NOVEL YET CLEAN: TOWARDS THEORY-GUIDED DATA AUGMENTATION FOR ROBOT IMITATION LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Data augmentation is a standard remedy for the demonstration-data bottleneck in robot imitation learning, yet existing methods are largely designed empirically and lack a unified theory of when and why augmentation helps. We develop a behavior-cloning analysis that jointly models real and augmented trajectories and reveals a fundamental principle for effective augmentation: augmented data should provide sufficiently new information while introducing as little augmentation error as possible—*Novel Yet Clean*. Guided by this principle, we propose **NYC**, a trajectory-wise, geometry-consistent, diffusion-based augmentation system that increases trajectory coverage through geometry-aware view synthesis while reducing augmentation bias through trajectory-wise consistency and diffusion-based visual repair. Extensive experiments on Push-T and five bimanual RLBench tasks validate both the theoretical predictions and the effectiveness of the resulting system. NYC improves task success rates by $+29\% \sim +180\%$ over a real-only baseline and achieves state-of-the-art performance on four of the five bimanual tasks, demonstrating that our theoretical principles can effectively translate into practical system design and downstream performance gains.

1 INTRODUCTION

Imitation learning has become a central paradigm for robot manipulation, underlying both classical behavior cloning (BC) and recent robot foundation models and vision–language–action systems such as RT-2, RT-X, Octo, and OpenVLA (Zitkovich et al., 2023; Open X-Embodiment Collaboration, 2024; Ghosh et al., 2024; Kim et al., 2025). These methods rely heavily on large-scale robot trajectory data (Firoozi et al., 2025; Kawaharazuka et al., 2024), yet high-quality demonstrations remain expensive to collect and difficult to scale across tasks, environments, and platforms (Open X-Embodiment Collaboration, 2024; Fu et al., 2025; Lin et al., 2025). Data scarcity therefore remains a major bottleneck for robot learning.

Data augmentation offers a natural solution. Existing methods range from simple image or action perturbations (Laskin et al., 2020; Corrado et al., 2024; Kobayashi et al., 2025) to generation-based approaches that synthesize new observations or trajectories using modern generative models (OpenAI, 2023; Ho et al., 2020; Rombach et al., 2022; Mandlekar et al., 2023; Liu et al., 2025; Zhang et al., 2024; Chen et al., 2025). Existing augmentation methods have largely evolved along two intuitive directions: making augmented data more novel, so that it introduces useful information beyond the original demonstrations, and making it more clean, so that it remains faithful to the underlying task distribution. In practice, their effectiveness also depends critically on choosing an appropriate amount of augmentation. These empirical observations naturally raise a fundamental theoretical question: *why, when, and how much does data augmentation help robot imitation learning?* Answering this question requires jointly characterizing the information gain brought by augmented data, the bias introduced in the augmentation process, and how their trade-off evolves as more augmented samples are added.

This question is especially important for generative augmentation, where realistic-looking samples may still introduce systematic task-relevant errors. A theoretical understanding of the interplay

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

between information gain and augmentation bias is therefore needed to explain both the benefits and the eventual saturation of augmentation.

To address this gap, we develop a unified theoretical framework for behavior cloning with mixed real and augmented trajectories. We characterize augmentation through two complementary quantities: *trajectory coverage efficiency* γ_{cover}, which measures the non-redundant information contributed by augmented trajectories, and *augmentation error*, which measures their deviation from expert-consistent trajectories. The former determines the effective sample size $N_{\text{eff}} = N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}$, while the latter controls the bias introduced by augmentation. Our generalization analysis separates these two effects, revealing an information–bias tradeoff and explaining why augmentation gains eventually saturate.

We validate these predictions in controlled Push-T experiments, independently varying augmentation quantity, coverage, and error. The observed effects of trajectory coverage, augmentation error, and augmentation quantity on model performance lead to a simple design principle:

*An effective augmentation system should be **novel yet clean**:*
(1) novel enough to provide new and useful information, and (2) clean enough to minimize the bias it introduces.

Guided by this principle, we propose **NYC** (**N**ovel **Y**et **C**lean), a trajectory-wise, geometry-consistent, diffusion-based augmentation system. Geometry-aware transformation generates larger and more diverse trajectory variations, while trajectory-wise consistency and diffusion-based repair reduce augmentation error. Experiments on five bimanual RLBench tasks show that these theory-derived design principles translate into improved augmentation quality and downstream imitation performance.

Our contributions are threefold:

- **A unified theory of augmentation for behavior cloning.** We jointly model real and augmented trajectories, characterize their information gain and bias through trajectory coverage efficiency, effective sample size, and augmentation error, and derive generalization and saturation results that reveal their fundamental tradeoff.
- **A theory-derived design principle and practical augmentation system.** Our analysis yields the *Novel Yet Clean* principle, which we instantiate in NYC using geometry-aware transformation, trajectory-wise consistency, and diffusion-based visual repair.
- **Validation from theory to downstream performance.** Controlled Push-T experiments validate the predicted effects of coverage, bias, and augmentation quantity, while five bimanual RLBench tasks demonstrate the effectiveness of the resulting system.

2 Related Work

Generalization theory for imitation learning. A long line of work studies the error behavior of imitation learning, typically assuming that all training data come from the real expert distribution (Osa et al., 2018). Classical analyses of behavior cloning (BC), from ALVINN (Pomerleau, 1988) to the compounding-error result of Ross & Bagnell (2010), show that distribution mismatch can yield an $O(T^2)$ performance loss, motivating interactive methods such as DAgger (Ross et al., 2011) and AggreVaTe (Ross & Bagnell, 2014) that recover linear $O(T)$ dependence. More recent work refines the statistical limits of offline imitation, including minimax analysis (Rajaraman et al., 2020), horizon-independent BC under suitable complexity conditions (Foster et al., 2024), and fast rates under concentrability assumptions (Rashidinejad et al., 2021). These analyses, however, consider datasets composed purely of real expert trajectories and do not model training on mixtures of real and augmented data, leaving the effectiveness, saturation, and bias of augmentation unexplained.

Data augmentation for robot imitation learning. A broad range of methods enrich robot demonstrations through image/state perturbation, trajectory-level synthesis, or direct observation generation. Random visual transforms and related heuristics are widely used in visual control and imitation (Laskin et al., 2020; Yarats et al., 2021; Corrado et al., 2024; Kobayashi et al., 2025); trajectory-level methods recombine or retarget demonstrations (Mandlekar et al., 2023; Wan et al., 2025), while

diffusion-based methods such as DMD and D-CODA generate "deviate-and-recover" samples for single- and dual-arm imitation (Zhang et al., 2024; Liu et al., 2025). Other approaches augment visual observations through inpainting, cross-view or cross-embodiment generation, and sim-to-real transfer (Yu et al., 2023; Chen et al., 2025; Samak et al., 2026). Despite their diversity, these methods are largely empirical, and no prior work has established a unified framework for understanding data augmentation.

Scaling laws and augmentation theory beyond robotics. Related work also studies scaling behavior and the statistical role of augmentation. Lin et al. (2025) find that manipulation performance scales strongly with environment and object diversity, suggesting that augmentation value depends on effective coverage rather than raw sample count. More broadly, augmentation has been analyzed through vicinal-risk minimization, implicit regularization, and variance reduction (Chapelle et al., 2000; Zhang et al., 2018; Dao et al., 2019; Chen et al., 2020). These theories mainly target i.i.d. classification or static vision settings and do not capture the temporal correlation, contact-sensitive dynamics, and accumulated error of robot trajectories, motivating our trajectory-level analysis.

3 Theoretical Analysis

Robot imitation data consists of correlated trajectories rather than independent samples, and augmented actions must remain consistent with task geometry and expert behavior. Thus, augmentation value depends on both the new trajectory information it contributes and the error it introduces. We develop a behavior-cloning analysis that separates these two effects; all detailed proofs are deferred to Appendix A.

3.1 Problem Formulation

We consider a finite-horizon deterministic Markov decision process $\mathcal{M} = (\mathcal{S}, \mathcal{A}, f, H)$ with state space $\mathcal{S} \subseteq \mathbb{R}^n$, action space $\mathcal{A} \subseteq \mathbb{R}^m$, transition map $f : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$, and horizon H. For manipulation, the end-effector pose lies in $SE(3)$; within a bounded workspace, we represent actions locally in $\mathbb{R}^6$ without affecting the scaling relations below.

An expert policy $\pi^\star : \mathcal{S} \to \mathcal{A}$ induces a trajectory $\tau = ((s_1, a_1^\star), \ldots, (s_T, a_T^\star))$ with $a_t^\star = \pi^\star(s_t)$ and $s_{t+1} = f(s_t, a_t^\star)$, where $T = H$. We measure trajectory discrepancy by

$$d_{\mathcal{T}}(\tau, \tau') \triangleq \frac{1}{T} \sum_{t=1}^{T} \|a_t - a_t'\|_2, \tag{1}$$

which captures task-relevant action deviation.

Let $\mathcal{D}_{\text{real}} = \{\tau_i\}_{i=1}^{N_{\text{real}}}$ denote real expert demonstrations and $\mathcal{D}_{\text{aug}} = \{\tilde{\tau}_j\}_{j=1}^{N_{\text{aug}}}$ the augmented trajectories produced by $\text{Aug}(\cdot)$. Training uses $\mathcal{D} = \mathcal{D}_{\text{real}} \cup \mathcal{D}_{\text{aug}}$ with $N = N_{\text{real}} + N_{\text{aug}}$. For policy class $\Pi = \{\pi_\theta : \mathcal{S} \to \mathcal{A} \mid \theta \in \Theta\}$ and squared imitation loss $\ell(\pi_\theta, (s, a)) = \|\pi_\theta(s) - a\|_2^2 \in [0, B^2]$, define

$$L(\pi_\theta) = \mathbb{E}_{s \sim \rho^{\pi^\star}}\big[\|\pi_\theta(s) - \pi^\star(s)\|_2^2\big], \qquad \hat{L}(\pi_\theta; \mathcal{D}) = \frac{1}{NT} \sum_{\tau \in \mathcal{D}} \sum_{t=1}^{T} \|\pi_\theta(s_t) - a_t\|_2^2, \tag{2}$$

where $\rho^{\pi^\star}$ is the expert state-visitation distribution. We study the ERM policy $\pi_\theta \in \arg\min_{\pi \in \Pi} \hat{L}(\pi; \mathcal{D})$.

3.2 Core Definitions

We characterize augmentation by two quantities: the *error* it introduces and the *new information* it contributes.

Definition 1 (Augmentation error)**.** *Let $\tau^\star_{\text{target}}$ denote the ideal expert-consistent trajectory at the augmented target configuration. The augmentation error is*

$$\varepsilon_{\text{aug}} \triangleq \Big(\mathbb{E}_{\tau \sim P_{\text{real}}}\big[d_{\mathcal{T}}(\text{Aug}(\tau), \tau^\star_{\text{target}})^2\big]\Big)^{1/2}. \tag{3}$$

A smaller ε_{aug} indicates less geometric, relabeling, or generative error.

To quantify new information, let $\varphi : \mathcal{T} \to \mathbb{R}^d$ be a trajectory feature encoder with normalized features $\hat{\varphi}(\tau) = \varphi(\tau)/\|\varphi(\tau)\|_2$, and let Φ_{real} and Φ_{aug} stack the real and augmented trajectory features.

Definition 2 (Trajectory Coverage Efficiency, TCE)**.**

$$\gamma_{\text{cover}} \triangleq 1 - \frac{\|\Phi_{\text{aug}}\Phi_{\text{real}}^{\top}\|_F}{\|\Phi_{\text{aug}}\|_F \, \|\Phi_{\text{real}}\|_F} \in [0, 1], \tag{4}$$

where $[\Phi_{\text{aug}}\Phi_{\text{real}}^{\top}]_{ji} = \hat{\varphi}(\tilde{\tau}_j)^{\top}\hat{\varphi}(\tau_i)$ *is the cosine similarity between augmented trajectory* j *and real trajectory* i*.*

A larger γ_{cover} means lower redundancy with the real trajectories and hence more new information. This motivates the *effective sample size*

$$N_{\text{eff}} \triangleq N_{\text{real}} + \gamma_{\text{cover}} \, N_{\text{aug}}. \tag{5}$$

Thus, an augmented trajectory contributes according to its non-redundant information rather than being counted automatically as a fresh real sample.

To account for trajectory correlation, we assume exponential β-mixing $\beta(k) \le \beta_0 e^{-k/\tau_{\text{mix}}}$ (Yu, 1994; Mohri & Rostamizadeh, 2010), with $\tau_{\text{mix}} + 1$ acting as the correlation factor in the bound below.

3.3 Effective Sample Size and Generalization

Theorem 1 (Effective-sample-size generalization bound)**.** *Let* Π *have pseudo-dimension* d *(Pollard, 1984), with induced squared-loss class uniformly bounded in* $[0, B^2]$*. Assume the exponential* β*-mixing condition with mixing time* τ_{mix}*, an* L_{aug}*-Lipschitz augmentation operator* Aug *(w.r.t.* $d_{\mathcal{T}}$*), and an* $L_{\pi^\star}$*-Lipschitz expert policy* $\pi^\star$*. With* $N_{\text{eff}} = N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}$ *and* $N = N_{\text{real}} + N_{\text{aug}}$*, the ERM policy* π_θ *satisfies, for any* $\delta \in (0, 1)$*, with probability at least* $1 - \delta$*,*

$$L(\pi_\theta) - L(\pi^\star) \;\le\; \underbrace{C_1 B^2 \sqrt{\frac{(\tau_{\text{mix}} + 1)\big(\mathsf{d} \ln \frac{N_{\text{eff}}}{\mathsf{d}} + \ln \frac{2}{\delta}\big)}{N_{\text{eff}}}}}_{\textit{statistical term}} + \underbrace{\frac{N_{\text{aug}}}{N} C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \, \varepsilon_{\text{aug}}^2}_{\textit{augmentation-bias term}}, \tag{6}$$

where $C_1, C_2 > 0$ *are absolute constants and* $C_2 = O(H)$*.*

Theorem 1 separates excess imitation risk into a statistical term controlled by N_{eff} and a bias term controlled by $\varepsilon_{\text{aug}}^2$ and the augmented fraction N_{aug}/N. Augmentation reduces statistical error only when it raises N_{eff}: high γ_{cover} makes augmented trajectories informative, whereas redundant augmentation provides little benefit. Meanwhile, nonzero ε_{aug} increases the bias term as the augmented fraction grows. Effective augmentation must therefore increase γ_{cover} while reducing $\varepsilon_{\text{aug}}^2$.

3.4 Bias–Variance Tradeoff and Saturation

Corollary 1 (Information gain vs. augmentation bias)**.** *Under the conditions of Theorem 1, the bound equation 6 equals* $S(N_{\text{aug}}) + B_{\text{aug}}(N_{\text{aug}})$ *with statistical term* S *decreasing in* N_{eff} *and bias term* $B_{\text{aug}}(N_{\text{aug}}) = \frac{N_{\text{aug}}}{N_{\text{real}} + N_{\text{aug}}} b$*. When* $\gamma_{\text{cover}} > 0$*, increasing* N_{aug} *lowers* S *(information gain) but, when* $b > 0$*, raises the weight of* B_{aug} *(augmentation bias).*

Thus, increasing augmentation reduces statistical error but simultaneously increases the influence of augmentation bias. This tradeoff leads directly to saturation.

Theorem 2 (Saturation of augmentation gains)**.** *Under the conditions of Theorem 1 with* $\gamma_{\text{cover}} > 0$*, as* $N_{\text{aug}} \to \infty$ *the statistical term tends to* 0 *and the bias term tends to*

$$\eta_{\text{sat}} \triangleq C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2, \tag{7}$$

so the bound equation 6 converges to η_{sat}*.*

Theorem 2 shows that scaling augmentation alone cannot eliminate the bound: once the statistical term becomes small, the remaining floor is determined by $\varepsilon_{\text{aug}}^2$. Further improvement must therefore come from better augmentation quality rather than quantity.

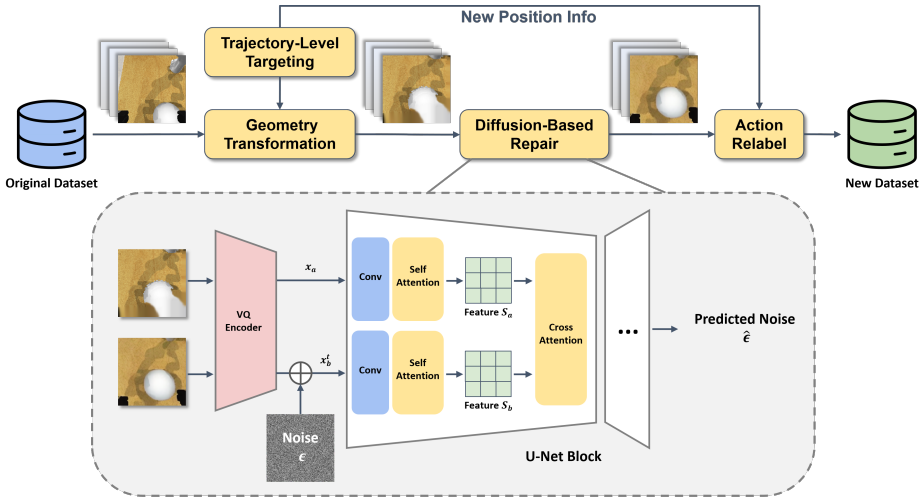

Figure 1: **Overview of NYC and its diffusion-based repair module.** NYC consists of four stages: **(1) trajectory-level targeting** constructs a new task-consistent trajectory; **(2) geometry transformation** maps observations to the target viewpoints using explicit geometry; **(3) diffusion-based repair** removes artifacts from the warped observations; and **(4) action relabel** recomputes action labels using a Levenberg–Marquardt (LM) inverse-kinematics solver. During diffusion training, the U-Net predicts noise conditioned on the warped observation; at inference, the warped observation conditions generation of the repaired result.

3.5 Summary

The analysis yields two design objectives: **(i)** increase trajectory coverage efficiency γ_{cover} so augmented trajectories provide more new information, and **(ii)** reduce augmentation error $\varepsilon^2_{\text{aug}}$ to limit systematic bias. We summarize these objectives as the *Novel Yet Clean* principle: augmentation should be *novel* enough to provide useful new information while remaining *clean* enough to avoid introducing excessive error. This principle directly motivates the design of NYC in Section 4.

4 The NYC Augmentation System

Section 3 shows that effective augmentation should provide new trajectory information while introducing minimal bias. Guided by this *Novel Yet Clean* principle, we propose **NYC** (**N**ovel **Y**et **C**lean), a trajectory-level augmentation system that combines geometry-aware transformation, trajectory-wise consistency, and diffusion-based visual repair.

4.1 System Overview

NYC operates on complete expert demonstrations rather than independently perturbing individual frames or action steps. It constructs a new task-consistent trajectory together with corresponding observations and action labels. Figure 1 shows the overall pipeline, and Figure 2 compares NYC with D-CODA. Further details on the system are provided in Appendix B.

The repair module follows the conditional diffusion formulation of DMD (Zhang et al., 2024), but differs in how target geometry is specified. DMD conditions generation directly on a perturbation offset and try to implicitly realize the viewpoint change, whereas NYC first explicitly warps the observation towards the target geometry and uses diffusion only for visual repair. This separation improves controllability and enables larger geometric transformations.

4.2 Geometry-aware Transformation for High Coverage

Augmentation provides statistical benefit only when it introduces sufficiently new trajectory information. Small image or action perturbations remain highly redundant with the original data, while

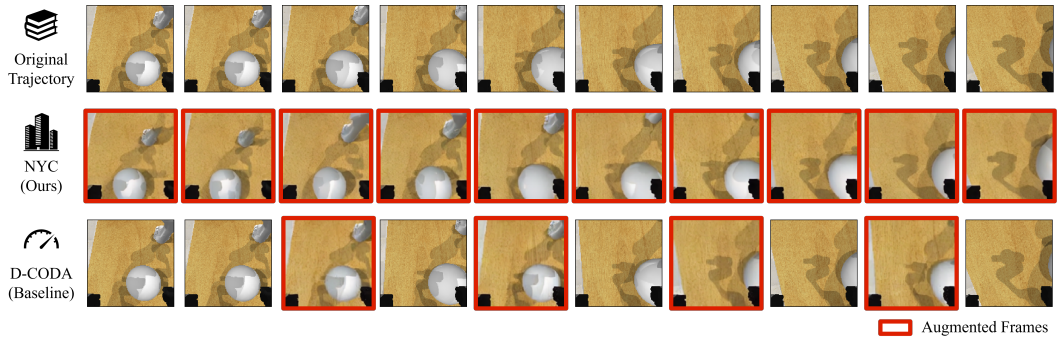

Figure 2: **Comparison of augmented observations generated by D-CODA and NYC.** NYC provides denser, more substantial, and more accurate trajectory-level augmentation, producing more diverse and high-quality trajectories.

directly conditioning diffusion models on large trajectory offsets can be difficult to control reliably (Zhang et al., 2024; Liu et al., 2025). NYC instead performs explicit geometry-aware transformation. Given depth and camera intrinsics, pixels are back-projected into 3D, transformed according to the target trajectory, and re-projected into the target view (Mildenhall et al., 2020). This allows larger and more structured task-space variations while preserving spatial relations among the robot, objects, and scene.

Geometric warping may introduce disocclusions, depth errors, and projection holes, which are subsequently corrected by the diffusion repair module. The resulting *warp-then-repair* pipeline therefore enables large geometric variations without excessive visual degradation.

4.3 Trajectory-wise Consistent Augmentation for Low Bias

High coverage alone is insufficient if augmentation introduces systematic trajectory errors. In particular, independently perturbing individual action steps, as in point-wise augmentation used by DMD and D-CODA (Zhang et al., 2024; Liu et al., 2025), can disrupt the temporal structure of robot motion. NYC therefore applies a smooth, task-consistent transformation to the complete trajectory and recomputes the corresponding action labels through inverse kinematics. This preserves temporal structure and maintains consistency among transformed robot poses, observations, and actions. Accordingly, trajectory-wise consistency reduces $d_{\mathcal{T}}(\mathrm{Aug}(\tau), \tau^{\star}_{\mathrm{target}})$ and hence $\varepsilon_{\mathrm{aug}}$, forming the "clean" side of NYC together with diffusion-based repair.

4.4 Empirical Measures of Coverage and Augmentation Bias

The theoretical quantities γ_{cover} and $\varepsilon_{\mathrm{aug}}$ are not directly observable in practical systems. We therefore introduce empirical measures for evaluating the two factors.

Trajectory coverage efficiency. We extract trajectory-level visual features using a pretrained ResNet-50 and compute the similarity between augmented and real trajectories according to Definition 2. Lower similarity corresponds to less redundancy and hence larger empirical trajectory coverage efficiency.

Spectral measure of augmentation bias. To estimate trajectory distortion, we compare the power spectral densities of real and augmented action sequences. Let $S_a(k)$ and $S_{\tilde{a}}(k)$ denote their power spectral densities. We define

$$D_{\mathrm{spec}}(\tilde{\tau}, \tau) \triangleq \frac{1}{T} \sum_{k=0}^{T-1} \big| \, \mathbb{E}[S_{\tilde{a}}(k)] - S_a(k) \, \big|. \tag{8}$$

A smaller D_{spec} indicates better preservation of expert trajectory structure and thus serves as an empirical proxy for lower augmentation bias. These measures are used in Section 5.2 to evaluate whether NYC achieves the two objectives identified by our theory.

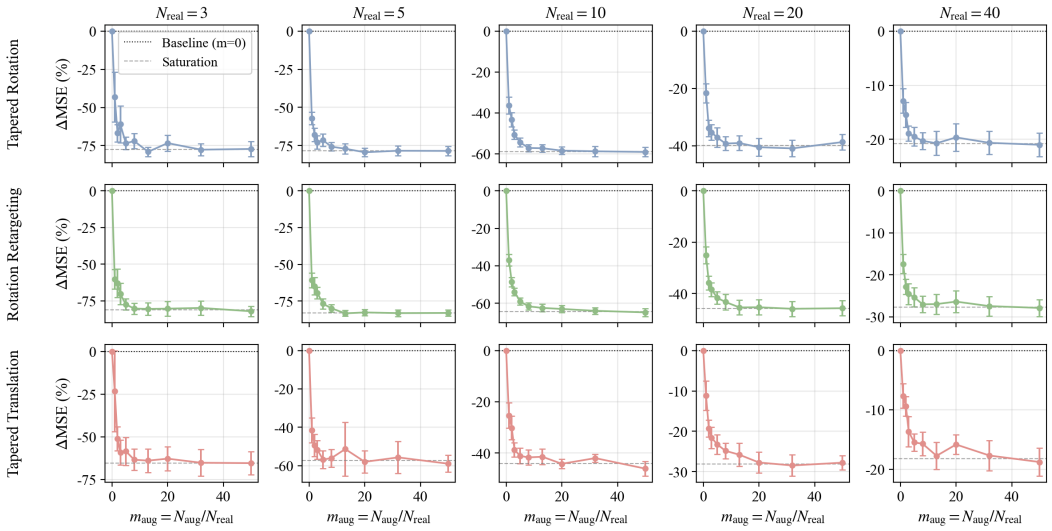

Figure 3: **Relative reduction in Push-T test MSE as the augmentation multiplier m_{aug} increases.** Gains are largest with few augmented trajectories, especially in low-data regimes, and progressively saturate as augmentation increases.

5 Experiments

Our experiments evaluate the framework at two levels. We first use controlled Push-T experiments to test the predicted effects of augmentation quantity, coverage, and error. We then evaluate NYC on five bimanual RLBench tasks to examine whether these principles translate into practical augmentation and downstream policy gains.

5.1 Theory Verification on Push-T

The bound in Section 3 can be summarized as

$$\mathcal{L}(\hat{\pi}) - \mathcal{L}(\pi^\star) \lesssim \underbrace{\sqrt{\frac{\mathsf{d} \log N_{\text{eff}}}{N_{\text{eff}}}}}_{\text{statistical term}} + \underbrace{\frac{N_{\text{aug}}}{N_{\text{real}} + N_{\text{aug}}} C\varepsilon_{\text{aug}}^2}_{\text{augmentation-bias term}}, \qquad N_{\text{eff}} = N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}, \tag{9}$$

where mixing and confidence factors are omitted for clarity. It predicts that augmentation gains should saturate with quantity, increase with coverage efficiency, and deteriorate with systematic augmentation error. We test these effects independently on Push-T.

Setup. Push-T (Chi et al., 2023) provides a low-dimensional continuous control setting in which N_{real}, N_{aug}, coverage, and augmentation error can be varied independently. We use a fixed BC policy and training protocol and evaluate action MSE on held-out expert trajectories. Results are averaged over 10 random seeds. Further details of the experimental setup are provided in Appendix B.

(1) Augmentation gains exhibit saturation. We vary the augmentation multiplier $m_{\text{aug}} = N_{\text{aug}}/N_{\text{real}}$ under different real-data budgets and geometric augmentation schemes. Figure 3 shows rapid initial improvement followed by diminishing gains, typically approaching a plateau around $m_{\text{aug}} \approx 5$–10. The effect is stronger when real data are scarce, consistent with Theorem 2: augmentation is most valuable while the statistical term remains large.

(2) Higher coverage improves statistical efficiency. We apply trajectory-level translation augmentation and vary its maximum displacement as a practical proxy for γ_{cover}, while keeping the augmentation procedure otherwise fixed. As shown in Figure 4, test MSE generally decreases with coverage; larger augmentation budgets also tend to further reduce MSE. This agrees with the effective-sample-size interpretation that both γ_{cover} and N_{aug} increase N_{eff}.

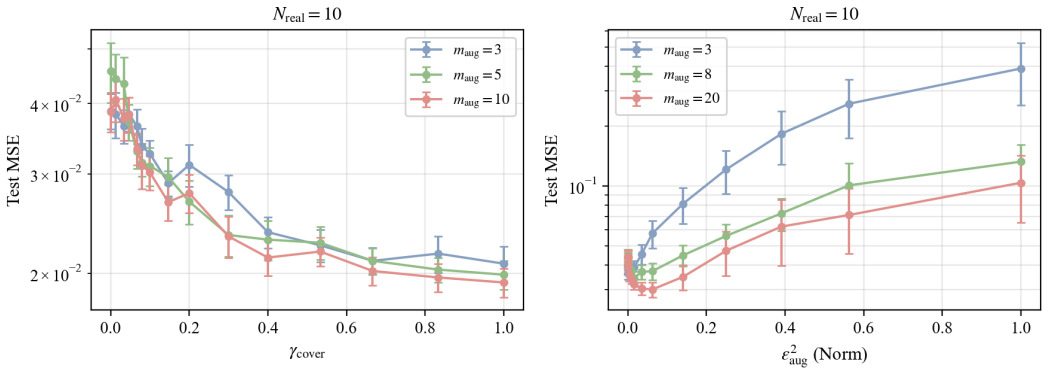

Figure 4: **Push-T test MSE versus trajectory coverage and augmentation error.** Higher coverage generally reduces test error, with stronger gains in lower-data regimes. Larger systematic error increases test loss, and additional augmented data does not eliminate the resulting excess error.

(3) Systematic augmentation error raises test loss. We introduce a trajectory-consistent action offset of controlled magnitude into augmented trajectories, providing a controllable realization of $\varepsilon^2_{\text{aug}}$. Figure 4 shows that MSE rises as augmentation error becomes appreciable, while increasing m_{aug} cannot remove the additional loss. This supports the predicted bias term: scaling augmented data reduces statistical uncertainty but not systematic augmentation error.

Together, these experiments reproduce the three predicted trends: augmentation exhibits diminishing returns, higher coverage provides greater statistical benefit, and systematic augmentation error limits the achievable gain.

5.2 System Evaluation: NYC on Bimanual RLBench

We evaluate NYC in CoppeliaSim on adapted versions of five bimanual tasks from RLBench and PerAct2 (Grotz et al., 2024; James et al., 2020). We adjust task-scene configurations and success criteria for more controlled evaluation, with full details provided in Appendix B. We compare **Real-only**, **More-real**, **Two-DMD** (Zhang et al., 2024), **D-CODA** (Liu et al., 2025), and **NYC** using the same ACT policy (Zhao et al., 2023) and evaluation protocol. Detailed task configurations, evaluation procedures, simulation settings, and policy-training details are provided in Appendix B.

Trajectory coverage efficiency. Using pretrained ResNet-50 trajectory features, we compute γ_{cover} according to Equation 4. Table 1 shows that NYC achieves $3.15\times$–$5.30\times$ higher coverage efficiency than D-CODA across all five tasks.

Table 1: Trajectory coverage efficiency γ_{cover} (higher is better).

Task	D-CODA	NYC	Gain
Coordinated Lift Ball	0.015	0.047	**3.15**×
Coordinated Push Box	0.013	0.050	**3.86**×
Coordinated Lift Tray	0.009	0.046	**5.30**×
Dual Push Buttons	0.014	0.051	**3.58**×
Bimanual Straighten Rope	0.015	0.050	**3.35**×

Table 2: Wasserstein-1 spectral distance between augmented and real trajectories (lower is better).

Task	PA vs Real	TA vs Real	PA/TA
Coordinated Lift Ball	0.143	0.001	**102.66**
Coordinated Push Box	0.193	0.002	**88.88**
Coordinated Lift Tray	0.151	0.002	**94.91**
Dual Push Buttons	0.286	0.004	**78.08**
Bimanual Straighten Rope	0.151	0.001	**152.83**

Spectral structure and augmentation bias. We compare the Wasserstein-1 spectral distance between augmented and real action sequences. DMD and D-CODA use point-wise augmentation (PA), whereas NYC uses trajectory-wise augmentation (TA). As shown in Table 2, TA reduces the spectral distance by factors of 78–153. Figure 5 further shows that TA better preserves the low-frequency motion structure of the expert trajectories, consistent with lower augmentation bias.

Downstream imitation performance. Figure 6 compares the downstream success rates. NYC improves over Real-only by $+50\%$, $+35\%$, $+180\%$, $+76\%$, and $+29\%$ across the five tasks, respectively. More-real does not consistently improve performance, likely because the larger dataset

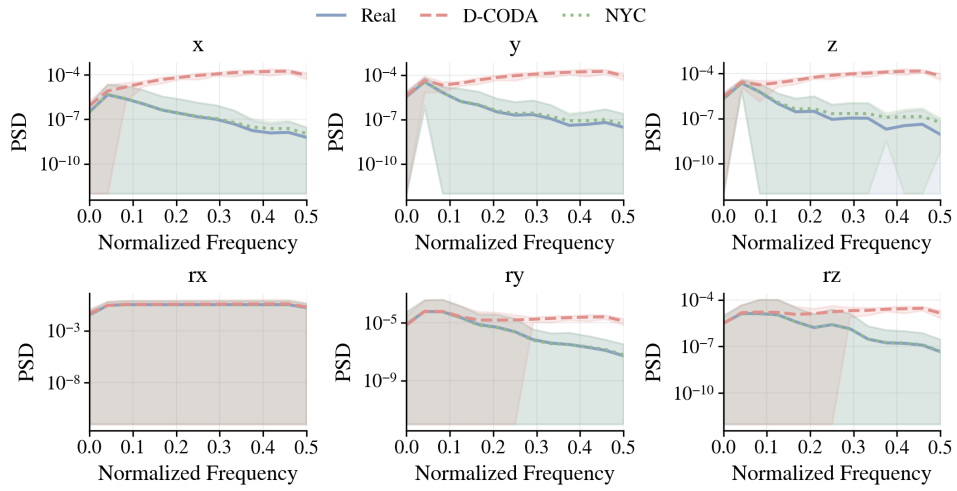

Figure 5: **Power spectral density of point-wise (PA) and trajectory-wise (TA) augmented actions.** TA better preserves the low-frequency structure of real trajectories, whereas PA introduces stronger high-frequency components.

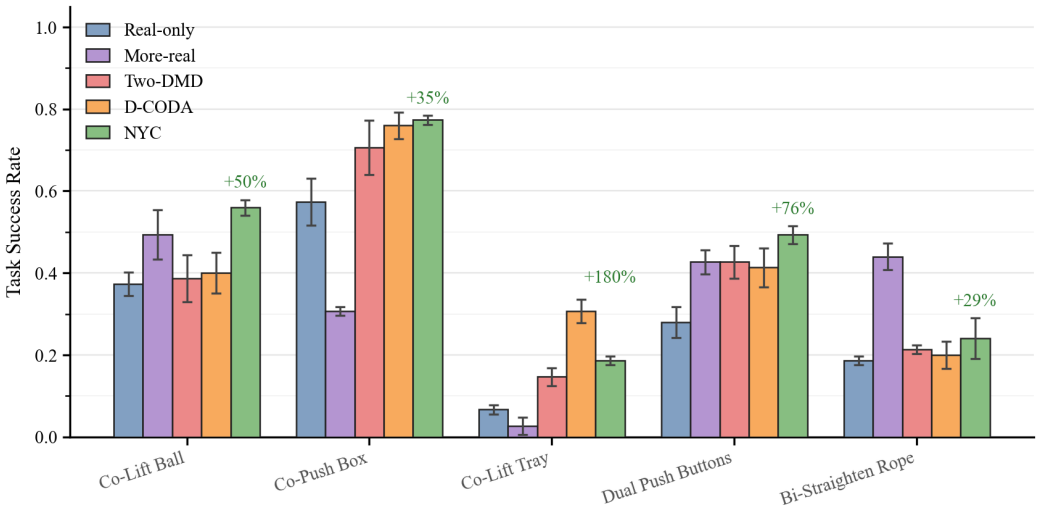

Figure 6: **Task success rates of different training-data settings on five bimanual RLBench tasks.** NYC consistently improves over Real-only and provides strong gains across the evaluated tasks.

is less fully optimized under the same training budget. Two-DMD and D-CODA generally provide smaller and less consistent gains. Together with the preceding analyses, these results demonstrate the effectiveness of translating the *Novel Yet Clean* principle into a practical augmentation system.

6 Conclusion

We presented a theory-guided framework for data augmentation in robot imitation learning, showing that effective augmentation should introduce new information while minimizing augmentation bias—the *Novel Yet Clean* principle. Guided by this principle, we developed NYC and validated both the theoretical predictions and system effectiveness through controlled Push-T and bimanual RLBench experiments. Our work provides theoretical foundations and practical guidance for designing robot imitation-learning augmentation systems. However, our analysis relies on idealized assumptions, and the current evaluation is primarily conducted in simulation. Future work will relax these assumptions and extend the framework to larger-scale real-robot settings.

AI USE STATEMENT

We used generative AI tools solely for basic code writing, literature search, and language polishing. Diffusion models were also used as part of our proposed methodology to generate augmented data for the experiments. All AI-generated content was carefully reviewed and verified by the authors. All research ideas, theoretical analyses, and experimental designs were developed by the authors, and all experimental results were obtained, analyzed, and verified by the authors. AI tools did not contribute to ideation or scientific content. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

P. L. Bartlett and S. Mendelson. Rademacher and Gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, 3:463–482, 2002.

O. Chapelle, J. Weston, L. Bottou, and V. Vapnik. Vicinal risk minimization. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, pp. 416–422, 2000.

L. Y. Chen, C. Xu, K. Dharmarajan, et al. RoVi-Aug: Robot and viewpoint augmentation for cross-embodiment robot learning. In *Proc. Conf. Robot Learning (CoRL)*, pp. 209–233, 2025.

S. Chen, E. Dobriban, and J. H. Lee. A group-theoretic framework for data augmentation. *Journal of Machine Learning Research*, 21(245):1–71, 2020.

C. Chi, S. Feng, Y. Du, et al. Diffusion policy: Visuomotor policy learning via action diffusion. In *Proc. Robotics: Science and Systems (RSS)*, 2023.

N. E. Corrado, Y. Qu, J. U. Balis, et al. Guided data augmentation for offline reinforcement learning and imitation learning. *Reinforcement Learning Journal*, 1:198–215, 2024.

T. Dao, A. Gu, A. Ratner, et al. A kernel theory of modern data augmentation. In *Proc. Int. Conf. Machine Learning (ICML)*, pp. 1528–1537, 2019.

R. Firoozi, J. Tucker, S. Tian, et al. Foundation models in robotics: Applications, challenges, and the future. *The International Journal of Robotics Research*, 44(5):701–739, 2025.

D. J. Foster, A. Block, and D. Misra. Is behavior cloning all you need? understanding horizon in imitation learning. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Z. Fu, T. Z. Zhao, and C. Finn. Mobile ALOHA: Learning bimanual mobile manipulation using low-cost whole-body teleoperation. In *Proc. Conf. Robot Learning (CoRL)*, pp. 4066–4083, 2025.

D. Ghosh, H. R. Walke, K. Pertsch, et al. Octo: An open-source generalist robot policy. In *Proceedings of Robotics: Science and Systems (RSS)*, 2024.

M. Grotz, M. Shridhar, T. Asfour, et al. PerAct2: Benchmarking and learning for robotic bimanual manipulation tasks. arXiv preprint arXiv:2407.00278, 2024.

J. Ho, A. Jain, and P. Abbeel. Denoising diffusion probabilistic models. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, pp. 6840–6851, 2020.

S. James, Z. Ma, D. R. Arrojo, et al. RLBench: The robot learning benchmark and learning environment. *IEEE Robotics and Automation Letters*, 5(2):3019–3026, 2020.

K. Kawaharazuka, T. Matsushima, A. Gambardella, et al. Real-world robot applications of foundation models: A review. *Advanced Robotics*, 38(18):1232–1254, 2024.

M. J. Kim, K. Pertsch, S. Karamcheti, et al. OpenVLA: An open-source vision-language-action model. In *Proc. Conf. Robot Learning (CoRL)*, pp. 2679–2713, 2025.

M. Kobayashi, T. Buamanee, and Y. Uranishi. DABI: Evaluation of data augmentation methods using downsampling in bilateral control-based imitation learning with images. In *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 16892–16898, 2025.

M. Laskin, K. Lee, A. Stooke, et al. Reinforcement learning with augmented data. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

F. Lin, Y. Hu, P. Sheng, et al. Data scaling laws in imitation learning for robotic manipulation. In *Proc. Int. Conf. Learning Representations (ICLR)*, 2025.

I.-C. A. Liu, J. Chen, G. S. Sukhatme, et al. D-CODA: Diffusion for coordinated dual-arm data augmentation. In *Proc. Conf. Robot Learning (CoRL)*, pp. 3569–3588, 2025.

A. Mandlekar, S. Nasiriany, B. Wen, et al. MimicGen: A data generation system for scalable robot learning using human demonstrations. In *Proc. Conf. Robot Learning (CoRL)*, pp. 1820–1864, 2023.

B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *Proc. Eur. Conf. Computer Vision (ECCV)*, pp. 405–421, 2020.

M. Mohri and A. Rostamizadeh. Stability bounds for stationary φ-mixing and β-mixing processes. *Journal of Machine Learning Research*, 11:789–814, 2010.

Open X-Embodiment Collaboration. Open X-Embodiment: Robotic learning datasets and RT-X models. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6892–6903, 2024.

OpenAI. GPT-4 technical report. arXiv preprint arXiv:2303.08774, 2023.

T. Osa, J. Pajarinen, G. Neumann, et al. An algorithmic perspective on imitation learning. *Foundations and Trends in Robotics*, 7(1–2):1–179, 2018.

D. Pollard. *Convergence of Stochastic Processes*. Springer Series in Statistics. Springer, New York, 1984.

D. A. Pomerleau. ALVINN: An autonomous land vehicle in a neural network. In *Advances in Neural Information Processing Systems (NIPS)*, pp. 305–313, 1988.

N. Rajaraman, L. F. Yang, J. Jiao, et al. Toward the fundamental limits of imitation learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

P. Rashidinejad, B. Zhu, C. Ma, et al. Bridging offline reinforcement learning and imitation learning: A tale of pessimism. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, pp. 11702–11716, 2021.

R. Rombach, A. Blattmann, D. Lorenz, et al. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695, 2022.

S. Ross and J. A. Bagnell. Efficient reductions for imitation learning. In *Proc. Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, pp. 661–668, 2010.

S. Ross and J. A. Bagnell. Reinforcement and imitation learning via interactive no-regret learning. arXiv preprint arXiv:1406.5979, 2014.

S. Ross, G. J. Gordon, and J. A. Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proc. Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, pp. 627–635, 2011.

C. V. Samak, T. V. Samak, B. Li, et al. Sim2real diffusion: Leveraging foundation vision language models for adaptive automated driving. *IEEE Robotics and Automation Letters*, 11(1):177–184, 2026.

W. Wan, J. Fu, X. Yuan, et al. LodeStar: Long-horizon dexterity via synthetic data augmentation from human demonstrations. In *Proc. Conf. Robot Learning (CoRL)*, pp. 4994–5021, 2025.

D. Yarats, I. Kostrikov, and R. Fergus. Image augmentation is all you need: Regularizing deep reinforcement learning from pixels. In *Proc. Int. Conf. Learning Representations (ICLR)*, 2021.

B. Yu. Rates of convergence for empirical processes of stationary mixing sequences. *The Annals of Probability*, 22(1):94–116, 1994.

T. Yu, T. Xiao, J. Tompson, et al. Scaling robot learning with semantically imagined experience. In *Proc. Robotics: Science and Systems (RSS)*, 2023.

H. Zhang, M. Cisse, Y. N. Dauphin, et al. mixup: Beyond empirical risk minimization. In *Proc. Int. Conf. Learning Representations (ICLR)*, 2018.

X. Zhang, M. Chang, P. Kumar, et al. Diffusion meets DAgger: Supercharging eye-in-hand imitation learning. In *Proc. Robotics: Science and Systems (RSS)*, 2024.

T. Z. Zhao, V. Kumar, S. Levine, et al. Learning fine-grained bimanual manipulation with low-cost hardware. In *Proc. Robotics: Science and Systems (RSS)*, 2023.

B. Zitkovich, T. Yu, S. Xu, et al. RT-2: Vision-language-action models transfer web knowledge to robotic control. In *Proc. Conf. Robot Learning (CoRL)*, pp. 2165–2183, 2023.

A PROOFS

This appendix provides detailed proofs of Theorem 1, Corollary 1, and Theorem 2. The proof of the main theorem consists of three steps. First, we decompose the excess behavior-cloning risk into a statistical generalization term and an augmentation-induced bias term. Second, we control the statistical term by combining a Rademacher generalization bound with a reduction for temporally dependent β-mixing trajectories. Finally, we bound the augmentation-bias term by relating the deviation of an augmented trajectory from its ideal expert-consistent counterpart to the augmentation error $\varepsilon_{\mathrm{aug}}$.

Throughout the appendix, constants denoted by $C, C', C_1, C_2, \ldots$ may change from line to line but do not depend on N_{real}, N_{aug}, or δ.

A.1 PRELIMINARIES AND SUPPORTING LEMMAS

For convenience, define the trajectory-level empirical loss

$$\ell_\pi(\tau) \triangleq \frac{1}{T}\sum_{t=1}^{T} \|\pi(s_t) - a_t\|_2^2. \tag{10}$$

Since the per-step squared loss is bounded by B^2, we also have

$$0 \le \ell_\pi(\tau) \le B^2. \tag{11}$$

The empirical risk can therefore be written compactly as

$$\hat{L}(\pi; \mathcal{D}) = \frac{1}{N}\sum_{\tau \in \mathcal{D}} \ell_\pi(\tau), \qquad N = N_{\mathrm{real}} + N_{\mathrm{aug}}. \tag{12}$$

The effective sample size introduced in Section 3 is

$$N_{\mathrm{eff}} = N_{\mathrm{real}} + \gamma_{\mathrm{cover}} N_{\mathrm{aug}}. \tag{13}$$

Thus an augmented trajectory contributes only γ_{cover} effective samples: when it is nearly redundant with the real demonstrations, $\gamma_{\mathrm{cover}} \approx 0$, whereas a highly non-redundant trajectory contributes nearly one additional effective sample.

Lemma 1 (Rademacher generalization bound for bounded classes (Bartlett & Mendelson, 2002))**.** *Let $\mathcal{F}$ be a class of functions satisfying $0 \le f(z) \le c$ for all $f \in \mathcal{F}$ and all z. Given m i.i.d. observations $Z_1, \ldots, Z_m$, let $\hat{\mathfrak{R}}_m(\mathcal{F})$ denote the empirical Rademacher complexity of $\mathcal{F}$. Then for every $\delta \in (0,1)$, with probability at least $1-\delta$,*

$$\sup_{f \in \mathcal{F}} \left(\mathbb{E}[f] - \frac{1}{m}\sum_{i=1}^{m} f(Z_i) \right) \le 2\hat{\mathfrak{R}}_m(\mathcal{F}) + c\sqrt{\frac{2\ln(1/\delta)}{m}}. \tag{14}$$

Remark. Lemma 1 is stated for independent samples. Robot trajectories, however, contain temporal dependence. We next show how this dependence modifies the effective number of independent observations appearing in the bound.

Lemma 2 (Effective sample size under exponential β-mixing)**.** *Suppose the trajectory process satisfies*

$$\beta(k) \le \beta_0 e^{-k/\tau_{\mathrm{mix}}}, \tag{15}$$

where τ_{mix} denotes the mixing time. Then, up to absolute constants, an empirical process based on N_{eff} correlated effective trajectory samples can be controlled as an independent empirical process with

$$m_{\mathrm{eff}} \asymp \frac{N_{\mathrm{eff}}}{\tau_{\mathrm{mix}} + 1} \tag{16}$$

independent observations. Consequently, replacing m by m_{eff} in a standard Rademacher or concentration bound introduces a multiplicative factor of order $\tau_{\mathrm{mix}} + 1$ in its variance term.

Proof. The purpose of the blocking argument is to separate observations far enough apart that their residual dependence is small.

For intuition, partition the trajectory sequence into blocks of length q. A standard independent-block construction separates the retained blocks by blocks of the same length, so that any two retained blocks are at least q time steps apart. Under β-mixing, the discrepancy between the joint law of these retained blocks and the product of their marginal laws is controlled by $\beta(q)$.

An equivalent way of expressing the cumulative dependence is through the dependence factor

$$\kappa_\beta \triangleq 1 + 2\sum_{k=1}^{\infty} \beta(k). \tag{17}$$

Under the exponential assumption equation 15,

$$\begin{aligned} \kappa_\beta &\le 1 + 2\beta_0 \sum_{k=1}^{\infty} e^{-k/\tau_{\mathrm{mix}}} \\ &= 1 + \frac{2\beta_0}{e^{1/\tau_{\mathrm{mix}}} - 1}. \end{aligned} \tag{18}$$

Using $e^x - 1 \ge x$ for $x > 0$ gives

$$\kappa_\beta \le 1 + 2\beta_0 \tau_{\mathrm{mix}} \le C_\beta(\tau_{\mathrm{mix}} + 1) \tag{19}$$

for a constant $C_\beta > 0$.

Standard blocking and coupling arguments for β-mixing empirical processes therefore replace the nominal sample count by an effective independent count proportional to

$$\frac{N_{\mathrm{eff}}}{\kappa_\beta} \ge \frac{N_{\mathrm{eff}}}{C_\beta(\tau_{\mathrm{mix}} + 1)}. \tag{20}$$

Since C_β is independent of the dataset sizes, it can be absorbed into the universal constants of the final bound. Hence we write

$$m_{\mathrm{eff}} = \Theta\left(\frac{N_{\mathrm{eff}}}{\tau_{\mathrm{mix}} + 1}\right), \tag{21}$$

which establishes the claimed scaling. □

Lemma 3 (Pseudo-dimension control of Rademacher complexity)**.** *Let the squared-loss class induced by Π have pseudo-dimension d and be uniformly bounded in $[0, B^2]$. Then for $m \gtrsim \mathsf{d}$,*

$$\hat{\mathfrak{R}}_m(\mathcal{L}(\Pi)) \le CB^2\sqrt{\frac{\mathsf{d}\ln(em/\mathsf{d})}{m}} \tag{22}$$

for an absolute constant $C > 0$.

Proof. The pseudo-dimension assumption gives the standard covering-number bound

$$\ln \mathcal{N}\left(\epsilon, \mathcal{L}(\Pi), \|\cdot\|_{L^2(\hat{P})}\right) \le c\mathsf{d} \ln \frac{cB^2}{\epsilon}, \qquad 0 < \epsilon \le B^2, \tag{23}$$

where $c > 0$ is an absolute constant.

Applying Dudley's entropy integral,

$$\hat{\mathfrak{R}}_m(\mathcal{L}(\Pi)) \le \inf_{\alpha>0} \left[4\alpha + \frac{12}{\sqrt{m}} \int_\alpha^{B^2} \sqrt{\ln \mathcal{N}\left(\epsilon, \mathcal{L}(\Pi), \|\cdot\|_{L^2(\hat{P})}\right)}\, d\epsilon \right]. \tag{24}$$

Substituting equation 23 gives

$$\hat{\mathfrak{R}}_m(\mathcal{L}(\Pi)) \le \inf_{\alpha>0} \left[4\alpha + \frac{C\sqrt{\mathsf{d}}}{\sqrt{m}} \int_\alpha^{B^2} \sqrt{\ln \frac{cB^2}{\epsilon}}\, d\epsilon \right]. \tag{25}$$

Choosing the usual cutoff $\alpha \asymp B^2\sqrt{\mathsf{d}/m}$ and evaluating the entropy integral yields

$$\hat{\mathfrak{R}}_m(\mathcal{L}(\Pi)) \le CB^2 \sqrt{\frac{\mathsf{d} \ln(em/\mathsf{d})}{m}}, \tag{26}$$

after absorbing numerical factors into C. □

A.2 Proof of Theorem 1

Proof. We divide the proof into three steps.

Step 1: Excess-risk decomposition. Recall that

$$\pi_\theta \in \arg\min_{\pi\in\Pi} \hat{L}(\pi; \mathcal{D}). \tag{27}$$

Hence

$$\hat{L}(\pi_\theta; \mathcal{D}) \le \hat{L}(\pi^\star; \mathcal{D}). \tag{28}$$

Add and subtract the empirical risks of π_θ and $\pi^\star$:

$$\begin{aligned} L(\pi_\theta) - L(\pi^\star) = {} & \big[L(\pi_\theta) - \hat{L}(\pi_\theta; \mathcal{D})\big] \\ & + \big[\hat{L}(\pi_\theta; \mathcal{D}) - \hat{L}(\pi^\star; \mathcal{D})\big] \\ & + \big[\hat{L}(\pi^\star; \mathcal{D}) - L(\pi^\star)\big]. \end{aligned} \tag{29}$$

By the ERM property equation 28, the middle term is non-positive. Therefore

$$\begin{aligned} L(\pi_\theta) - L(\pi^\star) \le {} & \sup_{\pi\in\Pi} \big[L(\pi) - \hat{L}(\pi; \mathcal{D})\big] \\ & + \hat{L}(\pi^\star; \mathcal{D}) - L(\pi^\star). \end{aligned} \tag{30}$$

Define

$$\mathrm{Gen}(\mathcal{D}) \triangleq \sup_{\pi\in\Pi} \big[L(\pi) - \hat{L}(\pi; \mathcal{D})\big] \tag{31}$$

and

$$\mathrm{Bias}(\mathcal{D}) \triangleq \hat{L}(\pi^\star; \mathcal{D}) - L(\pi^\star). \tag{32}$$

Then

$$L(\pi_\theta) - L(\pi^\star) \le \mathrm{Gen}(\mathcal{D}) + \mathrm{Bias}(\mathcal{D}). \tag{33}$$

This decomposition isolates the two effects of augmentation: additional non-redundant trajectories reduce $\mathrm{Gen}(\mathcal{D})$, whereas imperfect augmented trajectories increase $\mathrm{Bias}(\mathcal{D})$.

Step 2: Controlling the statistical term. Consider the squared-loss class

$$\mathcal{L}(\Pi) = \left\{(s,a) \mapsto \|\pi(s) - a\|_2^2 : \pi \in \Pi\right\}. \tag{34}$$

By assumption,

$$0 \le \ell_\pi(s,a) \le B^2. \tag{35}$$

The nominal number of trajectories in the mixed dataset is $N = N_{\text{real}} + N_{\text{aug}}$. However, augmented trajectories need not provide statistically independent information. By the definition of trajectory coverage efficiency, their statistical contribution is discounted by γ_{cover}, giving

$$N_{\text{eff}} = N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}. \tag{36}$$

The temporal dependence further reduces the effective independent count. Lemma 2 gives

$$m_{\text{eff}} = \Theta\left(\frac{N_{\text{eff}}}{\tau_{\text{mix}} + 1}\right). \tag{37}$$

Applying Lemma 1 with confidence $\delta/2$ yields

$$\text{Gen}(\mathcal{D}) \le 2\hat{\mathfrak{R}}_{m_{\text{eff}}}(\mathcal{L}(\Pi)) + B^2\sqrt{\frac{2\ln(2/\delta)}{m_{\text{eff}}}}. \tag{38}$$

Using Lemma 3,

$$\hat{\mathfrak{R}}_{m_{\text{eff}}}(\mathcal{L}(\Pi)) \le CB^2\sqrt{\frac{\mathsf{d}\ln(em_{\text{eff}}/\mathsf{d})}{m_{\text{eff}}}}. \tag{39}$$

Substitution into equation 38 gives

$$\begin{aligned}\text{Gen}(\mathcal{D}) \le{}& 2CB^2\sqrt{\frac{\mathsf{d}\ln(em_{\text{eff}}/\mathsf{d})}{m_{\text{eff}}}} \\ &+ B^2\sqrt{\frac{2\ln(2/\delta)}{m_{\text{eff}}}}.\end{aligned} \tag{40}$$

Using

$$\sqrt{x} + \sqrt{y} \le \sqrt{2(x+y)}, \qquad x, y \ge 0, \tag{41}$$

and absorbing numerical factors into an absolute constant,

$$\text{Gen}(\mathcal{D}) \le CB^2\sqrt{\frac{\mathsf{d}\ln(em_{\text{eff}}/\mathsf{d}) + \ln(2/\delta)}{m_{\text{eff}}}}. \tag{42}$$

Finally, substituting $m_{\text{eff}} = \Theta(N_{\text{eff}}/(\tau_{\text{mix}} + 1))$ and using $m_{\text{eff}} \le N_{\text{eff}}$ inside the logarithm, we obtain, after absorbing constant and logarithmic numerical factors,

$$\text{Gen}(\mathcal{D}) \le C_1 B^2\sqrt{\frac{(\tau_{\text{mix}} + 1)\left(\mathsf{d}\ln\frac{N_{\text{eff}}}{\mathsf{d}} + \ln\frac{2}{\delta}\right)}{N_{\text{eff}}}}. \tag{43}$$

Equation equation 43 makes explicit why the raw number N_{aug} is not the relevant quantity: augmentation reduces the statistical term only through the increment $\gamma_{\text{cover}} N_{\text{aug}}$ in effective sample size.

Step 3: Controlling the augmentation-bias term. Because the population risk is measured against the expert policy,

$$L(\pi^\star) = \mathbb{E}_{s\sim\rho^{\pi^\star}}\left[\|\pi^\star(s) - \pi^\star(s)\|_2^2\right] = 0. \tag{44}$$

Therefore

$$\text{Bias}(\mathcal{D}) = \hat{L}(\pi^\star; \mathcal{D}). \tag{45}$$

Separating real and augmented trajectories gives

$$\hat{L}(\pi^\star;\mathcal{D}) = \frac{1}{NT}\sum_{\tau\in\mathcal{D}_{\text{real}}}\sum_{t=1}^{T}\|\pi^\star(s_t)-a_t\|_2^2 + \frac{1}{NT}\sum_{\tilde{\tau}\in\mathcal{D}_{\text{aug}}}\sum_{t=1}^{T}\|\pi^\star(\tilde{s}_t)-\tilde{a}_t\|_2^2. \tag{46}$$

For real expert demonstrations, $a_t = \pi^\star(s_t)$, hence the first term is exactly zero. Thus

$$\text{Bias}(\mathcal{D}) = \frac{N_{\text{aug}}}{N}\hat{L}(\pi^\star;\mathcal{D}_{\text{aug}}), \tag{47}$$

where

$$\hat{L}(\pi^\star;\mathcal{D}_{\text{aug}}) = \frac{1}{N_{\text{aug}}T}\sum_{\tilde{\tau}\in\mathcal{D}_{\text{aug}}}\sum_{t=1}^{T}\|\pi^\star(\tilde{s}_t)-\tilde{a}_t\|_2^2. \tag{48}$$

Fix an augmented trajectory

$$\tilde{\tau} = \big((\tilde{s}_1,\tilde{a}_1),\dots,(\tilde{s}_T,\tilde{a}_T)\big) \tag{49}$$

and let

$$\tau^\star_{\text{target}} = \big((s_1^\star,a_1^\star),\dots,(s_T^\star,a_T^\star)\big), \qquad a_t^\star = \pi^\star(s_t^\star), \tag{50}$$

be the corresponding ideal expert-consistent trajectory at the augmented target configuration.

Insert the ideal expert action $a_t^\star$:

$$\begin{aligned}\|\pi^\star(\tilde{s}_t)-\tilde{a}_t\|_2 &\le \|\pi^\star(\tilde{s}_t)-\pi^\star(s_t^\star)\|_2 + \|\pi^\star(s_t^\star)-\tilde{a}_t\|_2\\ &= \|\pi^\star(\tilde{s}_t)-\pi^\star(s_t^\star)\|_2 + \|a_t^\star-\tilde{a}_t\|_2.\end{aligned} \tag{51}$$

By the $L_{\pi^\star}$-Lipschitzness of the expert,

$$\|\pi^\star(\tilde{s}_t)-\pi^\star(s_t^\star)\|_2 \le L_{\pi^\star}\|\tilde{s}_t-s_t^\star\|_2. \tag{52}$$

Using $(x+y)^2 \le 2x^2+2y^2$,

$$\|\pi^\star(\tilde{s}_t)-\tilde{a}_t\|_2^2 \le 2L_{\pi^\star}^2\|\tilde{s}_t-s_t^\star\|_2^2 + 2\|\tilde{a}_t-a_t^\star\|_2^2. \tag{53}$$

The L_{aug}-Lipschitz assumption on the augmentation operator controls the state–action deviation induced by an imperfect augmentation. At trajectory level, this gives

$$\frac{1}{T}\sum_{t=1}^{T}\big(\|\tilde{s}_t-s_t^\star\|_2^2+\|\tilde{a}_t-a_t^\star\|_2^2\big) \le C_H L_{\text{aug}}^2 d_{\mathcal{T}}(\tilde{\tau},\tau^\star_{\text{target}})^2, \tag{54}$$

where C_H captures the possible accumulation of trajectory-level deviation over the horizon. For finite-horizon dynamics, $C_H = O(H)$ under the assumptions of the theorem.

Without loss of generality we may take $L_{\pi^\star} \ge 1$, since any larger constant remains a valid Lipschitz constant. Averaging equation 53 over time and applying equation 54 therefore gives

$$\frac{1}{T}\sum_{t=1}^{T}\|\pi^\star(\tilde{s}_t)-\tilde{a}_t\|_2^2 \le C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 d_{\mathcal{T}}(\tilde{\tau},\tau^\star_{\text{target}})^2, \tag{55}$$

where $C_2 = O(H)$.

Averaging over augmented trajectories and using the definition of augmentation error,

$$\varepsilon_{\text{aug}}^2 = \mathbb{E}_{\tau\sim P_{\text{real}}}\left[d_{\mathcal{T}}\left(\text{Aug}(\tau),\tau^\star_{\text{target}}\right)^2\right], \tag{56}$$

yields

$$\hat{L}(\pi^\star;\mathcal{D}_{\text{aug}}) \le C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2. \tag{57}$$

Consequently,

$$\text{Bias}(\mathcal{D}) \le \frac{N_{\text{aug}}}{N} C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2. \tag{58}$$

Combining equation 43 and equation 58 with equation 33 gives

$$L(\pi_\theta) - L(\pi^\star) \le C_1 B^2 \sqrt{\frac{(\tau_{\text{mix}} + 1)\left(\mathsf{d} \ln \frac{N_{\text{eff}}}{\mathsf{d}} + \ln \frac{2}{\delta}\right)}{N_{\text{eff}}}} + \frac{N_{\text{aug}}}{N} C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2, \tag{59}$$

which is exactly equation 6. □

A.3 Proof of Corollary 1

Proof. Define

$$b \triangleq C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2 \ge 0. \tag{60}$$

Using

$$N_{\text{eff}} = N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}, \tag{61}$$

the bound in Theorem 1 can be written as

$$L(\pi_\theta) - L(\pi^\star) \le S(N_{\text{aug}}) + B_{\text{aug}}(N_{\text{aug}}), \tag{62}$$

where

$$S(N_{\text{aug}}) = C_1 B^2 \sqrt{\frac{(\tau_{\text{mix}} + 1)\left[\mathsf{d} \ln \frac{N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}}{\mathsf{d}} + \ln \frac{2}{\delta}\right]}{N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}}}}, \tag{63}$$

$$B_{\text{aug}}(N_{\text{aug}}) = \frac{N_{\text{aug}}}{N_{\text{real}} + N_{\text{aug}}} b. \tag{64}$$

We analyze the two terms separately.

When $\gamma_{\text{cover}} > 0$,

$$\frac{dN_{\text{eff}}}{dN_{\text{aug}}} = \gamma_{\text{cover}} > 0, \tag{65}$$

so increasing N_{aug} increases the effective sample size. Ignoring constants, the statistical term scales as

$$S(N_{\text{aug}}) = O\left(\sqrt{\frac{\ln N_{\text{eff}}}{N_{\text{eff}}}}\right). \tag{66}$$

More explicitly, define

$$g(x) = \frac{\mathsf{d} \ln(x/\mathsf{d}) + c_\delta}{x}, \qquad c_\delta = \ln \frac{2}{\delta}. \tag{67}$$

Then

$$g'(x) = \frac{\mathsf{d} - \mathsf{d} \ln(x/\mathsf{d}) - c_\delta}{x^2}. \tag{68}$$

Hence $g'(x) < 0$ once the effective sample size is in the usual non-degenerate regime

$$\mathsf{d} \ln(x/\mathsf{d}) + c_\delta > \mathsf{d}; \tag{69}$$

in particular, $x \ge e\mathsf{d}$ is sufficient. Thus $S(N_{\text{aug}})$ decreases with augmentation once the standard generalization regime is reached.

For the bias term,

$$\frac{d}{dN_{\text{aug}}} B_{\text{aug}}(N_{\text{aug}}) = \frac{b N_{\text{real}}}{(N_{\text{real}} + N_{\text{aug}})^2}. \tag{70}$$

Therefore, whenever $b > 0$,

$$\frac{d}{dN_{\text{aug}}} B_{\text{aug}}(N_{\text{aug}}) > 0, \tag{71}$$

and moreover

$$\lim_{N_{\text{aug}} \to \infty} B_{\text{aug}}(N_{\text{aug}}) = b. \tag{72}$$

Increasing augmentation thus has two competing effects: it increases N_{eff} and decreases statistical estimation error, but it simultaneously increases the weight assigned to biased augmented data. This is precisely the stated information-gain versus augmentation-bias tradeoff. □

A.4 Proof of Theorem 2

Proof. Let

$$\mathcal{U}(N_{\text{aug}}) \triangleq S(N_{\text{aug}}) + B_{\text{aug}}(N_{\text{aug}}) \tag{73}$$

denote the right-hand side of the generalization bound.

Because $\gamma_{\text{cover}} > 0$,

$$N_{\text{eff}} = N_{\text{real}} + \gamma_{\text{cover}} N_{\text{aug}} \longrightarrow \infty \qquad \text{as } N_{\text{aug}} \to \infty. \tag{74}$$

Since

$$\lim_{x \to \infty} \frac{\ln x}{x} = 0, \tag{75}$$

we obtain

$$\lim_{N_{\text{aug}} \to \infty} \frac{\mathsf{d} \ln(N_{\text{eff}}/\mathsf{d}) + \ln(2/\delta)}{N_{\text{eff}}} = 0. \tag{76}$$

Therefore

$$\lim_{N_{\text{aug}} \to \infty} S(N_{\text{aug}}) = 0. \tag{77}$$

For the augmentation-bias term,

$$\frac{N_{\text{aug}}}{N_{\text{real}} + N_{\text{aug}}} = \frac{1}{N_{\text{real}}/N_{\text{aug}} + 1} \longrightarrow 1. \tag{78}$$

Hence

$$\begin{aligned} \lim_{N_{\text{aug}} \to \infty} B_{\text{aug}}(N_{\text{aug}}) &= C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2 \\ &\triangleq \eta_{\text{sat}}. \end{aligned} \tag{79}$$

Combining equation 77 and equation 79,

$$\boxed{\lim_{N_{\text{aug}} \to \infty} \mathcal{U}(N_{\text{aug}}) = \eta_{\text{sat}} = C_2 L_{\pi^\star}^2 L_{\text{aug}}^2 \varepsilon_{\text{aug}}^2}. \tag{80}$$

Thus increasing the number of augmented trajectories can asymptotically eliminate the statistical component of the derived generalization bound, but it cannot eliminate the augmentation-bias component when $\varepsilon_{\text{aug}} > 0$. Consequently, the *generalization guarantee* exhibits a non-vanishing asymptotic floor determined by augmentation quality. Breaking this floor requires reducing augmentation error rather than merely generating additional samples. □

B Implementation Details

Push-T (theory verification). We conduct controlled theory-verification experiments on the LeRobot Push-T dataset, which contains 206 expert demonstrations with an 18-dimensional observation consisting of the agent position and eight object keypoints, and a 2-dimensional action. We use a fixed random split of 30 trajectories for testing and the remaining 176 as the training pool, shared across all controlled experiments. All spatial coordinates are normalized to $[-1, 1]^2$. Given a real-data budget N_{real}, we randomly sample the corresponding number of trajectories from the training pool and derive multiple augmented trajectories from each sampled demonstration.

We train a lightweight behavior-cloning policy implemented as a two-layer MLP using the squared action loss. Each policy is optimized with Adam at a learning rate of 10^{-3} for 4000 gradient steps with batch size 256. Performance is measured by the action MSE on the held-out trajectories in the normalized coordinate space. Each configuration is repeated with 10 random seeds, and we report the mean with standard error.

To isolate the quantities appearing in our theoretical analysis, we construct three controlled studies. For the augmentation-error study, we inject a constant directional bias into the augmented actions while keeping the observations unchanged, allowing ε_{aug} to be varied directly. For the coverage

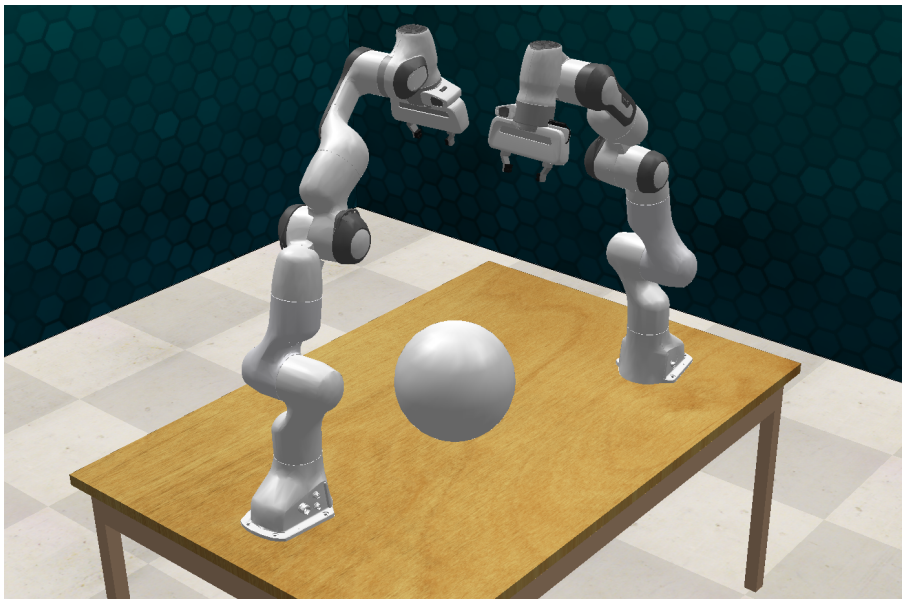

Figure 7: **Simulation environment and task setup for *Coordinated Lift Ball*.** The two robotic arms are required to cooperatively lift the target ball from the tabletop to a specified height while preventing it from falling.

study, we apply trajectory-level tapered translations with different maximum displacement ranges while smoothly reducing the displacement to zero at the end of the trajectory; the displacement range serves as a practical proxy for varying γ_{cover}. Finally, for the augmentation-quantity study, we vary the number of augmented trajectories generated by three terminal-state-preserving geometric augmentation schemes: tapered rotation, rotation retargeting, and tapered translation. These controlled interventions allow us to examine the effects of augmentation error, trajectory coverage, and augmentation quantity largely independently.

RLBench bimanual tasks (system evaluation). We evaluate NYC in the CoppeliaSim-based RLBench environment using five representative bimanual manipulation tasks from the PerAct2 benchmark Grotz et al. (2024); James et al. (2020). The environment is instantiated with two Franka Panda arms (`dual_panda`) and follows the standard RLBench/PyRep simulation pipeline. The five tasks cover different forms of spatial, temporal, and physical coordination between the two robot arms: *Coordinated Lift Ball* requires both arms to jointly grasp and lift a large ball; *Coordinated Push Box* requires coordinated pushing of a box into a target region; *Coordinated Lift Tray* requires both arms to lift a tray while maintaining its stability; *Dual Push Buttons* requires the two arms to press two specified buttons simultaneously; and *Bimanual Straighten Rope* requires manipulating the two ends of a deformable rope into their corresponding target regions. All available task variations are enabled, and the variation used for each episode is randomly sampled from the task-specific variation set. Taking *Coordinated Lift Ball* as an example, the simulation environment and task setup are shown in Figure 7.

Following prior bimanual evaluation protocols Liu et al. (2025), we modify several aspects of the benchmark, including the task-scene configuration and success-condition criteria, to reduce excessive environmental variation and enable a clearer comparison among augmentation methods.

Data collection and evaluation protocol. We collect successful demonstrations in CoppeliaSim using the built-in RLBench motion planner. Although the data-generation pipeline records observations from six cameras, the policy and augmentation pipeline use only the two wrist-mounted eye-in-hand cameras, corresponding to the left and right arms. Each camera provides RGB and depth observations at a resolution of 128×128, together with the corresponding camera intrinsics and poses. The demonstrations also contain the robot low-dimensional state, including joint and gripper information.

We compare five training-data settings: **Real-only**, containing 100 original demonstrations; **More-real**, containing 200 original demonstrations; and **Two-DMD**, **D-CODA**, and **NYC**, each constructed by augmenting the same 100 Real-only demonstrations with an additional 100 generated trajectories, resulting in 200 training trajectories in total. Thus, all augmentation-based settings contain the same number of trajectories as More-real, allowing improvements due to augmentation quality to be distinguished from those due simply to a larger training set.

To isolate the effect of the training data, all settings use the same ACT policy architecture and training protocol, with each policy trained for 260k gradient steps. At test time, we generate 25 task instances for each task according to the corresponding RLBench task specification and roll out the trained policy once in each instance. We repeat the complete training and evaluation procedure with three random seeds and report the mean success rate across the three runs.

Training the diffusion repair model. We adopt a latent diffusion architecture similar to those used in DMD and D-CODA (Zhang et al., 2024; Liu et al., 2025), but use a substantially different training procedure tailored to visual restoration. Specifically, we remove their original perturbation-offset conditioning and train the model to recover a clean observation from a geometrically transformed image containing blur, holes, and projection artifacts. The diffusion model operates in the latent space of a frozen VQGAN encoder–decoder, with a latent representation of size $32 \times 32 \times 4$. We use an NCSN++-based score network with a base channel width of 128 and two residual blocks per resolution level, together with a variance-exploding SDE (VE-SDE) with $\sigma_{\min} = 0.01$.

To construct paired training data without additional annotation, we start from the original trajectories of each task and apply two successive geometry-based viewpoint transformations to the RGB images using the corresponding depth maps. The two transformations have the same magnitude but opposite directions, so that the final image is approximately aligned with the original viewpoint while retaining artifacts introduced by geometric warping. During data generation, the translational perturbation magnitude is sampled between 0.01m and 0.05m, while each Euler-angle perturbation is bounded by 0.5°. This procedure yields a natural training pair in which the twice-warped image serves as the corrupted input and the original RGB image serves as the clean target. For each task, we construct approximately 50,000 such training pairs from 100 demonstrations and train a separate diffusion repair model.

The model is optimized using Adam with a learning rate of 2×10^{-4}, a global batch size of 32, and 3,000 warm-up steps. Each diffusion repair model is trained for a total of 39,000 optimization steps. We apply gradient clipping with a maximum norm of 1.0 and maintain an exponential moving average of the model parameters with a decay rate of 0.999. Training is performed with distributed data parallelism on four NVIDIA A800 GPUs. The main architecture and training hyperparameters are summarized in Table 3.

Table 3: Main configuration of the diffusion repair model.

Component	**Configuration**
Latent representation	$32 \times 32 \times 4$
VQGAN codebook size	16,384
Score network	NCSN++, base width 128
Residual blocks	2 per resolution level
SDE	VE-SDE
$\sigma_{\min}$	0.01
Training pairs per task	$\sim$50,000
Demonstrations per task	100
Translation perturbation	0.01–0.05 m
Rotation perturbation	$\pm 0.5^\circ$ per Euler angle
Optimizer	Adam
Learning rate	2×10^{-4}
Global batch size	32
Warm-up	3,000 steps
Training steps	39,000
Gradient clipping	1.0
EMA decay	0.999
Training hardware	$4\times$ NVIDIA A800

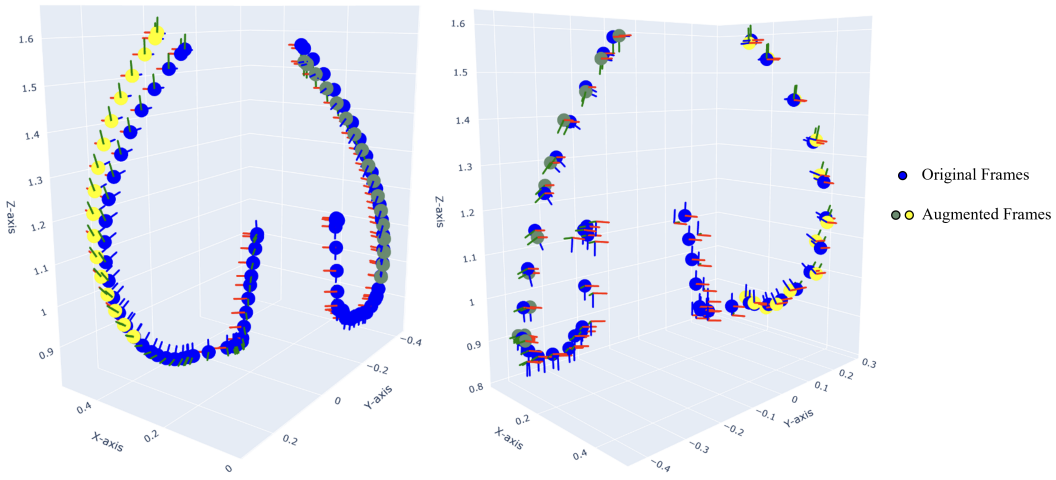

Figure 8: **Comparison between trajectory-wise and point-wise augmentation.** Trajectory-wise augmentation constructs a new and smooth trajectory from the original one while preserving its underlying action characteristics, whereas point-wise augmentation perturbs individual frames independently, resulting in incoherent and noisy trajectories. Moreover, our trajectory-wise augmentation supports denser augmentations with larger offsets from the original trajectory.

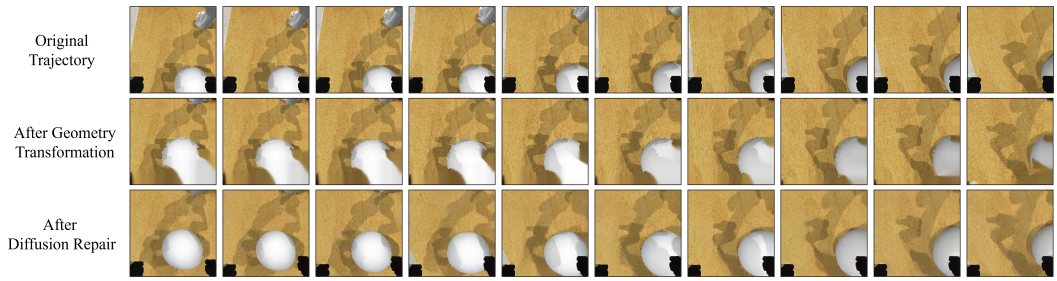

Figure 9: **Examples of the NYC data-generation process.** From top to bottom, the first row shows the original observations from the real dataset; the second row shows the geometry-transformed observations, which contain holes, projection artifacts, and blur introduced by the geometric transformation; and the third row shows the corresponding diffusion-repaired observations. The repair module removes these visual defects while preserving the prescribed target geometry. The repaired observations are then used to construct the augmented dataset for policy training.

NYC data generation pipeline. To generate the augmented dataset, we start from each original demonstration and specify a trajectory-level transformation with a prescribed offset magnitude and direction. For each observation, we first apply the corresponding geometry-aware transformation using the depth map and camera parameters, producing a warped observation at the target viewpoint. This geometric transformation preserves the desired spatial displacement but may introduce holes, occlusion artifacts, and local projection distortions. We then use the conditional diffusion repair module to restore these corrupted regions and obtain the final augmented observation. Applying this procedure consistently along the trajectory produces a new sequence of augmented observations corresponding to the prescribed trajectory transformation; the associated action labels are subsequently recomputed using a Levenberg–Marquardt (LM) inverse-kinematics solver, as described in Section 4.1. Figure 8 shows the difference between trajectory-wise and point-wise augmentation, while Figure 9 visualizes several examples at different stages.

ACT policy training and evaluation. We use ACT (Zhao et al., 2023) as the downstream policy learner for all data settings and keep its architecture and optimization configuration identical across

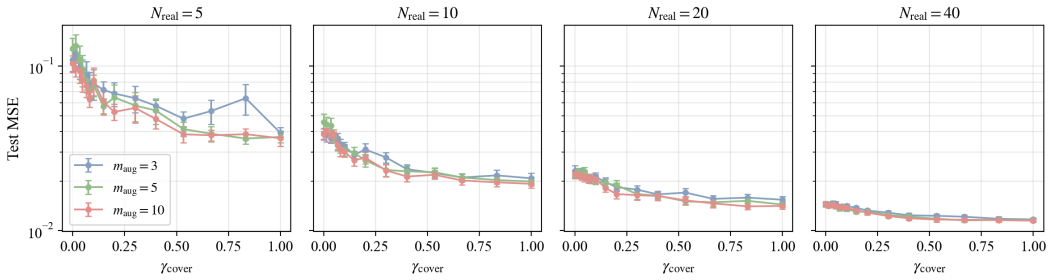

Figure 10: **Push-T test MSE versus trajectory coverage.** Higher coverage generally reduces test error, with stronger gains in lower-data regimes.

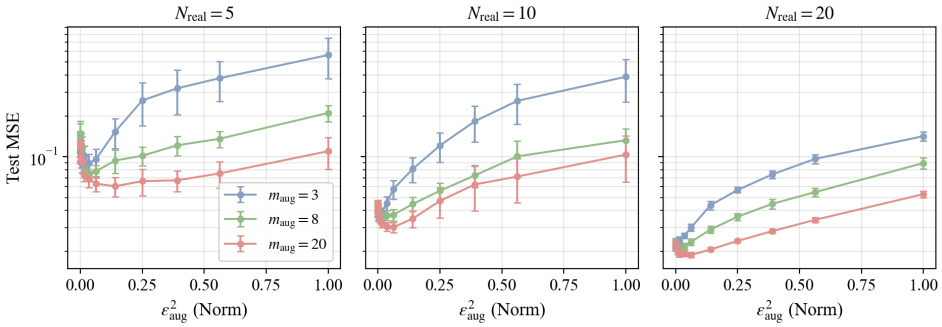

Figure 11: **Push-T test MSE versus augmentation error.** Larger systematic error increases test loss, and additional augmented data does not eliminate the resulting excess error.

methods. The policy receives RGB observations from the left and right wrist-mounted cameras at a resolution of 128×128, together with a 16-dimensional proprioceptive state consisting of the seven joint positions and one gripper state for each arm. Visual features are extracted using a ResNet-18 backbone with sine positional embeddings and are processed by a Transformer with four encoder layers and seven decoder layers. The Transformer has a hidden dimension of 512, eight attention heads, a feed-forward dimension of 3200, and a dropout rate of 0.1.

Following the standard ACT formulation, the policy predicts action chunks rather than individual actions. We use a chunk size of 10 with temporal aggregation enabled, conditioning on one previous action and predicting the next 10 actions. The KL regularization weight is set to 100. We train each policy using a batch size of 16 for 260k gradient steps with a learning rate of 1×10^{-4}, weight decay of 1×10^{-6}, and gradient clipping at 0.1. The ResNet-18 backbone uses a learning rate of 1×10^{-5}. The same built-in demonstration augmentation configuration is applied to all compared methods, ensuring that the only difference between experiments is the composition of the training trajectories.

Policies are evaluated by closed-loop rollout in the same RLBench bimanual environment used for data collection. Both arms are controlled in joint-position space, and each rollout is allowed a maximum horizon of 400 environment steps. For each task and trained policy, we evaluate on 25 independently generated task instances and report the fraction of successful rollouts as the task success rate. The evaluation uses the same two wrist-mounted cameras and observation configuration as during training. To reduce sensitivity to the exact optimization stage, we evaluate multiple late-training checkpoints saved between 200k and 260k gradient steps. All data settings use exactly the same evaluation protocol and environment configuration.

C Detailed Experimental Results

Figures 10 and 11 present the complete results of the Push-T experiments, showing test MSE versus trajectory coverage and test MSE versus augmentation error, respectively.

Table 4: Detailed success rates on all five tasks × five methods × three runs.

Task	Method	Run 1	Run 2	Run 3	Mean	Std
Co-Lift Ball	Real-only	0.44	0.36	0.32	0.37	0.05
	More-real	0.64	0.44	0.40	0.49	0.10
	Two-DMD	0.52	0.36	0.28	0.39	0.10
	D-CODA	0.52	0.36	0.32	0.40	0.09
	NYC (Ours)	0.60	0.52	0.56	**0.56**	**0.03**
Co-Push Box	Real-only	0.68	0.60	0.44	0.57	0.10
	More-real	0.32	0.32	0.28	0.31	0.02
	Two-DMD	0.84	0.72	0.56	0.71	0.11
	D-CODA	0.84	0.72	0.72	0.76	0.06
	NYC (Ours)	0.80	0.76	0.76	**0.77**	**0.02**
Co-Lift Tray	Real-only	0.08	0.08	0.04	0.07	0.02
	More-real	0.08	0.00	0.00	0.03	0.04
	Two-DMD	0.20	0.12	0.12	0.15	0.04
	D-CODA	0.36	0.32	0.24	**0.31**	**0.05**
	NYC (Ours)	0.20	0.20	0.16	0.19	0.02
Dual Push Buttons	Real-only	0.20	0.36	0.28	0.28	0.07
	More-real	0.48	0.44	0.36	0.43	0.05
	Two-DMD	0.52	0.40	0.36	0.43	0.07
	D-CODA	0.40	0.52	0.32	0.41	0.08
	NYC (Ours)	0.52	0.52	0.44	**0.49**	**0.04**
Bi-Straighten Rope	Real-only	0.20	0.20	0.16	0.19	0.02
	More-real	0.52	0.40	0.40	**0.44**	**0.06**
	Two-DMD	0.24	0.20	0.20	0.21	0.02
	D-CODA	0.28	0.16	0.16	0.20	0.06
	NYC (Ours)	0.36	0.20	0.16	0.24	0.09

Table 4 reports the complete success rates of all five training-data settings across the five bimanual RLBench tasks.

D Additional Experiments

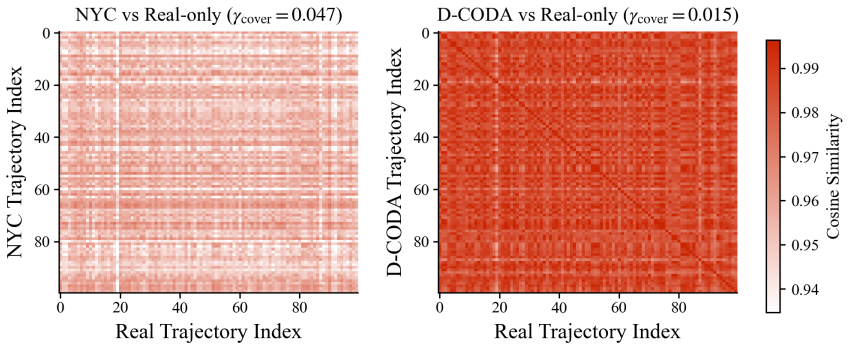

Figure 12: **Pairwise cosine similarity between augmented and real trajectories.** NYC shows substantially lower similarity than D-CODA, consistent with its higher γ_{cover}.

Similarity heatmap. To more intuitively assess the ability of NYC to introduce novel trajectory information compared with other data augmentation systems, we use ResNet-50 to extract visual features from trajectory images. We then compute the feature similarity between the augmented datasets generated by NYC and D-CODA and the original dataset, respectively. The resulting similarity heatmaps are shown in Figure 12. NYC exhibits substantially lower similarity to the original trajectories than D-CODA, indicating that it introduces more diverse trajectory information rather

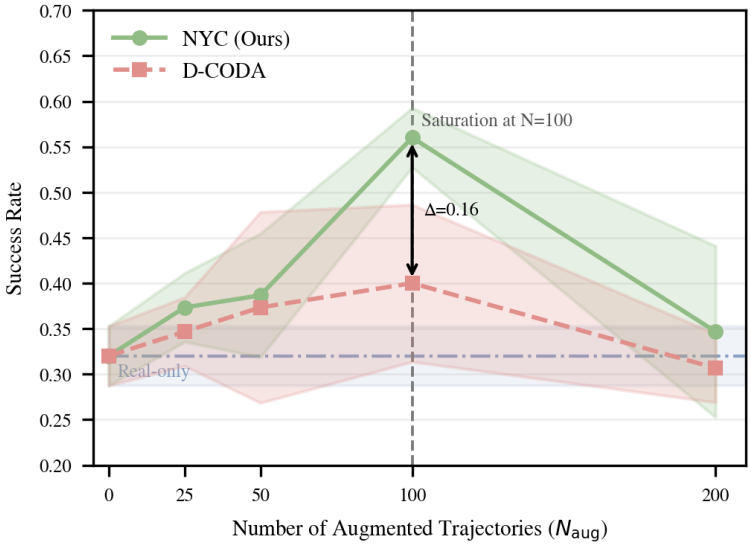

Figure 13: **Success rate versus augmentation count on Coordinated Lift Ball.** Both methods exhibit improve-then-saturate behavior, while NYC reaches a higher performance level.

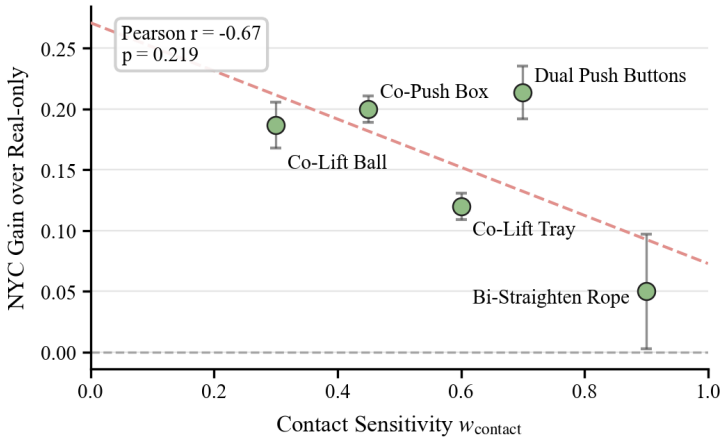

Figure 14: **Success-rate gain over Real-only versus contact sensitivity w_{contact}.** The improvement over Real-only gradually decreases as the task's contact sensitivity increases.

than merely reproducing variations close to the original data. These results further demonstrate that NYC provides richer novel information for robot policy training.

Saturation with augmentation quantity. To further validate the predicted saturation behavior in a practical robotic system, we additionally evaluate the effect of augmentation quantity on training performance in the *Coordinated Lift Ball* task. Figure 13 shows that both NYC and D-CODA initially improve and then saturate, while NYC reaches a substantially higher performance level. D-CODA eventually falls below the no-augmentation baseline, consistent with a larger bias contribution. Thus, the saturation behavior predicted by our analysis also appears in the full bimanual setting.

Contact sensitivity analysis. To further investigate the remaining performance bottlenecks of NYC and identify directions for future improvement, we introduce a task-level contact-sensitivity score w_{contact}. The score characterizes how sensitive task success is to small errors around physical interaction, considering several factors such as whether explicit grasping is required, the size of the

manipulated object or contact region, and the tolerance of the task success condition to positional deviations. Tasks involving tighter contact constraints, smaller objects, or more precise bimanual coordination are assigned higher contact sensitivity.

We then analyze the relationship between w_{contact} and the success-rate improvement of NYC over the Real-only baseline. As shown in Figure 14, the performance gain tends to decrease as contact sensitivity increases. This trend suggests that residual visual inaccuracies around contact-critical regions become increasingly important for tasks requiring precise physical interaction. The analysis therefore highlights a promising direction for further improving NYC: enhancing the diffusion-based visual repair module, particularly its ability to preserve fine-grained geometry and object boundaries in contact-sensitive regions.