Variance-Reduced MPPI: Workshop Version

Fabian Schramm[1], Franki Nguimatsia Tiofack[1], Nicolas Perrin-Gilbert[2], Marc Toussaint[3] and Justin Carpentier[1]
[1]Inria and DI-ENS, PSL Research University [2]Sorbonne University [3]TU Berlin

***Abstract*—Sampling-based controllers like Model Predictive Path Integral (MPPI) offer substantial flexibility but suffer from high variance and low sample efficiency. To address this, we introduce a variance-reduced MPPI framework that decomposes the objective into an approximate model and a residual term. By adopting a quadratic approximation, we derive a closed-form, model-guided prior that concentrates samples in informative regions. Crucially, this framework is agnostic to the source of geometric information, accommodating exact derivatives, structural approximations (e.g., Gauss-Newton), or gradient-free randomized smoothing. We validate our approach on optimization benchmarks, underactuated cart-pole control, and contact-rich manipulation and show that it achieves faster convergence in low-sample regimes compared to standard MPPI.**

I. Introduction

Formulating control and motion planning as optimization problems is a standard paradigm in robotics. Elevating this problem to the *space of probability measures* has unified several frameworks in stochastic control and sampling-based planning. A prominent approach is the Model Predictive Path Integral (MPPI) [1, 2] algorithm. MPPI formulates the optimal control distribution as a KL-regularized Boltzmann distribution over trajectories. This offers compelling practical advantages: it is derivative-free, accommodates non-smooth costs (e.g., contact events), and parallelizes efficiently on modern hardware accelerators. However, MPPI approximates the optimal distribution via Monte Carlo sampling, introducing a central challenge: high estimator variance. Standard isotropic sampling fails to adapt to landscape geometry, resulting in poor sample efficiency. While recent work incorporates curvature information via Gauss–Newton updates [3], such approaches often remain in a high-sample regime, relying on thousands of parallel rollouts to achieve stability.

While variance reduction and covariance adaptation are well-studied in broader stochastic optimization (e.g., CMA-ES [4], CEM [5], importance sampling [6], control variates [7], RL [8, 9]), real-time control imposes strict computational budgets that force algorithms to operate in a low-sample regime. This raises an important question: *Can we retain the flexibility of sampling-based control while achieving faster convergence with significantly fewer samples?*

We introduce a hybrid approach that guides sampling using a model of the objective function, accelerating the convergence of MPPI-based methods. Our main contributions are:

1) We derive a variance-reduced MPPI framework by decomposing the objective into a known model and a residual term.
2) We propose a specific instantiation using a quadratic model approximation, deriving a closed-form, model-guided proposal distribution.
3) We demonstrate that this model can be computed via exact derivatives, structural approximations, or gradient-free randomized smoothing.
4) We validate the approach numerically, showing superior convergence in low-sample regimes.

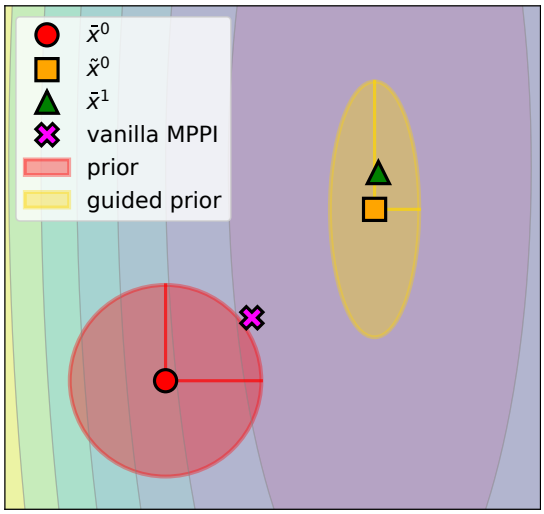


Fig. 1: **Illustration of covariance adaptation via Newton-like approximation.** Model-guided MPPI and vanilla MPPI start at the same state (circle) and use 100 samples from an isotropic Gaussian prior. The cross indicates the vanilla MPPI update. Our method exploits the gradient and Hessian at $\bar{x}^0$ to construct a quadratic model. This leads to a guided prior centered at $\tilde{x}^0$ (rectangle), concentrating the sampling distribution (yellow) along the valley and enabling an update $\bar{x}^1$ (triangle) by sampling the residual toward the optimum.

II. Preliminaries

We consider optimization problems of the form:

$$x^* \in \arg\min_{x \in \Omega} f(x), \tag{1}$$

where $f : \Omega \to \mathbb{R}$ denotes the trajectory cost objective and Ω represents the feasible set of control sequences. From a measure-theoretic perspective, the problem (1) can be reformulated as an equivalent problem over measures [10]. The

Function	Model-Guided MPPI	Vanilla MPPI	CMA-ES
Rosenbrock	6.9 ± 2.5 —	21.4 ± 7.3 —	12.0 ± 1.3 —
Styblinski–Tang	2.1 ± 0.4 —	9.2 ± 0.8 (14)	7.8 ± 2.9 —
Rastrigin	3.0 ± 1.1 —	6.9 ± 6.4 —	5.5 ± 4.6 —
Ackley	3.3 ± 1.6 (1)	3.8 ± 2.1 —	3.5 ± 1.4 —

TABLE I: Mean iterations $\pm$ std (failures) over 100 seeds.

goal is to find a probability distribution μ^*, minimizing the expected loss over distributions whose support lies in Ω:

$$\mu^* \in \underset{\substack{\mu \\ \text{supp}(\mu)\subseteq\Omega}}{\arg\min}\ \mathcal{J}(\mu) \quad \text{with} \quad \mathcal{J}(\mu) = \int_\Omega f(x)\mu(x)\,dx. \quad (2)$$

Depending on the setup, the objective function f can be convex, smooth, or arbitrarily complicated (non-convex, non-smooth). For instance, if f is strictly convex with a unique global minimum, working on (2) will converge to the optimal solution, which in the limit is a Dirac $\mu^*(x) \to \delta(x - x^*)$ concentrating the distribution mass entirely on the global optimal solution x^*. While Problem (2) theoretically allows for finding global optima [10], it is an infinite-dimensional problem. As common practice, we avoid solving for the global measure directly and instead adopt a sequential optimization strategy. At each step k, we seek a new distribution μ^{k+1} that minimizes the original objective $\mathcal{J}(\mu)$ in (2) subject to a proximity penalization to stay close to the previous estimate μ^k:

$$\mu^{k+1} = \underset{\mu}{\arg\min}\Big(\mathcal{J}(\mu) + \lambda \cdot \text{KL}(\mu \,||\, \mu^k)\Big). \quad (3)$$

The analytical solution to this problem yields the closed-form Boltzmann update:

$$\mu^*(x) \propto \mu^k(x) \cdot \exp\left(-\tfrac{1}{\lambda} f(x)\right). \quad (4)$$

While μ^* is the optimal update, it is in general complex and difficult to sample from directly.

To achieve tractability, a common strategy is to constrain μ to a parametric family of distributions, typically a Gaussian $p_\theta = \mathcal{N}(\bar{x}, \Sigma)$, where $\bar{x}$ denotes the mean and Σ the covariance. The objective then becomes to identify the finite-dimensional parameters θ such that p_θ best approximates the optimal Boltzmann distribution μ^* in (4). This particular instantiation of sequential KL control with a Gaussian parametrization underlies sampling-based algorithms such as Model Predictive Path Integral (MPPI) control [11, 1], which have gained significant traction in the robotics community in recent years [12, 13, 14, 15]. These approaches typically rely solely on zeroth-order information (function evaluations). While such derivative-free formulations naturally handle non-smooth dynamics, they do not exploit available gradient or curvature information. As a result, these sampling-based methods often exhibit high variance and low sample efficiency in complex cost landscapes, in contrast to classical optimization techniques — such as Newton-based methods — that can achieve second-order convergence rates near a local optimum.

In the following section, we address these limitations by embedding geometric information directly into the MPPI update. By decomposing the objective into an approximate model and a residual term, we enable the sampling process to exploit local curvature, bridging the gap between sampling-based flexibility and Newton-like convergence speeds.

III. Methodology

A. Model-guided distribution update

Consider the problem of finding the optimal distribution $p^*(x)$ at iteration $k+1$, given a prior estimate $p_{\theta^k}(x)$ obtained at iteration k. Following 4, the information-theoretic optimal update takes the Boltzmann form:

$$p^*(x) \propto p_{\theta^k}(x) \exp\left(-f(x)/\lambda\right). \quad (5)$$

At iteration k, we decompose the objective function f into a model m^k and a residual term r^k, such that $f(x) = m^k(x) + r^k(x)$. Here, m^k acts as a control variate intended to capture the dominant landscape geometry, while r^k accounts for unmodeled discrepancies. Substituting this decomposition into (5) yields the following factorization:

$$p^*(x) \propto \underbrace{\left[p_{\theta^k}(x) \exp\left(-m^k(x)/\lambda\right)\right]}_{\text{model-guided prior } \tilde{p}_{\theta^k}(x)} \cdot \exp\left(-r^k(x)/\lambda\right). \quad (6)$$

Equation (6) reveals a fundamental separation of concerns. The term in brackets, called model-guided prior, combines the previous prior with the explicit model to form a new intermediate distribution, which we denote as the *model-guided prior* $\tilde{p}_{\theta^k}(x)$. Rather than sampling from an uninformative prior and weighing by the full complex objective f as done classically in MPPI-based approaches, we sample from the structurally informed $\tilde{p}_{\theta^k}$ and weight only by the residual r^k.

Importantly, this formulation is valid for *any* choice of model. This generality offers a powerful degree of freedom: we can design m^k such that the resulting model-guided prior $\tilde{p}_{\theta^k}$ admits a closed-form solution. By selecting a model structure compatible with the prior, we ensure that the new sampling distribution remains analytically tractable.

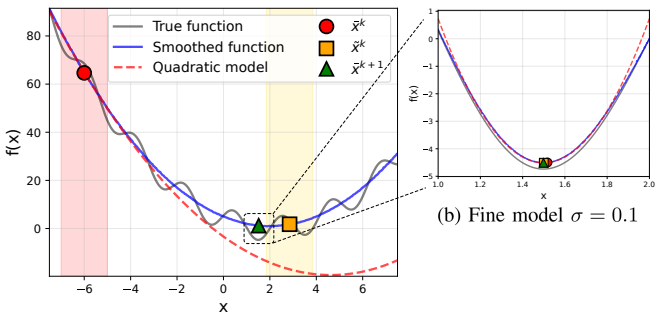
B. Closed-form model-guided prior via quadratic model approximation

To derive an efficient closed-form algorithm, we instantiate the general framework using Gaussian approximations. We assume the current estimate is a Gaussian distribution $p_{\theta^k}(x) = \mathcal{N}(x \,|\, \bar{x}^k, \Sigma^k)$ parametrized by $\theta^k = \{\bar{x}^k, \Sigma^k\}$. We choose m^k to be a quadratic expansion of the cost around the current mean $\bar{x}^k$:

$$m^k(x) = f(\bar{x}^k) + (g^k)^\top (x - \bar{x}^k) + \frac{1}{2}(x - \bar{x}^k)^\top H^k (x - \bar{x}^k), \quad (7)$$

where g^k and H^k denote the gradient and Hessian parameters at the expansion point, see Sec. III-C for details. Since the product of a Gaussian PDF and the exponential of a quadratic function is itself a Gaussian, the model-guided prior $\tilde{p}_{\theta^k}$ derived in (6) remains in the Gaussian family. We can therefore solve for its parameters $\tilde{\theta}^k = \{\tilde{x}^k, \tilde{\Sigma}^k\}$ analytically. The guided prior $\tilde{p}_{\theta^k}(x) = \mathcal{N}(x \,|\, \tilde{x}^k, \tilde{\Sigma}^k)$ is characterized by the updated covariance:

$$\tilde{\Sigma}^k = \left((\Sigma^k)^{-1} + \tfrac{1}{\lambda} H^k\right)^{-1}. \quad (8)$$



(a) Coarse model $\sigma = 1.0$

(b) Fine model $\sigma = 0.1$

Fig. 2: **Comparison coarse vs fine model.** (a) A large smoothing kernel $\sigma = 1.0$ suppresses sinusoidal disturbances (grey), yielding a coarse approximation (blue) capturing global geometry. The resulting quadratic model (red) generates a proposal $\tilde{x}^1$ moving toward the global minimum despite local nonconvexities. (b) Near the optimum, reducing the scale to $\sigma = 0.1$ captures the local curvature for final convergence.

To guarantee a valid positive-definite covariance matrix $\tilde{\Sigma}^k$, appropriate regularization or convexification is applied to H^k whenever it is indefinite. Furthermore, because incorporating positive-definite Hessian information strictly decreases variance, we apply an eigenvalue-based lower bound to $\tilde{\Sigma}^k$ and employ Polyak averaging across iterations to ensure stable exploration and prevent variance collapse. Consequently, the mean is given by:

$$\tilde{x}^k = \tilde{\Sigma}^k \left((\Sigma^k)^{-1} \bar{x}^k + \tfrac{1}{\lambda} \left(H^k \bar{x}^k - g^k \right) \right). \tag{9}$$

The final optimal distribution is then obtained by reweighting this model-guided Gaussian $\tilde{p}_{\theta^k}$ with the residual:

$$p^*(x) \propto \underbrace{\mathcal{N}(x \mid \tilde{x}^k, \tilde{\Sigma}^k)}_{\tilde{p}_{\theta^k}(x)} \cdot \exp\!\left(-r^k(x)/\lambda\right). \tag{10}$$

Interestingly, this reformulation naturally lends itself to applying MPPI on the residual function r^k while sampling trajectories $\{x^{(i)}\}_{i=1}^N$ from the model-guided prior $\tilde{p}_{\theta^k}$. The update is now given by:

$$\bar{x}^{k+1} \approx \sum_{i=1}^{N} \tilde{w}^{(i)}\, x^{(i)}, \text{ with } \tilde{w}^{(i)} = \frac{\exp\left(-r^k(x^{(i)})/\lambda\right)}{\sum_{j=1}^{N} \exp\left(-r^k(x^{(j)})/\lambda\right)}. \tag{11}$$

Compared to vanilla MPPI, which samples directly from p_θ^k, this model-guided approach fundamentally alters the update mechanism. First, the analytical shift to $\tilde{p}_{\theta^k}$ integrates geometry information, blending the prior mean $\bar{x}^k$ with a Newton-like step derived from the model m^k. Second, the sampling process is reserved for exploring the residual error r^k around this new center, rather than exploring the full cost f.

C. Quadratic model approximations

To construct the quadratic model (7), we require g^k and H^k. A key strength of our framework is that it decouples sampling from the source of this geometric information, supporting various approximation strategies depending on the problem structure.

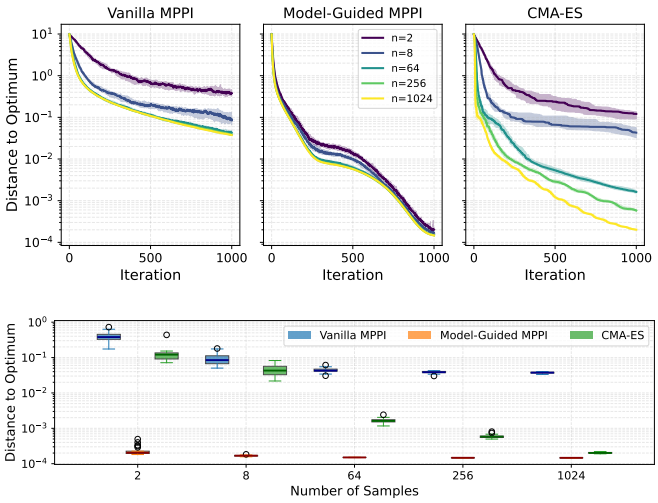


Fig. 3: **Cart-pole swing-up results.** Comparison across sample budgets $N \in \{2, \ldots, 1024\}$. **(Top)** Median distance to the reference optimum over iterations; shaded areas indicate the interquartile range (IQR). **(Bottom)** Final distance to the optimum after 1000 iterations. The log-scale highlights the severe variance and error increase for baselines when $N < 64$.

Gradient-based approximations. If f is differentiable, g^k and H^k can be exact analytical derivatives, acting as a precise local Newton step. For non-linear least-squares control objectives ($f(x) = \frac{1}{2}\|R(x)\|^2$), H^k can be efficiently approximated via the Gauss-Newton method as $H^k \approx J(x)^\top J(x)$. Alternatively, if second-order information is unavailable, Quasi-Newton methods like BFGS can estimate curvature using only a history of gradient updates.

Randomized smoothing (RS). For non-smooth or black-box functions, we estimate the model parameters via RS with a Gaussian kernel, defining the surrogate $f_\sigma(x) = \mathbb{E}_{z\sim\mathcal{N}(0,\sigma^2 I)}[f(x+z)]$. Exploiting Stein's identity and a centered control variate, we obtain the Monte Carlo estimators:

$$g^k \approx \frac{1}{M\sigma^2} \sum_{j=1}^{M} \left(f(\bar{x}^k + z_j) - f(\bar{x}^k) \right) z_j, \tag{12}$$

$$H^k \approx \frac{1}{M\sigma^4} \sum_{j=1}^{M} \left(f(\bar{x}^k + z_j) - f(\bar{x}^k) \right) (z_j z_j^\top - \sigma^2 I), \tag{13}$$

where $\{z_j\}_{j=1}^M \sim \mathcal{N}(0, \sigma^2 I)$. As illustrated in Fig. 2, modulating the noise scale σ allows us to transition from a *coarse model* that captures global structure to a *fine model* that captures precise local curvature.

IV. Experiments

We evaluate how guidance from a quadratic model affects the convergence, sample efficiency, and robustness of MPPI. As baselines, we consider vanilla MPPI and CMA-ES [16].

A. Illustration of quadratic model guidance

We first illustrate the mechanism on a 2D narrow-valley objective (Fig. 1). Second-order information reshapes the

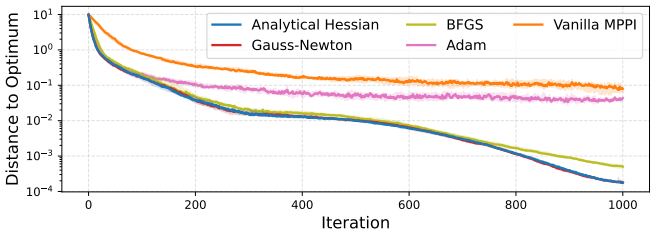


Fig. 4: **Hessian approximations (Cart-pole).** Comparison of dense vs. diagonal curvature models against the Vanilla MPPI baseline. Plot shows median distance to the reference optimum across 10 random seeds (shaded IQR).

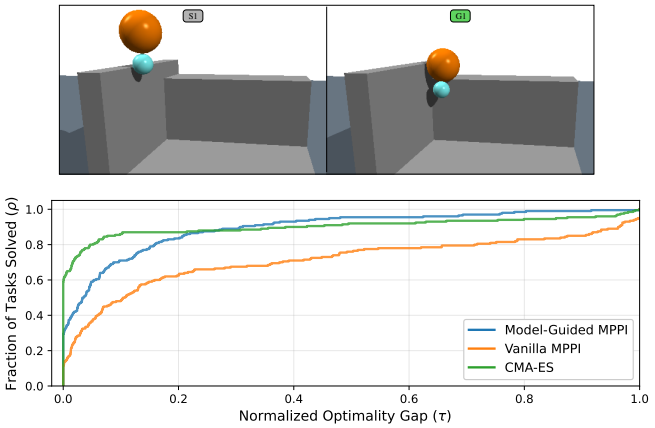


Fig. 5: **Single-Finger Sphere Manipulation. (Top)** Example start (left) and goal (right) states. The actuated finger (blue) must navigate to its target while manipulating the passive object (orange) to a specified goal pose. **(Bottom)** Performance profile comparing methods across 200 randomized tasks.

sampling distribution to align with the cost geometry, performing a Newton-like update for the distribution mean. Using 100 samples, vanilla MPPI collapses to an Effective Sample Size [17] (ESS) of 1.1, indicating severe weight degeneracy. In contrast, our model-guided prior maintains an ESS of 23.4, ensuring efficient state-space coverage.

B. Performance with analytical gradients

To isolate the benefits of the update rule from gradient estimation errors, we first test settings where exact analytical derivatives are available.

Static benchmarks. Tab. I reports the iterations required to reach the global optimum. Our framework consistently converges faster than the baselines and exhibits lower variance across 100 random seeds. For example, on the Rastrigin benchmark, we achieve a tight convergence distribution (3.0 ± 1.1), whereas Vanilla MPPI (6.9 ± 6.4) and CMA-ES (5.5 ± 4.6) display higher volatility.

Continuous control. We assess a nonlinear, underactuated cart-pole swing-up task using automatic differentiation derivatives via CasADi/Pinocchio. Fig. 3 evaluates convergence across sample budgets $N \in \{2, \dots, 1024\}$. In the high-sample regime ($N = 1024$), both CMA-ES and Model-Guided MPPI converge rapidly. However, below $N = 64$, baselines suffer from higher variance or instability. In contrast, our method maintains negligible variance and high-precision convergence even at $N = 2$. The analytical model successfully decouples optimization performance from the sampling budget.

Hessian approximations. Fig. 4 compares the exact local model against Gauss-Newton (GN), BFGS, and Adam-based (diagonal) approximations using $N = 8$. Dense structural approximations (GN, BFGS) achieve high-precision convergence that closely tracks the exact analytical Hessian.

C. Coarse and fine model approximations

Fig. 2 illustrates how modulating the smoothing scale σ balances exploration and precision on non-convex objectives. A coarse model (large σ) filters high-frequency noise to capture dominant global geometry, guiding the iterate toward the correct basin. A fine model (small σ) then resolves accurate local curvature for high-precision final convergence.

D. Randomized smoothing on non-smooth dynamics

To demonstrate versatility in contact-rich, non-differentiable scenarios, we apply the framework to a single-finger sphere manipulation task (Fig. 5) in MuJoCo. We use randomized smoothing with 128 auxiliary samples to extract the quadratic model guidance. We evaluate 200 diverse problem instances with challenging contact configurations, enforcing a strict planning budget of 64 trajectory samples per iteration.

As shown in the performance profile (Fig. 5), Model-Guided MPPI consistently achieves lower normalized optimality gaps across a large fraction of problem instances. This confirms that our formulation can reliably handle the hybrid dynamics inherent to rich-contact manipulation, outperforming purely sampling-based baselines in their higher variability.

V. Discussion and Conclusion

We presented a variance-reduced extension of MPPI based on a model–residual decomposition. By factoring the Boltzmann update into a model-guided prior and a residual correction, our method injects structural information into the sampling process. The resulting closed-form quadratic update can incorporate curvature information from any source, achieving superior sample efficiency, convergence speed, and robustness compared to vanilla MPPI and CMA-ES in low-sample regimes.

While more sample-efficient, our approach relies on the local validity of the quadratic approximation and introduces a computational trade-off. The overhead of constructing the local model must be weighed against the speedup gained from requiring fewer samples. This construction cost is almost negligible when exact derivatives are available, but can become larger when relying on zero-order techniques like randomized smoothing. Future work will explore adaptive smoothing strategies, the integration of learned data-driven priors and scaling the framework to high-dimensional dexterous manipulation tasks.

ACKNOWLEDGMENTS

This work has received support from the French government, managed by the National Research Agency, under the France 2030 program with the references Organic Robotics Program (PEPR O2R), "PR[AI]RIE-PSAI" (ANR-23-IACL-0008) and RODEO (ANR-24-CE23-5886). The European Union also supported this work through the ARTIFACT project (GA no.101165695) and the AGIMUS project (GA no.101070165). The Paris Île-de-France Région also supported this work in the frame of the DIM AI4IDF. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the funding agencies.

REFERENCES

[1] G. Williams, A. Aldrich, and E. A. Theodorou, "Model predictive path integral control: From theory to parallel computation," *Journal of Guidance Control and Dynamics*, vol. 40, pp. 344–357, 2017.

[2] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Information-theoretic model predictive control: Theory and applications to autonomous driving," *IEEE Transactions on Robotics*, vol. 34, pp. 1603–1622, 2017.

[3] H. Homburger, K. Baumgärtner, M. Diehl, and J. Reuter, "Gauss-newton accelerated mppi control," 2026.

[4] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evolutionary Computation*, vol. 9, no. 2, pp. 159–195, 2001.

[5] R. Y. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Methodology And Computing In Applied Probability*, vol. 1, pp. 127–190, 1999.

[6] R. Y. Rubinstein, "Simulation and the monte carlo method," in *Wiley series in probability and mathematical statistics*, 1981.

[7] P. Glasserman, "Monte carlo methods in financial engineering," 2003.

[8] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine Learning*, vol. 8, pp. 229–256, 2004.

[9] J. Peters and S. Schaal, "Policy gradient methods for robotics," *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2219–2225, 2006.

[10] J. B. Lasserre, "Global Optimization with Polynomials and the Problem of Moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[11] E. A. Theodorou and E. Todorov, "Relative entropy and free energy dualities: Connections to path integral and kl control," *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 1466–1473, 2012.

[12] T. Howell, N. Gileadi, S. Tunyasuvunakool, K. Zakka, T. Erez, and Y. Tassa, "Predictive sampling: Real-time behaviour synthesis with mujoco," 2022.

[13] G. Turrisi, V. Modugno, L. Amatucci, D. Kanoulas, and C. Semini, "On the Benefits of GPU Sample-Based Stochastic Predictive Controllers for Legged Locomotion," *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 13757–13764, 2024.

[14] H. Xue, C. Pan, Z. Yi, G. Qu, and G. Shi, "Full-order sampling-based mpc for torque-level locomotion control via diffusion-style annealing," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4974–4981, 2025.

[15] Y. Zhai, R. Reiter, and D. Scaramuzza, "Pa-mppi: Perception-aware model predictive path integral control for quadrotor navigation in unknown environments," *ArXiv*, vol. abs/2509.14978, 2025.

[16] N. Hansen, Y. Akimoto, and P. Baudis, "CMA-ES/pycma on Github." Zenodo, DOI:10.5281/zenodo.2559634, Feb. 2019.

[17] L. Martino, V. Elvira, and F. Louzada, "Effective sample size for importance sampling based on discrepancy measures," *Signal Processing*, vol. 131, pp. 386–401, 2017.