Adaptive Compliance Policy: Learning Approximate Compliance for Diffusion Guided Control

Yifan Hou[1] Zeyi Liu[1] Cheng Chi[1] Eric Cousineau[2] Naveen Kuppuswamy[2]
Siyuan Feng[2] Benjamin Burchfiel[2] Shuran Song[1]
[1] Stanford University [2] Toyota Research Institute
adaptive-compliance.github.io

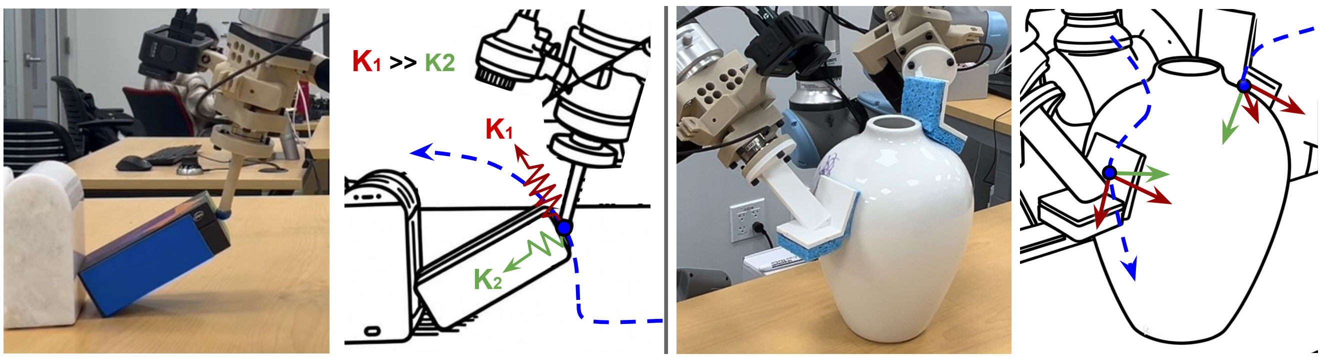

Fig. 1: **Compliance Requirements.** [Left] Flipping an item requires the robot to follow an arc trajectory (blue) while maintaining contact force. This demands low stiffness in pushing directions (K_2) and high stiffness elsewhere (K_1). [Right] Wiping a vase necessitates 3D compliance adjustments in both end-effectors to 1) hold the vase, 2) trace the marking, and 3) apply appropriate force without damage. Our algorithm aims to model these spatial-, temporal-, and task-dependent compliance requirements from human demonstration data.

***Abstract*— Compliance plays a crucial role in manipulation, as it balances between the concurrent control of position and force under uncertainties. Yet compliance is often overlooked by today's visuomotor policies that solely focus on position control. This paper introduces Adaptive Compliance Policy (ACP), a novel framework that learns to dynamically adjust system compliance both spatially and temporally for given manipulation tasks from human demonstrations, improving upon previous approaches that rely on pre-selected compliance parameters or assume uniform constant stiffness. However, computing full compliance parameters from human demonstrations is an ill-defined problem. Instead, we estimate an approximate compliance profile with two useful properties: avoiding large contact forces and encouraging accurate tracking. Our approach enables robots to handle complex contact-rich manipulation tasks and achieves over 50% performance improvement compared to state-of-the-art visuomotor policy methods. Project website with result videos: `adaptive-compliance.github.io`.**

I. INTRODUCTION

Manipulation often requires the concurrent control of both position and force to achieve the desired outcome. This joint objective can be captured by the concept of mechanical compliance [16, 18, 7], where a low compliance prioritizes position accuracy regardless of external forces, while a high compliance allows large position deviation in response to external forces, making the system "soft" during interaction.

The desired compliance for a robotics system is not a static property; rather, it varies drastically depending on the task objectives and the system's state. For instance, consider the flipping task in Fig. 1, the desired compliance:

- **Varies temporally**. For example, The system needs to be less compliant before contact to prioritize precise position tracking while becoming compliant upon contact.
- **Varies spatially**. For example, during the pivoting stage, the system should be compliant only in the pushing directions (i.e., K_2 direction) while maintaining high stiffness in other directions to follow the arc motion (e.g., low compliance in K_1 direction).
- **Varies from task to task**. If we change to a different task, such as wiping a vase in Fig. 1 [Right], both temporal and spatial properties of the compliance will change in order to satisfy the unique 3D motion and force requirements.

While the desired compliance can be obtained from optimization given physical measurements of the manipulation problem [10, 8], it remains a challenge to obtain compliance parameters directly from human demonstration. Prior work often requires known dynamics parameters [6] or repeated demonstrations to statistically estimate human compliance [1, 13, 4, 5]. These approaches cannot handle new scene configurations or unexpected perturbations. As a result, compliant policies either rely on pre-selected compliance parameters for the target tasks [11] or assume uniform constant stiffness across all directions [14].

In this work, we introduce **Adaptive Compliance Policy (ACP)**, a sensorimotor policy that learns to dynamically adjust the system compliance both *spatially* and *temporally* for a *given manipulation task* from human demonstrations.

Specifically, we derive a simple compliance labeling method from classical mechanical analysis that eliminates excessive forces while encourages precise tracking, under mild assumptions about the tasks. This simple rule allows us to approximate varying stiffness for every demonstration

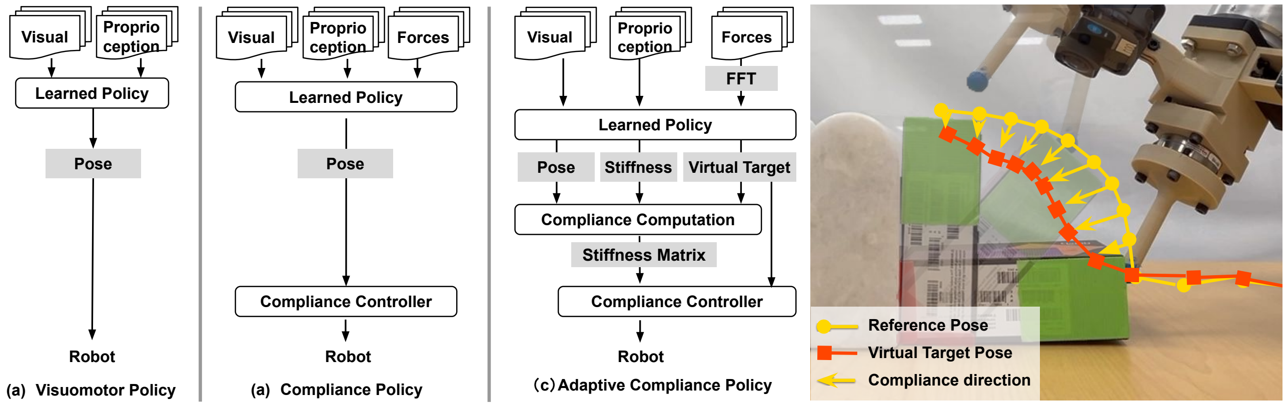

Fig. 2: **Method Comparisons.** [LEFT] shows the comparison between a) a typical visuomotor policy [2], b) a typical force-based compliant policy [14], and c) Adaptive Compliance Policy. [Right] Visualization of virtual target (orange sqaures) and reference poses (yellow circles) inferred by Adaptive Compliance Policy. The directional difference (orange arrows) between the virtual and reference poses encodes compliance direction.

episode with different object variations and scene configurations. Our policy encodes reference pose, force and full stiffness compactly using a virtual target.

We systematically evaluate the performance of our algorithm on two real world contact-rich manipulation tasks: object flipping and vase wiping. Our method achieves over 50% increase in performance compared to state-of-the-art visuomotor policy methods. In summary, the main contribution of the paper includes:

- Adaptive Compliance Policy formulation that is able to dynamically adjust compliance to maintain desired contact modes despite uncertainties and disturbances.
- A model-based method to compute spatial-, temporal-varying compliance labels from human demonstrations, making ACP training practical and scalable.
- A kinesthetic teaching system that allows demonstrations with varying compliance profiles.

II. Method

A. Demonstration Collection and Compliance Control

We choose kinesthetic teaching instead of teleoperation to collect human demonstrations in order for the operator to easily demonstrate variable compliance behavior under direct haptic feedback (see Fig. 3). The setup per arm includes one robot manipulator to provide accurate position feedback, one RGB camera to record visual information, and one force torque sensor mounted near the robot hand.

We use admittance control [15] to give our robot compliance. During demonstration, we specify zero stiffness, low damping and low mass for the controller so the operator can move the robot freely and demonstrate their own compliance profile. During testing, we set a high damping and mass value for the robot to maintain stability of the admittance controller. Our learned policy predicts the stiffness profile of the robot compliance, which has the most significant effect on manipulation, especially at low speed.

B. Estimating Stiffness from Demonstrations

Human uses varying stiffness during manipulation. High stiffness provides position accuracy under force disturbances,

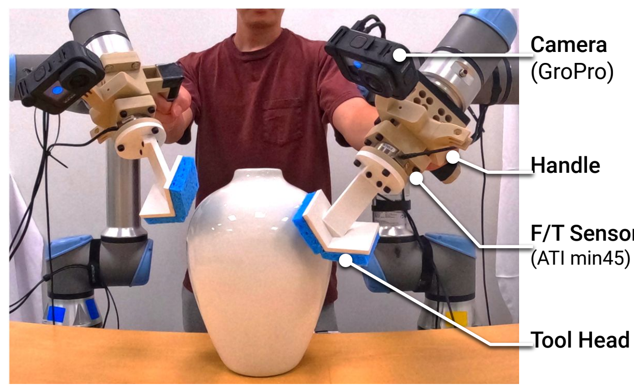

Fig. 3: **Data Collection with Haptic Feedback.** We designed a kinesthetic teaching system with low-stiffness compliance that allows the operator to demonstrate variable compliance behavior with direct haptic feedback.

while low stiffness allows safe contact engaging and maintaining [9]. However, it is often an ill-conditioned problem to estimate compliance parameters from a single human demonstration due to the lack of variations [1]. Instead of estimating the true human stiffness, we propose to find a stiffness matrix with the following properties:

- It avoids huge internal forces in manipulation.
- It encourages accurate tracking of the desired motion.

We obtain such a stiffness matrix by first deciding the direction of high and low stiffnesses, then specifying the stiffness magnitude in those directions.

C.1 Stiffness direction: We propose the following simple strategy to choose stiffness direction in the generalized space: *Use a low stiffness k_{low} in the direction of the force feedback, and a high stiffness k_{high} in all other directions.*

We show in II-C that such single-axis low stiffness control is sufficient to eliminate excessive contact force under mild assumptions. Then we can use high stiffness in other directions to improve position tracking. Let $K_0 \in \mathbb{R}^N$ be a diagonal matrix with $[k_{low}, k_{high}, \ldots, k_{high}]$ on its diagonal, and $S \in \mathbb{R}^N$ be a matrix whose columns form an orthonormal basis of $\mathbb{R}^N$ with its first column as $f/|f|$. The stiffness matrix can be written as:

$$K = SK_0S^{-1} \tag{1}$$

We use k_{high} in all directions when $|f|$ is small.

C.2 Stiffness Magnitude The high stiffness k_{high} is set empirically to support accurate position tracking. Since the low stiffness direction is estimated from noisy force signal, we found it helpful to let the low stiffness value decrease continuously with the force magnitude:

$$k_{low} = \begin{cases} k_{max}, & |f| < f_{min} \\ k_{max} - (k_{max} - k_{min}) \frac{|f| - f_{min}}{f_{max} - f_{min}}, & f_{min} \le |f| \le f_{max} \\ k_{min}, & |f| > f_{max} \end{cases} \tag{2}$$

where $k_{max}, k_{min}, f_{max}$ and f_{min} are parameters determined by the hardware system.

C. Effectiveness of single axis low stiffness control

To explain why it suffice to use only one axis compliance control to avoid excessive contact forces, we make the following assumptions:

Assumption I: *Contact force dominates all forces on the robot. Other types of force, such as inertia force, friction and gravity, are negligible comparing with contact force.*

We ensure Assumption I by avoiding fast robot motion and using lightweight objects. Denote $v, f \in \mathbb{R}^N$ as the generalized velocity and force vector of a manipulation system, $\lambda \in \mathbb{R}^n$ as the vector of contact normal forces. With Assumption I, the contact imposes constraints on the system through the Contact Jacobian matrix J as follows:

$$J^T \lambda = f, \tag{3}$$
$$Jv \ge 0. \tag{4}$$

We then make two more assumptions:

Assumption II: *Nonzero contact force: all made contacts should have nonzero contact forces.*

Assumption III: *No pinching contacts: the cone formed by rows of the Contact Jacobian J is contained in its dual cone.*

Assumption II can be satisfied by making contacts clearly in demonstrations. Assumption III means the contacts on the robot are not too restrictive. Fig. 4 shows examples:

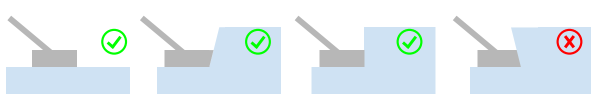

Fig. 4: **Pinching Examples.** Grey shape represents a robot tool, blue shape represents a frictionless environment. First three examples are not pinching contact, the last one is.

Excessive contact force happens in a manipulation system when the high stiffness control conflict with the contact constraints, both can be described as velocity constraints. In other words, no excessive force will occur if all the velocity constraints in the manipulation system form a feasible system. By doing low stiffness control, our method gives up control of velocity in the force feedback direction. We show in the following theorem that this can guarantee the feasibility of all velocity constraints, thus eliminating excessive contact force:

Theorem 1. *For a robot under external contact described by Eq. 3, there exists a solution v that satisfies the contact constraint 4 as long as it does not control its velocity in the direction of feedback force f in the generalized space.*

Proof. Not controlling velocity in the force direction means the velocity has a free component:

$$v = v_0 + kf = v_0 + kJ^T \lambda, \tag{5}$$

where k is an arbitrary scaling factor, v_0 denotes the components of generalized velocity in other directions. Due to Assumption II, the contact force λ must have all positive components, $J^T \lambda$ represents a ray strictly inside the cone formed by rows of J. Assumption III says this cone is contained in its dual cone $\{x \in \mathbb{R}^N | Jx \ge 0\}$, so

$$JJ^T \lambda > 0. \tag{6}$$

Then $JV = JV_0 + kJJ^T \lambda > 0$ for large enough k.

D. Adaptive Compliance Policy

We formulate the policy as a diffusion process [2] for both reference action and target stiffness.

1) Inputs and Encoding: We implement two encoding strategies for the force/torque data: 1) temporal encoding via causal convolution [17], which helps capture causal relations from sequential data like force. 2) FFT encoding, where we convert each dimension of the 6D force/torque readings into a 2D spectrogram then pass to a ResNet-18 model with a modified input channel of 6. Both image and force encodings are combined using self-attention then concatenated with robot end-effector poses and fed to the downstream diffusion policy head as a condition following [2].

2) Outputs and decoding: Our policy output encodes the position target, the stiffness matrix, and the reference force in a 19-dimensional vector per robot arm:

- Reference pose: 9D pose vector following convention in [2];
- Virtual target pose: 9D pose representing the actual target for the low-level compliance controller to track;
- A scalar stiffness value k_{low}.

During training, we compute the stiffness matrix from filtered wrench data using Eq. 1, then compute the virtual target following a 3D mechanical spring. The benefit of using virtual target is to have a uniform target representation across different robots: an impedance-controlled robot without FT sensor can also execute the virtual target. During inference, the full stiffness matrix is reconstructed following Eq. 1, then sent to the low-level compliance controller for execution together with the virtual target.

III. Experiments

We evaluate our method in two contact-rich manipulation tasks whose success depends on the maintenance of suitable contact modes. We evaluate the following four policies, all trained on the same dataset with the same number of epochs:

- ACP: Adaptive Compliance Policy, our approach;
- ACP w.o. FFT: same as ACP but with force encoded using temporal convolution [17, 14] instead of FFT.

- Stiff policy: Diffusion policy [2, 3] with additional force input. Outputs target positions.
- Compliant policy: Same as the [Stiff policy] except that the low level controller has a uniform stiffness $k = 500$N/m. Relying on low level robot compliance is common in visuomotor policies [14, 19, 20, 12].

A. Task I: Item Flipping

The task is to flip up an item with a point finger by pivoting it against a corner of a fixture (i.e., a wall), as exemplified in Fig. 2. This task evaluates our method's generalizability towards noval items. The task has three main failure modes: 1) The finger loses contact and drops the item; 2) Extensive pushing force violates robot force limit; 3) Motion gets stuck at a bad pose. Since the item weight is light, the friction from the item to the robot finger is also small, thus do not violate Assumption I.

Test Scenarios. We ran 20 tests in each of the five scenarios below, making 100 tests per algorithm:

- Training Items: Items appeared in training data.
- Unseen Items: Items not seen in training.
- Push&Flip: Items start 5cm away from the fixture and needs to be pushed before flipping.
- Varied Fixture Pose: Two different fixture poses.
- Unstable Fixture: Lighter fixture that causes unstable flipping movements. Require the policy to quickly adapt.

Results. The success rate is shown in Tab. I. Success is defined as the item being rotated greater than 70 degrees.

TABLE I
FLIPPING-UP SUCCESS RATES (%)

	Train Items	Unseen Items	Push &Flip	Fixture Pose	Unstable Fixture	All
ACP	**90**	95	95	**100**	**100**	**96**
ACP w.o. FFT	**90**	**100**	**100**	95	90	95
Compliant Policy	80	15	15	5	0	23
Stiff Policy	20	0	5	35	10	14

Findings. The two baselines [Compliant Policy] and [Stiff Policy] have a few successes when they can exploit the passive compliance in the system. They effectively applies a force when the predicted trajectory is in collision with a deformable item. When the item is rigid, or when the position uncertainty is not in a convenient direction, the baseline polices break the contacts and fail the task. On the contrary, both variations of ACP tolerates a large range of position uncertainties while maintain the needed contacts.

B. Task II: Vase Wiping

The robot needs to wipe off random markings on a vase that is randomly placed on the table. This task tests our method for handling large disturbance forces (friction). For this task, each robot arm has two pieces of kitchen sponges as wipers. We collected 200 demonstrations with various vase poses, marking shapes, and colors. Each demonstration includes one to five wipes to fully clean the markings. Although the vase is heavy, it conveys force to the robot only through contact force and friction, so Assumption I still holds.

Test Scenarios. The following scenarios are tested:

- Small Mark×5: easier cases that need only one wipe.
- Large Mark×5: require multiple wipes.
- Perturbation before contact (PbC)×4: move the vase right before the tool comes in contact with the vase.
- Perturbation after contact (PaC)×2: move the vase after the tool is engaged to disturb the wiping motion.

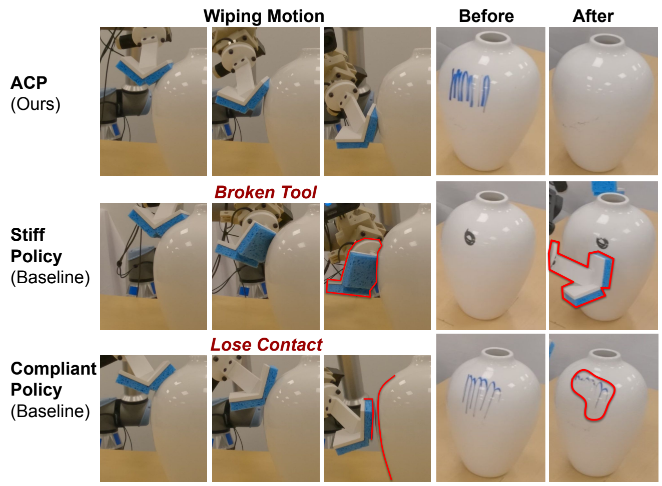

Fig. 5: **Wiping Comparisons**. [Top] APC: maintains contact and follows desired trajectory. [Middle] Stiff Policy: Position noise causes excessive force that breaks the tool. [Bottom] Compliant Policy: Safe contact, but friction hinders wiping position accuracy and eventually loses contact.

Results. The quantitative results are summarized in Tab. II. All policies demonstrated wiping behaviors. We define success as the mark being cleaned (the remaining marks are within 1cm×1cm) within three wipes.

TABLE II
WIPING SUCCESS RATES (%)

	Small	Large	PbC	PaC	All
ACP	**100**	**80**	**100**	**100**	**93.75**
ACP w.o. FFT	100	60	75	100	81.25
Compliant Policy	60	20	25	100	43.75

Findings. In all scenarios, [ACP] safely engages and maintains contacts, while [Stiff Policy]'s contact force magnitude varies greatly and broke its tool during the fourth test. [ACP] maintains accurate tracking of the desired motion, while the wiping motion of the [Compliant Policy] deviates from the position target under friction. We also observe that our policy with the FFT encoding wipes more efficiently than [ACP w.o. FFT], suggesting that FFT encoding helps the policy to make better decision on the next best wiping location.

IV. CONCLUSIONS

In this work, we show that Adaptive Compliance Policy is an effective visuomotor policy for compliant manipulation. Extensive real-world results show that our approach is able to extract useful compliance from human demonstration, and thereby significantly improve the success rate of two contact-rich manipulation tasks.

References

[1] Sylvain Calinon, Irene Sardellitti, and Darwin G Caldwell. "Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies". In: *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2010, pp. 249–254.

[2] Cheng Chi et al. "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion". In: *Proceedings of Robotics: Science and Systems (RSS)*. 2023.

[3] Cheng Chi et al. "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion". In: *The International Journal of Robotics Research* (2024).

[4] Jianghua Duan et al. "Learning compliant manipulation tasks from force demonstrations". In: *2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*. IEEE. 2018, pp. 449–454.

[5] Jianghua Duan et al. "Sequential learning unification controller from human demonstrations for robotic compliant manipulation". In: *Neurocomputing* 366 (2019), pp. 35–45.

[6] Giorgio Grioli and Antonio Bicchi. "A non-invasive, real-time method for measuring variable stiffness". In: (2011).

[7] Neville Hogan. "Impedance control: An approach to manipulation". In: *1984 American control conference*. IEEE. 1984, pp. 304–313.

[8] Yifan Hou, Zhenzhong Jia, and Matthew T Mason. "Manipulation with shared grasping". In: *arXiv preprint arXiv:2006.02996* (2020).

[9] Yifan Hou and Matthew T Mason. "An efficient closed-form method for optimal hybrid force-velocity control". In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2021, pp. 11125–11131.

[10] Yifan Hou and Matthew T Mason. "Robust execution of contact-rich motion plans by hybrid force-velocity control". In: *2019 International Conference on Robotics and Automation (ICRA)*. IEEE. 2019, pp. 1933–1939.

[11] Tatsuya Kamijo, Cristian C Beltran-Hernandez, and Masashi Hamaya. "Learning Variable Compliance Control From a Few Demonstrations for Bimanual Robot with Haptic Feedback Teleoperation System". In: *arXiv preprint arXiv:2406.14990* (2024).

[12] Colin Kohler et al. "Symmetric models for visual force policy learning". In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2024, pp. 3101–3107.

[13] Klas Kronander and Aude Billard. "Online learning of varying stiffness through physical human-robot interaction". In: *2012 IEEE International Conference on Robotics and Automation*. Ieee. 2012, pp. 1842–1849.

[14] Michelle A Lee et al. "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks". In: *2019 International conference on robotics and automation (ICRA)*. IEEE. 2019, pp. 8943–8950.

[15] J Maples and Joseph Becker. "Experiments in force control of robotic manipulators". In: *Proceedings. 1986 IEEE International Conference on Robotics and Automation*. Vol. 3. IEEE. 1986, pp. 695–702.

[16] Matthew T Mason. "Compliance and force control for computer controlled manipulators". In: *IEEE Transactions on Systems, Man, and Cybernetics* 11.6 (1981), pp. 418–432.

[17] Aaron van den Oord. "WaveNet: A Generative Model for Raw Audio". In: *arXiv preprint arXiv:1609.03499* (2016).

[18] Marc H Raibert and John J Craig. "Hybrid position/force control of manipulators". In: (1981).

[19] Chen Wang et al. "Mimicplay: Long-horizon imitation learning by watching human play". In: *arXiv preprint arXiv:2302.12422* (2023).

[20] Yifeng Zhu et al. "Viola: Imitation learning for vision-based manipulation with object proposal priors". In: *Conference on Robot Learning*. PMLR. 2023, pp. 1199–1210.