# Optimal Control of Piecewise-Smooth Control Systems
# via Singular Perturbations

Tyler Westenbroek[1]     Xiaobin Xiong[2]     Aaron D. Ames[2]     S. Shankar Sastry[1]

*Abstract*— This paper investigates optimal control problems formulated over a class of piecewise-smooth controlled vector fields. Rather than optimizing over the discontinuous system directly, we instead formulate optimal control problems over a family of regularizations which are obtained by "smoothing out" the discontinuity in the original system using tools from singular perturbation theory. Standard, efficient derivative-based algorithms are immediately applicable to solve these smooth approximations to local optimally. Under standard regularity conditions, it is demonstrated that the smooth approximations provide accurate derivative information about the non-smooth problem in the limiting case. The utility of the technique is demonstrated in an in-depth example, where we utilize recently developed reduced-order modeling techniques from the dynamic walking community to generate motion plans across contact sequences for a 18-DOF model of a lower-body exoskeleton.

## I. INTRODUCTION

Non-smooth dynamical systems naturally arise when modelling a vast array of engineering systems, and are familiar to researchers in areas ranging from the dynamic walking community [1], [2] to the domain of power systems analysis [3]. However, despite the pervasive nature of discontinuous dynamics in systems theory and control, optimizing system trajectories through unplanned sequences of discontinuities remains a distinct technical challenge.

In particular, non-smooth systems are well known to display non-differentiability with respect to initial conditions and inputs [4, Chapter 2, Section 11]. This fundamental departure from smooth control theory has meant that the powerful optimization-based techniques designed for smooth control systems [5] cannot be readily applied to the non-smooth setting. As an example, in the context of optimal control, this has traditionally led to the implementation of restrictive practices, such as forming large mixed-integer programs to reason over all possible "mode" sequences the trajectories may undergo [6], or limiting the class of system behaviors by fixing the number and sequence of discontinuities the trajectory of the system encounters *a priori* [7], [8].

In this work, we formulate an optimal control problem over a class of piecewise-smooth vector fields. Rather than reasoning through the discontinuities of the system directly, we instead study a family of regularizations which are obtained by "smoothing" the original non-smooth system along the discontinuity. We then formulate approximate optimal control problems over these smooth control systems which can be solved using standard derivative-based techniques, without pre-specifying the number of times the discontinuity needs to be crossed. Theoretical guarantees are discussed, demonstrating when the regularized problem can be used to obtain accurate derivative information about the non-smooth problem. This result enables the use of mature iterative schemes to approach minimizers of the discontinuous problem using a sequence of the regularizations [5]. Finally, we use the smoothing approach to generate trajectories for a model of a bipedal robot, specifically a powered exoskeleton, which intermittently makes contact with the ground. In particular, we build on [9], [10], which introduced a class of embedded reduced-order modeling techniques to simplify the complexity of the motion planning problem.

### A. Relation to the Literature

The roots of the smoothing technique considered in this paper can be traced back to the literature on *geometric singular perturbation theory* (see e.g. [11] or [12]). The precise smoothing approach considered in this paper was first considered for autonomous piecewiese-smooth systems in work pioneered by Teixeira and others [13], [14], [15]. Emerging applications of these techniques in the context of control can be found in [16], [17], [18] and the works that follow them. A growing number of authors have studied variational problems over this class of regularization's [19], [17], [16]. Our contribution is distinct from the above literature in that we formally prove convergence between variations on the discontinuous and regularized problems in the case when piecewise-continuous controls are applied to the system. This is a key step for analyzing optimization problems formulated over the regularization's, as solutions returned by numerical optimal control schemes generally return discontinuous input signals [5].

Finally we mention a recent contribution [20], which builds on the geometric theory of hybrid systems [21],[22],[23], in which the smoothing techniques discussed in this paper were extended to a class of hybrid dynamical systems where the trajectories of the system undergo autonomous "jumps". We hope to extend the results presented here to this more general setting in a forthcoming article.

### B. Notation

We now fix notation used throughout the document. Unless otherwise noted, the 2-norm is our choice of norm for finite dimensional spaces. We say that a function $f\colon \mathbb{R}^n \to \mathbb{R}^p$ is

1 Dept of EECS., University of California, Berkeley {westenbroekt, sastry}@berkeley.edu
2 Dept. of Mechanical and Civil Engineering, California Institute of Technology {xxiong, ames}@caltech.edu

*Lipschitz continuous* if there exists $L > 0$ such that for each $x_1, x_2 \in R^n$ we have $||f(x_1) - f(x_2)|| \le L||x_1 - x_2||$. Given a normed space $V$, $\mathcal{P}(V)$ denotes the set of all subsets of $V$. Given a subset $S \subset V$, $\overline{co}S$ denotes the convex closure of $S$ in V.

Throughout the paper all of the control systems we consider will have states belonging to $\mathbb{R}^n$ and admissible inputs taking values in $\mathbb{R}^m$, where $n$ and $m$ are positive integers. Moreover, we will assume that the admissible set of initial conditions for these control systems belong to a set $D \subset \mathbb{R}^n$, and that the control supplied to these systems takes on values in the set $U \subset \mathbb{R}^m$. Both $D$ and $U$ are assumed to be compact, connected and convex. Fixing a finite time horizon $T > 0$, we let $PC([0,T],U)$ denote the set of piece-wise continuous functions from $[0,T]$ into $U$. Recall that piece-wise continuous functions have a finite number of points of discontinuity on a bounded domain. When considering optimization problems for the discontinuous system and our smooth regularizations, we will optimize over the set

$$\mathcal{X} = D \times PC([0,T],U).$$

We define the space of tangent vectors on $\mathcal{X}$ by

$$\mathcal{X}' = \mathbb{R}^n \times PC([0,T],R^m).$$

Finally, we endow both $\mathcal{X}$ and $\mathcal{X}'$ with the norm $||\cdot||: \mathbb{R}^n \times PC([0,T],R^m) \to \mathbb{R}$ defined for each $\xi = (x_0, u) \in \mathbb{R}^n \times PC([0,T],R^m)$ by

$$||\xi|| = ||x_0||_2 + ||u||_2.$$

We emphasize that $PC([0,T],\mathbb{R}^m) \subset L^2([0,T],\mathbb{R}^m)$ to motivate our choice of norm. Note that our chosen norm induces an obvious inner product.

## II. PIECEWISE SMOOTH CONTROL SYSTEMS

We now introduce the class of bimodal piecewise-smooth control systems considered in this paper, briefly review Filippov's convention for defining the dynamics of this class of systems, and introduce the smooth approximations we study throughout the paper.

### A. Bimodal Piecewise-smooth Control Systems

To begin, let $g: \mathbb{R}^n \to \mathbb{R}$ be a regular, smooth map with Lipschitz continuous first and second partial derivatives. Consider the two disjoint domains

$$D_1 = \{x \in \mathbb{R}^n : g(x) < 0\}$$

and

$$D_2 = \{x \in \mathbb{R}^n : g(x) > 0\},$$

which are separated by the co-dimension-1 sub-manifold

$$\Sigma = \{x \in \mathbb{R}^n : g(x) = 0\}.$$

Let $f_1, f_2: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be smooth vector fields with Lipschitz continuous first and second partial derivatives, and then consider the piecewise-smooth control system:

$$\dot{x} = f(x,u) = \begin{cases} f_1(x,u) & \text{if } x \in D_1 \\ f_2(x,u) & \text{if } x \in D_2 \end{cases} \quad (1)$$
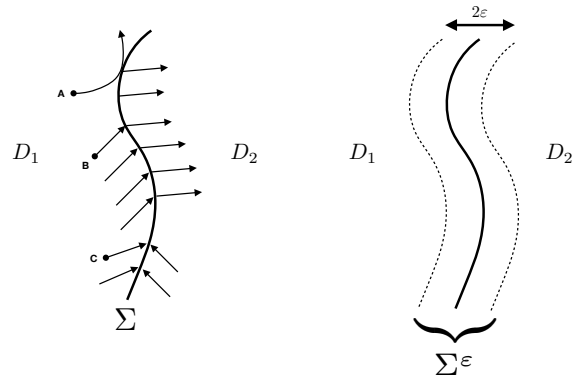


Fig. 1: (left) Trajectories of the discontinuous system with the input fixed. The trajectories starting from points $B$ and $C$ satisfy Assumption 1 and are thus unique. There are multiple solutions starting from point $A$, which passes through a point not satisfying the Assumption. (right) A depiction of the region of regularization.

Note that $f$ is undefined along $\Sigma$, and in general may also be discontinuous along this surface. Due to the discontinuity in $f$, classical (or $Carath\acute{e}odory$) solutions for the differential equation (1) may fail to exists along the surface of discontinuity. Thus, we turn to Filippov's convention to define the dynamics of the system in the following section.

For economy of notation we define the *Lie Derivatives* $L_{f_1}g, L_{f_2}g: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$L_{f_1}g(x,u) := \nabla g(x) \cdot f_1(x,u)$$

$$L_{f_2}g(x,u) := \nabla g(x) \cdot f_2(x,u).$$

### B. Filippov Solutions

We now review Filippov's convention, referring the interested reader to standard texts [4], [11], [24] for a more thorough introduction. Throughout the document we will make the following regularity assumption, which ensures the (right) existence and uniqueness of Filippov solutions for the discontinuous system (1) [4, Chapter 2, Section 10, Theorem 2].

**Assumption 1.** *There exists $\delta > 0$ such that for each point $(x,u) \in \Sigma \times U$ at least one of the following inequalities holds (with the understanding that the same inequality need not hold at every point):*

$$L_{f_1}g(x,u) > \delta \quad or \quad L_{f_2}g(x,u) < \delta \quad (2)$$

Roughly speaking, the assumption ensures that at least one of the vector fields $f_1, f_2$ points into the switching surface for each choice of $(x,u) \in \Sigma \times U$. The assumption rules out pathological cases where $L_{f_1}(x,u) < 0$ and $L_{f_2}(x,u) > 0$ when $x \in \Sigma$, in which case a solution passing through this point may continue into either $D_1$ or $D_2$.

More concretely, Filippov's convention defines the dynamics of the discontinuous system using the following

differential inclusion:

$$\dot{x} \in F(x,u) = \begin{cases} \{f_1(x,u)\} & \text{if } x \in D_1 \\ \overline{co}\{f_1(x,u), f_2(x,u)\} & \text{if } x \in \Sigma \\ \{f_2(x,u)\} & \text{if } x \in D_2, \end{cases} \quad (3)$$

where $F\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ is a multivalued map. Note that when $x \notin \Sigma$ $F(x,u)$ is a singleton taking on a value equal to that of $f(x,u)$. However, when $x \in \Sigma$, $F(x,u)$ may be multi-valued. For each data $\xi = (x_0, u) \in \mathcal{X}$, we say that the absolutely continuous $x^\xi\colon [0,T] \to \mathbb{R}^n$ is a *Filippov solution* corresponding to this data if $x^\xi(0) = x_0$ and it satisfies the differential inclusion (3) almost everywhere. With Assumption 1 in place, for each $t \in [0,T]$ we then define the map $\phi_t\colon \mathcal{X} \to \mathbb{R}^n$ by

$$\phi_t(\xi) = x^\xi(t). \quad (4)$$

A number of Filippov solutions for our chosen class of piecewise smooth control systems are depicted in Figure 1, in the case where a constant input $u \in U$ is applied to the system. In the Figure, solutions starting from the point $A$ pass through a point that violates Assumption 1 leading to non-uniqueness. Our analysis is restricted to cases such as those encountered by solutions starting from the points $B$ and $C$, wherein the trajectories either simply *cross* the surface of discontinuity or *slide* along the surface. We briefly outline conditions under which these two behaviors arise.

For a fixed $\xi \in \mathcal{X}$, suppose that $\phi_{t'}(\xi) \in \Sigma$ for some $t' \in [0,T]$. If $L_{f_1} g(\phi_t(\xi), u(t)) > 0$ and $L_{f_2} g(\phi_t(\xi), u(t)) > 0$ (or $L_{f_1} g(\phi_t(\xi), u(t)) < 0$ and $L_{f_2} g(\phi_t(\xi), u(t)) < 0$) for each $t$ in some open interval containing $t'$, then the solution simply crosses the surface of discontinuity. Note that in this case classical ($Carath\acute{e}odory$) solutions to the non-smooth differential equation (1) are well-defined.

On the other hand, in the case that $L_{f_1} g(\phi_t(\xi), u(t)) > 0$ and $L_{f_2} g(\phi_t(\xi), u(t)) < 0$ for each $t$ in some open interval of time containing $t'$, there will be a unique Fillippov solution which *slides* along the surface of discontinuity. This motion can be described by the differential equation

$$\dot{x} = f^s(x,u), \quad (5)$$

where Filippov's *sliding vector field* is defined by

$$f^s(x,u) = (1 - \alpha(x,u)) f_1(x,u) + \alpha(x,u) f_2(x,u)$$

where

$$\alpha(x,u) = \frac{\nabla g(x) \cdot f_1(x,u)}{\nabla g(x) \cdot (f_1(x,u) - f_2(x,u))}$$

selects the unique convex combination of $f_1$ and $f_2$ which keeps the trajectory on $\Sigma$.

We next introduce a family of smooth regularizations for the discontinuous vector field (1) which are used to approximate these dynamics.

### C. Smooth Approximations via Singlular Perturbations

The family of smooth approximations condisered in this paper are parameterized by $\varepsilon > 0$ and are obtained by smoothing (1) along the *region of regularization*

$$\Sigma^\varepsilon = \{x \in \mathbb{R}^n\colon -\varepsilon < g(x) < \varepsilon\}.$$

As depicted on the right of Figure 1, $\Sigma^\varepsilon$ is a $2\varepsilon$-thick strip centered around $\Sigma$. The main idea behind the smoothing approach is to use the following class of functions to transition between $f_1$ and $f_2$ along $\Sigma^\varepsilon$.

**Definition 1.** *We say that $\varphi \in C^\infty(\mathbb{R}, [0,1])$ is a transition function if 1) $\varphi(a) = 0$ if $a \leq -1$, 2) $\varphi(a) = 1$ if $a \geq 1$, 3) $\varphi$ is monotonically increasing on $(-1,1)$, and 4) each of the first and second partial derivatives of $\varphi$ are Lipschitz continuous.*

For the rest of the paper, we assume that a single transition function $\varphi$ has been selected. For each $\varepsilon > 0$ we then define the $\varepsilon$-relaxation of (1) to be

$$\dot{x} = f^\varepsilon(x,u) \quad (6)$$

where $f^\varepsilon\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is given by

$$f^\varepsilon(x,u) = \left(1 - \varphi\left(\frac{g(x)}{\varepsilon}\right)\right) f_1(x,u) + \varphi\left(\frac{g(x)}{\varepsilon}\right) f_2(x,u).$$

Note that $f^\varepsilon(x,u)$ is a convex combination of $f_1(x,u)$ and $f_2(x,u)$ if $x \in \Sigma^\varepsilon$, and that $f^\varepsilon(x,u) = f_1(x,u)$ if $x \in D_1 \setminus \Sigma^\varepsilon$ and $f^\varepsilon(x,u) = f_2(x,u)$ if $x \in D_2 \setminus \Sigma^\varepsilon$.

It is straightforward to see that for each $\varepsilon > 0$ the vector field $f^\varepsilon$ is smooth and Lipschitz continuous, and thus the trajectories corresponding to the smooth system will be unique. Thus for each $\varepsilon > 0$ and $\xi = (x_0, u) \in \mathcal{X}$ let $x^{(\varepsilon,\xi)}\colon [0,T] \to \mathbb{R}^n$ be the solution to (6) with initial condition $x^{(\xi,\varepsilon)}(0) = x_0$ and for each $t \in [0,T]$ define the flow $\phi_t^\varepsilon\colon \mathcal{X} \to \mathbb{R}^n$ by

$$\phi_t^\varepsilon(\xi) = x^{(\varepsilon,\xi)}(t).$$

The following result from [20] establishes that the trajectories of the relaxed system converge to the Filippov solutions of the discontinuous system as $\varepsilon \to 0$.

**Lemma 1.** *([20]) Let Assumption 1 hold for* (1). *Then there exists $C > 0$ and $\varepsilon_0 > 0$ such that for each $\varepsilon \leq \varepsilon_0$, $t \in [0,T]$ and $\xi = (x_0, u) \in \mathcal{X}$*

$$\|\phi_t(\xi) - \phi_t^\varepsilon(\xi)\| \leq C\varepsilon.$$

In other words, the solutions to the regularized system converge to the solutions of the non-smooth system at a rate that is linear in $\varepsilon$ under our standing regularity condition.

### III. OPTIMAL CONTROL PROBLEMS

In this section we formulate an optimal control problem over the piecewise-smooth system introduced in Section II and a family of approximate problems over the smooth regularizations introduced in Section II-C. We then discuss

the challenges associated with applying standard gradient-based algorithms to arrive at minimizers for these problems. While Assumption 1 is enough to ensure uniqueness of solutions for the discontinuous system, stronger assumptions will need to be put into place to ensure differentiablility with respect to initial conditions and inputs.

### A. Optimal Control Problems

To begin, let $\ell\colon \mathbb{R}^n \to \mathbb{R}$ be a cost function which is assumed to have Lipschitz continuous first and second partial derivatives, and define the cost functional $L\colon \mathcal{X} \to \mathbb{R}$ by

$$L(\xi) = \ell(\phi_T(\xi)).$$

We then define the following optimal control problem:

$$(\mathbf{P}) \quad \inf_{\xi \in \mathcal{X}} L(\xi)$$

Note that this problem formulation is quite general, since there exists well-known transformations to include additional terms, such as running costs, into the terminal cost functional we consider here [5, Chapter 5].

Next, for each $\varepsilon > 0$ we define the regularized cost functional $L^\varepsilon\colon \mathcal{X} \to \mathbb{R}$ by

$$L^\varepsilon(\xi) = \ell(\phi_T^\varepsilon(\xi)) \tag{7}$$

and subsequently define the $\varepsilon-$relaxation of $P$ by:

$$(\mathbf{P}^\varepsilon) \quad \inf_{\xi \in \mathcal{X}} .L^\varepsilon(\xi)$$

As a straightforward consequence of Lemma 1, we see that for each $\xi \in \mathcal{X}$

$$\lim_{\varepsilon \to 0} L^\varepsilon(\xi) = L(\xi). \tag{8}$$

While this condition is sufficient to ensure that global minimizers of $\mathbf{P}^\varepsilon$ converge to global minimizers of $\mathbf{P}$ as $\varepsilon \to 0$, it is generally not sufficient to say the same of local minimizers for the problems [5, Chapter 3]. Frameworks for analyzing this more difficult condition rely on variational analysis of the approximate and limiting problems. This motivates the utility of the convergence results presented in this article.

### B. Directional Derivatives

The majority of practical tools for finding local minimizers of optimal control problems rely on calculating the derivatives of the cost functional with respect to changes in the initial condition an inputs applied to the system. Once these derivatives are obtained, standard optimization procedures can be applied to iteratively improve performance using numerical discretization schemes [5].

*1) Directional Derivatives of Regularized Problems:* Since the vector field $f^\varepsilon$ is smooth for each $\varepsilon > 0$, we can compute directional derivatives of $L^\varepsilon$ in the usual way. In particular, given $\xi = (x_0, u) \in \mathcal{X}$, we define $DL^\varepsilon(\xi; \cdot)\colon \mathcal{X}' \to \mathbb{R}$ by

$$DL^\varepsilon(\xi; \delta\xi) = \lim_{\lambda \downarrow 0} \frac{L^\varepsilon(\xi + \lambda \cdot \delta\xi) - L^\varepsilon(\xi)}{\lambda} \tag{9}$$

for each direction $\delta\xi = (\delta x_0, \delta u) \in \mathcal{X}'$, which by [5, Theorem 5.6.8] is given by

$$DL^\varepsilon(\xi; \delta\xi) = p(0)^T \cdot \delta x_0 + \int_0^T p(\tau)^T B(\tau) \delta u(\tau) d\tau$$

where $p\colon [0,T] \to \mathbb{R}^n$ is the solution to the *co-state* or *adjoint* equations

$$-\dot{p}(t) = A(t)^T p(t) \text{ a.e. } t \in [0,T], \; p(T)^T = \frac{\mathrm{d}}{\mathrm{d}x} l(\phi_T^\varepsilon(\xi))$$

and for each $t \in [0,T]$ we have

$$A(t) = \frac{\partial}{\partial x} f^\varepsilon(\phi_t^\varepsilon(\xi), u(t))$$

$$B(t) = \frac{\partial}{\partial u} f^\varepsilon(\phi_t^\varepsilon(\xi), u(t))$$

In particular, the directional derivative $DL^\varepsilon(\xi; \delta\xi)$ provides a first-order approximation to how $L^\varepsilon$ changes by perturbing the data supplied to the system in the direction $\delta\xi$.

However, inspecting $f^\varepsilon$, we see that the partial derivative $\frac{\partial}{\partial x} f^\varepsilon(x, u)$ will be of the order $\frac{1}{\varepsilon}$ when $x \in \Sigma^\varepsilon$. Thus, we should be concerned that the adjoint equations associated to the regularized system will "blow up" as we take $\varepsilon \to 0$ along trajectories which pass through the region of regularization. In other words, in the numerical setting we should be concerned that discretizations of $P^\varepsilon$ will become ill-conditioned for small values of the regularization parameter. However, our subsequent analysis will demonstrate that this is not the case, and that the gradients of $P^\varepsilon$ remain bounded even as we take $\varepsilon \to 0$. Further discussion on numerical considerations can be found in [19] and [25].

*2) Directional Derivatives of Non-smooth Problem:* It is well-known that Filippov solutions for the discontinuous system (1) may have a non-smooth dependence on the initial conditions and inputs supplied to the system [4, Chapter 2, Section 11]. In particular, the map $\phi_t(\cdot)$ may fail to be differentiable at the point $\xi \in \mathcal{X}$ if the trajectory associated to this data arrives at the surface of discontinuity at time $t$. In the context of optimal control, this means that the adjoint equations associated to the discontinuous system may under go discrete "jumps" at time instances when the nominal trajectory reaches the surface of discontinuity[1]. Due to these discontinuities, the adjoint equations, and consequently the gradients of $L$, are difficult to approximate numerically using standard integration techniques such as Euler integration [19]. As discussed in Section I-A, this has typically meant that derivative-based methods for solving optimal control problems over discontinuous systems have required restrictive practices such as fixing the number of times the trajectory crosses the surface of discontinuity *a priori*. This is our primary motivation for employing the smooth approximations considered in the paper.

---

[1] The "jump conditions" associated with the adjoint equations of the discontinuous system have a rich geometric interpretation which we do not have space to discuss in this article. However, the interested reader is referred to [4, Chapter 2, Section 11]) for a comprehensive discussion regarding these "jump conditions" in the context of sensitivity analysis.

Before proving our convergence results in Section IV, we first formalize conditions which will ensure the differentiability of $L$ at a nominal choice of $\xi = (x_0, u) \in \mathcal{X}$.

**Assumption 2.** *The nominal data $\xi = (x_0, u) \in \mathcal{X}$ is such that $x_0 \notin \Sigma$.*

When the nominal Filippov solution begins on the surface of discontinuity, solutions may fail to be differentiable with respect to the supplied initial condition at this point since small perturbations to $x_0$ can cause nearby trajectories to start in either $D_1$ or $D_2$, with some of these solutions crossing the surface of discontinuity while others do not. This assumption can be relaxed if the nominal solution undergoes a stable sliding motion on some open interval $[0, t_1) \subset [0, T]$, since all nearby trajectories will reach the surface of discontinuity. This point is discussed in more detail in [26].

Our next assumption excludes pathological cases where the nominal solution arrives at the surface of discontinuity an infinite number of times. To formalize this notion, for each $\xi \in \mathcal{X}$ we let $\mathcal{T}^\xi$ denote the set of time instances at which the corresponding Filippov solution arrives at the surface of discontinuity:

$$\mathcal{T}^\xi = \{t' \in (0, T] : \phi_{t'}(\xi) \in \Sigma \text{ and } \exists \delta > 0 \text{ s.t.}$$
$$\phi_t(\xi) \notin \Sigma, \; \forall t \in (t' - \delta, t')\}$$

**Assumption 3.** *The nominal data $\xi = (x_0, u) \in \mathcal{X}$ is such that $\mathcal{T}^\xi$ is a finite set.*

We also assume that the nominal trajectory does not arrive at the surface of discontinuity exactly at time $T$. As discussed above, this may lead $\phi_T$ to be non-differentiable at $\xi$, which in turn may cause $L$ to be non-differentiable.

**Assumption 4.** *The nominal data is such that $T \notin \mathcal{T}^\xi$.*

Our final assumption ensures that the nominal trajectory arrives at the surface of discontinuity transversely, without "skimming" the surface, a situation which is well known to cause non-differentiability [4, Chapter 2, Section 11].

**Assumption 5.** *The data $\xi = (x_0, u) \in \mathcal{X}$ is such that if $\hat{t} \in \mathcal{T}^\xi$ then $\exists \gamma > 0$ and interval $I = (\hat{t} - \gamma, \hat{t} + \gamma) \cap [0, T]$ such that*
1) $\phi_{\hat{t}^-}(\xi) \in D_1 \implies L_{f_1} g(\phi_t(\xi), u(t)) > 0, \; \forall t \in I$
2) $\phi_{\hat{t}^-}(\xi) \in D_2 \implies L_{f_2} g(\phi_t(\xi), u(t)) < 0, \; \forall t \in I$

When the data $\xi \in \mathcal{X}$ satisfies the above assumptions the map $\phi_T$ will be differentiable at $\xi$, and thus the directional derivatives of $L$ will be well-defined at this point. In particular we define $DL(\xi; \cdot) : \mathcal{X}' \to \mathbb{R}$ by

$$DL(\xi; \delta\xi) = \lim_{\lambda \downarrow 0} \frac{L(\xi + \lambda \cdot \delta\xi) - L(\xi)}{\lambda}, \qquad (10)$$

for each direction $\delta\xi \in \mathcal{X}'$. When the above limit exists (i.e. when the above assumptions hold), $DL(\xi; \delta\xi)$ provides a first order approximation to how $L$ will change when perturbed in the direction $\delta\xi$.

## IV. Convergence Results

We are now ready to present the theoretical contribution of this work. As noted in Section III-B.1, we first need to establish that the derivatives of our regularized problems do not "blow up" as we take $\varepsilon \to 0$. The following intermediate result ensures that our regularized problems remain well-conditioned even for small values of the regularization parameter. Proof of the following Lemma and subsequent Theorem can be found in [26], where a number of other useful regularity properties of the regularized optimization problem are derived.

**Lemma 2.** *Let Assumption 1 hold for the discontinuous system* (1). *Then there exists $C > 0$ and $\varepsilon_0 > 0$ such that for each $\delta\xi \in \mathcal{X}'$ and $\varepsilon \le \varepsilon_0$*

$$\|DL^\varepsilon(\xi; \delta\xi)\| \le C\|\delta\xi\| \qquad (11)$$

The result is proven by demonstrating that the co-state equations associated to the nominal data remain bounded even when passing through the region of regularization. In cases where the nominal trajectory simply crosses the surface of discontinuity (without sliding) this is unsurprising since, for $\varepsilon$ sufficiently small, the regularized solution will take on the order of $\varepsilon$ units of time to cross the region of regularization, effectively "cancelling out" the effect of $\left\| \frac{\partial}{\partial x} f^\varepsilon(x, u) \right\|$ being on the order of $\frac{1}{\varepsilon}$ in this region. The case in which the nominal trajectory of the discontinuous system slides along the surface of discontinuity must be approached with more care. Nonetheless, our analysis uses a decomposition similar to the technique presented in [16] to demonstrate that the co-state equations are stable in regions where sliding occurs for the discontinuous system.

Having established that the derivatives of our regularized problems remain well defined even as $\varepsilon \to 0$, we now proceed to our main result:

**Theorem 1.** *Let Assumption 1 hold. Let $\xi = (x_0, u) \in \mathcal{X}$ satisfy Assumptions 2-5. Then for each $\delta\xi = (\delta x_0, \delta u) \in \mathcal{X}'$*

$$\lim_{\varepsilon \to 0} DL^\varepsilon(\xi; \delta\xi) = DL(\xi; \delta\xi)$$

In practice, solutions to difficult optimization problems are often approximated by solving a *sequence* of approximate problems. Theorem 1 is useful because it demonstrates that, when Assumptions 1- 5 hold, a descent direction calculated for $\mathbf{P}^\varepsilon$ is also a descent direction for $\mathbf{P}$, when $\varepsilon$ is sufficiently small. This enables the use of mature iterative schemes [5, Chapter 3] to arrive at a local minima for $\mathbf{P}$ using a sequence of regularized problems $\{\mathbf{P}^{\varepsilon_i}\}_{i=1}^\infty$ which are such that $\varepsilon_i \to 0$ as $i \to \infty$.

Of course, when using the standard derivative based techniques discussed in Section III-B.1 to solve a regularized problem (or a sequence of regularized problems) to local optimality we should not expect Assumptions 2 and 4 to hold at every iterate. Indeed, our stated goal was to develop a technique which allows us to freely optimize trajectories across the surface of discontinuity. Nevertheless, so long as an iterative procedure which solves a sequence of problems

$\{P^{\varepsilon_i}\}_{i \in \mathbb{N}}$ eventually produces iterates which all lie in a region of $\mathcal{X}$ which satisfies the Assumptions of Theorem 1, the above approach can be used to analyze convergence of the algorithm to stationary points of $\mathbf{P}$. A more thorough and formal discussion of this point can be found in [26].

## V. TRAJECTORY OPTIMIZATION ON A REDUCED-ORDER MODEL

To demonstrate the utility of the smoothing technique we have investigated, we first use the approach to generate hopping trajectories for an actuated Spring-mass hopper, which was first proposed in [9]. Its nonlinear leg spring model has shown great value for robotic applications such as under-actuated bipedal walking [1] and fully-actuated humanoid walking [10]. We note, however, that the trajectory generation procedures proposed in these prior works required fixing the sequence of contacts the robot makes with the ground *a priori*. Here, we demonstrate that the regularization techniques we consider can be used to effectively generate optimal trajectories on the Spring-mass model without pre-specifying the contact sequence ahead of time.

### A. Actuated Spring-mass Model of Hopping

The actuated Spring-mass hopper is depicted in Fig. 2 (a). The model has four states: the height of the mass, $z$, and its time derivative $\dot{z}$, as well as the natural length of the leg spring, $L$, and its time derivative, $\dot{L}$. The model is actuated by controlling $L$, i.e.,

$$\ddot{L} = u, \tag{12}$$

where $u$ the actuation on the leg. The *Flight* phase is ballistic, thus the dynamics of the mass is,

$$\ddot{z} = -g \tag{13}$$

where $g$ is the gravitational constant. When the hopper is on the ground, the dynamics of the mass is,

$$\ddot{z} = \frac{F(z, \dot{z}, L, \dot{L})}{m} - g \tag{14}$$

where $m$ is the value of the mass and $F(z, \dot{z}, L, \dot{L}) = K(L)(L - z) + D(L)(\dot{L} - z)$ is the leg force from the spring deflection. Here, $K(L)$ and $D(L)$ are the stiffness and damping coefficient, respectively. Further details of the model can be found in [9]. In practice $K(L)$ is much larger than $D(L)$, thus we make the simplifying assumption that the robot lifts off when $L = z$. The validity of this assumption follows from the close tracking between the reduced and full-order models exhibited in Section VI.

Collecting the states of the robot as $x = (z, \dot{z}, L, \dot{L})$, we represent the dynamics using the piecewise-smooth vector field $f : \mathbb{R}^4 \times \mathbb{R} \to \mathbb{R}^4$ where,

$$\dot{x} = f(x, u) = \begin{cases} f_{\mathrm{F}}(x, u) & \text{if } g(x) > 0 \\ f_{\mathrm{G}}(x, u) & \text{if } g(x) < 0, \end{cases} \tag{15}$$
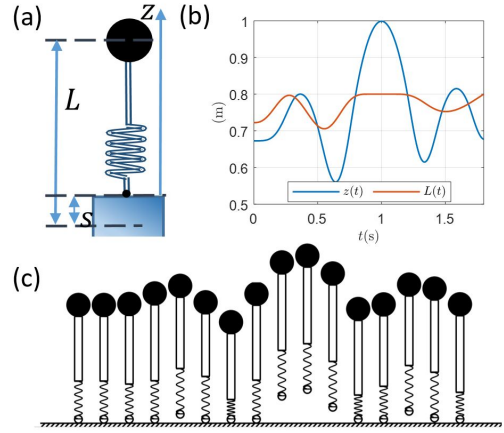


Fig. 2: (a) The actuated Spring-mass model. (b) The optimized hopping trajectories. (c) Visualization of the hopping.

where,

$$f_{\mathrm{F}}(x, u) = (\dot{z}, -g, \dot{L}, u)^T \tag{16}$$

$$f_{\mathrm{G}}(x, u) = (\dot{z}, F(x, \dot{z}, L, \dot{L})/m - g, \dot{L}, u)^T \tag{17}$$

$$g(x) = z - L. \tag{18}$$

### B. Trajectory Optimization Via Smoothing

Now we apply the proposed method by optimizing a hopping motion without specifying the contact modes. The task is to reach an apex height $z_{\mathrm{apex}} = 1\mathrm{m}$ at $t_{\mathrm{apex}} = 1\mathrm{s}$ from a static standing configuration, and then settle back to its original height at $t_{\mathrm{f}} = 1.8\mathrm{s}$. The initial condition is set as $x_0 = (z(0), \dot{z}(0), L(0), \dot{L}(0))^T = (.65, 0, .75, 0)^T$. The input is bounded, i.e., $u \in U = [-10, 10]$. The cost function is

$$J(x) = (z(t_{\mathrm{apex}}) - z_{\mathrm{apex}})^2 + \dot{z}(t_{\mathrm{apex}})^2 +$$
$$(z(t_{\mathrm{f}}) - z(0))^2 + \dot{z}(t_{\mathrm{f}})^2 + \int u^2 dt.$$

To approximate a solution to the problem numerically, we use the regularization parameter $\varepsilon = .01$, and Euler integration with 200 uniformly spaced gridpoints. The resulting finite dimensional optimization problem is solved using the Matlab fmincon function. The optimized Spring-mass trajectory is shown in Fig. 2 (b). Due to the actuation limitations of the model, the optimized hopping actually requires two hops to reach to the apex height and one additional hop to settle.

## VI. HOPPING EMBEDDING ON THE EXOSKELETON

Despite the simplicity of the Spring-mass hopper, the optimized dynamics can be embedded onto complex robotic systems. Here we briefly describe the process of embedding the hopping dynamics of the Spring-mass onto the exoskeleton (Fig. 3 (a)), using the trajectory of the spring-mass hopper to define outputs to track on the full-order model. A QP-CLF optimization based controller is implemented to track the desired outputs online. More examples of dynamics embedding from simple models to full robot models can be found in [9] [1] [10] and the references therein.

## A. Robot Model and Hybrid Dynamical Model

The exoskeleton (Exo) has two legs with 12 motor joints in total. Its equations of motion can be described by the floating-base Euler-Lagrange equation,

$$M\ddot{q} + h = Bu + J_v^T F_v, \quad (19)$$

$$J_v\ddot{q} + \dot{J}_v\dot{q} = 0, \quad (20)$$

where $q \in SE(3) \times \mathbb{R}^{n=12}$. The exact meaning of each term can be found in [9]. For the hopping behavior, we assume the two feet always have same contact, i.e. either both feet are or are not in contact with the ground. Thus the hopping is an alternation between two domains, i.e. *Ground* and *Flight* domain. We use subscript $v$ to denote different domains. The guards are defined from the transition conditions, i.e. ground normal forces→0 and foot position→0. We model the impact as plastic impact from *Flight→Ground*, where the joint velocities have jumps, i.e., $\dot{q}_{\text{Ground}}^+ = \Delta(q)\dot{q}_{\text{Flight}}^-$.

## B. Output Selection and CLF-QP based Force Control.

For *Ground*, the robot is fully actuated and 6 outputs are required to be defined. We first select the center of mass (COM) position and the yaw and roll angles of the pelvis/torso of the robot as relative degree 2 outputs, i.e.,

$$\mathcal{Y}_2^{\text{Ground}}(q,t) = \begin{bmatrix} p_{\text{COM}^x}(q) \\ p_{\text{COM}^y}(q) \\ p_{\text{COM}^z}(q) \\ \varphi_{\text{pelvis}}(q) \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ z^{\text{opt}}(t) \\ \mathbf{0} \end{bmatrix}, \quad (21)$$

where $z^{\text{opt}}(t)$ is the mass trajectory generated by the trajectory optimization in the previous section. It is also necessary to have zero centrodial angular momentum before *Flight* phase [9]. We select the last output as,

$$\mathcal{Y}_1^{\text{Ground}}(q,t) = H_{\text{pitch}}(q,\dot{q}) - 0, \quad (22)$$

where $H_{\text{pitch}}(q,\dot{q})$ is the pitch centrodial momentum.

In *Flight*, the robot is underactuated. Here, 12 outputs are required to be defined since $n_u = 12$. We selected the outputs as the positions of the feet to the COM and foot orientations. The desired vertical positions between the COM and the feet are the real leg length $L - z$ of the Spring-mass hopper,

$$\mathcal{Y}_2^{\text{Flight}}(q,t) = \begin{bmatrix} \mathbf{p}_{\text{Feet}\to\text{COM}}^x(q) \\ \mathbf{p}_{\text{Feet}\to\text{COM}}^y(q) \\ \mathbf{p}_{\text{Feet}\to\text{COM}}^z(q) \\ \varphi_{\text{Feet}}(q) \end{bmatrix} - \begin{bmatrix} \mathbf{0} \\ \mathbf{c} \\ L^{\text{opt}}(t) - z^{\text{opt}}(t) \\ \mathbf{0} \end{bmatrix}, \quad (23)$$

where $\mathbf{c}$ is a constant vector and its value depends on the initial state of the flight phase.

We employ control Lyapunov function based Quadratic programs (CLF-QP) [27] [1] for feedback-zeroing the outputs. Exponential convergence on the output dynamics is enforced by an inequality condition on the constructed Lyapunov function, i.e.,

$$\dot{V}(u,q,\dot{q}) \leq -\gamma V(q,\dot{q}), \quad (24)$$

with $\gamma > 0$. This inequality is affine with respect to $u$,

$$A_v^{\text{CLF}}(q,\dot{q})u \leq b_v^{\text{CLF}}(q,\dot{q}). \quad (25)$$

External holonomic constraints, such as the holonomic forces $F_v$, are affine with respect to $u$ and can be collected as

$$F_v = A_v u + b_v, \quad (26)$$

where $A_v, b_v$ can be found in [10]. Thus contact constraints such as friction cones and non-negative normal forces can be formulated as an inequality constraints on $u$,

$$C_v A_v u \leq -C_v b_v, \quad (27)$$

where $C_v$ is a constant matrix. It is also desirable to track the ground reaction forces experienced by the reduced model on the full-order order robot so that the two systems lift off from the ground at approximately the same time. This desired constraint is of the form

$$S_v A_v u = F^{\text{opt}} - S_v b_v, \quad (28)$$

where $S_v$ is a selection matrix to extract the vertical normal forces from the holonomic force vector and $F^{\text{opt}}$ is the leg force from the Spring-mass optimization (Fig. 3 (b)). It is necessary to relax the force control since the vertical COM position is one of the outputs for control [10]. Thus the equality in Eq. (28) becomes,

$$\underbrace{(1-c)F^{\text{opt}} - S_v b_v}_{c_{lb}} \leq S_v A_v u \leq \underbrace{(1+c)F^{\text{opt}} - S_v b_v}_{c_{ub}}, \quad (29)$$

where $c \in (0,1)$ is a coefficient of relaxation.

The main control law for the full-order model in each domain is thus formulated as a quadratic program as follows:

$$u^* = \operatorname*{argmin}_{u \in \mathbb{R}^{12}, \delta \in \mathbb{R}} u^T H u + 2Fu + p\delta^2, \quad (30)$$

$$\text{s.t.} \quad A_v^{\text{CLF}}(q,\dot{q})u \leq b_v^{\text{CLF}}(q,\dot{q}) + \delta, \quad \text{(CLF)}$$

$$C_v A_v u \leq -C_v b_v, \quad \text{(Contact)}$$

$$u_{lb} \leq u \leq u_{ub}, \quad \text{(Torque Limit)}$$

$$c_{lb} \leq S_v A_v u \leq c_{ub}, \quad \text{(Force Control)}$$

where $\delta$ is a relaxation term for increasing the instantaneous feasibility of the QP, and $p$ is a positive penalty constant.

## C. Simulation Result

Fig. 3 (c)(d) shows the resulting hopping motion of the exo when the CLF-QP controller is used to track the nominal outputs generated by the spring-mass model. The tracking controller is able to closely track the reference generated by the active SLIP model as the exo repeatedly makes and breaks contact with the ground.

## VII. CONCLUSION

This paper investigated a method for smoothing out a discontinuous differential equation. We studied optimal control problems over the discontinuous system and its regularizations, derived a number of useful properties for these problems and examined conditions under which the regularizations can be used to obtain accurate derivative information about the non-smooth problem. Finally, we demonstrated the efficacy of the smoothing approach by using it in conjunction
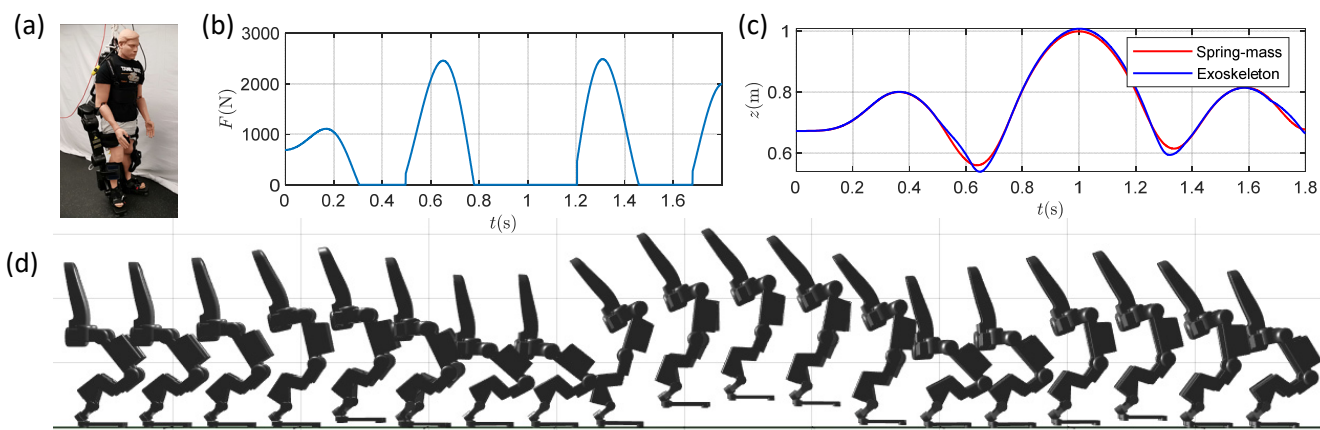
Fig. 3: (a) The Exoskeleton. (b) The optimized ground force of the hopping. (c) Comparison of the COM trajectories between the Spring-mass and the Exo. (d) The snapshots of the realized hopping.

with recent reduced-order modelling techniques to generate non-trivial motion plans on a 18-DOF exoskeleton robot in simulation. In the near future we hope to extend the results of this paper to the case of fully hybrid dynamics, building on the work of [20].

## VIII. Acknowledgement

## References

[1] X. Xiong and A. D. Ames, "Coupling reduced order models via feedback control for 3d underactuated bipedal robotic walking," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–9.

[2] E. R. Westervelt, J. W. Grizzle, and D. E. Koditschek, "Hybrid zero dynamics of planar biped walkers," *IEEE transactions on automatic control*, vol. 48, no. 1, pp. 42–56, 2003.

[3] I. A. Hiskens and M. Pai, "Trajectory sensitivity analysis of hybrid systems," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, no. 2, pp. 204–220, 2000.

[4] A. F. Filippov, *Differential equations with discontinuous righthand sides: control systems*. Springer Science & Business Media, 2013, vol. 18.

[5] E. Polak, *Optimization: algorithms and consistent approximations*. Springer Science & Business Media, 2012, vol. 124.

[6] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[7] M. Kamgarpour and C. Tomlin, "On optimal control of non-autonomous switched systems with a fixed mode sequence," *Automatica*, vol. 48, no. 6, pp. 1177–1181, 2012.

[8] M. S. Shaikh and P. E. Caines, "On the hybrid optimal control problem: theory and algorithms," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1587–1603, 2007.

[9] X. Xiong and A. D. Ames, "Bipedal hopping: Reduced-order model embedding via optimization-based control," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3821–3828.

[10] ——, "Motion decoupling and composition via reduced order model optimization for dynamic humanoid walking with clf-qp based active force control," in *Submitted to 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

[11] S. Sastry, *Nonlinear systems: analysis, stability, and control*. Springer Science & Business Media, 2013, vol. 10.

[12] H. K. Khalil, "Noninear systems."

[13] J. Sotomayor, "Regularization of discontinuous vector fields." World Scientific, 1996.

[14] J. Llibre, P. R. da Silva, M. A. Teixeira *et al.*, "Sliding vector fields via slow–fast systems," *Bulletin of the Belgian Mathematical Society-Simon Stevin*, vol. 15, no. 5, pp. 851–869, 2008.

[15] J. Llibre and M. A. Teixeira, "Regularization of discontinuous vector fields in dimension three," *Discrete & Continuous Dynamical Systems-A*, vol. 3, no. 2, pp. 235–241, 1997.

[16] D. Fiore, M. Coraggio, and M. di Bernardo, "Observer design for piecewise smooth and switched systems via contraction theory," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2959–2964, 2017.

[17] D. Fiore, S. J. Hogan, and M. di Bernardo, "Contraction analysis of switched systems via regularization," *Automatica*, vol. 73, pp. 279–288, 2016.

[18] D. Angulo-Garcia, F. Angulo, G. Osorio, and G. Olivar, "Control of a dc-dc buck converter through contraction techniques," *Energies*, vol. 11, no. 11, p. 3086, 2018.

[19] D. E. Stewart and M. Anitescu, "Optimal control of systems with discontinuous differential equations," *Numerische Mathematik*, vol. 114, no. 4, pp. 653–695, 2010.

[20] T. Westenbroek, H. Gonzalez, and S. S. Sastry, "A new solution concept and family of relaxations for hybrid dynamical systems," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 743–750.

[21] S. N. Simic, K. H. Johansson, J. Lygeros, and S. Sastry, "Towards a geometric theory of hybrid systems," *Dynamics of Continuous, Discrete and Impulsive Systems Series B: Applications and Algorithms*, vol. 12, no. 5-6, pp. 649–687, 2005.

[22] S. A. Burden, H. Gonzalez, R. Vasudevan, R. Bajcsy, and S. S. Sastry, "Metrization and simulation of controlled hybrid systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2307–2320, 2015.

[23] A. D. Ames and S. Sastry, "A homology theory for hybrid systems: Hybrid homology," in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2005, pp. 86–102.

[24] J.-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory*. Springer Science & Business Media, 2012, vol. 264.

[25] T. Westenbroek, X. Xiong, A. D. Ames, and S. S. Sastry, "Technical report: Optimal control of piece-wise smooth control systems via singular perturbations," *Arxiv Preprint*.

[26] T. Westenbroek and S. S. Sastry, "A new solution concept and family of relaxations for hybrid dynamical systems (extended edition)," Arxiv Preprint., 2017.

[27] A. D. Ames and M. Powell, "Towards the unification of locomotion and manipulation through control lyapunov functions and quadratic programs," in *Control of Cyber-Physical Systems*. Springer, 2013, pp. 219–240.