Equivariant Learning for Robotic Manipulation

Dian Wang, Northeastern University

1 Introduction

As a researcher in Machine Learning and Robotics, my ultimate goal is to build intelligent robotic systems that can generalize across diverse tasks and environments with minimal training, thereby enabling their deployment in complex real-world scenarios. However, current machine learning models are generally not sample-efficient, requiring vast amounts of training data before they can be deployed. This is often too expensive for many problems, especially in robotics where collecting real-robot data requires a large amount of time running the robot. Consequently, improving sample efficiency and generalizability—i.e., enabling policy learning with a minimal number of samples, and allowing the model to generalize to unseen scenarios—is crucial for the widespread adoption of real-world robotics applications.

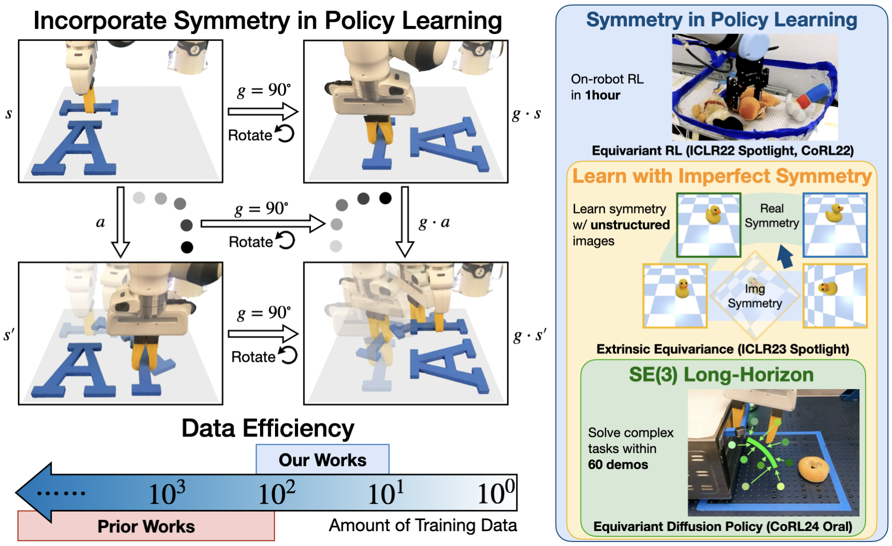


Figure 1: Left: symmetry in the transition of a letter alignment task. When the state is rotated, the learned action should be rotated accordingly. Our method embeds such symmetries in the policy to achieve learning within only tens to hundreds of data, significantly fewer than prior works require. Right: selection of my prior works.

My research leverages the fundamental geometric symmetries in the environment as an inductive bias to improve policy learning in robotics. The key idea is to use the mathematical study of symmetry to design equivariant neural network architectures whose layers are constrained to be symmetric, allowing policies to automatically generalize across symmetrical transformations. As shown in Figure 1 (left), the learned policy can generalize to a rotated state by rotating the action accordingly. This approach results in sample efficiency that is orders of magnitude greater than previous methods. While similar ideas have been explored in other domains, our work is the first to apply equivariant learning in robotics.

My work makes several contributions to design equivariant policies in robot learning. First, we introduce group-invariant MDPs [1], the theoretical grounding of equivariant policy learning. Second, we develop equivariant reinforcement learning algorithms under different problem formulations [2, 1, 3, 4] (Figure 1 right top). We have also investigated equivariant learning in behavior cloning [5], grasping [6, 7, 8, 9], and pick-place [10, 11, 12, 13] in various settings, and achieve the state-of-the-art performance. Third, we demonstrate theoretically and empirically that our models are robust to the degradation of model assumptions: they continue to work well even with symmetry-breaking factors like camera angles, occlusions, etc [14, 15] (Figure 1 right middle). Moreover, my recent paper [16] combines the strength of equivariant learning with diffusion models to learn sample-efficient policies in $\mathrm{SE}(3)$ long-horizon control (Figure 1 right bottom). In the remainder of this document, I will first describe my selected prior works in Section 2, then outline my proposed future works in Section 3.

2 Prior Research

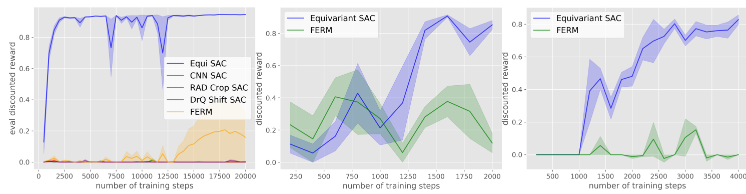
2.1 Theory of Equivariant Policy Learning

Our first work in this area theoretically characterizes the problem settings where equivariant learning can be applied. Consider a group G that represents a space of transformations (e.g., $G = \mathrm{SO}(2)$ for all planar rotations), we define G-invariant Markov decision processes (MDPs) [1] as a class of MDPs whose reward and transition functions remain unchanged under these transformations. Formally, $\forall g \in G, \mathcal{R}(s,a) = \mathcal{R}(gs, ga), T(s,a,s') = T(gs, ga, gs')$. We prove that the optimal solution of a G-invariant MDP is inherently G-equivariant:

Theorem 2.1 (Wang et al. [1]). *For any G-invariant MDP. Its optimal Q-function is G-invariant, $Q^*(s,a) = Q^*(gs, ga)$; its optimal policy is G-equivariant, $\pi^*(gs) = g\pi^*(s)$, for any $g \in G$.*

A function f is *equivariant* with respect to G if it commutes with all transformations $g \in G, f(gx) = gf(x)$. This is a mathematical way of expressing that f is symmetric with respect to G: if we evaluate f for differently transformed versions of the same input, we should obtain transformed versions of the same output. Theorem 2.1 establishes the theoretical foundation of equivariant policy learning: whenever our problem can be formulated as a G-invariant MDP, we can use equivariant networks, a class of networks that guarantees equivariance through weight sharing [17, 18], to model the value or policy functions to improve sample efficiency and generalization.

2.2 Equivariant SAC



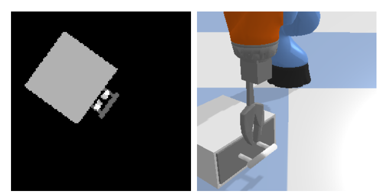
(a) Drawer Opening (Sim) (b) Grasping (Real) (c) Block in Bowl (Real)

Figure 2: Comparison of Equivariant SAC (blue) with baselines.

In our ICLR 2022 spotlight paper [1], we apply Theorem 2.1 to develop Equivariant Soft Actor-Critic (SAC), a novel reinforcement learning model for robotic manipulation. This model demonstrates significantly improved sample efficiency compared to baseline methods. In particular, for $g \in \mathrm{SO}(2)$, we model the policy network $\pi : S \to A$ as an equivariant network that satisfies the equivariant constraint $\pi(gs) = g\pi(s)$, and model the critic network $q : S \times A \to \mathbb{R}$ as an invariant network with the invariant constraint $q(gs, ga) = q(s, a)$. In the subsequent CoRL 2022 work [3], we use this method to learn real-world manipulation policies from scratch within one to two hours, a drastic improvement over previous methods that required hundreds of hours. Figure 2 shows the comparison between Equivariant SAC and the baselines in both simulation (a) and the real world (bc), where our method dramatically outperforms the baselines. My works established the foundation of equivariant robot learning and equivariant reinforcement learning, inspiring many follow-up works along this direction, e.g., [19, 20, 21, 22, 23].

2.3 Equivariant Learning with Extrinsic Symmetry

Equivariant models usually rely on structured observations, such as the top-down view in Figure 3a, where the state transformation can be directly derived from the image transformation. However, in many real-world scenarios, these ideal observations are unavailable due to factors like occlusion or varying camera angles (e.g., Figure 3b), leading to a mismatch between image transformation and state transformation. My ICLR 2023 spotlight paper [14] demonstrates that equivariant models are robust to these mismatches and can still provide a significant performance boost (we call it *extrinsic equivariance*). This finding significantly broadens the applicability of equivariant learning by eliminating the need for structured input. We further categorize the relationship between the problem symmetry and the model symmetry as *correct*, *incorrect*, and *extrinsic*, where our follow-up NeurIPS 2023 paper [15] proposes a general theory about them with the theoretical lower bounds of equivariant models under symmetry mismatch.



(a) Top-Down Obs (b) Side-View Obs

Figure 3: (a) Ideal top-down image (b) Side-view image where the transformation of the image is out-of-distribution.

2.4 Equivariant Diffusion Policy

Since robotic tasks often require multi-modal policies for learning from diverse human demonstrations, my most recent CoRL 2024 Best Paper Award Finalist work [16] explores the combination of equivariant policy learning with diffusion models to capture the multi-modality and thereby tackle complex long-horizon tasks. We show that the policy denoising process is equivariant when the policy itself is equivariant, and further provide a thorough demonstration of leveraging $\mathrm{SO}(2)$-equivariance in the full 6-DoF $\mathrm{SE}(3)$ control. We propose Equivariant Diffusion Policy, and demonstrate in simulation that it outperforms three state-of-the-art imitation learning baselines [24, 25, 26] by 22%. In the real world, our method can learn policies within only tens of demonstration data. For example, our method successfully solves a long-horizon bagel baking task (Figure 4) with 80% success rate using only 58 demonstrations, whereas the baseline only reaches a 10% success rate.

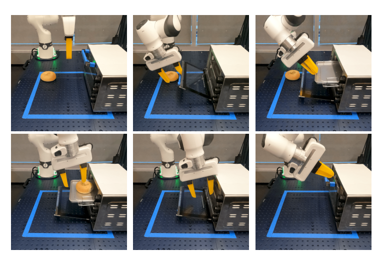


Figure 4: Our method can learn a long-horizon bagel baking task with 58 demonstrations.

3 Proposed Work: Toward Generalizable and Efficient Robot Learning Systems

My long-term vision is to develop robotic systems that generalize across diverse tasks and environments with minimal training, enabling their deployment in complex real-world scenarios. To achieve this, I believe the sample efficiency and generalizability offered by equivariant learning are essential. In this section, I outline key components for realizing this vision, addressing both practical challenges and promising extensions of equivariant learning.

3.1 General Robotic Equivariant Learning Framework

The first step toward this goal is to design a flexible and adaptive equivariant learning framework tailored specifically for robotics. This will simplify the integration of equivariant learning with other methods, broadening its applicability across the field. As a starting point, I propose to extend equivariant learning to a broader range of symmetries and robotic tasks.

Exploiting Local Symmetries with Adaptive Equivariant Attention: While my previous work has focused primarily on global symmetries, robotic tasks often involve local symmetries that need to be considered for more fine-grained manipulation. To achieve this, I propose integrating equivariant learning with an object-centric approach to reason about the object-level local symmetries. First, to facilitate object-centric reasoning, a region proposal or segmentation network will identify objects in the scene. An equivariant network will process each individual object and generate an equivariant object descriptor that encodes the local symmetry of each object. These object descriptors will be processed by an equivariant graph neural network or Transformer, which will consider both local and global symmetries to output an action.

Extending Equivariant Learning to other Robotic Platforms: Building upon my foundational work in robotic manipulation, I plan to extend equivariant learning to other domains, such as mobile manipulation, navigation, and locomotion. This expansion will validate the robustness of equivariant learning across a wider range of robotic tasks. The key research question is identifying the mathematical definition of symmetry in those systems, thereby implementing it with appropriate equivariant neural networks.

3.2 Resolving Symmetry-Breaking in Robotics

One immediate challenge in broadly applying equivariant policies is the symmetry-breaking caused by the physical constraints in robotics. Equivariant policies generally assume that the environment's dynamics are fully symmetric across a group (e.g., $\mathrm{SE}(3)$). However, this assumption could be violated by factors like fixed obstacles or kinematic constraints. As is shown in my prior work [15], this type of symmetry-breaking will create a theoretical upper bound on equivariant models' performance, posing a fundamental problem on the scalability of equivariant policy learning.

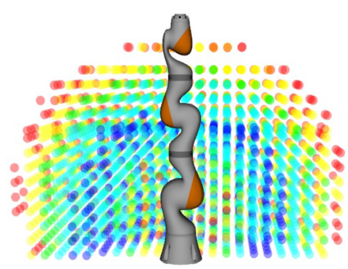


Figure 5: The reachability map of a robot arm. Color represents increasing rotational reachability: Red, Yellow, Green, SkyBlue and Blue. Image adopted from [27].

Handling Kinematic Constraint Symmetry Breaking: As shown in Figure 5, the robot arm can only reach a subset of $\mathrm{SE}(3)$ poses. As a result, an equivariant model might generate kinematically infeasible actions by generalizing from feasible actions in a symmetric manner. To address this issue, I propose using relaxed or approximate equivariant models [28, 29], which have been developed in our lab, to learn policies that are aware of kinematic constraints. These relaxed equivariant models are capable of capturing symmetry-breaking conditions in the dataset and adaptively relaxing the equivariant constraint to account for kinematic feasibility. Alternatively, I will explore pre-training or pre-computing a kinematic feasibility map that can be used to mask out infeasible actions from an equivariant policy. This approach will enhance the feasibility and practicality of using equivariant policies in real-world robotic applications, clearing the obstacles for the broad adoption of equivariant learning in robotics.

3.3 Bridging Equivariant Learning and Foundation Models

The ultimate goal of a general-purpose robotic system is being able to learn new tasks quickly while maintaining the knowledge of solving the previously learned tasks. This requires not only geometric-level generalization, as addressed in my previous work, but also task-level generalization. With the growing prominence of large-scale models, I believe this could be achieved by creating synergy between these large models and smaller, more efficient equivariant models, potentially enabling fast adaptation and generalization in new domains through the use of equivariant learning.

Learning General Policies via Equivariant Skill Composition: As the first step in this direction, I propose combining equivariant skills with a high-level planner to enable faster adaptation across different tasks. In particular, the low-level skills in robotic manipulation (e.g., picking, placing, pushing, etc.) often observe the most geometric symmetries. Although a certain high-level task might not fully leverage the symmetry due to the fixed poses of objects, an equivariant low-level skill could significantly ease skill generalization. For example, an equivariant skill learned for opening a drawer on the left could easily generalize to a task requiring the opening of a drawer on the right. Consequently, I propose to use my prior works to learn low-level skills, and compose them with a high-level planner using foundation models. In this approach, the symmetry in the low-level skills could dramatically improve cross-task generalization.

References

[1] **D. Wang**, R. Walters, and R. Platt, "$\mathrm{SO}(2)$-equivariant reinforcement learning," in *International Conference on Learning Representations (ICLR)*, 2022.

[2] **D. Wang**, R. Walters, X. Zhu, and R. Platt, "Equivariant Q learning in spatial action spaces," in *Conference on Robot Learning (CoRL)*, 2021.

[3] **D. Wang**, M. Jia, X. Zhu, R. Walters, and R. Platt, "On-robot learning with equivariant models," in *Conference on Robot Learning (CoRL)*, 2022.

[4] H. H. Nguyen, A. Baisero, D. Klee, **D. Wang**, R. Platt, and C. Amato, "Equivariant reinforcement learning under partial observability," in *Conference on Robot Learning (CoRL)*, 2023.

[5] M. Jia*, **D. Wang***, G. Su, D. Klee, X. Zhu, R. Walters, and R. Platt, "SEIL: Simulation-augmented Equivariant Imitation Learning," in *International Conference on Robotics and Automation (ICRA)*, 2023.

[6] X. Zhu, **D. Wang**, O. Biza, G. Su, R. Walters, and R. Platt, "Sample efficient grasp learning using equivariant models," in *Robotics: Science and Systems (RSS)*, 2022.

[7] H. Huang, **D. Wang**, X. Zhu, R. Walters, and R. Platt, "Edge Grasp Network: A Graph-Based $\mathrm{SE}(3)$-invariant Approach to Grasp Detection," in *International Conference on Robotics and Automation (ICRA)*, 2023.

[8] X. Zhu, **D. Wang**, G. Su, O. Biza, R. Walters, and R. Platt, "On robot grasp learning using equivariant models," *Autonomous Robots*, 2023.

[9] B. Hu, X. Zhu, **D. Wang**, Z. Dong, H. Huang, C. Wang, R. Walters, and R. Platt, "Orbitgrasp: $SE(3)$-equivariant grasp learning," in *Conference on Robot Learning (CoRL)*, 2024.

[10] H. Huang, **D. Wang**, R. Walters, and R. Platt, "Equivariant transporter network," in *Robotics: Science and Systems (RSS)*, 2022.

[11] H. Huang, **D. Wang**, A. Tangri, R. Walters, and R. Platt, "Leveraging Symmetries in Pick and Place," *The International Journal of Robotics Research (IJRR)*, 2024.

[12] H. Huang, O. L. Howell, **D. Wang**, X. Zhu, R. Platt, and R. Walters, "Fourier transporter: Bi-equivariant robotic manipulation in 3d," in *International Conference on Learning Representations (ICLR)*, 2024.

[13] H. Huang, K. Schmeckpeper, **D. Wang**, O. Biza, Y. Qian, H. Liu, M. Jia, R. Platt, and R. Walters, "Imagination policy: Using generative point cloud models for learning manipulation policies," in *Conference on Robot Learning (CoRL)*, 2024.

[14] **D. Wang**, J. Y. Park, N. Sortur, L. L. Wong, R. Walters, and R. Platt, "The Surprising Effectiveness of Equivariant Models in Domains with Latent Symmetry," in *International Conference on Learning Representations (ICLR)*, 2023.

[15] **D. Wang**, X. Zhu, J. Y. Park, M. Jia, G. Su, R. Platt, and R. Walters, "A general theory of correct, incorrect, and extrinsic equivariance," in *Neural Information Processing Systems (NeurIPS)*, 2023.

[16] **D. Wang**, S. Hart, D. Surovik, T. Kelestemur, H. Huang, H. Zhao, J. Wang, R. Walters, and R. Platt, "Equivariant diffusion policy," in *Conference on Robot Learning (CoRL)*, 2024.

[17] T. Cohen and M. Welling, "Group equivariant convolutional networks," in *International Conference on Machine Learning (ICML)*, 2016.

[18] M. Weiler and G. Cesa, "General e (2)-equivariant steerable cnns," in *Neural Information Processing Systems (NeurIPS)*, 2019.

[19] S. Liu, M. Xu, P. Huang, X. Zhang, Y. Liu, K. Oguchi, and D. Zhao, "Continual vision-based reinforcement learning with group symmetries," in *Conference on Robot Learning (CoRL)*, 2023.

[20] A. Deac, T. Weber, and G. Papamakarios, "Equivariant muzero," *Transactions on Machine Learning Research*, 2023.

[21] X. Yu, R. Shi, P. Feng, Y. Tian, S. Li, S. Liao, and W. Wu, "Leveraging partial symmetry for multi-agent reinforcement learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024.

[22] S. Yan, B. Zhang, Y. Zhang, J. Boedecker, and W. Burgard, "Learning continuous control with geometric regularity from robot intrinsic symmetry," in *International Conference on Robotics and Automation (ICRA)*, 2024.

[23] D. Chen and Q. Zhang, "E(3)-equivariant actor-critic methods for cooperative multi-agent reinforcement learning," in *International Conference on Machine Learning (ICML)*, 2024.

[24] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song, "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion," in *Robotics: Science and Systems (RSS)*, 2023.

[25] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware," in *Robotics: Science and Systems (RSS)*, 2023.

[26] Y. Ze, G. Zhang, K. Zhang, C. Hu, M. Wang, and H. Xu, "3d diffusion policy: Generalizable visuomotor policy learning via simple 3d representations," in *Robotics: Science and Systems (RSS)*, 2024.

[27] A. Makhal and A. K. Goins, "Reuleaux: Robot base placement by reachability analysis," in *International Conference on Robotic Computing (IRC)*, 2018.

[28] R. Wang, R. Walters, and R. Yu, "Approximately equivariant networks for imperfectly symmetric dynamics," in *International Conference on Machine Learning (ICML)*, 2022.

[29] R. Wang, E. Hofgard, H. Gao, R. Walters, and T. E. Smidt, "Discovering symmetry breaking in physical systems with relaxed group convolution," in *International Conference on Machine Learning (ICML)*, 2024.