On Weak Topology for Optimal Control of Switched Nonlinear Systems

Hua Chen and Wei Zhang

Abstract

Optimal control of switched systems is challenging due to the discrete nature of the switching control input. The embedding-based approach is one of the most effective approaches for addressing such a challenge. It tries to compute the optimal switching input by solving a corresponding relaxed optimal control problem with only continuous inputs, and then projecting the relaxed solution back to obtain the optimal switching solution of the original problem. This paper presents a novel idea that views the embedding-based approach as a change of topology over the optimization space, resulting in a general procedure to construct a switched optimal control algorithm with guaranteed convergence to a local optimizer. Our result provides a unified topology-based framework for the analysis and design of various embedding-based algorithms in solving the switched optimal control problem and includes many existing methods as special cases.

I. INTRODUCTION

Switched systems consist of a family of subsystems and a switching control signal determining the active subsystem (mode) at each time instant. Optimal control of switched systems involves finding both the continuous control input and the switching control input to jointly optimize certain system performance index. This problem has attracted considerable research attention due to its diverse engineering applications in power electronics [12], automotive systems [10], [15], [19], robotics [22], and manufacturing [4].

Optimal control of switched systems is in general challenging due to the discrete nature of the switching control input, which prevents us from directly applying the classical optimal control

Hua Chen and Wei Zhang are with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, Ohio, 43210. (e-mails: chen.3824@osu.edu; zhang.491@osu.edu)
techniques to solve the problem. To address this challenge, the maximum principle was extended in the literature to characterize optimal hybrid control solutions [13], [17], [18]. However, it is still very difficult to numerically compute the optimal solutions based on these abstract necessary conditions [24].

Among the rich literature, one well-known method is the so-called bilevel optimization [24], [25]. This approach involves solving two optimization problems at different levels. At the lower level, the approach fixes a switching mode sequence and optimizes the cost over the space of switching time instants through the classical variational approach. At the upper level, the switching mode sequence is updated to optimize the cost. Although various heuristic schemes have been proposed for the upper level [6], [7], [8], solutions obtained via this method may still be unsatisfactory due to the restriction on possible mode sequences.

More recently, an alternative approach based on the so-called embedding principle has been proposed [2], [20], [21], [23]. The embedding-based approach in general involves three key steps. The first step is to embed the switched systems into a larger class of classical nonlinear systems with only continuous control inputs. Then, the optimal control of the relaxed system is obtained using the classical optimal control algorithms. Once the relaxed optimal solution is obtained, the solution to the original problem can be computed by projecting the relaxed solution back to the original input space. This approach has been successfully applied to numerous applications, such as power electronics [12], automotive systems [15], [19], and robotics [22].

Several different versions of the embedding-based approach have been developed in the literature. These methods can be extended in their specific ways of embedding the switched trajectories, solving the associated classical optimal control problem, or projecting the relaxed solutions back to the original space. The goal of this paper is not to study these specific extensions. Instead, we aim to develop a general framework for analyzing and designing various embedding-based optimal control algorithms. The proposed framework is based on a novel idea that views the embedding-based approach as a change of topology over the optimization space. From this perspective, our framework adopts the weak topology structure and develops a general procedure to construct switched optimal control algorithms. Convergence of the constructed algorithm is guaranteed by specifications on several key components involved in the framework. Our result includes many existing results as special cases.

The rest of this paper is organized as follows: Section [11] formulates the optimal control
problem of switched systems. Preliminaries on topological space and weak topology are given in Section III, along with an example showing the importance of weak topology in optimal control of switched systems. Section IV presents our main results, establishing the unified framework and convergence analysis. Illustrating examples of our framework are presented in Section V. Concluding remarks are given in Section VI.

II. PROBLEM FORMULATION AND PRELIMINARIES

Consider a switched nonlinear system model

\[ \dot{x}(t) = f_{\sigma(t)}(t, x(t), u(t)), \text{ for a.e. } t \in [0, t_f], \]

where \( x(t) \in \mathbb{R}^{n_x} \) is the system state, \( u(t) \in U \subset \mathbb{R}^{n_u} \) is the continuous control input constrained in a compact and convex set \( U \), and \( \sigma(t) \in \Sigma \overset{\Delta}{=} \{1, 2, \ldots, n_\sigma\} \) is the switching control input determining the active subsystem (mode) among a finite number \( n_\sigma \) of subsystems at time \( t \).

The cost function considered in our optimal control problem is given by \( h(x(t_f)) \), i.e. only terminal state is penalized. Optimal control problems with nontrivial running costs can be transformed into this form by introducing an additional state variable [14]. It is also assumed that system (2) is subject to the following state constraints:

\[ h_j(x(t)) \leq 0, \quad \forall t \in [0, t_f], \quad \forall j \in J \overset{\Delta}{=} \{1, 2, \ldots, n_c\}. \]

The following assumption is adopted to ensure the existence and uniqueness of the state trajectory and the well-posedness of our optimal control problem.

**Assumption 1:**

1. \( f_i(t, x, u), \frac{\partial f_i}{\partial x}(t, x, u), \frac{\partial f_i}{\partial u}(t, x, u) \) are Lipschitz continuous with respect to all arguments for all \( i \in \Sigma \) with a common Lipschitz constant \( L \),

2. \( h_j(x), h(x), \frac{\partial h_j}{\partial x}(x), \frac{\partial h}{\partial x}(x) \) are Lipschitz continuous with respect to all argument for all \( j \in J \) with a common Lipschitz constant \( L \).

**Remark 1:** We assume a common Lipschitz constant to simplify notation. All the results in this paper extend immediately to the case where all these functions have different Lipschitz constants.
Following similar notations used in [2], [20], we rewrite the system dynamics as follows

\[ \dot{x} = \sum_{i=1}^{n_\sigma} d_i(t) f_i(t, x(t), u(t)) \]

\[ \triangleq f(t, x(t), u(t), d(t)), \text{ for a.e. } t \in [0, t_f], \]

where \( d(t) = [d_1(t), d_2(t), \ldots, d_{n_\sigma}(t)] \in D \triangleq \{ (d_1, \ldots, d_{n_\sigma}) \in \{0, 1\}^{n_\sigma} \mid \sum_{i=1}^{n_\sigma} d_i = 1 \} \), and \( D \) is the set of corners of the \( n_\sigma \) simplex. The continuous input \( u \) and switching input \( d \) can be viewed as mappings from \([0, t_f]\) to \( U \) and \( D \), respectively. In this paper, we assume these mappings to be elements of the \( L^2 \) space, defined as follows.

**Definition 1:** We say a function \( g : [0, t_f] \rightarrow G \subseteq \mathbb{R}^n \) belongs to \( L^2([0, t_f], G) \), if

\[ \|g\|_{L^2} \triangleq \left( \int_0^{t_f} \|g(t)\|^2 dt \right)^{\frac{1}{2}} < \infty, \]

where the integration is with respect to the Lebesgue measure.

Let \( U = L^2([0, t_f], U) \) be the space of continuous control inputs and let \( D = L^2([0, t_f], D) \) be the space of switching control inputs. We denote by \( \mathcal{X}_p = U \times D \) the overall input space and call \( \xi \in \mathcal{X}_p \) a pure input signal. Suppose the initial state \( x(0) = x_0 \in \mathbb{R}^{n_x} \) is given and fixed, we denote by \( x(t; \xi) \) the state at time \( t \) driven by \( \xi \) with initial state \( x_0 \). In order to emphasize the dependence on \( \xi \), the following notations are adopted in this paper:

\[ \phi_t(\xi) \triangleq x(t; \xi), J(\xi) \triangleq h(x(t_f; \xi)), \psi_{j,t}(\xi) \triangleq h_j(x(t; \xi)). \]

We further define \( \Psi(\xi) \triangleq \max_{j \in \mathcal{J}, t \in [0, t_f]} \{ \psi_{j,t}(\xi) \} \) and the state constraints in (2) can then be rewritten as \( \Psi(\xi) \leq 0 \), since \( \Psi(\xi) = \max_{j \in \mathcal{J}, t \in [0, t_f]} \{ \psi_{j,t}(\xi) \} \leq 0 \) if and only if \( \psi_{j,t}(\xi) \leq 0 \) for all \( j \in \mathcal{J} \) and \( t \in [0, t_f] \).

Adopting the above notations, the optimal control problem of switched systems considered in this paper is defined by:

\[ \mathcal{P}_{\mathcal{X}_p} : \begin{cases} \inf_{\xi \in \mathcal{X}_p} J(\xi), \\ \text{subj. to } \Psi(\xi) \leq 0. \end{cases} \]

The problem \( \mathcal{P}_{\mathcal{X}_p} \) can be viewed as a constrained optimization problem in function space \( \mathcal{X}_p \). However, the classical optimization techniques cannot be applied directly to solve this problem due to the discrete nature of \( D \). The embedding-based approach is one of the most
effective methods proposed in the literature for addressing this issue. This approach first embeds the switched systems into a larger class of traditional nonlinear systems with only continuous control inputs. Then, it solves an associated relaxed optimal control problem through the classical numerical optimal control algorithms. Lastly, it projects the relaxed optimal control back to the original input space to obtain the solution to the original problem. In this paper, we devise a novel idea that views the embedding-based approach as a change of topology, resulting in a general procedure for developing switched optimal control algorithms under the new topology. In the sequel, we shall first review some important concepts in topology in Section III and then establish the topology-based framework in Section IV.

III. TOPOLOGICAL SPACE AND WEAK TOPOLOGY

This section provides a brief review of several important concepts in topology and illustrates the importance of weak topology in switched optimal control problems through an numerical example.

Definition 2 (Topology [16]): Given a nonempty set $X$, a topology $\mathcal{T} \subset 2^X$ on $X$ is a collection of subsets of $X$, called open sets, such that:

1. $\emptyset$ and $X$ are open;
2. An arbitrary union of open sets is open;
3. A finite intersection of open sets is open.

A topological space is commonly denoted by $(X, \mathcal{T})$. Many different topologies can be defined on a given set $X$, and these topologies admit a partial order.

Definition 3 (Comparison of Topologies [16]): Let $\mathcal{T}_1, \mathcal{T}_2$ be two topologies defined on $X$. We say $\mathcal{T}_1$ is stronger (finer) than $\mathcal{T}_2$ if $\mathcal{T}_2 \subset \mathcal{T}_1$, and $\mathcal{T}_1$ is weaker (coarser) than $\mathcal{T}_2$ if $\mathcal{T}_1 \subset \mathcal{T}_2$.

The weak topology is an important topological notion utilizing the structure of continuous functions.

Definition 4 (Continuous Functions [5]): Given topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$. A function $g : X \mapsto Y$ from $X$ to $Y$ is continuous if and only if for any $V \in \mathcal{T}_Y$, $g^{-1}(V) \in \mathcal{T}_X$.

With the above definition of continuous functions, the weak topology is formally defined as follows.

Definition 5 (Weak Topology [16]): Let $\{g_i\}_{i \in \mathcal{I}}$ be a family of functions $g_i : X \mapsto Y_i$, $\forall i \in \mathcal{I}$, mapping from a set $X$ to several topological spaces $Y_i$, respectively. The weak topology on $X$
induced by \( \{g_i\}_{i \in I} \), denoted by \( T_{\{g_i\}_{i \in I}} \), refers to the weakest topology on \( X \) which makes all \( g_i \) continuous.

**Remark 2:** The structure of weak topology \( T_{\{g_i\}_{i \in I}} \) is determined by the family of functions \( \{g_i\}_{i \in I} \). In particular, the family may contain only one element. For example, the metric topology on a space \( X \) is defined to be the weak topology induced by a norm function \( \| \cdot \| \), denoted by \( T_{\| \cdot \|} \).

The topology selected over the optimization space plays a critical role in characterizing local optimizers of the underlying optimization problem. Before providing the formal definition of a local minimizer, we first define a neighborhood around a point \( \xi_p \in X_p \) under a topology \( T_g \) as follows:

**Definition 6:** Given the topological space \( (X_p, T_g) \), we say \( N_{T_g}(\xi_p, r) \subset (X_p, T_g) \) is a neighborhood around \( \xi_p \) under \( T_g \) if \( \exists O \in T_g \) such that \( \xi_p \in O \subset N_{T_g}(\xi_p) \). In particular, if we assume \( g : X_p \mapsto Y \) with \( Y \) a topological space endowed with a metric topology \( T_{\| \cdot \|_Y} \), a neighborhood around \( \xi_p \in X_p \) under \( T_g \) with radius \( r \) is defined by:

\[
N_{T_g}(\xi_p, r) = \{ \xi'_p \in X_p \| g(\xi_p) - g(\xi'_p) \|_Y \leq r \}. \tag{7}
\]

Employing the above definition, a local minimizer of \( P_{X_p} \) under a topology \( T_g \) is defined below.

**Definition 7:** We say \( \xi^*_p \in X_p \) is a local minimizer of \( P_{X_p} \) under the topology \( T_g \), if there exists a neighborhood \( N_{T_g}(\xi^*_p) \) such that \( J(\xi^*_p) \leq J(\xi'_p), \forall \xi'_p \in N_{T_g}(\xi^*_p) \cap \{ \xi_p \in X_p \| \Psi(\xi_p) \leq 0 \} \).

The following example is presented to illustrate that \( T_{\| \cdot \|} \) is not appropriate for \( P_{X_p} \) as it may lead to a solution with an unsatisfactory performance.

**Example:** Consider a switched system consisting two subsystems in the domain given by \( \{ x = (x_1, x_2)^T \in \mathbb{R}^2, x_1 \in [0, 4], x_2 \in [-1, 3] \} \). Dynamics of each mode is given by \( \dot{x} = f_1(x_1, x_2) = \begin{bmatrix} q_1(x_2) & 0 \end{bmatrix}^T \) and \( \dot{x} = f_2(x_1, x_2) = \begin{bmatrix} 0 & q_2(x_1) \end{bmatrix}^T \) respectively, where \( q_1 \) and \( q_2 \) are defined

\(^1\)Many norms can be defined on a space \( X \) and each of them induces a metric topology. In this paper, we assume \( T_{\| \cdot \|_X} \) to be the metric topology induced by the \( L^2 \) norm defined in \( \| \cdot \| \) on a function space \( X \) and the metric topology induced by the 2-norm on an Euclidean space \( X \).
as follows (see Fig. 1).

\[
q_1(x_2) = \begin{cases} 
  x_2 + 2, & \text{if } x_2 \in [-1, 0), \\
  x_2^2 - \frac{9}{7}x_2 + 2 & \text{if } x_2 \in [0, 1), \\
  \frac{12}{9-2x_2}, & \text{if } x_2 \in [1, 3].
\end{cases}
\]

\[
q_2(x_1) = \begin{cases} 
  2x_1 + 1, & \text{if } x_1 \in [0, 1), \\
  -2x_1 + 5, & \text{if } x_1 \in [1, 2), \\
  -\frac{2}{x_1-4}, & \text{if } x_1 \in [2, 3), \\
  -x_1 + 5, & \text{if } x_1 \in [3, 4].
\end{cases}
\]

\[\text{(8)}\]

![Fig. 1: Illustration of the functions } q_1(x_2) \text{ and } q_2(x_1) \]

For simplicity, we assume that neither continuous input nor state constraints are involved. Let the time horizon be \([0, 2]\) and let the initial condition be \(x(0) = A = (0, 0)^T\). The cost function is given by \(h(x(2; \xi)) = \|x(2; \xi) - B\|_2\) where \(B = (3, 3)^T\). In other words, we want to minimize the distance between the terminal state and the point \(B\). Due to the fact that the control input only consists of switching times, we assume the initial mode is 1 and define \(\xi = \{t_i\}_{i=1}^N\) to be a switching signal, where \(N\) is the total number of switches and \(0 \leq t_1 < t_2 < \ldots < t_N \leq 2\) are the switching times. Note that if \(t_1 = 0\), it means the initial mode is 2.

Let \(\xi_1 = \{1.5\}\) and \(\xi_2 = \{0.5, \frac{5}{6}\}\) be two switching control inputs. Given \(\xi_1\) as the initial point, if we adopt the classical metric topology \(T_{\|\cdot\|}\), \(\xi_1\) is the local minimizer with a cost of \(J(\xi_1) = 2\). To see this, we first consider a perturbation of \(\xi_1\) given by \(\hat{\xi}_1 = \{\hat{t}_1, \hat{t}_1 + \Delta t_1, 1.5\}\). By standard calculation, it can be shown that for \(\Delta t_1\) small enough and for any \(\hat{t}_1 \in (0, 1.5)\), we have \(x_2(2, \hat{\xi}_1) < 1\) and hence \(\|x_2(2, \hat{\xi}_1) - B\|_2 > 2\). For the other case where \(\hat{\xi}_1 = \{1.5, \hat{t}_1, \hat{t}_1 + \Delta t_1\}\), the same argument applies. Consequently, \(\xi_1\) is a local minimizer under \(T_{\|\cdot\|}\) with \(J(\xi_1) = 2\).

However, if we adopt the weak topology induced by the terminal state function, \(\xi_2\) is arbitrarily
close to $\xi_1$ under this new topology as they result in the same terminal state. If we define a perturbation of $\xi_2$ given by $\hat{\xi}_2 = \{0.5, \hat{t}_2\}$ parameterized by $\hat{t}_2 \in \left[\frac{5}{6}, \frac{3}{2}\right]$. Under $\hat{\xi}_2$, the terminal state will be $(3, 3\hat{t}_2 - \frac{3}{2})^T$ and the corresponding cost is $J(\hat{\xi}_2) = -3\hat{t}_2 + \frac{9}{2}$ for $\hat{t} \in \left[\frac{5}{6}, \frac{3}{2}\right]$. Clearly, the cost can be reduced to 0 by choosing $\hat{t}_2 = \frac{3}{2}$ under the weak topology. Notice that such a perturbation procedure is not valid under the metric topology $T_{||\cdot||}$ as $\xi_2$ is outside a sufficiently small neighborhood of $\xi_1$ and $\xi_1$ is already a local minimizer under $T_{||\cdot||}$. Hence, it is clear that

![State diagram](image.png)

Fig. 2: State trajectory generated by $\xi_1$ is highlighted by solid red path ending at point $C$. State trajectory generated by $\xi_2$ is highlighted by dotted green path which also ends at $C$. Additionally, one case of the state trajectory generated by $\xi$ with $\hat{t} = \frac{7}{6}$ is also shown as dash and dot blue path, ending at the midpoint of $CB$.

the metric topology is not appropriate for this switched optimal control problem and the weak topology induced by terminal state function is a better choice.

IV. A Unified Framework for Switched Optimal Control Problem

In this section, we establish a unified topology-based framework to solve $P_{X_p}$. We want to solve $P_{X_p}$ by constructing an algorithm $\Gamma_p : X_p \rightarrow X_p$ which generates a sequence of hybrid control inputs $\{\xi_i\}_{i \in \mathbb{N}}$ with guaranteed convergence to a local minimizer of $P_{X_p}$.

However, it is difficult in general to directly check the local minimizer condition in Definition 7 even for classical optimal control problems. In this paper, we adopt the following optimality function concept [14] to encode a necessary condition for local minimizers.

October 10, 2014 DRAFT
**Definition 8:** A function $\theta_p(\cdot) : \mathcal{X}_p \rightarrow \mathbb{R}$ satisfying the following conditions is called an optimality function for $\mathcal{P}_{\mathcal{X}_p}$:

1. $\theta_p(\xi) \leq 0$ for all $\xi \in \mathcal{X}_p$;
2. if $\xi_p^*$ is a local minimizer of $\mathcal{P}_{\mathcal{X}_p}$, then $\theta_p(\xi_p^*) = 0$.

**Remark 3:** Often times, the optimality function is required to be continuous (or upper semi-continuous) [14]. Such a condition is introduced to ensure that in a topological space, if $\xi^*$ is an accumulation point of the sequence $\{\xi_p^i\}_{i \in \mathbb{N}}$ and if $\liminf_{i \to \infty} \theta_p(\xi_p^i) = 0$, then $\theta_p(\xi^*) = 0$. However, in our problem we do not assume the existence of accumulation points of the sequence $\{\xi_p^i\}_{i \in \mathbb{N}}$. Hence, the continuity (or upper semi-continuity) condition is not necessary. Detailed discussion can be found in [1].

Employing this optimality function definition and the necessary optimality condition encoded therein, our goal becomes constructing the algorithm $\Gamma_p$ such that $\theta_p(\xi_p^i) \to 0$ as $i \to \infty$. In the following subsection, we establish the unified topology-based framework for constructing such an algorithm.

### A. Solution Framework

Our topology-based framework involves three key steps and several important components given as follows.

1. Relax the optimization space $\mathcal{X}_p$ to a vector space $\mathcal{X}_r$, select a weak topology function $g : \mathcal{X}_r \mapsto Y$ and construct a projection operator $R_k$ associated with the weak topology $T_g$.
2. Solve the relaxed optimal control problem $\mathcal{P}_{\mathcal{X}_r}$ defined in (10) below by designing a relaxed optimality function $\theta_r$ and selecting (or constructing) a relaxed optimal control algorithm $\Gamma_r$.
3. Set $\theta_p = \theta_r |_{\mathcal{X}_p}$ and $\Gamma_p = R_k \circ \Gamma_r$. Then a local optimizer can be approached through the following iteration:

$$\xi_p^{i+1} = \begin{cases} 
\Gamma_p(\xi_p^i), & \text{if } \theta_p(\xi_p^i) < 0, \\
\xi_p^i, & \text{if } \theta_p(\xi_p^i) = 0.
\end{cases}$$

(9)

For simplicity, we denote by $\{\xi_p^i\}_{i \in \mathbb{N}}$ the sequence generated by (9).

The relaxed optimal control problem $\mathcal{P}_{\mathcal{X}_r}$ in the above framework is given by

$$\mathcal{P}_{\mathcal{X}_r} : \begin{cases} 
\inf_{\xi \in \mathcal{X}_r} J(\xi) , \\
\text{subj. to } \Psi(\xi) \leq 0,
\end{cases}$$

(10)
and the relaxed optimality function $\theta_r$ is defined by replacing $X_p$ and $P_{X_p}$ in Definition 8 with $X_r$ and $P_{X_r}$.

The main underlying idea of the proposed framework is to transform the switched optimal control problem $P_{X_p}$ to a classical optimal control problem $P_{X_r}$ which can be solved through the classical gradient-based methods in functional spaces [11], [14]. The solution of $P_{X_r}$ will then be used to construct the solution to the original problem $P_{X_p}$. The key components of the framework include the relaxed optimization space $X_r$, the weak topology $T_g$, the projection operator $R_k$, and the relaxed optimization algorithm characterized by $\theta_r$ and $\Gamma_r$.

In the rest of this section, we will first show that $\theta_p$ is an optimality function for $P_{X_p}$ and then derive conditions for the aforementioned key components of our framework to guarantee that the sequence $\{\xi^i_{\theta_p}\}_{i \in \mathbb{N}}$ converges to a point satisfying the necessary optimality condition encoded in $\theta_p$.

**B. Convergence Analysis and Proofs**

Before stating our main results, we first impose the following assumptions on $X_r$, $T_g$ and $R_k$ in the framework to ensure its validity.

**Assumption 2:**

1. $J$ and $\Psi$ are Lipschitz continuous under topology $T_g$ with a common Lipschitz constant $L$.
2. $X_p$ is dense in $X_r$ under $T_g$, i.e. $\forall \xi_r \in X_r$, $\forall \epsilon > 0$, $\exists \xi_p \in X_p$ s.t. $\|g(\xi_r) - g(\xi_p)\|_Y \leq \epsilon$.
3. There exists a projection operator $R_k : X_r \to X_p$ associated with $T_g$ and parametrized by $k = 1, 2, \ldots$, such that $\forall \xi_r \in X_r$, $\forall \epsilon > 0$, there exists a $\hat{k} \in \mathbb{N}$, such that
   \[ \|g(R_k(\xi_r)) - g(\xi_r)\|_Y \leq e_{R_k}(k) \leq \epsilon, \quad \forall k \geq \hat{k}. \]  

Assumption 2.1 is a standard Lipschitz continuity condition that ensures the well-posedness of the relaxed problem $P_{X_r}$. Assumption 2.2 assigns a relationship between $X_r$ and $X_p$ under the weak topology structure which guarantees performance of the projection step in the embedding-based approach. However, it gives no instruction on how to construct a pure control input close enough to a given relaxed control input. Assumption 2.3 is adopted to provide a specific way of constructing such a pure control input with an known approximation error.

In the following lemma, we show that $\theta_p = \theta_r|_{X_p}$ is an optimality function for $P_{X_p}$.
Lemma 1: If $\theta_r$ is a valid optimality function for $\mathcal{P}_{\mathcal{X}_r}$, then $\theta_p = \theta_r|_{\mathcal{X}_p}$ is a valid optimality function for $\mathcal{P}_{\mathcal{X}_p}$.

Proof: To prove this lemma, we need to show $\theta_p$ satisfies the conditions in Definition 8. The first condition is trivially satisfied. For the second condition, suppose it does not hold, i.e. suppose $\xi^* \in \mathcal{X}_p$ is a local minimizer for $\mathcal{P}_{\mathcal{X}_p}$ but $\theta_p(\xi^*) < 0$. Since $\theta_r(\xi^*) = \theta_p(\xi^*)$, by the definition of local minimizers for $\mathcal{P}_{\mathcal{X}_r}$, it follows that there exists a $\xi_r$ and a positive number $C$, such that $J(\xi_r) - J(\xi^*) \leq -C$ and $\Psi(\xi_r) \leq -C$. By Assumption 2, we have $|J(\mathcal{R}_k(\xi_r)) - J(\xi_r)| \leq L \|g(\mathcal{R}_k(\xi_r)) - g(\xi_r)\|_Y \leq Le_{\mathcal{R}_k}(k)$. By adding and subtracting $J(\xi_r)$, it follows that

$$
J(\mathcal{R}_k(\xi_r)) - J(\xi^*) \leq |J(\mathcal{R}_k(\xi_r)) - J(\xi_r)| + J(\xi_r) - J(\xi^*) \leq Le_{\mathcal{R}_k}(k) - C
$$

(12)

For any given $\xi_r \in \mathcal{X}_r$, choose $\epsilon = \frac{C}{2L}$ in Assumption 3. For $k \geq \hat{k}$, it follows that $Le_{\mathcal{R}_k}(k) - C \leq -\frac{C}{2} < 0$, hence $J(\mathcal{R}_k(\xi_r)) - J(\xi^*) < 0$. A similar argument can be applied on $\Psi$, yielding that $\Psi(\mathcal{R}_k(\xi_r)) \leq 0$. These statements contradict that $\xi^*$ is a local minimizer for $\mathcal{P}_{\mathcal{X}_p}$. □

To show the convergence of $\{\xi^*_i\}_{i \in \mathbb{N}}$, we adopt a similar idea of the sufficient descent property presented in [11]. In order to handle the projection step in our framework and the state constraints considered in our problem, we define two functions $Q : \mathcal{X}_p \times \mathbb{N} \mapsto \mathbb{R}$ and $P : \mathcal{X}_r \times \mathcal{X}_r \mapsto \mathbb{R}$ below.

$$
Q(\xi, k) \triangleq \begin{cases} 
\max\{J(\mathcal{R}_k \circ \Gamma_r(\xi)) - J(\xi), 
\Psi(\mathcal{R}_k \circ \Gamma_r(\xi)) - \Psi(\xi)\} & \text{if } \Psi(\xi) \leq 0, \\
\Psi(\mathcal{R}_k \circ \Gamma_r(\xi)) - \Psi(\xi) & \text{if } \Psi(\xi) > 0.
\end{cases}
$$

(13)

$$
P(\xi_1, \xi_2) \triangleq \begin{cases} 
\max\{J(\xi_2) - J(\xi_1), \Psi(\xi_2)\} & \text{if } \Psi(\xi_1) \leq 0, \\
\Psi(\xi_2) - \Psi(\xi_1) & \text{if } \Psi(\xi_1) > 0.
\end{cases}
$$

(14)

We introduce the function $Q$ to compactly characterize the change of the value of the cost $J$ and the constraint $\Psi$ at a point $\xi$ under the projection operator $\mathcal{R}_k$. For a feasible point, we care about both the changes of the cost and the constraint under $\mathcal{R}_k$ and for an infeasible point, we only care about the change of the constraint.
The function $P$ characterizes the value difference of $J$ and $\Psi$ between two points $\xi_1$ and $\xi_2$. If $\xi_1$ is feasible and $P < 0$, it means the cost can be reduced while maintaining feasibility by moving from $\xi_1$ to $\xi_2$. Similarly, if $\xi_1$ is infeasible and $P < 0$, it is possible to reduce the infeasibility by moving from $\xi_1$ to $\xi_2$.

Exploiting Assumption 2.3, a bound for the function $Q$ is derived in the following lemma.

**Lemma 2:** There exists a $k^* \in \mathbb{N}$ such that given $\omega \in (0, 1)$, for any $C > 0$, $\gamma_C > 0$, and for any $\xi \in X_p$ with $\theta_p(\xi) < -C$, we have

$$Q(\xi, k) \leq (\omega - 1)\gamma_C\theta_p(\xi), \quad \forall k \geq k^*.$$  \hfill (15)

**Proof:** This is a straightforward result from Assumption 2.1 and Assumption 2.2. \hfill \blacksquare

Employing the definition of the function $P$ and the above two lemmas, our main result on the convergence of $\{\xi^i_p\}_{i \in \mathbb{N}}$ is presented below.

**Theorem 1:** If for each $C > 0$, there exists a $\gamma_C > 0$ such that for any $\xi_r \in X_r$ with $\theta_r(\xi_r) < -C$,

$$P(\xi_r, \Gamma_r(\xi_r)) \leq \gamma_C\theta_r(\xi_r) < 0,$$  \hfill (16)

then for an appropriate choice of $k$ for $R_k$, we have

1. if there exists a $i_0 \in \mathbb{N}$ such that $\Psi(\xi^i_p) \leq 0$, then $\Psi(\xi^i_p) \leq 0$ for all $i \geq i_0$,

2. $\lim_{i \to \infty} \theta_p(\xi^i_p) = 0$.

**Proof:**

1. Suppose there exists an $i_0$ such that $\Psi(\xi^i_p) \leq 0$, then we have for $k \geq k^*$

$$\Psi(\xi^{i_0+1}_p)$$

$$= \Psi(\mathcal{R}_k(\Gamma_r(\xi^i_p))) - \Psi(\Gamma_r(\xi^i_p)) + \Psi(\Gamma_r(\xi^i_p))$$

$$\leq (\omega - 1)\gamma_C\theta_p(\xi^i_p) + \gamma_C\theta_r(\xi^i_p)$$

$$= \omega\gamma_C\theta_p(\xi^i_p) < 0$$  \hfill (17)

2. We need to consider two cases due to different form of $P$ for different values of $\Psi$.

- $\Psi(\xi^i_p) > 0$ for all $i \in \mathbb{N}$, i.e. the entire sequence is infeasible.

  Suppose $\lim_{i \to \infty} \theta_p(\xi^i_p) \neq 0$, since $\theta_p(\cdot)$ is a non-positive function, we know there must exists $C > 0$ such that $\liminf_{i \to \infty} \theta_p(\xi^i_p) = -2C$. Hence, there exists an infinite subsequence $\{\xi^{i_m}_p\}$ and
an $m_1 \in \mathbb{N}_+$ such that $\theta_p(\xi_p^{i_{m_1}}) < -C$ for all $m \geq m_1$. Then, it follows that for all $m \geq m_1$, and for $k \geq k^*$, we have

$$\Psi(\xi_p^{i_{m+1}}) - \Psi(\xi_p^{i_m}) = \Psi(\mathcal{R}_k \circ \Gamma_r(\xi_p^{i_{m+1}})) - \Psi(\Gamma_r(\xi_p^{i_m})) + \Psi(\Gamma_r(\xi_p^{i_m})) - \Psi(\xi_p^{i_m}) \leq (\omega - 1) \gamma C \theta_p(\xi_p^{i_0}) + \gamma C \theta_r(\xi_p^{i_0})$$

$$= \omega \gamma C \theta_p(\xi_p^{i_0}) < 0$$

This leads to the fact that $\liminf_{m \to \infty} \Psi(\xi_p^{i_m}) = -\infty$, which contradicts the lower boundedness of $\Psi$ implied by Assumption 1.

- There exists an $i_0$ such that $\Psi(\xi_p^{i_0}) \leq 0$. By conclusion (1), it follows that $\Psi(\xi_p^{i_{i_0}}) \leq 0$ for all $i \geq i_0$.

An analogous argument to the previous case applies. Suppose $\liminf_{i \to \infty} \theta_p(\xi_p^{i_i}) \neq 0$, then there exists $C > 0$ such that $\liminf_{i \to \infty} \theta_p(\xi_p^{i_i}) = -2C$. Hence, there exists an infinite subsequence $\{\xi_p^{i_m}\}$ and a $m_1 \in \mathbb{N}_+$ such that $\theta_p(\xi_p^{i_m}) < -C$ for all $m \geq m_1$. Then, it follows that for all $m \geq m_1$ and for all $k \geq k^*$, we have:

$$J(\xi_p^{i_{m+1}}) - J(\xi_p^{i_m}) = J(\mathcal{R}_k \circ \Gamma_r(\xi_p^{i_{m+1}})) - J(\xi_p^{i_m})$$

$$= J(\mathcal{R}_k \circ \Gamma_r(\xi_p^{i_{m+1}})) + \gamma C \theta_r(\xi_p^{i_0}) - J(\xi_p^{i_m}) \leq (\omega - 1) \gamma C \theta_p(\xi_p^{i_0}) + \gamma C \theta_r(\xi_p^{i_0})$$

$$= \omega \gamma C \theta_p(\xi_p^{i_0}) < 0$$

This leads to the fact that $\liminf_{m \to \infty} J(\xi_p^{i_m}) = -\infty$, which contradicts with the lower boundedness of $J$ implied by Assumption 1.

In the following section, an example of specific selections of the components in the framework is presented to show how a switched optimal control algorithm can be constructed.

V. ILLUSTRATING EXAMPLE

Numerous embedding-based switched optimal control algorithms proposed in the literature can be analyzed using this framework. In this section, we first provide a concrete example for each
component in an embedding-based optimal control algorithm and illustrate how the framework can be used to analyze this algorithm. Then, alternative possible selections of the components are also provided to show the effectiveness of the framework in designing various switched optimal control algorithms.

A. Specific Selections of Framework Components

In this subsection, we will show the particular algorithm developed in [20] is a special case of our framework.

1. \( \mathcal{X}_r \): We relax the optimization space by letting \( D_r \triangleq \{ (d_1, \ldots, d_n) \in [0, 1]^n \sigma \big| \sum_{i=1}^n d_i = 1 \} \), and define \( D_r = L^2([0, t_f], D_r) \) and \( \mathcal{X}_r = \mathcal{U} \times D_r \) accordingly.

2. \( \mathcal{T}_\phi \): We select \( \mathcal{T}_\phi \) to be the weak topology induced by the state trajectory function \( \phi_t(\cdot) \) defined in (5) and denote this topology by \( \mathcal{T}_{\phi_t} \).

3. \( \mathcal{R}_k \): We construct the projection operator to be the composition of Haar wavelet approximation operator \( \mathcal{H}_k(\cdot) \) [9], and the Pulse-Width Modulation (PWM) operator \( \mathcal{W}_k(\cdot) \).

4. \( \theta_r \): We define the optimality function \( \theta_r \) and the corresponding descent direction \( \vartheta \) as follows.

\[
\theta_r(\xi) = \min_{\xi_r \in \mathcal{X}_r} \zeta(\xi, \xi_r), \quad \vartheta(\xi) = \arg\min_{\xi_r \in \mathcal{X}_r} \zeta(\xi, \xi_r), \quad (20)
\]

where

\[
\zeta(\xi, \xi_r) = \max \left\{ DJ(\xi; \xi_r - \xi) - \Psi_+(\xi), \max_{j \in J, t \in [0, t_f]} \{ D\psi_{j,t}(\xi; \xi_r - \xi) \} + \gamma \Psi_-(\xi) \right\} + \| \xi_r - \xi \|_{\mathcal{X}_r}
\]

with \( \gamma > 0 \) a design parameter, \( \Psi_+(\xi_r) = \max \{ 0, \Psi(\xi_r) \} \), \( \Psi_-(\xi_r) = \min \{ 0, \Psi(\xi_r) \} \) and \( \text{DH}(x; x') = \lim_{\lambda \downarrow 0} \frac{H(x + \lambda x') - H(x)}{\lambda} \) the directional derivative for function \( H \) at \( x \) along direction \( x' \).

5. \( \Gamma_r \): We select \( \Gamma_r \) to be the line search algorithm with Armijo step size rule.

Now, we briefly show that the above selections satisfy the framework requirements that guarantee convergence of the switched optimal control algorithm.

1. Assumption 2.2: The construction of \( \mathcal{X}_r \) and selection of \( \mathcal{T}_{\phi_t} \) guarantee Assumption 2.2 due to an immediate extension of chattering lemma (Corollary 2 in [2]).

2. Assumption 2.3: To show \( \mathcal{R}_k = \mathcal{H}_k \circ \mathcal{W}_k \) satisfies Assumption 2.3, we need to show 1) \( \forall \xi_r \in \mathcal{X}_r, \mathcal{R}_k(\xi_r) \in \mathcal{X}_p \) and 2) \( \forall \xi_r \in \mathcal{X}_r \) and \( \forall \epsilon > 0, \exists k^* \in \mathbb{N} \) such that \( \| g(\mathcal{R}_k(\xi_r)) - g(\xi_r) \|_Y \leq \epsilon \).
The Haar wavelet approximation of a function \( \varphi \) of order \( k \) is given by
\[
\mathcal{H}_k(\varphi) = \langle \varphi, 1_{[0,t_f]} \rangle > + \sum_{j=0}^{k} \sum_{n=0}^{2^j-1} \langle \varphi, \rho_{j,n} \rangle \rho_{j,n},
\]
where the Haar basis functions \( \{\rho_{j,n}\} \) are given by
\[
\rho_{j,n}(t) = 2^j \rho(2^j t - n)
\]
with
\[
\rho(t) = \begin{cases} 
1, & t \in [0, \frac{t_f}{2}) , \\
-1, & t \in [\frac{t_f}{2}, t_f) , \\
0, & \text{otherwise .}
\end{cases}
\]

The PWM operator of a discrete control input \( d \in D_r \) with frequency \( 2^{-k} \) is given by
\[
[W_k(d)]_i(t) = \begin{cases} 
1, & \text{if } t \in [S_{k,j,i-1}, S_{k,j,i}) , \\
0, & \text{otherwise ,}
\end{cases}
\]

where \( S_{k,j,i} = 2^{-k} \left( j + \sum_{n=1}^{i} d_n \left( \frac{j}{2^k} \right) \right) t_f \) and \( j \in \{0, 1, \ldots, 2^k - 1\} \).

It can be shown that \( \mathcal{H}_k(d) \in D_r \) for any \( d \in D_r \) and any \( k \in \mathbb{N} \), due to Section 3.3 in [9] and definition and linearity of \( \mathcal{H}_k \). Furthermore, for any \( \xi_r = (u, d) \in \mathcal{X}_r \), \( R_k(\xi_r) = (\mathcal{H}_k(u), W_k(\mathcal{H}_k(d))) \in \mathcal{X}_p \) due to the definition of \( W_k \).

The following error bound applies for the projection operator \( P_k \):
\[
\| \phi_t(\mathcal{R}_k(\xi_r)) - \phi_t(\xi_r) \|_2 \leq M(\xi_r) \left( \frac{1}{\sqrt{2}} \right)^k , \forall t \in [0,t_f] ,
\]

where \( M(\xi_r) \) is a positive real-valued function defined on \( \mathcal{X}_r \) which takes finite values when \( \xi_r \) is of bounded variation due to the result in [20]. Therefore, fix \( \xi_r \in \mathcal{X}_r \) and for all \( \epsilon > 0 \), there exists a \( k^* \in \mathbb{N} \) such that \( M(\xi_r)(\frac{1}{\sqrt{2}})^k < \epsilon \) for all \( k \geq k^* \).

3. Validity of \( \theta_r \): We need to show 1) \( \theta_r(\xi_r) \leq 0, \forall \xi_r \in \mathcal{X}_r \) and 2) if \( \xi^*_r \) is a local minimizer for \( \mathcal{P}_{\mathcal{X}_r} \), then \( \theta_r(\xi_r) = 0 \). The first condition is simple to verify, as \( \theta_r(\xi_r) \leq \zeta(\xi_r, \xi_r) = 0 \) for any \( \xi_r \in \mathcal{X}_r \). The proof of the second condition follows from a similar argument in the proof of Lemma [1] by taking advantage of the definition of \( \theta_r \).

4. Validity of \( \Gamma_r \): The line search algorithm is given by \( \Gamma_r : \mathcal{X}_r \to \mathcal{X}_r \) such that
\[
\Gamma_r(\xi_r) = \xi_r + \beta(\xi_r)(\partial(\xi_r) - \xi_r), \forall \xi_r \in \mathcal{X}_r
\]
where \( g(\xi_r) \) is defined in (20), \( \beta \in (0, 1) \) and the step size \( \beta^{\mu(\xi_r)} \) is determined by the Armijo step size rule given in (28)

\[
\mu(\xi_r) = \begin{cases} 
\min\{k \in \mathbb{N} \mid J(\xi_r + \beta^k(\vartheta(\xi_r) - \xi_r)) - J(\xi_r) \leq \alpha \beta^k \theta(\xi_r), \\
\psi(\xi_r + \beta^k(\vartheta(\xi_r) - \xi_r)) - \psi(\xi_r) \leq \alpha \beta^k \theta(\xi_r) \}, & \text{if } \theta(\xi_r) \leq 0, \\
\min\{k \in \mathbb{N} \mid \psi(\xi_r + \beta^k(\vartheta(\xi_r) - \xi_r)) - \psi(\xi_r) \leq \alpha \beta^k \theta(\xi_r) \}, & \text{if } \theta(\xi_r) > 0,
\end{cases}
\]

with \( \alpha \in (0, 1) \). It can be proved that for any \( \xi_r \in \mathcal{X}_r \) with \( \theta_r(\xi_r) < 0 \), \( \mu(\xi_r) < \infty \). Hence, it can be easily verified that the line search algorithm with Armijo step size rule meets the condition specified in Theorem 1.

**B. Alternative Selections**

Depending on the underlying applications, one may want to choose different relaxed spaces \( \mathcal{X}_r \), topologies \( T_g \), optimality functions \( \theta_r \), projection operators \( R_k \), or relaxed optimization algorithms \( \Gamma_r \). Each combination of these components will lead to a different switched optimal control algorithm that may have a better performance. To be concise, we consider the same relaxed space \( \mathcal{X}_r \) and present several alternative choices of other components which will preserve the convergence of the sequence of hybrid control inputs that generated by the newly constructed algorithms.

1. \( T_g \): The weak topology can also be chosen as the weak topology induced by the terminal state function \( \phi_{t_f}(\cdot) \), denoted by \( T_{\phi_{t_f}} \). Assumption 2.2 is satisfied under this selection since \( T_{\phi_{t_f}} \) is weaker than \( T_{\phi_t} \).

2. \( R_k \): The projection operator can be constructed as \( R_k = F_k \circ \mathcal{W}_k \) where \( F_k \) can be any wavelet approximation operator of order \( k \) with one vanishing moment, defined as follows

\[
F_k(\varphi) = \langle \varphi, \mathbb{1}_{[0,t_f]} \rangle + \sum_{j=0}^{k} \sum_{i=1}^{2^j} \langle \varphi, \eta_{j,i} \rangle \eta_{j,i},
\]

where the basis functions \( \{\eta_{j,i}\} \) are orthonormal, oscillatory signals with the following property

\[
\int_0^{t_f} \eta_{j,i}(t) dt = 0, \forall j, i.
\]

3. \( \Gamma_r \): We assume that the relaxed optimal control algorithm \( \Gamma_r \) generates a sequence by the following rule

\[
\xi_{r}^{i+1} = \xi_{r}^{i} + \alpha^i(\vartheta(\xi_{r}^{i}) - \xi_{r}^{i}),
\]
with $\alpha^i$ to be determined. The algorithms can be selected as the steepest descent algorithm with exact line search, the line search algorithm with Goldstein rule or the line search algorithm with limited minimization rule $[3]$.  

The validity of these selections can be verified using similar arguments to those presented in the previous subsection. Different combinations of the components presented here and before will result in different algorithms and all the algorithms solve $\mathcal{P}_{X_p}$.

As discussed in this section, the proposed framework can be used to analyze and design various switched optimal control algorithms, which demonstrates its importance in optimal control of switched systems.

VI. Conclusions

In this paper, we present a unified topology-based framework that is used for designing and analyzing various embedding-based switched optimal control algorithms. Our framework is based on a novel viewpoint which considers the embedding-based methods as a change of topology over the optimization space. From this viewpoint, our framework adopts the weak topology structure and develops a general procedure to construct a switched optimal control algorithm. Convergence property of the algorithm is guaranteed by specifications on several key components involved in the framework. Examples are shown to illustrate how the proposed framework can be used to analyze and design various switched optimal control algorithms.

REFERENCES

[1] H. Axelsson, Y. Wardi, M. Egerstedt, and E. Verriest, “Gradient descent approach to optimal mode scheduling in hybrid dynamical systems,” *Journal of Optimization Theory and Applications*, vol. 136, no. 2, pp. 167–186, 2008.

[2] S. C. Bengea and R. A. DeCarlo, “Optimal control of switching systems,” *Automatica*, vol. 41, no. 1, pp. 11 – 27, 2005.

[3] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[4] C. Cassandras, D. Pepyne, and Y. Wardi, “Optimal control of a class of hybrid systems,” *Automatic Control, IEEE Transactions on*, vol. 46, no. 3, pp. 398–415, Mar 2001.

[5] R. M. Dudley, *Real analysis and probability*. Cambridge University Press, 2002, vol. 74.

[6] M. Egerstedt, Y. Wardi, and H. Axelsson, “Transition-time optimization for switched-mode dynamical systems,” *Automatic Control, IEEE Transactions on*, vol. 51, no. 1, pp. 110–115, 2006.

[7] H. Gonzalez, R. Vasudevan, M. Kamgarpour, S. S. Sastry, R. Bajcsy, and C. J. Tomlin, “A descent algorithm for the optimal control of constrained nonlinear switched dynamical systems,” in *Proceedings of the 13th ACM international conference on Hybrid systems: computation and control*. ACM, 2010, pp. 51–60.
[8] ——, “A numerical method for the optimal control of switched systems,” in Decision and Control. Proceedings of the 49th IEEE Conference on, 2010, pp. 7519–7526.
[9] A. Haar, “Zur theorie der orthogonalen funktionensysteme,” Mathematische Annalen, vol. 69, no. 3, pp. 331–371, 1910.
[10] S. Hedlund and A. Rantzer, “Optimal control of hybrid systems,” in Decision and Control. Proceedings of the 38th IEEE Conference on, vol. 4, 1999, pp. 3972–3977.
[11] D. Kirk, “Optimal control theory: an introduction,” Prentice-Hall network series, 1970.
[12] F. M. Oettmeier, J. Neely, S. Pekarek, R. DeCarlo, and K. Uthaichana, “MPC of switching in a boost converter using a hybrid state model with a sliding mode observer,” Industrial Electronics, IEEE Transactions on, vol. 56, no. 9, pp. 3453–3466, 2009.
[13] B. Piccoli, “Hybrid systems and optimal control,” in Decision and Control. Proceedings of the 37th IEEE Conference on, vol. 1, 1998, pp. 13–18.
[14] E. Polak, Optimization: Algorithms and Consistent Approximations, ser. Applied mathematical sciences. Springer-Verlag, 1997.
[15] M. Rinehart, M. Dahleh, D. Reed, and I. Kolmanovsky, “Suboptimal control of switched systems with an application to the disc engine,” Control Systems Technology, IEEE Transactions on, vol. 16, no. 2, pp. 189–201, March 2008.
[16] W. Rudin, Functional analysis, 2nd ed. McGraw-Hill, Inc., New York, 1991.
[17] H. J. Sussmann, “A maximum principle for hybrid optimal control problems,” in Decision and Control. Proceedings of the 38th IEEE Conference on, vol. 1, 1999, pp. 425–430.
[18] ——, “Set-valued differentials and the hybrid maximum principle,” in Decision and Control. Proceedings of the 39th IEEE Conference on, vol. 1, 2000, pp. 558–563.
[19] K. Uthaichana, R. DeCarlo, S. Bengea, S. Pekarek, and M. Žefran, “Hybrid optimal theory and predictive control for power management in hybrid electric vehicle,” Journal of Nonlinear Systems and Applications, vol. 2, no. 1-2, pp. 96–110, 2011.
[20] R. Vasudevan, H. Gonzalez, R. Bajcsy, and S. Sastry, “Consistent approximations for the optimal control of constrained switched systems—part 1: A conceptual algorithm,” SIAM Journal on Control and Optimization, vol. 51, no. 6, pp. 4463–4483, 2013.
[21] ——, “Consistent approximations for the optimal control of constrained switched systems—part 2: An implementable algorithm,” SIAM Journal on Control and Optimization, vol. 51, no. 6, pp. 4484–4503, 2013.
[22] S. Wei, K. Uthaichana, M. Žefran, and R. DeCarlo, “Hybrid model predictive control for the stabilization of wheeled mobile robots subject to wheel slippage,” Control Systems Technology, IEEE Transactions on, vol. 21, no. 6, pp. 2181–2193, Nov 2013.
[23] S. Wei, K. Uthaichana, M. Žefran, R. DeCarlo, and S. Bengea, “Applications of numerical optimal control to nonlinear hybrid systems,” Nonlinear Analysis: Hybrid Systems, vol. 1, no. 2, pp. 264–279, 2007.
[24] X. Xu and P. J. Antsaklis, “Results and perspectives on computational methods for optimal control of switched systems,” in Hybrid Systems: Computation and Control. Springer, 2003, pp. 540–555.
[25] ——, “Optimal control of switched systems based on parameterization of the switching instants,” Automatic Control, IEEE Transactions on, vol. 49, no. 1, pp. 2–16, 2004.