Unbiased First-Order Randomized Smoothing for Differentiable Simulators

Mathis Scheffler[1] **Wilson Jallet**[2] **Cordelia Schmid**[1] **Justin Carpentier**[1]

Abstract

Computing informative gradients through nonsmooth physical models—such as those involving hard contact or bouncing—is a cornerstone challenge in robotics and machine learning. Randomized smoothing offers a principled way to obtain well-behaved surrogate gradients, yet standard first-order estimators of these smoothed gradients are biased whenever the underlying function is discontinuous. In this work, we show that this bias arises from neglected discontinuity contributions and derive a corrected first-order formula that eliminates it. When the discontinuity structure is known, we exploit it to build discontinuity-aware gradient estimators, which we instantiate for differentiable collision detection on 3D meshes. When it is unknown, we propose a complementary estimator that leverages classical gradients to reduce variance. We validate this estimator on nonsmooth trajectory optimization instances, demonstrating variance reductions of several orders of magnitude over zeroth-order baselines in the smooth regime, while remaining unbiased under discontinuities where standard first-order estimators fail.

1 Introduction

Gradient-based optimization is central to modern machine learning, yet many real-world objectives are nonsmooth or discontinuous, which undermines standard gradient methods. A classical remedy is *randomized smoothing* [Nesterov, 2011]: given a nonsmooth function f, one considers its smoothed surrogate $f_p(x) = \mathbb{E}_{\omega\sim p}\big[f(x+\omega)\big]$, whose gradient can be estimated via Monte Carlo sampling. This enables smooth optimization of nonsmooth objectives [Duchi et al., 2012], with applications in machine learning: from differentiable combinatorial solvers such as shortest paths [Berthet et al., 2020], to certified robustness and explainability in deep learning [Cohen et al., 2019, Liu et al., 2020].

In robotics, where differentiable simulation is a key enabler for gradient-based methods [Freeman et al., 2021, Howell et al., 2022a, Werling et al., 2021, Le Lidec et al., 2024], phenomena such as hard contact or friction introduce severe nonsmoothness. Randomized smoothing has thus been applied to a wide range of robotics problems from trajectory optimization [Suh et al., 2022a, Le Lidec et al., 2022, Suh et al., 2022b] to differentiable collision detection [Montaut et al., 2023]. A key challenge is the sensitivity of downstream methods to the bias and variance of the smooth surrogate's gradient estimator [Ghadimi and Lan, 2013]. Zeroth-order estimators, which only query function values [Williams, 1992], are always unbiased, but typically exhibit high variance. Trying to achieve lower variance, Le Lidec et al. [2022] introduced the first-order bundled gradient estimator, an estimator based on the original function's derivative, which promises lower variance, yet Suh et al. [2022b] show this estimator is biased on discontinuous data. This bias is not a mere theoretical concern: Suh et al. [2022b] demonstrate, across a variety of robotics tasks, that it causes optimization to fail in practice. In the presence of such discontinuities, to maintain unbiased estimators, one is thus forced to resort to zeroth-order estimators, thereby forgoing the variance reduction that first-order information could provide.

[1]Inria and École Normale Supérieure, PSL Research University, Paris, France. Correspondence to: Mathis Scheffler <mathis.scheffler@inria.fr>

[2]Capital Fund Management, Paris, France. Work done while at Inria.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

To address this trade-off, Suh et al. [2022b] propose the AoBG estimator, which interpolates between the zero- and first-order estimators by minimizing empirical variance subject to an accuracy constraint, partially mitigating the bias issue in practice. Recent work revisits the bias of first-order estimators under discontinuous dynamics by switching or weighting estimators [Onoda et al., 2026].

Our work takes a complementary perspective: rather than interpolating between biased and unbiased estimators, we correct the first-order estimator directly, remaining unbiased under discontinuities while retaining low variance in smooth regions. To achieve this correction, we draw on the idea of using distribution theory to obtain unbiased gradients through discontinuities, an approach pioneered in differentiable rendering. Li et al. [2018] first identified that hard edges in ray-traced scenes give rise to Dirac delta contributions and proposed edge-sampling to handle them. Bangaru et al. [2021] and Michel et al. [2024] subsequently systematized this approach to arbitrary parametric discontinuities in Monte Carlo integrands. Outside rendering, Lee et al. [2018] derived unbiased reparameterization gradients for non-differentiable models by adding boundary contributions estimated through manifold sampling. Our work transfers this perspective to the randomized smoothing setting in robotics, and more particularly in the context of differentiable simulation. Specifically, we explain the failure of the first-order gradient estimator and derive corrected estimators by leveraging distributional derivatives, treating nonsmooth functions as Schwartz distributions [Rudin, 1974].

Building on this analysis, we derive a corrected formula for the gradient of the smoothed surrogate that explicitly accounts for discontinuity contributions that are missed by standard first-order estimators. When the discontinuity structure is known, this formulation yields unbiased estimators. We instantiate this as a deterministic quadrature scheme (*QFoBG*) for differentiable collision detection on 3D meshes, achieving lower error and computational cost than standard Monte Carlo methods. For environments where discontinuities are unknown, we propose the Corrected First-order Bundled Gradient (*CFoBG*). This estimator requires no knowledge of the discontinuity structure yet remains fully unbiased. On locomotion tasks, *CFoBG* reduces estimator variance by 3–4 orders of magnitude compared to zeroth-order baselines when f is smooth, matches their variance in the presence of discontinuities, and avoids the bias inherent to standard first-order estimators.

The remainder of this paper is structured as follows. Section 2 establishes the necessary background on randomized smoothing and the standard first-order bundled gradient. Building on this foundation, Section 3 introduces our two corrected first-order gradient estimators. Section 4 then validates these estimators through concrete applications in differentiable simulation, before we conclude in Section 5.

2 Background

2.1 Randomized smoothing

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a (possibly nonsmooth) function. Randomized smoothing replaces f with a smooth surrogate f_p defined as the convolution of f with a probability density p:

$$f_p(x) \;:=\; \mathbb{E}_{\omega\sim p}\big[f(x+\omega)\big] \;=\; \int_{\mathbb{R}^n} f(x+\omega)\,p(\omega)\,\mathrm{d}\omega = (f\star p)(x). \tag{1}$$

Throughout, we assume that p is a smooth and symmetric density (i.e., $p(-x) = p(x)$); a common choice is the isotropic Gaussian $p_\sigma(x) \propto \exp(-\|x\|^2/(2\sigma^2))$. Intuitively, f_p is obtained by averaging f over random perturbations of its input, which has the effect of "blurring out" any sharp features such as kinks or discontinuities. In particular, even when f itself is not differentiable, f_p inherits the smoothness of p, so that ∇f_p is well-defined and can serve as a smooth surrogate gradient for optimization problems involving f. Since f_p is generally not available in closed form, both f_p and its gradient ∇f_p are usually estimated via Monte Carlo sampling. To derive expressions for ∇f_p, when f is $\mathcal{C}^1$, we rely on a classical property of convolutions: the derivative of a convolution can be computed by differentiating either factor, i.e.

$$\nabla f_p = (\nabla f)\star p = f\star(\nabla p). \tag{2}$$

Each choice leads to a different Monte Carlo estimator with distinct properties, as we discuss next. However, when f is discontinuous, the identity $\nabla f_p = (\nabla f)\star p$ breaks down—a point we address in Section 3.

2.2 Zeroth-order bundled gradient (ZoBG)

Starting from $\nabla f_p = f \star (\nabla p)$, which is always valid when p is smooth, one obtains the zeroth-order bundled gradient used in Suh et al. [2022a], Le Lidec et al. [2022], Suh et al. [2022b]:

$$\nabla f_p(x) = -\,\mathbb{E}_{\omega\sim p}\big[f(x+\omega)\,\nabla \log p(\omega)\big]. \tag{3}$$

This expression only requires evaluations of f and is valid regardless of the smoothness of f. A simple and effective variance reduction technique is to subtract the control variate $f(x)\,\nabla \log p(\omega)$ inside the expectation [Owen, 2013]:

$$\nabla f_p(x) = -\,\mathbb{E}_{\omega\sim p}\big[\big(f(x+\omega) - f(x)\big)\,\nabla \log p(\omega)\big], \tag{4}$$

since $\mathbb{E}_{\omega\sim p}[f(x)\nabla \log p(\omega)] = 0$ for any smooth density p, so the subtracted term vanishes in expectation. This is a classic control variate which significantly reduces variance [Suh et al., 2022a, Le Lidec et al., 2022, Suh et al., 2022b]. In practice, $\nabla f_p(x)$ is approximated using Monte Carlo estimation by averaging over N i.i.d. samples $\omega_1, \ldots, \omega_N \sim p$. We also consider the antithetic ZoBG estimator:

$$\widehat{\nabla}^{\,\mathrm{anti}}_{\mathrm{ZoBG}} f_p(x) = -\,\frac{1}{N}\sum_{k=1}^{N} \frac{f(x+\omega_k) - f(x-\omega_k)}{2}\,\nabla \log p(\omega_k), \tag{5}$$

which requires $2N$ evaluations of f but cancels all even-order terms in the Taylor expansion when f is smooth, also contributing to variance reduction [Owen, 2013]. It is worth noting that, as further discussed by Suh et al. [2022b], Le Lidec et al. [2022], the score-function identity underlying ZoBG also appears in reinforcement learning as the REINFORCE policy gradient [Williams, 1992].

2.3 First-order bundled gradient (FoBG)

When derivatives of f are available, one can instead attempt to differentiate f under the integral sign. When f is $\mathcal{C}^1$ and under a domination condition on ∇f, differentiating under the integral sign yields the first-order bundled gradient used in Suh et al. [2022a], Le Lidec et al. [2022], Suh et al. [2022b]:

$$\nabla f_p(x) \;=\; \mathbb{E}_{\omega\sim p}\big[\nabla f(x+\omega)\big], \tag{6}$$

which in practice is approximated by averaging $\nabla f(x+\omega_k)$ over N i.i.d. samples $\omega_k \sim p$. As with ZoBG, we also consider an antithetic variant:

$$\widehat{\nabla}^{\,\mathrm{anti}}_{\mathrm{FoBG}} f_p(x) = \frac{1}{N}\sum_{k=1}^{N} \frac{\nabla f(x+\omega_k) + \nabla f(x-\omega_k)}{2}, \tag{7}$$

which requires $2N$ gradient evaluations but cancels all odd-order terms in the Taylor expansion of ∇f, reducing variance when f is smooth.

These estimators approximate ∇f_p by averaging the classical gradient ∇f, and are expected to have lower variance when ∇f is well-behaved [Suh et al., 2022b, Le Lidec et al., 2022]. However, in the presence of discontinuities, the smoothed gradient is still well-defined as $\nabla f_p = f \star (\nabla p)$, but the identity $\nabla f_p = (\nabla f) \star p$ no longer holds, so naively applying (6) yields a biased estimator of ∇f_p. This is precisely the failure mode identified by Suh et al. [2022b]: on problems with discontinuous data (e.g. hard contact, Heaviside-like maps), the first-order estimator (6) can return "wrong", biased gradient estimates, which might hinder downstream exploitation of these gradients [Suh et al., 2022b].

3 Method

In this section, we develop our framework for unbiased first-order randomized smoothing. We first apply generalized derivative calculus to identify bias in standard estimators, deriving a formally corrected gradient that explicitly captures discontinuity terms. Since tracking these discontinuities is often intractable, Section 3.2 introduces the Corrected First-order Bundled Gradient (*CFoBG*). *CFoBG* requires no prior knowledge of the discontinuity structure: it evaluates smooth gradient components directly and captures jump corrections via an unbiased zeroth-order residual. Finally, we analyze the bounds on its variance and its asymptotic behavior as the smoothing parameter σ decays.

3.1 Gradient correction via discontinuity terms

The classical gradient ∇f does not capture the full derivative of a discontinuous function: it misses the contribution of the discontinuities. The appropriate notion is that of distributional (or generalized) derivative, in the sense of Schwartz [Rudin, 1974]. For a piecewise-$\mathcal{C}^1$ function f with discontinuity surfaces $\{\Sigma_i\}$, its generalized derivative Df is given by Rudin [1974, chap. 6], [Golse, 2010, Thm. 3.5.7]:

$$Df \;=\; \underbrace{\nabla f}_{\text{classical gradient}} \;+\; \underbrace{\sum_i s_f^{\Sigma_i}\,\xi_i}_{\text{discontinuity terms}}\;, \tag{8}$$

where ξ_i is the surface measure on Σ_i, and $s_f^{\Sigma_i}(x) \in \mathbb{R}^n$ is the vector-valued signed jump of f across Σ_i at $x \in \Sigma_i$, defined by

$$s_f^{\Sigma_i}(x) = \vec{n}(x) \lim_{t\to 0^+} \big[f(x + t\vec{n}(x)) - f(x - t\vec{n}(x))\big], \tag{9}$$

where $\vec{n}(x) \in \mathbb{R}^n$ is the unit normal to Σ_i at x and $\lim_{t\to 0^+}[f(x + t\vec{n}(x)) - f(x - t\vec{n}(x))]$ is the scalar jump amplitude. In dimension $n = 1$, each surface Σ_i reduces to a point x_i and ξ_i is a Dirac mass δ_{x_i}. Convolving this generalized derivative with p, we obtain the corrected first-order gradient of the smoothed surrogate [Rudin, 1974, Thm. 6.30]:

Theorem 1. *Let f be piecewise-$\mathcal{C}^1$ with a finite collection of discontinuity surfaces $\{\Sigma_i\}$, each an oriented hypersurface, continuous and differentiable almost everywhere, and intersecting only on sets of $(n-1)$-dimensional measure zero. Assume that:*

1. $\mathbb{E}_{\omega\sim p}[\|\nabla f(x+\omega)\|] < \infty$ *locally uniformly in x,*
2. $\displaystyle\int_{\Sigma_i} p(z-x)\,\|s_f^{\Sigma_i}(z)\|\,\mathrm{d}\xi_i(z) < \infty$ *for each i.*

Then,

$$\nabla f_p(x) \;=\; \underbrace{\mathbb{E}_{\omega\sim p}\big[\nabla f(x+\omega)\big]}_{\text{smooth part (standard FoBG)}} \;+\; \underbrace{\sum_i \int_{\Sigma_i} p(z-x)\, s_f^{\Sigma_i}(z)\,\mathrm{d}\xi_i(z)}_{\text{discontinuity corrections}}. \tag{10}$$

The first-order estimator used in Suh et al. [2022b], Le Lidec et al. [2022] corresponds to keeping only the first term of (10) while the discontinuity correction terms are entirely absent. This is the precise source of failure. For instance, for the Heaviside step function, $\mathcal{H}(x) = \mathbf{1}_{x>0}$, the classical gradient $\nabla\mathcal{H}$ is zero a.e., so the standard FoBG (6) returns $\nabla\mathcal{H}_p(x) = 0$. By contrast, applying Theorem 1, we have a discontinuity surface $\Sigma = \{0\}$ with jump $s_{\mathcal{H}}^{\Sigma}(0) = 1$ so $\nabla\mathcal{H}_p(x) = \int_{\{0\}} p(z-x)\,\delta_0(z) = p(0-x)$, which is exactly the true smoothed gradient $\nabla\mathcal{H}_p(x) = p(-x)$.

Validity of the assumptions in simulators. Over a finite horizon, a fixed-step simulator composes finitely many piecewise-smooth maps, so the resulting map is piecewise smooth with finitely many pieces, and fixed-step integration rules out event accumulation (Zeno behavior, see Nurkanović et al. [2023]). On a compact region, and when trajectories cross the discontinuity surfaces transversally (i.e., with nonzero normal velocity, see Nurkanović et al. [2023]), the discontinuity set is thus a finite union of smooth hypersurfaces of finite area. Degenerate events such as simultaneous contacts lie on intersections of these surfaces, which have zero surface measure and do not contribute to (10), while active-set changes in complementarity-based contact only create kinks, leaving the map continuous and piecewise-$\mathcal{C}^1$. Near grazing contacts, gradients may blow up as an integrable inverse square-root singularity, which is why Theorem 1 requires integrable rather than bounded gradient terms.

Estimation with known discontinuity surfaces. When the discontinuity surfaces $\{\Sigma_i\}$ are known and can be sampled via a proposal density q_i on Σ_i, the correction terms can be estimated by importance sampling, drawing M i.i.d. samples $z_1, \ldots, z_M \sim q_i$:

$$\widehat{\nabla} f_p(x) = \frac{1}{N}\sum_{k=1}^{N} \nabla f(x+\omega_k) + \sum_i \frac{1}{M}\sum_{k=1}^{M} \frac{p(z_{i,k}-x)\, s_f^{\Sigma_i}(z_{i,k})}{q_i(z_{i,k})}, \quad \omega_k \overset{\text{i.i.d.}}{\sim} p,\;\; z_{i,k} \overset{\text{i.i.d.}}{\sim} q_i. \tag{11}$$

This is the case in structured settings where the discontinuity surfaces are accessible by construction. In differentiable rendering, Bangaru et al. [2021], Michel et al. [2024], Li et al. [2018] compute derivatives similarly, requiring boundary sampling along discontinuities. In robotics, discontinuity sampling can also be leveraged. For instance, in differentiable collision detection, randomized smoothing is applied to mesh support functions to obtain derivatives of inter-geometry distances or witness points [Montaut et al., 2023]. There, the discontinuity surfaces are cones attached to the edges of the mesh, and the surface integrals of (10) can even be computed by deterministic quadrature rather than sampling. We call the resulting zero-variance estimator QFoBG and detail it in Section 4.1. However, in the general case, several challenges limit the direct applicability of Theorem 1:

Combinatorial explosion. Even in simple physical systems, the number of discontinuity surfaces can be huge. For instance, chaotic systems such as the Sinai billiard [Sinai, 1970] exhibit an exponentially growing number of discontinuity surfaces as the simulation horizon increases.

Composition of simulation steps. In many robotic applications, we consider a composition of T time steps, for instance in simulation: $\mathrm{sim} = \mathrm{step}_T \circ \cdots \circ \mathrm{step}_1$. Even if the discontinuity surfaces of each step are known, recovering those of the full composition is costly because discontinuities introduced at different time steps may cancel one another.

As an illustration, consider a ball dropped from a height h above a flat ground (Figure 1). At time t, it hits the ground and bounces. Increasing h slightly causes the ball to miss the ground at time t (crossing one discontinuity surface of step_t) but to hit it at time $t{+}1$ instead (crossing a discontinuity surface of step_{t+1}). The map from initial height to final position at time T remains continuous: the two discontinuities cancel each other. Sampling one discontinuity surface without its compensating counterpart does not merely leave the estimator incomplete: it actively injects a spurious correction term.

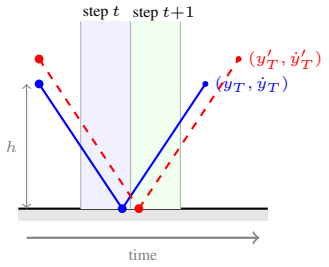

Figure 1: Increasing h by ε shifts the impact from step t to step $t{+}1$. A discontinuity surface is crossed at each step, but the final state $(y_T, \dot{y}_T)$ varies continuously.

Non-commutativity with composition. Randomized smoothing applied to a composition of functions is not equivalent to composing the individually smoothed functions: $(f \circ g)_p \neq f_p \circ g_p$ in general. For instance, taking $f = g = \mathcal{H}$ (Heaviside) and $p = \mathcal{N}(0, \sigma^2)$, we have $(\mathcal{H} \circ \mathcal{H})_p(x) = \Phi(x/\sigma)$ while $\mathcal{H}_p \circ \mathcal{H}_p(x) = \Phi(\Phi(x/\sigma)/\sigma)$, which differ by a factor of 22 at $x = -2$, $\sigma = 1$. This means that smoothing a full simulator end-to-end is fundamentally different from smoothing each step independently, further complicating the use of per-step discontinuity information.

3.2 *CFoBG* estimator

When the discontinuity surfaces $\{\Sigma_i\}$ are unknown, the corrected estimator (11) is not directly applicable. We propose the Corrected First-order Bundled Gradient estimator (*CFoBG*) for smoothing with a Gaussian kernel that relies solely on pointwise evaluations of f and ∇f. Following the decomposition of Theorem 1, it estimates the smooth part $\mathbb{E}_\omega[\nabla f(x{+}\omega)]$ via first-order evaluations of ∇f, and approximates the unknown discontinuity corrections via a zeroth-order residual. Throughout this section, to bound variance, we assume that f is piecewise-$\mathcal{C}^3$ with derivatives of at most polynomial growth on each smooth piece. This hypothesis is not necessary to obtain the unbiasedness of the following estimator.

The CFoBG estimator aims to estimate the smooth term and the discontinuity term of (10) separately. The smooth part is estimated using the antithetic FoBG estimator (7):

$$\widehat{g}^{\mathrm{smooth}} = \frac{1}{N}\sum_{k=1}^{N} \frac{\nabla f(x+\omega_k) + \nabla f(x-\omega_k)}{2}. \tag{12}$$

The discontinuity corrections are estimated by a zeroth-order residual. To preserve unbiasedness, we use a leave-one-out construction: for each k, let $\widehat{g}^{\mathrm{smooth}}_{-k}$ denote (12) computed on all samples except

k, ensuring independence between ω_k and $\widehat{g}^{\text{smooth}}_{-k}$. Then:

$$\widehat{g}^{\text{jump}} \;=\; \frac{1}{N}\sum_{k=1}^{N} \frac{\omega_k \left(f(x+\omega_k) - f(x-\omega_k) - 2\,\omega_k^\top \widehat{g}^{\text{smooth}}_{-k}\right)}{2\sigma^2}, \tag{13}$$

and the CFoBG estimator is

$$\widehat{\nabla}_{\text{CFoBG}} f_p(x) \;=\; \widehat{g}^{\text{smooth}} + \widehat{g}^{\text{jump}}. \tag{14}$$

Since ω_k is independent of $\widehat{g}^{\text{smooth}}_{-k}$, the estimator $\widehat{\nabla}_{\textit{CFoBG}} f_p(x)$ is unbiased and achieves variance $O(\sigma^4/N)$ away from discontinuities (Propositions 1 and 2, Figure 2), which is significantly better than ZoBG whose variance is $O(\|\nabla f\|^2/N)$ away from discontinuities. Furthermore, since $\widehat{g}^{\text{smooth}}$ is an unbiased estimator of the smooth term of (10), it follows that $\widehat{g}^{\text{jump}}$ is an unbiased estimator of the jump correction of (10). Unlike AoBG [Suh et al., 2022b], which retains bias near discontinuities (Figure 7, Appendix C), CFoBG carries no such bias by construction.

CFoBG requires $2N$ evaluations of both f and ∇f. With a fused value-and-gradient pass, as provided by differentiable simulators, this is the same cost as the antithetic FoBG estimator, since computing the leave-one-out means is a negligible post-processing of the samples.

3.3 Further analysis

Behavior near discontinuities. Near a discontinuity surface, the variance of the ZoBG and *CFoBG* estimators is dominated by the jump contribution, which is $O(1/(N\sigma^2))$. In particular, the improvements of *CFoBG* over ZoBG only manifest away from discontinuities (Proposition 3, Figure 2): near a discontinuity, the local gradient $\nabla f(x)$ is a poor predictor of the function's behavior, so first-order information does not provide significant variance reduction.

Relevance as $\sigma \to 0$. In randomized smoothing pipelines, σ is typically annealed toward zero as optimization progresses [Le Lidec et al., 2022]. In this regime, ZoBG variance stays at $O(\|\nabla f(x)\|^2/N)$, while *CFoBG* achieves $O(\sigma^4/N)$ away from discontinuities, a gain of order $\|\nabla f\|^2/\sigma^4$ that grows as $\sigma \to 0$.

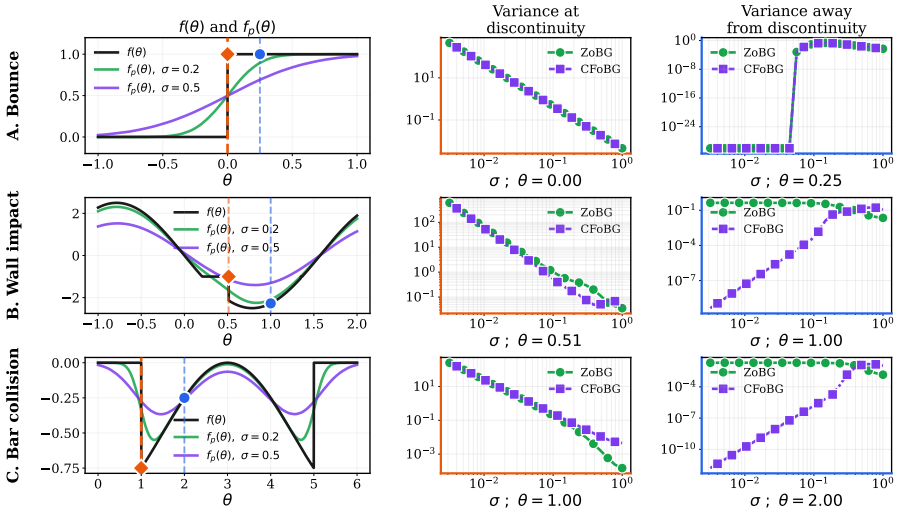

Figure 2: Estimator variance on three tasks involving discontinuity from Suh et al. [2022b]. Columns: $f(\theta)$ and $f_p(\theta)$, variance at and away from discontinuities as a function of σ. *CFoBG* matches ZoBG near discontinuities but achieves $O(\sigma^4/N)$ variance away from them. Contrary to Suh et al. [2022b], our first-order estimator is unbiased.

4 Applications to differentiable simulation

4.1 Differentiable collision detection

In robotics, differentiable collision detection enables the computation of derivatives of geometric quantities such as the minimum distance between two convex bodies or their witness points with respect to the poses of the bodies. These derivatives are critical in applications ranging from differentiable simulation [Howell et al., 2022a, Le Lidec et al., 2024, Freeman et al., 2021] to trajectory optimization [Dietz et al., 2026, Howell et al., 2022b] and can be obtained by evaluating the Hessian of the support function of each geometry [Montaut et al., 2023].

We consider a convex geometry $\mathcal{G}$ approximated by the convex hull A of a set of vertices $\{v_1, \dots, v_K\} \subset \mathbb{R}^3$. The support function of A is $\sigma_A(x) = \max_{1 \le i \le K} \langle v_i, x \rangle$. This function is piecewise affine: wherever the maximizer is unique, $\nabla \sigma_A(x) = \arg\max_{v_i} \langle v_i, x \rangle$, so the gradient is piecewise constant. Its classical Hessian is therefore either zero or undefined. This is a direct consequence of the polytope being a piecewise-flat approximation of $\mathcal{G}$: unlike the original smooth geometry, A carries no curvature information. To recover a meaningful Hessian, following Montaut et al. [2023], Berthet et al. [2020], we smooth $\nabla \sigma_A$ by convolution with an isotropic Gaussian p_σ and differentiate, yielding $\widetilde{H}(x) = \nabla_x \mathbb{E}_{\omega \sim p_\sigma}[\nabla \sigma_A(x + \omega)]$.

Smoothed Hessian. Let $R_i = \{x : \langle v_i, x \rangle \ge \langle v_k, x \rangle, \forall k\}$ be the region where v_i is the maximizer of σ_A, so that $\nabla \sigma_A = v_i$ on the interior of R_i. Applying Theorem 1 componentwise to $\nabla \sigma_A$, and noting that the smooth part vanishes since $\nabla \sigma_A$ is piecewise constant, only the discontinuity contributions remain:

$$\widetilde{H}(x) = \sum_{(i,j) \in \mathcal{E}} (v_i - v_j) \otimes n_{ij} \underbrace{\int_{\Sigma_{ij}} p_\sigma(z - x)\, d\xi(z)}_{I_{ij}(x)}, \tag{15}$$

where $\mathcal{E}$ is the edge set of A, Σ_{ij} is the boundary between regions R_i and R_j, and $n_{ij} = (v_i - v_j)/\|v_i - v_j\|$ is its unit normal pointing from R_j to R_i. Although we instantiate this formula with a Gaussian kernel, equation (15) holds for any smooth, symmetric probability density p; the specific kernel only enters in the evaluation of $I_{ij}(x)$. Note that the naive FoBG applied to this problem returns $\widetilde{H}(x) = 0$ for all x, completely missing the curvature information. This is a direct instance of the failure mode described in Section 3.

Quadrature evaluation. Each ingredient of (15) can be obtained in closed form (see Appendix A), with only $I_{ij}(x)$ being approximated by a Gauss–Legendre quadrature. We refer to this as: Quadrature First-order Bundled Gradient estimator (**QFoBG**). Because it involves no Monte Carlo sampling, QFoBG has zero variance; its only source of error is the quadrature approximation. Computation time ranges from about $0.3\,\mu$s (tens of vertices) to a few μs (700 vertices) on a single CPU core, leading to practical timing performance for differentiable simulation, aligned with benchmarks in Le Lidec et al. [2024].

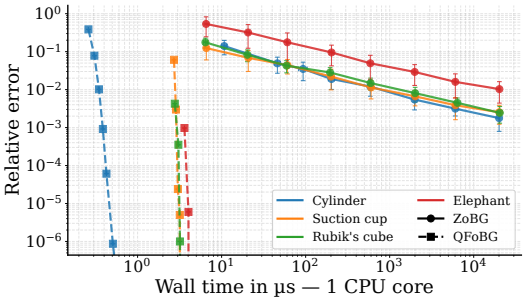

Figure 3: Average relative error of ZoBG and QFoBG as a function of computation time, for four different meshes. ZoBG computation time scales linearly with the number of MC samples; QFoBG computation time scales with the number of quadrature points.

Experiments. We compare QFoBG against the zeroth-order baseline ZoBG, which estimates $\widetilde{H}$ by Monte Carlo following equation (3).

Convergence comparison. Figure 3 reports the relative Frobenius error $\|\widetilde{H} - \widetilde{H}^\star\|_F / \|\widetilde{H}^\star\|_F$ as a function of computation time, where the ground truth $\widetilde{H}^\star$ is computed by QFoBG with 64 quadrature points. Both estimators are evaluated on four meshes of increasing complexity (see Figure 4) at a randomly drawn unit direction x. ZoBG converges as $O(M^{-1/2})$, where M is the number of Monte

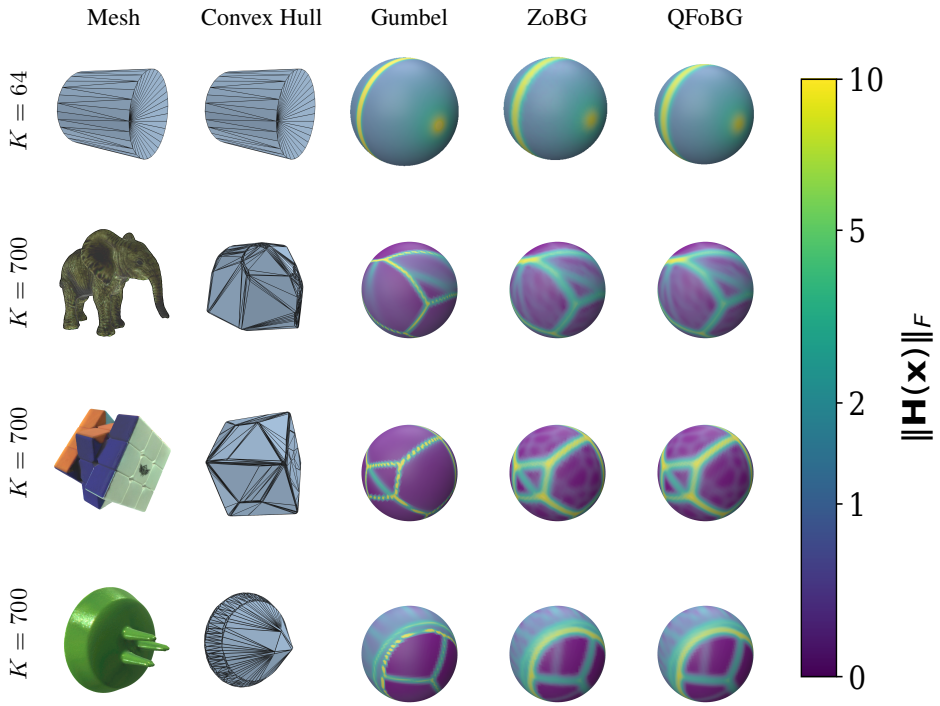

Figure 4: **Smoothed support-function Hessian visualized on $\mathbb{S}^2$.** Each point x is colored by $\log(1 + \|H(x)\|_F)$. Columns show the original mesh and its number of vertices, its convex hull, and the Hessian computed by Gumbel smoothing, ZoBG (10^5 samples), and QFoBG (closed form, ours). Gumbel smoothing perturbs vertex scores rather than the query point, producing dotted artifacts along edges and missing curvature where ZoBG and QFoBG agree on non-zero values. Meshes (except the cylinder) are from the GraspNet dataset [Fang et al., 2020].

Carlo samples, and requires on the order of 10^3 μs to reach a relative error of 10^{-2}. QFoBG achieves 10^{-5} in at most a few microseconds through exponential convergence in the number of quadrature points, yielding several orders of magnitude improvement at equal budget.

Validation on GraspNet. To confirm the correctness of QFoBG on a large and diverse set of geometries, we evaluate it on all meshes of the GraspNet dataset [Fang et al., 2020]. In every case, ZoBG with a large number of samples converges to the value returned by QFoBG, confirming correctness.

Qualitative comparison with Gumbel smoothing. To visualize a Hessian map $x \mapsto H(x)$ across all directions simultaneously, we render a unit sphere where each point $x \in \mathbb{S}^2$ is colored by $\log(1 + \|H(x)\|_F)$: high values indicate directions of large curvature of the support function, while low values correspond to flat regions. A closed-form Hessian approximation also exists under Gumbel smoothing [Montaut et al., 2023], which replaces the $\arg\max$ in $\nabla\sigma_A$'s expression by a softmax. However, Gumbel smoothing perturbs vertex scores $\langle v_i, x\rangle$ rather than the query point x, so it does not correspond to a convolution with a fixed kernel. As visible in Figure 4, this produces dotted artifacts along edges and misses curvature where ZoBG and QFoBG agree on non-negligible values. By contrast, $\widetilde{H}$ relies on an isotropic Gaussian of standard deviation σ, which has a clear geometric interpretation as the spatial averaging scale. Figure 4 confirms that QFoBG produces smooth, faithful Hessians consistent with high-sample ZoBG, while being orders of magnitude faster to compute.

4.2 Trajectory optimization with CFoBG

We evaluate the CFoBG estimator on trajectory optimization tasks, aiming to show that it inherits the best of both worlds: the unbiasedness of ZoBG at discontinuities, and the low variance of FoBG

in smooth regimes. DDP [Li and Todorov, 2004, Tassa et al., 2012] is a trajectory optimization algorithm that requires differentiable dynamics: $x_{t+1} = f(x_t, u_t)$. At each iteration, the backward pass builds a local quadratic model of a value function from the Jacobian $\partial f/\partial u$. When the dynamics are nonsmooth, that Jacobian can be undefined or non-informative, causing DDP to fail. We circumvent this by replacing the true Jacobian with the one of the smoothed dynamics $f_p(x, u) = \mathbb{E}_{\omega\sim p}[f(x, u + \omega)]$ in the backward pass, while the forward pass still rolls out the true dynamics f.

Tasks. We consider three tasks spanning the two regimes of interest: a discontinuous environment (Platformer) and two smooth ones (Hopper variants).

- **2D platformer.** A point mass must climb a two-step staircase to reach a target position (Figure 5). The state is $x = (p_x, p_y, v_x, v_y) \in \mathbb{R}^4$ and the control $u = (u_x, u_y) \in \mathbb{R}^2$ applies a horizontal force and a jump command respectively. The jump is triggered by a hard threshold: when $u_y > 0.1$ and the agent is on the ground, a fixed vertical impulse is applied. Additionally, the staircase steps are solid blocks whose vertical walls cause velocity reversals upon collision. The agent must jump twice to reach the target, requiring the optimizer to discover the jump action despite the discontinuity.
- **Hopper.** We follow Le Lidec et al. [2022] and consider a hopper standing on a plane that must jump to a target altitude z, using MuJoCo [Todorov et al., 2012] as simulator. Since MuJoCo's soft-contact model yields continuous, differentiable dynamics, FoBG is unbiased here, allowing us to evaluate CFoBG's variance reduction against both ZoBG and FoBG in the smooth setting. We also consider a harder variant, **land & jump**, where the hopper starts in the air, must land on the ground, and then jump to the target altitude z (Figure 6).

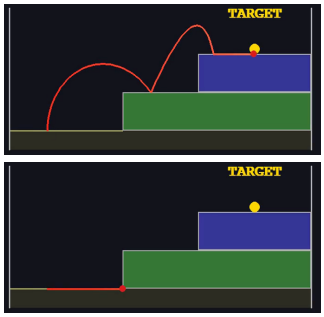

Figure 5: **2D Platformer.** *Top:* ZoBG/CFoBG discover the jump and reach the target. *Bottom:* FoBG never jumps and remains stuck at ground level.

Estimators. At each backward step, we estimate $\partial f_p/\partial u$ using ZoBG, FoBG, and *CFoBG* as defined in Sections 2.2–3.2. The smoothing parameter σ is annealed from 1 to 0.01 following Le Lidec et al. [2022], and all stochastic estimators use $N = 4$ antithetic samples per time step.

Variance measurement. To measure the variance of the Jacobian estimator, we fix an optimized trajectory and, at each time step t and smoothing level σ, compute the empirical variance of the Jacobian estimate over 400 independent trials. We summarize variance by the area under the mean-variance curve (AUC) over the annealing schedule (on a log σ scale). We refer the reader to Appendix D for additional variance plots.

Results. On the platformer (Table 1), since $\partial f/\partial u_y = 0$ almost everywhere due to the Heaviside jump trigger, FoBG never discovers the jump action and the agent remains stuck at ground level. This is a direct instance of the bias identified in Theorem 1: the classical gradient fails to capture the discontinuity contribution entirely. ZoBG and CFoBG both successfully reach the target by discovering the jump as they are unbiased.

Table 1: **Optimization results.** Median trajectory loss over 400 random seeds, and variance AUC. *Top:* Platformer: FoBG never discovers the jump ($\partial f/\partial u_y = 0$ a.e.). *Bottom:* Hopper: CFoBG reduces variance by 3–4 orders of magnitude vs. ZoBG.

Task	Method	Median trajectory loss	Variance AUC
Platformer	ZoBG	**9.29**	8.0×10^{1}
	FoBG	100.7	0
	CFoBG	**9.29**	8.0×10^{1}
Hopper	ZoBG	0.085	9.05
	FoBG	**0.055**	$\mathbf{1.4\times10^{-3}}$
	CFoBG	**0.056**	$\mathbf{2.6\times10^{-3}}$
Hopper (land & jump)	ZoBG	0.153	1.6×10^{1}
	FoBG	**0.107**	$\mathbf{8.5\times10^{-4}}$
	CFoBG	**0.105**	$\mathbf{1.8\times10^{-3}}$

On the hopper tasks (Table 1), where smooth dynamics make FoBG unbiased, CFoBG recovers FoBG-level performance, confirming it incurs little to no overhead in the smooth regime; both estimators outperform ZoBG in terms of trajectory loss and Jacobian estimator variance. This positions CFoBG as a robust alternative to FoBG: unbiased under discontinuities, and competitive in the smooth setting.

Additional benchmarks. We further evaluate CFoBG on three contact-rich tasks from Suh et al. [2022b]: planar pushing with soft contact, Coulomb friction with stick-slip transitions, and Tennis, a feedback-policy optimization problem. Without knowing whether the problem is smooth or discontinuous, CFoBG is best or tied-best on every task, whereas the bias of FoBG makes it fail on Friction and Tennis, and the higher variance of ZoBG degrades it on soft Pushing (Appendix E, Table 2).

5 Conclusion

In this paper, we examined the limitations of first-order bundled gradients when applied to discontinuous problem data. Using generalized derivative calculus, we showed that standard first-order estimators exhibit bias because they omit the contributions of discontinuities, which appear as corrective surface integral terms.

To address this, we introduced unbiased gradient estimators adapted to the available structure of the problem. When discontinuity surfaces are accessible, our QFoBG estimator explicitly approximates these surface integrals to yield unbiased first-order smoothed gradients, as demonstrated in our 3D mesh collision detection experiments. For systems where isolating these surfaces is intractable, our *Corrected FoBG* estimator uses a mixture of pointwise evaluation of the original function and its derivatives to reduce variance in the smooth regime. In the context of trajectory optimization of nonsmooth dynamical systems, this approach reduced estimator variance compared to zeroth-order baselines in the smooth regime while remaining unbiased near discontinuities.

Finally, while our analysis focuses on the randomized smoothing framework, the treatment of generalized derivatives and the use of first-order information for nonsmooth dynamics may have broader utility. These principles apply to other domains involving nonsmooth environments, such as first-order reinforcement learning (e.g., SHAC [Xu et al., 2022]), where discontinuities often impact gradient quality and policy convergence.

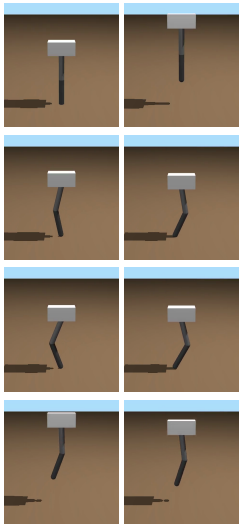

Figure 6: Optimized trajectories on the Hopper tasks. **Left:** jump. **Right:** land & jump.

Limitations and future work. First, CFoBG does not reduce variance at discontinuities, where it has the same asymptotic variance as ZoBG (Proposition 3). Since gradients carry no information about the jump amplitude, it is unclear whether any first-order estimator can do better there, and characterizing this lower bound is an open question. Second, using Theorem 1 directly, as QFoBG does, requires enumerable discontinuity surfaces. This is realistic for structured components such as collision detection, but likely intractable for full simulators, where surfaces multiply and cancel across time steps (Section 3). Two directions remain open: locating discontinuities on the fly, or estimating their contributions without enumerating them explicitly. Third, like all randomized smoothing estimators, CFoBG degrades in high dimensions. For neural policies, one can smooth in the action or state space of the simulator rather than in weight space and backpropagate through the network, which keeps the smoothing dimension small. Finally, the same construction could benefit other methods: it could inject first-order information into sampling-based controllers such as MPPI without introducing bias, and it could correct the bias of first-order reinforcement learning under discontinuous dynamics.

Acknowledgments

This work was supported in part by the French government, under the management of the *Agence Nationale de la Recherche* (ANR), as part of the "France 2030" program (grant ANR-23-IACL-0008, PR[AI]RIE-PSAI project; and grant ANR-22-EXOD-0006, PEPR O2R–AS2 project), and under the PIQ program, managed by the *Agence de Programme du Numérique*. It also received support from the European Union through the ARTIFACT project (Grant Agreement No. 101165695) and the AGIMUS project (Grant Agreement No. 101070165). The authors gratefully acknowledge the support and resources provided by the CLEPS infrastructure at Inria Paris. The views and opinions expressed are those of the authors and do not necessarily reflect those of the funding agencies.

Competing interests. The authors declare no competing interests.

References

Sai Praveen Bangaru, Jesse Michel, Kevin Mu, Gilbert Bernstein, Tzu-Mao Li, and Jonathan Ragan-Kelley. Systematically differentiating parametric discontinuities. *ACM Trans. Graph.*, 40(4): 107:1–107:18, 2021.

Quentin Berthet, Mathieu Blondel, Olivier Teboul, Marco Cuturi, Jean-Philippe Vert, and Francis R. Bach. Learning with differentiable perturbed optimizers. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Jeremy Cohen, Elan Rosenfeld, and J. Zico Kolter. Certified adversarial robustness via randomized smoothing. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019*, pages 1310–1320, 2019.

Christian Dietz, Sebastian Albrecht, Gianluca Frison, Moritz Diehl, and Armin Nurkanović. Robust rigid body assembly via contact-implicit optimal control with exact second-order derivatives, 2026. URL `https://arxiv.org/abs/2601.22849`.

John C. Duchi, Peter L. Bartlett, and Martin J. Wainwright. Randomized smoothing for stochastic optimization. *SIAM J. Optim.*, 22(2):674–701, 2012.

Hao-Shu Fang, Chenxi Wang, Minghao Gou, and Cewu Lu. GraspNet-1Billion: A large-scale benchmark for general object grasping. In *CVPR*, 2020.

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 *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, 2021.

Saeed Ghadimi and Guanghui Lan. Stochastic first- and zeroth-order methods for nonconvex stochastic programming. *SIAM J. Optim.*, 23(4):2341–2368, 2013.

François Golse. *Distributions, analyse de Fourier, équations aux dérivées partielles*. École polytechnique, 2010. ISBN 978-2-7302-1689-0.

Taylor A. Howell, Simon Le Cleac'h, Jan Brüdigam, Qianzhong Chen, Jiankai Sun, J. Zico Kolter, Mac Schwager, and Zachary Manchester. Dojo: A differentiable physics engine for robotics. *arXiv preprint arXiv:2203.00806*, 2022a.

Taylor A. Howell, Kevin Tracy, Simon Le Cleac'h, and Zachary Manchester. CALIPSO: A differentiable solver for trajectory optimization with conic and complementarity constraints. In *Robotics Research - The 20th International Symposium ISRR 2022*, pages 504–521, 2022b.

Quentin Le Lidec, Louis Montaut, Cordelia Schmid, Ivan Laptev, and Justin Carpentier. Leveraging randomized smoothing for optimal control of nonsmooth dynamical systems. *Nonlinear Analysis: Hybrid Systems*, 2022.

Quentin Le Lidec, Louis Montaut, Yann de Mont-Marin, and Justin Carpentier. End-to-end and highly-efficient differentiable simulation for robotics. *arXiv preprint arXiv:2409.07107*, 2024.

Wonyeol Lee, Hangyeol Yu, and Hongseok Yang. Reparameterization gradient for non-differentiable models. In *NeurIPS*, 2018.

Tzu-Mao Li, Miika Aittala, Frédo Durand, and Jaakko Lehtinen. Differentiable Monte Carlo ray tracing through edge sampling. *ACM Trans. Graph.*, 37(6):222, 2018.

Weiwei Li and Emanuel Todorov. Iterative linear quadratic regulator design for nonlinear biological movement systems. In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, pages 222–229, 2004.

Sijia Liu, Pin-Yu Chen, Bhavya Kailkhura, Gaoyuan Zhang, Alfred O Hero III, and Pramod K Varshney. A primer on zeroth-order optimization in signal processing and machine learning: Principals, recent advances, and applications. *IEEE Signal Processing Magazine*, 37(5):43–54, 2020.

Jesse Michel, Kevin Mu, Xuanda Yang, Sai Praveen Bangaru, Elias Rojas Collins, Gilbert Bernstein, Jonathan Ragan-Kelley, Michael Carbin, and Tzu-Mao Li. Distributions for compositionally differentiating parametric discontinuities. *Proc. ACM Program. Lang.*, 8:893–922, 2024.

Louis Montaut, Quentin Le Lidec, Antoine Bambade, Vladimír Petrík, Josef Sivic, and Justin Carpentier. Differentiable collision detection: a randomized smoothing approach. In *IEEE International Conference on Robotics and Automation, ICRA 2023*, pages 3240–3246, 2023.

Yurii Nesterov. Random gradient-free minimization of convex functions. Technical report, Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE), 2011.

Armin Nurkanović, Jonathan Frey, Anton Pozharskiy, and Moritz Diehl. Finite elements with switch detection for direct optimal control of nonsmooth systems with set-valued step functions. In *62nd IEEE Conference on Decision and Control, CDC 2023, Singapore, December 13-15, 2023*, 2023.

Ku Onoda, Paavo Parmas, Manato Yaguchi, and Yutaka Matsuo. Does "do differentiable simulators give better policy gradients?" give better policy gradients? In *ICLR*, 2026.

Art B. Owen. *Monte Carlo theory, methods and examples*. `https://artowen.su.domains/mc/`, 2013.

W. Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1974.

Yakov Sinai. Dynamical systems with elastic reflections. *Russian Mathematical Surveys*, 25:137, 1970.

Hyung Ju Terry Suh, Tao Pang, and Russ Tedrake. Bundled gradients through contact via randomized smoothing. *IEEE Robotics and Automation Letters*, 7(2):4000–4007, 2022a.

Hyung Ju Terry Suh, Max Simchowitz, Kaiqing Zhang, and Russ Tedrake. Do differentiable simulators give better policy gradients? In *International Conference on Machine Learning, ICML 2022*, Proceedings of Machine Learning Research, pages 20668–20696, 2022b.

Yuval Tassa, Tom Erez, and Emanuel Todorov. Synthesis and stabilization of complex behaviors through online trajectory optimization. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4906–4913, 2012.

Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2012*, pages 5026–5033, 2012.

Keenon Werling, Dalton Omens, Jeongseok Lee, Ioannis Exarchos, and C. Karen Liu. Fast and feature-complete differentiable physics for articulated rigid bodies with contact. In *Proceedings of Robotics: Science and Systems (RSS)*, 2021.

Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8:229–256, 1992.

Jie Xu, Viktor Makoviychuk, Yashraj Narang, Fabio Ramos, Wojciech Matusik, Animesh Garg, and Miles Macklin. Accelerated policy learning with parallel differentiable simulation. In *The Tenth International Conference on Learning Representations (ICLR 2022)*, 2022.

A Proofs for Section 4.1

Setting. We assume that A has nonempty interior and that $v_1, \dots, v_K$ are its distinct vertices; this is no loss of generality, since σ_A only depends on the extreme points of A, and non-extreme points can be discarded. Each region R_i is then the normal cone of A at v_i, a closed polyhedral cone with nonempty interior. The regions R_i cover $\mathbb{R}^3$, $\nabla\sigma_A = v_i$ on the interior of R_i, and the discontinuities of $\nabla\sigma_A$ therefore lie in the union of the boundaries Σ_{ij} between pairs of regions.

Discontinuity surfaces. For $i \neq j$, the regions R_i and R_j share the boundary

$$\Sigma_{ij} = \{x \in \mathbb{R}^3 \mid \langle v_i, x\rangle = \langle v_j, x\rangle \geq \langle v_k, x\rangle \ \forall k\}. \tag{16}$$

Lemma 1. *Let $i \neq j$ and $P := (v_i - v_j)^\perp$. Then:*

(i) Σ_{ij} is a cone contained in the plane P;

(ii) if $[v_i, v_j]$ is not an edge of A, then Σ_{ij} has zero surface measure.

Proof. (i) Every $x \in \Sigma_{ij}$ satisfies $\langle v_i - v_j, x\rangle = 0$, hence $\Sigma_{ij} \subset P$. Moreover, the conditions defining Σ_{ij} are invariant under $x \mapsto \lambda x$ for any $\lambda > 0$, so Σ_{ij} is a cone.

(ii) Let $x \in \Sigma_{ij} \setminus \{0\}$, and let $F(x) = \arg\max_{y\in A}\langle y, x\rangle$. Since $x \neq 0$, $F(x)$ is a proper face of A, i.e., a vertex, an edge, or a facet. It contains the two distinct vertices v_i and v_j, so it is not a vertex, and it is not an edge, since the only edge that could contain both is $[v_i, v_j]$, which is not an edge by assumption. Hence $F(x)$ is a facet of A, i.e., a polygon having v_i and v_j among its vertices; it therefore has at least one other vertex v_k, $k \notin \{i, j\}$, which does not lie on the line through v_i and v_j. As v_i and v_k both maximize $\langle \cdot, x\rangle$ over A, we have $\langle v_i - v_k, x\rangle = 0$, so x belongs to $L_k := (v_i - v_j)^\perp \cap (v_i - v_k)^\perp$. Since $v_i - v_j$ and $v_i - v_k$ are linearly independent, L_k is a line. Consequently,

$$\Sigma_{ij} \subseteq \{0\} \cup \bigcup_{k\notin\{i,j\}} L_k, \tag{17}$$

which is a finite union of lines and therefore has zero surface measure. □

By Lemma 1(ii), only the edges of A contribute to the sum in (15).

Lemma 2. *Let $[v_i, v_j]$ be an edge of A, and let F_1 and F_2 be the two facets of A containing it, with outward unit normals m_1 and m_2. Then*

$$\Sigma_{ij} = \operatorname{cone}(m_1, m_2) := \{\alpha\, m_1 + \beta\, m_2 : \alpha, \beta \geq 0\}. \tag{18}$$

Proof. For $r \in \{1, 2\}$, let $c_r = \max_{y\in A}\langle m_r, y\rangle$, so that $F_r = A \cap \{y : \langle m_r, y\rangle = c_r\}$. Since $v_i \in F_r$, we have $\langle m_r, v_i\rangle = c_r$, hence for every vertex v_k of A:

$$v_k \in F_r \ \Rightarrow\ \langle m_r, v_i - v_k\rangle = 0, \qquad v_k \notin F_r \ \Rightarrow\ \langle m_r, v_i - v_k\rangle > 0. \tag{19}$$

In particular, $v_j \in F_r$ gives $m_r \in P = (v_i - v_j)^\perp$. Moreover, m_1 and m_2 are linearly independent: $m_1 \neq m_2$ since $F_1 \neq F_2$, and if $m_2 = -m_1$, then $v_i \in F_1 \cap F_2$ gives $c_2 = -c_1$, so that $A \subset \{y : c_1 \leq \langle m_1, y\rangle \leq c_1\}$ would be contained in a plane, contradicting the nonempty interior of A. As $\dim P = 2$, (m_1, m_2) is therefore a basis of P.

($\supseteq$) Let $x = \alpha m_1 + \beta m_2$ with $\alpha, \beta \geq 0$. Then $x \in P$, so $\langle v_i - v_j, x\rangle = 0$, and by (19), for every k, $\langle v_i - v_k, x\rangle = \alpha\langle m_1, v_i - v_k\rangle + \beta\langle m_2, v_i - v_k\rangle \geq 0$. Hence $x \in \Sigma_{ij}$.

($\subseteq$) Let $x \in \Sigma_{ij}$. By Lemma 1(i), $x \in P$, so $x = \alpha m_1 + \beta m_2$ for some $\alpha, \beta \in \mathbb{R}$. Since $F_1 \cap F_2 = [v_i, v_j]$ and F_1 has at least three vertices, there exists a vertex $v_k \in F_1 \setminus F_2$. By (19), $\langle m_1, v_i - v_k\rangle = 0$ and $\langle m_2, v_i - v_k\rangle > 0$, so $0 \leq \langle v_i - v_k, x\rangle = \beta\,\langle m_2, v_i - v_k\rangle$ yields $\beta \geq 0$. Exchanging the roles of F_1 and F_2, a vertex $v_\ell \in F_2 \setminus F_1$ yields $\alpha \geq 0$. □

Triangulated meshes. In practice, meshes are triangulated, so a facet of A may be split into several coplanar triangles. An edge shared by two coplanar triangles is not an edge of A; by Lemma 1(ii), it does not contribute to (15) and can be skipped.

Jump and normal. Across Σ_{ij}, the gradient jumps by $[\nabla\sigma_A]_{ij} = v_i - v_j$, and since $\Sigma_{ij} \subset (v_i - v_j)^\perp$, the unit normal to Σ_{ij} is $n_{ij} = (v_i - v_j)/\|v_i - v_j\|$.

Evaluation of $I_{ij}(x)$. We parameterize Σ_{ij} in polar coordinates: let (e_1, e_2) be the orthonormal basis of $P = (v_i - v_j)^\perp$ obtained by Gram–Schmidt on (m_1, m_2), and $\theta_{ij} = \arccos\langle m_1, m_2\rangle$. Writing $z = s\,u(\theta)$ with $u(\theta) = \cos\theta\, e_1 + \sin\theta\, e_2$, $s > 0$, $\theta \in [0, \theta_{ij}]$:

$$I_{ij}(x) = \int_0^{\theta_{ij}} \int_0^{\infty} p_\sigma(s\,u(\theta) - x)\, s\, ds\, d\theta. \tag{20}$$

Radial integral (closed form). Decomposing $\|su - x\|^2 = (s - a)^2 + r^2$ with $a = u^\top x$, $r^2 = \|x\|^2 - a^2$:

$$\int_0^{\infty} p_\sigma(su - x)\, s\, ds = (2\pi\sigma^2)^{-3/2}\, e^{-r^2/(2\sigma^2)} \Big[\sigma^2 e^{-a^2/(2\sigma^2)} + a\sigma\sqrt{2\pi}\,\Phi(a/\sigma)\Big], \tag{21}$$

where Φ is the standard normal CDF.

Angular integral (numerical quadrature). The remaining one-dimensional integral over $\theta \in [0, \theta_{ij}]$ does not admit a closed form in general. We evaluate it numerically using Gauss–Legendre quadrature, which achieves high accuracy for smooth integrands with a small number of quadrature points.

B Proofs for Section 3.2

Setting. We write $\omega = \sigma u$ with $u \sim \mathcal{N}(0, I_n)$, and assume that f and ∇f have at most polynomial growth, so that all expectations below are finite. Let $x \in \mathbb{R}^n$ be a fixed point. For a random vector X, $\|X\|_2 := \mathbb{E}[\|X\|^2]^{1/2}$ and $\mathrm{Var}(X) := \mathbb{E}\|X - \mathbb{E}X\|^2$. We use the antithetic first- and zeroth-order samples

$$a(\omega) := \frac{\nabla f(x + \omega) + \nabla f(x - \omega)}{2}, \qquad z(\omega) := \frac{\omega\,\big(f(x + \omega) - f(x - \omega)\big)}{2\sigma^2}, \tag{22}$$

with $a_k := a(\omega_k)$ and $z_k := z(\omega_k)$, so that $\widehat{g}^{\mathrm{smooth}}$ and the antithetic ZoBG estimator are the means of the a_k and of the z_k. We write $\widehat{g}_{-k}$ and $\widehat{g}_{-j,-k}$ for the means of the a_i over $i \neq k$ and over $i \notin \{j, k\}$. Finally, $\mathbb{E}[uu^\top] = I_n$, and for any random vector w independent of u,

$$\big\|(I_n - uu^\top)\, w\big\|_2 = \sqrt{n + 1}\, \|w\|_2. \tag{23}$$

Proposition 1 (Unbiasedness)**.** *For every $\sigma > 0$ and $N \geq 2$, $\mathbb{E}\big[\widehat{\nabla}_{\mathrm{CFoBG}} f_p(x)\big] = \nabla f_p(x)$.*

Proof. Since each a_i appears in $N - 1$ leave-one-out means and $\omega_k\omega_k^\top/\sigma^2 = u_k u_k^\top$, (14) rewrites as

$$\widehat{\nabla}_{\mathrm{CFoBG}} f_p(x) = \frac{1}{N}\sum_{k=1}^{N} Y_k, \qquad Y_k := z_k + (I_n - u_k u_k^\top)\,\widehat{g}_{-k}. \tag{24}$$

By the ZoBG identity (3) and the symmetry of p, $\mathbb{E}[z(\omega)] = \nabla f_p(x)$. Since $\widehat{g}_{-k}$ is independent of u_k and $\mathbb{E}[I_n - uu^\top] = 0$, we get $\mathbb{E}[Y_k] = \nabla f_p(x)$ conditionally on $(\omega_i)_{i \neq k}$, hence the result. □

Variance. We now bound the variance of CFoBG. If the Y_k in (24) were i.i.d., this variance would simply be $\mathrm{Var}(Y_k)/N$. This is not quite the case: although each Y_k is unbiased, for $j \neq k$ the sample Y_j depends on ω_k through $\widehat{g}_{-j}$, so Y_j and Y_k are correlated. The following lemma shows that these correlations are weak, because ω_k enters $\widehat{g}_{-j}$ with weight $1/(N - 1)$: the variance still decreases as $1/N$, with a slightly larger constant.

Lemma 3. *Let $N \geq 3$, $\mathbb{E}[\alpha(\omega)] = 0$, and $V_k := \alpha(\omega_k) + (I_n - u_k u_k^\top)\,\widehat{g}_{-k}$. With $v := \|V_k\|_2$ and $\delta := \|a_k - \widehat{g}_{-j,-k}\|_2$,*

$$\mathbb{E}\Big\|\frac{1}{N}\sum_k V_k\Big\|^2 \leq \frac{v^2 + \sqrt{n + 1}\, v\, \delta}{N}. \tag{25}$$

Proof. Expanding the square, the diagonal terms give v^2/N, and it remains to bound the correlations $\mathbb{E}[V_k^\top V_j]$ for $j \neq k$. Let $\widetilde{V}_j$ be V_j with $\widehat{g}_{-j}$ replaced by $\widehat{g}_{-j,-k}$, so that $\widetilde{V}_j$ no longer depends on ω_k. As in Proposition 1, V_k has zero mean given $(\omega_i)_{i\neq k}$, hence $\mathbb{E}[V_k^\top \widetilde{V}_j] = 0$. The remainder $V_j - \widetilde{V}_j = (I_n - u_j u_j^\top)(a_k - \widehat{g}_{-j,-k})/(N-1)$ has norm $\sqrt{n+1}\,\delta/(N-1)$ by (23), and Cauchy–Schwarz concludes. □

We apply this lemma away from discontinuities, where CFoBG improves on ZoBG, and on a discontinuity, where both estimators behave alike.

Proposition 2 (Variance away from discontinuities)**.** *Assume that f is $\mathcal{C}^3$ on a ball $B(x,d)$ that meets no discontinuity surface Σ_i, and $N \geq 3$. Then, as $\sigma \to 0$,*

$$\mathrm{Var}\big(\widehat{\nabla}^{\mathrm{anti}}_{\mathrm{ZoBG}} f_p(x)\big) = \frac{(n+1)\,\|\nabla f(x)\|^2 + O(\sigma^2)}{N}, \qquad \mathrm{Var}\big(\widehat{\nabla}_{\mathrm{CFoBG}} f_p(x)\big) = O\Big(\frac{\sigma^4}{N}\Big). \quad (26)$$

Proof. Step 1: local approximations. We first compare a and z with their values for the linearization of f at x. When $\|\omega\| < d$, a Taylor expansion of ∇f gives $\nabla f(x \pm \omega) = \nabla f(x) \pm D^2 f(x)\,\omega + O(\|\omega\|^2)$; the first-order terms cancel in a, so $a(\omega) - \nabla f(x) = O(\sigma^2\|u\|^2)$. Similarly, expanding f to third order, the even-order terms cancel in $f(x+\omega) - f(x-\omega) = 2\,\omega^\top \nabla f(x) + O(\|\omega\|^3)$, so $z(\omega) - uu^\top \nabla f(x) = O(\sigma^2\|u\|^4)$. The event $\|\omega\| \geq d$ has probability $O(e^{-d^2/(4\sigma^2)})$, so by polynomial growth its contribution is negligible. Hence, in $\|\cdot\|_2$,

$$a(\omega) = \nabla f(x) + O(\sigma^2), \qquad z(\omega) = uu^\top \nabla f(x) + O(\sigma^2). \quad (27)$$

Taking expectations in the second estimate, with $\mathbb{E}[z] = \nabla f_p(x)$ and $\mathbb{E}[uu^\top] = I_n$, also gives $\nabla f_p(x) = \nabla f(x) + O(\sigma^2)$.

Step 2: ZoBG. The z_k are i.i.d., so the variance is $\mathbb{E}\|z - \nabla f_p(x)\|^2/N$. By Step 1, $z - \nabla f_p(x) = (uu^\top - I_n)\,\nabla f(x) + O(\sigma^2)$, and the first term has squared norm $(n+1)\,\|\nabla f(x)\|^2$ by (23).

Step 3: CFoBG. We apply Lemma 3 with $\alpha := z - \nabla f_p(x)$, which has zero mean, so that $V_k = Y_k - \nabla f_p(x)$ and the error of CFoBG is the mean of the V_k. It remains to show that v and δ are $O(\sigma^2)$. Substituting the estimates of Step 1 for z_k and $\nabla f_p(x)$ in (24) gives

$$V_k = (I_n - u_k u_k^\top)\big(\widehat{g}_{-k} - \nabla f(x)\big) + O(\sigma^2). \quad (28)$$

Here $\widehat{g}_{-k} - \nabla f(x)$ is a mean of terms $a_i - \nabla f(x) = O(\sigma^2)$, and it is independent of u_k, so the first term is $O(\sigma^2)$ by (23). Hence $v = O(\sigma^2)$. Likewise, $\delta \leq \|a_k - \nabla f(x)\|_2 + \|\widehat{g}_{-j,-k} - \nabla f(x)\|_2 = O(\sigma^2)$, and Lemma 3 concludes. □

Proposition 3 (Variance at a discontinuity)**.** *Let x lie on a single $\mathcal{C}^1$ discontinuity surface with unit normal $\vec{n}$, across which f jumps by J in the direction of $\vec{n}$, and let $N \geq 3$. Then, as $\sigma \to 0$, $N\sigma^2$ times the variance of both the antithetic ZoBG and the CFoBG estimators converges to $\kappa_n J^2$, with $\kappa_n := \frac{n}{4} - \frac{1}{2\pi} > 0$.*

Proof. ZoBG. For almost every u, $x+\sigma u$ and $x-\sigma u$ lie on opposite sides of the surface for small σ, so $\sigma z(\sigma u) \to \frac{J}{2}\,\mathrm{sgn}(\vec{n}^\top u)\,u$. By dominated convergence, $\mathbb{E}\|\sigma z\|^2 \to nJ^2/4$ and $\sigma\,\mathbb{E}[z] \to J\vec{n}/\sqrt{2\pi}$, so $N\sigma^2\,\mathrm{Var} \to \kappa_n J^2$.

CFoBG. By (24), CFoBG equals ZoBG plus $B := \frac{1}{N}\sum_k (I_n - u_k u_k^\top)\,\widehat{g}_{-k}$, and Lemma 3 with $\alpha = 0$ gives $\mathbb{E}[B] = 0$ and $\|B\|_2 = O(1/\sqrt{N})$ uniformly in σ. The two standard deviations thus differ by $O(1/\sqrt{N})$, which vanishes after multiplication by $\sigma\sqrt{N}$. □

Fallback when the gradient vanishes. Finally, if ∇f vanishes almost everywhere, then $\widehat{g}_{-k} = 0$ and $Y_k = z_k$ in (24): CFoBG coincides with the antithetic ZoBG estimator.

C Bias of the AoBG estimator

Figure 7 illustrates the empirical relative bias of the AoBG estimator from Suh et al. [2022b] around a given discontinuity point on several scenarios, as a function of σ, with several values of its hyperparameter γ. The bias decreases with σ and reaches the noise level at $\approx 10^{-3}$. As γ decreases, the bias decreases as AoBG converges to the zero-order estimator (ZoBG).

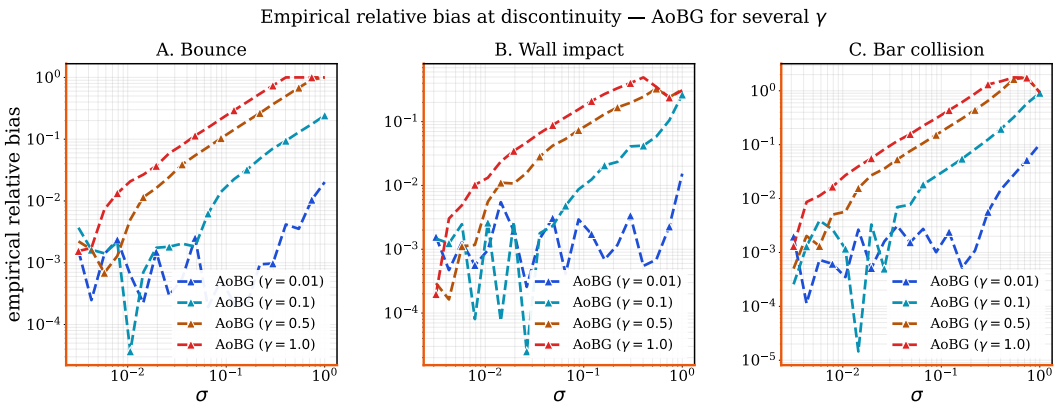

Figure 7: Empirical relative bias at a discontinuity point for AoBG with several values of γ.

D Variance results from Section 4.2

Figures 8 and 9 show the mean variance of the control Jacobian estimator on the *Jump* and *Land & Jump* tasks from Section 4.2, as functions of the smoothing parameter σ and number of antithetic sample pairs N respectively.

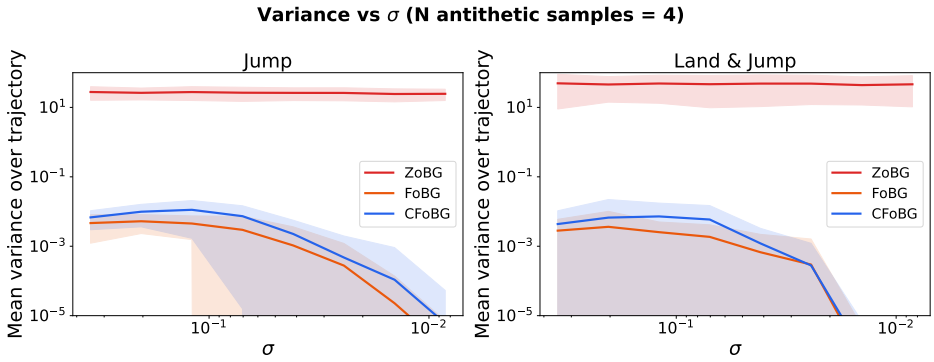

Figure 8: Mean variance of the control Jacobian estimator, averaged over trajectory steps, as a function of the smoothing parameter σ on the MuJoCo hopper *Jump* (left) and *Land & Jump* (right) tasks. Each estimator uses $N = 4$ antithetic sample pairs. Shaded bands denote ± 1 standard deviation across trajectory steps. ZoBG (red) exhibits σ-independent variance, consistent with its $\mathcal{O}(\|\nabla f\|^2/N)$ rate. The variances of the FoBG (orange) and CFoBG (blue) estimators decrease as $\sigma \to 0$, reflecting the $\mathcal{O}(\sigma^4/N)$ smooth regime.

Figure 10 shows the mean variance of the control Jacobian on the *Platformer* task. The major variance contribution here is the discontinuity of f with respect to u. Hence, the ZoBG and CFoBG estimators are virtually the same and have the same variance. FoBG has low variance but is biased, which leads to inefficient optimization.

E Additional benchmarks

We use three tasks from the benchmark suite of Suh et al. [2022b]: planar pushing with a soft-contact model with low contact stiffness; Coulomb friction with stick-slip transitions; and Tennis, a policy-optimization problem. We follow their hyperparameters and protocol and use antithetic sampling for all estimators. Table 2 reports the final cost, as the median and interquartile range over 100 seeds.

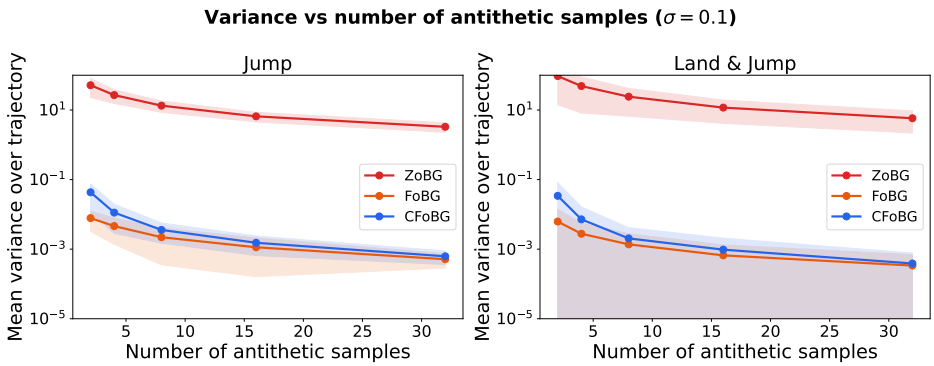

Figure 9: Mean variance of the control Jacobian estimator as a function of the number of antithetic sample pairs N, at $\sigma = 0.1$, on the *Jump* (left) and *Land & Jump* (right) hopper tasks. The FoBG and CFoBG estimators achieve variance 3 to 4 orders of magnitude lower than ZoBG at all values of N. In other words, ZoBG would require on the order of 10^3 to 10^4 samples to match the variance attained by FoBG or CFoBG with only $N = 4$.

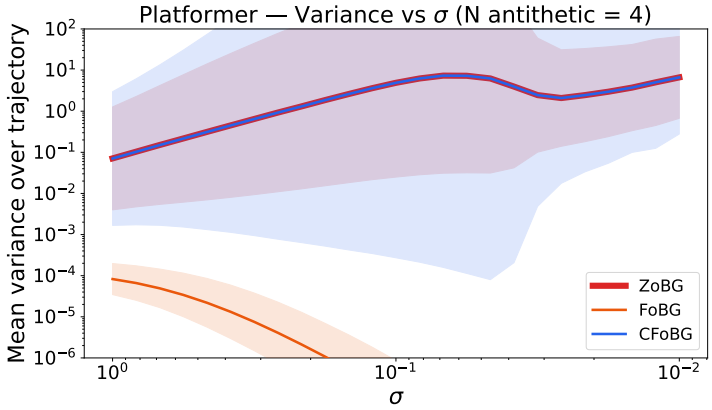

Figure 10: Mean variance of the control Jacobian estimator vs σ on the *Platformer* task ($N = 4$ antithetic samples).

Table 2: Final cost on tasks from Suh et al. [2022b]: median [Q25, Q75] over 100 seeds.

Task	ZoBG	FoBG	AoBG	CFoBG (ours)
Tennis	**3.211** [3.208, 3.213]	5.576 [5.576, 5.576]	**3.211** [3.208, 3.213]	**3.211** [3.208, 3.213]
Friction	8.858 [8.855, 8.860]	9.081 [9.080, 9.086]	8.858 [8.855, 8.860]	**8.853** [8.850, 8.857]
Pushing (soft ; N=4)	7.064 [7.060, 7.070]	**7.028** [7.028, 7.028]	7.064 [7.060, 7.070]	**7.028** [7.028, 7.028]
Pushing (soft ; N=100)	**7.029** [7.029, 7.029]	**7.028** [7.028, 7.028]	**7.029** [7.029, 7.029]	**7.028** [7.028, 7.028]

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction accurately reflect the paper's contributions. The bias of standard first-order estimators is formally established (Theorem 1) and empirically confirmed (Table 1). CFoBG is proven unbiased (Proposition 1), with a variance reduction of 3–4 orders of magnitude over ZoBG in the smooth regime (Proposition 2, Table 1, Fig. 2). QFoBG is validated on GraspNet meshes (Fig. 3).

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Limitations are discussed at the end of Section 5, each with a direction for future work: CFoBG does not reduce variance at discontinuities, where it has the same asymptotic variance as ZoBG (Proposition 3); using Theorem 1 directly, as QFoBG does, requires enumerable discontinuity surfaces, which is likely intractable for full simulators; and, like all randomized smoothing estimators, CFoBG degrades in high dimension.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best

judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: All theoretical results include explicit assumptions and complete proofs. Theorem 1 follows from classical distributional derivative calculus, with proof delegated to Rudin [1974] and Golse [2010], who establish it in a more general setting. The structure of the discontinuity surfaces used by QFoBG is proved in Appendix A. The unbiasedness and variance of CFoBG and ZoBG, away from and at discontinuities, are proved in Appendix B.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The estimators (CFoBG, QFoBG, ZoBG, FoBG) are fully described in Sections 2–3, with details sufficient for reimplementation. Experimental settings are specified: tasks (Platformer, Hopper), simulator (MuJoCo), annealing schedule ($\sigma : 1 \to 0.01$), number of samples ($N = 4$ antithetic), number of seeds, and evaluation metrics (AUC, relative Frobenius error). The additional benchmarks follow the hyperparameters and protocol of Suh et al. [2022b] (Appendix E). The GraspNet dataset is publicly available.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example

(a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
(b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
(c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: The code is not publicly available for institutional reasons. All estimators are described in full mathematical detail in Sections 2–3 and Appendices A and B, and all experimental settings are given in Section 4 and Appendix E, which is sufficient to reproduce the results independently.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Experimental details are specified in Section 4: tasks (Platformer, Hopper, GraspNet), simulator (MuJoCo), smoothing schedule ($\sigma : 1 \to 0.01$), number of samples ($N = 4$ antithetic), optimizer (DDP), and evaluation metrics (variance AUC, relative Frobenius error). Quadrature details for QFoBG are given in Appendix A.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.

- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Figure 3 reports error bars on the relative error of ZoBG. Appendix D provides variance plots for all estimators across tasks and smoothing levels. For trajectory optimization, Table 1 reports median performance over 400 random seeds, and Table 2 reports the median and interquartile range over 100 seeds. Variance is measured empirically over 400 independent trials per time step and summarized via AUC (Section 4.2).

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Compute resources are specified where relevant: wall-clock times for QFoBG and ZoBG are reported on a single CPU core (Section 4.1 and Figure 3). The remaining experiments (trajectory optimization) are lightweight and do not require significant compute resources beyond a standard CPU.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

Answer: [Yes]

Justification: This work presents theoretical and algorithmic contributions to gradient estimation for differentiable simulation. It raises no ethical concerns regarding data, privacy, or societal impact.

Guidelines:

- The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [N/A]

Justification: This paper presents foundational algorithmic and theoretical contributions to gradient estimation for differentiable simulation. It has no direct societal impact or foreseeable negative applications.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: This paper poses no such risks. No pre-trained models, image generators, or scraped datasets are released.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.

- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: All assets are properly credited. MuJoCo [Todorov et al., 2012] and the GraspNet dataset [Fang et al., 2020] are cited. The benchmark tasks from Suh et al. [2022b] are also properly referenced.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or research with human subjects.

    Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs are not used as part of the core methodology of this paper.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.