InterPrior: Scaling Generative Control for Physics-Based Human-Object Interactions

Supplementary Material

In this supplementary, we provide additional details of our InterPrior framework with extended experiments:

- **(i)** Sec. A describes the organization of the demo video.
- **(ii)** Sec. B details the overall simulation configuration.
- **(iii)** Sec. C provides additional information on our goal representation, *e.g.*, how snapshot, trajectory, and contact goals are constructed at training and evaluation time with the masks.
- **(iv)** Sec. D gives a comprehensive explanation on: (**I**) the detailed formulation of the reference-free hand reward; (**II**) the losses used for variational distillation and latent shaping, and (**III**) RL finetuning.
- **(v)** Sec. E specifies additional implementation details, including network architectures, training schedules, and how we apply data augmentation to expert training, as well as additional techniques we use during G1 training for sim-to-sim experiments.
- **(vi)** Sec. F presents further qualitative results, *e.g.*, the integration of InterPrior with kinematic HOI generators, additional details of metrics, and failure cases.
- **(vii)** Sec. G examines the limitations of our current system and its potential societal implications.

Contents

A. Demo Video

The demo video on the webpage visualizes behaviors produced by InterPrior across settings detailed in the following. All sequences are rendered from the physics simulator [42, 63] using the same SMPL [33, 52] and G1 [65] model as for training. No post-processing is applied other than camera selection and cropping for visualization.

Core Capability. We show examples of snapshot, trajectory, and contact-conditioned control corresponding to the scenarios illustrated in Figure 1 of the main paper, for objects with diverse shapes.

Failure Recovery and Regrasping. We visualize rollouts perturbed or initialized from failure states. The video highlights re-approaching, re-grasping, and recovery from falls as described in Sec. 3.4.

Long-Horizon Multi-Goal Chains. We include long sequences where three canonicalized sub-goals are chained (Sec. 4, "Chain" tasks) and the policy must transition smoothly between different interaction while maintaining task success.

Diverse Task Execution from the Same Goal. We show that our model is able to control the simulated human achieving the same task with different execution.

Baseline Comparison. We demonstrate that InterPrior achieves superior performance compared to existing baseline methods [59, 60, 88].

Novel Interaction Generalization. We visualize qualitative results on BEHAVE [3] and HODome [96], as a complementary to Figure 5 and Figure 7 in the main paper.

Interaction with multiple objects. We showcase that InterPrior supports human interactions with multiple objects, without requiring any task-specific training.

Sim-to-Sim for G1. We include more examples of the G1 humanoid with sim-to-sim transfer, as a complementary to Figure 6, for controlling a humanoid only based on object future snapshot goal.

Sim-to-Real for G1. Our work, ULTRA [17], extends InterPrior with additional perception modules and enables perception-in-the-loop loco-manipulation on the Unitree G1. More details are available on the webpage.

Interactive Steering Control. Finally, we show real-time keyboard control where a user steers high-level goals and InterPrior produces coherent whole-body motion online.

B. Simulation

All experiments are performed in IsaacGym [42] with the GPU PhysX backend. Control policies run at 30Hz, while the simulator is stepped at 60Hz with two internal substeps per control step. The main simulation hyperparameters are summarized in Table A.

Table A. Simulation hyperparameters used in IsaacGym [42]. We largely follow the settings from prior work [72, 88].

Hyperparameter	Value
Simulation step Δt	$1/60$ s
Control step Δt	$1/30$ s
Physics substeps per control step	2
Position solver iterations	4
Velocity solver iterations	1
Contact offset	0.02
Rest offset	0.0
Max depenetration velocity	100
Object & ground restitution	0.7
Object & ground friction	0.9
Object density	200
Max convex hulls per object	64
Object rest offset	0.01

We introduce a small object rest offset to reduce human-object interpenetration, especially for thin geometries. Although this slightly enlarges the effective collision boundary, it avoids the substantial cost associated with increasing solver accuracy to compensate for collision handling.

C. Goal Formulation

This section details the construction of snapshot, trajectory, and contact goals and the associated masks used. Specifically, a goal state $\boldsymbol{y}_t$ shares the same structure as the observation $\boldsymbol{x}_t$, and a binary mask $\boldsymbol{m}_t$ indicates which components of $\boldsymbol{y}_t$ are provided to the policy.

C.1. Horizon for Goals

Short-Horizon Preview. We use a small set of offsets $K = \{1, 2, 4, 16\}$ to provide short-horizon previews relative to the current timestep t. For each offset $k \in K$, we construct a goal pair $(\boldsymbol{y}_{t+k}, \boldsymbol{m}_{t+k})$.

Long-Horizon Snapshot. A long-horizon offset sampled by $L \in [1, 128]$ defines a single far-future goal $(\boldsymbol{y}_{t+L}, \boldsymbol{m}_{t+L})$. During training, L is initialized randomly at the start of each episode and then decremented each timestep, being resampled once it reaches zero. Although termed a long-horizon snapshot, its value naturally decreases at each step and may temporarily fall below the short-horizon offsets.

C.2. Stochastic Mask Sampling during Training

During training, masks are not tied to specific tasks (snapshot; trajectory; contact). Instead, we randomly decide which parts of the future state are revealed to the policy, so that the policy is exposed to a *wide variety* of partial and sparse goals, following [59]. We operate at the level of rigid bodies, including objects with following three rules:

Body-Wise Masking. Visibility is enforced at the body level. For each rigid body, we maintain a single binary variable. If it is *false*, all all state features associated with that body at time $t+k$ are masked out, positions, orientations, and linear and angular velocities. The same rule applies to the entries in the interaction vectors D_{t+k} and the contact state C_{t+k}, defined in Sect. 3.1, which are masked or revealed together.

Independent Sampling in Rigid Bodies. At each horizon offset k, each body is sampled independently according to a fixed Bernoulli distribution: human-state and interaction components are revealed with probability 0.1, and object components with probability 0.5. This procedure produces diverse, randomly constructed combinations of visible and masked human, object, and contact features, rather than relying on any task-specific mask templates.

Temporal Consistency of Masks. To avoid flickering visibility, masks evolve over time with a high probability of staying the same and a small probability of being resampled. Concretely, for $k > 1$ we define a first-order Markov process:

$$\boldsymbol{m}_{t+k} = \begin{cases} \boldsymbol{m}_{t+k-1}, & \text{with probability } 1 - p_{\text{reset}}, \\ \text{Bernoulli}(\boldsymbol{p}_{\text{vis}}), & \text{with probability } p_{\text{reset}}. \end{cases}$$

Here $p_{\text{reset}} = 0.01$ ensures that once a body is masked or unmasked, it tends to remain in that state for multiple steps, while occasional resets still diversify the masks. The visibility probabilities $\boldsymbol{p}_{\text{vis}}$ follow the design above.

C.3. Task Definition for Inference

During inference, masks are constructed according to the target task. For a given task, the visibility pattern remains fixed throughout the rollout. The only exception is the multi-goal chaining setting, where we resample a new mask whenever the controller transitions to the next sub-goal.

Snapshot-Conditioned Control. We unmask the long-horizon snapshot. We still apply the consistent per-body sampling to determine which body or object components are revealed. All short-horizon preview are fully masked.

Trajectory-Conditioned Control. We unmask the short-horizon preview. Following the same per-body sampling, we reveal only a subset of the joint or object components. The long-horizon snapshot goal is retained.

Contact-Conditioned Control. Contact goals are implemented as a special case of snapshot conditioning in which we reveal only contact-related information. Specifically, we unmask the contact entries of $\boldsymbol{C}_t$, the associated signed-distance fields $\boldsymbol{D}_t$ (defined in Sec. 3.1), and the relevant human body parts. To avoid ambiguity in the target, we additionally unmask the object pose in the snapshot frame.

Multi-Goal Chaining. For multi-goal chains, we extract data by concatenating different data sequences. Specifically, we canonicalize each subsequent first frame with respect to the previous last frame. Canonicalization is performed by aligning the human root position (excluding height), and heading, *i.e.*, rotation around the vertical z–axis only, rather than the full $SO(3)$ orientation. Because this transformation is applied with respect to the human frame only, the object frame may become partially misaligned after canonicalization. As a result, we do not expect the policy to perfectly satisfy all chained goals, especially when object-relative alignment becomes extremely inconsistent. Nevertheless, the presence of a long horizon makes the policy possibly compensate for canonicalization artifacts.

D. Additional Details on Methodology

This section expands the reward and loss formulations, as well as additional details for the three stages of our framework: (**I**) InterMimic+ expert training (extending Sec. 3.2), (**II**) variational distillation (extending Sec. 3.3), and (**III**) RL post-training (extending Sec. 3.4).

D.1. InterMimic+: Full-Reference Imitation Expert

Reference-Free Reward for Expert. Here we introduce the detailed formulation of the hand reward r_{h}. Let $\boldsymbol{p}_T$ denote the position of the thumb fingertip and $\{\boldsymbol{p}_j\}_{j\in S}$ the positions of the other fingertips, with $\boldsymbol{q}_T$ and $\{\boldsymbol{q}_j\}_{j\in S}$ being their respective nearest surface points on the object. We define unit bearing vectors from the object surface toward the fingertips as $\boldsymbol{u}_T = (\boldsymbol{p}_T - \boldsymbol{q}_T)/\|\boldsymbol{p}_T - \boldsymbol{q}_T\|$ and $\boldsymbol{u}_j = (\boldsymbol{p}_j - \boldsymbol{q}_j)/\|\boldsymbol{p}_j - \boldsymbol{q}_j\|$, $j \in S$. The reward is defined as $r_{\text{h}} = \exp(-w_{\text{h}} e_{\text{h}})$, where $e_{\text{h}} = 1 - \frac{1}{|S|}\sum_{j\in S}\frac{1-\boldsymbol{u}_T^\top \boldsymbol{u}_j}{2}$, and w_{h} increases as the hand-object distance decreases, activating only when the reference indicates an upcoming interaction. This reward encourages all five fingers to maximize upcoming surface contact with the object.

D.2. InterPrior: Variational Distillation

Here we introduce the formulation for our proposed losses for variational Distillation. Let $\boldsymbol{\mu}_{p,t}$ and $\boldsymbol{\Sigma}_{p,t}$ denote the prior's mean and covariance at time t, *i.e.*, $\mathcal{N}(\boldsymbol{\mu}_{p,t}, \boldsymbol{\Sigma}_{p,t}) \equiv p_\psi(\boldsymbol{z}_t \mid \boldsymbol{x}_{t-\ell:t}, \mathcal{G}_t)$.

(**I**) *Scale loss.* We regularize the prior mean to lie on the unit hypersphere. This is to prevent the output mean from collapsing or exploding, with the use of latent normalization:

$$\mathcal{L}_{\text{scale}} = \mathbb{E}_t\big[\big(\|\boldsymbol{\mu}_{p,t}\|_2 - 1\big)^2\big].$$

(**II**) *Temporal consistency loss.* To obtain a smooth latent prior over time, we use $\mathcal{L}_{\text{tc}}$ to penalize changes in the prior distribution across consecutive timesteps using the squared 2-Wasserstein distance between Gaussians.

(**III**) *Goal reconstruction loss.* The decoder includes an additional head that predicts future goal features conditioned on the latent. Let $\widehat{\boldsymbol{y}}_{t+k}$ denote the predicted goal at offset k and $\boldsymbol{m}_{t+k}$ the input mask used to construct the masked residual goal. We train this head to complete the *masked* entries of the goal, *i.e.*, those that were hidden from the policy input. Formally, the goal reconstruction loss is

$$\mathcal{L}_{\text{goal}} = \mathbb{E}_{t,k}\big[\big\|\big(\mathbf{1} - \boldsymbol{m}_{t+k}\big) \odot \big(\widehat{\boldsymbol{y}}_{t+k} - \boldsymbol{y}_{t+k}\big)\big\|_2^2\big],$$

where $\odot$ denotes element-wise multiplication and $\mathbf{1}$ is an all-ones vector. This loss encourages the latent $\boldsymbol{z}_t$ to capture intent and context sufficient to reconstruct the missing parts of the goal, given only the visible subset provided by the mask. In practice, we reconstruct short future with $k = 1$.

D.3. InterPrior: Post-Training Beyond Reference

Get-Up Training. To learn the get-up behavior, in addition to the new learnable token as discussed in Sec. 3.4, we introduce an auxiliary reward that becomes active, with episodes initialized from a fallen state. The reward encourages both elevation of the pelvis and reorientation of the torso toward an upright configuration:

$$r^{\text{getup}} = w_{\text{height}}\,\sigma\big(h_t - h_{\text{target}}\big) \;+\; w_{\text{upright}}\,\sigma\big(\mathbf{n}_t \cdot \mathbf{n}_{\text{up}}\big), \quad (1)$$

where h_t is the pelvis height, h_{target} is set as 0.7, $\mathbf{n}_t$ is the torso's up vector, $\mathbf{n}_{\text{up}}$ is the world up direction, and $\sigma(\cdot)$ denotes a clipped linear shaping function.

Distributed Training. To mitigate catastrophic forgetting, we divide the parallel simulation environments into three groups: (**I**) *RL environments*, optimized solely with the post-training reward r_t^{PT}; (**II**) *Distillation environments*, optimized using the ELBO objective and supervised by the expert policy, as described in Sec. 3.3. The policy parameters are shared across all environments. Gradients are aggregated synchronously to update the shared policy.

Mask Prompt Engineering during Inference. To further enhance robustness during inference without additional learning, we apply lightweight *mask-based prompting* over the goal specification $\mathcal{G}_t$ (Sec. 3.1): (**I**) When following a trajectory and the state lags behind, we remove the trajectory goal but redefine the nearest waypoint as the snapshot goal. (**II**) For snapshot goals with distant target joints ($>$1 m), we retain only the root translation goal while masking out all other components, prompting locomotion before fine manipulation. (**III**) When human-object targets are contradictory, *e.g.*, both are moving but no grasp is established, we set the human root goal to the current object position while maintaining root height, masking all other joints. This encourages natural re-approach and regrasping behaviors. These inference-time edits operate solely on the goal $\mathcal{G}_t$, while the policy parameters remain fixed.

Finetuning on Additional HOI Datasets. The same finetuning mechanism naturally extends to absorbing new interaction datasets. Given any additional HOI corpus (*e.g.*, BEHAVE [3] or HODome [96] in Sec. 4), states from such new dataset are treated as additional sources of long-horizon goals and initializations for RL rollouts, while the distillation group continues to regularize the policy toward the original prior. This allows InterPrior to incrementally acquire new object categories and interaction styles without retraining from scratch.

E. Implementation Details

This section summarizes key implementation details, including network configurations, hyperparameters, randomization settings used for expert training, and additional techniques used during G1 training for sim-to-sim experiments.

PPO Setup. For both the expert and RL finetuning stages, we use PPO with generalized advantage estimation (GAE) and a clipped surrogate objective, and train with Adam. Following common practice [47], we keep the PPO discount factor γ, GAE parameter λ, clip ratio, and entropy regularization as shown in Table B, and apply gradient clipping.

InterMimic+: Full-Reference Imitation Expert. The InterMimic+ expert policy and critic are MLPs with three hidden layers of sizes $(1024, 1024, 512)$, using ReLU activations. Actor and critic are parameterized separately, and the critic outputs a scalar value with full observation and reference as input. Please refer to [88] for more details.

InterPrior: Variational Distillation. The encoder and decoder used for variational distillation share the same MLP backbone with hidden sizes $(1024, 1024, 512)$. The prior p_ψ is implemented as a 4-layer Transformer encoder with 4 attention heads, a latent dimension of 512, and a feedforward width of 1024. For the distillation objective (Sec. D), we use unit weight for the action reconstruction loss, and assign a weight of 10^{-3} to all auxiliary terms (goal reconstruction, scale loss, and temporal consistency loss). The KL regularizer follows a β-VAE style schedule: the KL weight β is annealed from 10^{-3} to 1.0 over the course of training. We first perform 500 epochs of warm-up using only teacher-controlled rollouts, and then gradually increase the fraction of student-controlled rollouts [53] until epoch $10,000$, at which point 95% of environments are driven by the student policy while the remaining 5% always use the teacher for fresh expert trajectories.

InterPrior: Post-Training Beyond Reference. For the post-training stage, we retain the same loss weights used for the distillation branch, and combine with the PPO loss weights specified in Table B for the RL branches.

Inference Efficiency. The runtime breakdown is: observation 20.16,ms, physics 19.02,ms, policy inference 0.43,ms, *SDF* 0.134,ms, and other overheads 0.057,ms, highlighting the policy's potential for real-world deployment.

Table B. Hyperparamters for training teacher and student policies.

Hyperparameters	value
Discount factor γ	0.99
Generalized advantage estimation λ	0.95
Learning rate	2e-5
Action loss weight	1
Critic loss weight	5
Action bounds loss weight	10
Minibatch size	16384
Horizon length H	32
Maximum episode length	300

Table C. **Additional reward terms for G1** used in Stage I expert training. Here, $\boldsymbol{\tau}$ denotes the vector of joint torques with elementwise limits $[\boldsymbol{\tau}_{\min}, \boldsymbol{\tau}_{\max}]$; $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ are joint degrees and velocities with limits $[\boldsymbol{q}_{\min}, \boldsymbol{q}_{\max}]$; $\boldsymbol{a}_t$ is the control action at time t; $\boldsymbol{\omega}$ and $\boldsymbol{v}$ are the base (root) angular and linear velocities; F_z^{feet} is the vertical ground-reaction force at the feet; $\boldsymbol{v}^{\text{feet}}$ is the tangential (ground-plane) velocity of the feet; d_{feet} is the horizontal distance between the two feet, with desired bounds $[d_{\min}, d_{\max}]$; $\boldsymbol{g}_{xy}^{\text{feet}}$ is the projection of the gravity direction onto the foot frame's ground plane; $\mathbb{1}(\cdot)$ and $\mathbb{1}_{\text{termination}}$ are indicator functions. All norms $\|\cdot\|$ and $\|\cdot\|_2$ are Euclidean.

TERM	EXPRESSION	WEIGHT
Penalty:		
Torque limits	$\mathbb{1}(\boldsymbol{\tau} \notin [\boldsymbol{\tau}_{\min}, \boldsymbol{\tau}_{\max}])$	2
DoF position limits	$\mathbb{1}(\boldsymbol{q} \notin [\boldsymbol{q}_{\min}, \boldsymbol{q}_{\max}])$	5
Energy	$\|\boldsymbol{\tau} \odot \dot{\boldsymbol{q}}\|$	10^{-4}
Termination	$\mathbb{1}_{\text{termination}}$	-30
Regularization:		
DoF velocity	$\|\dot{\boldsymbol{q}}\|_2^2$	4×10^{-4}
Action rate	$\|\boldsymbol{a}_t\|_2^2$	0.1
Torque	$\|\boldsymbol{\tau}\|$	2×10^{-3}
Angular velocity	$\|\boldsymbol{\omega}\|^2$	0.01
Base velocity	$\|\boldsymbol{v}\|^2$	0.1
Foot slip	$\mathbb{1}(F_z^{\text{feet}} > 5.0) \cdot \sqrt{\|\boldsymbol{v}^{\text{feet}}\|}$	0.03
Feet distance reward	$\frac{1}{2}\exp\left(-100\,\lvert\max(d_{\text{feet}} - d_{\min}, -0.5)\rvert\right) + \frac{1}{2}\exp\left(-100\,\lvert\max(d_{\text{feet}} - d_{\max}, 0)\rvert\right)$	0.5
Feet orientation	$\sqrt{\|\boldsymbol{g}_{xy}^{\text{feet}}\|}$	1

F. Additional Experimental Results

In this section, we introduce metric details, provide supplementary qualitative results, and discuss failure cases.

Additional Details on Evaluation Metrics. For *trajectory-following* tasks, we evaluate the policy at each timestep by comparing the rollout state with the corresponding reference, and compute pose and object errors only over the unmasked components. For *snapshot goal-following* tasks, there is no time-aligned reference trajectory. Instead, we compute the error between the rollout state and the snapshot goal at every timestep and report the *minimum* of this

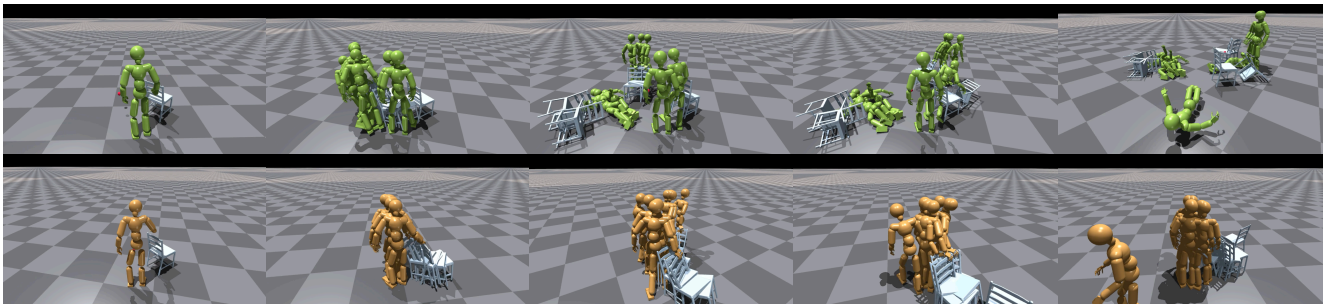

Figure A. **Additional qualitative comparisons** with baseline method [59, 60] (**Top**). Our InterPrior shows higher success rate under the same task goal.

Table D. **Range of dynamics randomization.** "default" refers to the parameter value from the unitree G1 official 29DoF model. v_{xy} is the planar (horizontal) push velocity.

Term	Range / Value
Dynamics randomization	
Friction coefficient	$\mathcal{U}(1.0,\ 3.0)$
Base CoM offset	$\mathcal{U}(-0.05,\ 0.05)$ m
Base mass offset	$\mathcal{U}(-3.0,\ 3.0)$ kg
P gain scaling	$\mathcal{U}(0.8,\ 1.2)\ \times$ default
D gain scaling	$\mathcal{U}(0.8,\ 1.2)\ \times$ default
External perturbation	
Push robot	interval $= 4$ s, $\quad v_{xy} = 1$ m/s

distance over the rollout. This reflects whether the policy is capable of reaching the target configuration.

Diverse Behaviors Under the Same Goal. Beyond the examples shown in the main paper, Figure B illustrates how InterPrior behaves diversely given the same goal, showing that our learned latent space is meaningful and is able to capture diverse behaviors.

Integration with Kinematic HOI Generators. To demonstrate that InterPrior's generalization, we integrate it with InterDiff [84] that produces physically unconstrained interaction trajectories. The integration proceeds as follows: (**I**) the kinematic generator produces a 25 frames of human-object poses given the past 15 frames following [84]; (**II**) we convert these sequences into our goal representation by extracting snapshot and trajectory goals; and (**III**) we feed these goals into InterPrior. The result is shown in Figure C.

G. Discussion

Limitations and Future Work. InterPrior is still bounded by the coverage and quality of its training data: highly corrupted or unseen interaction patterns are not reliably recovered, and in such cases the policy often defaults to conservative strategies, maintaining balance without fully solving the task. Our model is tailored to rigid object, and we still observe occasional artifacts such as shallow interpenetrations, foot skating, or failure cases such as object drop over long rollouts. The current hand and contact representation is also not designed for fine-grained finger dexterity or in-hand manipulation. Finally, our three-stage training introduces additional complexity and hyperparameters. Future work includes expanding dataset diversity, incorporating richer hand models, and simplifying or unifying the training scheme.

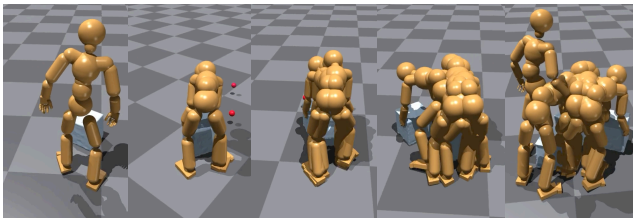

Figure B. **Qualitative results** given the same goal. Our framework produces multiple valid yet distinct interaction trajectories.

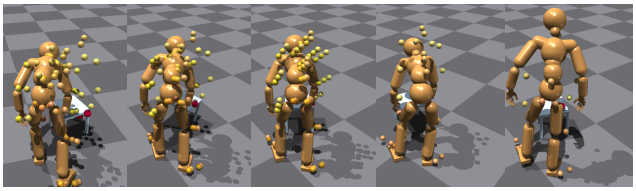

Figure C. **Qualitative results** of InterPrior following the targets generated by InterDiff [84] (*yellow* and *red* dots). InterPrior adaptively completes the task without strictly adhering to the targets, using only sparse inputs of wrist, feet, and object target.

Societal and Ethical Considerations. InterPrior enables more general-purpose, physically grounded humanoid controller, which can be beneficial for animation, simulation, and robotics, but also raises potential risks. More capable humanoid controllers could be deployed in unsafe settings or for applications that conflict with societal norms (*e.g.*, surveillance or coercive scenarios). We therefore encourage careful consideration of safety mechanisms, usage policies, and ethical guidelines when applying this type of model beyond controlled research environments.

References

[1] Jinseok Bae, Jungdam Won, Donggeun Lim, Cheol-Hui Min, and Young Min Kim. Pmp: Learning to physically interact with environments using part-wise motion priors. In *SIGGRAPH*, 2023. 3

[2] Donghoon Baek, Amartya Purushottam, Jason J Choi, and Joao Ramos. Whole-body bilateral teleoperation with multi-stage object parameter estimation for wheeled humanoid locomanipulation. *arXiv preprint arXiv:2508.09846*, 2025. 2

[3] Bharat Lal Bhatnagar, Xianghui Xie, Ilya Petrov, Cristian Sminchisescu, Christian Theobalt, and Gerard Pons-Moll. BEHAVE: Dataset and method for tracking human object interactions. In *CVPR*, 2022. 2, 6, 7, 8, 1, 4

[4] Yu-Wei Chao, Jimei Yang, Weifeng Chen, and Jia Deng. Learning to sit: Synthesizing human-chair interactions via hierarchical control. In *AAAI*, 2021. 3

[5] Peishan Cong, Ziyi Wang, Yuexin Ma, and Xiangyu Yue. Semgeomo: Dynamic contextual human motion generation with semantic and geometric guidance. In *CVPR*, 2025. 2

[6] Jieming Cui, Tengyu Liu, Nian Liu, Yaodong Yang, Yixin Zhu, and Siyuan Huang. AnySkill: Learning open-vocabulary physical skill for interactive agents. In *CVPR*, 2024. 3

[7] Zekai Deng, Ye Shi, Kaiyang Ji, Lan Xu, Shaoli Huang, and Jingya Wang. Human-object interaction via automatically designed vlm-guided motion policy. *arXiv preprint arXiv:2503.18349*, 2025. 3

[8] Christian Diller and Angela Dai. CG-HOI: Contact-guided 3d human-object interaction generation. In *CVPR*, 2024. 2

[9] Zhiyang Dou, Xuelin Chen, Qingnan Fan, Taku Komura, and Wenping Wang. C· ase: Learning conditional adversarial skill embeddings for physics-based characters. In *SIGGRAPH Asia*, 2023. 2

[10] Yuhui Fu, Feiyang Xie, Chaoyi Xu, Jing Xiong, Haoqi Yuan, and Zongqing Lu. DemoHLM: From one demonstration to generalizable humanoid loco-manipulation. *arXiv preprint arXiv:2510.11258*, 2025. 2

[11] Levi Fussell, Kevin Bergamin, and Daniel Holden. Supertrack: Motion tracking for physically simulated characters using supervised learning. *ACM Transactions on Graphics (TOG)*, 40(6):1–13, 2021. 3

[12] Jiawei Gao, Ziqin Wang, Zeqi Xiao, Jingbo Wang, Tai Wang, Jinkun Cao, Xiaolin Hu, Si Liu, Jifeng Dai, and Jiangmiao Pang. CooHOI: Learning cooperative human-object interaction with manipulated object dynamics. In *NeurIPS*, 2024. 3

[13] Zichen Geng, Zeeshan Hayder, Wei Liu, and Ajmal Saeed Mian. Auto-regressive diffusion for generating 3d human-object interactions. In *AAAI*, 2025. 2

[14] Anindita Ghosh, Rishabh Dabral, Vladislav Golyanik, Christian Theobalt, and Philipp Slusallek. IMoS: Intent-driven full-body motion synthesis for human-object interactions. In *Computer Graphics Forum*, 2023. 2

[15] Mohamed Hassan, Yunrong Guo, Tingwu Wang, Michael Black, Sanja Fidler, and Xue Bin Peng. Synthesizing physical character-scene interactions. In *SIGGRAPH*, 2023. 2, 3

[16] Wenkun He, Yun Liu, Ruitao Liu, and Li Yi. Syncdiff: Synchronized motion diffusion for multi-body human-object interaction synthesis. In *ICCV*, 2025. 2

[17] Xialin He, Sirui Xu, Xinyao Li, Runpei Dong, Liuyu Bian, Yu-Xiong Wang, and Liang-Yan Gui. ULTRA: Unified multimodal control for autonomous humanoid whole-body loco-manipulation. *arXiv preprint arXiv:2603.03279*, 2026. 8, 1

[18] Xiaoyu Huang, Takara Truong, Yunbo Zhang, Fangzhou Yu, Jean Pierre Sleiman, Jessica Hodgins, Koushil Sreenath, and Farbod Farshidian. Diffuse-cloc: Guided diffusion for physics-based character look-ahead control. *ACM Transactions on Graphics (TOG)*, 44(4):1–12, 2025. 3

[19] Yinghao Huang, Omid Taheri, Michael J. Black, and Dimitrios Tzionas. InterCap: Joint markerless 3D tracking of humans and objects in interaction. In *GCPR*, 2022. 2

[20] Kai Jia, Tengyu Liu, Mingtao Pei, Yixin Zhu, and Siyuan Huang. PrimHOI: Compositional human-object interaction via reusable primitives. In *ICCV*, 2025. 2

[21] Nan Jiang, Tengyu Liu, Zhexuan Cao, Jieming Cui, Yixin Chen, He Wang, Yixin Zhu, and Siyuan Huang. CHAIRS: Towards full-body articulated human-object interaction. In *ICCV*, 2023. 2

[22] Nan Jiang, Zimo He, Zi Wang, Hongjie Li, Yixin Chen, Siyuan Huang, and Yixin Zhu. Autonomous character-scene interaction synthesis from text instruction. In *SIGGRAPH Asia*, 2024. 2

[23] Nan Jiang, Zhiyuan Zhang, Hongjie Li, Xiaoxuan Ma, Zan Wang, Yixin Chen, Tengyu Liu, Yixin Zhu, and Siyuan Huang. Scaling up dynamic human-scene interaction modeling. In *CVPR*, 2024. 2

[24] Jordan Juravsky, Yunrong Guo, Sanja Fidler, and Xue Bin Peng. SuperPADL: Scaling language-directed physics-based control with progressive supervised distillation. In *SIGGRAPH*, 2024. 2

[25] Dvij Kalaria, Sudarshan S Harithas, Pushkal Katara, Sangkyung Kwak, Sarthak Bhagat, Shankar Sastry, Srinath Sridhar, Sai Vemprala, Ashish Kapoor, and Jonathan Chung-Kuan Huang. DreamControl: Human-inspired whole-body humanoid control for scene interaction via guided diffusion. *arXiv preprint arXiv:2509.14353*, 2025. 2

[26] Hyeonwoo Kim, Sangwon Beak, and Hanbyul Joo. DAViD: Modeling dynamic affordance of 3d objects using pre-trained video diffusion models. *arXiv preprint arXiv:2501.08333*, 2025. 2

[27] Jeonghwan Kim, Jisoo Kim, Jeonghyeon Na, and Hanbyul Joo. ParaHome: Parameterizing everyday home activities towards 3d generative modeling of human-object interactions. *arXiv preprint arXiv:2401.10232*, 2024. 2

[28] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 3, 5

[29] Jiaman Li, Jiajun Wu, and C Karen Liu. Object motion guided human motion synthesis. *ACM Transactions on Graphics (TOG)*, 42(6):1–11, 2023. 2, 6, 7, 8

[30] Yitang Li, Mingxian Lin, Zhuo Lin, Yipeng Deng, Yue Cao, and Li Yi. Learning physics-based full-body human reaching and grasping from brief walking references. In *CVPR*, 2025. 3

[31] Libin Liu and Jessica Hodgins. Learning to schedule control fragments for physics-based characters using deep q-learning. *ACM Transactions on Graphics (TOG)*, 36(3): 1–14, 2017. 3

[32] Yun Liu, Chengwen Zhang, Ruofan Xing, Bingda Tang, Bowen Yang, and Li Yi. Core4d: A 4d human-object-human interaction dataset for collaborative object rearrangement. In *CVPR*, 2025. 2

[33] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. SMPL: A skinned multi-person linear model. *ACM transactions on graphics*, 2015. 4, 1

[34] Jintao Lu, He Zhang, Yuting Ye, Takaaki Shiratori, Sebastian Starke, and Taku Komura. CHOICE: Coordinated human-object interaction in cluttered environments for pick-and-place actions. *arXiv preprint arXiv:2412.06702*, 2024. 2

[35] Jiaxin Lu, Chun-Hao Paul Huang, Uttaran Bhattacharya, Qixing Huang, and Yi Zhou. HUMOTO: A 4d dataset of mocap human object interactions. In *ICCV*, 2025. 2

[36] Zhengyi Luo, Jinkun Cao, Kris Kitani, Weipeng Xu, et al. Perpetual humanoid control for real-time simulated avatars. In *ICCV*, 2023. 3

[37] Zhengyi Luo, Jinkun Cao, Josh Merel, Alexander Winkler, Jing Huang, Kris Kitani, and Weipeng Xu. Universal humanoid motion representations for physics-based control. *arXiv preprint arXiv:2310.04582*, 2023. 3, 5

[38] Zhengyi Luo, Jinkun Cao, Sammy Christen, Alexander Winkler, Kris Kitani, and Weipeng Xu. Grasping diverse objects with simulated humanoids. In *NeurIPS*, 2024. 2, 3, 5

[39] Zhengyi Luo, Jiashun Wang, Kangni Liu, Haotian Zhang, Chen Tessler, Jingbo Wang, Ye Yuan, Jinkun Cao, Zihui Lin, Fengyi Wang, et al. SMPLOlympics: Sports environments for physically simulated humanoids. *arXiv preprint arXiv:2407.00187*, 2024. 3

[40] Zhengyi Luo, Chen Tessler, Toru Lin, Ye Yuan, Tairan He, Wenli Xiao, Yunrong Guo, Gal Chechik, Kris Kitani, Linxi Fan, et al. Emergent active perception and dexterity of simulated humanoids from visual reinforcement learning. *arXiv preprint arXiv:2505.12278*, 2025. 3

[41] Xintao Lv, Liang Xu, Yichao Yan, Xin Jin, Congsheng Xu, Shuwen Wu, Yifan Liu, Lincheng Li, Mengxiao Bi, Wenjun Zeng, et al. HIMO: A new benchmark for full-body human interacting with multiple objects. In *ECCV*, 2024. 2

[42] Viktor Makoviychuk, Lukasz Wawrzyniak, Yunrong Guo, Michelle Lu, Kier Storey, Miles Macklin, David Hoeller, Nikita Rudin, Arthur Allshire, Ankur Handa, et al. Isaac gym: High performance gpu-based physics simulation for robot learning. In *NeurIPS*, 2021. 6, 7, 1, 2

[43] Josh Merel, Saran Tunyasuvunakool, Arun Ahuja, Yuval Tassa, Leonard Hasenclever, Vu Pham, Tom Erez, Greg Wayne, and Nicolas Heess. Catch & carry: reusable neural controllers for vision-guided whole-body tasks. *ACM Transactions on Graphics (TOG)*, 39(4):39–1, 2020. 3

[44] Liang Pan, Jingbo Wang, Buzhen Huang, Junyu Zhang, Haofan Wang, Xu Tang, and Yangang Wang. Synthesizing physically plausible human motions in 3d scenes. In *3DV*, 2024. 3

[45] Liang Pan, Zeshi Yang, Zhiyang Dou, Wenjia Wang, Buzhen Huang, Bo Dai, Taku Komura, and Jingbo Wang. TokenHSI: Unified synthesis of physical human-scene interactions through task tokenization. In *CVPR*, 2025. 2, 3, 6

[46] Xiaogang Peng, Yiming Xie, Zizhao Wu, Varun Jampani, Deqing Sun, and Huaizu Jiang. HOI-Diff: Text-driven synthesis of 3d human-object interactions using diffusion models. *arXiv preprint arXiv:2312.06553*, 2023. 2

[47] Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel Van de Panne. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions On Graphics (TOG)*, 37(4):1–14, 2018. 2, 4

[48] Xue Bin Peng, Ze Ma, Pieter Abbeel, Sergey Levine, and Angjoo Kanazawa. Amp: Adversarial motion priors for stylized physics-based character control. *ACM Transactions on Graphics (ToG)*, 40(4):1–20, 2021. 2

[49] Xue Bin Peng, Yunrong Guo, Lina Halper, Sergey Levine, and Sanja Fidler. Ase: Large-scale reusable adversarial skill embeddings for physically simulated characters. *ACM Transactions On Graphics (TOG)*, 41(4):1–17, 2022. 2, 3, 5

[50] Ilya A Petrov, Vladimir Guzov, Riccardo Marin, Emre Aksan, Xu Chen, Daniel Cremers, Thabo Beeler, and Gerard Pons-Moll. ECHO: Ego-centric modeling of human-object interactions. *arXiv preprint arXiv:2508.21556*, 2025. 2

[51] Ilya A Petrov, Riccardo Marin, Julian Chibane, and Gerard Pons-Moll. Tridi: Trilateral diffusion of 3d humans, objects, and interactions. In *ICCV*, 2025. 2

[52] Javier Romero, Dimitrios Tzionas, and Michael J. Black. Embodied hands: Modeling and capturing hands and bodies together. *ACM Transactions on Graphics*, 36(6), 2017. 4, 1

[53] Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 627–635. JMLR Workshop and Conference Proceedings, 2011. 5, 4

[54] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017. 4

[55] Yutong Shen, Hangxu Liu, Lei Zhang, Penghui Liu, Ruizhe Xia, Tianyi Yao, and Tongtong Feng. Detach: Cross-domain learning for long-horizon tasks via mixture of disentangled experts. *arXiv preprint arXiv:2508.07842*, 2025. 3

[56] Wandong Sun, Luying Feng, Baoshi Cao, Yang Liu, Yaochu Jin, and Zongwu Xie. Ulc: A unified and fine-grained controller for humanoid loco-manipulation. *arXiv preprint arXiv:2507.06905*, 2025. 2

[57] Omid Taheri, Nima Ghorbani, Michael J Black, and Dimitrios Tzionas. GRAB: A dataset of whole-body human grasping of objects. In *ECCV*, 2020. 2

[58] Chen Tessler, Yoni Kasten, Yunrong Guo, Shie Mannor, Gal Chechik, and Xue Bin Peng. Calm: Conditional adversarial latent models for directable virtual characters. In *SIGGRAPH*, 2023. 2

[59] Chen Tessler, Yunrong Guo, Ofir Nabati, Gal Chechik, and Xue Bin Peng. Maskedmimic: Unified physics-based character control through masked motion inpainting. *ACM Transactions on Graphics (TOG)*, 43(6):1–21, 2024. 3, 4, 5, 6, 8, 1, 2

[60] Chen Tessler, Yifeng Jiang, Erwin Coumans, Zhengyi Luo, Gal Chechik, and Xue Bin Peng. MaskedManipulator: Versatile whole-body control for loco-manipulation. *arXiv preprint arXiv:2505.19086*, 2025. 2, 3, 5, 6, 1

[61] Guy Tevet, Sigal Raab, Setareh Cohan, Daniele Reda, Zhengyi Luo, Xue Bin Peng, Amit H Bermano, and Michiel van de Panne. CLoSD: Closing the loop between simulation and diffusion for multi-task character control. In *ICLR*, 2025. 2

[62] Emanuel Todorov and Michael I Jordan. Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5(11):1226–1235, 2002. 1

[63] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *IROS*, 2012. 7, 1

[64] Takara Everest Truong, Michael Piseno, Zhaoming Xie, and Karen Liu. Pdp: Physics-based character animation via diffusion policy. In *SIGGRAPH Asia*, 2024. 3

[65] Unitree. Unitree g1 humanoid agent ai avatar. https://www.unitree.com/g1/. 2, 3, 4, 1

[66] Ron Vainshtein, Zohar Rimon, Shie Mannor, and Chen Tessler. Task Tokens: A flexible approach to adapting behavior foundation models. *arXiv preprint arXiv:2503.22886*, 2025. 6

[67] Jingbo Wang, Sijie Yan, Bo Dai, and Dahua Lin. Scene-aware generative network for human motion synthesis. In *CVPR*, 2021. 2

[68] Jiashun Wang, Jessica Hodgins, and Jungdam Won. Strategy and skill learning for physics-based table tennis animation. In *SIGGRAPH*, 2024. 3

[69] Jiashun Wang, Yifeng Jiang, Haotian Zhang, Chen Tessler, Davis Rempe, Jessica Hodgins, and Xue Bin Peng. Hil: Hybrid imitation learning of diverse parkour skills from videos. *arXiv preprint arXiv:2505.12619*, 2025. 3

[70] Tingwu Wang, Yunrong Guo, Maria Shugrina, and Sanja Fidler. Unicon: Universal neural controller for physics-based character motion. *arXiv preprint arXiv:2011.15119*, 2020. 2

[71] Wenjia Wang, Liang Pan, Zhiyang Dou, Zhouyingcheng Liao, Yuke Lou, Lei Yang, Jingbo Wang, and Taku Komura. SIMS: Simulating human-scene interactions with real world script planning. In *ICCV*, 2025. 3

[72] Yinhuai Wang, Jing Lin, Ailing Zeng, Zhengyi Luo, Jian Zhang, and Lei Zhang. PhysHOI: Physics-based imitation of dynamic human-object interaction. *arXiv preprint arXiv:2312.04393*, 2023. 2, 3, 6

[73] Jungdam Won, Deepak Gopinath, and Jessica Hodgins. A scalable approach to control diverse behaviors for physically simulated characters. *ACM Transactions on Graphics (TOG)*, 39(4):33–1, 2020. 2

[74] Jungdam Won, Deepak Gopinath, and Jessica Hodgins. Physics-based character controllers using conditional vaes. *ACM Transactions on Graphics (TOG)*, 41(4):1–12, 2022. 3

[75] Lin Wu, Zhixiang Chen, and Jianglin Lan. HOI-Dyn: Learning interaction dynamics for human-object motion diffusion. *arXiv preprint arXiv:2507.01737*, 2025. 2

[76] Yan Wu, Korrawe Karunratanakul, Zhengyi Luo, and Siyu Tang. UniPhys: Unified planner and controller with diffusion for flexible physics-based character control. In *ICCV*, 2025. 3

[77] Zhen Wu, Jiaman Li, Pei Xu, and C Karen Liu. Human-object interaction from human-level instructions. In *ICCV*, 2025. 2, 3

[78] Zeqi Xiao, Tai Wang, Jingbo Wang, Jinkun Cao, Wenwei Zhang, Bo Dai, Dahua Lin, and Jiangmiao Pang. Unified human-scene interaction via prompted chain-of-contacts. In *ICLR*, 2024. 3

[79] Xianghui Xie, Jan Eric Lenssen, and Gerard Pons-Moll. InterTrack: Tracking human object interaction without object templates. In *3DV*, 2024. 2

[80] Zhaoming Xie, Sebastian Starke, Hung Yu Ling, and Michiel van de Panne. Learning soccer juggling skills with layer-wise mixture-of-experts. In *SIGGRAPH*, 2022. 3

[81] Zhaoming Xie, Jonathan Tseng, Sebastian Starke, Michiel van de Panne, and C Karen Liu. Hierarchical planning and control for box loco-manipulation. *arXiv preprint arXiv:2306.09532*, 2023. 3

[82] Liang Xu, Chengqun Yang, Zili Lin, Fei Xu, Yifan Liu, Congsheng Xu, Yiyi Zhang, Jie Qin, Xingdong Sheng, Yunhui Liu, et al. Perceiving and acting in first-person: A dataset and benchmark for egocentric human-object-human interactions. In *ICCV*, 2025. 2

[83] Michael Xu, Yi Shi, KangKang Yin, and Xue Bin Peng. Parc: Physics-based augmentation with reinforcement learning for character controllers. In *SIGGRAPH*, 2025. 2

[84] Sirui Xu, Zhengyuan Li, Yu-Xiong Wang, and Liang-Yan Gui. InterDiff: Generating 3d human-object interactions with physics-informed diffusion. In *ICCV*, 2023. 2, 5

[85] Sirui Xu, Ziyin Wang, Yu-Xiong Wang, and Liang-Yan Gui. Interdreamer: Zero-shot text to 3d dynamic human-object interaction. In *NeurIPS*, 2024. 2

[86] Sirui Xu, Yu-Wei Chao, Liuyu Bian, Arsalan Mousavian, Yu-Xiong Wang, Liangyan Gui, and Wei Yang. Dexplore: Scalable neural control for dexterous manipulation from reference scoped exploration. In *CoRL*, 2025. 4

[87] Sirui Xu, Dongting Li, Yucheng Zhang, Xiyan Xu, Qi Long, Ziyin Wang, Yunzhi Lu, Shuchang Dong, Hezi Jiang,

Akshat Gupta, Yu-Xiong Wang, and Liang-Yan Gui. InterAct: Advancing large-scale versatile 3d human-object interaction generation. In *CVPR*, 2025. 6

[88] Sirui Xu, Hung Yu Ling, Yu-Xiong Wang, and Liang-Yan Gui. InterMimic: Towards universal whole-body control for physics-based human-object interactions. In *CVPR*, 2025. 1, 3, 4, 6, 7, 8, 2

[89] Mengqing Xue, Yifei Liu, Ling Guo, Shaoli Huang, and Changxing Ding. Guiding human-object interactions with rich geometry and relations. In *CVPR*, 2025. 2

[90] Heyuan Yao, Zhenhua Song, Baoquan Chen, and Libin Liu. Controlvae: Model-based learning of generative controllers for physics-based characters. *ACM Transactions on Graphics (TOG)*, 41(6):1–16, 2022. 3, 5

[91] Heyuan Yao, Zhenhua Song, Yuyang Zhou, Tenglong Ao, Baoquan Chen, and Libin Liu. MoConVQ: Unified physics-based motion control via scalable discrete representations. *arXiv preprint arXiv:2310.10198*, 2023. 3

[92] Runyi Yu, Yinhuai Wang, Qihan Zhao, Hok Wai Tsui, Jingbo Wang, Ping Tan, and Qifeng Chen. Skillmimic-v2: Learning robust and generalizable interaction skills from sparse and noisy demonstrations. In *SIGGRAPH*, 2025. 3, 6

[93] Ling-An Zeng, Guohong Huang, Yi-Lin Wei, Shengbo Gu, Yu-Ming Tang, Jingke Meng, and Wei-Shi Zheng. Chain-HOI: Joint-based kinematic chain modeling for human-object interaction generation. In *CVPR*, 2025. 2

[94] Haotian Zhang, Ye Yuan, Viktor Makoviychuk, Yunrong Guo, Sanja Fidler, Xue Bin Peng, and Kayvon Fatahalian. Learning physically simulated tennis skills from broadcast videos. *ACM Transactions on Graphics (TOG)*, 42(4):1–14, 2023. 3

[95] Haozhuo Zhang, Jingkai Sun, Michele Caprio, Jian Tang, Shanghang Zhang, Qiang Zhang, and Wei Pan. HumanoidVerse: A versatile humanoid for vision-language guided multi-object rearrangement. *arXiv preprint arXiv:2508.16943*, 2025. 3

[96] Juze Zhang, Haimin Luo, Hongdi Yang, Xinru Xu, Qianyang Wu, Ye Shi, Jingyi Yu, Lan Xu, and Jingya Wang. NeuralDome: A neural modeling pipeline on multi-view human-object interactions. In *CVPR*, 2023. 2, 6, 7, 8, 1, 4

[97] Juze Zhang, Jingyan Zhang, Zining Song, Zhanhe Shi, Chengfeng Zhao, Ye Shi, Jingyi Yu, Lan Xu, and Jingya Wang. Hoi-mˆ 3: Capture multiple humans and objects interaction within contextual environment. In *CVPR*, 2024. 2

[98] Jinlu Zhang, Yixin Chen, Zan Wang, Jie Yang, Yizhou Wang, and Siyuan Huang. InteractAnything: Zero-shot human object interaction synthesis via llm feedback and object affordance parsing. In *CVPR*, 2025. 2

[99] Xiaohan Zhang, Bharat Lal Bhatnagar, Sebastian Starke, Ilya Petrov, Vladimir Guzov, Helisa Dhamo, Eduardo Pérez-Pellitero, and Gerard Pons-Moll. FORCE: Dataset and method for intuitive physics guided human-object interaction. In *3DV*, 2024. 2

[100] Xiaohan Zhang, Sebastian Starke, Vladimir Guzov, Zhensong Zhang, Eduardo Pérez Pellitero, and Gerard Pons-Moll. SCENIC: Scene-aware semantic navigation with instruction-guided control. *arXiv preprint arXiv:2412.15664*, 2024. 2

[101] Yunbo Zhang, Deepak Gopinath, Yuting Ye, Jessica Hodgins, Greg Turk, and Jungdam Won. Simulation and retargeting of complex multi-character interactions. In *SIGGRAPH*, 2023. 3

[102] Ziyu Zhang, Sergey Bashkirov, Dun Yang, Michael Taylor, and Xue Bin Peng. ADD: Physics-based motion imitation with adversarial differential discriminators. *arXiv preprint arXiv:2505.04961*, 2025. 2

[103] Chengfeng Zhao, Juze Zhang, Jiashen Du, Ziwei Shan, Junye Wang, Jingyi Yu, Jingya Wang, and Lan Xu. I'M HOI: Inertia-aware monocular capture of 3d human-object interactions. In *CVPR*, 2024. 2

[104] Kaifeng Zhao, Yan Zhang, Shaofei Wang, Thabo Beeler, and Siyu Tang. Synthesizing diverse human motions in 3d indoor scenes. In *ICCV*, 2023. 2

[105] Siheng Zhao, Yanjie Ze, Yue Wang, C Karen Liu, Pieter Abbeel, Guanya Shi, and Rocky Duan. ResMimic: From general motion tracking to humanoid whole-body loco-manipulation via residual learning. *arXiv preprint arXiv:2510.05070*, 2025. 2