AN INTERACTION IS MORE THAN TWO MOTIONS: UNIFYING MOTION GENERATION AND POLICY CONTROL FOR DYADIC HUMANOID INTERACTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent advances in physics-based humanoid motion imitation have improved individual robot agility, motivating growing interest in physically coupled humanoid interactions. However, control-oriented approaches often rely on predefined motion templates, while 3D motion generators synthesize diverse interactions but give limited consideration to policy-based execution on physical robots. Our key insight is that *generation, embodiment adaptation, and execution should be guided by a shared representation of interaction relations*. We present DyadLink, a generation-driven framework that synthesizes a compatible follower motion from a leader's reference motion under multimodal conditioning, such as language or music. DyadLink uses a time-varying, directed Shared Relational Graph (SRG) to maintain interaction relations throughout the pipeline. Its role-asymmetric leader–follower formulation prioritizes fidelity to the leader's prescribed motion while allowing the follower to adapt its motion to preserve the intended interaction. To make it actionable across heterogeneous stages, we design *a series of stage-specific, learnable and differentiable relational operators* that translate its shared specification into generation guidance, embodiment-aware optimization objectives, and policy-learning signals. Experiments on InterHuman and CoDance show improved contact fidelity and pairwise geometric consistency across generation, retargeting, and policy execution. We further demonstrate the resulting interactions on physical humanoids. Source code will be made available.[1]

1 INTRODUCTION

Recent advances in humanoid motion control have substantially improved the agility and expressiveness of individual robots (Fu et al., 2025; He et al., 2025b;a). Building on progress in physics-based imitation (Peng et al., 2018; Luo et al., 2023), recent systems retarget human motions and learn feedback policies in simulation to realize increasingly diverse whole-body skills on physical humanoids (Peng et al., 2021; Tessler et al., 2024; He et al., 2025a). These advances have further prompted growing interest in physically coupled interactions between multiple humanoids, which require joint coordination of motion and interaction across agents (Qi et al., 2026; Wang et al., 2026; Xue et al., 2026; Li et al., 2026).

Despite initial progress in humanoid interaction, existing approaches face two main limitations. Policy-oriented methods for dual-humanoid control typically focus on retargeting and tracking complete paired demonstrations, leaving the creation of new interaction references as a separate problem (Chen et al., 2026). Meanwhile, 3D interaction generators can synthesize diverse motion pairs jointly or generate a reaction conditioned on one partner (Liang et al., 2024; Xu et al., 2024; Ghosh et al., 2025), but generally do not address policy-based execution on physical humanoid robots. Although interaction-aware retargeting and control have incorporated some physical priors to improve physical coordination (Zhang et al., 2023; Ji et al., 2025; Huang et al., 2026), a principled framework for unifying generation, retargeting, and control remains lacking. However, naively implementing the unified framework introduces errors.

[1] https://anonymous.4open.science/r/DyadLink-23E0/README.md

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

Table 1: **Capability comparison.** DyadLink spans motion generation, retargeting, and policy control in both human- and robot-motion spaces.

Method	No. of agents	Motion domain	Condition	Generation	Retargeting	Policy control
Human Motion and Character Interaction						
InterGen[2024]	2	Human	Text	✓	—	—
DuoLando[2024]	2	Human	Music + leader	✓	—	—
InterAgent[2026]	2+	Character	Text	✓	—	✓
Humanoid Robot Motion and Control						
Whole-body locomotion[2026]	1	Robot	Terrain + goal	✓	—	✓
UniAct[2025]	1	Robot	Multimodal	✓	—	✓
Harmanoid[2025]	2	Robot→Robot	Paired motion	—	✓	✓
Rhythm[2026]	2	Robot→Robot	Paired motion	—	✓	✓
PAIR + D-STAR[2026]	2	Human→Robot	Human motion	—	✓	✓
DyadLink (ours)	**2**	**Human/Robot→Robot**	**Text/music + leader**	✓	✓	✓

Definitions. Agent counts describe the reported interaction setting; 2+ denotes two or more agents, and → denotes source-to-target embodiment transfer.

Our key insight is that *upstream motion errors can affect downstream execution, while stage-wise motion fidelity alone does not guarantee the preservation of interaction relations* (Yang et al., 2026; Zhang et al., 2023; Chen et al., 2026). Each transition can introduce accumulated errors and erase part of the interaction as a motion pair passes from human-space generation to embodiment adaptation and closed-loop execution. For instance, a generator may produce a plausible response but miss the intended contact; retargeting may preserve both poses while changing their separation, and the controller may track both references yet drift in relative timing under impact. A successful pipeline must therefore preserve an explicit relational specification consistently across all stages.

Enforcing a unified relational constraint across stages, however, poses a practical challenge: *how to impose the same relational constraint on the learning of fundamentally heterogeneous models and representations*. To this end, each stage requires a dedicated relational operator that translates the unified specification into a form compatible with its own optimization process. These operators must be learnable and differentiable, allowing the relational constraint to be embedded directly into the generative model, the embodiment-retargeting model, and policy learning.

In this work, we present DyadLink, a generation-driven framework that synthesizes and executes dual-humanoid interactions from one participant's reference motion and multimodal conditions, e.g., language and music for partnered dance (Figure 1). Specifically, we adopt a role-asymmetric leader–follower formulation: one participant provides the reference behavior, and a compatible partner motion is synthesized in response. We represent the intended inter-agent relations using a Shared Relational Graph (SRG), which provides corresponding relational targets throughout the pipeline. DyadLink first predicts a relational plan $\mathcal{G}^{\text{plan}}$ from the complete reference motion and available conditions. To make this shared specification actionable across heterogeneous stages, we carefully design **a series of stage-specific, learnable and differentiable relational operators** that embed its constraints into motion generation, embodiment retargeting, and policy learning. Across these stages, graph attributes adapt to the corresponding motions and embodiments, while the topology and relational semantics remain shared. This formulation 1) preserves the prescribed behavior while enabling new interactions to be constructed; 2) provides shared relational guidance for subsequent retargeting and control; and 3) reduces the tracking errors induced by paired motion control policies.

Our contributions are fourfold:

Generation-to-execution dyadic interaction. We introduce DyadLink, a motion imitation framework that unifies generation and control for Follower/Leader humanoid interaction under multimodal conditions.

Shared relational representation. We introduce a time-varying, directed SRG with shared topology and attribute semantics across generation, embodiment adaptation, and execution.

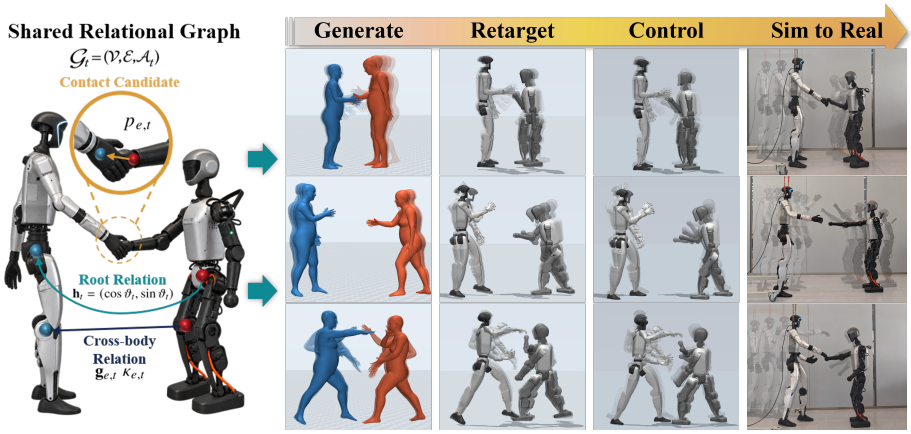

Figure 1: **DyadLink: from generation to execution.** Left: selected root, contact-candidate, and cross-body relations in SRG. Right: interaction examples across generation, retargeting, simulated control and hardware demonstrations.

Stage-specific relational operators. We carefully design a series of operators that instantiate this shared specification through learned generation guidance, morphology-aware retargeting objectives, and relational policy supervision with contact-gated compliant feedback.

Evaluation from synthesis to execution. We introduce PRPE and DRGE to quantify the preservation of semantic cross-agent geometry throughout the pipeline. Evaluations on InterHuman and CoDance, including simulations across diverse humanoid embodiments and real-robot deployment, show that shared relational guidance improves contact fidelity and pairwise geometric consistency.

2 Related Work

Dyadic Motion Generation Models. Recent dyadic generators inherit much of their modeling power from Transformers and diffusion models (Vaswani et al., 2017; Ho et al., 2020; Tevet et al., 2023). One line generates both participants jointly, allowing attention to couple their trajectories throughout denoising (Liang et al., 2024; Javed et al., 2025). Another treats one motion as context and generates the partner's response (Xu et al., 2024; Ghosh et al., 2025). Dance-oriented systems additionally use music and choreography cues (Li et al., 2024), while hierarchical models separate global pair structure or root trajectories from local body motion (Geng et al., 2026; Xu et al., 2025; Lin et al., 2026). These designs improve coordination in the source human-motion space, but their interaction features usually remain internal to the generator. Once the sampled motions are exported, downstream stages receive trajectories rather than an explicit account of the semantic coherence.

Motion Retargeting. Motion retargeting traditionally balances source fidelity against the kinematics, limits, and geometry of the target body (Gleicher, 1998; Araujo et al., 2026). This balance is already delicate for one character: matching every source joint can make the target infeasible, whereas aggressive regularization can erase expressive details. For an interacting pair, there is a second requirement. Retargeting must preserve relations between bodies even when their proportions, reachable sets, and collision shells change. Prior work addresses this through landmark constraints, contact terms, or agent–environment relations (Zhang et al., 2023; Yang et al., 2026). Human-to-humanoid pipelines further turn retargeted demonstrations into references for policy learning (He et al., 2024; 2025b). PAIR preserves contact semantics when converting human–human demonstrations to human–humanoid references, while Rhythm uses interaction geometry to retarget paired motion to dual humanoids (Huang et al., 2026; Chen et al., 2026).

Learning-Based Humanoid Whole-Body Control. Physics-based imitation turns kinematic references into feedback policies that tolerate perturbations and model error (Peng et al., 2018; 2021; Luo et al., 2023; Tessler et al., 2024). Recent systems transfer increasingly diverse retargeted skills from simulation to humanoid hardware (Li et al., 2025; Cheng et al., 2024; He et al., 2024; Fu et al., 2025; He et al., 2025b;a; Ze et al., 2025). Tracking two reference motions, however, does not

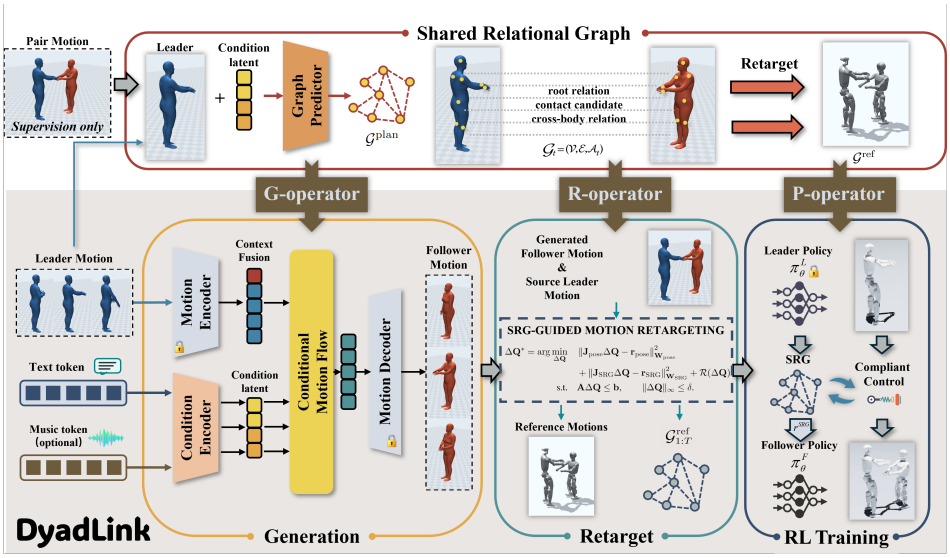

Figure 2: **DyadLink overview.** SRG guides motion generation, retargeting, and control through three stage-specific operators to preserve the intended interaction relations.

guarantee that their relation survives closed-loop execution. Small phase errors can break a short contact; independent pose errors can accumulate into a large pairwise displacement; an impact can require one agent to yield without abandoning the interaction. Recent interactive controllers address these effects through interaction-aware human–humanoid policies, graph-based rewards for dual humanoids, or interaction-graph exteroception in a text-conditioned diffusion policy (Huang et al., 2026; Chen et al., 2026; Li et al., 2026). These signals are often designed only for the control task.

3 METHOD

DyadLink generates and executes a compatible Follower motion from a complete Leader trajectory and the available conditions (Figure 2). Its Shared Relational Graph (SRG) carries interaction targets through generation, retargeting, and control. The stages share the graph topology and attribute semantics, while adapting its values to the corresponding embodiment.

3.1 SHARED RELATIONAL GRAPH

At frame t, SRG is $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}, \mathcal{A}_t)$, with fixed semantic topology and time-varying attributes. Its nodes are two role-labeled sets of 22 body anchors. Within-agent skeleton edges provide anatomical context. Cross-agent relations are directed from Leader to Follower and describe root placement, candidate contacts, and selected cross-body geometry. Human joints and robot link-attached proxies instantiate the same semantic anchors.

For a cross-agent edge $e = (i, j)$, we define

$$\mathbf{g}_{e,t} = (\mathbf{B}_t^{\mathrm{L}})^{\top} \left(\mathbf{x}_{j,t}^{\mathrm{F}} - \mathbf{x}_{i,t}^{\mathrm{L}}\right), \qquad \Delta\mathbf{g}_{e,t} = \mathbf{g}_{e,t} - \mathbf{g}_{e,t-1}, \tag{1}$$

where $\mathbf{B}_t^{\mathrm{L}}$ is the Leader heading frame. This expresses relative geometry independently of the pair's common translation and yaw. SRG also records relative heading $\mathbf{h}_t = (\cos\vartheta_t, \sin\vartheta_t)$, contact probabilities $p_{e,t}$ on candidate contacts, and signed clearances $\kappa_{e,t}$ on avoidance pairs. The predicted plan, retargeted reference, and executed graph retain these meanings without requiring their numerical values to remain unchanged. Appendix B specifies their construction.

3.2 G-OPERATOR FOR MOTION GENERATION

The G-operator guides Follower synthesis through relational conditioning and bounded motion refinement. Graph-based supervision links the predicted plan to the relations realized by the generated

motion. We build on UMF's reaction-flow backbone (Huang & Celiktutan, 2026) with conditional flow matching (Lipman et al., 2023). Before sampling the Follower, a graph–temporal predictor estimates the intended cross-agent relations from the Leader motion $\mathbf{X}^{\mathrm{L}}_{1:T}$ and condition tokens $\mathbf{c}$:

$$\left(\mathcal{G}^{\mathrm{plan}}, \mathbf{z}^{\mathcal{G}}\right) = \mathcal{P}_\phi\left(\mathbf{X}^{\mathrm{L}}_{1:T}, \mathbf{c}\right), \qquad \mathbf{z}^{\mathrm{ctx}} \quad = \mathcal{F}\left(\mathbf{z}^{\mathrm{L}}, \mathbf{z}^{\mathcal{G}}, \mathbf{c}\right). \tag{2}$$

The predictor combines Leader-anchor features, learned Follower queries, and condition tokens through graph and temporal processing. Shared hidden features produce both the planned relations and $\mathbf{z}^{\mathcal{G}}$. This relational latent is fused with the Leader latent and predicted relative-root context to condition reaction-flow sampling. After decoding, the planned root, heading, and wrist-contact relations guide bounded motion refinement.

Let $\mathcal{G}^*$ be the graph of the ground-truth pair and $\widehat{\mathcal{G}}$ the graph reconstructed from the Leader and the Follower prediction after decoding and refinement. We optimize

$$\mathcal{L}_{\mathrm{gen}} = \mathcal{L}_{\mathrm{FM}} + \mathcal{L}_{\mathrm{aux}} + \mathcal{D}_{\mathrm{plan}}\left(\mathcal{G}^{\mathrm{plan}}, \mathcal{G}^*\right) + \mathcal{D}_{\mathrm{state}}\left(\widehat{\mathcal{G}}, \mathcal{G}^*\right) + \mathcal{D}_{\mathrm{cons}}\left(\widehat{\mathcal{G}}, \mathrm{sg}[\mathcal{G}^{\mathrm{plan}}]\right). \tag{3}$$

Here, $\mathcal{L}_{\mathrm{FM}}$ is the flow-matching objective, sg stops gradients through the plan target, and $\mathcal{L}_{\mathrm{aux}}$ collects auxiliary motion and root supervision. The plan loss uses masked Smooth-L_1 regression for continuous relations and confidence-weighted focal classification for contact prediction. The state loss supervises the realized geometry and contacts while penalizing clearance deficits. The consistency loss compares realized attributes with the detached plan.

During active contact intervals, supervision for each active contact edge uses its relative offset at contact onset. Together, these objectives supervise the predicted plan and the realized interaction while encouraging their consistency. Appendix C gives their calculation.

3.3 R-OPERATOR FOR MOTION RETARGETING

The R-operator translates the generated motion pair and SRG plan into embodiment-aware relational targets for constrained differential IK(Araujo et al., 2026). Source relations are blended with their planned counterparts in $\mathcal{G}^{\mathrm{plan}}$ within a calibrated Leader frame, then adapted using source surface gaps and the target collision geometry.

Let $\mathbf{Q}$ collect the optimized configurations and $\Delta\mathbf{Q} \in \mathbb{R}^{n_v}$ their tangent-coordinate increment. Suppressing frame and iteration indices, each local solve is

$$\begin{aligned} \Delta\mathbf{Q}^\star = \arg\min_{\Delta\mathbf{Q}} \quad & \|\mathbf{J}_{\mathrm{pose}}\Delta\mathbf{Q} - \mathbf{r}_{\mathrm{pose}}\|^2_{\mathbf{W}_{\mathrm{pose}}} + \|\mathbf{J}_{\mathrm{SRG}}\Delta\mathbf{Q} - \mathbf{r}_{\mathrm{SRG}}\|^2_{\mathbf{W}_{\mathrm{SRG}}} + \mathcal{R}(\Delta\mathbf{Q}) \\ \text{s.t.} \quad & \mathbf{A}\Delta\mathbf{Q} \le \mathbf{b}, \qquad \|\Delta\mathbf{Q}\|_\infty \le \delta. \end{aligned} \tag{4}$$

Here, $\|\mathbf{v}\|^2_{\mathbf{W}} = \mathbf{v}^\top\mathbf{W}\mathbf{v}$, with nonnegative task weights. The residuals specify target-minus-current task errors, and the Jacobians map tangent increments to local task changes. The pose term tracks mapped body positions and orientations. The SRG term tracks adapted root, contact, and cross-body displacements together with relative heading. $\mathcal{R}$ contains quadratic damping, posture, and temporal regularization; the constraints collect the enabled joint, motion-rate, terrain, and collision limits, with step size bounded by $\delta > 0$.

Leader tracking receives greater weight, while active Follower contact endpoints allow more positional deviation from the source. This asymmetric weighting lets the Follower accommodate relational targets while favoring preservation of the Leader's prescribed behavior. For a prescribed human Leader, only the Follower is optimized. After fitting, $\mathcal{G}^{\mathrm{ref}}$ contains geometry, heading, and clearance reconstructed from the retargeted motion, with temporal increments recomputed from its edge vectors. Planned contact probabilities are retained as the contact intent. Appendix D describes target adaptation and the local solve.

3.4 P-OPERATOR FOR MOTION CONTROL

The P-operator converts the reference SRG and executed graph state into relational rewards and privileged Critic inputs. Planned contact probabilities additionally gate an external compliant response during training and deployment. We train reference-tracking policies with PPO (Schulman et al., 2017). For two robots, the Leader policy is trained first and frozen during Follower training.

The Actor observes its reference, proprioception, and previous action, while partner information and SRG residuals provide privileged Critic inputs and reward supervision.

We use the reward $r_t = r_t^{\text{base}} + r_t^{\text{SRG}}$, where r_t^{base} builds on the tracking reward in BeyondMimic (Liao et al., 2025), and r_t^{SRG} provides relational supervision. The SRG reward compares executed relations with the robot-scale reference:

$$r_t^{\text{SRG}} = \lambda_g e^{-D_{g,t}/\sigma_g^2} + \lambda_\Delta e^{-D_{\Delta,t}/\sigma_\Delta^2} + \lambda_h \rho_h(\epsilon_{h,t}) + \lambda_p e^{-D_{p,t}/\sigma_p^2} + \lambda_\kappa e^{-D_{\kappa,t}/m^2}. \quad (5)$$

Let $\mathcal{E}_{\text{geo}}$ contain the root, contact-candidate, and cross-body edges, and $\mathcal{E}_{\text{avoid}}$ the avoidance pairs. The geometry and clearance errors are

$$D_{g,t} = \frac{\sum_{e\in\mathcal{E}_{\text{geo}}} \omega_{e,t} \left\|\mathbf{g}_{e,t}^{\text{exec}} - \mathbf{g}_{e,t}^{\text{ref}}\right\|_2^2}{\sum_{e\in\mathcal{E}_{\text{geo}}} \omega_{e,t}}, D_{\kappa,t} \quad = \frac{1}{|\mathcal{E}_{\text{avoid}}|} \sum_{e\in\mathcal{E}_{\text{avoid}}} \left[\max(\kappa_{e,t}^{\text{ref}}, m) - \kappa_{e,t}^{\text{exec}}\right]_+^2. \quad (6)$$

Here, $[x]_+ = \max(x, 0)$, $m > 0$ is a clearance margin, and the geometric weights emphasize root relations and high-probability planned contacts. $D_{\Delta,t}$ is the unweighted mean squared Euclidean error between executed and reference edge increments, and $D_{p,t}$ is the mean squared difference between executed contact scores and planned probabilities. The heading term combines cosine and inverse Smooth-L_1 agreement for the wrapped angular error $\epsilon_{h,t}$. Appendix E gives the remaining definitions.

To allow the Follower to yield during physical interaction, we apply an external admittance-style correction (Hogan, 1985; Lee et al., 2022; Hartmann et al., 2024). Bounded hand forces produce damped joint-position offsets, with planned contact probabilities gating the response. Large unexpected forces override this gate. The offsets modify the Actor's position targets during both training and execution, complementing the learned relational behavior with force-responsive motion.

4 Experiments

We evaluate whether the interaction established during generation can be preserved through embodiment adaptation and physical execution. Our analysis distinguishes motion-distribution quality from contact fidelity and cross-agent geometric consistency.

4.1 Experimental Setup

Datasets and settings. We evaluate on InterHuman (Liang et al., 2024) and CoDance, covering general dyadic interactions and partnered dance. CoDance combines DD100 (Li et al., 2024) and CoMPAS3D (Burkanova et al., 2025).

DyadLink receives Leader motion and text on InterHuman, with music additionally provided on CoDance. Motion sequences are represented at 30 Hz, and neural models are trained on a single NVIDIA A100 GPU. All retargeting methods receive the same motion pairs generated by DyadLink.

Evaluation. Generation is assessed using the frozen InterCLIP evaluator (Liang et al., 2024). We report FID (Heusel et al., 2017), R-precision, Matching Score, and Diversity (Guo et al., 2022). These metrics assess distributional quality and text–motion alignment. Contact F1 and contact error on the generated pairs assess agreement with the reference contact states and timing.

For embodiment transfer and execution, we use Pairwise Relative Position Error (PRPE) and Dyadic Relation Graph Error (DRGE). Using the Leader-frame displacement from Eq. (1), we define

$$\begin{aligned} E_{\text{PRPE}}^a &= \frac{1}{T}\sum_{t=1}^{T} \left\|\mathbf{g}_{\text{root},t}^a - s_{\text{pair}}\mathbf{g}_{\text{root},t}^*\right\|_2, \\ E_{\text{DRGE}}^a &= \frac{\sum_{t=1}^{T}\sum_{e\in\mathcal{E}_{\text{eval}}} w_e \left\|\mathbf{g}_{e,t}^a - s_e\mathbf{g}_{e,t}^*\right\|_2}{T\sum_{e\in\mathcal{E}_{\text{eval}}} w_e}. \end{aligned} \quad (7)$$

Here, $a \in \{\text{ref}, \text{exec}\}$ denotes the retargeted reference or executed motion, and $*$ denotes the source pair. The scale factors s_{pair} and s_e account for source-to-robot morphology. PRPE measures the root relation, while DRGE averages errors over a fixed set of weighted semantic edges. Additional experimental results and implementation details are provided in Appendix F.

Table 2: **Performance comparison for motion generation.** Contact F1 and contact error assess interaction fidelity in the generated motion pairs. For Diversity, bold indicates the value closest to the ground-truth value shown in the header.

InterHuman								
Method	FID ↓	R@1 ↑	R@2 ↑	R@3 ↑	Matching ↓	Diversity ≈ 7.6867	Contact F1 ↑	Contact error (s) ↓
InterMask[2025]	5.2796	0.1577	0.2429	0.3046	3.9036	7.0670	0.0821 ± 0.0050	5.2778 ± 0.0645
ReMoGen[2026]	5.9911	0.0703	0.1278	0.1710	4.1056	6.4739	0.1345 ± 0.0055	5.6735 ± 0.0625
InterGen[2024]	6.0842	0.1205	0.1839	0.2302	3.9659	6.9344	0.0851 ± 0.0075	4.8101 ± 0.0545
UMF[2026]	4.8046	0.3681	0.5120	0.5896	**3.8335**	7.3160	0.0491 ± 0.0065	5.7665 ± 0.0345
Ours (DyadLink)	**3.7594**	**0.3689**	**0.5620**	**0.6350**	3.9018	**7.7764**	**0.1941** ± 0.0045	**4.1542** ± 0.0912
CoDance								
Method	FID ↓	R@1 ↑	R@2 ↑	R@3 ↑	Matching ↓	Diversity ≈ 3.0038	Contact F1 ↑	Contact error (s) ↓
DuoLando[2024]	23.9967	0.0789	0.1316	0.1579	4.2141	3.6963	0.0653 ± 0.0047	4.7230 ± 0.2392
ReMoS[2025]	28.3648	0.0000	0.0526	0.1316	4.1508	**2.9935**	0.0722 ± 0.0053	5.56234 ± 0.2341
UMF[2026]	18.7038	0.1445	**0.2578**	0.3789	4.2293	4.0120	0.0111 ± 0.0012	6.5722 ± 0.4258
Ours (DyadLink)	**6.1514**	**0.1859**	0.2188	**0.4023**	**4.1324**	3.1365	**0.1041** ± 0.0032	**3.8341** ± 0.3789

Table 3: **Performance comparison for Interaction-aware motion retargeting.** The first four metrics assess relational geometry and penetration of the retargeted references. The final three report execution tracking errors against those references.

InterHuman							
Method	PRPE (m) ↓	DRGE (m) ↓	Pair pen. ↓	Max pen. (cm) ↓	Global pos. (m) ↓	Local pos. (m) ↓	Rotation (°) ↓
OmniRetarget[2026]	0.21 ± 0.08	0.16 ± 0.06	0.10 ± 0.13	2.13 ± 2.48	0.78 ± 0.08	0.37 ± 0.07	88.62 ± 10.97
GMR[2026]	0.14 ± 0.04	0.17 ± 0.06	0.14 ± 0.10	3.36 ± 2.94	**0.78** ± 0.09	0.35 ± 0.06	**84.38** ± 5.84
PHC-style fitting[2023]	0.46 ± 0.18	0.41 ± 0.19	0.04 ± 0.05	1.49 ± 2.36	0.98 ± 0.11	0.39 ± 0.05	92.51 ± 20.76
Ours (DyadLink)	**0.13** ± 0.02	**0.12** ± 0.06	**0.04** ± 0.16	**1.33** ± 2.38	0.84 ± 0.12	**0.32** ± 0.04	85.22 ± 34.42
CoDance							
Method	PRPE (m) ↓	DRGE (m) ↓	Pair pen. ↓	Max pen. (cm) ↓	Global pos. (m) ↓	Local pos. (m) ↓	Rotation (°) ↓
OmniRetarget[2026]	**0.07** ± 0.02	0.33 ± 0.14	0.64 ± 0.09	9.39 ± 0.81	1.17 ± 0.22	0.40 ± 0.07	100.47 ± 17.86
GMR[2026]	0.22 ± 0.10	0.59 ± 0.07	0.47 ± 0.12	8.25 ± 0.95	1.17 ± 0.14	0.46 ± 0.04	109.88 ± 3.80
PHC-style fitting[2023]	0.42 ± 0.11	0.60 ± 0.03	0.29 ± 0.29	6.51 ± 1.02	1.30 ± 0.23	0.37 ± 0.02	**96.45** ± 7.80
Ours (DyadLink)	0.15 ± 0.04	**0.32** ± 0.11	**0.29** ± 0.11	**6.00** ± 1.89	**1.15** ± 0.02	**0.36** ± 0.03	104.51 ± 6.98

4.2 Comparison with State-of-the-Art Methods

Motion generation. Table 2 compares DyadLink with joint interaction generators, response-conditioned models, and dance-accompaniment methods. DyadLink combines the lowest FID on both datasets with the highest Contact F1 and lowest contact timing error, showing gains in contact fidelity already at the generation stage. On InterHuman, UMF achieves competitive retrieval and the best Matching Score, yet substantially lower Contact F1. This contrast illustrates why condition alignment alone is insufficient to assess interaction fidelity. DyadLink also leads retrieval on InterHuman and comes closest to its reference Diversity, although other methods remain better on individual motion metrics on CoDance.

SRG-guided retargeting. We compare against OmniRetarget (Yang et al., 2026), GMR (Araujo et al., 2026), and a PHC-style fitting baseline (Luo et al., 2023). DyadLink obtains the lowest mean DRGE on both datasets, together with favorable penetration results (Table 3). The distinction between root placement and broader body-part relations is particularly clear on CoDance: OmniRetarget has lower PRPE, but this advantage does not extend to DRGE or penetration. Accurate root placement alone is therefore insufficient to preserve the spatial structure of an interaction. The results support adapting semantic cross-agent relations during embodiment transfer, rather than relying solely on the participants' individual reference motions.

Physics-based execution. The relative advantage in pairwise geometry remains visible after execution (Table 4). DyadLink achieves the lowest executed PRPE and DRGE on both datasets, alongside higher contact recall and lower contact timing error. This improvement is not uniform: the PHC-style baseline slightly exceeds DyadLink in CoDance contact F1.

Table 4: **Interaction preservation after execution.** Each method's retargeted references are executed under a shared protocol and evaluated against the common source interaction. Contact error measures onset/offset timing discrepancies.

InterHuman

Method	Executed PRPE (m) ↓	Executed DRGE (m) ↓	Contact Recall ↑	Contact F1 ↑	Contact error (s) ↓
OmniRetarget[2026]	0.771 ± 0.047	0.971 ± 0.146	0.615 ± 0.371	0.698 ± 0.329	0.536 ± 0.204
GMR[2026]	0.770 ± 0.017	0.941 ± 0.103	0.526 ± 0.295	0.639 ± 0.274	0.568 ± 0.197
PHC-style fitting[2023]	0.977 ± 0.089	1.117 ± 0.097	0.496 ± 0.258	0.622 ± 0.267	0.567 ± 0.219
Ours (DyadLink)	**0.701** ± 0.152	**0.909** ± 0.088	**0.689** ± 0.312	**0.724** ± 0.273	**0.520** ± 0.216

CoDance

Method	Executed PRPE (m) ↓	Executed DRGE (m) ↓	Contact Recall ↑	Contact F1 ↑	Contact error (s) ↓
OmniRetarget[2026]	1.180 ± 0.254	1.362 ± 0.301	0.322 ± 0.014	0.479 ± 0.017	0.888 ± 0.034
GMR[2026]	1.078 ± 0.163	1.356 ± 0.243	0.207 ± 0.201	0.316 ± 0.274	0.954 ± 0.028
PHC-style fitting[2023]	1.154 ± 0.213	1.191 ± 0.109	0.346 ± 0.126	**0.499** ± 0.135	0.838 ± 0.030
Ours (DyadLink)	**1.055** ± 0.009	**1.166** ± 0.035	**0.408** ± 0.056	0.492 ± 0.089	**0.830** ± 0.002

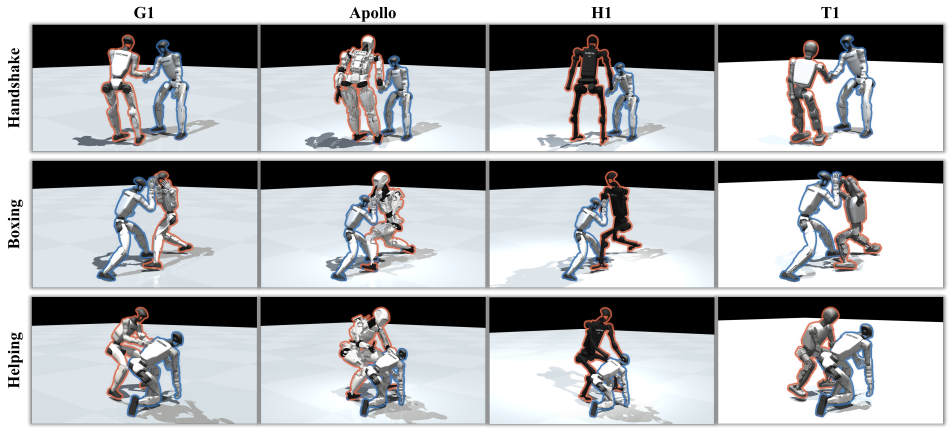

Figure 3: **Cross-embodiment control in simulation.** Each row shows an interaction across different orange Follower embodiments, with a fixed blue G1 Leader.

Cross-embodiment transfer and real-robot demonstrations. Figure 3 shows interactions transferred across different Follower embodiments with a fixed G1 Leader, demonstrating that DyadLink preserves inter-agent coordination under morphological variation. We further validate the learned interactions on physical humanoids (Figure 1). For human–robot interaction, both roles use the G1 embodiment during policy training to reduce morphological variation. These results demonstrate feasibility in the tested settings; additional sequences are provided in Appendix F.7. More results are available on our demo website.

4.3 Ablation Study

We ablate stage-specific uses of SRG in generation, retargeting, and policy learning. Generation and retargeting variants remove relational guidance or its contact-specific components. The policy-learning ablation changes reward terms while holding the retargeted references, policy inputs, and compliant module fixed. Appendix F.4 specifies the interventions.

Motion quality does not imply interaction fidelity. The generation ablation exposes a limitation of distributional metrics (Table 5). Without SRG, FID improves on both datasets, yet contact F1 decreases and errors in inter-hand distance and relative-root position increase. The two criteria therefore rank the models differently: a better match to the motion distribution can coincide with poorer agreement between the partners. Disabling contact guidance also reduces contact F1 and increases hand-distance error relative to the full model. Explicit contact guidance thus complements relative geometry when generating a compatible response.

Table 5: **SRG ablation in motion generation.** Hand MAE measures inter-hand distance error on reference-contact frames; Root RMSE measures relative-root position error. Both are in meters.

Dataset	Variant	FID ↓	R@1 ↑	R@2 ↑	R@3 ↑	Contact F1 ↑	Hand MAE ↓	Root RMSE ↓
InterHuman	w/o SRG	3.542	0.256	0.346	0.410	0.110	0.486	0.547
	w/o contact guidance	**3.171**	0.258	0.449	0.582	0.096	0.497	0.539
	Full	3.759	**0.369**	**0.562**	**0.635**	**0.194**	**0.325**	**0.491**
CoDance	w/o SRG	**5.569**	0.167	**0.286**	0.382	0.056	0.408	0.438
	w/o contact guidance	6.682	0.148	0.250	0.398	0.074	0.427	0.410
	Full	6.151	**0.185**	0.219	**0.402**	**0.104**	**0.338**	**0.378**

Table 6: **SRG ablation in motion retargeting.** PRPE, DRGE, and pair penetration assess the retargeted references. Position tracking and contact metrics assess subsequent execution. Position errors are in meters and contact error in seconds.

Variant	PRPE ↓	DRGE ↓	Pair pen. ↓	Global pos. ↓	Local pos. ↓	Contact F1 ↑	Contact error ↓
w/o SRG	0.1153	0.0723	0.0065	0.8742	0.3728	0.7954	0.5734
w/o contact guidance	0.1056	0.0711	0.0081	0.8574	0.3641	0.7672	0.5789
Full	**0.1052**	**0.0710**	**0.0027**	**0.8568**	**0.3636**	**0.8663**	**0.5096**

Contact guidance during embodiment adaptation. The retargeting ablation reveals a related distinction (Table 6). Removing all SRG guidance increases both relational errors. Removing only contact guidance leaves PRPE and DRGE close to the full model, but worsens contact F1 and contact timing. Average geometric agreement can therefore conceal errors in whether and when the partners interact. These results suggest that geometric alignment alone is insufficient: contact-specific objectives during retargeting also help preserve contact states and timing in subsequent execution.

Relational guidance during policy learning. With retargeted references held fixed, Table 7 evaluates the contribution of SRG reward supervision. The full policy achieves the best reported means in relational geometry, contact agreement, and contact timing. Removing contact-related rewards preserves better root alignment than removing all SRG-related rewards, but DRGE remains close to the SRG-free result and contact agreement is lower than with the full policy. These results support contact-aware reward supervision beyond root tracking alone.

Table 7: **SRG reward ablation in reinforcement learning.** Results are reported as mean ± sample standard deviation across 17 InterHuman motion clips. All variants use the same retargeted references and differ only in the ablated reward terms.

Variant	PRPE (m) ↓	DRGE (m) ↓	Contact Recall ↑	Contact F1 ↑	Contact error (s) ↓
w/o SRG rewards	0.2537 ± 0.0201	0.3022 ± 0.0321	0.8279 ± 0.0124	0.8288 ± 0.0212	0.3802 ± 0.2308
w/o contact rewards	0.1924 ± 0.0324	0.2982 ± 0.0121	0.9038 ± 0.0235	0.8742 ± 0.0178	0.3302 ± 0.2506
Full	**0.1705** ± 0.0418	**0.2466** ± 0.0401	**0.9689** ± 0.0135	**0.9598** ± 0.0347	**0.2491** ± 0.2608

Limitations. DyadLink currently assumes a complete Leader reference and motion-specific policies, with predefined relational structure and role assignments. On-the-fly adaptation should distinguish transient tracking deviations from genuine changes in interaction intent. A promising direction is to revise SRG relations and targets from partial observations while preserving temporal continuity, so that adaptation accommodates changing contact requirements without disrupting coordination.

5 Conclusion

We presented DyadLink, a generation-to-execution framework for dyadic humanoid interaction. A shared relational specification connects Leader-conditioned generation, embodiment retargeting, and policy execution through stage-specific operators. Comparisons and ablations on InterHuman and CoDance support explicit relational guidance for contact fidelity and pairwise geometric consistency, and show that distributional motion quality alone does not capture interaction quality. Hardware demonstrations further illustrate feasibility in the tested settings.

AI USE STATEMENT

We used generative AI tools to improve the clarity and fluency of the manuscript, assist with table preparation and formatting, and scaffold the project demonstration website. These tools also provided feedback on methodological descriptions and the interpretation of experimental results. The authors reviewed and revised all AI-assisted content and take full responsibility for the accuracy and integrity of the final manuscript and accompanying materials.

REFERENCES

Joao Pedro Araujo, Yanjie Ze, Pei Xu, Jiajun Wu, and C. Karen Liu. Retargeting matters: General motion retargeting for humanoid motion tracking. In *IEEE International Conference on Robotics and Automation*, June 2026. URL https://jaraujo98.github.io/retargeting_matters/.

Bermet Burkanova, Yasaman Etesam, Payam Jome Yazdian, Trinity Evans, Chuxuan Zhang, Zoe Stanley, Paige Tuttösí, and Angelica Lim. CoMPAS3D: A dataset and benchmark for interactive motion. *arXiv preprint arXiv:2507.19684*, 2025. URL https://arxiv.org/abs/2507.19684.

Hongjin Chen, Wei Zhang, Pengfei Li, Shihao Ma, Ke Ma, Yujie Jin, Zijun Xu, Xiaohui Wang, Yupeng Zheng, Zining Wang, Jieru Zhao, Yilun Chen, and Wenchao Ding. Rhythm: Learning interactive whole-body control for dual humanoids. *arXiv preprint arXiv:2603.02856*, 2026. doi: 10.48550/arXiv.2603.02856. URL https://arxiv.org/abs/2603.02856.

Xuxin Cheng, Yandong Ji, Junming Chen, Ruihan Yang, Ge Yang, and Xiaolong Wang. Expressive whole-body control for humanoid robots. In *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024. doi: 10.15607/RSS.2024.XX.107. URL https://roboticsproceedings.org/rss20/p107.html.

Zipeng Fu, Qingqing Zhao, Qi Wu, Gordon Wetzstein, and Chelsea Finn. HumanPlus: Humanoid shadowing and imitation from humans. In *Proceedings of The 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pp. 2828–2844. PMLR, 2025. URL https://proceedings.mlr.press/v270/fu25a.html.

Zichen Geng, Zeeshan Hayder, Bo Miao, Jian Liu, Wei Liu, and Ajmal Mian. Disentangled hierarchical VAE for 3D human-human interaction generation. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/ffd4f5a2ea6b93e9bf5af9264d568cf2-Abstract-Conference.html.

Anindita Ghosh, Rishabh Dabral, Vladislav Golyanik, Christian Theobalt, and Philipp Slusallek. ReMoS: 3D motion-conditioned reaction synthesis for two-person interactions. In *Computer Vision – ECCV 2024*, volume 15094 of *Lecture Notes in Computer Science*, pp. 418–437. Springer, Cham, 2025. doi: 10.1007/978-3-031-72764-1_24. URL https://doi.org/10.1007/978-3-031-72764-1_24.

Michael Gleicher. Retargeting motion to new characters. In *Proceedings of ACM SIGGRAPH 98*, pp. 33–42, 1998. doi: 10.1145/280814.280820. URL https://graphics.cs.wisc.edu/Papers/1998/Gle98/.

Chuan Guo, Shihao Zou, Xinxin Zuo, Sen Wang, Wei Ji, Xingyu Li, and Li Cheng. Generating diverse and natural 3D human motions from text. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5152–5161, 2022. URL https://openaccess.thecvf.com/content/CVPR2022/html/Guo_Generating_Diverse_and_Natural_3D_Human_Motions_From_Text_CVPR_2022_paper.html.

Adrian Hartmann, Dongho Kang, Fatemeh Zargarbashi, Miguel Zamora, and Stelian Coros. Deep compliant control for legged robots. In *IEEE International Conference on Robotics and Automation*, pp. 11421–11427, 2024. doi: 10.1109/ICRA57147.2024.10611209.

Tairan He, Zhengyi Luo, Wenli Xiao, Chong Zhang, Kris Kitani, Changliu Liu, and Guanya Shi. Learning human-to-humanoid real-time whole-body teleoperation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 8944–8951, 2024. doi: 10.1109/IROS58592.2024.10801984. URL https://human2humanoid.com/.

Tairan He, Jiawei Gao, Wenli Xiao, Yuanhang Zhang, Zi Wang, Jiashun Wang, Zhengyi Luo, Guanqi He, Nikhil Sobanbabu, Chaoyi Pan, Zeji Yi, Guannan Qu, Kris Kitani, Jessica K. Hodgins, Linxi Fan, Yuke Zhu, Changliu Liu, and Guanya Shi. ASAP: Aligning simulation and real-world physics for learning agile humanoid whole-body skills. In *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2025a. doi: 10.15607/RSS.2025.XXI.066. URL https://www.roboticsproceedings.org/rss21/p066.html.

Tairan He, Zhengyi Luo, Xialin He, Wenli Xiao, Chong Zhang, Weinan Zhang, Kris M. Kitani, Changliu Liu, and Guanya Shi. OmniH2O: Universal and dexterous human-to-humanoid whole-body teleoperation and learning. In *Proceedings of The 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pp. 1516–1540. PMLR, 2025b. URL https://proceedings.mlr.press/v270/he25b.html.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a local nash equilibrium. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a1d694707eb0fefe65871369074926d-Abstract.html.

Jonathan Ho, Ajay N. Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html.

Neville Hogan. Impedance control: An approach to manipulation: Part I—theory. *Journal of Dynamic Systems, Measurement, and Control*, 107(1):1–7, 1985. doi: 10.1115/1.3140702. URL https://doi.org/10.1115/1.3140702.

Guanhe Huang and Oya Celiktutan. Unified number-free text-to-motion generation via flow matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 23570–23580, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Huang_Unified_Number-Free_Text-to-Motion_Generation_Via_Flow_Matching_CVPR_2026_paper.html.

Wei-Jin Huang, Yue-Yi Zhang, Yi-Lin Wei, Zhi-Wei Xia, Juantao Tan, Yuan-Ming Li, Zhilin Zhao, and Wei-Shi Zheng. Beyond mimicry: Learning whole-body human-humanoid interaction from human-human demonstrations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 30740–30749, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Huang_Beyond_Mimicry_Learning_Whole-Body_Human-Humanoid_Interaction_from_Human-Human_Demonstrations_CVPR_2026_paper.html.

Muhammad Gohar Javed, Chuan Guo, Li Cheng, and Xingyu Li. InterMask: 3D human interaction generation via collaborative masked modeling. In *International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=ZAyuwJYN8N.

Kaiyang Ji, Ye Shi, Zichen Jin, Kangyi Chen, Lan Xu, Yuexin Ma, Jingyi Yu, and Jingya Wang. Towards immersive Human-X interaction: A real-time framework for physically plausible motion synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10173–10183, 2025. URL https://openaccess.thecvf.com/content/ICCV2025/html/Ji_Towards_Immersive_Human-X_Interaction_A_Real-Time_Framework_for_Physically_Plausible_ICCV_2025_paper.html.

Nan Jiang, Zimo He, Wanhe Yu, Lexi Pang, Yunhao Li, Hongjie Li, Jieming Cui, Yuhan Li, Yizhou Wang, Yixin Zhu, and Siyuan Huang. UniAct: Unified motion generation and action streaming for humanoid robots. *arXiv preprint arXiv:2512.24321*, 2025. URL https://arxiv.org/abs/2512.24321.

Harold W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1–2):83–97, 1955. doi: 10.1002/nav.3800020109.

Seunghwan Lee, Phil Sik Chang, and Jehee Lee. Deep compliant control. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings*, pp. 1–9, 2022. doi: 10.1145/3528233.3530719. URL https://doi.org/10.1145/3528233.3530719.

Bin Li, Ruichi Zhang, Han Liang, Jingyan Zhang, Juze Zhang, Xin Chen, Lan Xu, Jingyi Yu, and Jingya Wang. InterAgent: Physics-based multi-agent command execution via diffusion on interaction graphs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15253–15265, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Li_InterAgent_Physics-based_Multi-agent_Command_Execution_via_Diffusion_on_Interaction_Graphs_CVPR_2026_paper.html.

Siyao Li, Tianpei Gu, Zhitao Yang, Zhengyu Lin, Ziwei Liu, Henghui Ding, Lei Yang, and Chen Change Loy. DuoLando: Follower GPT with off-policy reinforcement learning for dance accompaniment. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=GW4j4n2cjH.

Zhongyu Li, Xue Bin Peng, Pieter Abbeel, Sergey Levine, Glen Berseth, and Koushil Sreenath. Reinforcement learning for versatile, dynamic, and robust bipedal locomotion control. *The International Journal of Robotics Research*, 44(5):840–888, 2025. doi: 10.1177/02783649241285161. URL https://doi.org/10.1177/02783649241285161.

Han Liang, Wenqian Zhang, Wenxuan Li, Jingyi Yu, and Lan Xu. InterGen: Diffusion-based multi-human motion generation under complex interactions. *International Journal of Computer Vision*, 132(9):3463–3483, 2024. doi: 10.1007/s11263-024-02042-6. URL https://link.springer.com/article/10.1007/s11263-024-02042-6.

Qiayuan Liao, Takara E. Truong, Xiaoyu Huang, Yuman Gao, Guy Tevet, Koushil Sreenath, and C. Karen Liu. BeyondMimic: From motion tracking to versatile humanoid control via guided diffusion. *arXiv preprint arXiv:2508.08241*, 2025. URL https://arxiv.org/abs/2508.08241.

Jingzhong Lin, Xinru Li, Yuanyuan Qi, Bohao Zhang, Wenxiang Liu, Kecheng Tang, Wenxuan Huang, Xiangfeng Xu, Bangyan Li, Changbo Wang, and Gaoqi He. ReactDance: Hierarchical representation for high-fidelity and coherent long-form reactive dance generation. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=FvMyAMbbX0.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Zuhong Liu, Junhao Ge, Minhao Xiong, Jiahao Gu, Bowei Tang, Wei Jing, and Siheng Chen. It takes two: Learning interactive whole-body control between humanoid robots. *arXiv preprint arXiv:2510.10206*, 2025. URL https://arxiv.org/abs/2510.10206.

Zhengyi Luo, Jinkun Cao, Alexander Winkler, Kris Kitani, and Weipeng Xu. Perpetual humanoid control for real-time simulated avatars. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10861–10870, 2023. doi: 10.1109/ICCV51070.2023.01000. URL https://doi.org/10.1109/ICCV51070.2023.01000.

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*, 37(4):143:1–143:14, 2018. doi: 10.1145/3197517.3201311.

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*, 40 (4):144:1–144:20, 2021. doi: 10.1145/3450626.3459670. URL https://xbpeng.github.io/projects/AMP/.

Zekun Qi, Xuchuan Chen, Jilong Wang, Chenghuai Lin, Yunrui Lian, Wenyao Zhang, Xinqiang Yu, He Wang, and Li Yi. Humanoid generative pre-training for zero-shot motion tracking. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20834–20844, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Qi_Humanoid_Generative_Pre-Training_for_Zero-Shot_Motion_Tracking_CVPR_2026_paper.html.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017. URL https://arxiv.org/abs/1707.06347.

Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, 2009. doi: 10.1016/j.ipm.2009.03.002.

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*, 43(6):209:1–209:21, 2024. doi: 10.1145/3687951. URL https://research.nvidia.com/labs/par/maskedmimic/.

Guy Tevet, Sigal Raab, Brian Gordon, Yonatan Shafir, Daniel Cohen-Or, and Amit H. Bermano. Human motion diffusion model. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=SJ1kSyO2jwu.

Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5026–5033, 2012. doi: 10.1109/IROS.2012.6386109.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

Yuxuan Wang, Haobin Jiang, Shiqing Yao, Ziluo Ding, and Zongqing Lu. End-to-end language-action model for humanoid whole body control. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 38399–38409, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Wang_End-to-End_Language-Action_Model_for_Humanoid_Whole_Body_Control_CVPR_2026_paper.html.

Liang Xu, Yizhou Zhou, Yichao Yan, Xin Jin, Wenhan Zhu, Fengyun Rao, Xiaokang Yang, and Wenjun Zeng. ReGenNet: Towards human action-reaction synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1759–1769, 2024. URL https://openaccess.thecvf.com/content/CVPR2024/html/Xu_ReGenNet_Towards_Human_Action-Reaction_Synthesis_CVPR_2024_paper.html.

Xiyan Xu, Sirui Xu, Yu-Xiong Wang, and Liang-Yan Gui. MoReact: Generating reactive motion from textual descriptions. *Transactions on Machine Learning Research*, 2025. URL https://openreview.net/forum?id=4zuT73heqm.

Haoru Xue, Tairan He, Zi Wang, Qingwei Ben, Wenli Xiao, Zhengyi Luo, Xingye Da, Fernando Castañeda, Guanya Shi, Shankar Sastry, Linxi Fan, and Yuke Zhu. Opening the sim-to-real door for humanoid pixel-to-action policy transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6642–6652, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Xue_Opening_the_Sim-to-Real_Door_for_Humanoid_Pixel-to-Action_Policy_Transfer_CVPR_2026_paper.html.

Lujie Yang, Xiaoyu Huang, Zhen Wu, Angjoo Kanazawa, Pieter Abbeel, Carmelo Sferrazza, C. Karen Liu, Rocky Duan, and Guanya Shi. OmniRetarget: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction. In *IEEE International Conference on Robotics and Automation*, June 2026. URL https://omniretarget.github.io/.

Yaoqin Ye, Yiteng Xu, Qin Sun, Xinge Zhu, Yujing Sun, and Yuexin Ma. ReMoGen: Real-time human interaction-to-reaction generation via modular learning from diverse data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16475–16485, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Ye_ReMoGen_Real-time_Human_Interaction-to-Reaction_Generation_via_Modular_Learning_from_Diverse_CVPR_2026_paper.html.

Yanjie Ze, Zixuan Chen, Joao Pedro Araujo, Zi-ang Cao, Xue Bin Peng, Jiajun Wu, and Karen Liu. TWIST: Teleoperated whole-body imitation system. In *Proceedings of the 9th Conference on Robot Learning*, volume 305 of *Proceedings of Machine Learning Research*, pp. 2143–2154. PMLR, 2025. URL https://proceedings.mlr.press/v305/ze25a.html.

Yunbo Zhang, Deepak Gopinath, Yuting Ye, Jessica Hodgins, Greg Turk, and Jungdam Won. Simulation and retargeting of complex multi-character interactions. In *ACM SIGGRAPH 2023 Conference Proceedings*, 2023. doi: 10.1145/3588432.3591491. URL https://arxiv.org/abs/2305.20041.

Zewei Zhang, Kehan Wen, Michael Xu, Junzhe He, Chenhao Li, Takahiro Miki, Clemens Schwarke, Chong Zhang, Xue Bin Peng, and Marco Hutter. Learning whole-body humanoid locomotion via motion generation and motion tracking. *arXiv preprint arXiv:2604.17335*, 2026. URL https://arxiv.org/abs/2604.17335.

A Overall Procedure

The generator is shared across motion pairs; control policies are trained for each resulting reference pair.

Algorithm 1 DyadLink: generation to compliant execution

Require: Paired motion data, conditions, and target embodiments

Generator training

1: **for** each training minibatch **do**
2: Predict $(\mathcal{G}^{\mathrm{plan}}, \mathbf{z}^{\mathcal{G}})$ from the Leader and conditions; form $\mathbf{z}^{\mathrm{ctx}}$.
3: Sample the flow-training pair in Eq. (9).
4: Decode the Follower endpoint estimate with the frozen VAE.
5: Apply bounded refinement; reconstruct $\widehat{\mathcal{G}}$.
6: Optimize Eq. (3).
7: **end for**

Reference construction

8: Predict the SRG plan and sample a Follower for the supplied Leader.
9: Retarget using Algorithm 2 for robot pairs, or the fixed-Leader variant in Appendix D.
10: Form $\mathcal{G}^{\mathrm{ref}}$ from retargeted spatial attributes and planned contact probabilities.

Policy learning and execution

11: Train and freeze the robot Leader policy, or replay a prescribed human Leader.
12: Train the Follower with PPO, the reward in Section 3.4, and compliant feedback.
13: Deploy the trained policies with the associated reference and compliant module.

B SRG Construction

Each role has 22 semantic anchors. Within-agent skeleton edges encode anatomy; cross-agent edges are directed from Leader to Follower. Table 8 specifies the relation sets.

For proxy radii r_i^{L} and r_j^{F},

$$d_{e,t}^{\mathrm{surf}} = \|\mathbf{x}_{j,t}^{\mathrm{F}} - \mathbf{x}_{i,t}^{\mathrm{L}}\|_2 - r_i^{\mathrm{L}} - r_j^{\mathrm{F}},$$
$$p_{e,t}^{\mathrm{geom}} = \sigma\left(\frac{d_e^{\mathrm{prob}} - d_{e,t}^{\mathrm{surf}}}{\tau_e}\right). \tag{8}$$

Table 8: Cross-agent SRG relations. $\mathcal{E}_{\text{con}} \subset \mathcal{E}_{\text{geo}}$.

Set	Count	Semantic pairs
$\mathcal{E}_{\text{geo}}$	17	1 pelvis–pelvis, 12 contact candidates, and 4 Leader-wrist–Follower-hip pairs.
$\mathcal{E}_{\text{con}}$	12	4 hand–hand and 8 hand–shoulder combinations, with either role supplying the hand.
$\mathcal{E}_{\text{avoid}}$	6	4 head/torso combinations and 2 same-side knee pairs.

The probability threshold d_e^{prob} and temperature τ_e set the contact transition; $\kappa_{e,t} = d_{e,t}^{\text{surf}}$ on avoidance pairs, with negative values indicating proxy overlap.

Contact labels use distance–velocity hysteresis: activate below the distance and relative-speed thresholds, release above the release-distance threshold, and discard short runs. Relative speed uses world-space endpoint differences; a detached distance–speed confidence weights supervision. During each active interval, the onset vector replaces the contact vector in $\mathbf{g}^{\text{sup}}$ before temporal differencing. Δ is a frame difference, not a velocity. Offline construction copies the first available difference to frame one; runtime differences reset at episode boundaries or reference-time jumps.

C Generation Details

The graph–temporal predictor uses fixed semantic connections, gated adaptive messages, and temporal convolutions. Shared node features feed the relation heads and the pooled projection $\mathbf{z}^{\mathcal{G}}$. Internal messages are bidirectional; exported relations remain Leader-to-Follower. Heads predict edge vectors, contact probabilities, nonnegative clearances, and normalized headings.

Flow training and sampling. Following UMF (Huang & Celiktutan, 2026), for the ground-truth Follower latent $\mathbf{z}^{\text{F}}$, $\tau \sim \mathcal{U}(0,1)$, and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$,

$$\begin{aligned} \mathbf{z}_\tau &= \left[(1-\tau)\boldsymbol{\epsilon} + \tau \mathbf{z}^{\text{ctx}},\ (1-\tau)\mathbf{z}^{\text{ctx}} + \tau \mathbf{z}^{\text{F}}\right], \\ \mathbf{u}_\tau &= \left[\mathbf{z}^{\text{ctx}} - \boldsymbol{\epsilon},\ \mathbf{z}^{\text{F}} - \mathbf{z}^{\text{ctx}}\right], \\ \mathcal{L}_{\text{FM}} &= \mathbb{E}\left[\frac{\|\mathbf{v}_\theta(\mathbf{z}_\tau, \tau, \mathbf{c}) - \mathbf{u}_\tau\|_2^2}{2d}\right], \end{aligned} \tag{9}$$

where each block has d scalar entries. Training decodes $\widehat{\mathbf{z}}^{\text{F}} = \mathbf{z}_\tau^{(2)} + (1-\tau)\mathbf{v}_\theta^{(2)}$; inference integrates from $[\boldsymbol{\epsilon}, \mathbf{z}^{\text{ctx}}]$ with RK4. Root placement and bounded contact refinement follow the frozen decoder; the refined pair defines $\widehat{\mathcal{G}}$.

Loss settings. Each graph loss is $\mathcal{D}_q = \sum_{a \in \mathcal{A}_q} \lambda_{q,a} \ell_{q,a}$, $q \in \{\text{plan}, \text{state}, \text{cons}\}$. Continuous attributes use

$$\begin{aligned} \mathcal{H}(\mathbf{U}, \mathbf{V}; \mathbf{m}) &= \frac{\sum_i m_i\, \ell_\beta(U_i - V_i)}{\max(1, \sum_i m_i)}, \\ \ell_\beta(x) &= \begin{cases} x^2/(2\beta), & |x| \le \beta, \\ |x| - \beta/2, & |x| > \beta, \end{cases} \qquad \beta = 0.1. \end{aligned} \tag{10}$$

Masks exclude padding and are broadcast over coordinates in both sums.

$\mathcal{L}_{\text{aux}}$ collects motion/root reconstruction, temporal/support terms, and enabled distillation/short-rollout losses. Graph and auxiliary terms are warmed up with configured weights; the VAE stays frozen. Consistency gradients pass only through the realized-motion branch.

D Retargeting Details

Source and planned relations are calibrated to the same Leader frame and interaction scale, then blended:

$$\overline{\mathbf{g}}_{e,t} = (1 - \beta_e)\mathbf{g}_{e,t}^{\text{base}} + \beta_e \mathbf{g}_{e,t}^{\text{plan}}, \qquad 0 \le \beta_e \le 1. \tag{11}$$

Table 9: Graph-supervision settings. $\mathcal{H}$ denotes masked Smooth-L_1 regression.

Term	Target and reduction
Geometry	Plan/state regress $\mathbf{g}^{\mathrm{sup}}$ and its first/second differences with $\mathcal{H}$. Root/contact edges have weight 1; other geometric edges, 0.5.
Heading	Plan/state regress the ground-truth heading vector.
Active contacts	Onset-anchor regression and maintenance loss on consecutive contact-vector changes.
Plan contacts	Confidence-weighted focal BCE (exponent 2); positive-class weight is the confidence-weighted negative/positive count ratio clipped to $[1, 20]$. $\mathcal{H}$ compares probability changes with label changes.
State contacts	Geometric contact scores are compared with binary labels.
Clearance	Plan targets $[\kappa^*]_+$; state penalizes $[m_e - \widehat{\kappa}_{e,t}]_+$ against zero, with edge-specific margin m_e.
Consistency	$\mathcal{H}$ compares realized geometry, temporal differences, heading, contacts, active-contact vectors, and clearance with detached planned values.

$\mathbf{g}^{\mathrm{base}}$ is the measured source relation, except that touch tasks may prescribe a contact distance along the source direction.

Robot-pair fitting uses a coarse pose/root/heading/contact stage; the fine stage adds non-contact relations. The geometric Jacobian includes Leader-frame variation when the Leader is optimized. Contact weights increase with planned probability and ramp at onset/release. Contacts are soft objectives; optional bounds for sustained high-probability contacts can be relaxed if infeasible. Fixed linearizations, nonnegative weights, and convex quadratic regularization yield a convex local QP, not a convex overall problem.

With a prescribed human Leader, only the Follower is optimized using root, heading, and contact tasks; the coupled shell adapter and fine-stage non-contact tasks are disabled. Reference contact vectors follow the retargeted motion, not fixed onset anchors.

Algorithm 2 SRG-guided target adaptation and retargeting for robot pairs

Require: Source motion pair, SRG plan, and two calibrated target robots
for each frame **do**
  Blend source and planned vectors using Eq. (11).
  For contact/cross-body tasks, use the blended direction and the original source-edge surface gap.
  Scale the nonnegative gap, cap it for touch contacts, and add the target shell or contact-patch separation.
  Adapt root spacing using horizontal clearance; temporally smooth targets.
  Correct heading by the wrapped planned–source difference, preserving source-to-robot calibration.
  for each fitting stage and local iteration **do**
    Stack target-minus-current geometry and rotation-vector heading residuals into $\mathbf{r}_{\mathrm{SRG}}$.
    Fix targets, Jacobians, and nonnegative weights; solve Eq. (4).
    Apply enabled collision checks; backtrack or re-solve if needed.
    Integrate the accepted tangent step on the configuration manifold.
  end for
end for
Build $\mathcal{G}^{\mathrm{ref}}$ from retargeted geometry, heading, and clearance plus planned contact probabilities.
Recompute temporal differences from the time-varying retargeted vectors.

E Control Details

E.1 Reward Definitions

Geometry weights in Eq. (6) are 2 for the root, $1 + b\, p_{e,t}^{\mathrm{plan}}$ for contact candidates, and 1 otherwise ($b \geq 0$). The remaining discrepancies are

$$\begin{aligned} D_{\Delta,t} &= \frac{1}{|\mathcal{E}_{\mathrm{geo}}|} \sum_{e \in \mathcal{E}_{\mathrm{geo}}} \|\Delta \mathbf{g}_{e,t}^{\mathrm{exec}} - \Delta \mathbf{g}_{e,t}^{\mathrm{ref}}\|_2^2, \\ D_{p,t} &= \frac{1}{|\mathcal{E}_{\mathrm{con}}|} \sum_{e \in \mathcal{E}_{\mathrm{con}}} (p_{e,t}^{\mathrm{exec}} - p_{e,t}^{\mathrm{plan}})^2, \qquad \sigma_p = 0.25. \end{aligned} \tag{12}$$

Force-enabled edges use the score $1 - \exp(-\|\mathbf{f}_{e,t}\|_2 / F_0)$, with configured contact-point proximity weighting; other edges use Eq. (8).

For $\epsilon_{h,t} = \mathrm{wrap}(\vartheta_t^{\mathrm{exec}} - \vartheta_t^{\mathrm{ref}})$,

$$\rho_h(\epsilon_{h,t}) = \alpha_h \frac{1 + \cos \epsilon_{h,t}}{2} + (1 - \alpha_h) \left(1 + \frac{\ell_{\delta_h}(\epsilon_{h,t})}{\delta_h}\right)^{-1}, \tag{13}$$

where $0 \leq \alpha_h \leq 1$, $\delta_h > 0$, and ℓ_{δ_h} follows Eq. (10). Reward weights absorb the control interval. The interaction curriculum scales SRG rewards and the auxiliary pair-tracking/contact-comfort terms in r^{base}, but not reference tracking or regularization. Contact comfort favors moderate forces at planned contacts and low forces elsewhere; it is not a hard safety constraint.

E.2 Compliant Execution

Apply a deadband and norm clipping to measured hand forces to obtain $\bar{\mathbf{f}}_a$. For incident contact edges $\mathcal{E}_a$,

$$\eta_a = \eta_{\min} + (1 - \eta_{\min}) \max_{e \in \mathcal{E}_a} p_e^{\mathrm{plan}}. \tag{14}$$

Raw force above the override threshold sets $\eta_a = 1$. The joint offset $\mathbf{d}$, with internal velocity $\mathbf{v} = \dot{\mathbf{d}}$, modifies the Actor's command:

$$\begin{aligned} \ddot{\mathbf{d}} &= \alpha (\mathbf{M} + \varepsilon_M \mathbf{I})^{-1} \sum_a \mathbf{J}_a^\top (\eta_a \bar{\mathbf{f}}_a) - \omega^2 \mathbf{d} - 2\zeta\omega \mathbf{v}, \\ \mathbf{q}^{\mathrm{cmd}} &= \mathrm{clip}_{[\mathbf{q}_{\min}, \mathbf{q}_{\max}]} \left(\mathbf{q}^{\mathrm{def}} + \mathbf{s} \odot \mathbf{a} + \mathbf{d}\right). \end{aligned} \tag{15}$$

$\mathbf{M}$ is the controlled-joint mass submatrix, $\mathbf{J}_a$ the hand-point Jacobian, and $\mathbf{a}$ and $\mathbf{s}$ the Actor output and action scale. $\varepsilon_M > 0$ regularizes the solve; α, ω, and ζ set response gain, angular recovery frequency, and damping. Semi-implicit integration updates velocity before position. Both states are clipped, restricted to compliant joints (waist/arms for G1), and reset per episode. The external module operates during training and deployment; a low-level PD controller executes $\mathbf{q}^{\mathrm{cmd}}$.

F Additional Experimental Details

F.1 Data and Comparison Protocol

Table 10: Experimental data and comparison protocol.

Item	Setting
Datasets and conditions	InterHuman: Leader motion and text. CoDance (DD100 + CoMPAS3D): additional music.
Motion rate / hardware	30 Hz / one NVIDIA A100 GPU.
Retargeting/execution comparison	15 InterHuman motion pairs and 8 CoDance motion pairs.
Upstream references	Retargeting methods receive the same DyadLink-generated pairs; execution uses each method's own retargeted references.
RL ablation	17 InterHuman clips with identical retargeted references across variants; mean $\pm$ sample standard deviation across clips.

F.2 Relational Geometry and Tracking Metrics

PRPE/DRGE use corresponding human joints and robot link proxies in the Leader frame (Eq. (1)). Embodiment scales are fixed, not fitted to evaluated outputs. Both retargeting and execution are evaluated against the same source pair supplied to retargeting.

DRGE includes 1 pelvis–pelvis, 4 hand–hand, and 8 hand–shoulder relations, weighted 2, 1.5, and 1, respectively. Heading, contact probability, and clearance are not included.

Global position error compares executed body positions with the retargeted reference; local error first subtracts each motion's root. Rotation error is the $\mathrm{SO}(3)$ geodesic distance in degrees. These measure execution tracking, not human-to-robot pose fitting. Pair penetration is the fraction of frames with inter-agent overlap; maximum penetration is the largest depth per sequence, computed from collision geometry (Todorov et al., 2012) and reported in centimeters in Table 3.

F.3 Contact and Generation-Specific Metrics

Contact agreement. Recall/F1 compare realized and source contact states, not predicted SRG probabilities. True/false positives and false negatives are pooled over valid frames and semantic channels (Sokolova & Lapalme, 2009): across generated sequences for generation, within each clip before clip-level aggregation for execution.

Contact timing (seconds). Events are maximal active intervals per semantic channel. Generation matches each reference onset/offset to the nearest predicted event of the same type/channel; a missing prediction incurs the valid clip duration. Average absolute timestamp errors over reference onsets/offsets. Matching is not one-to-one; extra predictions affect F1 but incur no separate timing penalty.

Execution uses Hungarian assignment (Kuhn, 1955) on absolute timestamp differences, separately for onsets/offsets within each channel. Each unmatched timestamp costs 1 s; thus errors above 2 s favor leaving both timestamps unmatched. Average the resulting onset/offset errors over channels.

Hand and root errors. Hand MAE averages absolute inter-hand anchor-distance errors over four hand–hand combinations, using valid frames with an active reference hand contact. Root RMSE averages squared Leader-relative Follower-position errors over valid frames and three coordinates before taking the square root; PRPE instead averages Euclidean distances.

F.4 Ablation Settings

Table 11: Ablation changes relative to Full. Geometric relations between candidate-contact anchors remain in *w/o contact guidance*.

Stage	w/o SRG	w/o contact guidance
Generation	Remove SRG prediction/conditioning, graph supervision, and relational refinement; retain the backbone.	Remove contact-state/transition/maintenance supervision and contact-specific alignment; retain geometric relations.
Retargeting	Remove interaction-specific relational objectives and SRG weighting.	Disable semantic contact objectives; retain root/geometric objectives, collision handling, and temporal optimization.
RL	Remove SRG-related rewards.	Remove contact-related rewards; retain other relational rewards.

Retargeting-ablation contact metrics (Table 6) are measured after executing the resulting references. RL variants alter rewards only, not policy inputs or the compliant module, and use the metrics in Sections F.2–F.3. These ablations do not separately isolate Leader–Follower asymmetry, compliance, or individual generator components.

F.5 Qualitative Multimodal Examples

Figure 4 illustrates text and music conditioning in our Leader-conditioned framework, alongside human-space motion, retargeted references, and simulated execution. Overlaid poses indicate temporal progression.

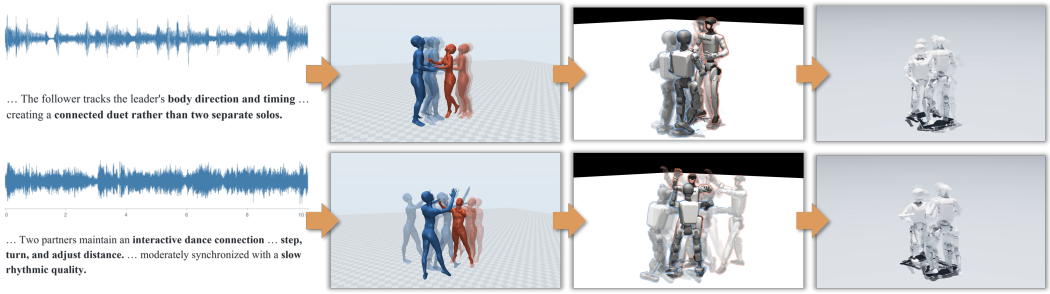

Figure 4: **Multimodal conditioning for dyadic interaction.** Text and music inputs are shown alongside stage-wise motion examples in human and robot embodiments.

F.6 Real-Robot Deployment

Platforms and communication. The hardware setup comprises Unitree G1-edu and Booster T1 platforms (Table 12). The deployment computer runs Ubuntu 22.04.5 LTS and is equipped with an NVIDIA RTX 4090D GPU. It connects to the robots over wired Ethernet, using the corresponding vendor's Python SDK and DDS communication middleware. The deployment program receives IMU measurements, joint positions, and joint velocities through the low-level state interface, constructs the policy observation, and converts the policy output into joint commands for the low-level controller. On G1, state feedback and commands use the `rt/lowstate` and `rt/lowcmd` topics. On T1, the SDK provides low-level state subscribers and command publishers, while `B1LocoClient` handles control-mode transitions.

Table 12: Hardware and software configurations for real-robot deployment. Platform degrees of freedom describe the robot, not the policy action dimension.

Configuration	Unitree G1-edu	Booster T1
Platform DoF	29	23
Python version	3.8	3.10.21
Robot interface	Unitree SDK2 Python	Booster Robotics SDK Python
Communication	CycloneDDS	DDS

Policy execution and reference timing. G1 policy inference and command updates run at a configured frequency of 50 Hz (20 ms per update), with 29-dimensional policy outputs. Its deployment references are also sampled at 50 Hz. For T1, the low-level state/command loop is configured at 500 Hz (2 ms per update). This low-level rate does not specify the policy-inference frequency, which is determined separately by the policy timing configuration. Reference progression begins after policy activation and follows the policy's time base.

Initialization and operational safeguards. The deployment procedure separates initialization from policy takeover. G1 smoothly transitions to the initial reference posture and waits for an explicit start command. The T1 takeover procedure first enters the preparation mode and establishes a hold target from the current joint measurements before switching to custom control. Pre-run checks cover communication and state validity, model/reference compatibility, observation and action dimensions, and the joint-order and scaling conventions used during training. Runtime guards monitor state freshness, IMU attitude, joint positions and velocities, and platform-specific command

or tracking limits. The exit procedure uses damping mode; on T1, a guard trigger stops reference progression before safe exit.

F.7 Qualitative Execution and Hardware Scope

Figure 5 supplements the execution results in Table 4. Cross-embodiment examples (Figure 3) keep a G1 Leader and vary the Follower, with policies trained per reference pair; they do not demonstrate zero-shot transfer of a universal policy. Policies for the human–robot hardware demonstrations are trained using G1 for both roles as a shared morphological and kinematic proxy.

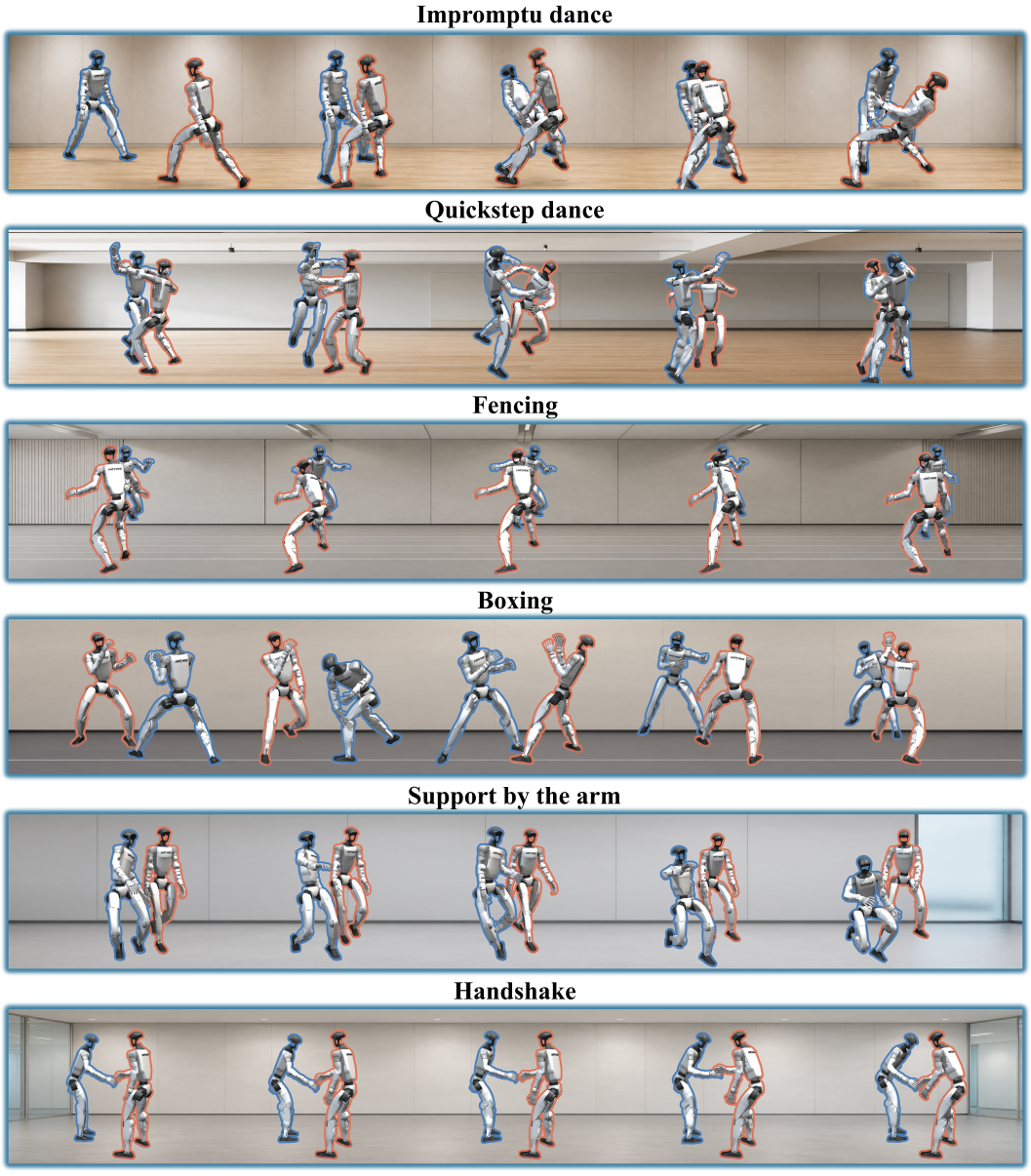

Figure 5: **Qualitative interaction.** Each row progresses from left to right; blue/orange denote the Leader/Follower.