Distributionally Robust Control via Stein Variational Inference for Contact-Rich Manipulation

Author Names Omitted for Double-blind Review

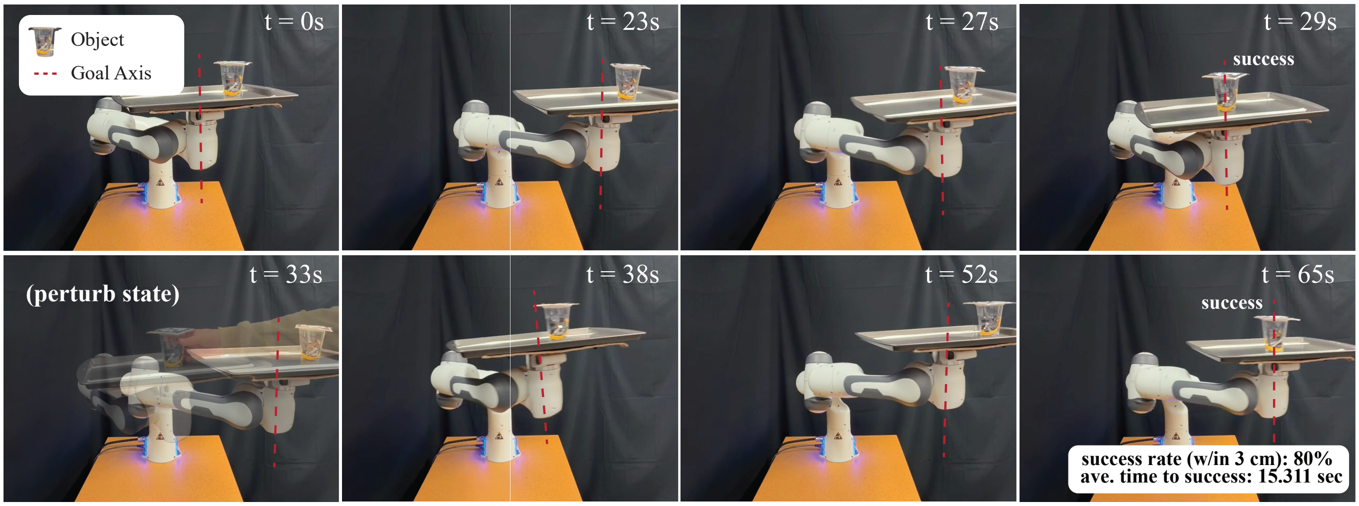

Fig. 1: Within-Hand Dynamics positioning of a cup with unknown mass distribution. SV-DRO reliably repositions a cup via within-hand sliding over continuous 3.5-minute trials, despite unknown mass distributions induced by objects inside the cup. The robot repeatedly repositions the cup upon reaching a goal tolerance (< 1.5 cm) without interrupting execution. The method achieves $\sim 80\%$ success at this tolerance (100% with relaxed tolerance) while persistently correcting deviations online.

***Abstract*—Reliable robotic manipulation requires control policies that accurately represent and adapt to uncertainty in contact-rich interactions. Modern data-driven methods rely on large-scale training to mitigate uncertainty but degrade under limited data, while classical model-based controllers are efficient yet struggle to capture task-relevant uncertainty. We address this gap by formulating manipulation as a distributionally robust control problem and introducing a deterministic Stein variational inference-based approach that explicitly models task-sensitive parameter uncertainty while preserving performance. The resulting controller adapts to task-critical uncertainty rather than reasoning conservatively over worst-case uncertainty, achieving high reliability without sacrificing efficiency, and demonstrate up to $3\times$ improved robustness across diverse contact-rich manipulation tasks under broad parametric uncertainty compared to existing model-based methods.**

I. Introduction

In-the-wild robotic manipulation requires operating under significant physical uncertainty, where reliable control hinges on the ability to exploit and adapt to uncertain interactions. Modern approaches address this through large-scale data and computation [8, 29, 4, 7, 5], effectively attempting to cover all possible scenarios at the expense of precision and interpretability, and degrading rapidly with limited samples [17, 6]. The central objective is to develop zero-shot manipulation controllers that adapt to physical uncertainty without sacrificing performance or robustness.

Prior work on uncertainty-aware manipulation spans data-driven and classical paradigms with distinct trade-offs in robustness [33, 12, 2]: data-driven methods [10, 24] provide adaptability but require extensive data, compute, and uncertainty coverage [30, 25, 26], ensemble approaches [1, 22, 23] optimize average performance but fail under limited samples and depend on assumed uncertainty ranges, and distributionally robust methods [20, 21, 28, 9] ensure guarantees via worst-case formulations [13] but are overly conservative in contact-rich settings. *Thus, we aim to extend model-based control through flexible uncertainty modeling that adapts to uncertainty while retaining manipulation performance.*

To achieve this, we formulate manipulation control as a distributionally robust optimization problem and introduce a deterministic framework based on Stein variational inference [27, 19]. Our approach extends conventional DRO by optimizing over a deterministic set of parameters evolved from a prior via Stein variational gradient descent (SVGD) [19]. This enables tractable, non-parametric approximation of task performance under parameter uncertainty, while remaining computationally efficient and parallelizable [32, 18]. By explicitly modeling how parameter variations impact task performance, the method performs exact inference over uncertainty-sensitive actions, which is critical in contact-rich manipulation.

Empirically, the proposed approach achieves optimal task performance under high parameter uncertainty with limited feedback, outperforming existing model-based baselines [1, 13]. Unlike worst-case formulations, our method optimizes for task sensitivity to parameter variations, prioritizing uncertainties that most affect performance and thereby improving expected outcomes. We validate this approach on contact-rich tasks including within-hand dynamic positioning and bi-manual manipulation, demonstrating enhanced robustness and efficiency.

II. Preliminaries

A. Distributionally Robust Control

We now formalize the control problem under parametric uncertainty using a distributionally robust perspective. In this setting, the objective is to synthesize a trajectory that minimizes a task-dependent cost while remaining robust to parameter uncertainty. To do so, we first consider the following optimization problem,

$$\min_{\tau=\{(x_j,u_j,c_j)\forall j\in[0,T]\}} \mathcal{J}_\theta(\tau) \quad \text{s.t} \begin{cases} x_0 \in \mathcal{X} & \text{(init. state)} \\ c_j \in \mathcal{C}_\theta & \text{(contact model)} \\ x_{j+1} = f_\theta(x_j, u_j) & \text{(dynamical model)} \end{cases} \tag{1}$$

where $\mathcal{J}_\theta(\cdot) : \mathcal{T} \times \Theta \to \mathbb{R}$, where $\mathcal{T} \subset \mathcal{X} \times \mathcal{U} \times \mathcal{C}_\theta$, is the task-relevant cost to minimize, $\tau = \{(x_j, u_j, c_j)\forall j \in [0, t_h]\}$ is a sequence of states $x \in \mathcal{X}$, control inputs $u \in \mathcal{U}$, contact impulses $c_j \in \mathcal{C}_\theta$, and $\theta \in \Theta$ are physics model parameters, where t_h is the planning horizon. Distributionally Robust Optimization (DRO) seeks to optimize over the Lagrangian via a bi-level optimization problem,

$$\min_{\tau} \max_{p\in\mathcal{P}} \mathbb{E}_{\theta\sim p(\theta)} \left[\mathcal{L}(\tau, \theta)\right] \tag{2}$$

where $\mathcal{L} : \mathcal{T} \times \Theta \to \mathbb{R}$ is the Lagrangian to the above optimization in Eq. 1, p is a surrogate distribution over θ, and $\mathcal{P}$ is the following ambiguity set,

$$\mathcal{P} = \{p \mid \mathbb{D}_{KL}\left[p\|q\right] \le \epsilon\} \tag{3}$$

where q is true unknown target distribution. Since finding an ambiguity set $\mathcal{P}$ is challenging to find a-priori, the above DRO problem is commonly approximated as shown in [20] as,

$$\min_{\tau,\lambda>0} \lambda\epsilon + \lambda \log \mathbb{E}_q \left[\exp\left(\frac{1}{\lambda}\mathcal{L}(\tau, \theta)\right)\right], \tag{4}$$

where $\lambda \in \mathbb{R}^{n_{eq}}$ is a dual auxiliary variable. Optimizing the second term is more commonly known as the Risk-Averse Optimal Control Problem [31, 11]. For sufficiently large λ, it holds true that,

$$\log \mathbb{E}_q \left[\exp\left(\frac{1}{\lambda}\mathcal{L}(\tau, \theta)\right)\right] \approx \mathbb{E}_q\big[\mathcal{L}(\tau, \theta)\big] + \frac{1}{2}\frac{1}{\lambda}\mathbb{V}_q\big[\mathcal{L}(\tau, \theta)\big] \tag{5}$$

Assumption 1. *(KL-ball ambiguity set)[20, 21] the surrogate distribution p exists in the ϵ-neighborhood of the target distribution, where $\mathbb{D}_{KL}(p\|q) \le \epsilon$ given small $\epsilon > 0$.*

To ensure the above assumption holds, prior work [20] estimates the target distribution via data-driven maximum-likelihood before performing k-nearest-neighbor sampling from the inferred model; however, this assumption is not guaranteed in general and is often violated in manipulation settings with high uncertainty where priors are insufficiently informative. Consequently, standard DRO becomes unreliable in such regimes, motivating our approach, which removes the need for Assumption 1 by directly and exactly characterizing uncertainty through a guided, task-aware variational inference framework.

B. Stein Variational Gradient Descent

Given a random variable $\xi \in \Xi$, variational inference seeks to minimize the Kullback-Leibler divergence between a surrogate $p(\xi)$ and target distribution $q(\xi)$. Stein Variational Gradient Descent (SVGD) [19] provides a non-parametric solution via kernelized gradient descent, initializing particles $\{\xi_0^t\}_{t=1}^N$ from a prior $p(\xi) = \mathcal{N}(\hat{\xi}, \Sigma_\xi)$ and evolving them through a deterministic mapping.

$$\xi_{s+1}^t \leftarrow \xi_s^t + \alpha\ \phi_{p,q}^*(\xi_s^t) \tag{6}$$

where $\phi_{p,q}^*(\xi_s^t)$ is a smooth function that captures the steepest descent direction that minimizes the KL-divergence measure at the s^{th} step, and $\alpha > 0$ is a sufficiently small perturbation magnitude. We consider $\phi_{p,q}^*$ to be the solution to the steepest descent problem shown in [27, 15, 3] and approximated via Monte-Carlo sampling over random variables ξ,

$$\phi_{q,p}^*(\cdot) \approx \frac{1}{N}\sum_{t=1}^{N} k(\xi_s^t, \cdot)\nabla_\xi \log p'(\xi_s^t) + \nabla_\xi k(\xi_s^t, \cdot) \tag{7}$$

where $k(\cdot)$ is a universal, positive definite kernel function that enforces particle repulsion to promote exploration over the approximated posterior $p'(\cdot)$. Over sufficient number of iterations of SVGD and number of samples N, the sampled parameters become a sufficient approximation of the target $q(\xi)$.

III. Stein Variational Distributionally Robust Control

In this section, we derive the main contribution of this work that formulates the Distributionally Robust Control problem via Stein variational inference.

A. Variational Approximation of Task-Dependent Parameter Posterior

We first formulate the parameter posterior distribution via SVI. Given a set of randomly drawn parameter points $\{\theta\}_{i=1}^N \sim p(\theta)$, where $p(\theta)$ is a prior distribution, we approximate the task-dependent parameter posterior for a given outcome $\mathcal{O}$ as,

$$q^*(\theta) = \arg\min_{q\in Q} D_{KL}\Big(q(\theta)\,||\,p(\mathcal{O}|\theta)\Big)$$

where $p(\mathcal{O}|\theta) = \exp\big(-\delta\mathcal{L}(\tau, \theta)\big)$ is the Boltzmann posterior of the task Lagrangian optimality gap modeled as $\delta\mathcal{L} = \mathcal{L}(\tau^*, \theta) - \mathcal{L}(\tau^*, \theta^*)$ with θ^* being the true parameter value. Because θ^* is not available to us in advance, we aim to approximate the true parameter through an empirical average

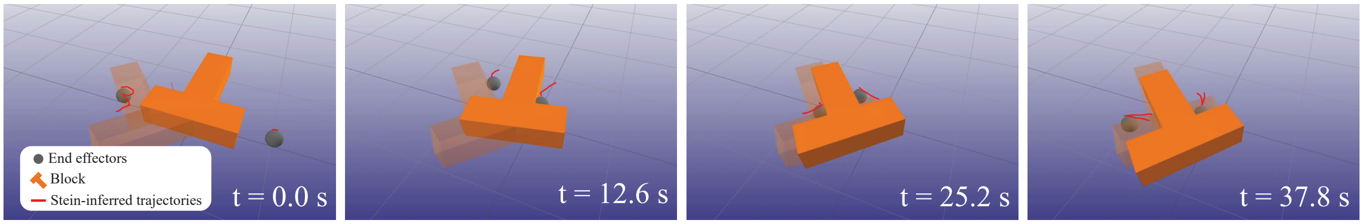

Fig. 2: Bimanual manipulation of Push-T with unknown mass distribution. Here, we demonstrate a control solution to the bimanual push-T task using our proposed SV-DRO. We observe that despite high uncertainty on mass, the bimanual end-effectors are able to successfully manipulate the T-block to the desired position while avoiding task-sensitive actions as a function of the uncertain parameters.

Method	(a) Bimanual Push-T Success % / Time to Completion (s) (≤ 10 **cm to goal)**	(≤ 1 **cm) to goal**	(b) Within-Hand dynamic positioning Success % / Time to Completion (s) (≤ 10 **cm to goal)**	(≤ 1 **cm) to goal**
SV-DRO (Ours)	**100 % / 14.69 $\pm$ 3.27 s**	**45 % / 15.52 $\pm$ 3.08 s**	**80 % / 8.93 $\pm$ 2.99 s**	**50 % / 8.95 $\pm$ 2.66 s**
EMPPI [1]	30 % / 16.26 $\pm$ 3.79 s	0 % / $\star$	40 % / 7.83 $\pm$ 0.93 s	20 % / 8.56 $\pm$ 0.92 s
MPC	80 % / 14.16 $\pm$ 4.32 s	15 % / 16.44 $\pm$ 3.87 s	40 % / 9.11 $\pm$ 1.42 s	5 % / 8.55 $\pm$ 0.28 s
DRO [13]	0 % / $\star$	0 % / $\star$	40 % / 10.567 $\pm$ 2.298	10 % / 11.438 $\pm$ 0.962

TABLE I: Simulated performance over 20 trials (mean $\pm$ std) for bimanual push-T and within-hand positioning under bounded uniform parameter priors (see supplementary); $\star$ denotes insufficient completions for timing statistics. SV-DRO significantly improves within-hand success despite high parameter uncertainty and limited state estimation.

over the sample points, $\frac{1}{N}\sum_{i=1}^{N}\theta_i$, which is commonly done in past Stein literature [15, 14]. Here, the likelihood model output $\mathcal{O}$ is an output that encodes model optimality variations with respect to the parameters of interest. The log posterior is then calculated as a parameter-conditioned optimality gap distribution,

$$\log p'(\theta) = \log p(\theta) - \delta\mathcal{L}(\tau,\theta) + \overset{\text{constant}}{\cancel{\log z}}. \quad (8)$$

We now seek to utilize the posterior distribution to evolve the j^{th} sampled set of estimators $\{\theta_i^j\}_{i=1}^N \forall j \in \mathbb{Z}^+$ drawn from a prior $p(\theta)$ by evaluating the steepest descent direction $\phi^* : \Theta \rightarrow \mathcal{S}_\Theta$ using Stein's identity. Given a prior $p(\theta)$ and a universal positive definite kernel $k : \Theta \times \Theta \rightarrow \mathbb{R}$, the resulting SGVD steepest descent solution to Eq. 7 is,

$$\phi^*(\cdot) \approx \frac{1}{N}\sum_{i=1}^{N} k(\theta_i^j,\cdot)\nabla_\theta \log p'(\theta_i^j) + \nabla_\theta k(\theta_i^j,\cdot) \quad (9)$$

where $k(\theta,\hat{\theta}) = \exp(-\|\theta-\hat{\theta}\|^2/h)$ is chosen as a Radius Basis Function (RBF) kernel [34].

Because the true physics parameters θ^* are unknown and inestimable, we extend existing DRO formulations [13, 20, 21] by diversifying parametric inference $\theta \sim p(\theta)$ through an evolving, task-aware surrogate posterior that explicitly adapts to uncertainty. Leveraging SVGD enables diverse, parameter-sensitive gradients across a broad range of θ, effectively reasoning over an ensemble of samples that accurately approximates the task-aware posterior. This focuses limited computation on task-critical uncertainties, yielding a robust controller that avoids the conservatism of worst-case methods while preserving the sample efficiency of model-based control.

B. Stein Variational DRO Formulation

Our goal in using Stein variational inference is to is to eliminate the restrictive Assumption 1 is not necessarily guaranteed for a weak prior $p(\theta)$. We therefore require a more

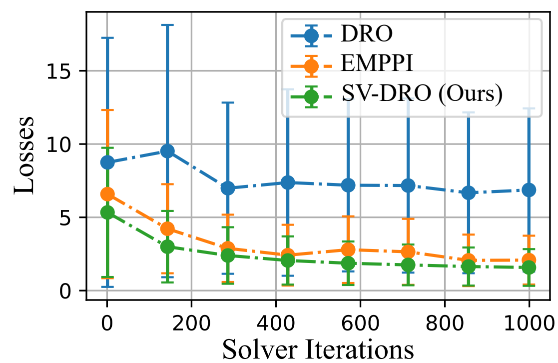

Fig. 3: Controller convergence based on prior distribution for Within-hand Positioning. The number of parameter samples, planning horizon, and prior distributions, are fixed. The results of this plot are evaluated over 5 trials per method with 5 random seeds. SV-DRO has improved reliability due to Stein inference.

robust method to optimize a control sequence u^* that reasons about *variations* of task sensitives given that θ^* is unknown. To do so, we adopt the Stein-Variational approach [19] in the DRO formulation on the perturbed set of evolving estimators $\{\theta_i^j\}_{i=1}^N \forall j \in \mathbb{Z}^+$.

We first define the evolving estimators $\hat{\theta}$ to follow the SVGD perturbation update,

$$\theta_{i+1}^j \leftarrow \theta_i^j + \alpha\phi^*(\theta_i^j) \quad (10)$$

where $\alpha > 0$ is the Stein evolve rate and kernel $k(\theta,\hat{\theta})$ is an RBF (commonly used in Stein Variational inference [14, 3]), and $\nabla_\theta \log p'(\theta)$ is the score of the log posterior in Eq. (8). The estimators are then implemented to solve the following optimization problem that extends Eq. 4,

$$\min_{\tau=\{(x_j,u_j,c_j)\forall j\in[0,t_h]\}} \mathcal{L}(\tau,\theta)\big|_{\theta=\bar{\theta}} + \gamma\,\frac{1}{N}\sum_{i=1}^{N}\delta\mathcal{L}(\tau,\theta_i^j) \quad (11)$$

where $\bar{\theta}$ is the empirically averaged parameter value given the current parameter points and γ is a design parameter that controls the trade-off between optimizing over the nominal parameters versus the variations of parameters. While Eq. 11 can be approximated via Eq. 5, we instead estimate the expected optimality gap using Monte Carlo sampling as in Eq. 11, evaluating task-aware parameter sensitivity gradients over diverse estimators $\{\theta_i^j\}_{t=1}^N$. This yields a more accurate representation of task-relevant sensitivities $\delta\mathcal{L}(\tau, \theta)$ and produces controllers that better mitigate the optimality gap; the full method is detailed in Alg. 1.

Algorithm 1 Stein Variational Distributional Robust Control (SV-DRO)

Require: planning horizon t_h, init. traj. $\tau = \{x_j, u_j, c_j\}_{j=1}^{t_h}$, prior $p(\theta)$, step size α, design param. γ, samples $\{\theta_i^0\}_{i=1}^N \sim p(\theta)$, parameterized dynamics $f_\theta(x_j, u_j)$, time duration T.

$j = 0$
while $j < T$ **do**
  $\tau^* \leftarrow \arg\min_\tau \mathcal{L}(\tau, \theta)\big|_{\theta=\bar{\theta}} + \gamma \ \frac{1}{N}\sum_{i=1}^N \delta\mathcal{L}(\tau, \theta_i^j)$
  Apply first control via MPC
  $x_{j+1} \leftarrow f_\theta(x_j, u^*[0])$ ▷ $\tau^* = \{(x^*_{0:t_h}, u^*_{0:t_h}, c^*_{0:t_h})\}$
  Update SVI
  $\{\theta_i^{j+1}\}_{i=1}^N \leftarrow \text{SVGD}(\{\theta_i^j\}_{i=1}^N)$ ▷ Eq. 10
  $j \leftarrow j + 1$
end while

IV. Results

We evaluate our method on two contact-rich manipulation tasks: (i) **Bimanual Push-T**, where a bimanual manipulator moves a T-shaped block with unknown mass and inertia to a goal pose via non-prehensile interactions, and (ii) **Dynamic Within-Hand Positioning**, where a robot repositions an object with unknown mass, inertia, and center of mass within a tray. Contact dynamics are modeled via a softplus/softmax Lagrangian formulation [16], with full implementation details provided in the supplementary material. We compare against EMPPI [1, 22], conventional DRO [13, 20, 21], and a naive MPC controller using the prior mean. All methods initialize parameters from a bounded uniform distribution; EMPPI and DRO resample parameters after each MPC update (Eq. 4), whereas our method samples once and subsequently evolves particles via SVGD.

Improved Performance Under Uncertainty. As shown in Table I, SV-DRO achieves higher success rates under a given tolerance—defined by distance to the goal state for both bimanual push-T and within-hand positioning—while maintaining comparable or lower completion times. Qualitative results (Fig. 2, Fig. 1) demonstrate coordinated, uncertainty-aware behavior, where SVGD particle diversity steers exploration toward performance-relevant regions, enabling richer contact interactions and adaptation to latent parameter uncertainty (e.g., unknown mass distribution). Overall, combining SVGD with DRO effectively approximates the task-aware posterior $p(\mathcal{O}|\theta)$, yielding controllers that robustly adapt to uncertainty in a task-sensitive manner.

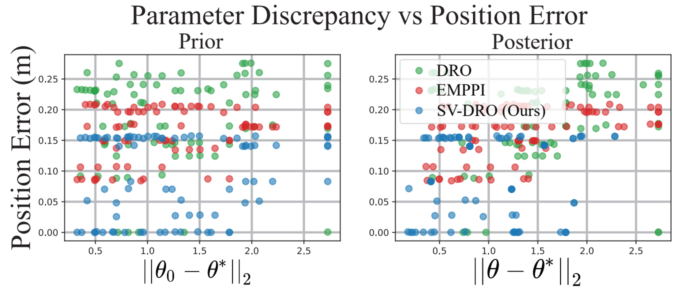

Fig. 4: Position Error vs. Param. Samples. We show the relationship between task performance (distance to goal) and initial parameter mean discrepancy in the within-hand positioning experiment; our method maintains lower error by adapting to task-sensitive uncertainty despite discrepancy, while EMPPI and DRO exhibit weaker success due to task-agnostic posterior samples.

SVI Improves Task-Aware Convergence. We analyze convergence of our method in solving the DRO problem for within-hand dynamic object positioning by isolating the trajectory optimization in Alg. 1 and iterating to convergence from parameter priors sampled uniformly within bounded sets, repeated over 16 trials. As shown in Fig. 3, our approach exhibits significantly lower variance due to the deterministic evolution of SVGD, yielding consistent and reliable controllers despite variability in prior initialization. In contrast, EMPPI displays high loss variance across iterations due to stochastic, task-agnostic sampling, while classical DRO converges to inferior solutions due to its inherently risk-averse formulation.

A3. Task-Aware Inference Improves Manipulation. We study the impact of parameter inference on manipulation success by relating parameter sample discrepancy to task performance in within-hand dynamic positioning. Our method evolves initially sampled parameters into a task-aware posterior via SVGD, yielding higher success rates (Table I) and, as shown in Fig. 4, improved performance irrespective of the initial L2 error to the unknown true parameters. This indicates that task-aware, non-parametric inference captures sensitivity-relevant uncertainty more effectively than baseline methods, whose performance degrades as their prior mean deviates from the true parameters due to conservative or task-agnostic responses to uncertainty.

V. Conclusion

We present a method for reliable manipulation under broad uncertainty by integrating Stein variational inference with DRO to model task-aware uncertainty and synthesize high-performing controllers, demonstrating accurate posterior approximation, significant gains in contact-rich tasks, and real-world robustness via continuous within-hand experiments.

References

[1] Ian Abraham, Ankur Handa, Nathan Ratliff, Kendall Lowrey, Todd D. Murphey, and Dieter Fox. Model-based generalization under parameter uncertainty using

path integral control. *IEEE Robotics and Automation Letters*, 5(2):2864–2871, April 2020. ISSN 2377-3774. doi: 10.1109/lra.2020.2972836. URL http://dx.doi.org/10.1109/LRA.2020.2972836.

[2] Bo Ai, Stephen Tian, Haochen Shi, Yixuan Wang, Tobias Pfaff, Cheston Tan, Henrik I. Christensen, Hao Su, Jiajun Wu, and Yunzhu Li. A review of learning-based dynamics models for robotic manipulation. *Science Robotics*, 10(106):eadt1497, 2025. doi: 10.1126/scirobotics.adt1497. URL https://www.science.org/doi/abs/10.1126/scirobotics.adt1497.

[3] Lucas Barcelos, Alexander Lambert, Rafael Oliveira, Paulo Borges, Byron Boots, and Fabio Ramos. Dual online stein variational inference for control and dynamics, 2021. URL https://arxiv.org/abs/2103.12890.

[4] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion, 2024. URL https://arxiv.org/abs/2303.04137.

[5] Embodiment Collaboration and Abby O'Neill et al. Open x-embodiment: Robotic learning datasets and rt-x models, 2025. URL https://arxiv.org/abs/2310.08864.

[6] Gabriel Dulac-Arnold, Nir Levine, Daniel J. Mankowitz, Jerry Li, Cosmin Paduraru, Sven Gowal, and Todd Hester. An empirical investigation of the challenges of real-world reinforcement learning, 2021. URL https://arxiv.org/abs/2003.11881.

[7] Rishi Bommasani et. al. On the opportunities and risks of foundation models, 2022. URL https://arxiv.org/abs/2108.07258.

[8] Christian Fiedler, Carsten W. Scherer, and Sebastian Trimpe. Learning-enhanced robust controller synthesis with rigorous statistical and control-theoretic guarantees, 2021. URL https://arxiv.org/abs/2105.03397.

[9] Astghik Hakobyan and Insoon Yang. Distributionally robust optimization with unscented transform for learning-based motion control in dynamic environments. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3225–3232, 2023. doi: 10.1109/ICRA48891.2023.10161246.

[10] Nicolas Heess, Greg Wayne, David Silver, Timothy Lillicrap, Yuval Tassa, and Tom Erez. Learning continuous control policies by stochastic value gradients, 2015. URL https://arxiv.org/abs/1510.09142.

[11] Ian, Rhodes, and David H. Jacobson. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control*, 18:124–131, 1973. URL https://api.semanticscholar.org/CorpusID:478053.

[12] Jaywant P. Kolhe, Md Shaheed, T.S. Chandar, and S.E. Talole. Robust control of robot manipulators based on uncertainty and disturbance estimation. *International Journal of Robust and Nonlinear Control*, 23(1):104–122, 2013. doi: https://doi.org/10.1002/rnc.1823. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rnc.1823.

[13] Daniel Kuhn, Soroosh Shafiee, and Wolfram Wiesemann. Distributionally robust optimization, 2025. URL https://arxiv.org/abs/2411.02549.

[14] Alexander Lambert, Adam Fishman, Dieter Fox, Byron Boots, and Fabio Ramos. Stein variational model predictive control, 2021. URL https://arxiv.org/abs/2011.07641.

[15] Darrick Lee, Cameron Lerch, Fabio Ramos, and Ian Abraham. Stein variational ergodic search, 2024. URL https://arxiv.org/abs/2406.11767.

[16] Yong-Hoon Lee, Keuntae Kim, Jaehyun Park, Chung Hyuk Park, and Hae-Won Park. Soft contact model for robust locomotion of legged robots.

[17] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-end training of deep visuomotor policies, 2016. URL https://arxiv.org/abs/1504.00702.

[18] Lei Li, Yingzhou Li, Jian-Guo Liu, Zibu Liu, and Jianfeng Lu. A stochastic version of stein variational gradient descent for efficient sampling. *Communications in Applied Mathematics and Computational Science*, 15(1):37–63, June 2020. ISSN 1559-3940. doi: 10.2140/camcos.2020.15.37. URL http://dx.doi.org/10.2140/camcos.2020.15.37.

[19] Qiang Liu and Dilin Wang. Stein variational gradient descent: A general purpose bayesian inference algorithm, 2019. URL https://arxiv.org/abs/1608.04471.

[20] Rui Liu, Guangyao Shi, and Pratap Tokekar. Data-driven distributionally robust optimal control with state-dependent noise, 2025. URL https://arxiv.org/abs/2303.02293.

[21] Kehan Long, Yinzhuang Yi, Zhirui Dai, Sylvia Herbert, Jorge Cortés, and Nikolay Atanasov. Sensor-based distributionally robust control for safe robot navigation in dynamic environments, 2025. URL https://arxiv.org/abs/2405.18251.

[22] Kendall Lowrey, Aravind Rajeswaran, Sham Kakade, Emanuel Todorov, and Igor Mordatch. Plan online, learn offline: Efficient learning and exploration via model-based control, 2019. URL https://arxiv.org/abs/1811.01848.

[23] Igor Mordatch, Kendall Lowrey, and Emanuel Todorov. Ensemble-cio: Full-body dynamic motion planning that transfers to physical humanoids. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5307–5314, 2015. doi: 10.1109/IROS.2015.7354126.

[24] OpenAI, Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafal Jozefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, Jonas Schneider, Szymon Sidor, Josh Tobin, Peter Welinder, Lilian Weng, and Wojciech Zaremba. Learning dexterous in-hand manipulation, 2019. URL https://arxiv.org/abs/1808.00177.

[25] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. In *2018 IEEE International Conference on Robotics and Automa-*

tion (ICRA), page 3803–3810. IEEE, May 2018. doi: 10.1109/icra.2018.8460528. URL http://dx.doi.org/10.1109/ICRA.2018.8460528.

[26] Aravind Rajeswaran, Sarvjeet Ghotra, Balaraman Ravindran, and Sergey Levine. Epopt: Learning robust neural network policies using model ensembles, 2017. URL https://arxiv.org/abs/1610.01283.

[27] Charles Stein, Persi Diaconis, Susan Holmes, and Gesine Reinert. Use of exchangeable pairs in the analysis of simulations. 46, 01 2004. doi: 10.1214/lnms/1196283797.

[28] Tyler Summers. Distributionally robust sampling-based motion planning under uncertainty. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6518–6523, 2018. doi: 10.1109/IROS.2018.8593893.

[29] TRI LBM Team and Jose Barreiros et. al. A careful examination of large behavior models for multitask dexterous manipulation, 2025. URL https://arxiv.org/abs/2507.05331.

[30] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world, 2017. URL https://arxiv.org/abs/1703.06907.

[31] P. Whittle. Risk-sensitive linear/quadratic/gaussian control. *Advances in Applied Probability*, 13(4):764–777, 1981. doi: 10.2307/1426972.

[32] Liang Yan and Tao Zhou. Stein variational gradient descent with local approximations. *Computer Methods in Applied Mechanics and Engineering*, 386:114087, 2021. ISSN 0045-7825. doi: https://doi.org/10.1016/j.cma.2021.114087. URL https://www.sciencedirect.com/science/article/pii/S0045782521004187.

[33] Dan Zhang and Bin Wei. A review on model reference adaptive control of robotic manipulators. *Annual Reviews in Control*, 43:188–198, 2017. ISSN 1367-5788. doi: https://doi.org/10.1016/j.arcontrol.2017.02.002. URL https://www.sciencedirect.com/science/article/pii/S1367578816301110.

[34] Jingwei Zhuo, Chang Liu, Jiaxin Shi, Jun Zhu, Ning Chen, and Bo Zhang. Message passing stein variational gradient descent, 2018. URL https://arxiv.org/abs/1711.04425.