Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

EfficientDexGrasp: Diverse and Physically Feasible Dexterous Grasping via Few-Step Diffusion Distillation

Hanyu Huang
City University of Macau
Geying Chen
Shenzhen University
Wen Zhao
zhaowen@nwpu.edu.cn

China University of Mining and Technology - Beijing

Article

Keywords:

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11533891/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

EfficientDexGrasp: Diverse and Physically Feasible Dexterous Grasping via Few-Step Diffusion Distillation

Hanyu Huang[1], **Geying Chen**[2], **and Wen Zhao**[3,4,*]

[1]Faculty of Data Science, City University of Macau, Macao Special Administrative Region, China
[2]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China
[3]State Key Laboratory for Fine Exploration and Intelligent Development of Coal Resources, China University of Mining & Technology, Beijing, China
[4]School of Civil Aviation, Northwestern Polytechnical University, Xi'an, China
[*]zhaowen@nwpu.edu.cn

ABSTRACT

Diffusion models dominate dexterous grasp generation, but their iterative sampling is too slow for latency-sensitive deployment. Distilling a multi-step diffusion teacher into a few-step student exposes two coupled failure modes: distinct grasp modes collapse into a single averaged pose that corresponds to no valid grasp, and the physical refinement the teacher performs at every denoising step is lost, so many student predictions drift into penetration or self-collision. We propose EfficientDexGrasp, a physics-guided distillation framework that addresses both. Contrastive energy distillation shapes an auxiliary energy landscape in which distinct grasp modes remain separate low-energy basins. Gated physics-guided consistency trajectory distillation scales correction with how much worse the student is than the teacher, so predictions at or beyond the teacher's level receive progressively less. The two modules reinforce each other: correction stays on genuinely degraded predictions and leaves the multimodal structure intact. On DexGraspNet, MultiDex, RealDex, and DexGRAB, EfficientDexGrasp achieves $11\times$ to $69\times$ inference speedups and surpasses the strongest prior baselines by 3.8 to 5.8 points on the stringent Suc.6 metric in the 8-step setting. We further demonstrate deployability on a real robot.

Introduction

Robotic dexterous grasping underpins many downstream manipulation skills, from tool use to in-hand reorientation. Classical methods build on force closure and contact-stability analysis[1,2]. These give closed-form quality metrics and theoretical guarantees. But they rely on explicit contact modeling, and they scale poorly to the high-dimensional, tightly coupled configuration space of multi-fingered hands. Large-scale datasets—DexGraspNet[3], MultiDex[4], RealDex[5], and DexGRAB[6,7]—have shifted the field toward data-driven methods. Among these, diffusion models[8–10] stand out. Iterative denoising naturally captures the multimodal grasp distribution, where one object admits several distinct yet equally valid hand configurations. SceneDiffuser[11], UGG[12], DexGrasp Anything[7], and EvolvingGrasp[13] add physics-aware sampling and preference alignment. They report the strongest results on standard benchmarks.

This expressiveness comes at a cost. Sequential denoising is slow. DexGrasp Anything, for example, needs about 100 network evaluations and tens of seconds per batch. That is two orders of magnitude above the sub-second budget of interactive deployment. Diffusion distillation[14,15] can compress these trajectories into a few evaluations. In dexterous grasping, however, naïve distillation triggers two failure modes.

Multimodal collapse. Distinct grasps of one object—around the handle, the rim, or the body of a mug—are separate modes of the distribution. Pointwise distillation losses regress toward the mode-average. That average is a pose lying between valid grasps, matching none. The student then returns nearly identical predictions per object, discarding the teacher's diversity.

Loss of physical plausibility. A diffusion teacher fixes physical violations gradually, one small correction per denoising step. Compressing the trajectory to two or eight steps removes most of these corrections. Many student outputs then end up in penetration or self-collision.

The two interact. The usual fix for the second is an absolute physics penalty, and it worsens the first. In the few-step regime, a uniform penalty over-corrects already-feasible samples. It wastes gradient capacity needed for the degraded ones. And it pulls every sample toward the single lowest-potential pose, erasing diversity. We therefore keep the physics term, but gate it and reference it to the teacher. Correction then concentrates on samples that have actually left the feasible manifold.

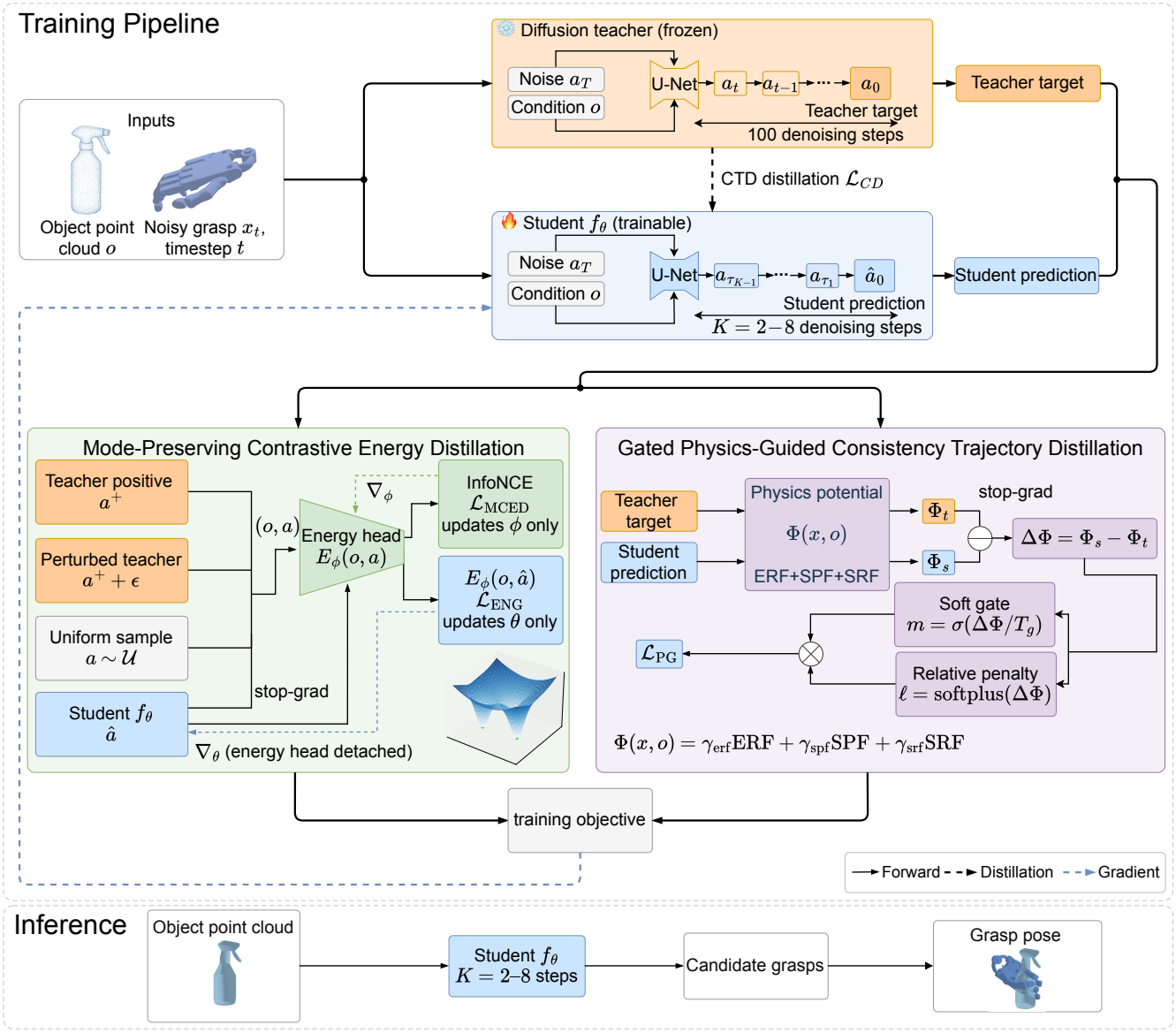


Figure 1. Overview of EfficientDexGrasp. We distill a frozen multi-step *teacher* U-Net$_\psi$ into a *few-step student* f_θ, and two modules are used here. On the left, **MCED** trains an auxiliary energy head $E_\phi(o, a)$, and it has two separate goals: teacher grasps serve as positives against a pool of three kinds of negatives, and the InfoNCE loss on them updates ϕ alone, where student predictions are detached; the student then gets the energy signal from $E_\phi(o, \hat{a})$, with the head detached this time, so these two goals touch different parameters. On the right, **PG-CTD** compares the physics potential of the student with that of the teacher, a teacher-relative penalty is applied, and the per-sample degradation $\Delta\Phi$ gates it, so correction goes mainly to predictions that are physically worse than the teacher. At the bottom, the student alone produces candidate grasps within 2–8 passes, and one of them is then executed on the hand.

We propose **EfficientDexGrasp**, a physics-guided distillation framework with three modules. *Consistency Trajectory Distillation* (CTD) is the acceleration backbone. It maps any point on the teacher's probability-flow trajectory to its clean endpoint in one pass, compressing the 100-step teacher into a 2–8-step student for an 11×–69× speedup. *Mode-preserving Contrastive Energy Distillation* (MCED) replaces pointwise regression with an InfoNCE energy objective. It carves one low-energy basin per grasp mode, so modes survive as separate attractors. *Gated Physics-Guided Consistency Trajectory Distillation* (PG-CTD) adds a degradation-aware gate that scales correction with how much worse the student is than the teacher. A teacher-relative penalty then corrects this gap, and predictions at or beyond the teacher's level receive progressively less. Because the gate references the teacher rather than a zero-violation target, it preserves the structure MCED builds.

The closest prior work is EvolvingGrasp[13], which also pairs a consistency model with physics terms. It differs from ours in two respects. Its physics-aware consistency model applies an *absolute* physics penalty with a fixed target, whereas PG-CTD applies a *teacher-relative*, gated penalty that withdraws as the student reaches the teacher's level. And it contains no mechanism against mode collapse, whereas MCED supervises the energy geometry of the action space explicitly. The ablation study isolates the contribution of each addition.

Our contributions are as follows.

1. We identify multimodal collapse and physical infeasibility as the two dominant failure modes of naïve diffusion distillation for dexterous grasping, and present a framework that addresses both under few-step inference, with an 11×–69× speedup over the 100-step teacher.

2. We develop *Mode-preserving Contrastive Energy Distillation* (MCED), which recasts distillation as conditional energy learning. Contrastive supervision carves multiple low-energy basins, suppressing mode-averaging.

3. We develop *Gated Physics-Guided Consistency Trajectory Distillation* (PG-CTD), which pairs a degradation-aware gate with a teacher-relative constraint. Physics is thus corrected selectively, avoiding the diversity loss of indiscriminate penalties.

4. On four benchmarks, EfficientDexGrasp achieves an 11×–69× speedup over the teacher. It also improves Suc.6 by 3.8–5.8 absolute points over the strongest baseline, while remaining competitive with the teacher on physical plausibility.

Dexterous grasp generation. Classical formulations rest on force-closure and contact-stability criteria[1,2]. These give tight quality measures, but depend on explicit contact modeling. They also do not scale as finger joints multiply and objects depart from primitive shapes. Large-scale benchmarks[3,16] reoriented the field toward learning from data. Regression networks map point clouds directly to grasp parameters[17–20] in one forward pass. They share a well-known flaw. When several valid grasps coexist, an L_2-type loss is minimized by their conditional mean. That mean sits between modes, and is neither geometrically reasonable nor mechanically stable.

Distribution-modeling generators avoid this collapse by representing the conditional grasp distribution explicitly[21,22]. UniDexGrasp[23,24] decouples proposal generation from a goal-conditioned RL policy. GraspTTA[4,25] refines hand–object contact at test time. Diffusion models have since become the dominant backbone. SceneDiffuser[11] conditions denoising on scene context. UGG and successors[12,26,27] separate finger articulation from wrist pose. DexGrasp Anything[7] folds three differentiable physics terms into the sampling loop. EvolvingGrasp[13] adds direct preference optimization on top of a physics-aware consistency model; its physics term is absolute and it has no explicit mechanism for mode preservation, which are the two points at which our method departs from it. One bottleneck persists across the diffusion-based methods: denoising needs dozens to hundreds of sequential evaluations, giving per-batch latencies of tens of seconds. This is fine for offline planning, but a bottleneck for interactive deployment. We keep the strengths of diffusion—multimodality and physics-aware sampling. We then compress these multi-step policies into few-step models via consistency distillation, without losing diversity or physical plausibility.

Accelerating diffusion models. Two lines reduce denoising cost. Solver-based methods take fewer integration steps along a fixed trajectory[28,29]. But they degrade sharply below roughly four steps—exactly the regime we target. Distillation-based methods instead train a student to reproduce the teacher's multi-step output in fewer evaluations. Progressive distillation[14] halves the step count round by round. Consistency models[15] force every point on a PF ODE trajectory to map to the same clean output, and consistency trajectory models[30] extend this to arbitrary trajectory jumps. This paradigm has entered robot policy learning. Diffusion Policy[10,31] frames action generation as conditional denoising, and Consistency Policy[32], ManiCM[33], and One-Step Diffusion Policy[34] bring consistency and one-step distillation to visuomotor networks. Our backbone shares this spirit, but differs decisively. We pair consistency trajectory distillation with two domain-specific mechanisms: energy-based mode preservation, and gated, teacher-relative physics guidance. Both target failure modes peculiar to dexterous grasping, where an averaged-mode output can drive fingers into the object or into each other rather than merely missing.

Energy-based models and contrastive learning. Energy-based models (EBMs) assign an unnormalized scalar energy to each input–output pair[35]. They handle disconnected modes more naturally than VAEs or normalizing flows. This fits grasp distributions, where the valid set splits into distinct clusters separated by wide infeasible regions. Du and Mordatch[36] showed EBMs generate sharp multi-modal images. Florence *et al.*[37] brought this to robotics via implicit behavioral cloning, learning an energy over the action space and acting at its minimum. Contrastive learning makes EBM training cheaper, since it shapes the energy surface without a partition function[38–40]. In knowledge distillation with a multimodal teacher, an L_2 objective is minimized by the conditional mean of the teacher's modes[41,42], and in the few-step regime we study this shows up as a loss of diversity (the ablation study). Energy-based formulations instead preserve the teacher's topology, because they train relative energies rather than a point prediction[22,43]. MCED builds on this. It supervises the energy geometry of the action space, so distinct grasp strategies survive as separate minima. A concurrent line, CED[44], also uses contrastive energy for diffusion distillation, but in the image domain and with a different design: its energy is a sample-distance metric inside a joint generator–score EBM, its negatives are the generator's past outputs only, and it targets a single-step generator. MCED instead learns an independent head $E_\phi(o,a)$ over observation–action pairs, draws negatives from three sources to cover the multimodal structure of grasp space, runs on a multi-step consistency backbone, and is co-designed with a physics module that has no image-domain analogue.

Results

Table 1. Grasping performance in terms of Suc.6, Suc.1, and Pen. comparison across different methods and datasets. Bold values highlight the best results and underlined values indicate the runner-up. Higher is better for Suc.6 and Suc.1; lower is better for Pen. and Time. Runtimes are as reported in the original papers and were obtained on different hardware; on the same RTX 5090 under identical settings, the teacher takes 17.30 s per batch and our student 0.25 s (2-step) and 1.55 s (8-step). EG-2 / EG-8: EvolvingGrasp at 2 / 8 steps.

Method	DexGraspNet			MultiDex			RealDex			DexGRAB			Time
	Suc.6	Suc.1	Pen.	Suc.6	Suc.1	Pen.	Suc.6	Suc.1	Pen.	Suc.6	Suc.1	Pen.	
UniDexGrasp	33.9	70.1	31.9	21.6	47.5	13.5	27.1	59.4	39.0	20.8	55.8	37.4	$0.46^{\pm 0.11}$
GraspTTA	18.6	67.8	24.5	30.3	62.8	19.0	13.3	46.4	40.1	14.4	51.0	51.4	$10.41^{\pm 0.32}$
SceneDiffuser	26.6	66.9	31.0	69.8	85.6	14.6	21.7	56.1	42.0	39.1	85.0	41.1	$3.41^{\pm 0.13}$
UGG	46.9	79.0	25.2	65.5	93.4	<u>10.3</u>	32.7	63.4	34.4	42.7	90.6	33.2	$38.34^{\pm 2.31}$
DexGrasp Any.	53.6	90.4	21.5	72.2	96.3	**9.6**	34.6	71.2	23.1	56.5	91.8	28.6	$32.91^{\pm 1.34}$
EG-2	62.4	90.5	19.3	65.9	97.2	15.3	44.0	77.8	**19.7**	53.3	92.5	24.3	$0.73^{\pm 0.05}$
EG-8	<u>65.4</u>	92.3	**15.9**	<u>80.3</u>	**98.7**	12.3	<u>64.4</u>	<u>89.1</u>	21.8	<u>60.8</u>	<u>96.4</u>	<u>22.3</u>	$2.71^{\pm 0.08}$
Ours (2-step)	63.1	92.5	19.6	67.6	96.8	14.8	45.8	79.5	<u>20.2</u>	55.3	93.6	23.6	$0.25^{\pm 0.02}$
Ours (8-step)	**70.3**	**94.7**	<u>16.1</u>	**86.1**	<u>98.6</u>	11.9	**68.2**	**90.1**	20.8	**65.2**	**97.3**	**22.5**	$1.55^{\pm 0.03}$

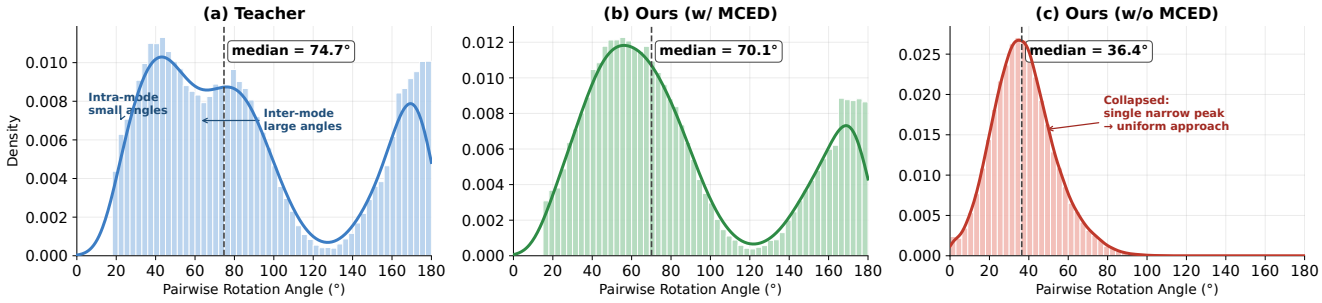


Figure 2. Pairwise rotation angle distribution of generated grasps on the same object. The teacher and our method (w/ MCED) exhibit a bimodal distribution: a left peak corresponding to intra-mode variations (small angles) and a right peak corresponding to inter-mode differences (large angles), indicating the preservation of multiple distinct approach directions. Without MCED, the distribution collapses into a single narrow peak concentrated at small angles (median 36.4° vs. 74.7° / 70.1°), indicating that the generated grasps converge to a uniform approach direction. Shown for one representative MultiDex object; *w/o MCED* is row d of Table 2. Div. over the full MultiDex test set is reported in the text of Results.

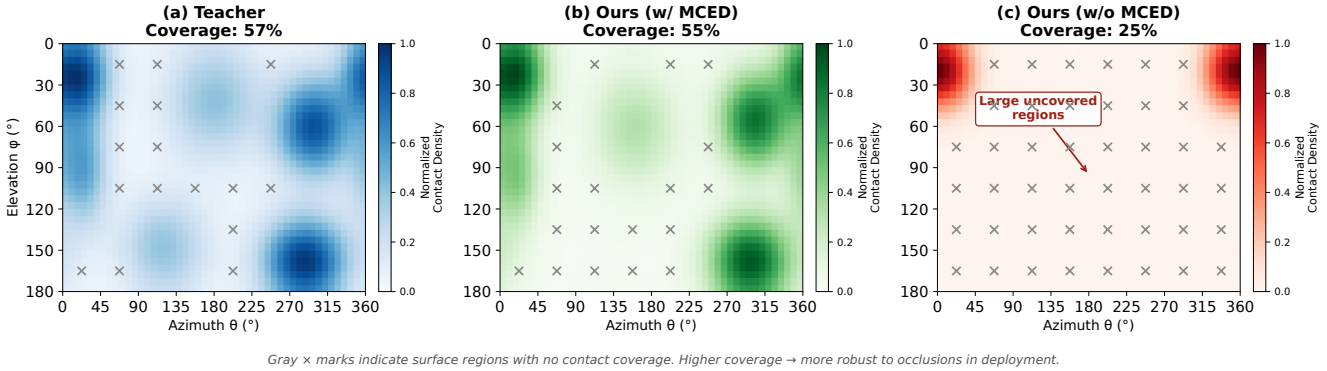


Figure 3. Object surface contact coverage of generated grasps. The object surface is unwrapped into an azimuth–elevation grid, where color intensity indicates contact density and gray $\times$ marks denote uncovered regions. The teacher (57%) and our method with MCED (55%) achieve comparable coverage, indicating that MCED preserves diverse approach directions during distillation. Without MCED (row d of Table 2), coverage drops to 25%, with large surface regions receiving no contact, indicating that the generated grasps collapse to a narrow set of approach directions. Same object as Fig. 2.

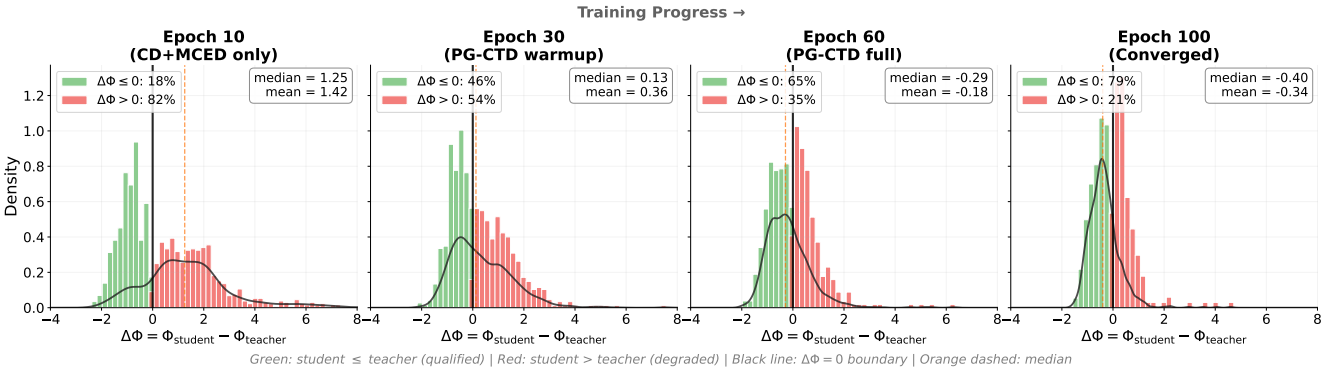


Figure 4. Evolution of the physics degradation distribution $\Delta\Phi = \Phi_{\text{student}} - \Phi_{\text{teacher}}$ across training stages. Green regions indicate samples where the student matches or surpasses the teacher in physical feasibility ($\Delta\Phi \leq 0$), while red regions indicate degraded samples ($\Delta\Phi > 0$). At Epoch 10 (before PG-CTD activation), 82% of samples are physically degraded relative to the teacher. As Gated PG-CTD progressively engages through the warmup schedule, the distribution shifts leftward: the degradation ratio drops to 54% at Epoch 30, 35% at Epoch 60, and 21% at convergence, with the median $\Delta\Phi$ moving from $+1.25$ to -0.40. This is consistent with PG-CTD acting mainly on physically degraded predictions, with the gate deactivating as the student improves.

Main results

Table 1 compares all methods on stability (Suc.6, Suc.1), plausibility (Pen.), and latency; grasp diversity is discussed at the end of this section.

Inference speedup. Table 1 lists runtimes as reported in the original papers, which were obtained on different hardware. For a controlled comparison we re-measured the teacher on the same RTX 5090 under identical settings (Table 1, caption): it takes 17.30 s per batch, against 0.25 s for our two-step student and 1.55 s at eight steps—a 69× and 11× reduction. Baseline runtimes in Table 1 are quoted from their papers and are therefore only indicative.

Grasp stability. At 8 steps, EfficientDexGrasp has the highest Suc.6 on all four benchmarks, beating EvolvingGrasp 8-step by +3.8 to +5.8 points. The largest margin is on MultiDex (+5.8), followed by DexGraspNet (+4.9), which fits the view that mode preservation matters most on geometrically diverse benchmarks, where a collapsed student forfeits the grasp strategies that suit a larger share of the objects. At 2 steps, it still beats EvolvingGrasp 2-step on all four datasets, so MCED and PG-CTD hold up under extreme compression.

At 8 steps, the student also exceeds the teacher's Suc.6 on all four benchmarks, under the same all-samples protocol. We attribute this to two training-time effects. MCED supervises the student against the teacher's grasps as positives and against the teacher's perturbed neighbors as negatives, which biases the student toward the centers of the teacher's modes rather than their fringes. PG-CTD penalizes every training sample that is physically worse than the teacher, so the student's output distribution has a thinner infeasible tail than the teacher's own samples (Fig. 4: at convergence 79% of student predictions are at or above the teacher's physical quality).

Physical plausibility. On DexGraspNet, EfficientDexGrasp (8-step) reaches 16.1 mm penetration—near EvolvingGrasp (15.9 mm) and well below the teacher (21.5 mm); both few-step methods train with an explicit physics term in the loss, whereas the teacher applies its physics correction only inside the sampling loop and still leaves a tail of high-penetration samples. On MultiDex, the teacher's 9.6 mm reflects its 100-step physics-correction budget, which few-step methods forgo by design; our 11.9 mm is second among the few-step methods, behind UGG (10.3 mm). On RealDex and DexGRAB, the top few-step methods cluster tightly (19.7–21.8 mm and 22.3–24.3 mm), showing PG-CTD prevents the catastrophic violations that removing the teacher's iterative correction would otherwise cause.

Grasp diversity. Few-step distillation risks collapsing the teacher's grasp modes, so we also report Div., the mean standard deviation of the pose parameters across successful grasps, following DexGrasp Anything[7], over the full MultiDex test set under the same all-samples protocol. The teacher reaches a Div. of 0.22 and our 8-step student 0.18, so the student retains about 82% of the teacher's diversity after compression from 100 steps to 8. Figs. 2 and 3 show the same effect on a single object, together with the collapse that occurs without MCED. EvolvingGrasp does not report a diversity metric, so no comparison is possible for EG-2 / EG-8.

Ablation study

Table 2. Ablation study on MultiDex at 8 steps, under the same all-samples protocol as Table 1. All variants share the same student network and the same inference procedure, so inference time is identical across rows (1.55 s per batch) and is not repeated here.

	CTM	PGD	MCED	PG-CTD	Suc.6↑	Suc.1↑	Pen.↓
a	✓				68.1	96.5	13.0
b	✓	✓			72.5	97.2	12.5
c	✓		✓		75.1	97.4	12.3
d	✓	✓		✓	78.5	98.5	12.1
e	✓	✓	✓		77.4	97.6	12.2
f	✓	✓	✓	✓	86.1	98.6	11.9

We ablate each module on MultiDex at 8 steps (Table 2), starting from a bare consistency trajectory model (CTM, row a).

Physics-guided distillation (PGD). A static physics penalty (row b) lifts Suc.6 from 68.1 to 72.5 (+4.4) and lowers Pen. from 13.0 to 12.5, with no inference overhead.

MCED. MCED alone (row c) gives the largest single-module gain, +7.0 Suc.6. Under mode collapse, all grasps of an object converge to one configuration, so objects for which that configuration is unstable fail on every sample; a student that retains several strategies succeeds on the subset of samples that use the appropriate one. Fig. 2 shows the diversity: pairwise rotation angles are bimodal (median 70.1°), close to the teacher's (74.7°), versus a single narrow peak at 36.4° without MCED. Surface coverage (Fig. 3) agrees—55% with MCED versus 25% without, leaving three-quarters of the surface unreachable

when contrastive supervision is absent. Supplementary Fig. S3 shows the same collapse per joint: without MCED the angles pile against the kinematic limits; with MCED they track the teacher's broad distribution.

Gated PG-CTD. On top of PGD, the gate (row d) adds +6.0 Suc.6 over static PGD (row b), measuring the value of adaptive, teacher-relative correction over uniform. Fig. 4 shows its self-adaptation: before activation (Epoch 10), 82% of predictions are already worse than the teacher—direct evidence that few-step compression strips the teacher's iterative correction. As gating engages, the degradation ratio falls from 82% to 21% at convergence, with median $\Delta\Phi$ moving from +1.25 to −0.40; the typical prediction now beats the teacher, and the gate withdraws on its own, no threshold tuning.

Static vs. gated physics guidance. Gating helps diversity too, beyond the +6.0 Suc.6. Comparing static absolute (row b, Suc.6 72.5, Pen. 12.5) against gated relative (row d, Suc.6 78.5, Pen. 12.1) on the same object as Fig. 2: at comparable physical quality, the static penalty cuts diversity (pairwise-rotation median 51°) while gating keeps it high (65°). Both exclude MCED (rows b, d), so absolute diversity is lower than the full model in Fig. 2; this isolates the gate's effect. Concentrating correction on degraded samples thus improves feasibility without flattening structure.

Synergy. The full model (row f) reaches 86.1 Suc.6, +18.0 over CTM. To test whether this is genuine synergy rather than addition, we remove each module from the full setup. Removing PG-CTD (row e) drops Suc.6 to 77.4 (−8.7); removing MCED (row d) drops it to 78.5 (−7.6). Each removal costs more than the module gives in isolation (+7.0 for MCED, row c; +6.0 for gated PG-CTD, row d vs. b), so the two are not independently additive but mutually dependent—each is far more effective with the other present. The leave-one-out pattern is consistent with the division of labor hypothesized in the PG-CTD subsection of Methods, though we note it is measured on one dataset.

Real-robot evaluation

Table 3. Real-robot grasping performance and inference efficiency over six household objects (10 attempts each).

Method	Bottle	Mug	Box	Irregular Toy	Cosmetics	Can	Overall SR↑	Time↓
Teacher	6/10	6/10	7/10	5/10	6/10	7/10	61.7%	17.30 s
Ours (2-step)	6/10	5/10	6/10	5/10	4/10	7/10	55.0%	**0.25 s**

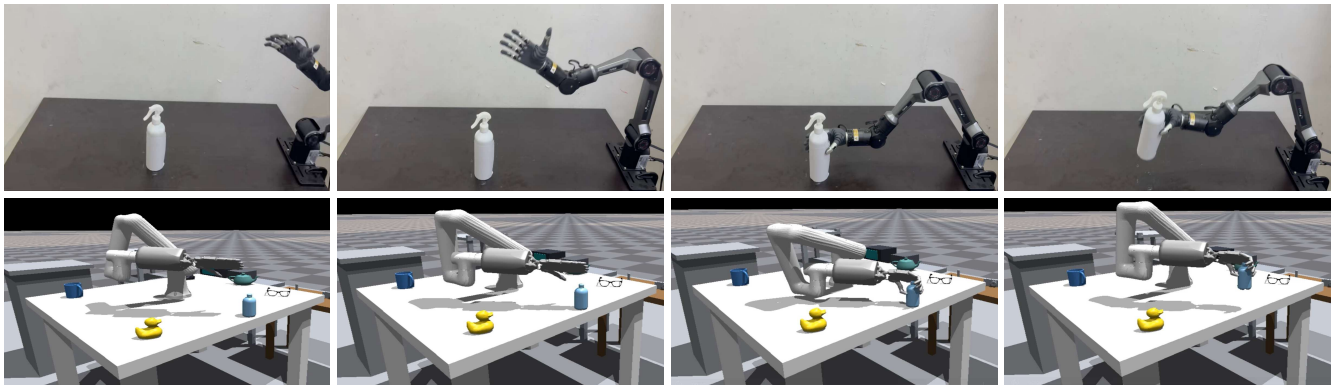


Figure 5. Deployment of the few-step student. **Top:** Grasps executed on a physical setup consisting of an AgileX Piper 6-DoF arm and a BrainCo Revo2 (basic) dexterous hand; the ShadowHand configuration predicted by the student is retargeted to the Revo2 kinematics before execution. **Bottom:** Grasps on objects of the same categories evaluated in IsaacGym.

On the physical platform (see Methods), the student reaches 55.0% overall success over 60 trials (Wilson 95% CI $[42.5, 66.9]$%); the teacher reaches 61.7% ($[49.0, 72.9]$%). The intervals overlap almost entirely, so at this sample size the two are statistically indistinguishable; the experiment establishes deployability, not a ranking. Per-batch inference time drops from 17.30 s to 0.25 s, matching the same-GPU measurement reported with Table 1, since the robot generates a single batch of N_g grasps for one object under the same setting. Fig. 5 shows a representative sequence, alongside grasps on objects of the same categories in simulation.

Discussion

EfficientDexGrasp solves a specific problem: compressing a multi-step diffusion grasping policy into a few-step one while keeping its multimodal structure and physical validity. Its gains should be read in this scope. We discuss limitations along three

axes—the distillation paradigm, the physics prior, and the evaluation setting—and the direction each implies.

Bounded by the teacher policy. Since MCED rebuilds the teacher's energy landscape and PG-CTD applies only a teacher-relative correction, the student inherits the teacher's mode coverage and biases (DexGrasp Anything[7]). By design it *preserves* modes the teacher already found rather than *recovering* ones it missed, and never generates grasps outside the teacher's support. Systematic gaps in the teacher—say, under-represented grasp types for thin or concave objects—thus propagate to the student, so the attainable Suc.6 is effectively capped by teacher quality. Mode preservation also depends on the coverage of MCED's three-source negative pool (student outputs, perturbed teacher samples, uniform samples): if it under-covers the multimodal structure, some mode collapse can persist. Coupling distillation with a lightweight exploratory or reward-based fine-tuning stage that reaches beyond the teacher's support is a natural way to lift this ceiling; reinforcement-learning fine-tuning of diffusion policies[45,46] has shown that such a stage can move a policy past the support of its demonstrations.

Reliance on hand-crafted physics priors. The physics potential Φ in PG-CTD uses three predefined geometric terms calibrated for rigid bodies: external repulsion (ERF, penetration), surface pulling (SPF, contact deficiency), and self-repulsion (SRF, self-collision). This has two consequences. First, all three are defined on static surface geometry, so Φ may misjudge deformable, articulated, or thin-shell objects, whose contact dynamics these terms do not capture. Second, the *same* Φ drives both the training-time gate $m_i = \sigma(\Delta\Phi_i/T_g)$ and the inference-time selection of the lowest-Φ candidate, so a mis-specified Φ compounds at both ends—mis-gating valid-but-unusual grasps in training and selecting them away at deployment. The method also assumes reasonably accurate geometry, since Φ relies on mesh-based distance queries; its robustness to partial, noisy, or occluded point clouds is not yet characterized. Replacing the hand-crafted terms with learned contact models or differentiable physics, and conditioning the gate on perceptual uncertainty, are promising next steps.

Limited scene complexity and closed-loop evaluation. Our experiments primarily isolate the grasp-synthesis stage. Evaluation focuses on single-object grasping and open-loop pose prediction. We also validate EfficientDexGrasp on a physical dexterous hand. However, the real-world experiments cover limited scene diversity and interaction complexity. Three gaps remain relative to full deployment. (i) *Limited scene-level reasoning*: cluttered environments introduce collision avoidance, reachability, and accessibility constraints. The current evaluation does not comprehensively cover these constraints. (ii) *No closed-loop reactivity*: the student predicts a static target. It does not act as a feedback controller, and the latency reductions we report are measured per batch of grasp proposals, not for a control loop. Object motion, contact slip, and disturbances during approach are therefore not explicitly handled; the reduced latency removes one obstacle to closed-loop use but does not by itself provide it. (iii) *End-to-end pipeline integration*: the 11×–69× speedups concern the grasping policy itself. Practical system latency also depends on perception, motion planning, and control. Our real-robot results provide initial evidence of sim-to-real transfer under sensing and actuation noise. Broader validation remains necessary. This should cover cluttered and dynamic environments, as well as diverse physical setups. Future work should extend EfficientDexGrasp to these settings. It should also integrate perception and closed-loop control to support robust deployment.

We presented **EfficientDexGrasp**, a physics-guided distillation framework that compresses multi-step diffusion grasping policies into efficient few-step ones. It rests on two complementary mechanisms. Mode-preserving Contrastive Energy Distillation (MCED) rebuilds the teacher's energy landscape by contrastive learning, keeping distinct grasp modes from collapsing into one averaged solution. Gated Physics-Guided Consistency Trajectory Distillation (PG-CTD) scales correction with the degree of physical degradation through a teacher-relative constraint, leaving the structure MCED builds intact. On DexGraspNet, MultiDex, RealDex, and DexGRAB, EfficientDexGrasp delivers 11×–69× speedups and beats the strongest baselines by 3.8–5.8 absolute Suc.6 points, and leave-one-out ablations show that each module contributes more in the presence of the other than alone.

Methods

EfficientDexGrasp distills a multi-step diffusion grasping policy into a few-step student that keeps the teacher's quality, diversity, and physical plausibility. It has three components. A consistency trajectory distillation backbone compresses the teacher's iterative denoising into a one- or few-pass mapping. On top of it, Mode-preserving Contrastive Energy Distillation (MCED) rebuilds the teacher's energy landscape in action space, keeping the multimodal structure from collapsing to a single averaged solution. Gated Physics-Guided Consistency Trajectory Distillation (PG-CTD) then enforces feasibility on predictions that drift off the feasible manifold, without flattening the diversity MCED preserves. The full objective and the inference procedure follow.

Problem formulation

We address conditional dexterous grasp generation. Given a point cloud observation $o \in \mathbb{R}^{N\times 3}$ of an object, the goal is a grasp pose $x \in \mathbb{R}^d$ specifying the full hand configuration for stable contact. The pose has three parts: joint angles $q \in \mathbb{R}^{24}$ for finger articulation, a global translation $\mathbf{t} \in \mathbb{R}^3$, and a global rotation $R \in \mathrm{SO}(3)$ for approach orientation. Together they span a high-dimensional space in which valid grasps occupy a small, geometrically complex region.

We use DexGrasp Anything[7] as the teacher, which models $p(x \mid o)$ as a diffusion process. From Gaussian noise $x_T \sim \mathcal{N}(0, I)$, it refines the sample over T steps via the learned reverse transition:

$$p_\psi(x_{t-1} \mid x_t, o) = \mathcal{N}\left(x_{t-1}; \mu_\psi(x_t, t, o), \sigma_t^2 I\right), \tag{1}$$

where ψ denotes the frozen teacher parameters and μ_ψ predicts the denoised mean. The teacher embeds differentiable physics in the sampling loop, nudging each step away from penetration and self-collision, so its grasps are high-quality and feasible. But this needs $T \gg 1$ evaluations, giving latencies of tens of seconds—fine for offline planning, too slow for interactive use. Our goal is a student f_θ that approximates the full trajectory in one or a few passes, $\hat{x}_0 = f_\theta(x_t, t, o)$, cutting inference time by an order of magnitude while keeping grasp quality.

Consistency trajectory distillation

Our acceleration backbone is consistency trajectory distillation[30], which exploits a property of the probability-flow ODE behind the teacher[47]: any two points on the same trajectory map to the same clean sample x_0 once the trajectory is followed to the end. A function f_θ with this self-consistency can jump from any noise level straight to the output, skipping intermediate steps. Formally,

$$f_\theta(x_{\tau_n}, \tau_n, o) = f_\theta(x_{\tau_m}, \tau_m, o), \quad \forall\, \tau_n, \tau_m \in [0, T], \tag{2}$$

with boundary condition $f_\theta(x_0, 0, o) = x_0$, so f_θ is the identity at zero noise. For stability we use a skip-connection parameterization:

$$f_\theta(x_{\tau_n}, \tau_n, o) = c_{\text{skip}}(\tau_n)\, x_{\tau_n} + c_{\text{out}}(\tau_n)\, F_\theta(x_{\tau_n}, \tau_n, o), \tag{3}$$

where F_θ is the learned network and $c_{\text{skip}}, c_{\text{out}}$ are differentiable schedules with $c_{\text{skip}}(0) = 1$, $c_{\text{out}}(0) = 0$. This enforces the boundary condition by construction: at $\tau = 0$ the skip path dominates and the output is x_0; at higher noise F_θ takes over.

Training enforces consistency between outputs at adjacent points on the teacher's ODE trajectory. Given a noisy x_{τ_n} and the teacher's one-step solution $\hat{x}^*_{\tau_{n-1}}$ (advancing $\tau_n \to \tau_{n-1}$), the loss penalizes their discrepancy:

$$\mathcal{L}_{\text{CD}} = \mathbb{E}_{x_0, \tau_n}\left[\left\| f_\theta(x_{\tau_n}, \tau_n, o) - f_{\theta'}(\hat{x}^*_{\tau_{n-1}}, \tau_{n-1}, o) \right\|_2^2\right], \tag{4}$$

where θ' is an exponential moving average of θ—a momentum reference, borrowed from self-supervised learning, that prevents collapse. Minimizing this across all intervals lets the student internalize the whole trajectory, producing in one pass what the teacher needs many steps for.

Applied directly to grasping, however, consistency distillation has two failure modes the original framework was not built for. The first is *multimodal collapse*. The loss acts on individual trajectory-sample pairs, so it cannot preserve the global structure of the teacher's output distribution. When one object has several valid modes—a mug grasped by the handle, rim, or body—the student may settle on an average between them that is no valid grasp. The second is *physical infeasibility*. The teacher reaches the feasible manifold gradually, one physics correction per step; compressing this to a single step removes that mechanism, and some predictions violate contact or collision constraints. The next two sections address each.

Mode-preserving contrastive energy distillation

Consistency distillation supervises only at discrete points along the teacher's trajectories. It teaches the student to map noisy inputs to clean outputs, but leaves the arrangement of grasp modes and the barriers between them unsupervised, and a regression loss on a multimodal target settles at the centroid of the modes[41]. MCED adds a continuous, structure-aware energy signal that shapes the student's output landscape.

Energy function and contrastive objective

MCED reframes distillation from point-wise regression to conditional energy learning. We learn an energy function $E_\phi(o, a)$, an auxiliary head with parameters ϕ separate from f_θ, scoring each observation–action pair; low energy marks high-quality grasps. The head is a training-time signal only and is discarded at inference. Only the relative ordering of energies matters, which lets MCED keep one low-energy basin per valid mode without forcing equal energies across them. We train the energy function with contrastive supervision. For each observation o, positives a^+ come from the teacher's grasps. Negatives a^- are drawn uniformly from a three-source pool covering different regions of action space: the student's current predictions (possibly collapsed), perturbed teacher outputs (probing the neighborhood of valid modes), and uniform samples (covering infeasible

Algorithm 1 MCED Loss

Input: Observation batch $\{o^{(i)}\}_{i=1}^{B}$; teacher π_{teacher}; student f_θ; energy head E_ϕ; negatives per sample M; temperature κ; perturbation scale σ_{neg}

Output: MCED loss $\mathscr{L}_{\text{MCED}}$ (updates ϕ only)

Sample positive: $a_i^+ \sim \pi_{\text{teacher}}(o^{(i)})$

Sample M negatives $\{a_{i,j}^-\}_{j=1}^{M}$ uniformly from:

$\quad \text{sg}\big[f_\theta(x_t, t, o^{(i)})\big] \cup \big\{a_i^+ + \mathscr{N}(0, \sigma_{\text{neg}}^2 I)\big\} \cup \mathscr{U}(\mathscr{A})$

Score energies: $e_i^+ = E_\phi(o^{(i)}, a_i^+),\ e_{i,j}^- = E_\phi(o^{(i)}, a_{i,j}^-)$

Let $p_i^+ = \exp(-e_i^+/\kappa),\ p_{i,j}^- = \exp(-e_{i,j}^-/\kappa)$

InfoNCE objective:

$\quad \mathscr{L}_{\text{MCED}} \leftarrow -\dfrac{1}{B}\displaystyle\sum_{i=1}^{B} \log \dfrac{p_i^+}{p_i^+ + \sum_{j=1}^{M} p_{i,j}^-}$

return $\mathscr{L}_{\text{MCED}}$

regions). The pool gives informative gradients across the whole space, not just near the teacher's outputs. We optimize an InfoNCE loss:

$$\mathscr{L}_{\text{MCED}} = -\mathbb{E}\left[\log \frac{e^{-E_\phi(o,a^+)/\kappa}}{e^{-E_\phi(o,a^+)/\kappa} + \sum_{j=1}^{M} e^{-E_\phi(o,a_j^-)/\kappa}}\right], \tag{5}$$

over observations o, positives a^+, and negative sets $\{a_j^-\}_{j=1}^{M}$, with temperature $\kappa > 0$ controlling contrast sharpness. Throughout, a denotes a grasp treated as an action scored by the energy head and x the same quantity as the state of the diffusion process; the two coincide, $a \equiv x_0$. The softmax turns energies into a distribution over candidates and drives the positive to the lowest energy. Versus margin losses like triplet loss, InfoNCE has two practical edges: its gradients decay smoothly rather than truncating, avoiding the dead zone that stalls late training; and the softmax bounds gradient magnitude, avoiding the unstable updates of unbounded margins. The result is one low-energy basin per valid mode, separated by high-energy barriers that keep modes from merging[35,38]. Student predictions are detached before they enter the negative pool, so $\mathscr{L}_{\text{MCED}}$ updates ϕ alone and never pushes the student itself toward higher energy.

Alternating optimization

The energy head and the student are trained with two separate objectives that we optimize in alternation. The head minimizes $\mathscr{L}_{\text{MCED}}$, which raises the energy of everything in the negative pool, and student predictions are detached here so only ϕ moves. The student minimizes the energy of its own output,

$$\mathscr{L}_{\text{ENG}} = \mathbb{E}\Big[\text{sg}_\phi\big[E_\phi\big](o, f_\theta(x_t, t, o))\Big], \tag{6}$$

where $\text{sg}_\phi[\cdot]$ detaches the head's parameters, so only θ moves. The two objectives are adversarial: the head raises the energy of the student's current outputs while the student lowers it, a min–max game of the same form as energy-based GANs[43]. Three ingredients keep it stable in our setting. The head is trained against teacher positives, which are fixed, so its landscape is anchored to the teacher rather than to the moving student; the student's outputs form only one third of the negative pool, so the head cannot fit the student alone; and the student objective is dominated by $\mathscr{L}_{\text{CD}}$, with $\mathscr{L}_{\text{ENG}}$ acting as a weighted regularizer (λ_{e}, Methods). Under these conditions the head converges to a landscape whose low-energy regions are the teacher's modes, and the student, driven by $\mathscr{L}_{\text{CD}}$ toward the teacher's trajectories and by $\mathscr{L}_{\text{ENG}}$ away from high-energy regions, settles in those modes. Nothing in the objective forces the student into a single basin: $\mathscr{L}_{\text{CD}}$ ties each noise sample to the trajectory it started on, and different trajectories end in different basins. We did not observe divergence or oscillation in any run.

Mode preservation mechanism

Unlike a regression loss that assumes a unimodal target, the MCED landscape is pushed to keep multiple local minima, one per viable strategy. InfoNCE allows this because it does not require the positive to be the unique global minimum—only to be lower than all negatives in the batch. When different batches present positives from different modes, the energy function learns to score all of them low, carving several valleys instead of one. The intended effect is that positive energies decrease and negative energies increase over training, so that the landscape forms low-energy regions for valid modes and higher-energy regions between them; the diversity results in the ablation study provide the quantitative evidence. MCED thus lifts distillation from sample-level imitation to distribution-level transfer, where the student inherits the teacher's full grasp repertoire.

Algorithm 2 Gated PG-CTD Loss

Input: Student prediction $\hat{\mathbf{x}}_0 = f_\theta(\mathbf{x}_t, t, \mathbf{o})$; teacher target $\mathbf{a}_{\text{teacher}}$; physics weights $\boldsymbol{\gamma} = (\gamma_{\text{erf}}, \gamma_{\text{spf}}, \gamma_{\text{srf}})$; gate temperature T_g
Output: Gated relative-physics loss $\mathscr{L}_{\text{PG}}$

Compute physics potentials:
  $\boldsymbol{\Phi}_s \leftarrow \Phi(\hat{\mathbf{x}}_0, \mathbf{o}; \boldsymbol{\gamma})$
  $\boldsymbol{\Phi}_t \leftarrow \text{sg}\big[\Phi(\mathbf{a}_{\text{teacher}}, \mathbf{o}; \boldsymbol{\gamma})\big]$
Degradation: $\Delta\boldsymbol{\Phi} \leftarrow \boldsymbol{\Phi}_s - \boldsymbol{\Phi}_t$
Gating mask: $\mathbf{m} \leftarrow \sigma(\Delta\boldsymbol{\Phi}/T_g)$
Per-sample penalty: $\boldsymbol{\ell} \leftarrow \text{softplus}(\Delta\boldsymbol{\Phi})$
$\mathscr{L}_{\text{PG}} \leftarrow \big(\mathbf{m}^\top \boldsymbol{\ell}\big)/\big(\mathbf{1}^\top \mathbf{m} + \varepsilon\big)$
return $\mathscr{L}_{\text{PG}}$

MCED preserves the teacher's multimodal structure but does not guarantee each mode stays feasible: the few-step student lacks the teacher's per-step physics correction, so even when its predictions match the teacher on average, a non-negligible fraction overshoot the feasible boundary. An absolute penalty $\Phi(\hat{x}_0, o)$ would fix this at the cost of MCED's diversity, since it treats all violations alike and pulls every prediction toward low-potential configurations. PG-CTD instead applies a *teacher-relative* constraint whose strength scales with how much worse the student is than the teacher.

Physics potential function

We quantify physical quality with a differentiable infeasibility potential, following DexGrasp Anything[7]:

$$\Phi(x, o) = \gamma_{\text{erf}}\,\text{ERF}(x, o) + \gamma_{\text{spf}}\,\text{SPF}(x, o) + \gamma_{\text{srf}}\,\text{SRF}(x, o), \tag{7}$$

each term capturing one aspect of infeasibility. The external repulsion force ERF measures hand–object penetration. The surface pulling force SPF measures contact deficiency, rising when fingers are too far for stable contact. The self-repulsion force SRF penalizes self-collision, such as overlapping fingers. The weights $\gamma_{\text{erf}}, \gamma_{\text{spf}}, \gamma_{\text{srf}}$ set their relative importance. Higher Φ means less feasible, and all three terms are differentiable in x for gradient-based training.

Degradation-based gating mechanism

The core principle of Gated PG-CTD: scale physics correction with the degree of degradation, rather than applying it uniformly. We define a per-sample degradation comparing the two potentials:

$$\Delta\Phi_i = \Phi(\hat{x}_0^{(i)}, o) - \Phi(a_{\text{teacher}}^{(i)}, o), \tag{8}$$

with $\hat{x}_0^{(i)} = f_\theta(x_t^{(i)}, t, o)$ the student's prediction and $a_{\text{teacher}}^{(i)}$ the teacher's. Positive $\Delta\Phi_i$ means the student is worse—more penetration, weaker contact, more self-collision; zero or negative means it matches or beats the teacher. A gate turns this into a soft mask:

$$m_i = \sigma\left(\frac{\Delta\Phi_i}{T_g}\right), \tag{9}$$

with $\sigma(\cdot)$ the sigmoid and T_g a temperature setting transition sharpness. The gate is computed from a detached copy of $\Delta\Phi_i$, $m_i = \sigma(\text{sg}[\Delta\Phi_i]/T_g)$, and the normalizer $\sum_i m_i$ in Eq. (10) is likewise detached, so the gate acts purely as a per-sample weight and contributes no gradient of its own. For severe degradation ($\Delta\Phi_i \gg 0$) the gate saturates near one and physics guidance fully activates. As the student approaches the teacher ($\Delta\Phi_i \to 0$) the gate falls to 0.5, and once the student is clearly better ($\Delta\Phi_i \ll 0$) it falls toward zero, leaving the prediction close to where MCED placed it.

This degradation-based gate self-adapts across training, unlike a fixed absolute-Φ threshold. Early on, the student is far from the teacher and most predictions are degraded, so the gate activates broadly and corrects fast. As the student converges, fewer samples are degraded, and the gate deactivates for more of them automatically—no manual threshold or schedule.

Relative physics constraint loss

With the mask in place, the physics loss is a gated, teacher-relative penalty on the degradation:

$$\mathscr{L}_{\text{PG}} = \frac{\sum_{i=1}^{B} m_i \cdot \text{softplus}(\Delta\Phi_i)}{\sum_{i=1}^{B} m_i + \varepsilon}, \tag{10}$$

where softplus$(z) = \ln(1+e^z)$ smoothly approximates the rectifier and ε is for stability. Softplus is monotonically increasing, so $\mathscr{L}_{\mathrm{PG}}$ is not minimized at $\Delta\Phi_i = 0$; at equality it equals $\ln 2$ and it continues to decrease as the student improves beyond the teacher. What makes the penalty effectively one-sided is the product of the two factors. When the student is worse ($\Delta\Phi_i > 0$), softplus acts like the identity and the gate is near one, so the sample is penalized roughly linearly with degradation. When the student matches or beats the teacher ($\Delta\Phi_i \le 0$), the softplus gradient $\sigma(\Delta\Phi_i)$ and the gate m_i both shrink, so the effective gradient on such samples is the product of two small factors and vanishes as $\Delta\Phi_i$ decreases. The gate thus concentrates the gradient budget on the worst samples rather than spreading it across the batch. Algorithm 2 summarizes one training iteration.

Interaction with consistency distillation

The relative constraint is designed to coexist with the consistency loss $\mathscr{L}_{\mathrm{CD}}$ rather than to be exactly aligned with it. Consistency asks the student to match the teacher's trajectory; the relative physics loss asks it to be physically no worse than the teacher, and continues to reward improvement beyond the teacher's level. The two can therefore disagree on individual samples. What limits the disagreement is the gate: the residual pressure below $\Delta\Phi_i = 0$ is scaled by $m_i \le 0.5$ and by a vanishing softplus gradient, whereas the pressure on degraded samples is at full strength. Physics guidance is thus strongest where the consistency approximation is imperfect—where few-step compression error shows up as measurable degradation—and weak where the student already matches the teacher. An absolute penalty $\Phi(\hat{x}_0)$, by contrast, targets $\Phi = 0$ regardless of the teacher's level, and applies full pressure to every sample, including those that are already as feasible as the teacher. The ablation study quantifies the difference.

Complementarity with MCED

MCED and Gated PG-CTD are designed to handle complementary failure modes. MCED supervises the energy geometry of the action space, carving low-energy basins for distinct modes. PG-CTD penalizes predictions that are physically worse than the teacher. Because the gate scales correction down as samples approach the teacher's quality, PG-CTD is intended to put little directional pressure on predictions that already sit in a valid mode, and to concentrate on those that have degraded. We treat this division of labor as a design hypothesis; the ablation study tests it with leave-one-out ablations and Fig. 4 tracks which samples the gate selects over training.

Training objective and inference

Combined training loss

Training alternates between two objectives. The student minimizes

$$\mathscr{L}_\theta = \mathscr{L}_{\mathrm{CD}} + \lambda_{\mathrm{e}}\mathscr{L}_{\mathrm{ENG}} + \lambda_{\mathrm{pg}}(e)\mathscr{L}_{\mathrm{PG}}, \tag{11}$$

and the energy head minimizes

$$\mathscr{L}_\phi = \mathscr{L}_{\mathrm{MCED}}, \tag{12}$$

with the detachments described in the MCED subsection of Methods: the head is frozen in $\mathscr{L}_{\mathrm{ENG}}$ and the student is frozen in $\mathscr{L}_{\mathrm{MCED}}$. Here λ_{e} is fixed and $\lambda_{\mathrm{pg}}(e)$ follows a three-phase warmup over epochs:

$$\lambda_{\mathrm{pg}}(e) = \begin{cases} 0, & e < e_{\mathrm{start}}, \\ \lambda_{\mathrm{pg}}^{\max}\dfrac{e - e_{\mathrm{start}}}{E_{\mathrm{warm}}}, & e_{\mathrm{start}} \le e < e_{\mathrm{start}} + E_{\mathrm{warm}}, \\ \lambda_{\mathrm{pg}}^{\max}, & e \ge e_{\mathrm{start}} + E_{\mathrm{warm}}. \end{cases} \tag{13}$$

The student must reach a basic level of consistency and energy structure before physics correction helps. Early on ($e < e_{\mathrm{start}}$), the student is far from the teacher and $\Delta\Phi$ is mostly noise, not systematic violation; activating physics now risks collapsing the student onto one low-penetration mode while MCED is still building diversity. The linear ramp introduces correction gradually, without disrupting the energy landscape already formed. After warmup, physics runs at full strength on the residual degraded predictions.

Inference

At inference, training infrastructure is stripped away: the teacher, the energy head E_ϕ, the gate, and the PG-CTD loss are all discarded, leaving only the student f_θ. Neither the teacher nor $\Delta\Phi$ is needed at deployment. Given an observation o, the student generates a batch of grasps from independent noise samples in a few passes—two steps for speed, up to eight for quality.

Two protocols use this batch. For the benchmark evaluation in Results, *every* generated grasp is evaluated and the reported metrics are averages over the full batch, for our method and for all baselines alike; no candidate selection is involved. For execution on a physical robot, where a single grasp must be chosen, we evaluate Φ on the batch in one batched pass (no gradients) and execute the lowest-Φ candidate. Since Φ needs only mesh-based distance queries, this adds negligible overhead. The full pipeline, from point cloud to executed grasp, runs at a fraction of the teacher's cost.

Datasets, evaluation protocol and baselines

Datasets. We evaluate on four benchmarks. DexGraspNet[3] gives optimization-based annotations over thousands of household objects; MultiDex[4] spans multiple hand morphologies and object categories; RealDex[5] captures human-like grasps from real demonstrations; and DexGRAB[6] provides whole-body human grasping with fine-grained functional, power, and precision contact annotations.

Evaluation protocol. For every test object, each method generates $N_g = 32$ grasps from independent noise samples, and *all* of them are evaluated. Reported metrics are averages over all generated grasps and all test objects; no candidate selection or filtering is applied to any method. The test split contains 1126 objects for DexGraspNet, 10 for MultiDex, 6 for RealDex, and 15 for DexGRAB, disjoint from the training objects.

Metrics. Following SceneDiffuser[11], we report four. A grasp counts as successful when the object's displacement under an applied external force stays below 2 cm. **Suc.6** (our primary metric) requires this in all six axis-aligned directions ($\pm X$, $\pm Y$, $\pm Z$) in IsaacGym[48]; **Suc.1** requires it in at least one direction; **Pen.** (mm) is the maximum hand–object penetration depth of a grasp, averaged over grasps and objects; the definitions of Suc.1, Suc.6 and Pen. match those used by the works from which baseline numbers are taken. **Time** is the wall-clock latency for generating one batch of N_g grasps. Qualitative examples of the six-direction stability test and of a full pick-and-place rollout in IsaacGym are given in Supplementary Figs. S1 and S2.

Baselines. We compare against six methods: UniDexGrasp[23], GraspTTA[25], SceneDiffuser[11], UGG[12], DexGrasp Anything[7] (also our teacher), and EvolvingGrasp[13]. EvolvingGrasp is reported at 2 and 8 steps for budget-matched comparison.

Implementation details. The student shares the teacher's U-Net architecture and is initialized from its weights; the energy head E_ϕ is a three-layer MLP ($539 \to 256 \to 128 \to 1$, ReLU, with a tanh output). All models are trained in PyTorch with AdamW (weight decay 10^{-6}, linear warmup followed by cosine decay) at a learning rate of 1×10^{-5} for 300 epochs with batch size 128 on a single NVIDIA GeForce RTX 5090, following the official train–test split of each dataset. MCED uses temperature $\kappa = 0.15$, $M = 8$ negatives per sample, and weight $\lambda_e = 1.0$; PG-CTD uses $T_g = 0.2$ with $\lambda_{pg}^{max} = 0.1$, and the physics weights γ follow DexGrasp Anything[7]. Latency is measured on the same GPU; baseline runtimes in Table 1 are cited from the original papers and were obtained on different hardware.

Real-robot setup

The physical platform pairs an AgileX Piper 6-DoF arm with a BrainCo Revo2 (basic) dexterous hand, and we deploy the 2-step student on it. The student predicts a ShadowHand configuration, which is retargeted to the Revo2 kinematics before execution by matching fingertip positions: the fingertip positions of the predicted ShadowHand pose are computed by forward kinematics and the Revo2 joint angles are solved by inverse kinematics to reproduce them in a least-squares sense; the wrist pose is applied directly. For each trial the method generates a batch of N_g grasps for the observed point cloud and the lowest-Φ candidate is executed. A trial counts as a success if the object is lifted 10 cm above the table and held for 5 s without slipping. We evaluate six household objects (bottle, mug, box, irregular toy, cosmetics container, and can) with 10 grasp attempts each, re-placing the object at a randomized pose before every attempt; the teacher runs under the same protocol.

Code availability

The code implementing EfficientDexGrasp is available from the corresponding author upon reasonable request.

Data availability

The datasets analysed in this study are publicly available: DexGraspNet[3], MultiDex[4], RealDex[5], and DexGRAB[6].

References

1. Ferrari, C. & Canny, J. Planning optimal grasps. In *Proceedings 1992 IEEE International Conference on Robotics and Automation*, 2290–2295 vol.3 (1992). DOI 10.1109/ROBOT.1992.219918.

2. Bicchi, A. & Kumar, V. Robotic grasping and contact: A review. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, vol. 1, 348–353 vol.1 (2000). DOI 10.1109/ROBOT.2000.844081.

3. Wang, R. *et al.* DexGraspNet: A large-scale robotic dexterous grasp dataset for general objects based on simulation. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 11359–11366 (2023). DOI 10.1109/ICRA48891.2023.10160982.

4. Li, P. *et al.* GenDexGrasp: Generalizable dexterous grasping. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 8068–8074 (2023). DOI 10.1109/ICRA48891.2023.10160667.

5. Liu, Y. *et al.* RealDex: Towards human-like grasping for robotic dexterous hand. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*, 6859–6867 (2024). DOI 10.24963/ijcai.2024/758.

6. Taheri, O., Ghorbani, N., Black, M. J. & Tzionas, D. GRAB: A dataset of whole-body human grasping of objects. In Vedaldi, A., Bischof, H., Brox, T. & Frahm, J.-M. (eds.) *Computer Vision – ECCV 2020*, 581–600 (Springer International Publishing, Cham, 2020). DOI 10.1007/978-3-030-58548-8_34.

7. Zhong, Y., Jiang, Q., Yu, J. & Ma, Y. DexGrasp anything: Towards universal robotic dexterous grasping with physics awareness. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 22584–22594 (2025). DOI 10.1109/CVPR52734.2025.02103.

8. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. & Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning*, 2256–2265 (2015).

9. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, vol. 33, 6840–6851 (2020).

10. Chi, C. *et al.* Diffusion policy: Visuomotor policy learning via action diffusion. *The Int. J. Robotics Res.* **44**, 1684–1704 (2025). DOI 10.1177/02783649241273668.

11. Huang, S. *et al.* Diffusion-based generation, optimization, and planning in 3D scenes. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 16750–16761 (2023). DOI 10.1109/CVPR52729.2023.01607.

12. Lu, J. *et al.* UGG: Unified generative grasping. In Leonardis, A. *et al.* (eds.) *Computer Vision – ECCV 2024*, 414–433 (Springer Nature Switzerland, Cham, 2025). DOI 10.1007/978-3-031-72855-6_24.

13. Zhu, Y. *et al.* Evolvinggrasp: Evolutionary grasp generation via efficient preference alignment. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, 11665–11674 (2025). DOI 10.1109/ICCV51701.2025.01085.

14. Salimans, T. & Ho, J. Progressive distillation for fast sampling of diffusion models. In *Proceedings of the International Conference on Learning Representations* (2022).

15. Song, Y., Dhariwal, P., Chen, M. & Sutskever, I. Consistency models. In *Proceedings of the 40th International Conference on Machine Learning*, vol. 202 of *ICML'23*, 32211–32252 (JMLR.org, Honolulu, Hawaii, USA, 2023).

16. Fang, H.-S., Wang, C., Gou, M. & Lu, C. GraspNet-1Billion: A large-scale benchmark for general object grasping. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11441–11450 (2020). DOI 10.1109/CVPR42600.2020.01146.

17. Liu, M., Pan, Z., Xu, K., Ganguly, K. & Manocha, D. Generating grasp poses for a High-DOF gripper using neural networks. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1518–1525 (2019). DOI 10.1109/IROS40897.2019.8968115.

18. Mousavian, A., Eppner, C. & Fox, D. 6-DOF GraspNet: Variational grasp generation for object manipulation. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2901–2910 (2019). DOI 10.1109/ICCV.2019.00299.

19. Lundell, J., Verdoja, F. & Kyrki, V. DDGC: Generative deep dexterous grasping in clutter. *IEEE Robotics Autom. Lett.* **6**, 6899–6906 (2021). DOI 10.1109/LRA.2021.3096239.

20. Wei, W. *et al.* DVGG: Deep variational grasp generation for dextrous manipulation. *IEEE Robotics Autom. Lett.* **7**, 1659–1666 (2022). DOI 10.1109/LRA.2022.3140424.

21. Kingma, D. P. & Welling, M. Auto-encoding variational bayes. In *Proceedings of the International Conference on Learning Representations* (2014).

22. Goodfellow, I. J. *et al.* Generative adversarial nets. In *Advances in Neural Information Processing Systems*, vol. 27 (2014).

23. Xu, Y. *et al.* UniDexGrasp: Universal robotic dexterous grasping via learning diverse proposal generation and goal-conditioned policy. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4737–4746 (2023). DOI 10.1109/CVPR52729.2023.00459.

24. Wan, W. *et al.* UniDexGrasp++: Improving dexterous grasping policy learning via geometry-aware curriculum and iterative generalist-specialist learning. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 3868–3879 (2023). DOI 10.1109/ICCV51070.2023.00360.

25. Jiang, H., Liu, S., Wang, J. & Wang, X. Hand-object contact consistency reasoning for human grasps generation. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 11087–11096 (2021). DOI 10.1109/ICCV48922.2021.01092.

26. Weng, Z., Lu, H., Kragic, D. & Lundell, J. DexDiffuser: Generating dexterous grasps with diffusion models. *IEEE Robotics Autom. Lett.* **9**, 11834–11840 (2024). DOI 10.1109/LRA.2024.3498776.

27. Xu, G.-H., Wei, Y.-L., Zheng, D., Wu, X.-M. & Zheng, W.-S. Dexterous grasp transformer. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 17933–17942 (2024). DOI 10.1109/CVPR52733.2024.01698.

28. Song, J., Meng, C. & Ermon, S. Denoising diffusion implicit models. In *Proceedings of the International Conference on Learning Representations* (2021).

29. Lu, C. *et al.* DPM-Solver: A fast ODE solver for diffusion probabilistic model sampling in around 10 steps. In *Advances in Neural Information Processing Systems*, vol. 35, 5775–5787 (2022). DOI 10.52202/068431-0418.

30. Kim, D. *et al.* Consistency trajectory models: Learning probability flow ODE trajectory of diffusion. In *Proceedings of the International Conference on Learning Representations* (2024).

31. Janner, M., Du, Y., Tenenbaum, J. & Levine, S. Planning with diffusion for flexible behavior synthesis. In *Proceedings of the 39th International Conference on Machine Learning*, 9902–9915 (2022).

32. Prasad, A., Lin, K., Wu, J., Zhou, L. & Bohg, J. Consistency policy: Accelerated visuomotor policies via consistency distillation. In *Proceedings of Robotics: Science and Systems* (2024).

33. Lu, G. *et al.* ManiCM: Real-time 3D diffusion policy via consistency model for robotic manipulation. *arXiv preprint arXiv:2406.01586* (2024).

34. Wang, Z. *et al.* One-step diffusion policy: Fast visuomotor policies via diffusion distillation. *arXiv preprint arXiv:2410.21257* (2024).

35. LeCun, Y., Chopra, S., Hadsell, R., Ranzato, M. & Huang, F. J. A tutorial on energy-based learning. *Predict. Struct. Data* **1**, 1–59 (2006).

36. Du, Y. & Mordatch, I. Implicit generation and modeling with energy based models. In *Advances in Neural Information Processing Systems*, vol. 32 (2019).

37. Florence, P. *et al.* Implicit behavioral cloning. In *Proceedings of the 5th Conference on Robot Learning*, 158–168 (2022).

38. Gutmann, M. & Hyvärinen, A. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 297–304 (2010).

39. He, K., Fan, H., Wu, Y., Xie, S. & Girshick, R. Momentum contrast for unsupervised visual representation learning. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9726–9735 (2020). DOI 10.1109/CVPR42600.2020.00975.

40. Chen, T., Kornblith, S., Norouzi, M. & Hinton, G. A simple framework for contrastive learning of visual representations. In *Proceedings of the 37th International Conference on Machine Learning*, 1597–1607 (2020).

41. Hinton, G., Vinyals, O. & Dean, J. Distilling the knowledge in a neural network. In *NIPS Deep Learning and Representation Learning Workshop* (2015).

42. Romero, A. *et al.* FitNets: Hints for thin deep nets. In *Proceedings of the International Conference on Learning Representations* (2015).

43. Zhao, J., Mathieu, M. & LeCun, Y. Energy-based generative adversarial network. In *Proceedings of the International Conference on Learning Representations* (2017).

44. Zhu, H. *et al.* Simple distillation for one-step diffusion models. In *Advances in Neural Information Processing Systems*, vol. 38, 92071–92094 (2025). DOI 10.52202/085713-3080.

45. Black, K., Janner, M., Du, Y., Kostrikov, I. & Levine, S. Training diffusion models with reinforcement learning. In *International Conference on Learning Representations (ICLR)* (2024).

46. Ren, A. Z. *et al.* Diffusion policy policy optimization. In *International Conference on Learning Representations (ICLR)* (2025).

47. Song, Y. *et al.* Score-based generative modeling through stochastic differential equations. In *Proceedings of the International Conference on Learning Representations* (2021).

48. Makoviychuk, V. *et al.* Isaac Gym: High performance GPU based physics simulation for robot learning. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks* (2021).

Acknowledgements

This work was supported in part by the Foundation of State Key Laboratory for Fine Exploration and Intelligent Development of Coal Resources (Grant No. SKLCRSM24KFA08), in part by the Open Fund of State Key Laboratory of Digital Intelligent Technology for Unmanned Coal Mining (Grant No. SZQZ2025-2-006), in part by the Basic Research Program of Taicang (Grant No. TC2024JC30), in part by the National Key Laboratory Foundation of Helicopter Aeromechanics (Grant No. 2024-ZSJ-LB-02-03), and in part by the National Natural Science Foundation of China (Grant No. 62303379).

Author contributions statement

H.H. developed the method, implemented the code, conducted the simulation experiments and wrote the manuscript. G.C. designed and conducted the real-robot experiments. W.Z. supervised the project, reviewed the methodology and revised the manuscript. All authors reviewed the manuscript.

Additional information

Competing interests: The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- EfficientDexGraspSciRepsupplementary.pdf