in Section 5, an exact penalty function to construct a corresponding unconstrained optimal parameter selection problem, which can be solved by standard gradient-based optimization methods, such as sequential quadratic programming method. The existing optimal control software packages, such as VISUAL MISER [8], can be used. In Section 6, the method proposed is applied to solve a nontrivial optimal control problem. Finally, some conclusions are given in Section 7.

2. Problem formulation. Consider a process described by the following system differential equations:

\[ \dot{x}(t) = f(t, x(t), u(t)), \quad t \in [0, T], \]  
\[ x(0) = x^0, \]  
\[ (1a) \]
\[ (1b) \]

where \( x(t) \in \mathbb{R}^n \) is the state vector at time \( t \), \( u(t) \in \mathbb{R}^r \) is the control vector at time \( t \), \( x^0 \in \mathbb{R}^n \) is the initial state, \( T > 0 \) is a given terminal time, and \( f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n \) is a given continuously differentiable function.

Define

\[ U = \{ \nu = [v_1, \ldots, v_r]^\top \in \mathbb{R}^r : \alpha_i \leq v_i \leq \beta_i, i = 1, \ldots, r \}, \]  
\[ (2) \]

where \( \alpha_i, i = 1, \ldots, r \), and \( \beta_i, i = 1, \ldots, r \), are given real numbers. Note that \( U \) is a compact and convex subset of \( \mathbb{R}^r \).

**Definition 2.1.** A function \( u : [0, T] \to \mathbb{R}^r \) is said to be almost smooth if it is continuous and is differentiable almost everywhere.

Let \( u \) be an almost smooth function defined on \([0, T]\) with values in \( U \), and let \( \dot{u} \) denote the derivative of \( u \). If

\[ c_i \leq \dot{u}_i(t) \leq d_i, \quad \forall t \in [0, T], \quad i = 1, \ldots, r, \]  
\[ (3) \]

where \( c_i, i = 1, \ldots, r \), and \( d_i, i = 1, \ldots, r \), are given real numbers, then \( u \) is called an admissible control. Let \( U \) be the class of all such admissible controls. Furthermore,
let \( \mathcal{U}^o \) be a subset of the set \( \mathcal{U} \) defined by
\[
\mathcal{U}^o = \{ u \in \mathcal{U} : \alpha_i < u_i(t) < \beta_i, \forall t \in [0, T], i = 1, \ldots, r \}. \tag{4}
\]

Note that the derivative \( \dot{u} \) of \( u \) is, in fact, only defined almost everywhere in \([0, T]\). However, we may assign appropriate values for the function \( \dot{u} \) at those points in \([0, T]\) at which the function \( \dot{u} \) is not defined so that the extended function satisfies the condition \( (3) \). Throughout this paper, the function \( \dot{u} \) is to be understood as its extended counterpart. For each \( u \in \mathcal{U} \), let \( x(\cdot | u) \) be the corresponding solution of system \( (1) \).

We assume that \( f : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n \) in system \( (1) \) satisfies the following linear growth condition: there exists a constant \( M > 0 \), such that
\[
\| f(t, x, u) \| \leq M (1 + \| x \|) \tag{5}
\]
for all \( (t, x, u) \in [0, T] \times \mathbb{R}^n \times \mathcal{U} \), where \( \| \cdot \| \) denotes a norm of \( \mathbb{R}^n \). This assumption is standard in control theory (see, for example \cite{1}). Condition \( (2) \) and \( (5) \) ensure that there exists a unique solution of system \( (1) \) corresponding to each \( u \in \mathcal{U} \) (see Theorem 3.3.3 of \cite{1}).

Now, we formally state our optimal control problem as follows.

**Problem \( (P_1) \).** Given the dynamic system \( (1) \), choose a feasible control \( u \in \mathcal{U} \) such that the cost function
\[
g_0(u) = \Phi_0(x(T|u)) + \int_0^T L_0(t, x(t|u), u(t)) dt \tag{6}
\]
is minimized subject to the following terminal state constraints
\[
\Phi_i(x(T|u)) \leq 0, \quad i = 1, \ldots, N_T, \tag{7}
\]
and the continuous inequality constraints
\[
h_i(t, x(t|u), u(t)) \leq 0, \quad \forall t \in [0, T], \quad i = 1, \ldots, N_S, \tag{8}
\]
where the functions \( \Phi_0 : \mathbb{R}^n \to \mathbb{R}, L_0 : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}, \Phi_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, N_T \) and \( h_i : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}, i = 1, \ldots, N_S \) are assumed to be continuously differentiable with respect to each of their arguments.

Note that control bounds can be easily incorporated into \( (8) \) and the admissible control is required to be almost smooth. Thus, it is allowed to appear in the continuous inequality constraints \( (8) \). This is a slight generalization of that considered in Chapter 8 in \cite{20}.

Define
\[
\Omega = \{ u \in \mathcal{U} : \Phi_i(x(T|u)) \leq 0, i = 1, \ldots, N_T \} \tag{9}
\]
and
\[
\mathcal{F} = \{ u \in \Omega : h_i(t, x(t|u), u(t)) \leq 0, \forall t \in [0, T], i = 1, \ldots, N_S \}. \tag{10}
\]
Controls in \( \mathcal{F} \) are called feasible controls, and \( \mathcal{F} \) is called the class of feasible controls. We assume that \( \mathcal{F} \) is not empty. For convenience, let \( \Omega^o \) (respectively, \( \mathcal{F}^o \)) be the interiors of \( \Omega \) (respectively, \( \mathcal{F} \)), i.e., \( \Omega^o \) (respectively, \( \mathcal{F}^o \)) are all those \( u \) such that \( (9) \) (respectively, \( (10) \)) are satisfied as strict inequalities.

To proceed further, we assume that the following condition which was first introduced in \cite{18} is satisfied.

**Assumption 2.1.** For any \( u \in \mathcal{F} \), there exists a \( \bar{u} \in \Omega^o \) such that
\[
\alpha \bar{u} + (1 - \alpha)u \in \mathcal{F}^o, \quad \forall \alpha \in (0, 1].
\]
This condition has been widely used in the literature.
3. Model transformation. In this section, we introduce an extra set of differential equations for the control $u$: 

$$\dot{u}(t) = v(t)$$

(11a)

with the initial conditions:

$$u(0) = \xi$$

(11b)

where $v = [v_1, \ldots, v_r]^T \in \mathbb{R}^r$ and $\xi = [\xi_1, \ldots, \xi_r]^T \in \mathbb{R}^r$.

In view of (11a), we see that $u$ is now a state function rather than a control function, and is determined by the new control function $v$ and the initial vector $\xi$. For convenience, let $\xi$ be referred to as the system parameter vector. Clearly, for a given system parameter vector $\xi$, if $v$ is approximated by a piecewise constant function, then $u$ will be a piecewise linear function.

Define

$$\hat{x} = [x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+r}]^T,$$

where $x_{n+i} = u_i, \ i = 1, \ldots, r$.

Then, by adding (11) to (1), we have the following state differential equations:

$$\dot{\hat{x}}(t) = \hat{f}(t, \hat{x}(t), v(t)), \quad t \in [0, T],$$

(12a)

$$\hat{x}(0) = \hat{x}^0(\xi),$$

(12b)

where $\hat{f} = [f_1, \ldots, f_n, v_1, \ldots, v_r]^T$ and $\hat{x}^0(\xi) = [x_1^0, \ldots, x_n^0, \xi_1, \ldots, \xi_r]^T$.

Considering the conditions of (2)-(3), the following constraints must be satisfied.

$$\alpha_i \leq \xi_i \leq \beta_i, \quad i = 1, \ldots, r,$$

(13)

$$\alpha_i \leq x_{n+i}(t) \leq \beta_i, \quad \forall t \in [0, T], \quad i = 1, \ldots, r,$$

(14)

and

$$c_i \leq v_i(t) \leq d_i, \quad \forall t \in [0, T], \quad i = 1, \ldots, r.$$  

(15)

For convenience, let $Z$ and $V$ be, respectively, the sets of $\xi$ and $v$ such that constraints (13) and (15) are satisfied. Then, in this section, $v$ is called an admissible control, and $V$ is called the set of admissible controls. Furthermore, $\xi$ is called a system parameter vector, and $Z$ is called the set of system parameter vectors. Let the trajectory $\hat{x}(\cdot|\xi, v)$ be the solution of the system (12) corresponding to $(\xi, v) \in Z \times V$.

We write the constraints (14) as:

$$\tilde{h}_{N_s+i}(t, \hat{x}(t|\xi, v)) = x_{n+i}(t|\xi, v) - \beta_i \leq 0, \quad \forall t \in [0, T], \quad i = 1, \ldots, r$$

(16)

and

$$\tilde{h}_{N_{s}+r+i}(t, \hat{x}(t|\xi, v)) = \alpha_i - x_{n+i}(t|\xi, v) \leq 0, \quad \forall t \in [0, T], \quad i = 1, \ldots, r.$$

Also, the terminal state constraints (7) and the continuous inequality constraints (8) can be written as:

$$\Phi_i(\hat{x}(T|\xi, v)) = \Phi_i(x(T|u)) \leq 0, \quad i = 1, \ldots, N_T,$$

and

$$\hat{h}_i(t, \hat{x}(t|\xi, v)) = h_i(t, x(t|u), u(t)) \leq 0, \quad \forall t \in [0, T], \quad i = 1, \ldots, N_s,$$

respectively.

Define

$$\tilde{\Omega} = \{ (\xi, v) \in Z \times V : \tilde{h}_i(\hat{x}(T|\xi, v)) \leq 0, i = 1, \ldots, N_T \},$$

(16)
and
\[
\tilde{F} = \{ (\xi, v) \in \hat{\Omega} : \tilde{h}_i(t, \tilde{x}(t|\xi, v)) \leq 0, \forall t \in [0, T], i = 1, \ldots, N_{S+2r} \}.
\] (17)

The combined optimal control and optimal parameter selection problem under consideration may now be stated formally as follows.

**Problem** \((P_2)\). Given the dynamic system (12), choose a feasible combined parameter vector and control \((\xi, v) \in \tilde{F}\) such that the cost function
\[
\tilde{G}_0(\xi, v) = \tilde{\Phi}_0(\tilde{x}(T|\xi, v)) + \int_0^T \tilde{L}_0(t, \tilde{x}(t|\xi, v))dt
\] (18)
is minimized over \(\tilde{F}\), where
\[
\tilde{\Phi}_0(\tilde{x}(T|\xi, v)) = \Phi_0(x(T|u))
\] and
\[
\tilde{L}_0(t, \tilde{x}(t|\xi, v)) = L_0(t, x(t|u), u(t)),
\]
where \(\Phi_0(x(T|u))\) and \(L_0(t, x(t|u), u(t))\) are given in (6) which is the cost function of Problem \((P_1)\).

We note that Problem \((P_1)\) is equivalent to Problem \((P_2)\). Even so, we present this trivial result in the following as a theorem for future reference.

**Theorem 3.1.** Problem \((P_1)\) is equivalent to Problem \((P_2)\) in the sense that for each \(u \in F\) there corresponds uniquely a \((\xi, v) \in \tilde{F}\) such that
\[
g_0(u) = \tilde{G}_0(\xi, v),
\]
and vice versa.

4. **Control parametrization and time scaling transformation.** In this section, we will introduce a solution scheme for solving Problem \((P_2)\). Our method involves two stages. The first stage is the control parametrization technique introduced in [19]. By partitioning the time horizon into a finite subintervals, and approximating the control function by a piecewise constant function consistent with the partition, Problem \((P_2)\) is transformed into an approximate optimization problem \((P_3)\), where both the heights and the partition points, known as switching points, are regarded as decision variables. The second stage is the application of the time scaling transformation introduced in [13]. The essential idea is to transform the approximate problem \((P_3)\) with variable switching times into a new problem with fixed switching times in a new time horizon.

4.1. **Control parametrization.** We approximate the control signal \(v\) by piecewise constant basis functions as follows:
\[
v^p(t|\tau, \sigma) = \sum_{k=1}^{p} \sigma^k \chi_{[\tau_{k-1}, \tau_k)}(t), \quad t \in [0, T),
\] (19)
where \(p \geq 1\) is a given number of the control subintervals, and \(\chi_{[\tau_{k-1}, \tau_k)} : \mathbb{R} \rightarrow [0, 1]\) is the characteristic function defined by
\[
\chi_{[\tau_{k-1}, \tau_k)}(t) = \begin{cases} 
1, & \text{if} \quad t \in [\tau_{k-1}, \tau_k), \\
0, & \text{otherwise}.
\end{cases}
\]
\(\tau \in \mathcal{T}\) is a switching time vector satisfying
\[
\mathcal{T} = \{ \tau \in \mathbb{R}^{p-1} : 0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq \tau_{p-1} \leq \tau_p = T \},
\]
and $\sigma \in Q$ is a vector consisting of the heights of the approximate piecewise constant function. It satisfies

$$Q = \{ \sigma = [\sigma^1, \ldots, \sigma^p]^\top \in \mathbb{R}^{p'} : \sigma^k \in \mathbb{R}, k = 1, \ldots, p \}.$$  

Note that $v^p(\cdot|\tau, \sigma)$ switches value at times $t = \tau_k, k = 1, \ldots, p-1$. Furthermore, we note that (19) does not define a value of $v^p(\cdot|\tau, \sigma)$ at $t = T$. For convenience, we define

$$v^p(T|\tau, \sigma) = v^p(T^-|\tau, \sigma) = \lim_{t \to T^-} v^p(t|\tau, \sigma).$$

Considering the condition on $v$ in (15), we obtain the bound constraints on $\sigma$ as follows:

$$c_i \leq \sigma^k_i \leq d_i, \quad k = 1, \ldots, p, \quad i = 1, \ldots, r.$$  

Let $Q^p$ be the set of all those $\sigma \in Q$ which satisfy the constraints (20).

Substituting the piecewise constant control (19) into the dynamic system (12) yields

$$\dot{x}(t) = \sum_{k=1}^{p} \tilde{f}(t, \tilde{x}(t), \sigma^k \chi_{[\tau_{k-1}, \tau_k]}(t), \quad t \in [0, T],$$  

$$\tilde{x}(0) = \tilde{x}^0(\xi).$$  

Define

$$\delta = (\xi, \sigma).$$

The constraints appearing in (16) and (17) become

$$\tilde{\Phi}_i(\tilde{x}^p(T|\tau, \delta)) \leq 0, \quad i = 1, \ldots, N_T,$$

and

$$\sum_{k=1}^{p} \tilde{h}_i(t, \tilde{x}^p(t|\tau, \delta), \sigma^k \chi_{[\tau_{k-1}, \tau_k]}(t) \leq 0, \quad \forall t \in [0, T], \quad i = 1, \ldots, N_{S+2r}.$$  

Define

$$\Lambda = \{ (\tau, \delta) \in T \times (Z \times Q^p) : \tilde{\Phi}_i(\tilde{x}^p(T|\tau, \delta)) \leq 0, i = 1, \ldots, N_T \}$$

and

$$\Xi = \{ (\tau, \delta) \in \Lambda : \sum_{k=1}^{p} \tilde{h}_i(t, \tilde{x}^p(t|\tau, \delta), \sigma^k \chi_{[\tau_{k-1}, \tau_k]}(t) \leq 0, \forall t \in [0, T], i = 1, \ldots, N_{S+2r} \}.$$  

Then, we can approximate Problem $(P_2)$ by the finite-dimensional optimization problem $(P_3)$ as follows.

**Problem $(P_3)$**. Given the dynamic system (21), choose a feasible pair $(\tau, \delta) \in \Xi$ such that the cost function

$$\tilde{G}_0^p(\tau, \delta) = \tilde{\Phi}_0(\tilde{x}^p(T|\tau, \delta)) + \sum_{k=1}^{p} \int_{\tau_{k-1}}^{\tau_k} \tilde{L}_0(t, \tilde{x}^p(t|\tau, \delta), \sigma^k)dt$$

is minimized over $\Xi$, where $\tilde{\Phi}_0$ (respectively, $\tilde{L}_0$) is obtained from $\Phi_0$ (respectively, $L_0$) in an obvious manner.
The following convergence properties of this piecewise constant approximation scheme are based on the arguments in [14]. In this paper, we concern more on the practical computation of a control policy, rather than the convergence results.

**Theorem 4.1.** Let \( v^{p,*} \) be an optimal control of Problem \((P_3)\), and \( v^* \) be an optimal control of Problem \((P_2)\). Then,

\[
\lim_{p \to +\infty} \tilde{G}_0(v^{p,*}) = \tilde{G}_0(v^*).
\]

**Theorem 4.2.** Let \( v^{p,*} \) be an optimal control of Problem \((P_3)\), and \( v^* \) be an optimal control of Problem \((P_2)\). Suppose that

\[
\lim_{p \to +\infty} v^{p,*} = \bar{v}, \text{ a.e. on } [0,T].
\]

Then, \( \bar{v} \) is an optimal control of Problem \((P_2)\), and

\[
\lim_{p \to +\infty} \tilde{G}_0(v^{p,*}) = \tilde{G}_0(v^*).
\]

Let \((\tau^*, \delta^*) \in \Xi\) be an optimal solution of Problem \((P_3)\), then, we can define the following piecewise constant control for Problem \((P_2)\):

\[
v^{p,*}(t) = \begin{cases} 
\sum_{k=1}^{p} \sigma_k \chi_{[\tau_{k-1}, \tau_k]}(t), & \text{if } t \in [0,T), \\
\nu^{p,*}(T^-), & \text{if } t = T,
\end{cases}
\]

where \(\tau_0 = 0\) and \(\tau_p = T\).

Based on the convergence properties stated in Theorem 4.1 and Theorem 4.2, we see that \(v^{p,*}\) minimizes (18) over the space of feasible piecewise constant controls. That is, a suboptimal control for Problem \((P_2)\) can be obtained by solving Problem \((P_3)\) with the switching times being considered as fixed. Clearly, using the switching times as decision variables will yield better solutions, as the search space is larger. However, the variable switching times may cause difficulties during numerical computation in the execution of any gradient-based optimization algorithm (see [14] for details). In the next subsection, we will use the novel time scaling transformation technique [13] for converting Problem \((P_3)\) into an equivalent optimization problem with fixed switching times in a new time horizon. Thus, existing gradient-based optimization procedures can be applied to solve it.

4.2. **Time scaling transformation.** To apply the time scaling transformation, we first introduce a new time variable \(s\), then construct a transformation from \(t \in [0,T]\) to \(s \in [0,p]\) which maps the variable switching time points, \(0 = \tau_0, \tau_1, \tau_2, \ldots, \tau_{p-1}, \tau_p = T\), into the pre-fixed switching time points, \(s = 0, 1, 2, \ldots, p-1, p\), in a new time scale.

Define

\[
\Theta = \{\theta = [\theta_1, \theta_2, \ldots, \theta_p]^T \in \mathbb{R}^p : \theta_k = \tau_k - \tau_{k-1} \geq 0, k = 1, 2, \ldots, p, \sum_{k=1}^{p} \theta_k = T\},
\]  

(26)

Then, the time scaling function \(\mu(\cdot|\theta) : [0,p] \to \mathbb{R}\) is defined by:

\[
\mu(s|\theta) = \begin{cases} 
\sum_{m=1}^{\lfloor s \rfloor} \theta_m + \theta_{\lfloor s \rfloor+1}(s - \lfloor s \rfloor), & \text{if } s \in [0,p), \\
T, & \text{if } s = p.
\end{cases}
\]  

(27)
It is easy to see that (27) defines \( \mu(-|\theta) \) as a non-negative, non-decreasing and continuous piecewise linear function of \( s \).

For each \( \theta \in \Theta \), define
\[
\hat{\tau}(\theta) = [\mu(1|\theta), \mu(2|\theta), \ldots, \mu(p - 1|\theta)]^T \in \mathbb{R}^{p-1}.
\]

Clearly, we have
\[
\mu(k|\theta) = \sum_{m=1}^{k} \theta_m = \tau_k, \quad k = 0, 1, \ldots, p.
\]

This means that the time scaling transformation defined by (27) maps \( s = k \) to the \( k \)th switching time \( t = \tau_k \). Then, \( \hat{\tau}(\theta) \in T \) is a valid switching time vector for Problem \((P_3)\).

Using the time substitution \( t = \mu(s|\theta) \), where \( \mu(s|\theta) \) is the time scaling function defined by (27), we obtain the following new state vector:
\[
\tilde{x}^p(s) := \tilde{x}^p(\mu(s|\theta)), \quad s \in [0, p].
\]

For notational simplicity, let \( \mu(s) = \mu(s|\theta) \). Then, applying (27) to the dynamic system (21), we have
\[
\begin{align*}
\hat{\dot{x}}^p(s) &= \sum_{k=1}^{p} \theta_k \tilde{f}(\mu(s), \tilde{x}^p(s), \sigma^k)\chi_{[k-1,k)}(s), \quad s \in [0, p], \quad (28a) \\
\hat{x}^0(0) &= \hat{x}^0(\xi), \quad (28b)
\end{align*}
\]

where the overhead dot now represents differentiation with respect to \( s \), \( \hat{x}^0(\xi) \) and \( \tilde{f} \) are defined as in Section 3, and \( \chi_{[k-1,k)} : \mathbb{R} \rightarrow \mathbb{R} \) is the characteristic function defined in Subsection 4.1.

Let \( \hat{x}^p(\cdot|\theta, \delta) \) be the solution of system (28) corresponding to \( \theta \in \Theta \) and \( \delta = (\xi, \sigma) \in Z \times Q^p \). Notice that the switching time points in system (28) occur at the pre-fixed locations \( s = 1, 2, \ldots, p - 1 \).

Applying the time scaling transformation (27), the constraints appearing in (23) and (24) become
\[
\hat{\Phi}_i(\hat{x}^p(p|\theta, \delta)) \leq 0, \quad i = 1, \ldots, N_T, \quad (29)
\]

and
\[
\sum_{k=1}^{p} \theta_k \hat{h}_i(\mu(s), \tilde{x}^p(s|\theta, \delta), \sigma^k)\chi_{[k-1,k)}(s) \leq 0, \quad \forall s \in [0, p], \quad i = 1, \ldots, N_{S+2r}. \quad (30)
\]

Similarly, the cost function becomes
\[
\hat{G}^0_{\hat{\eta}}(\theta, \delta) := \hat{\Phi}_0(\hat{x}^p(p|\theta, \delta)) + \sum_{k=1}^{p} \int_{k-1}^{k} \theta_k \hat{L}_0(\mu(s), \tilde{x}^p(s|\theta, \delta), \sigma^k)ds.
\]

Define
\[
\hat{\Lambda} = \{(\theta, \delta) \in \Theta \times (Z \times Q^p) : \hat{\Phi}_i(\hat{x}^p(p|\theta, \delta)) \leq 0, \quad i = 1, \ldots, N_T \}
\]

and
\[
\hat{\Xi} = \{(\theta, \delta) \in \hat{\Lambda} : \sum_{k=1}^{p} \theta_k \hat{h}_i(\mu(s), \tilde{x}^p(s|\theta, \delta), \sigma^k)\chi_{[k-1,k)}(s) \leq 0, \quad \forall s \in [0, p], \quad i = 1, \ldots, N_{S+2r} \}.
\]

Then, Problem \((P_3)\) can be transformed into Problem \((P_4)\) as follows.
Theorem 4.3. Assume that $\text{(P)}$ is equivalent to Problem (4), considered as a state variable. The next theorem reveals that Problem (4) is equivalent to Problem ($P_3$).

**Theorem 4.3.** Assume that $\theta \in \Theta$. Furthermore, let $\hat{x}^p(\cdot|\theta, \delta)$ be the solution of system (28) corresponding to $(\theta, \delta)$, and let $\tilde{x}^p(\cdot|\tilde{\tau}(\theta), \delta)$ be the solution of system (21) corresponding to $(\tilde{\tau}(\theta), \delta)$. Then

$$
\hat{x}^p(\theta, \delta) = \tilde{x}^p(\mu(s)|\tilde{\tau}(\theta), \delta).
$$

**Proof.** For notational simplicity, we write

$$
\mu(\cdot) = \mu(\cdot|\theta), \quad \hat{x}^p(\cdot) = \hat{x}^p(\cdot|\tilde{\tau}(\theta), \delta), \quad \tilde{y}^p(\cdot) = \tilde{x}^p(\mu(\cdot)|\tilde{\tau}(\theta), \delta).
$$

This notation will not cause confusion because $\theta$ and $\delta$ are considered fixed throughout this proof.

Let $k \in \{1, \ldots, p\}$ be a fixed integer. Then

$$
\mu(s) = \sum_{m=1}^{k-1} \theta_m + \theta_k (s-k+1), \quad s \in [k-1, k].
$$

Consider the following two cases:

1. If $\theta_k = 0$, then

$$
\mu(s) = \mu(k-1) = \theta_1 + \ldots + \theta_{k-1}, \quad s \in [k-1, k],
$$

and thus

$$
\tilde{y}^p(s) = \tilde{y}^p(k-1), \quad s \in (k-1, k).
$$

This implies that

$$
\hat{y}^p(s) = 0 = \theta_k \hat{f}(\mu(s), \tilde{y}^p(s), \sigma_k), \quad s \in (k-1, k).
$$

2. If $\theta_k > 0$, then $\mu(\cdot)$ is strictly increasing on $[k-1, k]$. We have

$$
\mu(k-1) < \mu(s) < \mu(k), \quad s \in (k-1, k),
$$

and

$$
\hat{\mu}(s) = \theta_k, \quad s \in (k-1, k).
$$

Thus, by taking differentiation of $\tilde{y}^p(\cdot)$ with respect to $s$ and then using the chain rule, we obtain

$$
\hat{y}^p(s) = \frac{d(\tilde{y}^p(s))}{ds} = \frac{d(\hat{x}^p(\mu(s)))}{ds} = \frac{d(\hat{x}^p(\mu(s)))}{dt} \cdot \frac{d(\mu(s))}{ds}
$$

$$
= \theta_k \hat{f}(\mu(s), \tilde{x}^p(\mu(s)), \sigma_k), \quad s \in (k-1, k)
$$

$$
= \theta_k \hat{f}(\mu(s), \tilde{y}^p(s), \sigma_k), \quad s \in (k-1, k).
$$

Combining the two cases, it follows that, regardless of whether $\theta_k = 0$ or $\theta_k > 0$,

$$
\hat{y}^p(s) = \theta_k \hat{f}(\mu(s), \tilde{y}^p(s), \sigma_k), \quad s \in (k-1, k).
$$

(33)
Also,
\[
\tilde{y}^P(0) = \tilde{x}^P(\mu(0|\theta)) = \tilde{x}^P(0) = \tilde{x}^0(\xi).
\] (34)

From (33) and (34), we conclude that \(\tilde{y}^P(\cdot)\) is the solution of system (28).

Thus,
\[
\tilde{y}^P(s) = \tilde{x}^P(s|\theta, \delta).
\]

However, we recall that
\[
\tilde{y}^P(s) = \tilde{x}^P(\mu(s|\theta)|\tilde{\tau}(\theta), \delta).
\]

Thus,
\[
\tilde{x}^P(s|\theta, \delta) = \tilde{x}^P(\mu(s|\theta)|\tilde{\tau}(\theta), \delta).
\]

The proof is completed.

Theorem 4.3 demonstrates how the time scaling function \(\mu(\cdot|\theta)\) links the dynamic system (21) in Problem \((P_3)\) with the dynamic system (28) in Problem \((P_4)\). In the following, Theorem 4.4 links the feasible regions between Problem \((P_3)\) and Problem \((P_4)\).

**Theorem 4.4.** Assume that \(\theta \in \Theta\). Then,
\[
\tilde{G}^P_0(\theta, \delta) = \tilde{G}^P_o(\tilde{\tau}(\theta), \delta),
\] (35)
\[
\tilde{\Phi}_i(\tilde{x}^P(p|\theta, \delta)) = \tilde{\Phi}_i(\tilde{x}^P(T|\tilde{\tau}(\theta), \delta)), \quad i = 1, \ldots, N_T,
\] (36)

and
\[
\sum_{k=1}^p \theta_k \tilde{h}_i(\mu(s), \tilde{x}^P(s|\theta, \delta), \sigma^k)\chi_{[k-1,k)}(s) = \sum_{k=1}^p \tilde{h}_i(t, \tilde{x}^P(t|\tilde{\tau}(\theta), \delta), \sigma^k)\chi_{[\tau_{k-1}, \tau_k)}(t),
\]
\[
i = 1, \ldots, N_S+2r.
\] (37)

**Proof.** For notational simplicity, similar to Theorem 4.3, we write
\[
\mu(\cdot) = \mu(\cdot|\theta), \quad \tilde{x}^P(\cdot) = \tilde{x}^P(\cdot|\tilde{\tau}(\theta), \delta), \quad \tilde{x}^P(\cdot) = \tilde{x}^P(\cdot|\theta, \delta).
\]

Using Theorem 4.3, we have
\[
\tilde{G}^P_0(\tilde{\tau}(\theta), \delta) = \tilde{\Phi}_0(\tilde{x}^P(T)) + \sum_{k=1}^p \int_{\tau_{k-1}(\theta)}^{\tilde{\tau}_k(\theta)} \tilde{L}_0(t, \tilde{x}^P(t), \sigma^k)dt
\]
\[
= \tilde{\Phi}_0(\tilde{x}^P(t)|_{t=\mu(p)}) + \sum_{k=1}^p \int_{\mu(k-1)}^{\mu(k)} \tilde{L}_0(t, \tilde{x}^P(t), \sigma^k)dt
\]
\[
= \tilde{\Phi}_0(\tilde{x}^P(p)) + \sum_{k=1}^p \int_{k-1}^k \theta_k \tilde{L}_0(\mu(s), \tilde{x}^P(s), \sigma^k)ds
\]
\[
= \tilde{G}^P_0(\theta, \delta).
\]

It implies that equation (35) is valid.

Equations (36) and (37) can be proved similarly. This completes the proof.

The next result shows that a solution of Problem \((P_4)\) can be used to generate a solution of Problem \((P_3)\).

**Theorem 4.5.** Assume that \(\theta^* \in \Theta\). Then \((\theta^*, \delta^*)\) is an optimal solution for Problem \((P_4)\) if and only if \((\tilde{\tau}(\theta^*), \delta^*)\) is an optimal solution for Problem \((P_3)\).
Proof. Suppose that \((\theta^*, \delta^*)\) is optimal for Problem \((P_4)\). Then
\[
\Phi_i(\hat{x}^p(p|\theta^*, \delta^*)) \leq 0, \quad i = 1, \ldots, N_T,
\]
\[
\sum_{k=1}^{p} \theta_k^* \hat{h}_k(\mu(s), \hat{x}^p(s|\theta^*, \delta^*), \sigma^k) \chi_{[k-1,k)}(s) \leq 0, \quad \forall s \in [0, p], \quad i = 1, \ldots, N_{S+2r}.
\]
Thus, by using Theorem 4.4, we have
\[
\Phi_i(\hat{x}^p(T|\hat{\tau}(\theta^*), \delta^*)) = \Phi_i(\hat{x}^p(p|\theta^*, \delta^*)) \leq 0, \quad i = 1, \ldots, N_T,
\]
\[
\sum_{k=1}^{p} \hat{h}_k(t, \hat{x}^p(t|\hat{\tau}(\theta^*), \delta^*), \sigma^k) \chi_{[\tau_{k-1}, \tau_k)}(t)
\]
\[
= \sum_{k=1}^{p} \theta_k^* \hat{h}_k(\mu(s), \hat{x}^p(s|\theta^*, \delta^*), \sigma^k) \chi_{[k-1,k)}(s) \leq 0, \quad i = 1, \ldots, N_{S+2r}.
\]
This shows that \((\hat{\tau}(\theta^*), \delta^*)\) is feasible for Problem \((P_3)\).

Now, let \((\tau, \delta)\) denote an arbitrary but fixed feasible pair for Problem \((P_3)\). Define \(\theta = [\theta_1, \ldots, \theta_p]^T\) as follows:
\[
\theta_k = \tau_k - \tau_{k-1}, \quad k = 1, \ldots, p.
\]

Then, \(\mu(k|\theta) = \tau_k\), implying that \(\tau = \hat{\tau}(\theta)\). Furthermore,
\[
\theta_k \geq 0, k = 1, \ldots, p \quad \text{and} \quad \theta_1 + \theta_2 + \ldots + \theta_p = \tau_p = T.
\]

Then, by Theorem 4.4, \((\theta, \delta)\) is feasible for Problem \((P_4)\). Therefore,
\[
\hat{G}_0^p(\hat{\tau}(\theta^*), \delta^*) = \hat{G}_0^p(\theta^*, \delta^*) \leq G_0^p(\theta, \delta) = \hat{G}_0^p(\hat{\tau}(\theta), \delta) = \hat{G}_0^p(\tau, \delta).
\]
Since \((\tau, \delta)\) is arbitrary, it is clear that \((\hat{\tau}(\theta^*), \delta^*)\) is an optimal solution for Problem \((P_3)\).

The proof of the reverse implication is similar.

From Theorems 4.4 and Theorem 4.5, the equivalence between Problem \((P_3)\) and Problem \((P_2)\) is readily seen.

**Corollary 1.** Problem \((P_3)\) and Problem \((P_4)\) are equivalent.

**Remark 1.** Given an optimal control problem \((P_2)\), we first apply the piecewise constant functions to obtain an approximate problem \((P_3)\) with variable switching times. The next step is to use the time scaling transformation \(t = \mu(s|\theta)\), where \(\mu(|\theta)\) is the time scaling function, to transform Problem \((P_3)\) into an equivalent problem \((P_4)\) with fixed switching times. Then, Problem \((P_4)\) is solved by using any existing gradient-based computational method to obtain a solution \((\theta^*, \delta^*)\).

In fact, if \((\theta^*, \delta^*)\) is a solution of Problem \((P_4)\), then the optimal switching times for Problem \((P_3)\) are:
\[
\tau_k^* = \sum_{m=1}^{k} \theta_m^*, \quad k = 1, 2, \ldots, p - 1.
\]
Then, the corresponding suboptimal control for Problem \((P_2)\) is:
\[
u^{p,*}(t) = \sum_{k=1}^{p} \sigma^{k,*} \chi_{[\tau_{k-1}, \tau_k^*)}(t).
\]
5. **An exact penalty method for Problem** \((P_4)\). Note that Problem \((P_4)\) can be solved indirectly through solving Problem \((P_3)\), which is an optimization problem with fixed switching times. The solution of Problem \((P_3)\) is used to generate a solution of Problem \((P_4)\). It is then used to construct a corresponding suboptimal solution of Problem \((P_3)\). However, Problem \((P_4)\) involves continuous state inequality constraints, which are required to be satisfied at every point in the time horizon (there are infinite number of time points). Thus, frequently used optimization algorithms are not directly applicable. In this section, we will introduce an exact penalty function method proposed in [24] for solving Problem \((P_4)\). The main idea is to approximate Problem \((P_4)\) by an unconstrained optimization problem by appending the constraint violation to the cost function to form an augmented cost function. This approximated unconstrained optimization problem is solvable by frequently used optimization algorithms, such as sequential quadratic programming method [10].

Let \(\mathbb{V}\) be the set of all \(\theta \in \mathbb{R}^p\) such that \(0 \leq \theta_k \leq T, k = 1, \ldots, p\). Clearly, \(\Theta \subset \mathbb{V}\). By adopting the idea introduced in [24], we define the following exact penalty function:

\[
F_p(\theta, \delta, \epsilon) = \begin{cases} 
\tilde{G}_0^p((\theta, \delta)), & \text{if } \epsilon = 0, \text{ and } (\theta, \delta) \text{ is feasible for } (P_4), \\
\tilde{G}_0^p((\theta, \delta)) + \epsilon^{-\alpha} \Delta(\theta, \delta, \epsilon) + \rho \epsilon^\beta, & \text{if } \epsilon > 0, \\
+\infty, & \text{otherwise},
\end{cases}
\]

(38)

where \(\epsilon > 0\) is a new decision variable, \(\alpha > 0\) and \(\beta > 0\) are fixed constants, \(\rho > 0\) is the penalty parameter, and the constraint violation \(\Delta(\theta, \delta, \epsilon)\) is constructed as follows:

\[
\Delta(\theta, \delta, \epsilon) = \left( \sum_{k=1}^{p} |\theta_k - T| \right)^2 + \sum_{i=1}^{N_T} \max\{\tilde{\Phi}_i((\tilde{x}^p(p|\theta, \delta)), 0)\}^2 \\
+ \sum_{s=1}^{N_{s+2r}} \sum_{k=1}^{p} \int_{k-1}^{k} \max\{\theta_k \tilde{h}_i(\mu(s), \tilde{x}^p(s|\theta, \delta), \sigma^k), 0\}^2.
\]

The term \(\epsilon^{-\alpha} \Delta(\theta, \delta, \epsilon)\) in (38) is constructed to penalize constraint violations, and the term \(\rho \epsilon^\beta\) is to penalize large values of \(\epsilon\). The smaller the value of \(\epsilon\) yields the larger value of coefficient \(\epsilon^{-\alpha}\) in the term \(\epsilon^{-\alpha} \Delta(\theta, \delta, \epsilon)\). This will cause constraint violations to be penalized severely. Therefore, minimizing the penalty function for large values of \(\rho\) will lead to feasible points of Problem \((P_4)\). Now, we consider the following penalty problem.

**Problem** \((P_5)\). Given the dynamic system (28), choose a feasible triple \((\theta, \delta, \epsilon) \in \mathbb{V} \times (\mathbb{Z} \times \mathbb{Q}^p) \times [0, \infty)\) such that the penalty function \(F_p(\theta, \delta, \epsilon)\) is minimized.

The following convergence theorems are based on the results in [24]. The proofs of these theorems are quite similar, thus they are omitted.

**Theorem 5.1.** Let \((\theta^*, \delta^*, \epsilon^*)\) be a local solution of Problem \((P_5)\). Then, \((\theta^*, \delta^*)\) is a local solution of Problem \((P_4)\) if and only if \(\epsilon^* = 0\).

**Theorem 5.2.** Assume that \(\{\rho_k\}_{k=1}^\infty\) is an increasing sequence of penalty parameters such that \(\rho_k \to \infty\) as \(k \to \infty\), and \((\theta_k^*, \delta_k^*, \epsilon_k^*)\) is a local solution of Problem \((P_5)\) with \(\rho = \rho_k\). Then, there exists a \(k_0 > 0\), such that for any \(k \geq k_0\), every local
minimizer \((\theta^k, \delta^k, \epsilon^k, \ast)\) of Problem \((P_5)\) has the form of \((\theta^*, \delta^*, 0)\) where \((\theta^*, \delta^*)\) is a local minimizer of Problem \((P_4)\).

From the results above, we can conclude that Problem \((P_5)\) is a good approximation of Problem \((P_4)\) when \(\rho\) is large. That is, for a sufficiently large \(\rho\), a local optimal solution of Problem \((P_5)\) is a local optimal solution of Problem \((P_3)\). Since Problem \((P_5)\) only involves bound constraints, it is much easier to solve.

Remark 2. Problem \((P_5)\) is a standard optimal control problem with fixed switching times and can be readily solved by gradient-based optimization methods such as sequential quadratic programming method or existing optimal control software packages such as the optimal control software VISUAL MISER [8]. However, many of these approaches are designed to find local solutions rather than global solutions. The quality of the local solutions depends on the initial guess used to start the optimization procedure. Thus, to obtain global or near-global solutions, these approaches should be used with several different initial guesses. Hence, further works on solving Problem \((P_5)\) may consider to combine our method with a global optimization technique, such as the filled function method [21] or simulated annealing approach [23].

6. Numerical simulations. In this section, we present some experiences with a nontrivial test example. The computational approach described in the previous sections have been implemented by the optimal control software VISUAL MISER [8].

Example. Transfer of container crane operation: minimum swing [17]

Consider the following optimal control problem which comes from realistic and complex problem of transferring containers from a ship to a cargo truck at the port of Kobe:

\[
\min g_0(u) = 4.5 \int_0^1 [(x_3(t))^2 + (x_6(t))^2] dt,
\]

subject to the dynamic system

\[
\begin{align*}
\dot{x}_1(t) &= 9x_4(t), \\
\dot{x}_2(t) &= 9x_5(t), \\
\dot{x}_3(t) &= 9x_6(t), \\
\dot{x}_4(t) &= 9(u_1(t) + 17.2656x_3), \\
\dot{x}_5(t) &= 9u_2(t), \\
\dot{x}_6(t) &= -9(u_1(t) + 27.0756x_3(t) + 2x_5(t)x_6(t)) / x_2(t).
\end{align*}
\]

the control constraints

\[
\begin{align*}
|\dot{u}_1(t)| &\leq 10, t \in [0, 1], \\
|\dot{u}_2(t)| &\leq 10, t \in [0, 1],
\end{align*}
\]

the terminal state constraints

\[
\begin{align*}
x_1(1) &= 10, \\
x_2(1) &= 14,
\end{align*}
\]
\begin{align*}
x_3(1) &= 0, & x_4(1) &= 2.5, \\
x_5(1) &= 0, & x_6(1) &= 0,
\end{align*}

and the continuous inequality state constraints
\[ |x_4(t)| \leq 2.5, \quad |x_5(t)| \leq 1.0, \]

with the initial condition
\[ x(0) = [0, 22, 0, 0, -1, 0]^T. \]

Note that the control functions \( u_1(t) \) and \( u_2(t) \) are required to be continuous, thus, we introduce the following new states for the original dynamic system:
\begin{align*}
x_7(t) &= u_1(t), & t &\in [0, 1], \\
x_8(t) &= u_2(t), & t &\in [0, 1],
\end{align*}

such that
\begin{align*}
\dot{x}_7(t) &= v_1(t), & t &\in [0, 1], \\
\dot{x}_8(t) &= v_2(t), & t &\in [0, 1],
\end{align*}

with the initial conditions
\begin{align*}
x_7(0) &= \xi_1, \\
x_8(0) &= \xi_2.
\end{align*}

This problem is solved by using the optimal control software VISUAL MISER \cite{8} with the planning time horizon \([0, 1]\) being divided into \( p = 10 \) variable subintervals, and we obtain a cost of 5.47E-03. For comparison, we also divide the planning time horizon into \( q = 10 \) equally-spaced subintervals, and the optimal cost obtained is 5.61E-03. This is expected, as the new approach provides an added flexibility to optimize the control switching times. Figures 1-6 show the states and controls of the optimal control problem. It is clear from these figures that all the constraints are satisfied.

\section{Conclusion.}
In the literature on optimal control, the control is usually assumed to be essentially bounded measurable. However, for many real practical problems, abrupt changes in the control signal are not allowed because they will induce an infinite jerk at the switching points, and hence exciting a large bandwidth of vibration modes causing structural vibration. To overcome this problem, as in \cite{20}, a class of optimal control problems subject to continuous state and control inequality constraints is considered in this paper, where the control is required to be almost smooth. This class of optimal control problems is certainly an important addition to the optimal control literature. This paper then goes on to develop a much improved computational method for solving this class of optimal control problems. Detailed contributions can be summarized as follows:

1. A class of optimal control problems subject to continuous state and control inequality constraints is considered in this paper, where the control is required to be almost smooth.

2. A new differential equation is introduced for each almost smooth control variable, where the control variable is treated as a new state with a new control being introduced. Then this new control can be taken as an essentially bounded measurable function, and hence it can be approximated by piecewise constant function.
The initial condition of these new differential equations are also decision variables to be determined optimally.

3. For the continuous state and control inequality constraints, they are dealt with the exact penalty function method. Then, they are appended to the cost function forming a new cost function, and hence a new class of optimal control problems without the continuous state and control inequality constraints is obtained. This approach is better than the constraint transcription method used in [20], as the
Figure 3. Optimal state trajectory $x_1(t)$.

Figure 4. Optimal state trajectory $x_2(t)$.

local optimal solution obtained by the exact penalty function approach for the transformed problem is a local solution of the original constrained optimal control problem. This property is not shared by the constraint transcription approach used in [20].

4. For these new optimal control problems, the control parameterization approach is used to approximate the new control functions which is further supplemented by the time scaling transformation. Thus, a sequence of approximate finite
dimensional dynamic optimization problems is obtained. The gradient formulas of the new cost functions are derived and therefore, a gradient based optimization technique can be used to solved each of these finite dimensional approximate dynamic optimization problems. The computational method is supported by convergence analysis. We wish to remark that the time scaling transformation is not used in [8], and consequently the number of the partition points used by the control parametrization method is required to be very large. In this paper, due to the use
of the time scaling transformation, the number of the partition points is much less and yet the accuracy is higher. This can be clearly seen from the numerical results obtained in [8] and in Section 5 of this paper.

5. A nontrivial numerical example involving a container crane is solved using the approach developed in this paper. The results obtained look convincing.
REFERENCES

[1] N. U. Ahmed, *Dynamic Systems and Control with Applications*, Singapore: World Scientific, 2006.

[2] N. Banihashemi and C. Y. Kaya, Inexact restoration and adaptive mesh refinement for optimal control, *Journal of Industrial and Management Optimization*, 10 (2014), 521–542.

[3] D. Chang and Z. Wu, Stochastic maximum principle for non-zero sum differential games of FBSDEs with impulse controls and its application to finance, *Journal of Industrial and Management Optimization*, 11 (2015), 27–40.

[4] M. Gerdis and M. Kunkel, A nonsmooth Newton’s method for discretized optimal control problems with state and control constraints, *Journal of Industrial and Management Optimization*, 4 (2008), 247–270.

[5] L. Göllmann and H. Maurer, Theory and applications of optimal control problems with multiple time-delays, *Journal of Industrial and Management Optimization*, 10 (2014), 413–441.

[6] Y. Han and Y. Gao, Determining the viability for hybrid control systems on a region with piecewise smooth boundary, *Numerical Algebra, Control and Optimization*, 5 (2015), 1–9.

[7] L. S. Jennings, M. E. Fisher, K. L. Teo, et al., *MISER 3 Optimal Control Software: Theory and User Manual*, version 3, University of Western Australia, 2004.

[8] L. Jennings, C. Yu, B. Li, V. Rehbock, R. Loxton and F. Yang, Visual miser: an efficient user-friendly visual program for solving optimal control problems, *Journal of Industrial and Management Optimization*, 12 (2015), 781–810.

[9] J. Kaartinen, J. Hätönen, H. Hyötyniemi and J. Miettunen, Machine-vision-based control of zinc flotation: A case study, *Control Engineering Practice*, 14 (2006), 1455–1466.

[10] C. T. Lawrence and A. L. Tits, A computationally efficient feasible sequential quadratic programming algorithm, *SIAM Journal on Optimization*, 11 (2006), 1092–1118.

[11] H. W. J. Lee, K. L. Teo, V. Rehbock and L. S. Jennings, Control parametrization enhancing technique for time optimal control problems, *Dynamic Systems and Applications*, 6 (1997), 243–261.

[12] B. Li, C. Xu, K. L. Teo and J. Chu, Time optimal Zermelo’s navigation problem with moving and fixed obstacles, *Applied Mathematics and Computation*, 224 (2013), 866–875.

[13] 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.

[14] R. Loxton, Q. Lin, V. Rehbock and K. L. Teo, Control parameterization for optimal control problems with continuous inequality constraints: New convergence results, *Numerical Algebra, Control and Optimization*, 2 (2012), 571–599.

[15] K. S. Peterson and A. G. Stefanopoulou, Extremum seeking control for soft landing of an electromechanical valve actuator, *Automatica*, 40 (2004), 1063–1069.

[16] V. Rehbock, *Tracking Control and Optimal Control*, PhD thesis, University of Western Australia, Perth, 1994.

[17] Y. Sakawa and Y. Shindo, Optimal control of container cranes, *Automatica*, 18 (1982), 257–266.

[18] K. Schittkowski, NLPQL: A fortran subroutine solving constrained nonlinear programming problems, *Ann. Oper. Res.*, 5 (1986), 485–500.

[19] K. L. Teo and L. S. Jennings, Nonlinear optimal control problems with continuous state inequality constraints, *Journal of Optimization Theory and Application*, 63 (1989), 1–22.

[20] K. L. Teo, C. J. Goh and K. H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Long-man Scientific and Technical, Essex, 1991.

[21] W. X. Wang, Y. L. Shang, L. S. Zhang, et. al., Global minimization of non-smooth unconstrained problems with filled function, *Optimization Letters*, 7 (2013), 435–446.

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

[23] K. F. C. Yiu, Y. Liu and K. L. Teo, A hybrid descent method for global optimization, *Journal of Global Optimization*, 28 (2004), 229–238.
[24] C. Yu, K. L. Teo, L. S. Zhang and Y. Q. Bai, A new exact penalty function method for continuous inequality constrained optimization problems, *Journal of Industrial and Management Optimization*, 6 (2010), 895–910.

[25] C. Yu, K. L. Teo, L. S. Zhang and Y. Q. Bai, On a refinement of the convergence analysis for the new exact penalty function method for continuous inequality constrained optimization problem, *Journal of Industrial Management and Optimization*, 8 (2012), 485–491.

Received January 2018; revised March 2018.

E-mail address: zmuy@zjnu.cn
E-mail address: yuchangjun@126.com
E-mail address: zmuxyt@126.com
E-mail address: yqbai@t.shu.edu.cn