Nonlinear Analysis: Hybrid Systems 37 (2020) 100902

Contents lists available at ScienceDirect

Nonlinear Analysis: Hybrid Systems

journal homepage: www.elsevier.com/locate/nahs

Hybrid parametric minimum principle

Jia-Jiang Lin, Xiong-Lin Luo *

Department of Automation, China University of Petroleum, Beijing 102249, PR China

ARTICLE INFO

Article history:
Received 14 May 2019
Received in revised form 21 November 2019
Accepted 6 April 2020
Available online xxxx

Keywords:
Hybrid systems
Pontryagin maximum principle
Optimal control problem
Discrete control
Hybrid minimum principle

ABSTRACT

This paper discusses a special kind of hybrid system, named hybrid parametric system, which subsystem is characterized by parameters of model function. The parameters of model function represent the batch operation (discrete control). The joint optimization of both continuous and discrete control variables is named hybrid parametric optimal control problem. The corresponding necessary optimality conditions are established in the form of the hybrid parametric minimum principles, which are proposed and proved in this paper. The first principle, i.e. single stage hybrid parametric minimum principle, is proved based on the similar technique used by the proof of the classical Pontryagin maximum principle, and the second principle, multistage hybrid parametric minimum principle, is proved by reducing it into a nominal case of first principle. Moreover, in the virtue of the continuity of parameter space, the hybrid parametric minimum principle only requires a prescribed number of switching times.

1. Introduction

Hybrid systems [1–3] are systems whose dynamics involve a combination of continuous evolution and discrete transitions. More specifically, hybrid systems are described by a collection of control systems and a finite sequence of times (called switching times) which partitions the time interval into subintervals. On each subinterval, the state of the system flows in accordance with one of the systems from a given collection; at a switching time, the state experiences an instantaneous jump, and another system from the collection is selected for the next subinterval. While providing a richer modeling framework than the continuous control systems, hybrid systems are definitely more difficult to analysis and control.

The optimal control problem is one of the most important issues in the study of control systems. The minimum principle, also called the maximum principle in the pioneering work of Pontryagin et al. [4], is a milestone of control theory. A general framework of hybrid optimal control problem [5] is established and the minimum principle is generalized for hybrid system, which is called the hybrid minimum principle [6–9]. To ensure the validity of the hybrid minimum principle in a non-degenerate form, some hypotheses are introduced [10,11]. Then by performing first order variational analysis via the needle variation methodology, the necessary optimality conditions are established.

When the continuous control is absent, the system can only be optimized by choosing switching sequence and switching times [12,13]. Another important special case arises when the dynamical system is piecewise affine and the cost functional is quadratic [14,15]. When the differentiability of hybrid optimal control problems takes a weaker form, the corresponding nonsmooth hybrid minimum principle could also be obtained [16]. Moreover, the minimum principle could also be extended to a general class of stochastic hybrid systems [17–19] with state dependent diffusion fields which are subject to autonomous and controlled switchings and state jumps.

* Corresponding author.
E-mail address: luoxl@cup.edu.cn (X.-L. Luo).

https://doi.org/10.1016/j.nahs.2020.100902
1751-570X/

Since the dynamic programming is of equal importance as the minimum principle, a unified general framework of hybrid dynamic programming [20–24] is also studied. Within this general framework, it is proved that along optimal trajectories of a hybrid system, the costate in the hybrid minimum principle, and the gradient of the value function in hybrid dynamic programming are equal almost everywhere, which is also true for classical control problem [25,26]. In analytic mechanics, if there are holonomic constraints, the dimensions of mechanical system could be reduced by generalized coordinates. In other words, the mechanical system is considered in a submanifold of original Euclidean space. This intrinsic representation has a higher level of generality. Hence, casting the maximum principle in the framework of manifolds greatly elucidates the essence of the maximum principle [27,28]. Conceivably, the hybrid minimum principle could also be described in geometric language [29,30].

The hybrid systems usually refer to the systems composed by finite subsystems. Such hybrid systems arise in various application domains, including robotics [31], production control [32], power converters [33], hybrid vehicle [34]. However, there is another kind of hybrid system in chemical industry, which switching is caused by the batch operations, such as fed-batch operation of bioprocess [35,36] and the addition of CO promoter in the fluid catalytic cracking unit (previous work by the authors) [37–39]. The batch operations, which take places only in some time instants, change system's model function during a time span. Since the batch operation could be regarded as the parameters of the main model function, the subsystem could be represented by these parameters. Therefore, it is called hybrid parametric system. In continuous chemical industry, the batch operation is usually treated as an auxiliary means to ensure system's security and stability, while the continuous operation is the main method to maximum the unit's profit. Since the batch operation means consumption of some resources, the hybrid parametric optimal control problem usually has an extra cost term, and the separated optimization of continuous and discrete control fails to exploit the potential profit. Thus, the joint optimization of both continuous and discrete control variables is needed, which requires the optimality conditions for the hybrid parametric optimal control problem, i.e. the hybrid parametric minimum principle (Theorem 3).

The existing hybrid minimum principle [6–9] is constructed for hybrid systems consisting finite subsystems. Since the topology of these hybrid systems is separate, variations of a model sequence generate trajectories "far" from the original one. Hence, the hybrid minimum principle only gives local necessary optimality conditions for a given model sequence. For hybrid parametric system, the subsystems are infinite and can be indexed by the discrete control. Since the integration optimization of continuous and batch operations wishes to provide an optimal trajectory of continuous operations and an optimal sequence of batch operations, the existing hybrid minimum principle is useless in this case, as it requires a given sequence of batch operations. In fact, the index of hybrid parametric system can be treated as a parameter of model function, and it is topologically continuous. Hence, stronger optimality conditions could be expected for the hybrid parametric optimal control problem, which is given by Theorem 3. Without requiring a given model sequence, hybrid parametric minimum principle gives globally necessary optimality conditions for the hybrid parametric optimal control problem under given switching times, but it has an extra property (d) than the hybrid minimum principle. Furthermore, the proof of Theorem 3 is based on Theorem 2, namely single stage hybrid parametric minimum principle, and Theorem 2 extends the classical Pontryagin maximum principle [4] to the hybrid parametric optimal control case. The canonical optimal control problem assumes that an admissible control could take any value of control set at any time, which is the reason that guarantees the needle-like variation works, while the hybrid parametric optimal control problem has an extra global constraint on the batch operation (discrete control), i.e. constancy constraint. Hence, the extension of Pontryagin maximum principle is needed, which gives rise to Theorem 2. As the requirements on the discrete control are stronger, the minimality condition for the discrete control (property (d) of Theorem 2) takes a weaker form than the minimality condition of continuous control (property (c) of Theorem 2).

The rest of this paper is organized as follows: In Section 2.1, the conception and mathematical description of hybrid parametric optimal control problem are proposed; In Section 2.2, the single stage hybrid parametric minimum principle is proposed and proved based on the similar technique used by the proof of the classical Pontryagin maximum principle. In addition, five remarks are given; In Section 2.3, multistage hybrid parametric minimum principle is proposed and proved by reducing it to a single stage hybrid parametric optimal control problem, and then four remarks are given; In Section 3, two analytic examples of hybrid parametric optimal control problems are provided, which corroborates the validity of the proposed principles.

2. Hybrid parametric minimum principle

The hybrid parametric optimal control Problem A is given at Section 2.1, then the corresponding optimality conditions described by Theorem 3 are given at Section 2.3, which is based on the intermediate Theorem 2 given at Section 2.2.

2.1. Statement of hybrid parametric optimal control problem

The hybrid control system consists a collection of (time-invariant) control systems

$$o\colon \{\dot{\boldsymbol{x}} = \boldsymbol{f}_q(\boldsymbol{x}(t), \boldsymbol{u}(t)), q \in Q\} \tag{1}$$

where Q is an index set; $\boldsymbol{x}(t) \in R^n$ is the state, and is an absolutely continuous function of t; $\boldsymbol{u}(t) \in U \subset R^m$ is the continuous control variable or continuous operation, where the control set U is a compact set. To guarantee local existence and uniqueness of solutions for (1), it is assumed that $\boldsymbol{f}(\cdot)$ is continuous in $\boldsymbol{u}(t)$ and C^1 in $\boldsymbol{x}(t)$; $\partial \boldsymbol{f}/\partial \boldsymbol{x}$ is continuous in $\boldsymbol{u}(t)$; and $\boldsymbol{u}(t)$ is a piecewise continuous function (Actually it only requires measurable and locally bounded). Obviously, $\boldsymbol{f}(\cdot)$ is locally Lipschitz in this case. For the time variant version, if $\boldsymbol{f}(\cdot)$ are C^1 in t, then it could be transformed into autonomous version by introducing the extra state variable $x_{n+1} := t$, with the dynamics $\dot{x}_{n+1} = 1$.

If the cardinality of set Q is finite, then the hybrid system is usually called switched system. The corresponding hybrid minimum principle [6–9] is given by assuming that the model sequence $\{q\}$ is defined and fixed in that the elements in control systems o are topologically separate. Variations of this sequence generate trajectories which are "far" from the given one. Hence, they cannot be compared by methods of analysis. In this paper, the cardinality of the set Q is $\aleph_1$. The model function $\boldsymbol{f}_q(\cdot)$ is distinguished by parameters $\overline{\boldsymbol{u}} \in \overline{U} \subset R^l$, where $\overline{U}$ is a compact set, and $\overline{\boldsymbol{u}}$ is also called discrete control variables or batch operation. Then control system o could be written as $\{\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}})\}$, named hybrid parametric system. Moreover, it is assumed that the control system o shares the same continuous control set U and discrete control set $\overline{U}$.

The second ingredient is a collection of switching constraint surfaces $S_{\boldsymbol{x}(t_i^-),\boldsymbol{x}(t_i^+)} \subset R^{2n}$, $i = 1, \ldots, k$, and an endpoint constraint surface $E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)} \subset R^{2n}$. In this paper, a surface D in R^n is defined by the equality constraints

$$h_1(\boldsymbol{x}) = h_2(\boldsymbol{x}) = \cdots = h_m(\boldsymbol{x}) = 0 \tag{2}$$

where $h_i(\boldsymbol{x})$, $i = 1, \ldots, m$, are C^1 functions from R^n to R. Assuming that the optimal solution $\boldsymbol{x}^*(t)$ is a regular point of constraint surface in the sense that the gradients $\nabla h_i(\boldsymbol{x})$, $i = 1, \ldots, m$, are linearly independent at $\boldsymbol{x}^*(t)$. Then, the tangent space of the constraint surface at $\boldsymbol{x}^*(t)$ can be describe as

$$T_{\boldsymbol{x}^*(t)}D = \{\boldsymbol{d} \in R^n \big| \left(\nabla h_i(\boldsymbol{x}^*(t))\right)^T \cdot \boldsymbol{d} = 0, i = 1, \ldots, m\} \tag{3}$$

where $\boldsymbol{d}$ is a tangent vector to D at $\boldsymbol{x}^*(t)$. The vector $\boldsymbol{v} \perp T_{\boldsymbol{x}^*(t)}D$ equals

$$\boldsymbol{v} \in \mathrm{span}\{\nabla h_i(\boldsymbol{x}), i = 1, \ldots, m\} \tag{4}$$

The cost functional has the form

$$J(\boldsymbol{u}(t), \overline{\boldsymbol{u}}_i, t_i) = \sum_{i=0}^{k} \left(\int_{t_i}^{t_{i+1}} L_i(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}}_i) dt + \phi_i(\overline{\boldsymbol{u}}_i) \right) + \sum_{i=1}^{k} \Phi_{i-1,i}(\boldsymbol{x}(t_i^-), \boldsymbol{x}(t_i^+)) \tag{5}$$

where $L_i : R^n \times U \times \overline{U} \to R$ is the running cost, which is continuously differentiable in $\boldsymbol{x}(t)$, $\phi_i : R^l \to R$ is the resource consumption term for discrete control variables $\overline{\boldsymbol{u}}_i$ and $\Phi_{i-1,i} : R^n \times R^n \to R$ is the switching cost between two periods, and is continuously differentiable. For simplicity, there is not a terminal cost, as terminal cost can be easily incorporated into the above running-plus-switching cost along the lines. Finite sequences $\{t_i\}$ and $\{\overline{\boldsymbol{u}}_i\}$ are referred as the time sequence and discrete control sequence, respectively.

A function $\boldsymbol{x} : [t_0, t_f] \to R^n$ is an admissible trajectory of hybrid parametric system corresponding to a continuous control $\boldsymbol{u} : [t_0, t_f] \to U$ and a sequence of discrete control $\overline{\boldsymbol{u}}_0, \overline{\boldsymbol{u}}_1, \ldots, \overline{\boldsymbol{u}}_k \in \overline{U}$ if there exist time instants $t_0 < t_1 < \cdots < t_k < t_{k+1} = t_f$ such that $\boldsymbol{x}(\cdot)$ satisfies

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}}_i), \forall \mathrm{t} \in (t_i, t_{i+1}), \quad i = 0, 1, \ldots, k \tag{6}$$

$$\begin{pmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_f) \end{pmatrix} \in E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)} \tag{7}$$

$$\begin{pmatrix} \boldsymbol{x}(t_i^-) \\ \boldsymbol{x}(t_i^+) \end{pmatrix} \in S_{\boldsymbol{x}(t_i^-),\boldsymbol{x}(t_i^+)}, \quad i = 1, \ldots, k \tag{8}$$

Here $\boldsymbol{x}(t_i^-)$ and $\boldsymbol{x}(t_i^+)$ are the values of $\boldsymbol{x}(\cdot)$ right before and right after t_i, respectively, and the value $\boldsymbol{x}(t_i)$ is taken to be equal to one of these one-sided limits, depending on the desired convention. The function $\overline{\boldsymbol{u}} : [t_0, t_f] \to \overline{U}$ defined by $\overline{\boldsymbol{u}}(t) := \overline{\boldsymbol{u}}_i$ for $t \in [t_i, t_{i+1})$ describes the evolution of $\overline{\boldsymbol{u}}$ along the trajectory. Then equation (6) could be rewritten more concisely as

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}}(t)) \text{ for } t \neq t_i, \quad i = 0, 1, \ldots, k \tag{9}$$

Hybrid parametric optimal control problem could be described as:

Problem A.

$$\min J(\boldsymbol{u}(t), \{\overline{\boldsymbol{u}}_i\}, \{t_i\}) = \sum_{i=0}^{k} \left(\int_{t_i}^{t_{i+1}} L_i(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}}_i) dt + \phi_i(\overline{\boldsymbol{u}}_i) \right) + \sum_{i=1}^{k} \Phi_{i-1,i}(\boldsymbol{x}(t_i^-), \boldsymbol{x}(t_i^+))$$

$$\text{s.t. } \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}}(t)) \text{ for } t \neq t_i, \quad i = 0, 1, \ldots, k.$$

$$\begin{pmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_f) \end{pmatrix} \in E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)}$$

$$\begin{pmatrix} \boldsymbol{x}(t_i^-) \\ \boldsymbol{x}(t_i^+) \end{pmatrix} \in S_{\boldsymbol{x}(t_i^-),\boldsymbol{x}(t_i^+)}, \; i = 1, \ldots, k$$

where the technical assumptions are satisfied as discussed before and t_0 is given (t_f can be free or given).

The hybrid parametric optimal control problem consists in finding a continuous control and two discrete sequences that minimizes the cost (5) subject to the constraints (7)–(9). To obtain optimality conditions for Problem A, the several theorems need to be established first.

2.2. Single stage hybrid parametric minimum principle

This section establishes a theorem, which gives the optimality conditions for a special case of Problem A ($k = 1$). The theorem that we are about to established is based on the following basic canonical autonomous optimal control problem and the corresponding minimum principle.

Problem B.

$$\begin{aligned} \min J(\boldsymbol{u}(t)) = & \int_{t_0}^{t_f} L(\boldsymbol{x}(t), \boldsymbol{u}(t))dt + \Phi(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)) \\ \text{s.t.} \;\; & \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) \\ & \begin{pmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_f) \end{pmatrix} \in E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)} \end{aligned}$$

where the technical assumptions are satisfied as discussed before and t_0 is given (t_f can be free or given).

The corresponding optimality conditions are given by the following theorem [4,40,41].

Theorem 1 (*Pontryagin Minimum Principle for Problem B*)**.** *Let $\boldsymbol{u}^*: [t_0, t_f] \to U$ be an optimal control (in the global sense) of Problem B and let $\boldsymbol{x}^*: [t_0, t_f] \to R^n$ be the corresponding optimal state trajectory. Then there exists a piecewise continuously differentiable vector function $\boldsymbol{p}^*: [t_0, t_f] \to R^n$, named costate, and having the following properties:*
(a) $\boldsymbol{p}^*(t)$ satisfies the adjoint equation:

$$\dot{\boldsymbol{p}}^* = -\frac{\partial}{\partial \boldsymbol{x}} H(\boldsymbol{x}^*, \boldsymbol{u}^*, \boldsymbol{p}^*) \tag{10}$$

where Hamiltonian $H: R^n \times U \times R^n \to R$ is defined as

$$H(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}) = L(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{p}^T(t) \cdot \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \tag{11}$$

(b) transversality condition:

$$\begin{pmatrix} \boldsymbol{p}^*(t_0) \\ -\boldsymbol{p}^*(t_f) \end{pmatrix} + \nabla \Phi(\boldsymbol{x}^*(t_0), \boldsymbol{x}^*(t_f)) \perp T_{\begin{pmatrix} \boldsymbol{x}^*(t_0) \\ \boldsymbol{x}^*(t_f) \end{pmatrix}} E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)} \tag{12}$$

(c) minimality condition: for all $t \in [t_0, t_f]$

$$\boldsymbol{u}^*(t) = \arg \min_{\boldsymbol{u}(t) \in U} H\left(\boldsymbol{x}^*, \boldsymbol{u}, \boldsymbol{p}^*\right) \tag{13}$$

(d) constancy condition ("energy preservation law"): for all $t \in [t_0, t_f]$

$$H\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \boldsymbol{p}^*\right) = c \tag{14}$$

where c is a constant. If t_f is free, then $c = 0$.

For hybrid parametric system, $\overline{\boldsymbol{u}}$ could be treated as the index of a subsystem or a discrete control variable for nominal control system. Instead of considering multistage hybrid parametric optimal control problem, single stage hybrid parametric optimal control Problem C is considered here, which is very similar to Problem B.

Problem C.

$$\begin{aligned} \min J(\boldsymbol{u}(t), \overline{\boldsymbol{u}}) = & \int_{t_0}^{t_f} L(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}})dt + \phi(\overline{\boldsymbol{u}}) + \Phi(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)) \\ \text{s.t.} \;\; & \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \overline{\boldsymbol{u}}) \\ & \begin{pmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_f) \end{pmatrix} \in E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)} \end{aligned}$$

where the technical assumptions are satisfied as discussed before and t_0 is given (t_f can be free or given).

The corresponding optimality conditions are given by Theorem 2.

Theorem 2 (*Single Stage Hybrid Parametric Minimum Principle for Problem C*)**.** *Let $\overline{\boldsymbol{u}}^* \in \overline{U}$ and $\boldsymbol{u}^*: [t_0, t_f] \to U$ be an optimal discrete and continuous control of Problem C (in the global sense) and let $\boldsymbol{x}^*: [t_0, t_f] \to R^n$ be the corresponding optimal state trajectory. Then there exists a piecewise continuously differentiable vector function $\boldsymbol{p}^*: [t_0, t_f] \to R^n$, named costate, and having the following properties:*
(a) $\boldsymbol{p}^*(t)$ satisfies the adjoint equation:

$$\dot{\boldsymbol{p}}^* = -\frac{\partial}{\partial \boldsymbol{x}} H(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*) \tag{15}$$

where Hamiltonian $H: R^n \times U \times \overline{U} \times R^n \to R$ is defined as

$$H(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}}, \boldsymbol{p}) = L(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}}) + \boldsymbol{p}^T(t) \cdot \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}}) \tag{16}$$

(b) transversality condition:

$$\begin{pmatrix} \boldsymbol{p}^*(t_0) \\ -\boldsymbol{p}^*(t_f) \end{pmatrix} + \nabla \Phi(\boldsymbol{x}^*(t_0), \boldsymbol{x}^*(t_f)) \perp T_{\begin{pmatrix} \boldsymbol{x}^*(t_0) \\ \boldsymbol{x}^*(t_f) \end{pmatrix}} E_{\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)} \tag{17}$$

(c) minimality condition for continuous control variable: for all $t \in [t_0, t_f]$

$$\boldsymbol{u}^*(t) = \arg \min_{\boldsymbol{u}(t) \in U} H\left(\boldsymbol{x}^*, \boldsymbol{u}, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*\right) \tag{18}$$

(d) minimality condition for discrete control variable:

$$\overline{\boldsymbol{u}}^* = \arg \min_{\overline{\boldsymbol{u}} \in \overline{U}} \left(\int_{t_0}^{t_f} H\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}, \boldsymbol{p}^*\right) dt + \phi(\overline{\boldsymbol{u}}) \right) \tag{19}$$

(e) constancy condition ("energy preservation law"): for all $t \in [t_0, t_f]$

$$H\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*\right) = c \tag{20}$$

where c is a constant. If t_f is free, then $c = 0$.

Theorem 2 will be proved based on similar techniques used by the proof of Theorem 1.

Proof of Theorem 2. Let function $\overline{\boldsymbol{u}}(t) := \overline{\boldsymbol{u}}$ for all $t \in [t_0, t_f]$, i.e. $\overline{\boldsymbol{u}}(t)$ is a constant function with value $\overline{\boldsymbol{u}}$. Introducing a new control variable

$$\boldsymbol{v}(t) = \left(\boldsymbol{u}^T(t) \quad \overline{\boldsymbol{u}}^T(t)\right)^T \in U \times \overline{U} \subset R^m \times R^l \tag{21}$$

then Problem C is transformed into Problem C′, which follows as

Problem C′.

$$\begin{aligned} \min J(\boldsymbol{v}(t)) = \int_{t_0}^{t_f} & L(\boldsymbol{x}, \boldsymbol{v}) dt + \phi(\boldsymbol{v}) + \Phi(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)) \\ \text{s.t. } & \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{v}(t)) \\ & \begin{pmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_f) \end{pmatrix} \in E_{\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)} \end{aligned}$$

where the technical assumptions are satisfied as Problem C, except that the last l components of control variable $\boldsymbol{v}(t)$ are constant.

Obviously, Problem C′ is very similar to Problem B. Following the same process as the proof of Theorem 1 [4,40,41], which introduces the Lagrange multiplier $\boldsymbol{p}(t)$, named costate, then the Hamiltonian, which is identical with (11), is given by

$$H(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{p}) = L(\boldsymbol{x}, \boldsymbol{v}) + \boldsymbol{p}^T(t) \cdot \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{v})$$

The optimality conditions are obtained by comparing optimal solution with perturbated trajectory, which requires applying variations $\delta \boldsymbol{v}(t)$, $\delta \boldsymbol{x}(t_0)$ and δt_f on the optimal solution. Since the original problem is a minimality problem, it follows

$$\delta J(\cdot) \geq 0 \tag{22}$$

By analyzing the (22) and the constraint conditions, the properties (a), (b), and (e) of Theorem 2 could be derived, because these properties have identical formations as Theorem 1. However, since the constraints of Problems C′ and B on the control variables are different, the deduction of properties (c) and (d) cannot be exactly copied. Fortunately, the following inequality still holds

$$\int_{t_0}^{t_f} H(\boldsymbol{x}^*, \boldsymbol{v}, \boldsymbol{p}^*)dt + \phi(\boldsymbol{v}) \geq \int_{t_0}^{t_f} H(\boldsymbol{x}^*, \boldsymbol{v}^*, \boldsymbol{p}^*)dt + \phi(\boldsymbol{v}^*) \tag{23}$$

where $\boldsymbol{v}(t)$ is the perturbated control variable. The properties (c) and (d) are obtained by applying two different kinds of perturbation.

For continuous control component $\boldsymbol{u}(t)$, which is the case considered in Theorem 1, the perturbation should be taken as needle-like variation (McShane variation), which has the form

$$\boldsymbol{u}(t) = \begin{cases} \boldsymbol{u}^*(t), \ \text{if}\, t_0 \leq t \leq t_i \\ \boldsymbol{u}', \ \text{if}\, t_i < t < t_i + \Delta t \\ \boldsymbol{u}^*(t), \ \text{if}\, t_i + \Delta t \leq t \leq t_f \end{cases} \tag{24}$$

where $\boldsymbol{u}' \in U$. Then the first kind of perturbated $\boldsymbol{v}(t)$ follows by

$$\boldsymbol{v}(t) = \left(\boldsymbol{u}^T(t) \quad \overline{\boldsymbol{u}}^{*T}(t)\right)^T \tag{25}$$

By (23), it follows that

$$\int_{t_i}^{t_i+\Delta t} H\left(\boldsymbol{x}^*, \left(\boldsymbol{u}'^T, \overline{\boldsymbol{u}}^{*^T}\right)^T, \boldsymbol{p}^*\right) dt \geq \int_{t_i}^{t_i+\Delta t} H\left(\boldsymbol{x}^*, \boldsymbol{v}^*, \boldsymbol{p}^*\right) dt \tag{26}$$

By mean value theorems for definite integrals, there exists $\tau \in [t_i, t_i + \Delta t]$,

$$H\left(\boldsymbol{x}^*(\tau), \left(\boldsymbol{u}'^T, \overline{\boldsymbol{u}}^{*^T}\right)^T, \boldsymbol{p}^*(\tau)\right) \Delta t \geq H\left(\boldsymbol{x}^*(\tau), \boldsymbol{v}^*(\tau), \boldsymbol{p}^*(\tau)\right) \Delta t \tag{27}$$

Let $\Delta t \to 0$ and t_i takes any time in $[t_0, t_f]$, it has

$$\boldsymbol{u}^*(t) = \arg \min_{\boldsymbol{u}(t)\in U} H\left(\boldsymbol{x}^*, \left(\boldsymbol{u}^T, \overline{\boldsymbol{u}}^{*^T}\right)^T, \boldsymbol{p}^*\right) \tag{28}$$

which completes the proof of property (c) of Theorem 2.

Since the $\overline{\boldsymbol{u}}(t)$ is a constant vector, the perturbation of this component cannot be taken arbitrarily at any time t. It must assure that the perturbated variable is still an element of the control set, which requires that the perturbation must be globally equal. Hence the perturbation of $\overline{\boldsymbol{u}}(t)$ should be described as

$$\overline{\boldsymbol{u}}(t) = \overline{\boldsymbol{u}}', \ \text{for}\, t_0 \leq t \leq t_f \tag{29}$$

where $\overline{\boldsymbol{u}}' \in \overline{U}$. Then the second kind of perturbated $\boldsymbol{v}(t)$ follows by

$$\boldsymbol{v}(t) = \left(\boldsymbol{u}^{*T}(t) \quad \overline{\boldsymbol{u}}^T(t)\right)^T \tag{30}$$

By (23), it follows that

$$\int_{t_0}^{t_f} H\left(\boldsymbol{x}^*, \left(\boldsymbol{u}^*, \overline{\boldsymbol{u}}'^T\right)^T, \boldsymbol{p}^*\right) dt + \phi(\overline{\boldsymbol{u}}') \geq \int_{t_0}^{t_f} H\left(\boldsymbol{x}^*, \boldsymbol{v}^*, \boldsymbol{p}^*\right) dt + \phi(\overline{\boldsymbol{u}}^*) \tag{31}$$

It is identical with the following expression

$$\overline{\boldsymbol{u}}^*(t) = \arg \min_{\substack{\overline{\boldsymbol{u}}(t) \in \overline{U} \\ \overline{\boldsymbol{u}}(t) = const}} \int_{t_0}^{t_f} H\left(\boldsymbol{x}^*, \left(\boldsymbol{u}^*, \overline{\boldsymbol{u}}\right)^T, \boldsymbol{p}^*\right) dt + \phi(\overline{\boldsymbol{u}}) \tag{32}$$

which completes the proof of property (d) of Theorem 2.

Remark 1. When $L(\cdot)$ and $\boldsymbol{f}(\cdot)$ is differentiable with respect to $\boldsymbol{u}(t)$ and (18) obtains the extremum at the interior point, which can be trivially satisfied when $\boldsymbol{u}(t) \in R^m$, then the (18) is identical with

$$\frac{\partial}{\partial \boldsymbol{u}} H\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*\right) = \boldsymbol{0} \tag{33}$$

Similarly, when $L(\cdot)$, $\boldsymbol{f}(\cdot)$ and $\phi(\cdot)$ is differentiable with respect to $\overline{\boldsymbol{u}}$ and (19) obtains the extremum at the interior point, which can be trivially satisfied when $\overline{\boldsymbol{u}} \in R^l$, then the (19) is identical with

$$\int_{t_0}^{t_f} \frac{\partial}{\partial \overline{\boldsymbol{u}}} H\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*\right) dt + \frac{\partial}{\partial \overline{\boldsymbol{u}}} \phi(\overline{\boldsymbol{u}}^*) = \boldsymbol{0} \tag{34}$$

Remark 2. Theorem 2 could be treated as an extension of Theorem 1 to hybrid parametric optimal control case. Traditionally, the control variable $\boldsymbol{u}(t)$ is a piecewise continuous function with range U, which means an admissible control of $\boldsymbol{u}(t)$ could take any value of U at any time $t \in [t_0, t_f]$. This property is exactly the reason that guarantees the needle-like variation works. Hence, even for time-variant $U(t)$, the minimality property (c) of Theorem 1 still holds. Unfortunately, there are some cases that the needle-like variation will not work, such as the constancy constraint discussed above. Those cases have extra global constraints besides pointwise constraint. For example, $\dot{\boldsymbol{u}} = \boldsymbol{0}$ for all $t \in [t_0, t_f]$, which is case considered by Problem C or $\int_{t_0}^{t_f} g(\boldsymbol{u})dt = c$, where $g(\cdot)$ is a measurable and essentially bounded function and c is a constant. The needle-like variation requires that the perturbation only takes place at small region while other regions remain unchanged. However, this perturbation will certainly break the global constraint of special control variable. Hence, the corresponding minimality property for those control variables only takes a weaker form, which is similar to the minimality property (d) of Theorem 2.

Remark 3. There are two well-known analysis methods to solve dynamic optimization problem, which are minimum principle and dynamical programming. The dynamical programming is based on the Bellman's principle of optimality, which states that an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. This classical theorem is based on the same assumption as the minimum principle that the control variable is free of global constraints. Therefor a sub-trajectory of a Problem C's optimal trajectory will not be guaranteed to be an optimal trajectory, i.e. Theorem 2 has not its counterpart in dynamical programming. The counterexample is given by the following example.

Considering a car travels through x_0 to x_f, which has the distance

$$S = x_f - x_0 \tag{35}$$

The speed, which is the model function, is described by

$$\dot{x} = \overline{u}u \tag{36}$$

where $\overline{u} \in [0, \overline{u}^u]$, i.e. the discrete control variable, is some sort of accelerator, and $u(t) \in [u_l, u^u]$ is the normal continuous control variable. The cost function follows as

$$\min J(u(t), \overline{u}) = \int_0^T dt + \overline{u} \tag{37}$$

which means to minimize travel time while saving accelerator resource. Apparently, for any $\overline{u} > 0$, the optimal continuous control variable is given by

$$u^*(t) = u^u > 0 \tag{38}$$

then cost function is given by

$$J(\overline{u}) = \frac{S}{\overline{u}u^u} + \overline{u} \tag{39}$$

Assuming $\overline{u}^u$ is big enough, then the optimal discrete control variable follows as

$$\overline{u}^* = \arg\min J(\overline{u}) = \sqrt{\frac{u^u}{S}} \tag{40}$$

Obvious, the optimal value of $\overline{u}$ depends on distance S, thus it will not be optimal at its sub-trajectory. Take half trajectory as an example, the optimal value of $\overline{u}$ is given by

$$\overline{u}^*_{half} = \sqrt{\frac{2u^u}{S}} \neq \sqrt{\frac{u^u}{S}} \tag{41}$$

Remark 4. For the state constrained cases, the corresponding Lagrange multipliers could be invited, and the corresponding properties could be obtained by comparing with the classical state constrained minimum principle.

Remark 5. $\overline{\boldsymbol{u}}$ has dual roles, which could be interpreted as a discrete control variable or a parameter of the model function. Hence there is also another routine to prove Theorem 2, which restates Problem C as a following problem.

Problem D.

$$\begin{aligned} \min J(\boldsymbol{u}(t); \overline{\boldsymbol{u}}) = & \int_{t_0}^{t_f} L(\boldsymbol{x}, \boldsymbol{u}; \overline{\boldsymbol{u}})dt + \phi(\overline{\boldsymbol{u}}) + \Phi(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)) \\ \text{s.t.} \;\; & \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t); \overline{\boldsymbol{u}}) \end{aligned}$$

$$\begin{pmatrix} \boldsymbol{x}(t_0) \\ \boldsymbol{x}(t_f) \end{pmatrix} \in E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)}$$

where the technical assumptions are satisfied as discussed before and t_0 is given (t_f can be free or given).

For simplicity, here gives a proof of Theorem 2 in a special case, which assumes that the conditions for (33) and (34) hold, $\Phi(\cdot) = 0$, $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, and $\boldsymbol{x}(t_f)$ is free.

Proof of Theorem 2. Treat $\bar{\boldsymbol{u}}$ as a parameter, and solve Problem D by Theorem 1, then the optimal solution is parameterized by $\bar{\boldsymbol{u}}$, and denote it by a upper star sign. Since Theorem 1 holds for any given parameter $\bar{\boldsymbol{u}}$, it holds for optimal parameter $\bar{\boldsymbol{u}}^*$. Therefore, the properties (a), (b), (c) and (e) of Theorem 2 hold by the properties (a), (b), (c) and (d) of Theorem 1.

Since conditions for (33) and (34) hold, and $\bar{\boldsymbol{u}}^{\boldsymbol{*}}$ gives the extremum of $J^*(\bar{\boldsymbol{u}})$, it follows that

$$\mathbf{0} = \frac{dJ^*(\bar{\boldsymbol{u}}^*)}{d\bar{\boldsymbol{u}}} = \int_{t_0}^{t_f} \left(\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \frac{\partial L^*}{\partial \boldsymbol{x}} + \left(\frac{\partial \boldsymbol{u}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \frac{\partial L^*}{\partial \boldsymbol{u}} + \frac{\partial L^*}{\partial \bar{\boldsymbol{u}}} \right) dt + \frac{\partial \phi}{\partial \bar{\boldsymbol{u}}} \tag{42}$$

By definition, it follows that

$$\frac{\partial H^*}{\partial \boldsymbol{x}} = \frac{\partial L^*}{\partial \boldsymbol{x}} + \left(\frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{x}}\right)^{\boldsymbol{T}} \boldsymbol{p}^* \tag{43}$$

$$\frac{\partial H^*}{\partial \boldsymbol{u}} = \frac{\partial L^*}{\partial \boldsymbol{u}} + \left(\frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{u}}\right)^{T} \boldsymbol{p}^* \tag{44}$$

By (42)–(44), it gives that

$$\mathbf{0} = \int_{t_0}^{t_f} \left[\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \left(\frac{\partial H^*}{\partial \boldsymbol{x}} - \left(\frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{x}}\right)^{\boldsymbol{T}} \boldsymbol{p}^* \right) + \left(\frac{\partial \boldsymbol{u}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \left(\frac{\partial H^*}{\partial \boldsymbol{u}} - \left(\frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{u}}\right)^{\boldsymbol{T}} \boldsymbol{p}^* \right) + \frac{\partial L^*}{\partial \bar{\boldsymbol{u}}} \right] dt + \frac{\partial \phi}{\partial \bar{\boldsymbol{u}}} \tag{45}$$

According to the property (a) of Theorem 1, (33) and (45), it gives that

$$\mathbf{0} = \int_{t_0}^{t_f} \left[-\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \dot{\boldsymbol{p}}^* - \left(\frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}} + \frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}^*}{\partial \bar{\boldsymbol{u}}} \right)^{\boldsymbol{T}} \boldsymbol{p}^* + \frac{\partial L^*}{\partial \bar{\boldsymbol{u}}} \right] dt + \frac{\partial \phi}{\partial \bar{\boldsymbol{u}}} \tag{46}$$

By means of integration by part, it follows that

$$\int_{t_0}^{t_f} \left(-\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \dot{\boldsymbol{p}}^* \right) dt = -\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \boldsymbol{p}^* \Bigg|_{t_0}^{t_f} + \int_{t_0}^{t_f} \left(\left(\frac{d}{dt} \left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \right) \boldsymbol{p}^* \right) dt \tag{47}$$

By the definition of the model function, it gives that

$$\frac{d}{dt}\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right) = \frac{\partial}{\partial \bar{\boldsymbol{u}}} \dot{\boldsymbol{x}}^* = \frac{d\boldsymbol{f}^*}{d\bar{\boldsymbol{u}}} = \frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}} + \frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}^*}{\partial \bar{\boldsymbol{u}}} + \frac{\partial \boldsymbol{f}^*}{\partial \bar{\boldsymbol{u}}} \tag{48}$$

Because $\boldsymbol{x}(t_0)$ is given, it follows that

$$\left.\frac{\partial \boldsymbol{x}}{\partial \bar{\boldsymbol{u}}}\right|_{t=t_0} = \mathbf{0} \tag{49}$$

By transversality conditions of Theorem 1, it follows that

$$\boldsymbol{p}^*(t_f) = \mathbf{0} \tag{50}$$

Combining (48), (49) and (50), (47) gives that

$$\int_{t_0}^{t_f} \left(-\left(\frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \dot{\boldsymbol{p}}^* \right) dt = \int_{t_0}^{t_f} \left(\frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}^*}{\partial \bar{\boldsymbol{u}}} + \frac{\partial \boldsymbol{f}^*}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}^*}{\partial \bar{\boldsymbol{u}}} + \frac{\partial \boldsymbol{f}^*}{\partial \bar{\boldsymbol{u}}} \right)^T \boldsymbol{p}^* dt \tag{51}$$

Due to (46) and (51), it follows that

$$\int_{t_0}^{t_f} \left(\frac{\partial L^*}{\partial \bar{\boldsymbol{u}}} + \left(\frac{\partial \boldsymbol{f}^*}{\partial \bar{\boldsymbol{u}}}\right)^T \boldsymbol{p}^* \right) dt + \frac{\partial \phi}{\partial \bar{\boldsymbol{u}}} = \int_{t_0}^{t_f} \frac{\partial H^*}{\partial \bar{\boldsymbol{u}}} dt + \frac{\partial \phi}{\partial \bar{\boldsymbol{u}}} = \mathbf{0} \tag{52}$$

which completes the proof of property (d) of Theorem 2.

2.3. Multistage hybrid parametric minimum principle

The optimality conditions for Problem A are given by Theorem 3.

Theorem 3 (*Multistage Hybrid Parametric Minimum Principle for Problem A*). *Let $\{t_i^*\}$, $i = 1, 2, \dots, k$, $\{\overline{\boldsymbol{u}}_i^*\} \in \overline{U}$, $i = 0, 1, \dots, k$ and $\boldsymbol{u}^*: [t_0, t_f] \to U$ be an optimal time sequence, optimal discrete and continuous control variable of Problem A and let $\boldsymbol{x}^*: [t_0, t_f] \to R^n$ be the corresponding optimal state trajectory. Then there exists a piecewise continuously differentiable vector function $\boldsymbol{p}^*: [t_0, t_f] \to R^n$, named costate, and having the following properties:*
(a) $\boldsymbol{p}^*(t)$ satisfies the adjoint equation in every stage:

$$\dot{\boldsymbol{p}}^* = -\frac{\partial}{\partial \boldsymbol{x}} H_i(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*), \forall t \in (t_i^*, t_{i+1}^*), \quad i = 0, 1, \dots, k \tag{53}$$

where Hamiltonian $H_i: R^n \times U \times \overline{U} \times R^n \to R$ is defined as

$$H_i(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}}, \boldsymbol{p}) = L_i(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}}) + \boldsymbol{p}^T(t) \cdot \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \overline{\boldsymbol{u}}) \tag{54}$$

(b) transversality condition:

$$\begin{pmatrix} \boldsymbol{p}^*(t_0) \\ -\boldsymbol{p}^*(t_f) \end{pmatrix} \perp T_{\begin{pmatrix} \boldsymbol{x}^*(t_0) \\ \boldsymbol{x}^*(t_f) \end{pmatrix}} E_{\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)} \tag{55}$$

$$\begin{pmatrix} -\boldsymbol{p}^*(t_i^{*-}) \\ \boldsymbol{p}^*(t_i^{*+}) \end{pmatrix} + \nabla \Phi_{i-1,i}(\boldsymbol{x}^*(t_i^{*-}), \boldsymbol{x}^*(t_i^{*+})) \perp T_{\begin{pmatrix} \boldsymbol{x}^*(t_i^{*-}) \\ \boldsymbol{x}^*(t_i^{*+}) \end{pmatrix}} S_{\boldsymbol{x}(t_i^-), \boldsymbol{x}(t_i^+)}, \quad i = 1, \dots, k \tag{56}$$

(c) minimality condition for continuous control variable:

$$\boldsymbol{u}^*(t) = \arg \min_{\boldsymbol{u}(t) \in U} H_i\left(\boldsymbol{x}^*, \boldsymbol{u}, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*\right) \forall t \in (t_i^*, t_{i+1}^*), \quad i = 0, 1, \dots, k \tag{57}$$

(d) minimality condition for discrete control variable:

$$\overline{\boldsymbol{u}}_i^* = \arg \min_{\overline{\boldsymbol{u}} \in \overline{U}} \left(\int_{t_i}^{t_{i+1}} H_i\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}, \boldsymbol{p}^*\right) dt + \phi_i(\overline{\boldsymbol{u}}) \right) i = 0, 1, \dots, k \tag{58}$$

(e) constancy condition ("energy preservation law"): for all $t \in [t_0, t_f]$

$$H_i\left(\boldsymbol{x}^*, \boldsymbol{u}^*, \overline{\boldsymbol{u}}^*, \boldsymbol{p}^*\right) = c \tag{59}$$

where c is a constant. If t_f is free, then $c = 0$.

To obtain optimality conditions of Problem A, the original problem is reduced into a Problem C by transforming all the state and control variables to a common fixed time interval, and establishing a correspondence between admissible trajectories of these problems.

Proof of Theorem 3. Introduce a new time $\tau \in [0, 1]$ and a series of transformations which map $[0, 1]$ into different time spans, i.e. $\eta_i: [0, 1] \to [t_i, t_{i+1}]$, $\quad i = 0, 1, \dots, k$, from the equations:

$$\frac{d\eta_i}{d\tau} = z_i(\tau) d_i, \;\; \eta_i(0) = t_i \tag{60}$$

where

$$d_i = t_{i+1} - t_i \in \Omega_d = [0, t_f - t_0] \tag{61}$$

and $z_i(\tau)$ are arbitrary measurable essentially bounded functions on $[0, 1]$. In order to maintain the invertibility of $\eta_i(\tau)$ and ensure $\eta_i(1) = t_{i+1}$, it requires that

$$z_i(\tau) \in \Im = \left\{ z(\tau) > 0 | \int_0^1 z(\tau) d\tau = 1 \right\} \tag{62}$$

Actually, $z(\tau) = 1$ is an obvious candidate.

By these transformations of t, the original piecewise function $\boldsymbol{x}(t)$, $\boldsymbol{u}(t)$ could be transformed into function series

$$\boldsymbol{y}_i(\tau) = \boldsymbol{x}(\eta_i(\tau)), \;\; \tau \in (0, 1), \;\; \boldsymbol{y}_i(0) = \boldsymbol{x}(t_i^-), \;\; \boldsymbol{y}_i(1) = \boldsymbol{x}(t_{i+1}^+), \;\; i = 0, 1, \dots, k \tag{63}$$

$$\boldsymbol{v}_i(\tau) = \boldsymbol{u}(\eta_i(\tau)), \;\; \tau \in (0, 1), \;\; \boldsymbol{v}_i(0) = \boldsymbol{u}(t_i^-), \;\; \boldsymbol{v}_i(1) = \boldsymbol{u}(t_{i+1}^+), \;\; i = 0, 1, \dots, k \tag{64}$$

The original discrete control sequence could be easily transformed into new control sequence:

$$\overline{\boldsymbol{u}}_i = \overline{\boldsymbol{u}}_i, \;\; i = 0, 1, \dots, k \tag{65}$$

Then the corresponding model functions are as follows

$$\frac{d\boldsymbol{y}_i}{d\tau} = d_i z_i(\tau)\boldsymbol{f}(\boldsymbol{y}_i(\tau), \boldsymbol{v}_i(\tau), \overline{\boldsymbol{u}}_i) \qquad i = 0, 1, \ldots, k \tag{66}$$

$$\frac{d\eta_i}{d\tau} = d_i z_i(\tau) \qquad i = 0, 1, \ldots, k \tag{67}$$

Let

$$L(\{\boldsymbol{y}_i\}, \{\boldsymbol{v}_i\}, \{\overline{\boldsymbol{u}}_i\}, \{d_i\}) = \sum_{i=0}^{k} d_i z_i(\tau) L_i(\boldsymbol{y}_i(\tau), \boldsymbol{v}_i(\tau), \overline{\boldsymbol{u}}_i) \tag{68}$$

$$\phi(\{\overline{\boldsymbol{u}}_i\}) = \sum_{i=0}^{k} \phi_i(\overline{\boldsymbol{u}}_i) \tag{69}$$

$$\Phi(\{\boldsymbol{y}_{i-1}(1)\}, \{\boldsymbol{y}_i(0)\}) = \sum_{i=1}^{k} \Phi_{i-1.i}(\boldsymbol{y}_{i-1}(1), \boldsymbol{y}_i(0)) \tag{70}$$

then the corresponding cost function is as follows

$$J(\{\boldsymbol{v}_i(\tau)\}, \left\{\overline{\boldsymbol{u}}_i\right\}, \{d_i\}) = \int_0^1 L(\{\boldsymbol{y}_i\}, \{\boldsymbol{v}_i\}, \{\overline{\boldsymbol{u}}_i\}, \{d_i\}) d\tau + \phi(\{\overline{\boldsymbol{u}}_i\}) + \Phi(\{\boldsymbol{y}_{i-1}(1)\}, \{\boldsymbol{y}_i(0)\}) \tag{71}$$

Suppose that the original endpoint constraint surface $E_{\boldsymbol{x}(t_0),\boldsymbol{x}(t_f)}$ is described by

$$\boldsymbol{h}_0(\boldsymbol{x}(t_0), \boldsymbol{x}(t_f)) = 0 \tag{72}$$

then the corresponding endpoint constraint surface $E_{\boldsymbol{y}_0(0),\boldsymbol{y}_k(1)}$ is described by

$$\boldsymbol{h}_0(\boldsymbol{y}_0(0), \boldsymbol{y}_k(1)) = 0 \tag{73}$$

Suppose that the switching constraint surface $S_{\boldsymbol{x}(t_i^{*-}),\boldsymbol{x}(t_i^{*+})}$ is described by

$$\boldsymbol{h}_i(\boldsymbol{x}(t_i^-), \boldsymbol{x}(t_i^+)) = 0, i = 1, \ldots, k \tag{74}$$

then the corresponding endpoint constraint surface $S_{\boldsymbol{y}_{i-1}(1),\boldsymbol{y}_i(0)}$ is described by

$$\boldsymbol{h}_i(\boldsymbol{y}_{i-1}(1), \boldsymbol{y}_i(0)) = 0, i = 1, \ldots, k \tag{75}$$

The extra constraints on $\eta_i(\tau)$ include endpoint constraints, which are given by

$$\eta_0(0) = t_0 \tag{76}$$

$$\eta_k(1) = t_f \tag{77}$$

and continuous constraints, which are given by

$$\eta_{i-1}(1) = \eta_i(0) = t_i, \ i = 1, \ldots, k \tag{78}$$

Hence, the extra endpoint constraint surfaces include $E_{\eta_0(0),\eta_k(1)}$, which is given by

$$l_0(\eta_0(0)) = \eta_0(0) - t_0 = 0 \tag{79}$$

$$l_{k+1}(\eta_k(1)) = \eta_k(1) - t_f = 0 \tag{80}$$

and $S_{\eta_{i-1}(1),\eta_i(0)}$, which is described by

$$l_i(\eta_{i-1}(1), \eta_i(0)) = \eta_{i-1}(1) - \eta_i(0) = 0, i = 1, \ldots, k \tag{81}$$

In order to assemble a bigger optimal control problem of type C, a bigger state variable is needed, which has the following form

$$\boldsymbol{X}(\tau) = \begin{pmatrix}\boldsymbol{Y}(\tau)^T & \boldsymbol{\Pi}(\tau)^T\end{pmatrix}^T \in R^{(n+1)(k+1)} = R^{n(k+1)} \times R^{k+1} \tag{82}$$

where

$$\boldsymbol{Y}(\tau) = (\boldsymbol{y}_0^T(\tau) \quad \cdots \quad \boldsymbol{y}_k^T(\tau))^T \in R^{n(k+1)} = \underbrace{R^n \times \cdots \times R^n}_{k+1} \tag{83}$$

$$\boldsymbol{\Pi}(\tau) = (\eta_0(\tau) \quad \cdots \quad \eta_k(\tau))^T \in R^{k+1} = \underbrace{R \times \cdots \times R}_{k+1} \tag{84}$$

Otherwise, bigger continuous and discrete control variables are also needed, which are given by

$$\boldsymbol{V}(\tau) = \begin{pmatrix} \boldsymbol{v}_0^T(\tau) & \cdots & \boldsymbol{v}_k^T(\tau) \end{pmatrix}^T \in \Omega_V = \underbrace{U \times \cdots \times U}_{k+1} \subset R^{m(k+1)} \tag{85}$$

$$\overline{\boldsymbol{V}} = \begin{pmatrix} \boldsymbol{\Lambda}^T & \boldsymbol{D}^T \end{pmatrix}^T \in \Omega_{\overline{V}} = \Omega_\Lambda \times \Omega_\Lambda \subset R^{(l+1)(k+1)} \tag{86}$$

where

$$\boldsymbol{\Lambda} = \begin{pmatrix} \overline{\boldsymbol{u}}_0^T & \cdots & \overline{\boldsymbol{u}}_k^T \end{pmatrix}^T \in \Omega_\Lambda = \underbrace{\overline{U} \times \cdots \times \overline{U}}_{k+1} \subset R^{l(k+1)} \tag{87}$$

$$\boldsymbol{D} = \begin{pmatrix} d_0 & \cdots & d_k \end{pmatrix}^T \in \Omega_D = \underbrace{\Omega_d \times \cdots \times \Omega_d}_{k+1} \subset R^{k+1} \tag{88}$$

By (66) and (67) as well as new notations, the corresponding model function is descried by

$$\dot{\boldsymbol{X}} = \boldsymbol{F}(\boldsymbol{X}, \boldsymbol{V}, \overline{\boldsymbol{V}}) = \begin{pmatrix} d_0 z_0(\tau)\boldsymbol{f}(\boldsymbol{y}_0(\tau), \boldsymbol{v}_0(\tau), \overline{\boldsymbol{u}}_0) \\ \vdots \\ d_k z_k(\tau)\boldsymbol{f}(\boldsymbol{y}_k(\tau), \boldsymbol{v}_k(\tau), \overline{\boldsymbol{u}}_k) \\ d_0 z_0(\tau) \\ \vdots \\ d_k z_k(\tau) \end{pmatrix} \tag{89}$$

By (71) and new notations, the corresponding cost function is given by

$$J(\boldsymbol{V}(\tau), \overline{\boldsymbol{V}}) = \int_0^1 L(\boldsymbol{X}, \boldsymbol{V}, \overline{\boldsymbol{V}})d\tau + \phi(\overline{\boldsymbol{V}}) + \Phi(\boldsymbol{X}(0), \boldsymbol{X}(1)) \tag{90}$$

Let

$$E_{\boldsymbol{Y}(0),\boldsymbol{Y}(1)} = E_{\boldsymbol{y}_0(0),\boldsymbol{y}_k(1)} \times S_{\boldsymbol{y}_0(1),\boldsymbol{y}_1(0)} \times \cdots \times S_{\boldsymbol{y}_{k-1}(1),\boldsymbol{y}_k(0)} \tag{91}$$

$$E_{\boldsymbol{\Pi}(0),\boldsymbol{\Pi}(1)} = E_{\eta_0(0),\eta_k(1)} \times S_{\eta_0(1),\eta_1(0)} \times \cdots \times S_{\eta_{k-1}(1),\eta_k(0)} \tag{92}$$

then the corresponding endpoint constraint surface is descried by

$$E_{\boldsymbol{X}(0),\boldsymbol{X}(1)} = E_{\boldsymbol{Y}(0),\boldsymbol{Y}(1)} \times E_{\boldsymbol{\Pi}(0),\boldsymbol{\Pi}(1)} \tag{93}$$

Hence the transformed Problem A has the following form:

Problem A′.

$$\begin{aligned} \min J(\boldsymbol{V}(\tau), \overline{\boldsymbol{V}}) = & \int_0^1 L(\boldsymbol{X}, \boldsymbol{V}, \overline{\boldsymbol{V}})d\tau + \phi(\overline{\boldsymbol{V}}) + \Phi(\boldsymbol{X}(0), \boldsymbol{X}(1)) \\ \text{s.t. } & \dot{\boldsymbol{X}} = \boldsymbol{F}(\boldsymbol{X}, \boldsymbol{V}, \overline{\boldsymbol{V}}) \\ & \begin{pmatrix} \boldsymbol{X}(0) \\ \boldsymbol{X}(1) \end{pmatrix} \in E_{\boldsymbol{X}(0),\boldsymbol{X}(1)} \end{aligned}$$

Obviously, Problem A′ is a nominal form of Problem C, then Theorem 2 holds. According to the construction of Problem A′, which could be called mapping H and is mainly described by (60), (61) and (63)–(65), it is easily to see that any admissible trajectory of Problem A will be an admissible trajectory of Problem A′. Since the transformation series $\eta_i(\tau)$, $i = 0, 1, \ldots, k$, are invertible, any admissible trajectory of Problem A′ will be an admissible trajectory of Problem A by following mapping G:

$$t = \eta_i(\tau), \forall t \in (t_i, t_{i+1}), \quad i = 0, 1, \ldots, k \tag{94}$$

$$t_0 = \eta_0(0),\ t_f = t_{k+1} = \eta_k(1),\ t_i = \eta_i(0),\ i = 1, \ldots, k \tag{95}$$

$$\overline{\boldsymbol{u}}_i = \overline{\boldsymbol{u}}_i,\ i = 0, 1, \ldots, k \tag{96}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{y}_0(0),\ \boldsymbol{x}(t_f) = \boldsymbol{y}_k(1),\ \boldsymbol{x}(t) = \boldsymbol{y}_i(\eta_i^{-1}(t)),\ \forall t \in (t_i, t_{i+1}), \quad i = 0, 1, \ldots, k \tag{97}$$

$$\boldsymbol{u}(t_0) = \boldsymbol{v}_0(0),\ \boldsymbol{u}(t_f) = \boldsymbol{v}_k(1),\ \boldsymbol{u}(t) = \boldsymbol{v}_i(\eta_i^{-1}(t)),\ \forall t \in (t_i, t_{i+1}), \quad i = 0, 1, \ldots, k \tag{98}$$

the value $\boldsymbol{x}(t_i)$ is taken to be equal to $\boldsymbol{y}_{i-1}(1)$ or $\boldsymbol{y}_i(0)$, and $\boldsymbol{u}(t_i)$ is taken to be equal to $\boldsymbol{v}_{i-1}(1)$ or $\boldsymbol{v}_i(0)$, $i = 1, 2, \ldots, k$, depending on the desired convention.

If this convention is identical with the original convention, which is used to determine $\boldsymbol{x}(t_i)$ of Problem A, then the mapping F and G are inverse to each other. Moreover, these mappings preserve the value of the cost functional, no matter what convention is. Hence the optimal solution of Problem A could be derived from the optimal solution of Problem A′, so are the optimality conditions.

For Problem A′, the Hamiltonian

$$H(\boldsymbol{X},\boldsymbol{V},\overline{\boldsymbol{V}},\boldsymbol{P})=L(\boldsymbol{X},\boldsymbol{V},\overline{\boldsymbol{V}})+\boldsymbol{P}^T\cdot\boldsymbol{F}(\boldsymbol{X},\boldsymbol{V},\overline{\boldsymbol{V}}) \tag{99}$$

where costate, composed of multi-components, is given by

$$\boldsymbol{P}(\tau)=\begin{pmatrix}\boldsymbol{p}_{\boldsymbol{y}_0}^T(\tau) & \cdots & \boldsymbol{p}_{\boldsymbol{y}_k}^T(\tau) & p_{\eta_0}(\tau) & \cdots & p_{\eta_k}(\tau)\end{pmatrix}^T \tag{100}$$

For $i=0,1,\ldots,k$, let

$$\hat{H}_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i,\boldsymbol{p}_{\boldsymbol{y}_i},p_{\eta_i})=L_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i)+\boldsymbol{p}_{\boldsymbol{y}_i}^T(\tau)\cdot f(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i)+p_{\eta_i}(\tau) \tag{101}$$

$$H_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i,\boldsymbol{p}_{\boldsymbol{y}_i})=L_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i)+\boldsymbol{p}_{\boldsymbol{y}_i}^T(\tau)\cdot f(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i) \tag{102}$$

then it follows that

$$\hat{H}_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i,\boldsymbol{p}_{\boldsymbol{y}_i},p_{\eta_i})=H_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i,\boldsymbol{p}_{\boldsymbol{y}_i})+p_{\eta_i}(\tau) \tag{103}$$

By (68), (89) and (101), the (99) gives that

$$H(\boldsymbol{X},\boldsymbol{V},\overline{\boldsymbol{V}},\boldsymbol{P})=\sum_{i=0}^{k}d_iz_i(\tau)\hat{H}_i(\boldsymbol{y}_i,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i,\boldsymbol{p}_{\boldsymbol{y}_i},p_{\eta_i}) \tag{104}$$

According to the property (a) of Theorem 2, the costate $\boldsymbol{P}(\tau)$ satisfies

$$\dot{\boldsymbol{P}}^*=-\frac{\partial}{\partial\boldsymbol{X}}H(\boldsymbol{X}^*,\boldsymbol{V}^*,\overline{\boldsymbol{V}}^*,\boldsymbol{P}^*) \tag{105}$$

which is formed by the following components

$$\dot{\boldsymbol{p}}_{\boldsymbol{y}_i}^*=-d_i^*z_i(\tau)\frac{\partial}{\partial\boldsymbol{y}_i}H_i(\boldsymbol{y}_i^*,\boldsymbol{v}_i^*,\overline{\boldsymbol{u}}_i^*,\boldsymbol{p}_{\boldsymbol{y}_i}^*)\qquad i=0,1,\ldots,k \tag{106}$$

$$\dot{p}_{\eta_i}^*=-\frac{\partial}{\partial\eta_i}H(\boldsymbol{X}^*,\boldsymbol{V}^*,\overline{\boldsymbol{V}}^*,\boldsymbol{P}^*)=0\qquad i=0,1,\ldots,k \tag{107}$$

According to the property (c) of Theorem 2, for all $\tau\in[0,1]$

$$\boldsymbol{V}^*(\tau)=\arg\min_{\boldsymbol{V}(\tau)\in\Omega_V}H(\boldsymbol{X}^*,\boldsymbol{V},\overline{\boldsymbol{V}}^*,\boldsymbol{P}^*) \tag{108}$$

which gives the following equations

$$\boldsymbol{v}_i^*(\tau)=\arg\min_{\boldsymbol{v}_i(\tau)\in U}H_i(\boldsymbol{y}_i^*,\boldsymbol{v}_i,\overline{\boldsymbol{u}}_i^*,\boldsymbol{p}_{\boldsymbol{y}_i}^*)\quad i=0,1,\ldots,k \tag{109}$$

According to the property (d) of Theorem 2,

$$\overline{\boldsymbol{V}}^*=\arg\min_{\overline{\boldsymbol{V}}\in\Omega_{\overline{V}}}\left(\int_0^1H(\boldsymbol{X}^*,\boldsymbol{V}^*,\overline{\boldsymbol{V}},\boldsymbol{P}^*)d\tau+\phi(\overline{\boldsymbol{V}})\right) \tag{110}$$

which gives the following equations

$$\overline{\boldsymbol{u}}_i^*=\arg\min_{\overline{\boldsymbol{u}}_i\in\overline{U}}\left(\int_0^1H_i(\boldsymbol{y}_i^*,\boldsymbol{v}_i^*,\overline{\boldsymbol{u}}_i,\boldsymbol{p}_{\boldsymbol{y}_i}^*)d\tau+\phi_i(\overline{\boldsymbol{u}}_i)\right)\qquad i=0,1,\ldots,k \tag{111}$$

$$d_i^*=\arg\min_{d_i\in\Omega_d}\left(\int_0^1d_iz_i(\tau)\hat{H}_i(\boldsymbol{y}_i^*,\boldsymbol{v}_i^*,\overline{\boldsymbol{u}}_i^*,\boldsymbol{p}_{\boldsymbol{y}_i}^*,p_{\eta_i}^*)d\tau\right)\qquad i=0,1,\ldots,k \tag{112}$$

According to the property (e) of Theorem 2,

$$H(\boldsymbol{X}^*,\boldsymbol{V}^*,\overline{\boldsymbol{V}}^*,\boldsymbol{P}^*)=\sum_{i=0}^{k}d_i^*z_i(\tau)\hat{H}_i(\boldsymbol{y}_i^*,\boldsymbol{v}_i^*,\overline{\boldsymbol{u}}_i^*,\boldsymbol{p}_{\boldsymbol{y}_i}^*,p_{\eta_i}^*)=\text{c} \tag{113}$$

where c is a constant. Since $z_i(\tau)$ is an arbitrary function in $\Im$, (113) follows that

$$\hat{H}_i(\boldsymbol{y}_i^*,\boldsymbol{v}_i^*,\overline{\boldsymbol{u}}_i^*,\boldsymbol{p}_{\boldsymbol{y}_i}^*,p_{\eta_i}^*)=0\text{ or }\frac{c_i}{z_i(\tau)}\text{ or }\frac{c'}{\sum_{i=0}^{k}d_i^*z_i(\tau)},\quad i=0,1,\ldots,k \tag{114}$$

where c_i, c' are nonzero constants. On the other hand, if d_i^* gives an extremum at its boundary point, it follows that if $d_i^*=0$ for some $i\in\{0,1,\ldots,k\}$, then these stages are meaningless and discardable; if $d_j^*=t_f-t_0$ for $j\in\{0,1,\ldots,k\}$, then it follows that

$$d_i^* = \begin{cases} t_f - t_0, & i = j \\ 0, & i \neq j \end{cases}, \; i \in \{0, 1, \ldots, k\} \tag{115}$$

which means Problem A degenerates into a single stage hybrid parametric optimal control problem, and it is inconsistent to the original setting. Hence d_i^* must give an extremum at its interior point. Furthermore, $H(\boldsymbol{X}, \boldsymbol{V}, \overline{\boldsymbol{V}}, \boldsymbol{P})$ is differentiable with respect to d_i, then by (112), it follows that, for $i = 0, 1, \ldots, k$

$$\int_0^1 \frac{\partial}{\partial d_i} H(\boldsymbol{X}^*, \boldsymbol{V}^*, \overline{\boldsymbol{V}}^*, \boldsymbol{P}^*) d\tau = \int_0^1 z_i(\tau) \hat{H}_i(\boldsymbol{y}_i^*, \boldsymbol{v}_i^*, \overline{\boldsymbol{u}}_i^*, \boldsymbol{p}_{\boldsymbol{y}_i}^*, p_{\eta_i}^*) d\tau = 0 \tag{116}$$

If $\hat{H}_i(\boldsymbol{y}_i^*, \boldsymbol{v}_i^*, \overline{\boldsymbol{u}}_i^*, \boldsymbol{p}_{\boldsymbol{y}_i}^*, p_{\eta_i}^*) = \frac{c_i}{z_i(\tau)}$, then $c_i = 0$ by $z_i(\tau) > 0$ and (116), which is inconsistent with hypothesis; if $\hat{H}_i(\boldsymbol{y}_i^*, \boldsymbol{v}_i^*, \overline{\boldsymbol{u}}_i^*, \boldsymbol{p}_{\boldsymbol{y}_i}^*, p_{\eta_i}^*) = \frac{c'}{\sum_{i=0}^k d_i^* z_i(\tau)}$, then $c' = 0$ by $d_i^* > 0$, $z_i(\tau) > 0$ and (116), which is also inconsistent with hypothesis; eventually, it gives that

$$\hat{H}_i(\boldsymbol{y}_i^*, \boldsymbol{v}_i^*, \overline{\boldsymbol{u}}_i^*, \boldsymbol{p}_{\boldsymbol{y}_i}^*, p_{\eta_i}^*) = 0 \tag{117}$$

According to the property (b) of Theorem 2,

$$\begin{pmatrix} \boldsymbol{P}^*(0) \\ -\boldsymbol{P}^*(1) \end{pmatrix} + \nabla \Phi(\boldsymbol{X}^*(0), \boldsymbol{X}^*(1)) \perp T_{\begin{pmatrix} \boldsymbol{X}^*(0) \\ \boldsymbol{X}^*(1) \end{pmatrix}} E_{\boldsymbol{X}_0(0), \boldsymbol{X}_k(1)} \tag{118}$$

which is formed by the following components

$$\begin{pmatrix} \boldsymbol{p}_{\boldsymbol{y}_0}^*(0) \\ -\boldsymbol{p}_{\boldsymbol{y}_k}^*(1) \end{pmatrix} \perp T_{\begin{pmatrix} \boldsymbol{y}_0^*(0) \\ \boldsymbol{y}_k^*(1) \end{pmatrix}} E_{\boldsymbol{y}_0(0), \boldsymbol{y}_k(1)} \tag{119}$$

$$\begin{pmatrix} -\boldsymbol{p}_{\boldsymbol{y}_{i-1}}^*(1) \\ \boldsymbol{p}_{\boldsymbol{y}_i}^*(0) \end{pmatrix} + \nabla \Phi_{i-1.i}(\boldsymbol{y}_{i-1}^*(1), \boldsymbol{y}_i^*(0)) \perp T_{\begin{pmatrix} \boldsymbol{y}_{i-1}^*(1) \\ \boldsymbol{y}_i^*(0) \end{pmatrix}} S_{\boldsymbol{y}_{i-1}(1), \boldsymbol{y}_i(0)}, \quad i = 1, \ldots, k \tag{120}$$

$$\begin{pmatrix} p_{\eta_0}^*(0) \\ -p_{\eta_k}^*(1) \end{pmatrix} \perp T_{\begin{pmatrix} \eta_0^*(0) \\ \eta_k^*(1) \end{pmatrix}} E_{\eta_0(0), \eta_k(1)} \tag{121}$$

$$\begin{pmatrix} -p_{\eta_{i-1}}^*(1) \\ p_{\eta_i}^*(0) \end{pmatrix} \perp T_{\begin{pmatrix} \eta_{i-1}^*(1) \\ \eta_i^*(0) \end{pmatrix}} S_{\eta_{i-1}(1), \eta_i(0)}, \quad i = 1, \ldots, k \tag{122}$$

By (79), (80) and (121), it follows that

$$p_{\eta 0}^*(0) = b_0 \frac{\partial l_0}{\partial \eta_0(0)} = b_0 \tag{123}$$

$$p_{\eta_k}^*(1) = -b_{k+1} \frac{\partial l_{k+1}}{\partial \eta_k(1)} = -b_{k+1} \tag{124}$$

By (81) and (122), it follows that

$$p_{\eta_i}^*(0) = b_i \frac{\partial l_i}{\partial \eta_i(0)} = -b_i, \, i = 1, 2, \ldots, k \tag{125}$$

$$p_{\eta_i}^*(1) = -b_{i+1} \frac{\partial l_{i+1}}{\partial \eta_i(1)} = -b_{i+1}, \quad i = 0, 1, \ldots, k-1 \tag{126}$$

where b_i, $i = 0, 1, \ldots, k+1$, are arbitrary constants. By (125) and (126), it follows that

$$p_{\eta_{i-1}}^*(1) = p_{\eta_i}^*(0), \; i = 1, 2, \ldots, k \tag{127}$$

According to (107), $p_{\eta_i}^*(\tau)$ is a constant for $i = 0, 1, \ldots, k$. Considering (123), (124) and (126), it follows that

$$p_{\eta_i}^*(\tau) = b_0, \;\; \tau \in [0, 1], \;\; i = 0, 1, \ldots, k \tag{128}$$

For free-time problems, there is not an endpoint constraint for $\eta_k(1)$, which gives that

$$p_{\eta_k}^*(1) = 0 \tag{129}$$

then

$$b_0 = 0 \tag{130}$$

By (103), (117) and (128)

$$H_i(\boldsymbol{y}_i^*, \boldsymbol{v}_i^*, \overline{\boldsymbol{u}}_i^*, \boldsymbol{p}_{\boldsymbol{y}_i}^*) = -b_0 \tag{131}$$

By far all the counterpart properties for Theorem 3 have been obtained, then the proof of Theorem 3 could be done by mapping G and extra transformation on costate, which is similar to the transformation of state. The forward transformation of costate is given by

$$\boldsymbol{p}_{\boldsymbol{y}_i}(\tau) = \boldsymbol{p}(\eta_i(\tau)),\ \tau \in (0, 1),\ \boldsymbol{p}_{\boldsymbol{y}_i}(0) = \boldsymbol{p}(t_i^-),\ \boldsymbol{p}_{\boldsymbol{y}_i}(1) = \boldsymbol{p}(t_{i+1}^+),\ i = 0, 1, \ldots, k \tag{132}$$

$$p_{\eta_i}(\tau) = -c,\ i = 0, 1, \ldots, k \tag{133}$$

where c is given by (59). The backward transformation of costate is given by

$$\boldsymbol{p}(t_0) = \boldsymbol{p}_{\boldsymbol{y}_0}(0),\ \boldsymbol{p}(t_f) = \boldsymbol{p}_{\boldsymbol{y}_k}(1),\ \boldsymbol{p}(t) = \boldsymbol{p}_{\boldsymbol{y}_i}(\eta_i^{-1}(t)),\ \forall t \in (t_i, t_{i+1}),\quad i = 0, 1, \ldots, k \tag{134}$$

the value of $\boldsymbol{p}(t_i)$ is taken to be equal to $\boldsymbol{p}_{\boldsymbol{y}_{i-1}}(1)$ or $\boldsymbol{p}_{\boldsymbol{y}_i}(0)$ by the same convention of state.

The property (a) of Theorem 3 could be obtained by (106); the property (b) of Theorem 3 could be obtained by (119) and (120); the property (c) of Theorem 3 could be obtained by (109); the property (d) of Theorem 3 could be obtained by (111); the condition (e) of Theorem 3 could be obtained by (130) and (131).

Remark 1. For some hybrid optimal control problems, the number of switch times k could be deduced by the realistic backgrounds, such as Example 2 of Section 3. In these cases, the systems experience certain number of switching times to reach a destination. In some other cases, the number of switching times is not known beforehand, especially when the switching event could happen anytime, anywhere in some special cases. Actually, this is the singular case, which state trivially satisfies switching surface. For example, the singular switching surface could be given by

$$S_{\boldsymbol{x}(t_i^-),\boldsymbol{x}(t_i^+)} = \{(\boldsymbol{x}(t_i^-), \boldsymbol{x}(t_i^+))^{\mathrm{T}} | \boldsymbol{x}(t_i^-) - \boldsymbol{x}(t_i^+) = 0\} \tag{135}$$

Since there is not jump at switching point, it seems that the switching surface is “missing”, which makes the switching cost $\Phi_{i-1,i}(\cdot)$ missing, and the $L_i(\cdot)$ and $\phi_i(\cdot)$ identical for all i.

Remark 2. The hybrid minimum principle where set Q has finite cardinality only provides necessary conditions for a trajectory $\boldsymbol{x}^*(\cdot)$ of the hybrid control system corresponding to a control $\boldsymbol{u}^*(\cdot)$, a time sequence $\{t_i\}$, and a model sequence $\{q_i\}$ to be locally optimal over trajectories $\boldsymbol{x}(\cdot)$ with the same model sequence $\{q_i\}$. Since the topology of hybrid control system with finite index set is separate, for a perturbated $\boldsymbol{x}(\cdot)$ to be close to $\boldsymbol{x}^*(\cdot)$ on each subinterval (t_i, t_{i+1}), the model sequence $\{q_i\}$ must remain unchanged. As the Pontryagin maximum principle, given the number of switching times k, the hybrid parametric minimum principle proposed in this paper gives globally necessary optimality conditions. That is because the perturbation of model sequence $\{\boldsymbol{f}_q(\cdot)\}$ with the same number of switching times can be compared by methods of analysis due to the continuity of parameters.

Remark 3. In classical mechanics the Newtonian approach and the least action principle are completely equivalent. Furthermore, by comparing the least action principle with the minimum principle, it indicates that there should be a hybrid least action principle in physics, and it is something that is never heard of. The reason is that the hybrid setting needs extra global messages, such as the global model sequence $\{q_i\}$ in finite cardinality case and global control variable constraints, i.e. constant constraints, in hybrid parametric control case. Hence, the least action principle is the combination of two locally based principles, which are variational principle and optimality principle. In details, at any given point the original extremal problem is transformed into differential equations by the variational principle, and those equations are locally “solved” based on the local information of a particle. Noted that this information must include all parameters it needs, thus there is not hybrid parametric concept here; since all the local infinitesimal trajectory the particle taking is locally optimal, the whole trajectory will be globally optimal by the optimality principle.

Remark 4. As most practical problems are too complex to allow for analytical solutions, the numerical algorithms are inevitable for solving optimal control problems. There are two practical methods, which are adaptive dynamic programming and direct methods. Adaptive dynamic programming based on the dynamic programming tries to solve the optimal control problem at any state of the system, which naturally is of the close-loop form, but this method will not work for hybrid parametric control problem by Remark 3 of Section 2.2. On the other hand, the direct methods, i.e. the sequential method (control vector parameterization (CVP)) and simultaneous method, parameterize control variables and transform the original problem into a nonlinear programming problem (NLP). Since the direct methods render continuous and discrete control variables equal in the transformed NLP, it can be directly used to solve hybrid parametric optimal problem when the number of switching times k is prescribed ($k = 1$ for single stage hybrid parametric problem). The multiphase algorithm [42,43] is needed when the number of switching times cannot be prescribed. The algorithm could be described as follows:

Step 0: predefine a number of switching times k, then apply a direct method to calculate continuous and discrete control variables;

Step 1: update the number of switching times k, and apply a direct method to calculate continuous and discrete control variables;

Step 2: evaluate the solution of Step 1. If the convergence is realized, stop the algorithm, otherwise go to Step 1.

3. Analytical examples

Two analytical examples are given here to corroborates the validity of the proposed principles, and one realistic example of hybrid parametric optimal control problem could be found in our previous work (fluid catalytic cracking unit with CO promoter [39]).

Example 1. Consider the following single stage hybrid parametric optimal problem:

$$\min J(u(t), \overline{u}) = \int_0^1 [x^2(t) + u^2(t)]dt \tag{136}$$

$$\text{s.t.} \;\; \dot{x} = x(t) - u(t) + \overline{u} \tag{137}$$

$$u(t) \in R, \;\; \overline{u} \in \overline{U} = [0.5, 1] \tag{138}$$

$$E_{x(0),x(1)} = \{(x(0), x(1)) | x(0) = 0\} \tag{139}$$

where the technical assumptions are satisfied as Problem C.

Solution. The Hamiltonian is formed as

$$H(x, u, \overline{u}, p) = x^2 + u^2 + p(x - u + \overline{u}) \tag{140}$$

By property (a) of Theorem 2, it follows that

$$\dot{p}^* = -p^* - 2x^* \tag{141}$$

By property (b) of Theorem 2, it follows that

$$p^*(1) = 0 \tag{142}$$

By property (c) of Theorem 2, it follows that

$$\frac{\partial H}{\partial u} = 2u^* - p^* = 0 \tag{143}$$

Differentiate (141) with the substitution of (137) and (143), it gives that

$$\ddot{p}^* - 2p^* = -2\overline{u}^* \tag{144}$$

By (139) and (141), it follows that

$$\dot{p}(0) + p(0) = 0 \tag{145}$$

Solving (144) with boundary condition (142) and (145), it gives that

$$p^*(t) = \overline{u}(-0.29e^{\sqrt{2}t} + 0.73e^{-\sqrt{2}t} + 1) \tag{146}$$

By property (d) of Theorem 2, it follows that

$$\overline{u}^* = \arg\min_{\overline{u}\in\overline{U}} \int_0^1 p^*(t)\overline{u}dt = \arg\min_{\overline{u}\in\overline{U}} 0.74\overline{u}^2 = 0.5 \tag{147}$$

With the substitution of (146) and (147), (143) gives that

$$u^*(t) = -0.0725e^{\sqrt{2}t} + 0.1825e^{-\sqrt{2}t} + 0.25 \tag{148}$$

Example 2. Consider the following multistage hybrid parametric optimal problem, which is about a car traveling through a designated point, and then returning to the original point. The control on the car is the acceleration, which is the product of continuous and discrete control. The continuous control can be regarded as the control of accelerator and brake, and discrete control can be regarded as some sort of accelerating mechanism, which is an expensive resource. Moreover, arriving a designated point allows user to reassign a new value of discrete control, and triggers a variation on the cost function. The objective is to minimize the expense of both continuous and discrete resource. The mathematics form is given as following:

$$\min J(u(t), \overline{u}, t_1) = \int_0^{t_1} \frac{1}{2}u^2(t)dt + 48\overline{u}_1 + \int_{t_1}^{2} \frac{1}{2}u^2(t)dt + 49\overline{u}_2 \tag{149}$$

$$\text{s.t.}\ \dot{x}_1 = x_2(t) \tag{150}$$

$$\dot{x}_2 = \bar{u}u(t) \tag{151}$$

$$u(t) \in R, \quad \bar{u} \in R$$

$$E_{x(0),x(2)} = \left\{ \begin{pmatrix} x_1(0) \\ x_2(0) \\ x_1(2) \\ x_2(2) \end{pmatrix} \middle| \begin{matrix} E_1\colon x_1(0) = 0 \\ E_2\colon x_2(0) = 0 \\ E_3\colon x_1(2) = 0 \\ E_4\colon x_2(2) = 0 \end{matrix} \right\} \tag{152}$$

$$S_{x(t_1^-),x(t_1^+)} = \left\{ \begin{pmatrix} x_1(t_1^-) \\ x_2(t_1^-) \\ x_1(t_1^+) \\ x_2(t_1^+) \end{pmatrix} \middle| \begin{matrix} S_1\colon x_1(t_1^-) - x_1(t_1^+) = 0 \\ S_2\colon x_2(t_1^-) - x_2(t_1^+) = 0 \\ S_3\colon x_1(t_1^-) - 2 = 0 \end{matrix} \right\} \tag{153}$$

where the technical assumptions are satisfied as Problem A.

Solution. Two stages' Hamiltonian have the same form, which is as follows

$$H(x_2, u, \bar{u}, p_1, p_2) = \frac{1}{2}u^2 + p_1 x_2 + p_2 \bar{u} u \tag{154}$$

The convention of values at discontinuity points is taken as right-sided limit. By property (a) of Theorem 3, it follows that

$$\dot{p}_1^* = -\frac{\partial H}{\partial x_1} = 0,\ \forall t \in [0, t_1^*) \cup [t_1^*, 2] \tag{155}$$

$$\dot{p}_2^* = -\frac{\partial H}{\partial x_2} = -p_1^*(t),\ \forall t \in [0, t_1^*) \cup [t_1^*, 2] \tag{156}$$

By property (b) of Theorem 3, it follows that

$$p_1^*(2) = \lambda_1 \tag{157}$$

$$p_2^*(2) = \lambda_2 \tag{158}$$

$$p_1^*(0) = \lambda_4 \tag{159}$$

$$p_2^*(0) = \lambda_5 \tag{160}$$

$$p_1^*(t_1^{*-}) = p_1^*(t_1^{*+}) + \lambda_3 \tag{161}$$

$$p_2^*(t_1^{*+}) = p_2^*(t_-^{*+}) \tag{162}$$

where λ_i, $i = 1, 2, 3, 4, 5$, are arbitrary constants. Solving (155) with boundary condition (157) and (161), it gives that

$$p_1^*(t) = \begin{cases} \lambda_1 + \lambda_3, & t \in [0, t_1^*) \\ \lambda_1, & t \in [t_1^*, 2] \end{cases} \tag{163}$$

Solving (156) with the substitution of (163) and boundary condition (158) and (162), it gives that

$$p_2^*(t) = \begin{cases} -(\lambda_1 + \lambda_3)t + \lambda_3 t_1^* + 2\lambda_1 + \lambda_2, & t \in [0, t_1^*) \\ -\lambda_1(t-2) + \lambda_2, & t \in [t_1^*, 2] \end{cases} \tag{164}$$

By property (c) of Theorem 3, it follows that

$$\frac{\partial H}{\partial u} = u^*(t) + p_2^* \bar{u}^*(t) = 0 \tag{165}$$

By (164) and (165), it gives that

$$u^*(t) = \begin{cases} \bar{u}_1^* \left((\lambda_1 + \lambda_3)t - (\lambda_3 t_1^* + 2\lambda_1 + \lambda_2) \right), & t \in [0, t_1^*) \\ \bar{u}_2^* \left(\lambda_1(t-2) - \lambda_2 \right), & t \in [t_1^*, 2] \end{cases} \tag{166}$$

Solving (151) with the substitution of (166) and boundary condition E_2 of (152) and S_2 of (153), it gives that

$$x_2^*(t) = \begin{cases} \bar{u}_1^{*2} \left(\frac{1}{2}(\lambda_1 + \lambda_3)t^2 - (\lambda_3 t_1^* + 2\lambda_1 + \lambda_2)t \right), & t \in [0, t_1) \\ \bar{u}_2^{*2} \left(\frac{1}{2}\lambda_1 (t-2)^2 - \lambda_2 t \right) + c_1, & t \in [t_1, 2] \end{cases} \tag{167}$$

where the constant c_1 is given by

$$c_1 = \bar{u}_1^{*2} \left(\frac{1}{2}(\lambda_1 + \lambda_3)t_1^{*2} - (\lambda_3 t_1^* + 2\lambda_1 + \lambda_2)t_1^* \right) - \bar{u}_2^{*2} \left(\frac{1}{2}\lambda_1 (t_1^* - 2)^2 - \lambda_2 t_1^* \right) \tag{168}$$

Solving (150) with the substitution of (167) and boundary condition E_1 of (152) and S_1 of (153), it gives that

$$x_1^*(t) = \begin{cases} \overline{u}_1^{*2}\left(\frac{1}{6}(\lambda_1+\lambda_3)t^3 - \frac{1}{2}(\lambda_3 t_1^* + 2\lambda_1 + \lambda_2)t^2\right), & t \in [0, t_1^*) \\ \overline{u}_2^{*2}\left(\frac{1}{6}\lambda_1(t-2)^3 - \frac{1}{2}\lambda_2 t^2\right) + c_1 t + c_2, & t \in [t_1^*, 2] \end{cases} \tag{169}$$

Where the constant c_2 is given by

$$c_2 = \overline{u}_1^{*2}\left(\frac{1}{6}(\lambda_1+\lambda_3)t_1^{*3} - \frac{1}{2}(\lambda_3 t_1^* + 2\lambda_1 + \lambda_2)t_1^{*2}\right) - \overline{u}_2^{*2}\left(\frac{1}{6}\lambda_1(t_1^*-2)^3 - \frac{1}{2}\lambda_2 t_1^{*2}\right) - c_1 t_1^* \tag{170}$$

By property (e) of Theorem 3, it follows that

$$H(x_2^*(t_1^{*-}), u^*(t_1^{*-}), \overline{u}_1^*, p_1^*(t_1^{*-}), p_2^*(t_1^{*-})) = H(x_2^*(t_1^{*+}), u^*(t_1^{*+}), \overline{u}_2^*, p_1^*(t_1^{*+}), p_2^*(t_1^{*+})) \tag{171}$$

which gives that

$$p_1^*(t_1^{*-})x_2^*(t_1^{*-}) - \frac{1}{2}\overline{u}_1^{*2}p_2^*(t_1^{*-}) = p_1^*(t_1^{*+})x_2^*(t_1^{*+}) - \frac{1}{2}\overline{u}_2^{*2}p_2^*(t_1^{*+}) \tag{172}$$

By property (d) of Theorem 3, it follows that

$$\int_0^{t_1^*} \frac{\partial H}{\partial \overline{u}_1}dt + \frac{\partial \phi_1}{\partial \overline{u}_1} = 0 \tag{173}$$

$$\int_{t_1^*}^2 \frac{\partial H}{\partial \overline{u}_2}dt + \frac{\partial \phi_2}{\partial \overline{u}_2} = 0 \tag{174}$$

which gives that

$$\int_0^{t_1^*} p_2^*(t)u^*(t)dt + 48 = 0 \tag{175}$$

$$\int_{t_1^*}^2 p_2^*(t)u^*(t)dt + 49 = 0 \tag{176}$$

Thus far, the solution of optimal control problem is parameterized by some constants. These constants could be obtained by simultaneously solving the (172), (175), (176), E_3, E_4 of (152) and S_3 of (153), then the solution of optimal control is given by

$$\overline{u}_1^* = 2.21756,\ \ \overline{u}_2^* = 0.57997 \tag{177}$$

$$u^*(t) = \begin{cases} -112.61t + 25.797, & t \in [0, 0.46518) \\ 9.6569t - 11.445, & t \in [0.46518, 2] \end{cases} \tag{178}$$

The optimal solution and trajectory are given by Fig. 1.

4. Conclusions

In this paper, we present a special kind of hybrid system and its corresponding hybrid minimum principle. This kind of system's subsystem is characterized by parameters of model function, named hybrid parametric system. The parameters of model function are the mathematics representation of the batch operation (discrete control) in chemical process. Hence, dynamic performance of system will slightly change in every period, while the batch operation itself will bring some extra cost. Due to the difficulty of infinite dimension optimization, the hybrid parametric optimal control problem could only be analytically solved by necessary optimality conditions, i.e. hybrid parametric minimum principle, which is proposed and proved in this paper as Theorem 3. Moreover, there are some remarks on each section to cover more aspects about hybrid parametric system, and the validity of proposed theorems is verified by two analytic examples.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (21676295).

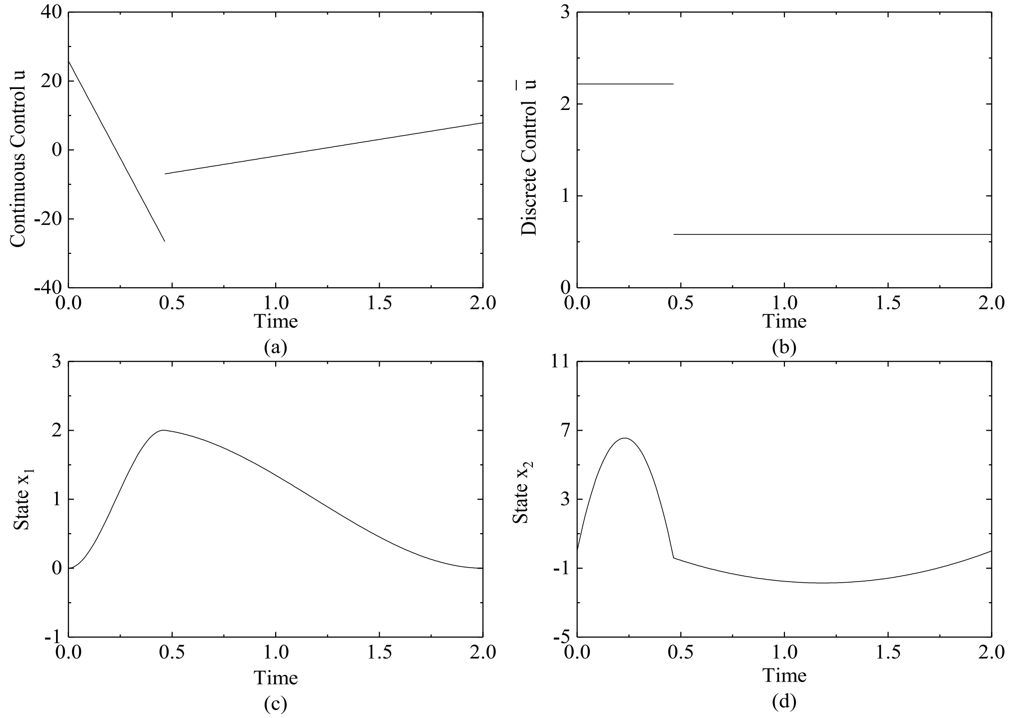

Fig. 1. Optimal solution and trajectory: (a) continuous control $u^*(t)$; (b) discrete control $\overline{u}^*$; (c) trajectory of state x_1^*; (d) trajectory of state x_2^*.

References

[1] S. Corner, C. Sandu, A. Sandu, Modeling and sensitivity analysis methodology for hybrid dynamical system, Nonlinear Anal.-Hybri. 31 (2019) 19–40.
[2] Z. Wang, H. An, X. Luo, Switch detection and robust parameter estimation for slowly switched hammerstein systems, Nonlinear Anal.-Hybri. 32 (2019) 202–213.
[3] A. Pakniyat, P.E. Caines, Hybrid optimal control of an electric vehicle with a dual-planetary transmission, Nonlinear Anal.-Hybri. 25 (2017) 263–282.
[4] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mishchenko, The Mathematical Theory of Optimal Processes, Wiley, New York, 1962.
[5] M. Branicky, V. Borkar, S. Mitter, A unified framework for hybrid control: Model and optimal control theory, IEEE Trans. Automat. Control 43 (1) (1998) 31–45.
[6] H.J. Sussmann, A maximum principle for hybrid optimal control problems, in: IEEE Conference on Decision and Control, 1999, pp. 425–430.
[7] H.J. Sussmann, Set-valued differentials and the hybrid maximum principle, in: IEEE Conference on Decision and Control, 2000, pp. 558–563.
[8] A. Pakniyat, P.E. Caines, The Hybrid Minimum Principle in the presence of switching costs, in: IEEE Conference on Decision and Control, 2013, pp. 3831–3836.
[9] A.V. Dmitruk, A.M. Kaganovich, The hybrid maximum principle is a consequence of pontryagin maximum principle, Syst. Control. Lett. 57 (11) (2008) 964–970.
[10] M.S. Shaikh, P.E. Caines, On the hybrid optimal control problem: theory and algorithms, IEEE Trans. Automat. Control 52 (9) (2007) 1587–1603.
[11] M. Garavello, B. Piccoli, Hybrid becessary principle, SIAM J. Control Optim. 43 (5) (2005) 1867–1887.
[12] M. Egerstedt, Y. Wardi, F. Delmotte, Optimal control of switching times in switched dynamical systems, in: IEEE Conference on Decision and Control, 2003, pp. 2138–2143.
[13] X. Xu, P.J. Antsaklis, Optimal control of switched systems via nonlinear optimization based on direct differentiations of value functions, Internat. J. Control 75 (16–17) (2002) 1406–1426.
[14] A. Bemporad, A. Giua, C. Seatzu, A master–slave algorithm for the optimal control of continuous-time switched affine systems, in: IEEE Conference on Decision and Control, 2002, pp. 1976–1981.
[15] A. Giua, C. Seatzu, C.V.D. Mee, Optimal control of switched autonomous linear systems, in: IEEE Conference on Decision and Control, 2001, pp. 2472–2477.
[16] H.J. Sussmann, A nonsmooth hybrid maximum principle, Lecture Notes Control Inf. Sci. 246 (1999) 325–354.
[17] A. Pakniyat, P.E. Caines, On the stochastic Minimum Principle for hybrid systems, in: IEEE Conference on Decision and Control, 2016, pp. 1139–1144.
[18] C. Aghayeva, Q. Abushov, The maximum principle for the nonlinear stochastic optimal control problem of switching systems, J. Global Optim. 56 (2) (2013) 341–352.
[19] A. Bensoussan, J. Menaldi, Stochastic hybrid control, J. Math. Anal. Appl. 249 (1) (2000) 261–288.
[20] S. Dharmatti, M. Ramaswamy, Hybrid control systems and viscosity solutions, SIAM J. Control Optim. 44 (4) (2005) 1259–1288.
[21] J.E.D. Silva, J.B.D. Sousa, F.L. Pereira, Dynamic programming based feedback control for systems with switching costs, in: IEEE International Conference on Control Applications, 2012, pp. 634–639.
[22] F. Borrelli, M. Baotic, A. Bemporad, M. Morari, Dynamic programming for constrained optimal control of discrete-time linear hybrid systems, Automatica 41 (10) (2005) 1709–1721.

[23] P.E. Caines, M. Egerstedt, R. Malhame, A. Schöllig, A hybrid Bellman equation for bimodal systems, in: International Conference on Hybrid Systems: Computation and Control, 2007, 656–659.
[24] Z. Yin, C. Gao, P. Li, Optimal control for hybrid systems based on dynamical programming, Control Theory Appl. 19 (1) (2002) 29–33.
[25] A. Pakniyat, P.E. Caines, On the relation between the minimum principle and dynamic programming for classical and hybrid control systems, IEEE Trans. Automat. Control 62 (9) (2017) 4347–4362.
[26] A. Pakniyat, P.E. Caines, On the minimum principle and dynamic programming for hybrid systems with low dimensional switching manifolds, IFAC Proc. 47 (3) (2014) 9629–9634.
[27] D. Chang, A simple proof of the pontryagin maximum principle on manifolds, Automatica 47 (3) (2011) 630–633.
[28] M. Barbero-Linan, M.C. Munoz-Lecanda, Geometric approach to Pontryagin's maximum principle, Acta Appl. Math. 108 (2) (2009) 429–485.
[29] F. Taringoo, P.E. Caines, On the optimal control of impulsive hybrid systems on riemannian manifolds, SIAM J. Control Optim. 51 (4) (2013) 3127–3153.
[30] F. Taringoo, P.E. Caines, Gradient-geodesic HMP algorithms for the optimization of hybrid systems based on the geometry of switching manifolds, in: IEEE Conference on Decision and Control, 2010, pp. 1534–1539.
[31] M.M. Sadeghi, E.F. Kececi, Hybrid power system for mobile robotics, in: International Conference on System Reliability and Science, 2017, pp. 64–67.
[32] M. Boccadoro, P. Valigi, A modelling approach for the dynamic scheduling problem of manufacturing systems with non-negligible setup times and finite buffers, in: IEEE Conference on Decision and Control, 2003, pp. 5472–5477.
[33] D. Flieller, J.P. Louis, J. Barrenscheen, General sampled data modeling of power systems supplied by static converter with digital and analog controller, Math. Comput. Simulation 46 (3) (1998) 373–385.
[34] G. Wang, Q. Wang, P. Wang, W. Wang, Engine control of power-split hybrid vehicle, Appl. Mech. Mater. 392 (5) (2013) 267–271.
[35] M.N. Pantano, M.C. Fernandez, M.E. Serrano, O.A. Ortiz, G. Scaglia, Tracking control of optimal profiles in a nonlinear fed-batch bioprocess under parametric uncertainty and process disturbances, Ind. Eng. Chem. Res. 57 (32) (2018) 11130–11140.
[36] A. Johnson, The control of fed-batch fermentation processes-a survey, Automatica 23 (6) (1987) 691–705.
[37] R. Wang, X. Luo, F. Xu, Economic and control performance of a fluid catalytic cracking unit: interactions between combustion air and CO promoters, Ind. Eng. Chem. Res. 53 (1) (2014) 287–304.
[38] R. Wang, X. Luo, F. Xu, Effect of CO combustion promoters on combustion air partition in FCC under nearly complete combustion, Chin. J. Chem. Eng. 22 (5) (2014) 531–537.
[39] J. Lin, F. Xu, X. Luo, Dynamic optimization of continuous chemical process with batch operation variable: a case study of a fluid catalytic cracking unit with CO promoter, Chem. Eng. Trans. 76 (2019) 751–756.
[40] V.G. Boltyanskii, Mathematical Methods of Optimal Control, Holt Rinehart & Winston, New York, 1971.
[41] Y. Gao, The proof of minimum principle and its deliberation, J. Wuhan Yejin Univ. Sci. Technol. 21 (1) (1998) 84–89.
[42] H. Axelsson, Y. Wardi, M. Egerstedt, E.I. Verriest, Gradient descent approach to optimal mode scheduling in hybrid dynamical systems, J. Optim. Theory Appl. 136 (2) (2008) 167–186.
[43] H. González, R. Vasudevan, M. Kamgarpour, S.S. Sastry, R. Bajcsy, C.J. Tomlin, A descent algorithm for the optimal control of constrained nonlinear switched dynamical systems, in: Acm International Conference on Hybrid Systems: Computation & Control, 2010, pp. 51–60.