

Certified Gradient-Based Contact-Rich Manipulation via Smoothing-Error Reachable Tubes

Wei-Chen Li and Glen Chou

Georgia Institute of Technology, Atlanta, Georgia 30308

Email: {wli777, chou}@gatech.edu

Project Website: (Link) | Code: (Github) | Video: (YouTube)

Abstract—Gradient-based methods can efficiently optimize controllers by leveraging differentiable simulation and physical priors. However, contact-rich manipulation remains challenging because hybrid contact dynamics often produce discontinuous or vanishing gradients. Although smoothing the dynamics can restore informative gradients, the resulting model mismatch can cause controller failures when deployed on real systems. We address this trade-off by planning with smoothed dynamics while explicitly quantifying and compensating for the induced error, providing formal guarantees on safety and task completion under the original nonsmooth dynamics. Our approach applies smoothing to both contact dynamics and contact geometry within a differentiable simulator based on convex optimization, allowing us to characterize the deviation from the nonsmooth dynamics as a set-valued discrepancy. We incorporate this discrepancy into the optimization of time-varying affine feedback policies through analytical reachable sets, enabling robust constraint satisfaction for the closed-loop hybrid system while relying solely on the informative gradients of the smoothed model. By bridging differentiable simulation with set-valued robust control, our method produces affine feedback policies that respect the unilateral nature of contact. We evaluate our method on several contact-rich tasks, including planar pushing, object rotation, and in-hand dexterous manipulation, achieving certified constraint satisfaction with lower safety violations and smaller goal errors than baseline approaches.

I. INTRODUCTION

Contact-rich manipulation involves controlling unactuated objects to reach a desired configuration, which requires making discrete decisions about when and how to make or break contact. Among current approaches, reinforcement learning (RL) [51] and zero-order planning methods such as cross-entropy method (CEM) [18, 31], covariance matrix adaptation evolution strategy (CMA-ES) [16, 19], and model predictive path integral (MPPI) control [45] have achieved notable success. However, these methods are fundamentally sampling-based and effectively require rediscovering the underlying physics through trial-and-error, resulting in low sample efficiency.

In contrast, gradient-based trajectory optimization [6, 50, 35] can exploit the problem’s physical structure via differentiable simulation [41, 40]. However, hybrid dynamics in contact-rich manipulation violate key assumptions of these methods. In particular, nonsmooth hybrid dynamics create fundamental challenges: gradients are often discontinuous across contact modes and frequently zero. For example, before the robot contacts the object, its actions do not affect the object or reward. Since the contact manifold occupies a zero-measure subset of the configuration space, standard gradients provide

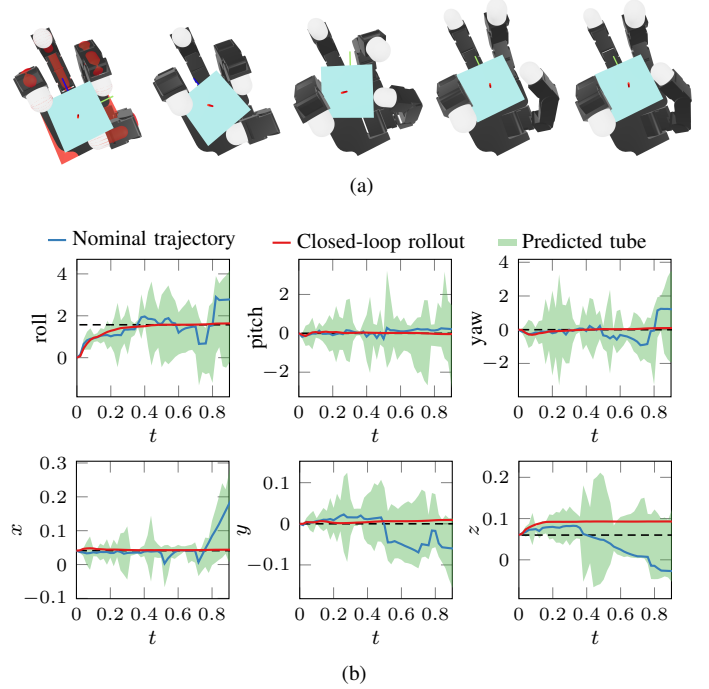


Figure 1. We evaluate on an in-hand cube reorientation task. Our method produces an affine feedback policy that respects the unilateral nature of contact and can be executed online. (a) Keyframes from the executed rollout. (b) The closed-loop trajectory remains within the reachable tube for all object degrees of freedom, thereby certifying constraint satisfaction.

little guidance for improving actions. As such, while differentiable simulators exist [11], the practical utility of the gradients they provide remains limited [56] for manipulation.

From the RL policy gradient formula [57] with policy π_θ ,

$$\nabla_\theta J(\theta) := \mathbb{E}_{\tau \sim \pi_\theta} [\nabla_\theta \log p_\theta(\tau) R(\tau)],$$

it might appear that the policy gradient $\nabla_\theta J(\theta)$ would also vanish, as most actions do not affect the manipulated object and yield zero episodic reward $R(\tau)$ for a sampled trajectory τ . RL addresses this through stochasticity in the policy and environment, often via domain randomization [44, 48], so that some rollouts accidentally induce contact, producing nonzero rewards and gradients. Domain randomization can be seen as randomized smoothing, where the dynamics are convolved with a probabilistic kernel to improve the optimization landscape [55, 9]. Later work showed that this is equivalent to analytical smoothing [43, 60], where the environment is

deterministic but contact dynamics are smoothed explicitly.

Despite the connection between RL and analytically-smoothed contact dynamics, naively planning with smoothed dynamics often fails due to premature contact loss [52] arising from the unilateral nature of contact, i.e., contact can only apply normal forces in one direction. Later work mitigates this issue by creating a local contact trust region (CTR) [54], replacing conventional elliptical trust regions [53] to keep smoothed dynamics close to the nonsmooth hybrid system. However, contact loss can still occur, and the CTR is heuristic, as the allowable deviation that ensures a sufficiently accurate model for planning is not formally characterized. Quantifying and propagating smoothing error through the dynamics produces reachable sets that can constrain controllers to ensure goal reachability and constraint satisfaction. However, no methods can quantify smoothing error, perform reachability analysis for high-dimensional hybrid contact dynamics, and perform efficient gradient-based policy optimization using these bounds.

To close these gaps, we propose a scalable *robust manipulation planner* that tightly bounds smoothing error, propagates it through the dynamics, and leverages the resulting reachable tubes to guarantee task constraint satisfaction and goal reaching. A key ingredient is a differentiable simulator that smooths contact dynamics and contact geometry in a modular manner via convex conic optimization. This simulator provides informative gradients for planning, as well as gradients with respect to the smoothing parameter κ , which are required to compute smoothing-error bounds. *Our key insight is that the deviation between the smoothed dynamics and the nonsmooth hybrid dynamics can be treated as a structured model mismatch, which we show can be tightly bounded by a set-valued, state-control-dependent function.* Building on recent advances in set-valued robust control [29], we jointly optimize a nominal trajectory and an affine feedback policy that admit analytical predictions of the nonsmooth hybrid dynamics behavior under feedback control. We compare with CTR in Section V-C, obtaining similar solve times while reducing constraint violations and certifying constraint satisfaction on *the nonsmooth hybrid system*. Our contributions are:

- 1) We present a differentiable simulator with contact dynamics and contact geometry smoothing building on modular differentiable convex conic programs, providing accurate gradients for planning, without differentiating through the entire computation graph.
- 2) We derive tight analytical bounds on the deviation between smoothed dynamics and nonsmooth hybrid dynamics, which depend on gradients provided by our differentiable simulator.
- 3) We propose a robust policy synthesis method that uses gradients from the smoothed dynamics for efficient optimization while using tube-valued predictions of the nonsmooth hybrid dynamics under feedback control to ensure robust constraint satisfaction.
- 4) We show the efficacy of our method on contact-rich manipulation problems both in simulation and on hardware, including planar pushing (2DOF), bimanual

non-prehensile manipulation (9DOF), and dexterous in-hand reorientation (22DOF), reducing constraint violation rate relative to baselines.

II. RELATED WORK

Differentiable simulators have been used to compute policy gradients via backpropagation through time [39, 47, 59, 12]. However, these approaches often suffer from vanishing or exploding gradients caused by repeatedly chaining derivatives over long horizons.

Trajectory optimization offers an alternative by avoiding long-horizon gradient chaining and instead using derivatives of the system dynamics to iteratively refine trajectories. For contact-rich systems, contact dynamics are often modeled using the Signorini condition, Coulomb’s law, and the maximum dissipation principle, yielding nonlinear complementarity program (NCP) embedded within the trajectory optimization problem [27, 22]. However, solution to NCP contact dynamics can be multi-valued [33], causing the dynamics to be inconsistent with its local linear approximations and undermining gradient reliability.

Convex formulations of contact dynamics address this issue by ensuring single-valued solutions, which ensures consistency of the gradients for trajectory optimization [54]. In settings where soft, spring-like contact models are sufficient, these convex formulations can even admit closed-form solutions [20, 25]. Regardless of the specific contact model, trajectory optimization is typically performed either via shooting methods, where only controls are decision variables, or via direct transcription, where both controls and states are optimized. In both cases, smoothing of the contact dynamics is typically required to mitigate discontinuous or vanishing gradients.

Another class of methods formulates trajectory optimization as mathematical program with complementarity constraints (MPCC), jointly optimizing controls, states, and contact forces. Early approaches restore constraint qualifications through Scholtes relaxation [49], which effectively smooths the contact complementarity constraints [46, 36]. More recent work instead avoids such relaxation by using solvers that do not introduce dual variables for the complementarity constraints [32, 5].

Gradient-based trajectory optimization can only guarantee convergence to local minima. Although outside the scope of this work, several approaches aim to address this limitation through global optimization, including graph of convex sets [15], polynomial optimization [21], mixed-integer programming [17], and rapidly exploring random tree [43]. While these methods can perform global search, they are typically restricted to low-dimensional systems and remain vulnerable to combinatorial growth in the number of possible contact modes.

Finally, trajectory optimization produces an open-loop trajectory and therefore typically requires continual replanning in a receding-horizon fashion to compensate for modeling errors and disturbances. In contrast, our method produces affine feedback policies that respect the unilateral nature of contact, enabling online deployment without repeated replanning.

III. PRELIMINARIES AND PROBLEM STATEMENT

Conic Programs. Throughout this paper, we make extensive use of convex conic programs of the form

$$\begin{aligned} & \underset{x}{\text{minimize}} && \frac{1}{2}x^\top Px + q^\top x \\ & \text{subject to} && s := b - Ax, \quad s \in \mathcal{K}, \end{aligned} \quad (1)$$

where $\mathcal{K}$ is a Cartesian product of symmetric cones, i.e., $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \dots \times \mathcal{K}_I$. Each $\mathcal{K}_i$ is either a non-negative cone or a second-order cone. A variety of numerical solvers can be used to solve (1), including [13].

The Karush-Kuhn-Tucker (KKT) conditions of (1) are

$$Px + q + A^\top z = 0 \quad (2a) \quad s \circ z = 0 \quad (2c)$$

$$Ax + s - b = 0 \quad (2b) \quad s \in \mathcal{K}, \quad z \in \mathcal{K}^* \quad (2d)$$

where $\mathcal{K}^*$ denotes the dual cone of $\mathcal{K}$, and $\kappa = 0$ corresponds to the complementarity condition.

Relaxing the complementarity condition (2c) to a nonzero value $\kappa > 0$ yields the KKT conditions of the following convex barrier-penalized problem [7, §11.2]:

$$\underset{x}{\text{minimize}} \quad \frac{1}{2}x^\top Px + q^\top x + \kappa \sum_{i=1}^I \psi_i(s_i), \quad (3)$$

where $s_i := (b - Ax)_i$ and $\psi_i(\cdot)$ is the logarithmic barrier associated with the symmetric cone $\mathcal{K}_i$.

Convex conic programs of the form (1) or (3) can be embedded as differentiable layers within computational graphs [1]. When P , q , A , and b depend on problem data θ , the sensitivity of the solution with respect to θ can be computed [2], as can its sensitivity with respect to the complementarity parameter κ . The corresponding implicit differentiation formulas are summarized in Appendix A.

Contact Model. We adopt the quasistatic model of contact dynamics from [42, 43], which assumes velocities are sufficiently small to be neglected and employs a convex formulation of contact. This assumption is appropriate for manipulation tasks in which departures from static equilibrium are brief, so that transient accelerations do not accumulate into significant velocities [37, §10.1]. Although the convex contact formulation introduces an artifact whereby slipping contact occurs with a nonzero offset [4], this offset is proportional to the tangential velocity and is therefore negligible under the quasistatic assumption.

In this model, the nonsmooth contact dynamics $f_0 : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{X}$, which maps the current state and control input (x, u) to the next state x^+ , is defined implicitly as the argmin of the conic program

$$\underset{x^+}{\text{minimize}} \quad \frac{1}{2}(x^+)^\top P(x) x^+ + q(x, u)^\top x^+ \quad (4a)$$

$$\begin{aligned} & \text{subject to} \quad \nu_i := J_i(x) (x^+ - x) + [\phi_i(x) \quad 0 \quad 0]^\top, \quad (4b) \\ & \quad \nu_i \in \mathcal{F}_i^*, \quad \forall i = 1, \dots, n_c, \quad (4c) \end{aligned}$$

where P includes the mass matrix, J_i is the contact Jacobian for the i th contact, and n_c denotes the total number of contact pairs. Additional details of the model are provided in

Appendix B. The κ -smoothed dynamics $f_\kappa : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{X}$ is defined analogously as the argmin of the relaxed problem

$$\underset{x^+}{\text{minimize}} \quad \frac{1}{2}(x^+)^\top P(x) x^+ + q(x, u)^\top x^+ + \kappa \sum_{i=1}^{n_c} \psi_i(\nu_i), \quad (5)$$

where ν_i is defined in (4b). This corresponds to the same complementarity relaxation used to transform (1) into (3). The primal solutions of (4) and (5), denoted by x^+ and x_κ^+ , represent the next state under the nonsmooth and smoothed dynamics, respectively. The corresponding dual solutions, denoted by λ and λ_κ , are the associated contact forces.

Problem Statement. We aim to 1) build a reachability-friendly differentiable simulator and bound the smoothing-error for reachability analysis, and 2) use it for certified contact-rich planning.

Problem 1: Reachability for contact dynamics. Find a tight, efficiently computable, and differentiable bound on the smoothing error $f_0(x, u) - f_\kappa(x, u)$, and propagate this error through the system dynamics to obtain reachable state and control tubes $\{\mathcal{R}_k^x\}_{k=0}^N$ and $\{\mathcal{R}_k^u\}_{k=0}^{N-1}$. These tubes must guarantee that the *true* executed trajectory, evolving under the nonsmooth hybrid dynamics (4), satisfies $x_k \in \mathcal{R}_k^x \subseteq \mathcal{X}$ and $u_k \in \mathcal{R}_k^u \subseteq \mathcal{U}$ for all k .

Problem 2: Robust feedback motion planning. Optimize a nominal state $\mathbf{z} := [z_k]_{k=0}^N$ and control $\mathbf{v} := [v_k]_{k=0}^{N-1}$ trajectory that is feasible under the κ -smoothed dynamics (5) and a causal feedback controller $\pi := (\pi_0, \dots, \pi_{N-1})$, where $\pi_k : (\mathbb{R}^{n_x})^{k+1} \rightarrow \mathbb{R}^{n_u}$, that stabilizes the nonsmooth contact dynamics (4) about $(\mathbf{z}, \mathbf{v})$ and ensures closed-loop satisfaction of state and control constraints.

Outline. Next, we will present our differentiable simulator and smoothing-error bound (Sec. IV) and our reachability-certified contact-rich feedback motion planner (Sec. V).

IV. DIFFERENTIABLE SIMULATOR WITH CONTACT SMOOTHING

A. Differentiable Simulator

Utilizing the ability to differentiate through convex conic programs, we can differentiate the dynamics $x_\kappa^+ = f_\kappa(x, u)$ implicitly defined by (5) to obtain $\partial x_\kappa^+ / \partial x$ and $\partial x_\kappa^+ / \partial u$. We can also compute the sensitivity with respect to κ , which is required for the smoothing-error bound, as derived next.

B. Contact Dynamics Smoothing and Smoothing-Error Bound

Transforming the nonsmooth hybrid dynamics $x^+ = f_0(x, u)$ defined by (4) into the smoothed dynamics $x_\kappa^+ = f_\kappa(x, u)$ defined by (5) via complementarity relaxation inevitably introduces a deviation between the two. Although this deviation may be small over a single time step, it can accumulate over time, leading to significant errors. This creates a tension: the smoothed dynamics is necessary for well-conditioned trajectory optimization, yet our ultimate objective is to provide guarantees with respect to the original nonsmooth hybrid dynamics. To address this issue, we first

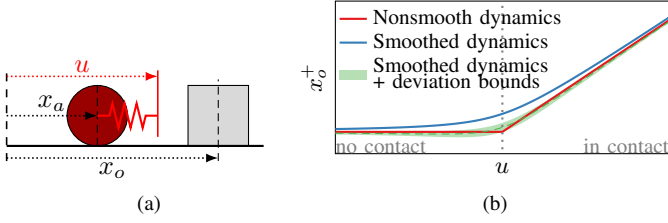


Figure 2. 1D pusher. (a) System schematic. (b) Discrete-time dynamics $x^+ = f(x, u)$, with x_o^+ plotted versus the input u .

quantify the deviation between the nonsmooth and smoothed dynamics, as formalized in the following theorem.

Theorem 1. Under the assumption that $J_i P^{-1} J_j^\top = 0$ for all $i \neq j$, the nonsmooth hybrid dynamics f_0 and the smoothed dynamics f_κ satisfy

$$f_0(x, u) = f_\kappa(x, u) - P(x)^{-1} \sum_{i=1}^{n_c} J_i(x)^\top \frac{\partial \lambda_{\kappa,i}}{\partial \kappa} \kappa w_i, \quad (6)$$

where $\lambda_{\kappa,i}$ denotes the contact force associated with contact i under the smoothed dynamics, and $w_i \in [1, 2]$. Moreover, the bound is tight: for every choice of $w_i \in \{1, 2\}$, there exist (x, u) pairs that satisfy (6).

The proof of Theorem 1 proceeds by integrating $\frac{\partial f_\kappa}{\partial \kappa}$ from 0 to κ ; details are provided in Appendix C. Inspection of (6) reveals that the deviation bound is induced by the terms $\frac{\partial \lambda_{\kappa,i}}{\partial \kappa} \kappa w_i$, which represent deviations in the contact forces. Because w_i is one-sided (i.e., not centered around zero), the resulting bound naturally reflects the unilateral nature of contact. The contact Jacobian $J_i^\top$ maps these contact force deviations into generalized coordinates, and the resulting contributions are aggregated and weighted by P^{-1} .

Equation (6) can be written compactly as

$$x^+ = f_\kappa(x, u) + E_\kappa(x, u) w, \quad (7)$$

where $w \in [1, 2]^{n_c}$ and the i th column of E_κ is given by

$$E_{\kappa,i} = -P(x)^{-1} J_i(x)^\top \frac{\partial \lambda_{\kappa,i}}{\partial \kappa} \kappa.$$

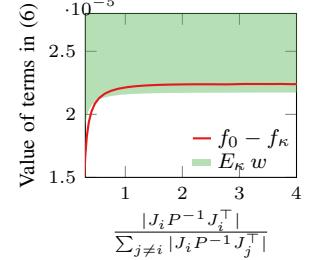
Both f_κ and E_κ are continuously differentiable. The following example illustrates this set-valued bound.

Example 1. Consider the 1D system of a stiffness-controlled red pusher pushing a gray object, shown in Fig. 2a. The corresponding discrete-time dynamics are plotted in Fig. 2b. Under the nonsmooth hybrid dynamics, when the pusher is not in contact with the object, the object position remains constant with respect to the input, resulting in flat regions and nonsmooth transitions between the non-contact and contact modes.

By smoothing the dynamics, the gradient $\partial x_o^+ / \partial u$ becomes positive for all values of u . These gradients are informative, as they guide the optimizer toward actions that induce contact, and they remain continuous across the entire input domain. The shaded region in Fig. 2b represents the bounds induced by the smoothing-error: its upper and lower envelopes correspond to $f_\kappa(x, u) + E_\kappa(x, u)$ and $f_\kappa(x, u) + 2 E_\kappa(x, u)$,

respectively. As predicted by (7), the original nonsmooth dynamics lie entirely within this shaded region. One may also wonder if the nominal dynamics could instead be chosen as $f_\kappa(x, u) + 1.5 E_\kappa(x, u)$, shown as the dashed line in Fig. 2b, to eliminate the bias. The answer is no. Such a choice yields zero or even negative $\partial x_o^+ / \partial u$ gradients over parts of the input domain, rendering them non-informative for action refinement.

Remark 1. On the right, we plot terms in (6) against the diagonal dominance factor (DDF) of $J P^{-1} J^\top$ for the bimanual planar bucket manipulation example in Sec. VI-A. Although Theorem 1 requires $J_i P^{-1} J_j^\top = 0$ for all $i \neq j$, the plot shows that (6) holds if the DDF exceeds approximately 0.5. This suggests, in practice, diagonal dominance is sufficient for (6) to hold numerically. In the Sec. VI examples, the DDF ranges from 1.3 to 2.1, exceeding the 0.5 cutoff.



C. Contact Geometry Smoothing

The complementarity between signed distance and contact force is not the only source of nonsmoothness in the mapping from (x, u) to x^+ . The geometry of contact primitives can also introduce nonsmoothness. For example, in contact between a sphere and a polyhedron, nonsmoothness arises from discontinuous changes in surface normals across the faces of the polyhedron. This results in discontinuities in the contact Jacobian and, consequently, in a nonsmooth dynamics. This issue has also been identified in prior work [38, 26].

In this work, we utilize the differentiability of conic programs by formulating contact point queries as quadratic programs (QPs). For instance, the contact points between a sphere $\{p_1 \mid \|p_1 - p_c\|_2 \leq r\}$ and a polyhedron $\{p_2 \mid A p_2 \leq b\}$ are obtained by solving a relaxed version of the QP

$$\underset{p_2}{\text{minimize}} \quad \frac{1}{2} \|p_2 - p_c\|_2^2 \quad \text{subject to} \quad A p_2 \leq b,$$

analogous to how (1) is relaxed to (3). The resulting contact points are then used to compute the contact Jacobian $J_i(x)$ and signed distance $\phi_i(x)$, which enter into (4).

Ideally, one would characterize how such geometry smoothing induces deviations in the contact Jacobian and signed distance, and thus in the dynamics mapping. We defer this analysis to future work, as the error introduced by contact geometry smoothing is less severe than that arising from contact dynamics smoothing. In particular, contact dynamics smoothing introduces a force-at-a-distance effect whose error can accumulate over time, whereas contact geometry smoothing only induces a local perturbation that manifests as a constant offset.

V. TRAJECTORY OPTIMIZATION AND POLICY SYNTHESIS

We aim to solve the trajectory optimization problem:

$$\text{minimize} \quad J(\mathbf{z}, \mathbf{v}) \quad (8a)$$

$$\text{subject to} \quad z_{k+1} = f_\kappa(z_k, v_k), \quad \forall k = 0, \dots, N-1, \quad (8b)$$

$$z_0 = x_0, \quad (8c)$$

where $\mathbf{z} = [z_0^\top \cdots z_N^\top]^\top$ is the nominal state trajectory and $\mathbf{v} = [v_0^\top \cdots v_{N-1}^\top]^\top$ is the nominal control sequence. The cost $J(\mathbf{z}, \mathbf{v})$ may include terms such as $\sum_{k=0}^N \frac{1}{2} \|z_k - x_{\text{goal}}\|_Q^2$. The dynamics constraint (8b) is based on the smoothed dynamics f_κ rather than the nonsmooth dynamics f_0 . This choice is necessary because f_0 does not provide informative gradients: its associated optimization landscape contains flat regions and sharp discontinuities that hinder gradient-based optimization.

However, when executing the nominal control sequence $\mathbf{v}$ on the nonsmooth hybrid dynamics, the resulting state trajectory $\mathbf{x} = [x_0^\top \cdots x_N^\top]^\top$ deviates from $\mathbf{z}$. As a consequence, there is no guarantee that $\mathbf{x}$ will reach the goal state x_{goal} . More critically, in the presence of state or control constraints, the true trajectory may violate constraints even if the nominal trajectory is feasible. The key idea is that, instead of executing nominal controls on the nonsmooth hybrid dynamics, we execute a closed-loop feedback *policy*. Specifically, we consider a class of affine feedback policies of the form

$$u_k = \pi(x_0, \dots, x_k) = v_k + \sum_{j=0}^k K_{kj} (x_j - z_j), \quad (9)$$

where K_{kj} are the feedback gains. Under such policies, the closed-loop behavior of the nonsmooth hybrid dynamics admits analytically tractable tube-valued predictions. This result is formalized in the following theorem.

Theorem 2. For the dynamics written as (7), consider strictly lower-triangular block matrices

$$\Phi^x = \begin{bmatrix} 0 & & \\ \Phi_{1,0}^x & & \\ \vdots & \ddots & \\ \Phi_{N,0}^x & \cdots & \Phi_{N,N-1}^x \end{bmatrix}, \Phi^u = \begin{bmatrix} 0 & & \\ \Phi_{1,0}^u & & \\ \vdots & \ddots & \\ \Phi_{N-1,0}^u & \cdots & \Phi_{N-1,N-2}^u & 0 \end{bmatrix},$$

that satisfy, for all $j = 0, \dots, N-1$,

$$\begin{aligned} \Phi_{k+1,j}^x &= A_k \Phi_{k,j}^x + B_k \Phi_{k,j}^u, \quad \forall k = j+1, \dots, N-1, \\ \Phi_{j+1,j}^x &= E_j, \end{aligned}$$

where $A_k = \partial f_\kappa(z_k, v_k)/\partial z$, $B_k = \partial f_\kappa(z_k, v_k)/\partial v$, and $E_j = E_\kappa(z_j, v_j)$. The set of all such block matrices Φ^x and Φ^u parameterize all possible closed-loop system responses

$$\mathbf{x} = \mathbf{z} + \Phi^x \mathbf{w} \quad \text{or} \quad x_k = z_k + \sum_{j=0}^{k-1} \Phi_{k,j}^x w, \quad (10a)$$

$$\mathbf{u} = \mathbf{v} + \Phi^u \mathbf{w} \quad \text{or} \quad u_k = v_k + \sum_{j=0}^{k-1} \Phi_{k,j}^u w, \quad (10b)$$

with $\mathbf{w} \in [1, 2]^{N \cdot n_c}$. Moreover, a feedback policy of the form

$$\mathbf{u} = \mathbf{v} + \Phi^u (\Phi^x)^{-1} (\mathbf{x} - \mathbf{z}) := \mathbf{v} + \mathbf{K}(\mathbf{x} - \mathbf{z})$$

achieves such a closed-loop response.

The proof of Theorem 2 is provided in Appendix D. By leveraging Theorem 2, we can formulate an optimization problem in which the closed-loop response of the nonsmooth hybrid dynamics is guaranteed to satisfy the stage and terminal constraints

$$G_k \begin{bmatrix} x_k \\ u_k \end{bmatrix} + g_k \leq 0, \quad \forall k = 0, \dots, N-1, \quad (11a)$$

$$G_f x_N + g_f \leq 0. \quad (11b)$$

The optimization problem that jointly optimizes the nominal trajectory and affine feedback policy is formulated as

$$\min_{\substack{\mathbf{z}, \mathbf{v}, \\ \Phi^x, \Phi^u}} J(\mathbf{z}, \mathbf{v}) + \mathcal{J}(\Phi^x, \Phi^u) \quad (12a)$$

$$\text{s.t.} \quad z_{k+1} = f_\kappa(z_k, v_k), \quad z_0 = x_0, \quad (12b)$$

$$\Phi_{k+1,j}^x = A_k \Phi_{k,j}^x + B_k \Phi_{k,j}^u, \quad \Phi_{j+1,j}^x = E_j, \quad (12c)$$

$$\underbrace{\sum_{j=0}^{k-1} G_k \begin{bmatrix} \Phi_{k,j}^x \\ \Phi_{k,j}^u \end{bmatrix} w_c + w_r \left\| G_k \begin{bmatrix} \Phi_{k,j}^x \\ \Phi_{k,j}^u \end{bmatrix} \right\|_{\text{row},q}}_{h_k(\Phi^x, \Phi^u)} + G_k \begin{bmatrix} z_k \\ v_k \end{bmatrix} + g_k \leq 0, \quad (12d)$$

$$\underbrace{\sum_{j=0}^{N-1} G_f \Phi_{N,j}^x w_c + w_r \left\| G_f \Phi_{N,j}^x \right\|_{\text{row},q}}_{h_f(\Phi^x)} + G_f z_N + g_f \leq 0, \quad (12e)$$

where $\|\cdot\|_{\text{row},q}$ denotes the row-wise q -norm. The constraints (12d) and (12e) correspond to the stage constraint (11a) and the terminal constraint (11b), respectively. Their form follows from the closed-loop system response representation in (10) and a standard robustification argument. Specifically, recall

$$w \in [1, 2]^{n_c} = \{w \in \mathbb{R}^{n_c} \mid \|w - w_c\|_p \leq w_r\},$$

where $w_c = 1.5 \mathbf{1}$, $w_r = 0.5$, and $p = \infty$. For any vector $a \in \mathbb{R}^{n_c}$, the worst-case value of the linear function $a^\top w$ over this uncertainty set admits the closed-form expression

$$\max_{w \in [1, 2]^{n_c}} a^\top w = a^\top w_c + w_r \|a\|_q,$$

where q is the dual norm exponent satisfying $1/p + 1/q = 1$. Applying this property row-wise to the affine dependence of (x_k, u_k) on w in (10) yields the robust constraints in (12).

We do not explicitly account for model mismatch introduced by estimating the nonlinear tracking error dynamics as a linear time-varying (LTV) system. This can be addressed via linearization error bounds [29] computed via interval arithmetic [34] or statistical methods [24, 23], which inflate the tubes to maintain guarantees.

A. Cost and Constraints

The cost (12a) includes two components:

- Nominal trajectory cost, encouraging goal reaching and smooth controls.

$$J(\mathbf{v}, \mathbf{z}) = \sum_{k=0}^N \frac{1}{2} \|z_k - x_{\text{goal}}\|_Q^2 + \sum_{k=0}^{N-2} \frac{1}{2} \|v_k - v_{k+1}\|_{\bar{R}}^2.$$

- Tube cost, which penalizes the size of the tube.

$$\mathcal{J}(\Phi^x, \Phi^u) = \sum_{j=0}^{N-1} \left(\sum_{k=j+1}^N \|\bar{Q}^{\frac{1}{2}} \Phi_{k,j}^x\|_{\mathcal{F}}^2 + \sum_{k=j+1}^{N-1} \|\bar{R}^{\frac{1}{2}} \Phi_{k,j}^u\|_{\mathcal{F}}^2 \right),$$

The constraints (12d) and (12e), or equivalently (11), enforce the following:

- Joint angle and torque limits.

- Non-collision constraints, locally approximated by affine inequalities.
- Object position constraints, which require the object state to lie within a polytopic set.

B. Solver Algorithm

We solve the optimization problem (12) using a sequential convex programming (SCP) approach. At each iteration, we linearize around the current nominal trajectory $(\mathbf{z}, \mathbf{v})$ and solve the resulting convex optimization problem:

$$\min_{\delta \mathbf{z}, \delta \mathbf{v}, \Phi^x, \Phi^u} J(\mathbf{z} + \delta \mathbf{z}, \mathbf{v} + \delta \mathbf{v}) + \mathcal{J}(\Phi^x, \Phi^u) \quad (13a)$$

$$\text{s.t. } \delta z_{k+1} = A_k \delta z_k + B_k \delta v_k, \quad \delta z_0 = 0, \quad (13b)$$

$$\Phi_{k+1,j}^x = A_k \Phi_{k,j}^x + B_k \Phi_{k,j}^u, \quad (13c)$$

$$\Phi_{j+1,j}^x = E_j + \frac{\partial E_j}{\partial z} \delta z_j + \frac{\partial E_j}{\partial v} \delta v_j, \quad (13d)$$

$$h_k(\Phi^x, \Phi^u) + G_k \begin{bmatrix} z_k + \delta z_k \\ v_k + \delta v_k \end{bmatrix} + g_k \leq 0, \quad (13e)$$

$$h_f(\Phi^x) + G_f(z_N + \delta z_N) + g_f \leq 0, \quad (13f)$$

$$\|\delta \mathbf{z}\| \leq \varepsilon, \quad \|\delta \mathbf{v}\| \leq \varepsilon.$$

The constraints (13f) ensure the validity of the local linearization. The resulting increments $\delta \mathbf{z}$ and $\delta \mathbf{v}$ are then used to update the nominal trajectory, and the procedure is repeated until convergence. The full algorithm is summarized in Alg. 1.

Algorithm 1: Certifiable Gradient-Based Policy Synthesis

Input: Initial control sequence $\mathbf{v}$, initial state x_0

Output: Nominal trajectory $(\mathbf{z}, \mathbf{v})$, feedback gains $\mathbf{K}$

// SCP iterations

1 **for** $iter = 1$ **to** $maxIters$ **do**

2 $z_0 \leftarrow x_0$

3 **for** $k = 0$ **to** $N - 1$ **do**

4 $z_{k+1} \leftarrow f_\kappa(z_k, v_k)$

5 **end**

6 $\mathbf{A}, \mathbf{B} \leftarrow \text{LinearizedDynamics}(f_\kappa, \mathbf{z}, \mathbf{v})$

7 $\mathbf{E}, \frac{\partial \mathbf{E}}{\partial \mathbf{z}}, \frac{\partial \mathbf{E}}{\partial \mathbf{v}} \leftarrow \text{EvaluateDisturbance}(E_\kappa, \mathbf{z}, \mathbf{v})$

8 $\delta \mathbf{z}, \delta \mathbf{v}, \Phi^x, \Phi^u \leftarrow \text{Solve}(13)(\mathbf{z}, \mathbf{v}, \mathbf{A}, \mathbf{B}, \mathbf{E}, \frac{\partial \mathbf{E}}{\partial \mathbf{z}}, \frac{\partial \mathbf{E}}{\partial \mathbf{v}})$

9 $\mathbf{v} \leftarrow \mathbf{v} + \delta \mathbf{v}$

10 **end**

// Post-processing

11 $\mathbf{K} \leftarrow \Phi^u(\Phi^x)^{-1}$

12 **for** $k = 0$ **to** $N - 1$ **do**

13 $v_k \leftarrow v_k + \sum_{j=0}^k K_{kj} (x_j - z_j)$

14 $z_{k+1} \leftarrow f_\kappa(z_k, v_k)$

15 $x_{k+1} \leftarrow f_0(x_k, v_k)$

16 **end**

17 **return** $\mathbf{z}, \mathbf{v}, \mathbf{K}$

While the optimization problem (13) is convex and can be solved by standard solvers, its structure can be exploited to drastically improve computational efficiency. A specialized solver is detailed in Appendix E.

C. Baseline

We compare our method against a baseline trajectory optimization with contact trust region (TO-CTR) proposed in

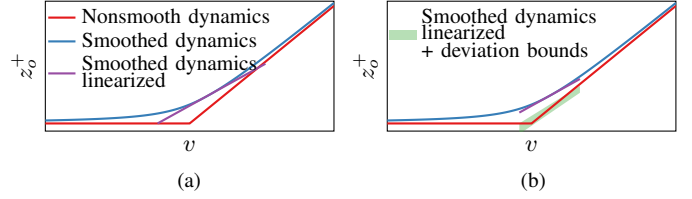


Figure 3. Linearization of the smoothed dynamics $z^+ = f_\kappa(z, v)$ for a 1D pusher as viewed by (a) TO-CTR and (b) our method.

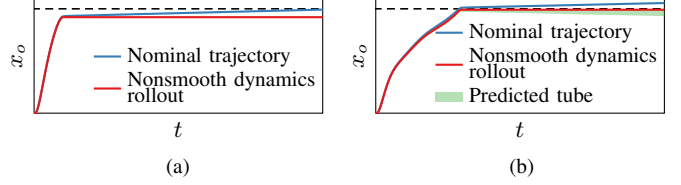


Figure 4. Rollouts of the 1D pusher example for pushing the object to reach a goal position (dashed line) using (a) TO-CTR and (b) our method.

[54]. The method is closely related to iterative linear quadratic regulator (iLQR), but augments the backward pass with a heuristic CTR constraint. Specifically, during the backward pass, instead of solving (13), TO-CTR solves the following convex optimization problem:

$$\min_{\delta \mathbf{z}, \delta \mathbf{v}} J(\mathbf{z} + \delta \mathbf{z}, \mathbf{v} + \delta \mathbf{v}) \quad (14a)$$

$$\text{s.t. } \delta z_{k+1} = A_k \delta z_k + B_k \delta v_k, \quad \delta z_0 = 0, \quad (14b)$$

$$G_k \begin{bmatrix} z_k + \delta z_k \\ v_k + \delta v_k \end{bmatrix} + g_k \leq 0, \quad (14c)$$

$$G_f(z_N + \delta z_N) + g_f \leq 0, \quad (14d)$$

$$\|\delta \mathbf{z}\| \leq \varepsilon, \quad \|\delta \mathbf{v}\| \leq \varepsilon, \quad (14e)$$

$$\lambda_\kappa + \frac{\partial \lambda_\kappa}{\partial z_k} \delta z_k + \frac{\partial \lambda_\kappa}{\partial v_k} \delta v_k \in \mathcal{F}. \quad (14f)$$

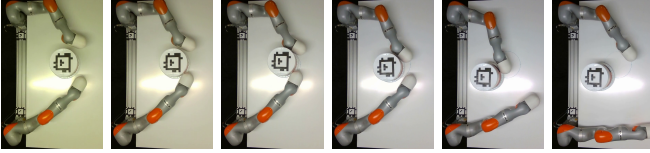
The constraint (14f) enforces the CTR by restricting the linearized contact force to remain within the friction cone $\mathcal{F}$. An intuitive explanation of the key difference between TO-CTR and our method is best illustrated through an example.

Example 2. Consider the 1D pusher-object system (Fig. 2). The linearized dynamics used by TO-CTR and our method are shown in Fig. 3. TO-CTR assumes that the linearization of the smoothed dynamics accurately approximates the nonsmooth hybrid dynamics over control perturbations δv satisfying the CTR constraint $\lambda_\kappa + \frac{\partial \lambda_\kappa}{\partial v} \delta v \geq 0$ and $|\delta v| \leq \varepsilon$. As illustrated, this approximation error remains non-negligible within the assumed region. In contrast, our method explicitly bounds the deviation between the smoothed and nonsmooth hybrid dynamics, guaranteeing that for all $|\delta v| \leq \varepsilon'$, the nonsmooth dynamics lies within a known deviation of the smoothed model.

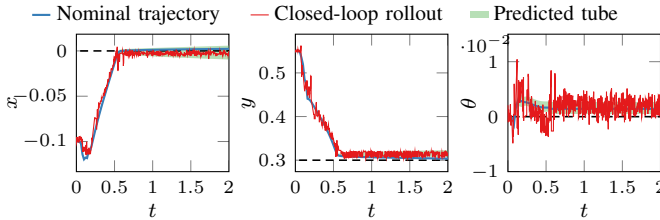
This difference is evident in the rollouts (Fig. 4). TO-CTR exploits a force-at-a-distance artifact of the smoothed dynamics to reach the goal, but fails when executed on the nonsmooth dynamics (Fig. 4a). Our method compensates for this discrepancy during optimization, overshooting the nominal trajectory so that rollout on the nonsmooth hybrid dynamics reaches the goal (Fig. 4b).

TABLE I
TASK PARAMETERS

	Time step size (s)	Time steps	Dynamics smoothing	Geometry smoothing
Bimanual Planar Bucket	0.02	100	10^{-3}	0
Bimanual Planar Box	0.02	100	10^{-3}	10^{-3}
In-hand cube	0.02	50	10^{-4}	0



(a)



(b)

Figure 5. Bimanual planar bucket manipulation on hardware. (a) Keyframes from the executed rollout. (b) Bucket state trajectory and predicted tube.

VI. EXPERIMENT RESULTS AND DISCUSSION

Through a series of simulated and real-world experiments, we seek to answer the following research questions (RQs):

- RQ1.** Do closed-loop rollouts of the nonsmooth hybrid dynamics remain within the predicted tube?
- RQ2.** Does our method achieve lower goal-reaching error than the baseline while strictly enforcing constraints?
- RQ3.** Does performance degrade due to strict constraint enforcement?
- RQ4.** Is contact geometry smoothing necessary in practice?
- RQ5.** Can our method scale to high-dimensional systems?

Parameters for the experimented tasks are listed in Table I. We perform all computations on a desktop computer with an Intel Core Ultra 9 285K processor and 64 GB RAM.

A. Bimanual Planar Bucket Manipulation

For this task, the goal is to reorient and translate the bucket to a desired target pose. We first evaluate our method on bimanual hardware, using two Kuka iiwa 7 arms, each acting as a 3-DoF finger (Fig. 5). In the hardware rollout, the bucket reaches the goal configuration with only a small steady-state error, in spite of sim-to-real modeling errors such as mass and joint stiffness. We believe this robustness is due in part to the deployed policy (9) having an integral structure, which enables the rejection of constant disturbances. Furthermore, aside from noise in the position estimates, the bucket remains largely within the computed tubes, affirming **RQ1**. Strict containment can be achieved by inflating $E_{\kappa}(x, u)$ with an additional bound to account for perception noise [30, 8].

TABLE II
COMPARISON OF ALGORITHMS ON BIMANUAL BUCKET MANIPULATION

	TO-CTR	Ours	TO-CTR with MPC	Ours with MPC
MPC horizon	-	-	1	1
Goal position error (m)	0.1522	0.1447	0.1558	0.1538
Goal angle error (rad)	0.1034	0.0946	0.0985	0.0987
Nominal traj. (z, v) cost	313.51	299.43	-	-
Real traj. (x, u) cost	314.66	296.41	318.02	309.39
Constraint violating rollouts	3.1%	0.0%	1.2%	0.0%
Compute time (s)	88.2	85.9	6.36	6.99

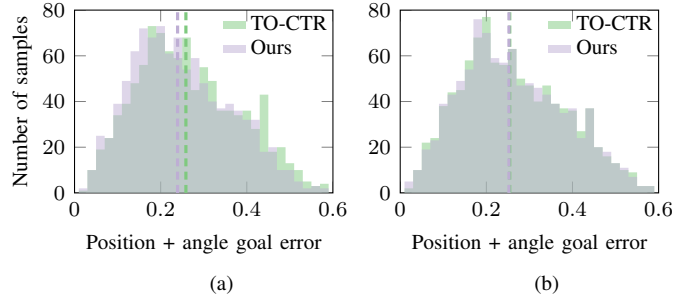


Figure 6. Goal position error plus goal angle error for bimanual planar bucket manipulation. (a) Without replanning. (b) With MPC.

We also use this task to compare against the baseline (Sec. V-C) in simulation. We sampled a set of initial and goal states and computed a trajectory or policy using both the baseline and our method. Samples deemed infeasible by either algorithm were discarded. A quantitative comparison between our method and the baseline is presented in Table II and Fig. 6. All reported metrics are averaged over 1000 experiments.

Without replanning, our method achieves lower average goal position and orientation error than both the open-loop baseline and the baseline with model predictive control (MPC). We therefore confirm **RQ2**: our method improves goal-reaching accuracy, although the improvement is modest. This is likely because the planning horizon is not long enough for smoothing errors to accumulate sufficiently to invalidate the TO-CTR trajectory.

Although our method can also be executed in an MPC fashion, it provides no clear advantage over TO-CTR with MPC, aside from reduced constraint violation. Because our method already yields an affine feedback policy that captures the unilateral nature of contact and enables inexpensive online rollouts, additional replanning through MPC is largely unnecessary unless the system encounters additional disturbances not accounted for during planning.

Table II also indicates that several TO-CTR rollouts on the nonsmooth hybrid dynamics result in constraint violations. These violations manifest either as robot collisions (Fig. 7) or as the bucket falling off the supporting table (Fig. 8). In both cases, the nominal trajectories generated by TO-CTR are feasible under the smoothed model; however, executing the nominal controls on the nonsmooth hybrid dynamics leads to constraint violations.

We also compare against a method that uses exact dynamics in the forward pass and smoothed gradients in the backward

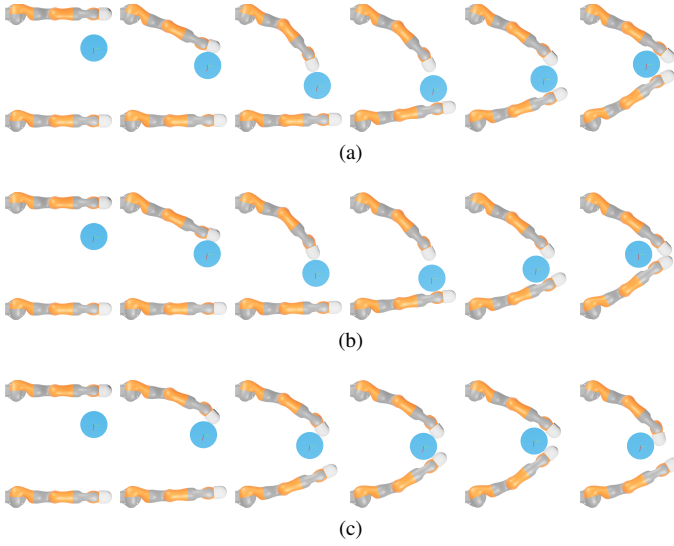


Figure 7. Rollouts of bimanual planar bucket manipulation illustrating robot collisions under baselines and their absence in our method. (a) The trajectory generated by TO-CTR under smoothed dynamics exhibits no collision. (b) Executing the TO-CTR nominal control sequence on the nonsmooth hybrid dynamics leads to a robot collision. (c) Executing the policy produced by our method on the nonsmooth hybrid dynamics avoids robot collision.

pass, as in [22] (Fig. 8b), which is likewise insufficient to guarantee constraint satisfaction. This failure arises because gradient bias prevents state constraints from being accurately translated into control constraints. Furthermore, the resulting linearization error is first-order in $\|\delta u\|$. As a result, the local model is never sufficiently accurate, and reducing $\|\delta u\|$ alone cannot recover feasibility guarantees.

In contrast, our method certifies constraint satisfaction for the nonsmooth hybrid dynamics. Furthermore, as shown in Fig. 8c, our method optimizes a policy that minimizes the reachable tube, thereby keeping the nominal trajectory and the true rollout closer than TO-CTR.

We also address **RQ3**. Strict constraint enforcement does not degrade performance. In fact, our method achieves lower nominal trajectory cost, requiring smaller control effort and incurring lower state cost than the baseline. This suggests that the tube-based reasoning enables the optimizer to select contacts less conservatively. In contrast, the baseline’s CTR condition appears overly conservative relative to the tube characterization, prematurely marking certain transitions as infeasible.

Overall, this experiment validates **RQ1–RQ3**. Closed-loop executions of the nonsmooth hybrid dynamics remain within the predicted tubes (RQ1). We obtain modest improvements in goal-reaching error without sacrificing performance (RQ2–RQ3), while providing a formal certificate of robust constraint satisfaction.

B. Bimanual Planar Box Manipulation

Next, we evaluate our method on a planar box manipulation problem to motivate the need for geometry smoothing (RQ4). In this setting, the discontinuous gradients that arise from the discontinuous contact normals between the actuated arms and the box destabilize the method without geometry smoothing

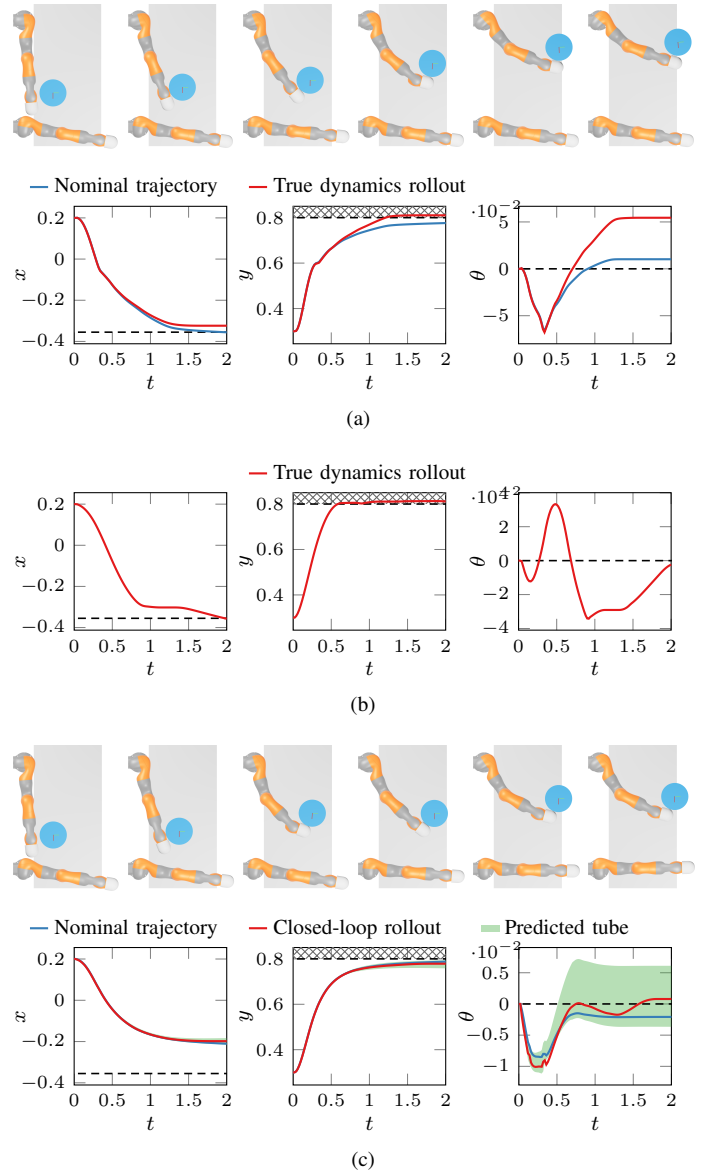


Figure 8. Rollouts of bimanual planar bucket manipulation illustrating state constraint violations under baselines and their absence in our method. Infeasible states are indicated by a cross-hatched pattern. (a) Executing the TO-CTR nominal control sequence on the nonsmooth hybrid dynamics leads to bucket falling out of the gray table. (b) Using exact dynamics for rollouts with smoothed gradients also violates the constraint, causing the bucket to fall off the table in the same manner as shown in the keyframes in (a). (c) Executing the policy produced by our method on the nonsmooth hybrid dynamics prevents the bucket from falling off the table.

(Fig. 9a,c). While the closed-loop rollout remains within the reachable tube, once again validating **RQ1**, the box fails to approach the goal, leading to a large final position error, which is in fact further from the goal than the initial state, due to the unstable gradients. In contrast, with geometry smoothing, we find a solution that gets much closer to the target, leading to a goal error of 9 mm. In this case, the closed-loop rollouts do not stay within the reachable tubes, as the geometry smoothing-error is not accounted for in the bound $E_{\kappa}(x, u)$. However, as noted earlier, the resulting error only leads to a small constant offset from the tubes and does not grow over time. Overall, this result

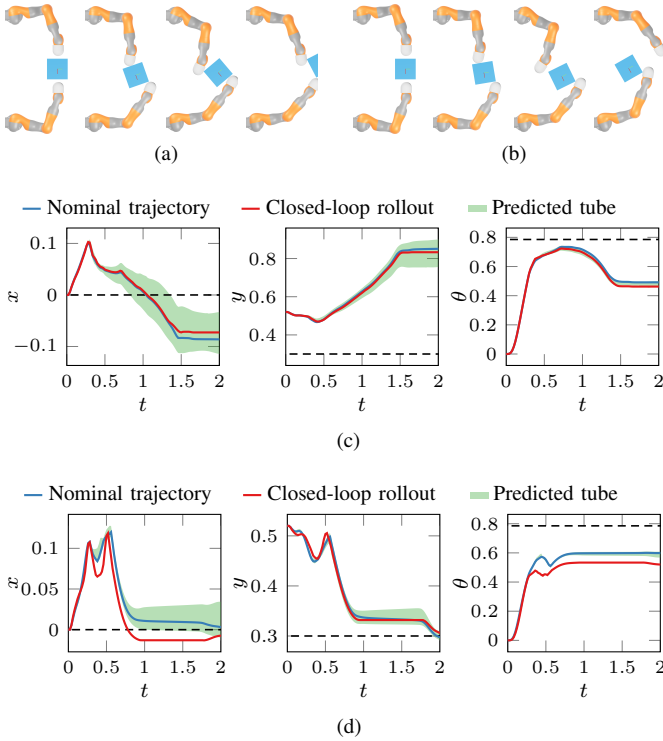


Figure 9. Rollout of bimanual planar box manipulation. Keyframes under (a) no geometry smoothing, and (b) has geometry smoothing. Trajectory and predicted tube per object DoF under (c) no geometry smoothing, and (d) has geometry smoothing.

demonstrates the need for contact geometry smoothing for manipulating objects with discontinuous surface normals, validating **RQ4**, though this is at the cost of tube validity (**RQ1**).

C. In-hand Cube Reorientation

Finally, we evaluate our method on a high-dimensional dexterous in-hand manipulation task using an Allegro hand manipulating a cube. The system consists of 16 actuated degrees of freedom and 6 unactuated object degrees of freedom. The objective is to rotate the cube by 90° (i.e., achieve a roll angle of $\pi/2$) while maintaining its center of mass position and the remaining Euler angles. This yields a 22-dimensional state and 16-dimensional control input system, providing a challenging test of scalability (**RQ5**). Despite the high dimensionality and contact-rich dynamics, Alg. 1 synthesizes a policy in 610 seconds with rollout result shown in Fig. 1. We again validate **RQ1**: the true closed-loop trajectory remains within the predicted reachable tubes. The synthesized policy is able to stabilize the nonsmooth hybrid system about the nominal trajectory, and the resulting goal error is 9 mm and 2.2° . Overall, this experiment demonstrates that Alg. 1 scales to high-dimensional, complex, contact-rich manipulation tasks while retaining a reachability-based certificate that guarantees robust constraint satisfaction.

D. Ablation on Smoothing Parameter

We perform an ablation study on κ and plot the goal error and tube size as functions of κ (Fig. 10). As expected, the

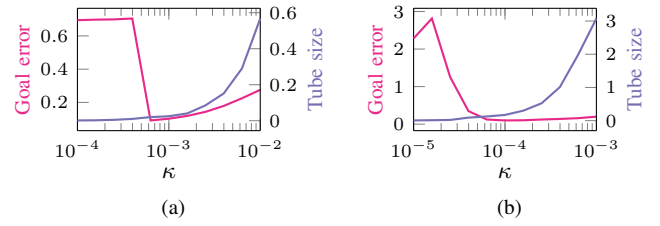


Figure 10. Goal error and tube size as a function of smoothing parameter. (a) Bimanual planar bucket manipulation task. (b) In-hand cube reorientation task.

tube size scales proportionally with κ . For small κ , insufficient smoothing prevents convergence to a good mode sequence, resulting in larger goal error. For sufficiently large κ , the solver converges to a favorable contact mode sequence, yielding a smaller goal error that varies monotonically with κ . Therefore, the smoothing parameter must exceed a task-dependent threshold to ensure reliable convergence, while remaining sufficiently small to avoid unnecessarily large tube sizes. Table I reports the nominal smoothing parameter κ used for each task.

VII. LIMITATIONS

Despite enabling gradient-based policy synthesis for contact-rich manipulation with formal guarantees on the non-smooth hybrid dynamics, our method has several limitations. First, our approach is susceptible to poor local minima, which we believe is partly due to the use of shooting method for optimization. In contrast, direct transcription often appears more robust to initialization and allows dynamically infeasible warm starts. Combining a more robust optimization method with stochastic exploration strategies similar to those used in RL may improve motion planning performance. Second, we account only for errors induced by contact dynamics smoothing and do not model uncertainty in physical parameters such as mass, friction, or geometry. Quantifying how parametric uncertainty propagates through convex programs [14] and incorporating such uncertainty into our framework is an important direction for future work. Third, our formulation relies on a quasistatic model and assumes contact only between actuated and unactuated bodies, limiting its applicability to dynamic manipulation and tool-use scenarios. Finally, our algorithm can benefit from GPU acceleration to reduce solve time [10].

VIII. CONCLUSION

We presented a principled method for handling hybrid contact dynamics that overcomes the nonsmoothness that has traditionally hindered gradient-based trajectory optimization, while retaining guarantees of closed-loop constraint satisfaction. The method builds on characterizing the smoothing error, propagating it to compute reachable tubes, and synthesizing affine feedback policies. Because the feedback gains are constructed from a one-sided deviation bound, they appropriately capture the unilateral nature of contact, while ensuring that the system remains within a region where those gains are valid by construction. The resulting policy enables significantly cheaper online execution than MPC, making formal feedback synthesis practical for high-dimensional contact-rich systems.

REFERENCES

- [1] Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J Zico Kolter. Differentiable convex optimization layers. *Advances in Neural Information Processing Systems*, 32, 2019.
- [2] Akshay Agrawal, Shane Barratt, Stephen Boyd, Enzo Busseti, and Walaa M Moursi. Differentiating through a cone program. *arXiv preprint arXiv:1904.09043*, 2019.
- [3] James Anderson, John C. Doyle, Steven H. Low, and Nikolai Matni. System level synthesis. *Annual Reviews in Control*, 47:364–393, 2019.
- [4] Mihai Anitescu. Optimization-based simulation of non-smooth rigid multibody dynamics. *Mathematical Programming*, 105(1):113–143, 2006.
- [5] Alp Aydinoglu, Adam Wei, Wei-Cheng Huang, and Michael Posa. Consensus complementarity control for multicontact MPC. *IEEE Transactions on Robotics*, 40: 3879–3896, 2024.
- [6] John T. Betts. Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control, and Dynamics*, 21(2):193–207, 1998.
- [7] Stephen Boyd and Lieven Vandenbergh. *Convex Optimization*. Cambridge University Press, 2004.
- [8] Glen Chou, Necmiye Ozay, and Dmitry Berenson. Safe output feedback motion planning from images via learned perception modules and contraction theory. In *International Workshop on the Algorithmic Foundations of Robotics*, pages 349–367, 2023.
- [9] John C. Duchi, Peter L. Bartlett, and Martin J. Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization*, 22(2):674–701, 2012.
- [10] Jeffrey Fang and Glen Chou. Safe large-scale robust nonlinear MPC in milliseconds via reachability-constrained system level synthesis on the GPU. *arXiv preprint arXiv:2604.07644*, 2026.
- [11] C. Daniel Freeman, Erik Frey, Anton Raichuk, Sertan Girgin, Igor Mordatch, and Olivier Bachem. Brax—A differentiable physics engine for large scale rigid body simulation. In *35th Conference on Neural Information Processing Systems*, 2021.
- [12] Ignat Georgiev, Krishnan Srinivasan, Jie Xu, Eric Heiden, and Animesh Garg. Adaptive horizon actor-critic for policy learning in contact-rich differentiable simulation. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.
- [13] Paul J. Goulart and Yuwen Chen. Clarabel: An interior-point solver for conic programs with quadratic objectives. *arXiv preprint arXiv:2405.12762*, 2024.
- [14] Brendan Gould, Chih-Yuan Chiu, Antoine P Leeman, Kyriakos G Vamvoudakis, Samuel Coogan, and Glen Chou. Over-approximating minimizer sets of constrained convex programs with parametric uncertainty via reachability analysis. *arXiv preprint arXiv:2604.27355*, 2026.
- [15] Bernhard Paus Graesdal, Shao Yuan Chew Chia, Tobia Marcucci, Savva Morozov, Alexandre Amice, Pablo A Parrilo, and Russ Tedrake. Towards tight convex relaxations for contact-rich manipulation. In *Proceedings of Robotics: Science and Systems*, 2024.
- [16] Nikolaus Hansen. The CMA evolution strategy: A tutorial. *arXiv preprint arXiv:1604.00772*, 2016.
- [17] Francois R Hogan and Alberto Rodriguez. Reactive planar non-prehensile manipulation with hybrid model predictive control. *The International Journal of Robotics Research*, 39(7):755–773, 2020.
- [18] Taylor Howell, Nimrod Gileadi, Saran Tunyasuvunakool, Kevin Zakka, Tom Erez, and Yuval Tassa. Predictive sampling: Real-time behaviour synthesis with MuJoCo. *arXiv preprint arXiv:2212.00541*, 2022.
- [19] Julius Jankowski, Lara Brudermüller, Nick Hawes, and Sylvain Calinon. VP-STO: Via-point-based stochastic trajectory optimization for reactive robot behavior. *arXiv preprint arXiv:2210.04067*, 2022.
- [20] Wanxin Jin. Complementarity-free multi-contact modeling and optimization for dexterous manipulation. In *Proceedings of Robotics: Science and Systems*, 2025.
- [21] Shucheng Kang, Guorui Liu, and Heng Yang. Global contact-rich planning with sparsity-rich semidefinite relaxations. In *Proceedings of Robotics: Science and Systems*, 2025.
- [22] Gijeong Kim, Dongyun Kang, Joon-Ha Kim, Seungwoo Hong, and Hae-Won Park. Contact-implicit model predictive control: Controlling diverse quadruped motions without pre-planned contact modes or trajectories. *The International Journal of Robotics Research*, 44(3): 486–510, 2025.
- [23] Craig Knuth, Glen Chou, Necmiye Ozay, and Dmitry Berenson. Planning with learned dynamics: Probabilistic guarantees on safety and reachability via Lipschitz constants. *IEEE Robotics and Automation Letters*, 6(3): 5129–5136, 2021.
- [24] Craig Knuth, Glen Chou, Jamie Reese, and Joseph Moore. Statistical safety and robustness guarantees for feedback motion planning of unknown underactuated stochastic systems. In *2023 IEEE International Conference on Robotics and Automation*, pages 12700–12706, 2023.
- [25] Vince Kurtz, Alejandro Castro, Aykut Özgün Önel, and Hai Lin. Inverse dynamics trajectory optimization for contact-implicit model predictive control. *The International Journal of Robotics Research*, 45(1):23–40, 2026.
- [26] Simon Le Cleac’h, Mac Schwager, Zachary Manchester, Vikas Sindhwani, Pete Florence, and Sumeet Singh. Single-level differentiable contact simulation. *IEEE Robotics and Automation Letters*, 8(7):4012–4019, 2023.
- [27] Simon Le Cleac’h, Taylor A. Howell, Shuo Yang, Chi-Yen Lee, John Zhang, Arun Bishop, Mac Schwager, and Zachary Manchester. Fast contact-implicit model predictive control. *IEEE Transactions on Robotics*, 40:1617–1629, 2024.
- [28] Antoine P. Leeman, Johannes Köhler, Florian Messerer,

- Amon Lahr, Moritz Diehl, and Melanie N. Zeilinger. Fast system level synthesis: Robust model predictive control using Riccati recursions. *IFAC-PapersOnLine*, 58(18): 173–180, 2024.
- [29] Antoine P. Leeman, Johannes Köhler, Andrea Zanelli, Samir Bennani, and Melanie N. Zeilinger. Robust nonlinear optimal control via system level synthesis. *IEEE Transactions on Automatic Control*, 70(7):4780–4787, 2025.
- [30] Antoine P Leeman, Shuyu Zhan, Melanie N Zeilinger, and Glen Chou. VISION-SLS: Safe perception-based control from learned visual representations via system level synthesis. *arXiv preprint arXiv:2604.24894*, 2026.
- [31] Albert H. Li, Preston Culbertson, Vince Kurtz, and Aaron D. Ames. DROP: Dexterous reorientation via online planning. In *2025 IEEE International Conference on Robotics and Automation*, pages 14299–14306, 2025.
- [32] Yulin Li, Haoyu Han, Shucheng Kang, Jun Ma, and Heng Yang. On the surprising robustness of sequential convex optimization for contact-implicit motion planning. In *Proceedings of Robotics: Science and Systems*, 2025.
- [33] Q. Le Lidec, W. Jallet, L. Montaut, I. Laptev, C. Schmid, and J. Carpentier. Contact models in robotics: A comparative analysis. *IEEE Transactions on Robotics*, 40: 3716–3733, 2024.
- [34] D. Limon, J.M. Bravo, T. Alamo, and E.F. Camacho. Robust MPC of constrained nonlinear systems based on interval arithmetic. *IEE Proceedings - Control Theory and Applications*, 152:325–332, 2005.
- [35] Danylo Malyuta, Taylor P. Reynolds, Michael Szmuk, Thomas Lew, Riccardo Bonalli, Marco Pavone, and Behçet Açıkmeşe. Convex optimization for trajectory generation: A tutorial on generating dynamically feasible trajectories reliably and efficiently. *IEEE Control Systems Magazine*, 42(5):40–113, 2022.
- [36] Zachary Manchester, Neel Doshi, Robert J. Wood, and Scott Kuindersma. Contact-implicit trajectory optimization using variational integrators. *The International Journal of Robotics Research*, 38(12-13):1463–1476, 2019.
- [37] Matthew T. Mason. *Mechanics of Robotic Manipulation*. The MIT Press, 2001.
- [38] Louis Montaut, Quentin Le Lidec, Antoine Bambade, Vladimir Petrik, Josef Sivic, and Justin Carpentier. Differentiable collision detection: A randomized smoothing approach. In *2023 IEEE International Conference on Robotics and Automation*, pages 3240–3246, 2023.
- [39] Miguel Angel Zamora Mora, Momchil Peychev, Sehoon Ha, Martin Vechev, and Stelian Coros. PODS: Policy optimization via differentiable simulation. In *Proceedings of the 38th International Conference on Machine Learning*, pages 7805–7817, 2021.
- [40] J. Krishna Murthy, Miles Macklin, Florian Golemo, Vikram Voleti, Linda Petrini, Martin Weiss, Brendan Considine, Jérôme Parent-Lévesque, Kevin Xie, Kenny Erleben, Liam Paull, Florian Shkurti, Derek Nowrouzezahrai, and Sanja Fidler. gradSim: Differentiable simulation for system identification and visuomotor control. In *International Conference on Learning Representations*, 2021.
- [41] Rhys Newbury, Jack Collins, Kerry He, Jiahe Pan, Ingmar Posner, David Howard, and Akansel Cosgun. A review of differentiable simulators. *IEEE Access*, 12: 97581–97604, 2024.
- [42] Tao Pang and Russ Tedrake. A convex quasistatic time-stepping scheme for rigid multibody systems with contact and friction. In *2021 IEEE International Conference on Robotics and Automation*, pages 6614–6620, 2021.
- [43] Tao Pang, H.J. Terry Suh, Lujie Yang, and Russ Tedrake. Global planning for contact-rich manipulation via local smoothing of quasi-dynamic contact models. *IEEE Transactions on Robotics*, 39(6):4691–4711, 2023.
- [44] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. In *2018 IEEE International Conference on Robotics and Automation*, pages 3803–3810, 2018.
- [45] Corrado Pezzato, Chadi Salmi, Elia Trevisan, Max Spahn, Javier Alonso-Mora, and Carlos Hernández Corbato. Sampling-based model predictive control leveraging parallelizable physics simulations. *IEEE Robotics and Automation Letters*, 10(3):2750–2757, 2025.
- [46] Michael Posa, Cecilia Cantu, and Russ Tedrake. A direct method for trajectory optimization of rigid bodies through contact. *The International Journal of Robotics Research*, 33(1):69–81, 2013.
- [47] Yi-Ling Qiao, Junbang Liang, Vladlen Koltun, and Ming C Lin. Efficient differentiable simulation of articulated bodies. In *Proceedings of the 38th International Conference on Machine Learning*, pages 8661–8671, 2021.
- [48] Aravind Rajeswaran, Sarvjeet Ghotra, Balaraman Ravindran, and Sergey Levine. EPOpt: Learning robust neural network policies using model ensembles. In *International Conference on Learning Representations*, 2017.
- [49] Stefan Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001.
- [50] John Schulman, Yan Duan, Jonathan Ho, Alex Lee, Ibrahim Awwal, Henry Bradlow, Jia Pan, Sachin Patil, Ken Goldberg, and Pieter Abbeel. Motion planning with sequential convex optimization and convex collision checking. *The International Journal of Robotics Research*, 33(9):1251–1270, 2014.
- [51] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.
- [52] Yuki Shirai, Tong Zhao, H.J. Terry Suh, Huaijiang Zhu, Xinpei Ni, Jiuguang Wang, Max Simchowitz, and Tao Pang. Is linear feedback on smoothed dynamics sufficient for stabilizing contact-rich plans? In *2025 IEEE International Conference on Robotics and Automation*, page

11926–11932, 2025.

- [53] D. C. Sorensen. Newton’s method with a model trust region modification. *SIAM Journal on Numerical Analysis*, 19(2):409–426, 1982.
- [54] H. J. Terry Suh, Tao Pang, Tong Zhao, and Russ Tedrake. Dexterous contact-rich manipulation via the contact trust region. *The International Journal of Robotics Research*, 2026.
- [55] H.J. Terry Suh, Tao Pang, and Russ Tedrake. Bundled gradients through contact via randomized smoothing. *IEEE Robotics and Automation Letters*, 7(2):4000–4007, 2022.
- [56] H.J. Terry Suh, Max Simchowitz, Kaiqing Zhang, and Russ Tedrake. Do differentiable simulators give better policy gradients? In *Proceedings of the 39th International Conference on Machine Learning*, pages 20668–20696, 2022.
- [57] Richard S. Sutton, David McAllester, Satinder Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. In *Advances in Neural Information Processing Systems*, volume 12, 1999.
- [58] Lieven Vandenberghe. *The CVXOPT linear and quadratic cone program solvers*, 2010.
- [59] Jie Xu, Viktor Makoviychuk, Yashraj Narang, Fabio Ramos, Wojciech Matusik, Animesh Garg, and Miles Macklin. Accelerated policy learning with parallel differentiable simulation. In *International Conference on Learning Representations*, 2022.
- [60] Sheno Zhang, Wanxin Jin, and Zhaoran Wang. Adaptive barrier smoothing for first-order policy gradient with contact dynamics. In *Proceedings of the 40th International Conference on Machine Learning*, pages 41219–41243, 2023.