



Research article

Sequential adaptive switching time optimization technique for maximum hands-off control problems

Sida Lin¹, Lixia Meng¹, Jinlong Yuan¹, Changzhi Wu², An Li³, Chongyang Liu^{4,5,*} and Jun Xie⁶

¹ Department of Applied Mathematics, School of Science, Dalian Maritime University, Dalian 116026, China

² Chongqing National Center for Applied Mathematics, Chongqing Normal University, Chongqing 404087, China

³ School of Mathematical Sciences, Xiamen University, Xiamen 361005, China

⁴ School of Mathematics and Information Science, Shandong Technology and Business University, Yantai 264005, China

⁵ Yantai Key Laboratory of Big Data Modeling and Intelligent Computing, Yantai 264005, China

⁶ Department of Basics, PLA Dalian Naval Academy, Dalian 116018, China

* **Correspondence:** Email: liu_chongyang@yahoo.com; Tel: +86-05356903824; Fax: +86-05356903824.

Abstract: In this paper, we consider maximum hands-off control problem governed by a nonlinear dynamical system, where the maximum hands-off control constraint is characterized by an L^0 norm. For this problem, we first approximate the L^0 norm constraint by a L^1 norm constraint. Then, the control parameterization together with sequential adaptive switching time optimization technique is proposed to approximate the optimal control problem by a sequence of finite-dimensional optimization problems. Furthermore, a smoothing technique is exploited to approximate the non-smooth maximum operator and an error analysis is investigated for this approximation. The gradients of the cost functional with respect to the decision variables in the approximate problem are derived. On the basis of these results, we develop a gradient-based optimization algorithm to solve the resulting optimization problem. Finally, an example is solved to demonstrate the effectiveness of the proposed algorithm.

Keywords: maximum hands-off control; control parameterization; sequential adaptive switching time optimization technique; smoothing technique; error analysis

1. Introduction

Within numerous control systems, various measures are implemented to mitigate costs or curtail energy consumption. For some special cases, specific control values are deliberately maintained at zero for designated durations using handoff control strategies, thereby effectively minimizing the associated control costs. In [1], a novel control methodology was introduced to solve the maximum hands-off control problem characterized by L^0 norm, which is to find the sparsest control among all permissible controls. This control methodology holds significant utility in minimizing electricity or fuel consumption, showcasing its practical relevance and applicability. In [1], within the car's start-stop system, the engine is automatically deactivated when the vehicle comes to a halt or operates at speeds below a predetermined threshold. This deliberate shutdown serves to diminish CO₂ emissions and decrease fuel consumption as part of the system's efficiency measures. Handoff control demonstrates effectiveness within networked control systems as well [2].

Sparse optimization has emerged as a prominent subject of interest among scholars in recent years. The technology of sparse optimization has garnered extensive attention and exploration across various domains, including compressed sensing [3], image and signal processing [4,5], machine learning [6,7], Alzheimer's disease biomarkers [8], and other fields. Maximum hands-off control shares a relationship with sparsity, a characteristic pertinent to numerous optimization problems. In the realm of sparse optimization, sparsity refers to a condition where a substantial portion of elements within a matrix or vector assumes zero values. Leveraging sparsity offers the potential to economize on both time and compression expenses. Moreover, it serves as a mechanism to sift through extensive datasets, extracting pertinent information and thereby streamlining the complexity of the problem [9]. Within various optimization problems, the L^0 norm of a vector commonly serves as a means to quantify and define sparsity [10, 11]. Due to its non-convex and non-continuous nature, the L^0 norm presents inherent complexity in analysis. Consequently, numerous scholars resort to exploring the variational property of the L^0 norm and employing subdifferentiation techniques to address its intricacies. In [12], a technique known as convex relaxation was introduced as a substitute for the L^0 norm, employing the more tractable L^1 norm. This transformation enables the formulation of the problem in a linear programming paradigm, amenable to resolution using methods such as the interior point method or the simplex method. In [13, 14], the approach revolves around the non-convex relaxation method, leveraging the L^p relaxation technique, where $0 \leq p \leq 1$, to yield solutions with increased sparsity. In [15, 16], a focus is directed toward the Difference of Convex functions (DC) algorithm. This method entails the representation of the L^0 norm as a difference between two convex functions. Subsequently, leveraging the DC algorithm facilitates the resolution of the resultant relaxation model. In [17], by addressing the non-convex nature of the logarithmic function, a proposal was made to substitute the logarithm function for the L^0 norm. In [18], the resolution involves addressing a sparse optimization problem that encapsulates two competing objectives: measurement error minimization and sparsity maximization. This problem is approached and solved through the utilization of a multi-objective evolutionary algorithm. In [19], an effective fault diagnosis framework was constructed by integrating L^0 norm sparse constraint optimization with principal component analysis, aimed at mitigating the extent of sparsity. However, in [12–19], the sparse optimization problems do not take into account the constraints of the dynamical system. The focus of the study presented in [20] revolves around the L^1 objective optimal control problem specifically tailored for linear discrete systems. This formulation leads to diverse

sparse solutions based on the selection of distinct problem parameters. The study detailed in [21] introduces a methodical approach for synthesizing sparse optimal control structures governed by linear discrete systems. This method guarantees sparsity within solutions, enabling the specification of a predetermined count of zero elements within the control structure. In [22], the utilization of the Smith iteration and Altering-Direction-Implicit (ADI) iteration methods was explored for obtaining numerical solutions in a large sparse optimal control problem governed by linear discrete-time Riccati equations, leveraging Newton's method. In [23], provably optimal sparse solutions to overdetermined linear systems with non-negativity constraints were researched in a least-squares sense by implicit enumeration. Nevertheless, in [20–23], the sparse optimal control problem governed by the linear dynamical system failed to incorporate the constraints associated with the nonlinear dynamical system. The work presented in [24] delved into the examination of the Newton-like method of undetermined equations. This exploration led to the discovery of sparse solutions pertaining to sparse optimal control problems governed by an affine nonlinear system. The analysis of optimal control using the value function was conducted employing the dynamic programming method. In [25], a dynamic programming approach was introduced to effectively approximate the sparse optimal control problem governed by an affine nonlinear system in a numerical context. The maximum hands off control problem governed by the class of affine nonlinear systems was studied in [26]. Nonetheless, in [24–26], the nonlinear dynamical systems involved in the sparse optimal control problem are affine nonlinear systems, where the state variables and control inputs are separable. This paper delves into the investigation of the sparse optimal control problem governed by a general nonlinear dynamical system where the state variables and control inputs are inseparable.

The conventional approach to optimal control involves determining the most effective control strategy within specified constraints [27]. This strategy aims to either maximize or minimize the performance index. Typically, deriving an analytical solution for the optimal control problem governed by a nonlinear dynamic system proves challenging. Therefore, the resolution of such problems often necessitates numerical methods to obtain an effective solution [28]. The numerical solutions to optimal control problems can be divided into two solution strategies: first optimize then discretize (indirect methods), and first discretize then optimize (direct methods) [29]. For problems that do not contain inequality constraints, indirect methods derive the first-order necessary conditions for the optimal control problem, namely, the Euler-Lagrange equations that include the initial and boundary value conditions. The solution to this initial-boundary value problem mainly involves two types of methods: the control vector iteration method and methods such as multiple shooting and collocation [30]. The direct method encompasses techniques such as the direct collocation method and control parameterization [31, 32]. Control parameterization involves discretizing solely the control function, approximating it using fixed basis functions within specific subintervals [33]. The coefficients within this linear combination of basis functions serve as the optimal decision variables. Typically, predetermined switching times govern the transitions between values for each control component, often evenly divided within specified intervals [34]. To enhance solution accuracy, the time range of the control function is frequently subdivided more finely, leading to a greater number of decision variables. However, this denser division amplifies computational costs. Simultaneously, to mitigate the need for extensive time range subdivisions, one must consider incorporating the switching times as additional decision variables [35]. The traditional time-scaling transformation operates by mapping variable switching times with fixed points within a redefined time horizon. This process yields an optimization problem wherein the revised

switching times remain constant. The application of the time-scaling transformation finds extensive utility across diverse domains, including mixed integer programming [36] and singular optimal control [37]. Yet, in practical scenarios, simultaneous switching of time poses notable challenges [38]. The sequential adaptive switching time optimization technique (SASTOT) presented by [39], suggests that the interval for control switching times can vary, affording the flexibility to freely select the number of segments for each approximate control component. The method initiates by applying control parameterization and the time-scaling transformation to a single control component initially, leaving the remaining components untouched. Subsequently, within the newly introduced time range, the time-scaling transformation introduces a subsequent new time range. This iterative process continues until the final control component undergoes processing in a similar manner. This modification introduces a significant level of flexibility into the control strategy. Compared with the traditional time-scale transformation techniques, the SASTOT can accurately identify sparse positions for each control.

In this study, our focus lies on the maximum hands-off control problem governed by a nonlinear dynamical system (MHCPNDS) with a maximum hands-off control constraint characterized by the L^0 norm. The optimization variables encompass the count of segments for each control component and the respective switching times for each of these components. The examination of the MHCPNDS poses challenges in obtaining an analytical solution. Hence, the adoption of a numerical solution becomes imperative. In practical scenarios, simultaneous switching of all control components is not optimal. Addressing this, this paper introduces the SASTOT, which allows for a flexible selection of the number of segments and switching times for each individual control element. The integration of control parameterization with the SASTOT serves as a transformative approach to addressing the MHCPNDS. The non-smooth term that involves the maximum operator is approximated through a smoothing function, and we further illustrate the convergence of this approximation technique. The attainment of a sparse solution for the MHCPNDS involves the utilization of a gradient-based algorithm. Empirical assessments through numerical experiments substantiate the efficacy of the algorithm put forth in the study.

The contributions of the paper are three-fold:

1) Distinguished from the above-mentioned sparse optimization problem, which is amenable to analytical solutions, we consider the sparse optimal control problem within a nonlinear dynamic system. Given the intrinsic complexity of the nonlinear dynamic system, pinpointing an analytical solution poses a significant challenge. Taking these into consideration, we propose a numerical solution method, based on the gradient formulae of the cost function (Theorems 5–8) and the gradient-based algorithm, without requiring the linearization of the nonlinear dynamical systems.

2) A smoothing function has been introduced to mitigate the roughness of the non-smooth term involving the maximum operator. Several significant theorems (Theorems 1–4) have been established, illustrating that the smoothing function effectively addresses the shortcomings associated with constraint qualification noncompliance.

3) Setting itself apart from the time-scaling transformation technique, this paper introduces the SASTOT method that enables flexible determination of the number of segments and the specific switching times for each control element individually.

The rest of the paper is organized as follows. In Section 2, we present the MHCPNDS. In Section 3, the MHCPNDS is transformed by using control parameterization and the SASTOT. In Section 4, we deal with the nonsmoothness of the objective function by using the smoothing technique. In Section 5, the gradient-based algorithm is used to solve the resulting smooth problem. In Section 6, numeri-

cal results are presented. In Section 7, we draw some conclusions and suggest some future research directions.

2. Maximum hands-off control problem

Let I_n be the set of $\{1, 2, \dots, n\}$. For a continuous time control $v : [0, T] \rightarrow R$, the L^p norm is defined by

$$\|v\|_p \triangleq \left(\int_0^T |v|_p dt \right)^{\frac{1}{p}}, p \in [1, +\infty). \quad (2.1)$$

The L^0 norm is defined by

$$\|v\|_0 \triangleq q_M(\text{supp}(v)), \quad (2.2)$$

where q_M is the Lebesgue measure and the set $\text{supp}(v)$ is called a support set of v defined by

$$\text{supp}(v) = \{t \in [0, T] : v(t) \neq 0\}. \quad (2.3)$$

In some cases, it is possible to significantly reduce the control effort by keeping the control value at exactly zero over a time interval. Maximum hands-off control can be called sparse control, which means that the more zeros the control value is in unit time, the more sparse it is. It makes the time interval of the control completely zero.

In order to obtain the sparse solution of maximum hands-off control, we consider a nonlinear dynamical system defined by

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), t \in [0, T], \\ \mathbf{x}(0) = \xi, \\ \mathbf{x}(T) = 0, \end{cases} \quad (2.4)$$

where $\mathbf{x}(t) := (x_1(t), \dots, x_n(t))^T \in \mathbb{R}^n$ denotes the state vector at time t , $\mathbf{u}(t) := (u_1(t), \dots, u_m(t))^T \in \mathbb{R}^m$ denotes control input vector at time t , ξ denotes the initial state value, $T \geq 0$ denotes a given terminal time, and $\mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))$ denotes a nonlinear function vector defined on \mathbb{R}^n .

Let $\mathbf{x}(\cdot|\mathbf{u})$ denote the solution of system (2.4) satisfying

$$\mathbf{x}(t|\mathbf{u}) \geq 0, t \in [0, T]. \quad (2.5)$$

For all t on $[0, T]$, there is also the following constraint on each component of the control input vector $\mathbf{u}(t)$

$$\max_{i \in I_m} |u_i(t)| \leq 1. \quad (2.6)$$

A control input vector $\mathbf{u}(t) \in \mathbb{R}^m$ that satisfies constraint (2.6) is called a candidate control input vector. Let $\mathcal{L}[T, \xi]$ be the set consisting of all candidate control input vectors.

Definition 1. The maximum hands-off control constraint characterized by the L^0 norm is defined by

$$A_0(\mathbf{u}) \triangleq \frac{1}{T} \sum_{i=1}^m \lambda_i \|u_i\|_0 \leq \mu, \quad (2.7)$$

where $\lambda_i \geq 0, i \in I_m$, are given weights, $\mathbf{u} \in \mathcal{L}[T, \xi]$ denotes the admissible control, and μ is a small positive number.

Definition 1 is used to characterize the sparsity of maximum hands-off control. The cost functional is defined by

$$J(u) = \pi_0(\mathbf{x}(T|u)) + \int_0^T \vartheta(\mathbf{x}(t|u)) dt, \quad (2.8)$$

where $\pi_0 : \mathbb{R}^m \rightarrow \mathbb{R}$ and $\vartheta : \mathbb{R}^m \rightarrow \mathbb{R}$ are continuously differentiable functions.

Then, our maximum hands-off control problem can be formulated as follows.

Problem A: Given system (2.4), choose an admissible control $\mathbf{u} \in \mathcal{L}[T, \xi]$ to minimize the cost functional defined in (2.8) subject to boundary condition (2.6) and maximum hands-off control constraint (2.7).

3. Solution to maximum hands-off control problem

The L^0 norm, owing to its non-convex and discontinuous characteristics, inherently introduces complexity into analytical procedures. The L^0 norm can be well approximated by the L^1 norm [20]. The maximum hands-off control constraint characterized by the L^0 norm is approximated by Definition 2.

Definition 2. The maximum hands-off control constraint characterized by the L^1 norm is defined by

$$A_1(\mathbf{u}) \triangleq \frac{1}{T} \sum_{i=1}^m \lambda_i \|u_i\|_1 = \frac{1}{T} \sum_{i=1}^m \lambda_i \int_0^T |u_i| dt \leq \mu, \quad (3.1)$$

where $\lambda_i \geq 0, i \in I_m$, are given weights, $\mathbf{u} \in \mathcal{L}[T, \xi]$ denotes the admissible control, and μ is a small positive number.

Definition 2 serves to delineate the computationally efficient sparsity of maximum hands-off control. By Definition 2, Problem A can be well approximated by the following Problem B.

Problem B: Given system (2.4), choose the control $\mathbf{u} \in \mathcal{L}[T, \xi]$ to minimize the objective functional defined in (2.8) subject to boundary condition (2.6) and maximum hands-off control constraint (3.1).

Control parameterization stands as an effective methodology for resolving the optimal control problem, commonly approached through approximation via piecewise constant functions [29, 40–42]. Increasing the level of detail in the time level partition leads to greater accuracy as it allows for a more intricate and nuanced representation of the time-dependent processes or phenomena under consideration [43]. Employing control parameterization enables the derivation of a finite-dimensional approximation for Problem B. Ultimately, the gradient-based algorithm is employed to resolve the resultant approximation problem.

The conventional time-scaling transformation necessitates simultaneous switching of all control components, a condition challenging to attain in practical applications [44]. Hence, the introduction of

SASTOT emerges as a solution, combining both time-scaling transformation and control parameterization methodologies [39]. Within this methodology, each control component possesses the capability to adaptively adjust and independently select its switching time, allowing for diverse switching points across components. Empirical validation showcases a notable reduction in computational complexity alongside an improvement in the accuracy of calculations. For a clearer exposition of this approach, this article examines two distinct control inputs. Let $\tilde{\mathbf{u}}(t) = [\tilde{u}_1(t), \tilde{u}_2(t)]^T$. Then, the dynamical system is

$$\tilde{\mathbf{f}}(t, \tilde{\mathbf{x}}(t), \tilde{u}_1(t), \tilde{u}_2(t)) = \tilde{\mathbf{f}}(t, \tilde{\mathbf{x}}(t), \tilde{\mathbf{u}}(t)). \quad (3.2)$$

3.1. Control parameterization and time-scale transformation for $\tilde{u}_1(t)$.

The interval $[0, T]$ is divided into q_1 subintervals $[\delta_1^{l_1-1}, \delta_1^{l_1}]$, $l_1 \in I_{q_1}$, where $\sigma_1 := [\delta_1^1, \delta_1^2, \dots, \delta_1^{q_1}]^T$ are variable switching time vector, and

$$0 = \delta_1^1 \leq \delta_1^2 \leq \dots \leq \delta_1^{q_1} = T. \quad (3.3)$$

The control component $\tilde{u}_1(t)$ can be approximated by

$$\tilde{u}_1(t) \approx \tilde{u}_1^{q_1}(t|\delta_1, \sigma_1) = \sum_{l_1=1}^{q_1} \eta_1^{l_1} \chi_{[\delta_1^{l_1-1}, \delta_1^{l_1}]}(t), \quad t \in [0, T], l_1 \in I_{q_1}, \quad (3.4)$$

where $\eta_1^{l_1}$ is the piecewise constant function value on the l_1 th subinterval of the control component $u_1(t)$ satisfying

$$a_1 \leq \eta_1^{l_1} \leq b_1, l_1 \in I_{q_1}, \quad (3.5)$$

and $\chi_{[\delta_1^{l_1-1}, \delta_1^{l_1}]}(t)$ is an indicator function defined by

$$\chi_{[\delta_1^{l_1-1}, \delta_1^{l_1}]}(t) = \begin{cases} 1, & \text{if } t \in [\delta_1^{l_1-1}, \delta_1^{l_1}), \\ 0, & \text{if } t \notin [\delta_1^{l_1-1}, \delta_1^{l_1}). \end{cases} \quad (3.6)$$

Then, we obtain the following system:

$$\begin{cases} \frac{d\tilde{\mathbf{x}}(t)}{dt} = \sum_{l_1=1}^{q_1} \tilde{\mathbf{f}}(t, \tilde{\mathbf{x}}(t), \eta_1^{l_1}, \tilde{u}_2(t)), & t \in [0, T], \\ \tilde{\mathbf{x}}(0) = \xi, \\ \tilde{\mathbf{x}}(T) = 0. \end{cases} \quad (3.7)$$

Let $\tilde{\mathbf{x}}(\cdot|\boldsymbol{\eta}_1, u_2)$ denote the solution of system (3.7).

After applying the time-scaling transformation [45] to $\tilde{u}_1(t)$ and mapping the variable switching times $\{\delta_1^0, \delta_1^1, \dots, \delta_1^{q_1}\}$ to a fixed number of switching times $\{0, 1, \dots, q_1\}$ in the new time horizon, we define the vector $\boldsymbol{\phi}_1 := [\phi_1^1, \dots, \phi_1^{q_1}]^T \in \mathbb{R}^{q_1}$, where $\phi_1^{l_1} = \delta_1^{l_1} - \delta_1^{l_1-1} \geq \rho$, ρ is an extremely small positive number, $l_1 \in I_{q_1}$, and $\phi_1^1 + \phi_1^2 + \dots + \phi_1^{q_1} = T$.

Then, we introduce a new time variable p and define a time scaling function $v_1(p, \boldsymbol{\phi}_1)$

$$t(p) \triangleq v_1(p, \boldsymbol{\phi}_1) = \sum_{l_1=1}^{\lfloor p \rfloor} \phi_1^{l_1} + \phi_1^{(\lfloor p \rfloor + 1)}(p - \lfloor p \rfloor), \quad p \in [0, q_1], \quad (3.8)$$

where $\theta(p) = \tilde{u}_2(v_1(p, \phi_1)) = \tilde{u}_2(t)$ and $\lfloor p \rfloor$ is a floor function of the time variable p . With the new time variable p , the dynamical system is redefined on the subinterval $[l_1 - 1, l_1)$, $l_1 \in I_{q_1}$,

$$\begin{cases} \frac{d\hat{x}(p)}{dp} = \phi_1^{\lfloor p \rfloor} \hat{f}(p, \hat{x}(p), \eta_1^{\lfloor p \rfloor}, \theta(p)), \\ \hat{x}(0) = \xi, \\ \hat{x}(T) = 0. \end{cases} \quad (3.9)$$

Let $\hat{x}(\cdot | \eta_1, \theta)$ denote the solution of system (3.9).

3.2. Control parameterization and time-scaling transformation for $\tilde{u}_2(t)$.

The interval $[0, q_1]$ is divided into q_2 subintervals $[p_2^{l_2-1}, p_2^{l_2}]$, $l_2 \in I_{q_2}$, where $\sigma_2 = [p_2^0, p_2^1, \dots, p_2^{q_2}]^T$ is the variable switching time vector, and

$$0 = p_2^0 \leq p_2^1 \leq \dots \leq p_2^{q_2} = T. \quad (3.10)$$

The control component $\tilde{u}_2(t)$ can be approximated by

$$\tilde{u}_2(t) \approx \theta(p) := \sum_{l_2=1}^{q_2} \eta_2^{l_2} \chi_{[p_2^{l_2-1}, p_2^{l_2})}(p), \quad p \in [0, q_1], \quad (3.11)$$

where $\eta_2^{l_2}$ is the piecewise constant function value on the l_2 th subinterval of the control component $\theta(p)$ satisfying

$$a_2 \leq \eta_2^{l_2} \leq b_2, \quad l_2 \in I_{q_2}. \quad (3.12)$$

With the new time variable p , system (3.9) is redefined on the subinterval $[p_2^{l_2-1}, p_2^{l_2})$, $l_2 \in I_{q_2}$,

$$\begin{cases} \frac{d\check{x}(p)}{dp} = \phi_1^{\lfloor p \rfloor} \check{f}(p, \check{x}(p), \eta_1^{\lfloor p \rfloor}, \eta_2^{l_2}), \\ \check{x}(0) = \xi, \\ \check{x}(0) = 0. \end{cases} \quad (3.13)$$

After applying a time-scaling transformation to $\check{u}_2(t)$ and mapping the variable switching times $\{\delta_2^0, \delta_2^1, \dots, \delta_2^{q_2}\}$ to a fixed number of switching times $\{0, 1, \dots, q_2\}$ in the new time horizon, we define the vector $\phi_2 := [\phi_2^1, \dots, \phi_2^{q_2}]^T \in \mathbb{R}^{q_2}$, where $\phi_2^{l_2} = \delta_2^{l_2} - \delta_2^{l_2-1} \geq \rho$, ρ is an extremely small positive number, $l_2 \in I_{q_2}$, and $\phi_2^1 + \phi_2^2 + \dots + \phi_2^{q_2} = T$.

Then, we introduce a new time variable w and define a time scaling function $v_2(w, \phi_2)$

$$p(w) \triangleq v_2(w, \phi_2) = \sum_{l_2=1}^{\lfloor w \rfloor} \phi_2^{l_2} + \phi_2^{(\lfloor w \rfloor + 1)}(w - \lfloor w \rfloor), \quad w \in [0, q_2], \quad (3.14)$$

where $\lfloor w \rfloor$ is the floor function of the time variable w . System (3.13) is redefined on the subinterval $[l_2 - 1, l_2)$, $l_2 \in I_{q_2}$,

$$\begin{cases} \frac{d\check{x}(w)}{dw} = \phi_2^{l_2} \phi_1^{\lfloor v_2(w) \rfloor} \check{f}(w, \check{x}(w), \eta_1^{\lfloor v_2(w) \rfloor}, \eta_2^{l_2}), \\ \check{x}(0) = \xi, \\ \check{x}(0) = 0. \end{cases} \quad (3.15)$$

Let $\tilde{x}(\cdot|\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$ be the solution of system (3.15). Then the continuous state constraint (2.5) becomes

$$\tilde{x}(w|\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \geq 0, \forall w \in [l_2 - 1, l_2), l_2 \in I_{q_2}. \quad (3.16)$$

The cost function (2.8) and the maximum hands-off control constraint (3.1) become

$$\bar{J}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) = \pi_0(\tilde{x}(q_2|\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2)) + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor l_2(w) \rfloor} \vartheta(\tilde{x}(w|\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2)) dw, \quad (3.17)$$

$$A_2(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) = \frac{1}{q_2} \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor l_2(w) \rfloor} [\lambda_1 |\eta_1^{\lfloor l_2(w) \rfloor}| + \lambda_2 |\eta_2^{l_2}|] dw \leq \mu. \quad (3.18)$$

With these in mind, Problem **B** can be approximated by the following Problem **C**.

Problem C: Given system (3.15), choose the quadruple $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \mathbb{R}^{q_2}$ to minimize the objective function defined by (3.17) subject to continuous inequality constraints (3.16), boundary conditions (3.5) and (3.12), and maximum hands-off control constraint (3.18).

Because of $|\eta_1^{\lfloor l_2(w) \rfloor}| = 2\max\{\eta_1^{\lfloor l_2(w) \rfloor}, 0\} - \eta_1^{\lfloor l_2(w) \rfloor}$, $|\eta_2^{l_2}| = 2\max\{\eta_2^{l_2}, 0\} - \eta_2^{l_2}$, the maximum hands-off control constraint (3.18) can be equivalently transformed as $A_3(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$ defined in (3.19).

$$A_3(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) = \frac{1}{q_2} \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor l_2(w) \rfloor} [\lambda_1 (2\max\{\eta_1^{\lfloor l_2(w) \rfloor}, 0\} - \eta_1^{\lfloor l_2(w) \rfloor}) + \lambda_2 (2\max\{\eta_2^{l_2}, 0\} - \eta_2^{l_2})] dw \leq \mu \quad (3.19)$$

Define

$$J_1(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \rho) = \bar{J}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) + \rho H(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2), \quad (3.20)$$

where ρ is the penalty parameter and

$$H(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) = \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \max\{-\tilde{x}(w|\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2), 0\} dw + \sum_{l_1=1}^{q_1} \left[\max\{a_1 - \eta_1^{l_1}, 0\} + \max\{\eta_1^{l_1} - b_1, 0\} \right] + \sum_{l_2=1}^{q_2} \left[\max\{a_2 - \eta_2^{l_2}, 0\} + \max\{\eta_2^{l_2} - b_2, 0\} \right] + A_3(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2).$$

Thus, Problem **C** can be equivalently transformed into Problem **D**.

Problem D: Given system (3.15), choose the quadruple $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \mathbb{R}^{q_2}$ to minimize the objective function defined in (3.20).

Remark 1. According to [29], it has been established that when the penalty parameter ρ exceeds its threshold value ρ^* , the solution derived for Problem **D** represents an exact solution for Problem **C**.

4. Smoothing process

4.1. Smoothing technique

The fundamental concept underlying the smooth technique involves the approximation of the non-smooth maximum operator through the utilization of a smoothing function [46], which is defined by

$$P\{G, \rho, q, \varepsilon\} = \begin{cases} 0, & \text{if } G < -\frac{\varepsilon}{\rho q}, \\ \frac{\rho q}{2\varepsilon}G^2 + G + \frac{\varepsilon}{2\rho q}, & \text{if } -\frac{\varepsilon}{\rho q} \leq G < 0, \\ G + \frac{\varepsilon}{2\rho q}, & \text{if } G \geq 0, \end{cases} \quad (4.1)$$

where ρ is a penalty factor, q is the number of continuous inequality constraints, and $\varepsilon > 0$ is the smoothing parameter.

Theorem 1. ([46]) For $\varepsilon > 0$, we have

$$0 \leq P\{G, \rho, q, \varepsilon\} - \max\{G, 0\} \leq \frac{\varepsilon}{2\rho q}. \quad (4.2)$$

By the smoothing process, the maximum hands-off control constraint $A_3(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2)$ defined in (3.19) can be approximated by

$$\begin{aligned} & \tilde{A}_3(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \rho, q_2, \varepsilon) \\ &= \frac{1}{q_2} \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor} \left[\lambda_1 \left(2P\{\eta_1^{\lfloor \nu_2(w) \rfloor}, \rho, q_2, \varepsilon\} - \eta_1^{\lfloor \nu_2(w) \rfloor} \right) + \lambda_2 \left(2P\{\eta_2^{l_2}, \rho, q_2, \varepsilon\} - \eta_2^{l_2} \right) \right] dw. \end{aligned} \quad (4.3)$$

Based on the smoothing function (4.1), the cost function (3.20) can be approximated by

$$J_2(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \rho, q_1, q_2, \varepsilon) = \bar{J}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) + \rho \tilde{H}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \rho, q_1, q_2, \varepsilon), \quad (4.4)$$

where

$$\begin{aligned} \tilde{H}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \rho, q_1, q_2, \varepsilon) &= \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} P\{-\tilde{x}(w) | \boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2\}, \rho, q_2, \varepsilon\} dw + \tilde{A}_3(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \rho, q_2, \varepsilon) \\ &+ \sum_{l_2=1}^{q_2} \left[P\{a_2 - \eta_2^{l_2}, \rho, q_2, \varepsilon\} + P\{\eta_2^{l_2} - b_2, \rho, q_2, \varepsilon\} \right] + \sum_{l_1=1}^{q_1} \left[P\{a_1 - \eta_1^{l_1}, \rho, q_1, \varepsilon\} + P\{\eta_1^{l_1} - b_1, \rho, q_1, \varepsilon\} \right]. \end{aligned} \quad (4.5)$$

Based on the smoothing function (3.20), Problem **D** can be approximated by Problem **E**.

Problem E: Given system (3.15), choose the quadruple $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \mathbb{R}^{q_2}$ to minimize the cost function defined in (4.4).

4.2. Error analysis

The application of smoothing techniques may introduce discrepancies between Problems **D** and **E**. This section entails the derivation of errors between Problems **D** and **E** concerning the smoothing function (3.20). It has been demonstrated that for a sufficiently small smoothing parameter ε , the solution to Problem **D** can be acquired by sequentially solving a series of Problem **E** while incrementing the values of the penalty factor ρ .

Theorem 2. If $\rho > 0$, $q > 0$, and $\varepsilon > 0$, then

$$0 \leq J_2(\eta_1, \eta_2, \phi_1, \phi_2, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1, \eta_2, \phi_1, \phi_2, \rho) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}. \quad (4.6)$$

Proof. Based on Theorem 1, we have for $\rho > 0$, $q_1, q_2 > 0$, and $\varepsilon > 0$,

$$0 \leq P\{G, \rho, q, \varepsilon\} - \max\{G, 0\} \leq \frac{\varepsilon}{2\rho q}.$$

Then,

$$\begin{aligned} 0 &\leq J_2(\eta_1, \eta_2, \phi_1, \phi_2, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1, \eta_2, \phi_1, \phi_2, \rho) \\ &= \rho \left\{ \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \left(P\{\tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2), \rho, q_2, \varepsilon\} - \max\{\tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2), 0\} \right) dw \right. \\ &\quad + \sum_{l_1=1}^{q_1} \left[\left(P\{\eta_1^{l_1} - a_1, \rho, q_1, \varepsilon\} - \max\{\eta_1^{l_1} - a_1, 0\} \right) + \left(P\{b_1 - \eta_1^{l_1}, \rho, q_1, \varepsilon\} - \max\{b_1 - \eta_1^{l_1}, 0\} \right) \right] \\ &\quad + \sum_{l_2=1}^{q_2} \left[\left(P\{\eta_2^{l_2} - a_2, \rho, q_2, \varepsilon\} - \max\{\eta_2^{l_2} - a_2, 0\} \right) + \left(P\{b_2 - \eta_2^{l_2}, \rho, q_2, \varepsilon\} - \max\{b_2 - \eta_2^{l_2}, 0\} \right) \right] \\ &\quad \left. + \frac{1}{q_2} \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor} \left[2\lambda_1 \left(P\{\eta_1^{\lfloor \nu_2(w) \rfloor}, \rho, q_2, \varepsilon\} - \max\{\eta_1^{\lfloor \nu_2(w) \rfloor}, 0\} \right) + 2\lambda_2 \left(P\{\eta_2^{l_2}, \rho, q_2, \varepsilon\} - \max\{\eta_2^{l_2}, 0\} \right) \right] dw \right\} \\ &\leq \rho \left[\frac{5\varepsilon}{2\rho} + \frac{2(\lambda_1 + \lambda_2)T^2\varepsilon}{2\rho q_2} \right] = \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}, \end{aligned}$$

which completes the proof.

Theorem 3. Let $(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*)$ be the solution of Problem **D**, and let $(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**})$ be the solution of Problem **E**. Then,

$$0 \leq J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}.$$

Proof. Based on Theorem 2, we have

$$\begin{aligned} 0 &\leq J_2(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}, \\ 0 &\leq J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}. \end{aligned}$$

Since $(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*)$ is the solution of Problem **D**, we have

$$J_1(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho) > J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho),$$

which yields

$$\begin{aligned} & J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho) \\ & \leq J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho). \end{aligned} \quad (4.7)$$

Since $(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**})$ is the solution of Problem **D**, we obtain

$$J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) < J_2(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho, q_1, q_2, \varepsilon),$$

which yields

$$\begin{aligned} & J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho) \\ & \leq J_2(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho, q_1, q_2, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho). \end{aligned} \quad (4.8)$$

Based on (4.7) and (4.8), we have

$$\begin{aligned} 0 & \leq J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q, \varepsilon) - J_1(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho) \\ & \leq J_2(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho) \\ & \leq J_2(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho, q, \varepsilon) - J_1(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*, \rho) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}, \end{aligned}$$

which completes this proof.

Remark 2. Theorem 3 states that when the smoothing parameter ε assumes a sufficiently small value, the solution of Problem **E** tends to approximate the solution of Problem **D**.

To obtain an error estimation between the solutions of Problems **E** and **C** outlined in Theorem 4, the definition of ε -feasibility to Problem **D** is given in Definition 3.

Definition 3. A vector $(\eta_1^\varepsilon, \eta_2^\varepsilon, \phi_1^\varepsilon, \phi_2^\varepsilon) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \mathbb{R}^{q_2}$ is called ε -feasible to Problem **D** if

$$\begin{aligned} & P\{(\eta_2^l)^\varepsilon - a_2, \rho, q_2, \varepsilon\} \leq \varepsilon, \quad P\{b_2 - (\eta_2^l)^\varepsilon, \rho, q_2, \varepsilon\} \leq \varepsilon, \quad P\{(\eta_1^l)^\varepsilon - a_1, \rho, q_1, \varepsilon\} \leq \varepsilon, \\ & P\{b_1 - (\eta_1^l)^\varepsilon, \rho, q_1, \varepsilon\} \leq \varepsilon, \quad P\{-\tilde{x}(w|\eta_1^\varepsilon, \eta_2^\varepsilon, \phi_1^\varepsilon, \phi_2^\varepsilon), \rho, q_2, \varepsilon\} \leq \varepsilon, \\ & P\{(\eta_1^{lv_2(w)})^\varepsilon, \rho, q_2, \varepsilon\} \leq \varepsilon, \quad P\{(\eta_2^l)^\varepsilon, \rho, q_2, \varepsilon\} \leq \varepsilon. \end{aligned}$$

Theorem 4. Let $(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*)$ and $(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**})$ be the solution of Problems **D** and **E**, respectively. Furthermore, let $(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*)$ be feasible to Problem **D** and $(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**})$ be ε -feasible to Problem **D**. Then

$$-\frac{5\varepsilon}{2} - \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2} \leq \bar{J}(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}) - \bar{J}(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}. \quad (4.9)$$

Proof. Since $(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*)$ is the solution of Problem **D**, we have

$$\begin{aligned} H(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*) & = \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \max\{-\tilde{x}(w|\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*), 0\} dw \\ & + \sum_{l_1=1}^{q_1} \left[\max\{a_1 - (\eta_1^{l_1})^*, 0\} + \max\{(\eta_1^{l_1})^* - b_1, 0\} \right] \\ & + \sum_{l_2=1}^{q_2} \left[\max\{a_2 - (\eta_2^{l_2})^*, 0\} + \max\{(\eta_2^{l_2})^* - b_2, 0\} \right] + A_3(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*) = 0, \end{aligned}$$

where

$$\begin{aligned} & A_3(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*) \\ &= \frac{1}{q_2} \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} (\phi_2^{l_2})^* (\phi_1^{\lfloor v_2(w) \rfloor})^* [\lambda_1 (2 \max\{(\eta_1^{\lfloor v_2(w) \rfloor})^*, 0\} - (\eta_1^{\lfloor v_2(w) \rfloor})^*) + \lambda_2 (2 \max\{(\eta_2^{l_2})^*, 0\} - (\eta_2^{l_2})^*)] dw \\ &= 0. \end{aligned}$$

Since $(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**})$ is ε -feasible to Problem **D**, we obtain

$$0 \leq \rho \tilde{H}(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) \leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}.$$

Based on Theorem 3, we get

$$\begin{aligned} 0 &\leq \bar{J}(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}) + \rho \tilde{H}(\eta_1^{**}, \eta_2^{**}, \phi_1^{**}, \phi_2^{**}, \rho, q_1, q_2, \varepsilon) - \bar{J}(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*) - \rho H(\eta_1^*, \eta_2^*, \phi_1^*, \phi_2^*) \\ &\leq \frac{5\varepsilon}{2} + \frac{(\lambda_1 + \lambda_2)T^2\varepsilon}{q_2}, \end{aligned}$$

which completes this proof.

Remark 3. In this scenario, as outlined in Theorem 4, an error estimation between the solutions of Problems **E** and **C** is provided, specifically under the condition of $\rho > \rho^*$. Consequently, the solution for SOP (Sum of Optimization Problems) can be approximately obtained by iteratively solving a sequence of Problem **E**.

5. Gradient computation

In this section, we will derive the gradient formulae of the cost function $J_2(\eta_1, \eta_2, \phi_1, \phi_2, \rho, q_1, q_2, \varepsilon)$ with respect to the quadruple $(\eta_1, \eta_2, \phi_1, \phi_2) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \mathbb{R}^{q_2}$ based on Theorems 5–8, whose proofs are similar to the proofs of Theorems 1 and 2 in [39].

Theorem 5. For each $(\eta_1, \eta_2, \phi_1, \phi_2)$, we have

$$\frac{\partial \tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2)}{\partial \eta_1} = M_1(w|\eta_1, \eta_2, \phi_1, \phi_2), \quad w \in [0, q_2], \quad (5.1)$$

where $M_1(\cdot|\eta_1, \eta_2, \phi_1, \phi_2)$ is the solution to the following system on $w \in [l_2 - 1, l_2], l_2 \in I_{q_2}$:

$$\begin{cases} M_1(w) = \phi_2^{l_2} \phi_1^{\lfloor v_2(w) \rfloor} \left\{ \frac{\partial f(w, \hat{x}(w), \eta_1^{\lfloor v_2(w) \rfloor}, \eta_2^{l_2})}{\partial \tilde{x}} M_1(w) + \frac{\partial f(w, \hat{x}(w), \eta_1^{\lfloor v_2(w) \rfloor}, \eta_2^{l_2})}{\partial \eta_1} \right\}. \\ M_1(0) = 0. \end{cases} \quad (5.2)$$

On the basis of Theorem 5, the gradient of J_2 with respect to η_1 is given as follows:

$$\begin{aligned} \frac{\partial J_2}{\partial \eta_1} &= \frac{\pi_0(\tilde{x}(q_2|\eta_1, \eta_2))}{\partial \tilde{x}} M_1(q_2) + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor} \frac{\partial(\tilde{x}(w|\eta_1, \eta_2))}{\partial \tilde{x}} M_1(w) dw \\ &+ \rho \left[\sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{P\{-\tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2), \rho, q_2, \varepsilon\}}{\partial \tilde{x}} M_1(w) dw \right. \\ &+ \sum_{l_1=1}^{q_1} \frac{(P\{a_1 - \eta_1^{l_1}, \rho, q_1, \varepsilon\} + P\{\eta_1^{l_1} - b_1, \rho, q_1, \varepsilon\})}{\partial \eta_1} \\ &\left. + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{\phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor}}{q_2} \left\{ \frac{\partial[\lambda_1(2P\{\eta_1^{\lfloor \nu_2(w) \rfloor}, c, q_2, \varepsilon\} - \eta_1^{\lfloor \nu_2(w) \rfloor})]}{\partial \eta_1} \right\} dw \right]. \end{aligned} \quad (5.3)$$

Theorem 6. For each $(\eta_1, \eta_2, \phi_1, \phi_2)$,

$$\frac{\partial \tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2)}{\partial \eta_2} = M_2(w|\eta_1, \eta_2, \phi_1, \phi_2), \quad w \in [0, q_2],$$

where $M_2(\cdot|\eta_1, \eta_2, \phi_1, \phi_2)$ is the solution to the following system on $w \in [l_2 - 1, l_2]$, $l_2 \in I_{q_2}$:

$$\begin{cases} \dot{M}_2(w) = \phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor} \left\{ \frac{\partial f(w, \hat{x}(w), \eta_1^{\lfloor \nu_2(w) \rfloor}, \eta_2^{l_2})}{\partial \tilde{x}} M_2(w) + \frac{\partial f(w, \hat{x}(w), \eta_1^{\lfloor \nu_2(w) \rfloor}, \eta_2^{l_2})}{\partial \eta_2} \right\}, \\ M_2(0) = 0. \end{cases} \quad (5.4)$$

On the basis of Theorem 6, the gradient of J_2 with respect to η_2 is given as follows:

$$\begin{aligned} \frac{\partial J_2}{\partial \eta_2} &= \frac{\pi_0(\tilde{x}(q_2|\eta_1, \eta_2))}{\partial \tilde{x}} M_2(q_2) + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor} \frac{\partial(\tilde{x}(w|\eta_1, \eta_2))}{\partial \tilde{x}} M_2(w) dw \\ &+ \rho \left[\sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{P\{-\tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2), \rho, q_2, \varepsilon\}}{\partial \tilde{x}} M_2(w) dw \right. \\ &+ \sum_{l_1=1}^{q_1} \frac{(P\{a_2 - \eta_2^{l_2}, \rho, q_2, \varepsilon\} + P\{\eta_2^{l_2} - b_2, \rho, q_2, \varepsilon\})}{\partial \eta_2} \\ &\left. + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{\phi_2^{l_2} \phi_1^{\lfloor \nu_2(w) \rfloor}}{q_2} \left\{ \frac{\partial[\lambda_2(2P\{\eta_2^{l_2}, \rho, q_2, \varepsilon\} - \eta_2^{l_2})]}{\partial \eta_2} \right\} dw \right]. \end{aligned}$$

Theorem 7. For each $(\eta_1, \eta_2, \phi_1, \phi_2)$, we have

$$\frac{\partial \tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2)}{\partial \phi_1} = N_1(w|\eta_1, \eta_2, \phi_1, \phi_2), \quad w \in [0, q_2],$$

where $N_1(\cdot|\eta_1, \eta_2, \phi_1, \phi_2)$ is the solution to the following system on $w \in [l_2 - 1, l_2]$, $l_2 \in I_{q_2}$:

$$\begin{cases} \dot{N}_1(w) = \phi_2^{l_2} \left\{ \frac{\phi_1^{\lfloor \nu_2(w) \rfloor} \partial f(w, \hat{x}(w), \eta_1^{\lfloor \nu_2(w) \rfloor}, \eta_2^{l_2})}{\partial \tilde{x}} N_1(w) + \frac{\phi_1^{\lfloor \nu_2(w) \rfloor} \partial f(w, \hat{x}(w), \eta_1^{\lfloor \nu_2(w) \rfloor}, \eta_2^{l_2})}{\partial \phi_1} \right\}, \\ N_1(0) = 0. \end{cases} \quad (5.5)$$

Based on Theorem 7, the gradient of J_2 with respect to ϕ_1 is given as follows:

$$\begin{aligned} \frac{\partial J_2}{\partial \phi_1} = & \frac{\pi_0(\tilde{x}(q_2|\eta_1, \eta_2, \phi_1, \phi_2))}{\partial \phi_1} N_1(q_2) + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \frac{\phi_1^{\lfloor v_2(w) \rfloor} \vartheta(\tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2))}{\partial \phi_1} N_1(w) dw \\ & + \rho \left[\sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{P\{-\tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2), \rho, q_2, \varepsilon\}}{\partial \phi_1} N_1(w) dw \right. \\ & \left. + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{\phi_2^{l_2}}{q_2} \left\{ \frac{\phi_1^{\lfloor v_2(w) \rfloor} \partial[\lambda_1(2P\{\eta_1^{\lfloor v_2(w) \rfloor}, \rho, q_2, \varepsilon\} - \eta_1^{\lfloor v_2(w) \rfloor}) + \lambda_2(2P\{\eta_2^{\lfloor v_2(w) \rfloor}, \rho, q_2, \varepsilon\} - \eta_2^{\lfloor v_2(w) \rfloor})]}{\partial \phi_1} \right\} dw \right]. \end{aligned}$$

Remark 4. For this, we introduce the inverse of $v_2(w)$. The gradient of J_2 with respect to ϕ_2 is complicated to derive because the switching time is in the time range of w . $v_2^{-1}(p)$ is a function of ϕ_2 , and at every connection point, this inverse function is not differentiable. $v_2^{-1}(r)$, $r \in I_{q_1}$ is represented by w_r , and the gradient of w_r with respect to $\phi_2^{l_2}$, $l_2 \in I_{q_2}$ at $r \in I_{q_1-1}$ is given in Remark 5 as follows.

Remark 5. If point w_r coincide with $w \in I_{q_2-1}$, then we have

$$\frac{\partial w_r^-}{\partial \phi_2^{l_2}} := \begin{cases} 0 & \text{if } l_2 > e, \\ (-p + \sum_{j=1}^{e-1} \phi_2^j) / (\phi_2^e)^2, & \text{if } l_2 = e, \\ -\frac{1}{\phi_2^{e+1}}, & \text{if } l_2 < e, \end{cases} \quad (5.6)$$

and

$$\frac{\partial w_r^+}{\partial \phi_2^{l_2}} := \begin{cases} 0 & \text{if } l_2 > e + 1, \\ (-p + \sum_{j=1}^e \phi_2^j) / (\phi_2^{e+1})^2, & \text{if } l_2 = e + 1, \\ -\frac{1}{\phi_2^{e+1}}, & \text{if } l_2 < e + 1. \end{cases} \quad (5.7)$$

Remark 6. If point w_h does not coincide with $w \in I_{q_2-1}$, then we have $\frac{\partial w_r^+}{\partial \phi_2} = \frac{\partial w_r^-}{\partial \phi_2}$.

With the above discussion, the gradient of states with respect to ϕ_2 is given below.

Theorem 8. For each $(\eta_1, \eta_2, \phi_1, \phi_2)$, we have

$$\frac{\partial \tilde{x}(w|\eta_1, \eta_2, \phi_1, \phi_2)}{\partial \phi_2} = N_2(w|\eta_1, \eta_2, \phi_1, \phi_2), \quad w \in [0, q_2],$$

where $N_2(\cdot|\eta_1, \eta_2, \phi_1, \phi_2)$ is the solution to the following system on $w \in [l_2 - 1, l_2]$, $l_2 \in I_{q_2}$:

$$\begin{cases} \dot{N}_2(w) = \phi_1^{\lfloor v_2(w) \rfloor} \left\{ \frac{\phi_2^{l_2} \partial f(w, \hat{x}(w), \eta_1^{\lfloor v_2(w) \rfloor}, \eta_2^{l_2})}{\partial \tilde{x}} N_2(w) + \frac{\phi_2^{l_2} \partial f(w, \hat{x}(w), \eta_1^{\lfloor v_2(w) \rfloor}, \eta_2^{l_2})}{\partial \phi_2} \right\}, \\ N_2(w_r^+) = N_2(w_r^-) - H_{r,f,h}^+ + H_{r,f,h}^-, \\ N_2(0) = 0, \end{cases} \quad (5.8)$$

with

$$H_{r,f,h}^- = \phi_1^{\lfloor \nu_2(w_h^-) \rfloor} \phi_2^{\lfloor w_h^- \rfloor} \left[f(\tilde{x}(w_h^-)) + h_1(\tilde{x}(w_h^-)) \eta_1^{\lfloor \nu_2(w_h^-) \rfloor} + h_2(\tilde{x}(w_h^-)) \eta_2^{\lfloor w_h^- \rfloor} \right] \frac{\partial w_r^-}{\partial \phi_2} \text{ and}$$

$$H_{r,f,h}^+ = \phi_1^{\lfloor \nu_2(w_h^+) \rfloor} \phi_2^{\lfloor w_h^+ \rfloor} \left[f(\tilde{x}(w_h^+)) + h_1(\tilde{x}(w_h^+)) \eta_1^{\lfloor \nu_2(w_h^+) \rfloor} + h_2(\tilde{x}(w_h^+)) \eta_2^{\lfloor w_h^+ \rfloor} \right] \frac{\partial w_r^+}{\partial \phi_2}.$$

Based on Theorem 8, the gradient of J_2 with respect to ϕ_2 is given as follows:

$$\begin{aligned} \frac{\partial J_2}{\partial \phi_2} = & \frac{\pi_0(\tilde{x}(q_2 | \eta_1, \eta_2, \phi_1, \phi_2))}{\partial \phi_2} N_2(q_2) + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \phi_2^{l_2} \frac{\phi_1^{\lfloor \nu_2(w) \rfloor} \vartheta(\tilde{x}(w | \eta_1, \eta_2, \phi_1, \phi_2))}{\partial \phi_2} N_2(w) dw \\ & + \rho \left[\sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{P\{-\tilde{x}(w | \eta_1, \eta_2, \phi_1, \phi_2), \rho, q_2, \varepsilon\}}{\partial \phi_2} N_2(w) dw \right. \\ & + \sum_{l_2=1}^{q_2} \int_{l_2-1}^{l_2} \frac{\phi_1^{\lfloor \nu_2(w) \rfloor}}{q_2} \left\{ \frac{\phi_2^{l_2} \partial [\lambda_1 (2P\{\eta_1^{\lfloor \nu_2(w) \rfloor}, \rho, q_2, \varepsilon\} - \eta_1^{\lfloor \nu_2(w) \rfloor}) + \lambda_2 (2P\{\eta_2^{l_2}, \rho, q_2, \varepsilon\} - \eta_2^{l_2})]}{\partial \phi_2} \right\} dw \\ & \left. - H_{r,f,h}^+ + H_{r,f,h}^- \right]. \end{aligned}$$

where H_{r,λ_i}^+ and H_{r,λ_i}^- are defined as $H_{r,f,h}^+$ and $H_{r,f,h}^-$, respectively.

Remark 7. The gradient formulae of the cost function $J_2(\eta_1, \eta_2, \phi_1, \phi_2, \rho, q_1, q_2, \varepsilon)$ with respect to the quadruple $(\eta_1, \eta_2, \phi_1, \phi_2) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \mathbb{R}^{q_2}$ is calculated by the variational method (Theorems 5–8). Therefore, the gradient-based algorithm easily obtains the optimal solution for Problem E.

6. Numerical results

We consider a nonlinear maximum hands-off control problem with the cost functional defined by

$$\min_i h_0 = \int_0^1 \{-6x_1^2(t) - 12x_2(t) + 3u_1(t) + u_2(t)\} dt, \quad (6.1)$$

governed by the nonlinear dynamical system

$$\begin{cases} \dot{x}_1(t) = u_2(t), \\ \dot{x}_2(t) = -x_1^2(t) + u_1(t), t \in [0, 1], \\ \mathbf{x} = (1, 0)^T. \end{cases} \quad (6.2)$$

with the maximum hands-off control constraint

$$A_1(u_1, u_2) = \sum_{i=1}^2 \lambda_i \int_0^1 |u_i(t)| dt \leq \mu = 10, \quad (6.3)$$

and the bound constraints of the control inputs

$$-1 \leq u_j(t) \leq 1, t \in [0, 1], j \in I_2.$$

Table 1. Type, control strategies, optimal cost, and $A_1(u_1, u_2)$.

Type	Control strategies	Optimal cost	$A_1(u_1, u_2)$
Case 1	\mathbf{u}^{1*}	$h_0(\mathbf{u}^{1*}) = 0.3392$	$A_1(\mathbf{u}^{1*}) = 21.2052$
Case 2	\mathbf{u}^{2*}	$h_0(\mathbf{u}^{2*}) = 0.3737$	$A_1(\mathbf{u}^{2*}) = 7.303$

Case 1: The optimal control problem governed by nonlinear dynamical system (6.2) without the maximum hands-off control constraint (6.3).

Case 2: The optimal control problem governed by nonlinear dynamical system (6.2) with the maximum hands-off control constraint (6.3).

For Cases 1 and 2, the number of segments for the two controls are, respectively, 10 and 20. For Cases 1 and 2, by using the proposed method we respectively obtain the optimal control strategies \mathbf{u}^{1*} and \mathbf{u}^{2*} plotted in Figures 1 and 2. Their corresponding optimal cost and sparsity levels are given in Table 1. From Figure 2, it is observed that for the first control, sparse control happened in the first and from the fourth to the tenth segments. For the second control, sparse control occurred in the second, the sixteenth, from the ninth to the thirteenth, and from the eighteenth to the twentieth segments.

The graphical representation in Figure 1 illustrates that the optimal control strategies demonstrate a considerable density. In contrast, Figure 2 portrays these strategies with a noticeable level of sparsity. Nevertheless, it is essential to highlight that the inclusion of the maximum hands-off control constraint (6.3) results in a marginal increase in the cost function value compared to the scenario without this constraint (6.3). Specifically, the cost function slightly exceeds its counterpart when the maximum hands-off control constraint is imposed.

The observed trend in Figure 2 reveals a noteworthy pattern in the sparsity of the sparse optimal control strategies. It becomes evident that there is a rapid initial increase in sparsity, accompanied by a relatively minor escalation in cost. Therefore, based on the findings presented in this paper, it can be inferred that the proposed method exhibits the capability to generate solutions of superior quality.

Furthermore, from Table 1, it is pertinent to note that the value of the cost function, specifically $h_0(\mathbf{u}^{2*})$, slightly exceeds that of $h_0(\mathbf{u}^{1*})$. The smaller the value of $A_1(\mathbf{u}^*)$, the sparser \mathbf{u}^* becomes. From Table 1, it follows that \mathbf{u}^{2*} is sparser than \mathbf{u}^{1*} . So, we can conclude that this proposed method achieves a balance between system performance and sparsity, thereby suggesting its efficacy in providing solutions that optimize both aspects effectively.

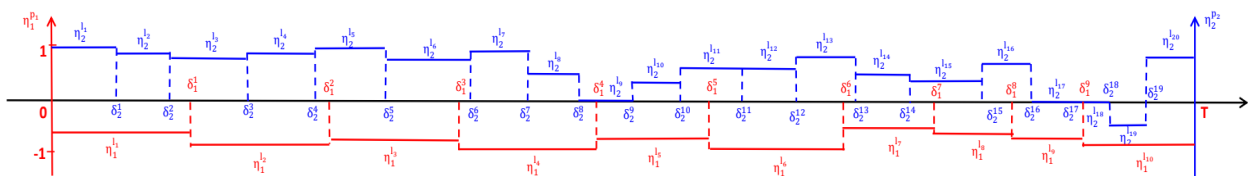


Figure 1. The optimal control strategies \mathbf{u}^{1*} without the maximum hands-off control constraint (6.3).

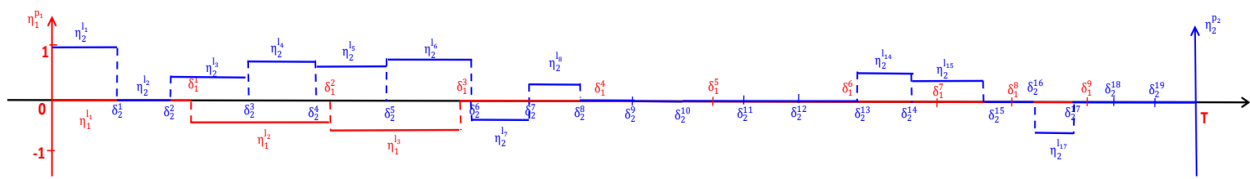


Figure 2. The sparse optimal control strategies \mathbf{u}^{2*} with the maximum hands-off control constraint (6.3).

7. Conclusions

Solving the sparse optimal control problem within the framework of linear dynamical systems often allows for an analytical solution. However, when addressing the maximum hands-off control problem within the domain of nonlinear dynamical systems, obtaining an analytical solution proves to be considerably challenging. This paper necessitates the numerical solution of the sparse optimal control problem which is governed by nonlinear dynamical systems. We employ control parameterization in conjunction with the sequential adaptive switching time optimization technique to approximate the maximum hands-off control problem by a sequence of finite-dimensional optimization problem. This approach allows for varying control switching times without imposing uniformity and offers flexibility in selecting control components. The resolution of the sparse optimal control problem relies on a gradient-based algorithm. To showcase the efficacy of the proposed methodology, an illustrative example is provided.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB3304600; in part by the National Natural Science Foundation of China under Grants 11901075, 12271307 and 12161076; in part by the China Postdoctoral Science Foundation under Grant 2019M661073; in part by the Fundamental Research Funds for the Central Universities under Grants 3132022201, 3132023206, 3132023535 and DUT22LAB305; in part by the Guangdong Province Natural Science Foundation of China under Grant 2022A1515011761; in part by Chongqing Natural Science Foundation Innovation and Development Joint Fund (CSTB2022NSCQ-LZX0040); in part by Ministry of Higher Education (MoHE) Malaysia through the Foundation Research Grant Scheme (FRGS/1/2021/STG06/SYUC/03/1); in part by the Shandong Province Natural Science Foundation of China under Grant ZR2023MA054; in part by the Fundamental Scientific Research Projects of Higher Education Institutions of Liaoning Provincial Department of Education under JYTMS20230165 (General Project); and in part by the Xinghai Project of Dalian Maritime University.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. M. Nagahara, D. E. Quevedo, D. Nesic, Maximum hands-off control: a paradigm of control effort minimization, *IEEE Trans. Autom. Control*, **61** (2016), 735–747. <https://doi.org/10.1109/TAC.2015.2452831>
2. J. Huang, Y. Shi, Guaranteed cost control for multi-sensor networked control systems using historical data, in *2012 American Control Conference (ACC)*, (2012), 4927–4932. <https://doi.org/10.1109/ACC.2012.6315279>
3. D. L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory*, **52** (2006), 1289–1306. <https://doi.org/10.1109/TIT.2006.871582>
4. J. Wu, F. Liu, L. C. Jiao, X. Wang, Compressive sensing SAR image reconstruction based on Bayesian framework and evolutionary computation, *IEEE Trans. Image Process.*, **20** (2011), 1904–1911. <https://doi.org/10.1109/TIP.2010.2104159>
5. L. Wang, Q. Wang, J. Wang, X. Zhang, Fast band-limited sparse signal reconstruction algorithms for big data processing, *IEEE Sens. J.*, **23** (2023), 13084–13099. <https://doi.org/10.1109/JSEN.2023.3268295>
6. Y. Shi, Y. Gao, S. Liao, D. Zhang, Y. Gao, D. Shen, A learning based CT prostate segmentation method via joint transductive feature selection and regression, *Neurocomputing*, **173** (2016), 317–331. <https://doi.org/10.1016/j.neucom.2014.11.098>
7. S. Weng, Editorial: spectroscopy, imaging and machine learning for crop stress, *Front. Plant Sci.*, **14** (2023). <https://doi.org/10.3389/fpls.2023.1240738>
8. S. Wang, H. Chen, W. Kong, X. Wu, Y. Qian, K. Wei, A modified FGL sparse canonical correlation analysis for the identification of Alzheimer’s disease biomarkers, *Electron. Res. Arch.*, **31** (2023), 882–903. <https://doi.org/10.3934/era.2023044>
9. F. Xu, Z. Lu, Z. Xu, An efficient optimization approach for a cardinality-constrained index tracking problem, *Optim. Methods Software*, **31** (2016), 258–271. <https://doi.org/10.1080/10556788.2015.1062891>
10. S. Bahmani, P. T. Boufounos, B. Raj, Learning model-based sparsity via projected gradient descent, *IEEE Trans. Inf. Theory*, **62** (2016), 2092–2099. <https://doi.org/10.1109/TIT.2016.2515078>
11. S. N. Negahban, P. Ravikumar, M. J. Wainwright, B. Yu, A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers, *Stat. Sci.*, **27** (2012), 538–557. <https://doi.org/10.1214/12-STS400>
12. E. J. Candes, T. Tao, Decoding by linear programming, *IEEE Trans. Inf. Theory*, **51** (2005), 4203–4215. <https://doi.org/10.1109/TIT.2005.858979>
13. S. Foucart, M. J. Lai, Sparsest solutions of underdetermined linear systems via l_q minimization for $0 \leq q \leq 1$, *Appl. Comput. Harmon. Anal.*, **26** (2009), 395–407. <https://doi.org/10.1016/j.acha.2008.09.001>

14. R. Chartrand, V. Staneva, Restricted isometry properties and nonconvex compressive sensing, *Inverse Probl.*, **24** (2008), 035020. <https://doi.org/10.1088/0266-5611/24/3/035020>
15. G. Gasso, A. Rakotomamonjy, S. Canu, Recovering sparse signals with a certain family of non-convex penalties and DC programming, *IEEE Trans. Signal Process.*, **57** (2009), 4686–4698. <https://doi.org/10.1109/TSP.2009.2026004>
16. H. A. L. Thi, H. M. Le, P. D. Tao, Feature selection in machine learning: an exact penalty approach using a difference of convex function algorithm, *Mach. Learn.*, 101 (2015), 163–186. <https://doi.org/10.1007/s10994-014-5455-y>
17. W. Peng, T. Gu, Y. Zhuang, Z. He, C. Han, Pattern synthesis with minimum mainlobe width via sparse optimization, *Digital Signal Process.*, **128** (2022), 103632. <https://doi.org/10.1016/j.dsp.2022.103632>
18. J. Liang, X. Zhu, C. Yue, Z. Li, B. Qu, Performance analysis on knee point selection methods for multi-objective sparse optimization problems, in *2018 IEEE Congress on Evolutionary Computation (CEC)*, (2018), 1–8. <https://doi.org/10.1109/CEC.2018.8477915>
19. X. Xiu, Z. Miao, W. Liu, A sparsity-aware fault diagnosis framework focusing on accurate isolation, *IEEE Trans. Ind. Inf.*, **19** (2022), 1356–1365. <https://doi.org/10.1109/TII.2022.3180070>
20. C. V. Rao, Sparsity of linear discrete-time optimal control problems with L^1 objectives, *IEEE Trans. Autom. Control*, **63** (2017), 513–517. <https://doi.org/10.1109/TAC.2017.2732286>
21. M. Babazadeh, Regularization for optimal sparse control structures: a primal-dual framework, in *2021 American Control Conference (ACC)*, (2021), 3850–3855. <https://doi.org/10.23919/ACC50511.2021.9482729>
22. P. Benner, H. Faßbender, On the numerical solution of large-scale sparse discrete-time Riccati equations, *Adv. Comput. Math.*, **35** (2011), 119–147. <https://doi.org/10.1007/s10444-011-9174-7>
23. F. S. Aktacs, O. Ekmekcioglu, M. C. Pinar, Provably optimal sparse solutions to overdetermined linear systems with non-negativity constraints in a least-squares sense by implicit enumeration, *Optim. Eng.*, **22** (2021), 2505–2535. <https://doi.org/10.1007/s11081-021-09676-2>
24. B. Polyak, A. Tremba, Sparse solutions of optimal control via Newton method for under-determined systems, *J. Global Optim.*, **76** (2020), 613–623. <https://doi.org/10.1007/s10898-019-00784-z>
25. D. Kalise, K. Kunisch, Z. Rao, Infinite horizon sparse optimal control, *J. Optim. Theory Appl.*, **172** (2017), 481–517. <https://doi.org/10.1007/s10957-016-1016-9>
26. P. Budhraja, A. S. A. Dilip, Maximum hands-off control for a class of nonlinear systems, in *2021 29th Mediterranean Conference on Control and Automation (MED)*, (2021), 1003–1006. <https://doi.org/10.1109/MED51440.2021.9480171>
27. W. Ji, Optimal control problems with time inconsistency, *Electron. Res. Arch.*, **31** (2023), 492–508. <https://doi.org/10.3934/era.2023024>
28. P. Yu, S. Tan, J. Guo, Y. Song, Data-driven optimal controller design for sub-satellite deployment of tethered satellite system, *Electron. Res. Arch.*, **32** (2024), 505–522. <https://doi.org/10.3934/era.2024025>

29. K. L. Teo, B. Li, C. Yu, V. Rehbock, *Applied and Computational Optimal Control: A Control Parametrization Approach*, Springer: Optimization and Its Applications, 2021. <https://doi.org/10.1007/978-3-030-69913-0>
30. S. Su, M. Shao, C. Yu, K. L. Teo, On the correlation of local collocation and control parameterization methods, *J. Ind. Manage. Optim.*, 2024. <https://doi.org/10.3934/jimo.2024004>
31. Q. Lin, R. Loxton, K. L. Teo, The control parameterization method for nonlinear optimal control: a survey, *J. Ind. Manage. Optim.*, **10** (2014), 275–309. <https://doi.org/10.3934/jimo.2014.10.275>
32. C. Xu, H. Li, Two-grid methods of finite element approximation for parabolic integro-differential optimal control problems, *Electron. Res. Arch.*, **31** (2023), 4818–4842. <https://doi.org/10.3934/era.2023247>
33. Y. Yuan, C. Liu, Optimal control for the coupled chemotaxis-fluid models in two space dimensions, *Electron. Res. Arch.*, **29** (2021), 4269–4296. <https://doi.org/10.3934/era.2021085>
34. Z. Z. Tao, B. Sun, A feedback design for numerical solution to optimal control problems based on Hamilton-Jacobi-Bellman equation, *Electron. Res. Arch.*, **29** (2021), 3429–3447. <https://doi.org/10.3934/era.2021046>
35. X. Pang, H. Song, X. Wang, J. Zhang, Efficient numerical methods for elliptic optimal control problems with random coefficient, *Electron. Res. Arch.*, **28** (2020), 1001–1022. <https://doi.org/10.3934/era.2020053>
36. H. W. J. Lee, K. L. Teo, X. Q. Cai, An optimal control approach to nonlinear mixed integer programming problems, *Comput. Math. Appl.*, **36** (1998), 87–105. [https://doi.org/10.1016/S0898-1221\(98\)00131-X](https://doi.org/10.1016/S0898-1221(98)00131-X)
37. A. Sibirian, V. Rehbock, Numerical procedure for solving a class of singular optimal control problems, *Optim. Methods Software*, **19** (2004), 413–426. <https://doi.org/10.1080/10556780310001656637>
38. G. Vossen, Switching time optimization for bang-bang and singular controls, *J. Optim. Theory Appl.*, **144** (2010), 409–429. <https://doi.org/10.1007/s10957-009-9594-4>
39. X. Zhu, C. Yu, K. L. Teo, Sequential adaptive switching time optimization technique for optimal control problems, *Automatica*, **146** (2022), 110565. <https://doi.org/10.1016/j.automatica.2022.110565>
40. C. Liu, R. Loxton, K. L. Teo, S. Wang, Optimal state-delay control in nonlinear dynamic systems, *Automatica*, **135** (2022), 109981. <https://doi.org/10.1016/j.automatica.2021.109981>
41. C. Liu, Z. Gong, C. Yu, S. Wang, K. L. Teo, Optimal control computation for nonlinear fractional time-delay systems with state inequality constraints, *J. Optim. Theory Appl.*, **191** (2021), 83–117. <https://doi.org/10.1007/s10957-021-01926-8>
42. C. Liu, C. Yu, Z. Gong, H. T. Cheong, K. L. Teo, Numerical computation of optimal control problems with Atangana-Caputo fractional derivatives, *J. Optim. Theory Appl.*, **197** (2023), 798–816. <https://doi.org/10.1007/s10957-023-02212-5>

43. C. Yu, K. H. Wong, An enhanced control parameterization technique with variable switching times for constrained optimal control problems with control-dependent time-delayed arguments and discrete time-delayed arguments, *J. Comput. Appl. Math.*, **427** (2023), 115106. <https://doi.org/10.1016/j.cam.2023.115106>
44. D. Wu, Y. Chen, C. Yu, Y. Bai, K. L. Teo, Control parameterization approach to time-delay optimal control problems: a survey, *J. Ind. Manage. Optim.*, **19** (2023), 3750–3783. <https://doi.org/10.3934/jimo.2022108>
45. D. Wu, Y. Bai, F. Xie, Time-scaling transformation for optimal control problem with time-varying delay, *Discrete Contin. Dyn. Syst. - Ser. S*, **13** (2020), 1683–1695. <https://doi.org/10.3934/dcdss.2020098>
46. P. Liu, X. Li, X. Liu, Y. Hu, An improved smoothing technique-based control vector parameterization method for optimal control problems with inequality path constraints, *Optim. Control. Appl. Methods*, **38** (2017), 586–600. <https://doi.org/10.1002/oca.2273>



AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)