

CoDance: Learning Reactive and Compliant Human-Humanoid Interaction from Video

Zhuoqun Chen¹, Shucheng Jia¹, and Boyuan Chen¹

<https://generalroboticslab.com/CoDance>

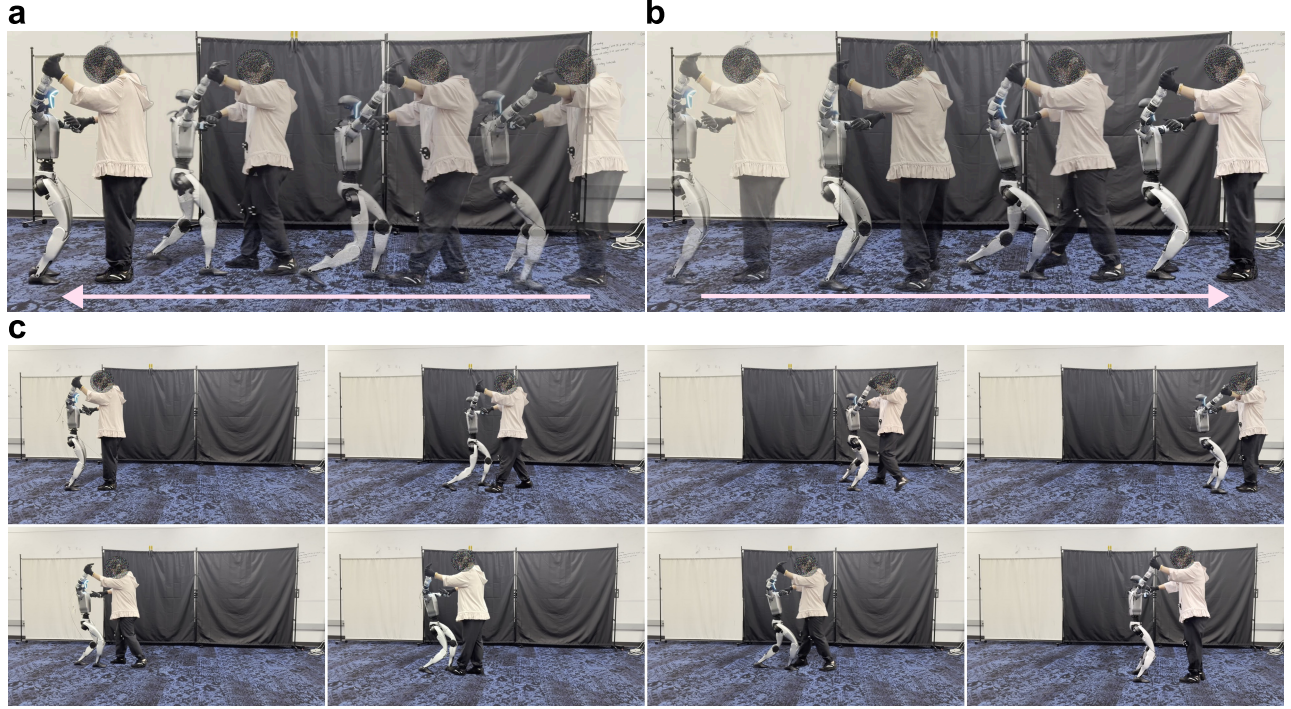


Fig. 1: **Reactive and compliant human-humanoid dancing with CODANCE.** Learned from a single video of two human dancers, the humanoid coordinates its locomotion with a human partner while maintaining compliant two-hand contact, and adapts its upper-body motion to forces transmitted through the hands. It steps backward as the partner advances (a) and forward as the partner retreats (b), and it repeats the transition between the two directions over one continuous dance (c).

Abstract— Partnered human-humanoid interaction couples locomotion with continuous physical contact. A humanoid needs to coordinate with a person’s motion while responding to interaction forces and maintaining stable and natural movement. We present CODANCE, a framework for learning reactive and compliant human-humanoid interaction from video. We study partnered dancing as a challenging instantiation, where a humanoid coordinates its footsteps with a moving partner and maintains continuous two-hand contact. Given a single video of two human dancers, CODANCE retargets their motions into a robot reference and a moving partner. We introduce a multi-link compliance augmentation that transforms the kinematic demonstration into force-aware training data by adapting the robot reference under structured forces at both hands. Policies trained on this data follow the observed partner while preserving the demonstrated locomotion style and responding compliantly to physical interaction. In simulation, the policies adapt their footsteps to changes in the partner and reproduce approximately 80% of the wrist displacement encoded by the augmented demonstrations. On a physical humanoid, CODANCE enables

sustained two-hand dancing with a human partner including repeated transitions between forward and backward motions.

I. INTRODUCTION

Humanoid robots are increasingly capable of reproducing dynamic and expressive human motion through advances in motion imitation, retargeting, and reinforcement learning [1]–[3]. However, most learned humanoid behaviors remain focused on the robot in isolation, tracking a motion reference or following a locomotion command while treating external contact primarily as a disturbance. Physical interaction with a person poses a unique challenge. When two partners walk, carry, guide, or dance together, their motions are continuously coupled through both observation and contact. The robot needs to adapt its motion to the partner and accommodate interaction forces while preserving balance and the natural movement.

We study this problem through partnered dancing, a demanding form of sustained physical human-humanoid

¹All authors are with Duke University.

interaction. The robot coordinates its footsteps with a moving partner while preserving the demonstrated locomotion style and maintaining continuous two-hand contact. Differences in timing, position, and body geometry therefore need to be accommodated through compliant motion. This creates a direct conflict with conventional motion imitation, where deviations from the reference are treated as tracking errors. Recent compliant tracking methods [4] address physical perturbations during motion imitation, but partnered interaction also requires the robot to adapt its locomotion to another person while continuously responding to physical contact.

We introduce CoDance, a framework for learning reactive and compliant human-humanoid interaction from a single video of two people dancing (Fig. 2). CoDance reconstructs and separately retargets the two dancers, using one as the robot reference and the other as a moving partner during training. The policy observes sparse partner keypoints and coordinates its footsteps directly with the partner’s motion, without relying on a prescribed velocity or trajectory. At deployment, the simulated partner is replaced by a real person whose corresponding keypoints are measured from a small set of motion-capture markers.

A central challenge is that video provides motion without the interaction forces. We address this with a multi-link compliance augmentation that transforms the kinematic demonstration into force-aware training data. Structured interaction events apply forces at one or both hands, with coupled displacements and force directions informed by the partner’s motion. Each event adapts the robot reference according to the sampled force and commanded stiffness, encoding how the robot should respond to physical interaction while preserving the demonstrated motion. Temporal reversal, force-free samples, and clip chaining further expand a short demonstration into long-horizon interaction trajectories. Our final data can support different locomotion objectives while sharing the same partner-following and compliance objectives.

We evaluate each component of CoDance in simulation and deploy the learned policy on a physical humanoid. Changing or removing the partner observation shows that the policy uses the observed human motion to adapt its footsteps. Under interaction forces, the learned policy reproduces approximately 80% of the wrist displacement encoded by the augmented demonstrations, compared with about 20% for a stiff tracking baseline. The demonstrated locomotion style can be learned through either adversarial imitation or reference tracking, while removing the locomotion objective substantially degrades the resulting motion. Clip chaining further enables transitions between forward and backward motions. We show that CoDance supports a humanoid for continuous two-hand dancing with a human partner across repeated forward and backward motions. Our contributions are:

- We introduce CoDance, a framework for learning partner-conditioned locomotion and compliant physical interaction from a single video of two human dancers.
- We develop a partner-informed multi-link compliance augmentation that transforms kinematic demonstrations

into force-aware training data for continuous two-hand interaction.

- We demonstrate reactive partner following, compliant interaction, and learned locomotion style in simulation, and sustained partnered dancing on a physical humanoid.

II. RELATED WORK

A. Partnered Dancing and Human-Humanoid Interaction

Physical robot dancing has been studied for more than two decades. Ms DanceR [5] follows a human partner through force sensing on an omnidirectional base. Model-based humanoids have performed box-step dances by adapting their steps to contact forces [6] or leading through compliant arm motion [7]. Recent learning-based methods generate dance motion on legged robots [8], interactive motions between two humanoids [9], and hand-holding dances under velocity commands [1]. Learned humanoids have also coordinated with people for object carrying [10], [11] and acquired interaction skills from human videos or paired human demonstrations [12], [13]. CoDance studies a different setting where a humanoid learns the locomotion style from paired human video, follows an observed human partner, and maintains compliant two-hand contact.

B. Whole-Body Motion Imitation for Humanoids

Physics-based character control learns natural motion through reference tracking [14] or adversarial objectives [3], [15]. These approaches have transferred to physical humanoids for expressive whole-body control [1], [2], teleoperation [16], [17], and diverse motion tracking [18]. These methods primarily consider single-agent motion, where tracking objectives keep the robot close to its reference. Physical interaction introduces contact forces that can require motion away from this reference. In CoDance, the locomotion objective preserves the demonstrated lower-body motion while the upper body adapts to interaction with the partner.

C. Compliant Control and Motion Tracking

Impedance and admittance control regulate robot motion under physical contact [19]. Learned controllers extend these ideas to legged robots through impedance-reference tracking [20], joint position-force control [21], and compliant task-space control for humanoid loco-manipulation [22]. For humanoid motion tracking, SoftMimic [4] augments references with inverse-kinematics solutions under sampled forces, while GentleHumanoid [23] learns upper-body compliance using guiding forces and human motion. Our augmentation takes one step further to target sustained partnered locomotion. Force events act on one or both wrists with coupled displacements and directions informed by the partner’s motion, producing force-aware references for compliant two-hand interaction during partner following.

III. METHOD

CoDance consists of three stages (Fig. 2). The data pipeline turns a paired-dancing video into force-aware training data (Section III-A). A policy then learns partner following and

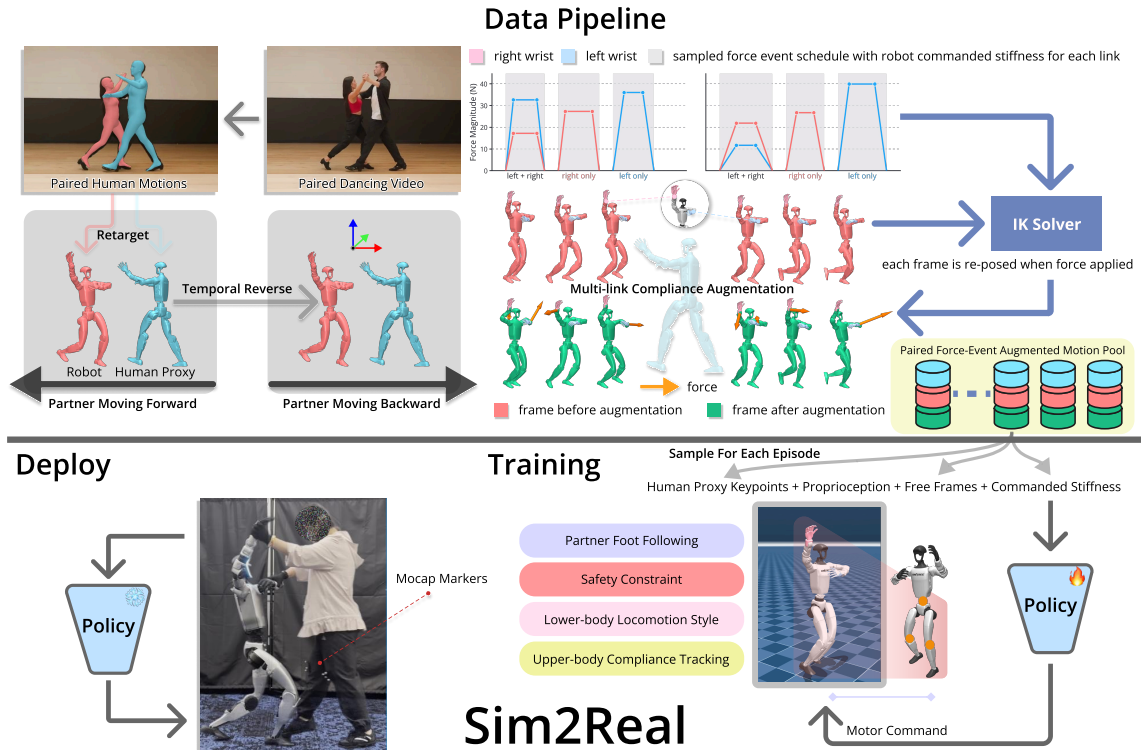


Fig. 2: Overview of CODANCE. A paired-dancing video is converted into force-aware training data through motion retargeting and multi-link compliance augmentation. The policy learns to follow the observed partner, maintain the demonstrated locomotion style, and remain compliant at both hands. At deployment, the simulated partner is replaced by a person. Training is illustrated with the decoupled policy.

two-hand compliance together with a locomotion objective (Section III-B). At deployment, a person replaces the simulated partner and the policy runs directly on a physical humanoid (Section III-C).

A. Data Pipeline

From Video to a Paired Clip. PromptHMR [24] recovers the SMPL-X [25] motions of both dancers, and GMR [26] retargets each motion separately to the humanoid. One dancer provides the robot reference and the other provides the partner motion. Since separate retargeting does not preserve their relative placement, we position the partner using a fixed following offset. During training, a second humanoid kinematically replays this motion as a human proxy. We also reverse the 9.7s paired clip in time to cover both directions of travel. The video does not carry contact forces, so the proxy’s upper body does not collide with the robot, and the hand forces come from the force augmentation below. Its lower body retains collision so that stepping into the partner can be penalized.

Multi-Link Compliance Augmentation. Our augmentation builds on the quasi-static compliance solve of SoftMimic [4]. Force events are distributed throughout the clip, with each event following a ramp, hold, and release profile (Fig. 2). For a force f and robot stiffness K , the contacted link is displaced by f/K . We sample K log-uniformly from 10 N m^{-1} to 1000 N m^{-1} . A whole-body inverse-kinematics then adapts

the reference while holding the feet fixed, constraining the center of mass, and weakly tracking the remaining body links. If a frame violates feasibility or self-collision constraints, the event is recomputed with the wrist force directed away from the torso or reduced in magnitude. We call the resulting motion adapted reference and the original motion the free reference. We extend the sampled interaction beyond the single-link perturbations used in SoftMimic in three ways.

Coupled two-hand interaction. Instead of SoftMimic’s left or right wrist, each event samples an interaction intent that sets the contacted wrists and their relative displacements. The intents include a single wrist (`single`), common motion of both wrists (`along`), mirrored motion (`open`), opposing vertical motion (`twist`), and motion toward or away from each other (`oppose`). Each contacted wrist has an independently sampled stiffness, except under `oppose`, where both wrists share the same stiffness to produce equal and opposite sampled forces.

Partner-guided force direction. The partner’s smoothed horizontal velocity sets the base direction of each event. Most events push along this velocity, and the rest push along the robot’s facing axis against the direction of travel. When the partner is nearly still, the direction is drawn at random. Directional jitter expands each base direction into a local cone.

Balance-aware force scaling. A force applied at the hands

Algorithm 1 Multi-link compliance augmentation. Black: the SoftMimic [4] single-link procedure as we implement it; blue: ours.

Require: reference clip; contactable links (the two wrists); **partner clip**; stiffness range; force ceiling; **balance budget**; displacement scale

Ensure: adapted clip; force schedule

```

1:  $t \leftarrow 0$ ; events  $\leftarrow \emptyset$ 
2: while  $t < \text{clip length}$  do
3:    $t \leftarrow t + \mathcal{U}(0.5, 1.5)$  s; skip if a wrist is busy
4:   draw the ramp and hold durations
5:   draw an intent; it fixes which links are contacted (one
   or both wrists) and how their displacements relate
6:   draw the force direction from the partner's motion;
   random if the partner is still
7:   for each contacted link do
8:     draw  $K$  log-uniformly (shared by the pair under
     oppose) and  $\delta$  from the intent and the displacement scale
9:     shrink  $\delta$  until  $K\|\delta\| \leq \text{ceiling}$ ;  $f \leftarrow K\delta$ 
10:  end for
11:  scale the event once so its CoM shift stays within
   the budget
12:  record the event
13: end while
14: for each frame do
15:   contacted links: target = reference +  $f/K$ 
16:   solve whole-body IK: feet held, CoM task, every
   other body weakly held at its reference pose
17:   if the frame fails a feasibility or self-collision check
   then
18:     turn the wrist's force outward or scale the event
     down, then re-solve from the event's start
19:   end if
20: end for
21: return adapted clip, force schedule

```

requires a corresponding center-of-mass shift for balance. Before solving the event, we scale its force once so that the required shift stays within a prescribed budget throughout the event. This preserves the sampled ramp, hold, and release profile while keeping the adapted motion feasible.

Each augmented clip stores its adapted reference and a force schedule containing the force, commanded stiffness and timing of every event. Force-free clips use the free reference with a commanded stiffness resampled every 2 to 5 s. Our dataset contains 200 clips with force events and 200 force-free clips, evenly divided between forward and reversed motion. Algorithm 1 and Table I provide the full augmentation procedure and parameters.

B. Policy Training

CoDance is not tied to one specific policy design. Humanoid motion imitation takes its motion style either from an adversarial reward [3] or from reference tracking [14], [18], so we train one policy of each kind on the same data. Both policies share the observations, the environment spring, every

TABLE I: Augmentation parameters.

Parameter	Value
Robot stiffness K	log-uniform, 10 to 1000 N m^{-1}
Environment stiffness	same range, independent draw
Force ceiling	140 N
Displacement scale $d_{\max}$	0.15 m
per-event draw	$\mathcal{U}(0.03, d_{\max})$ m
Balance budget (CoM shift)	0.10 m
Intent probability, single and oppose	0.25 and 0.075
along, open, twist (each)	0.225
Vertical direction scale	0.5
Push probability p_{push}	0.85
Partner velocity smoothing window	0.4 s
Partner speed lower bound	0.1 m s $^{-1}$
Direction jitter	0.4
Facing-axis displacement cap	0.08 m
CoM task weights (XY, Z)	0.1, 10^{-6}
Loading rate limit	1.0 m s $^{-1}$
Ramp, hold (s)	$\mathcal{U}(0.2, 0.6)$, $\mathcal{U}(0.4, 0.8)$
Clips per direction	100 with force events, 100 force-free

reward term except the locomotion objective, and the training setup. We call the policy with the adversarial locomotion objective the decoupled policy, since only its upper body tracks the reference, and the policy with lower-body reference tracking the whole-body tracker.

Observations. The actor observes proprioception with a short history, namely the base angular velocity, the joint positions and velocities, the previous actions and the projected gravity. It also observes the free-reference anchor orientation and the joint states, which provide the phase of the source motion. Partner observations contain the pelvis and both knees in the robot pelvis frame over the last second. It also observes the commanded stiffness of each wrist. The actor does not observe the adapted reference or applied force. Its response to physical interaction therefore depends on the commanded stiffness and the resulting proprioceptive motion. The critic additionally observes base linear velocity, the adapted reference, applied and scheduled forces, environment stiffness, and the robot body poses. The policy does not take a velocity command. It follows the partner from the partner's keypoints.

Interaction Force. During a force event, the force on each contacted wrist is modeled as a spring toward a setpoint x_t^* ,

$$f_t = K_{\text{env}}(x_t^* - x_t), \quad (1)$$

where x_t is the wrist position and K_{env} is drawn log-uniformly for each contacted wrist, independently of the robot stiffness. The setpoint x_t^* is placed so that the spring delivers the scheduled force at the adapted reference. Deviations from the adapted reference will then change the interaction force. An analogous rotational spring acts on the wrist orientation with zero scheduled torque.

Rewards. The reward contains partner following, compliance, locomotion, and regularization terms,

$$r_t = r_t^{\text{partner}} + r_t^{\text{comp}} + r_t^{\text{style}} + r_t^{\text{reg}}, \quad (2)$$

where r_t^{style} is the locomotion objective. Tracking terms use exponential kernels $\exp(-e^2/\sigma^2)$ [14], [18].

Partner following. The robot follows the partner through corresponding foot targets. Each robot foot is rewarded for staying near the opposite-side partner foot with a fixed following offset,

$$r_t^{\text{foot}} = \sum_{s \in \{L, R\}} \exp\left(-\frac{\|x_t^s - (c_t^{\bar{s}} - o)\|^2}{\sigma_f^2}\right), \quad (3)$$

where x_t^s is robot foot s , $c_t^{\bar{s}}$ is the opposite-side partner foot, and o is a 0.5 m following offset. A feet air-time term rewards natural step duration. Contact with the partner’s lower body is penalized.

Compliance. The upper body tracks the adapted reference under the scheduled interaction forces. For link positions,

$$r_t^{\text{arm}} = \exp\left(-\frac{1}{|\mathcal{B}_{\text{up}}|} \sum_{b \in \mathcal{B}_{\text{up}}} \frac{\|p_t^b - \tilde{p}_t^b\|^2}{\sigma_u^2}\right), \quad (4)$$

where $\mathcal{B}_{\text{up}}$ contains the shoulder, elbow, and wrist links. The adapted reference $\tilde{p}_t^b$ is re-anchored to the ground-plane position and heading of the robot torso, making the tracking error invariant to global locomotion. Link orientations are treated similarly. Additional terms place higher tracking weight on the wrists and encourage the applied wrist forces and torques to match their scheduled values.

Locomotion objective. The decoupled policy uses an adversarial motion reward [3]. Its discriminator compares two-frame transitions from policy rollouts with transitions from the training references using lower-body and torso poses in the pelvis frame. Arm states are excluded so that physical interaction does not enter the locomotion-style objective. Expert transitions use adapted references for clips with force events and free references for force-free clips.

The whole-body tracker uses lower-body reference tracking in place of the adversarial reward, following the motion-tracking formulation of prior work [4]. It tracks lower-body poses relative to the torso, their global velocities, and the global torso pose. Both policies use force-aware reference state initialization [4], [14], including initialization within active force events.

Regularization. r^{reg} penalizes action rate, joint-limit violations, self-collisions, and the joint accelerations of the hips and knees.

Clip-Chaining and Training Details. A dance moves forward and backward in turn, so dancing longer than one clip requires the transition from a clip in one direction to a clip in the other. During training, clips are chained at their ends. When a clip runs out, a clip drawn at random from the training data starts from the partner’s current position, so an episode spans more than one clip. Episodes last 20 s, about two clips. We use the domain randomization of the BeyondMimic in mujoco [18], [27], including torso center-of-mass offsets, foot friction, and joint encoder bias. Random pushes are disabled because the scheduled interaction events already perturb the robot. Both policies are trained with PPO in MuJoCo through mujoco [27] using 16384 parallel environments at 50 Hz control rate. Actor and critic networks have three hidden layers with 512,

256 and 128 units. Training converges in approximately 37 h for the decoupled policy and 29 h for the whole-body tracker on one NVIDIA L40S.

C. Hardware Deployment

We deployed the converged decoupled policy on a Unitree G1 at 50 Hz. Joint position targets were sent through the Unitree SDK at 200 Hz. A Vicon system tracked the person’s pelvis and knees and the robot pelvis, providing the same partner keypoints used during training. Marker heights were calibrated once to match the simulated proxy. Both wrists used a commanded stiffness of 140 N m^{-1} . Physical forces came directly from the person’s hands during the two-hand hold.

IV. EXPERIMENTS

We evaluated four aspects of CoDance in simulation: reactive partner following, compliant interaction, locomotion style, and transitions across chained clips. Targeted ablations isolated the components responsible for each behavior. We then evaluated partnered dancing with a human on the physical robot.

A. Experimental Setup

All policies were trained on the same half of the augmented clips (seen). The remaining half of the clips (unseen) were generated independently by the same augmentation pipeline with newly sampled event timing, interaction intents, stiffness commands, and forces.

We evaluated the **Decoupled** and **Whole-body** policies from Section III-B and three baselines. **Whole-body w/o foot constraint** removed the foot-following reward in (3) and the foot-distance termination. **Stiff tracker** tracked the free reference under the same force events, without force matching or the tighter wrist-tracking terms. **Decoupled w/o adversarial term** removed the adversarial locomotion objective.

Simulation results were averaged over three evaluation seeds with 400 trials per seed. A trial terminates if the torso height drops below 0.3 m, its orientation deviates from the reference by more than 1.2 rad, or either foot moves more than 1.2 m from its partner-relative target. Hardware evaluation follows Section III-C.

B. Evaluation Metrics

- **Partner following.** Foot error is the ground-plane distance of each foot from its follow target averaged over both feet and over time before the first termination, or read at 1 s after the trial starts in Table II. Root distance is the ground-plane distance between the two pelvises, with a target of 0.5 m (Table III). Success rate is the fraction of trials that reach the end of the clip, or of a chain of clips, without a termination (Tables II and IV).
- **Compliance.** Wrist displacement and force error were read over the last 0.2 s of each force event’s hold, in the torso-anchored frame. The wrist displacement measures how far the wrist moves from the free reference along the applied

TABLE II: Evaluation results on different partner observation and foot constraint settings. Masking out the partner observation or reversing the partner leads to a large foot error and a large drop in success rate. Training without the foot constraint leads to a large foot error.

Policy	Partner	Foot Err at 1 s (m)	Success Rate
Decoupled	observed	0.057	99.6 %
	masked out	0.421	0.0 %
	reversed	0.267	28.6 %
Whole-body	observed	0.072	99.7 %
	masked out	0.430	0.0 %
	reversed	0.255	24.3 %
Whole-body w/o foot constraint	observed	0.179	99.6 %

force, for the policy and for the adapted reference. Mean displacement error is the absolute difference between the two wrist displacements, averaged over events (Fig. 3). Relative adapted ratio is the policy’s wrist displacement summed over events, divided by the same sum for the adapted reference (Fig. 3). Force error is the magnitude of the difference between the applied and the scheduled force, averaged over events (Table III). Upper-body tracking error is the torso-relative mean per-keybody position and rotation error of the upper body against the adapted reference (R-MPKPE, R-MPKRE), with the reference re-anchored to the robot’s torso at each step as in Equation (4) (Table III).

- **Locomotion style.** Lower-body tracking error is the R-MPKPE and R-MPKRE of the lower body against the adapted reference (Table III).
- **Long horizon.** We report over a chain of two clips (Section IV-F), the success rate and the foot error before the first termination (Table IV) .

C. Reactive Partner Following

We first tested whether the robot responded to the observed partner instead of simply replaying its reference motion. We evaluated two changes to the partner observation. In the masked condition, all partner keypoints are set to zero. In the reversed condition, the partner motion is temporally reversed while the robot reference remains unchanged. We also evaluated the whole-body policy trained without the foot-following constraint.

As shown in Table II, both policies completed nearly every trial with the original partner observation. Masking the partner observation caused all trials to fail, with foot error reaching approximately 0.42m after 1s. Since zero-valued observations were outside the training distribution, we additionally used the reversed condition to separate partner dependence from this distribution shift. Reversing the partner avoided this problem, because the policies observed the partner moving in both directions in training and only its mismatch with the robot’s reference was new. With this in-distribution but mismatched input, the foot error was about 0.26m after one second and about a quarter of the trials finished the clip. Without the foot constraint, the whole-



Fig. 3: Evaluation results on compliance interaction. Under force, both compliant policies move the wrist most of the distance encoded in the adapted reference, while the stiff tracker moves it far less. Error bars show the standard deviation over evaluation seeds.

body tracker finished nearly every trial but did not follow the partner, with 2.5 times the foot error of the constrained whole-body tracker after 1s. These results show that partner-conditioned locomotion depends on both the observed partner motion and the foot-following objective. The supplementary video further shows hardware behavior when the human partner pauses during the dance.

D. Compliance Interaction

We next evaluated whether the force-aware training data produced compliant wrist motion during partner-conditioned locomotion, on both seen and unseen clips. The baseline was the stiff tracker, which we compared with the adapted reference of the same events.

As shown in Fig. 3, the two compliant policies had similar displacement error, and their relative adapted ratios differed by 0.03. On seen clips, their wrists moved about 80 % as far as in the adapted reference, with a mean displacement error of 2.6 cm. The stiff tracker moved its wrist about 20 % as far, with more than twice the displacement error. On unseen clips, the relative adapted ratio was higher for every policy, including the stiff tracker. The displacement error of the two compliant policies grew by less than 1 cm. The stiff tracker remains the farthest from the adapted reference on both measures. These results show that training with the adapted references produced the compliant wrist motion encoded by the augmentation and transferred to unseen force-event schedules.

E. Locomotion Style

We tested whether the demonstrated locomotion style depended on the locomotion objective by comparing the decoupled policy with the whole-body tracker on the unseen clips and by removing the adversarial reward from the decoupled policy, and whether changing it affects following or compliance (Table III).

The rows in Table III differ most in the locomotion style columns. The whole-body tracker tracked the lower body most closely, as its objective required, and the decoupled

TABLE III: Evaluation results on different locomotion objective settings. Removing the adversarial term from the decoupled policy leads to large lower-body tracking errors at comparable foot error. The whole-body tracker leads to tracking errors and foot error similar to the decoupled policy. The standard deviation is below 0.003 in the tracking and following columns and 0.1 N in the force error. Bold marks the best per column, and for Root Distance the value closest to 0.5 m.

Methods	Locomotion Style				Compliance					Partner Following			
	Lower		Lower		Upper		Upper		Force Err (N) ↓	Foot Err (m) ↓		Root Distance (m)	
	R-MPKPE (m) ↓		R-MPKRE (rad) ↓		R-MPKPE (m) ↓		R-MPKRE (rad) ↓						
	w/ force	force-free	w/ force	force-free	w/ force	force-free	w/ force	force-free	w/ force	w/ force	force-free	w/ force	force-free
Decoupled	0.089	0.082	0.225	0.198	0.041	0.034	0.144	0.084	7.8	0.072	0.053	0.490	0.470
Decoupled w/o adversarial term	0.143	0.142	0.587	0.590	0.034	0.024	0.142	0.083	8.8	0.072	0.047	0.533	0.517
Whole-body	0.069	0.065	0.197	0.177	0.034	0.025	0.137	0.079	7.7	0.074	0.059	0.431	0.419

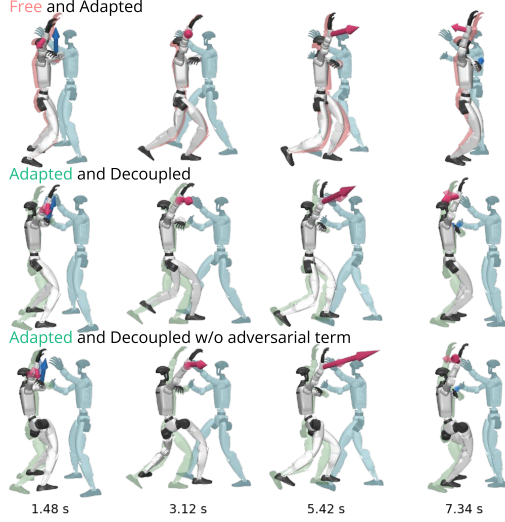


Fig. 4: Visualization of the free and adapted references (top) and of the decoupled policy with and without the adversarial term, at the middle of four force holds. The partner is in teal, and arrows mark the force at the left wrist (blue) and right wrist (pink), with the applied force solid and the scheduled force faded. Without the adversarial term, the pelvis turns away from the dancer’s heading, the waist and hips twist back toward the direction of travel, and the stance widens into a crouch. Top row also shows that separate retargeting does not preserve relative position in the paired motion clip.

policy was within 2 cm and 0.03 rad of it while keeping the root distance closer to the following distance. Without the adversarial reward, the decoupled policy still followed the partner with comparable foot error, but its lower-body rotation error was 2.6 times as large on clips with force events and 3.0 times as large on force-free clips. Fig. 4 visualized one seen clip with force events. Across the three rows in Table III, the foot error, force error and upper-body tracking error stayed within about 1 cm, 0.01 rad and 1 N of each other, while the root distance changed by up to 10 cm. Overall, the locomotion style came from the locomotion objective, while foot error and compliance change little when the objective was swapped or removed.

TABLE IV: Evaluation results on different clip-chaining settings. Training without clip-chaining leads to a large drop in success rate at comparable foot error.

Policy	Foot Err (m)	Success Rate
Decoupled trained w/ chaining	0.073	97.6 %
Decoupled trained w/o chaining	0.076	48.0 %
Whole-body trained w/ chaining	0.074	96.2 %

F. Benefit of Clip-Chaining

Clip-chaining trained the policy to continue from the end of one clip into another. We tested whether following holds across the transition between directions, on chains of two unseen clips with force events, one in each direction, played back to back in both orders. We evaluated the decoupled policy trained with and without clip-chaining, both with clip-chaining at evaluation, and the whole-body tracker also trained with clip-chaining (Table IV).

Both policies trained with clip-chaining succeeded in nearly every trial with similar foot error as on a single clip (Table III). The policy trained without clip-chaining succeeded in about half of the trials. Its foot error before termination matched the other two, so it did not lose the partner before failing. It failed at the clip switch while still following. Overall, following held over an unseen chain for both locomotion objectives when training with clip-chaining and this mechanism maintained partner following across changes in travel direction.

G. Hardware Results

We evaluated the Decoupled policy (Fig. 5 and the supplementary video) with a human partner using the setup in Section III-C. The robot maintains two-hand contact and follows the partner without falling in all three repeated runs. One forward run lasts 12.8 s, while two 38.9 s runs alternate forward and reversed clips until the reference ends. The root distance remains within 0.1 m of the 0.5 m target for 93%, 83%, and 73% of the three runs, with mean distances between 0.48 m and 0.50 m. The pair travels approximately 3.8 m during the run in Fig. 5. We also qualitatively evaluated a standing policy trained with the same augmentation pipeline without the partner guidance as shown in Fig. 6. It illustrated the compliant response at the hands. The supplementary video shows the complete hardware interactions.

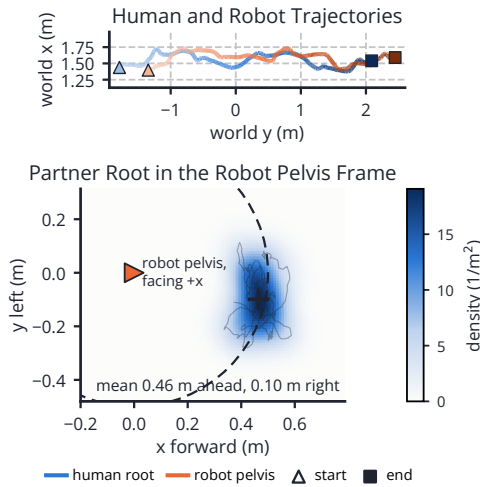


Fig. 5: Visualization of one real-robot dance with a human partner. The robot keeps close to the following distance for most of the dance. Top: the trajectories of the robot and the partner. Bottom: the partner’s root in the robot’s pelvis frame.

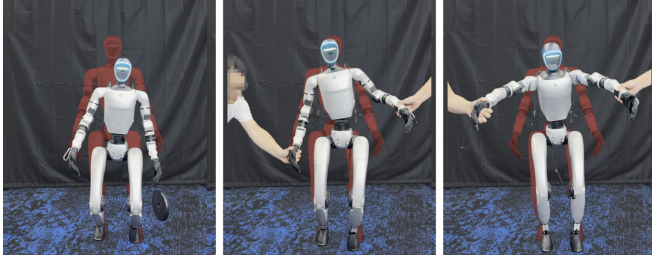


Fig. 6: Visualization of a standing policy under force on one or both hands. The red overlay is the reference standing pose. Left: one hand loaded by a hanging weight. Middle: the two hands pulled in different directions. Right: both hands pulled outward. The arms move with the pull and the feet stay still.

V. CONCLUSION, LIMITATIONS, AND FUTURE WORK

We presented CODANCE, a framework for learning reactive and compliant human-humanoid dancing from a single video of two human dancers. CODANCE converts kinematic demonstrations into force-aware training data and learns a policy that coordinates its footsteps with an observed partner while maintaining compliant two-hand interaction and the demonstrated locomotion style. In simulation, the learned policies reproduce most of the wrist displacement encoded by the augmented references and maintain partner following across chained clips. On a physical humanoid, CODANCE enables sustained two-hand dancing with a human partner across forward and backward motions.

Our study is currently limited to one video and one dance. The simulated partner replays a fixed motion, with interaction forces generated by the augmentation. The policies still depend on their reference clip, so only about a quarter of the trials finish a clip when the partner moves opposite to it. The foot-following target also keeps the robot at a fixed offset along a single direction from the partner, so dances

that change the partners’ relative placement are not covered. In future work, we will study rise and fall, turning and the box step, and extend the augmentation to any selected link.

ACKNOWLEDGMENT

This work is supported by DARPA FoundSci program under award HR00112490372, DARPA TIAMAT program under award HR00112490419, ARO under award W911NF2410405, ARL STRONG program under awards W911NF2320182, W911NF2220113, and W911NF242021. The authors thank Kuai Yu for dancing with the robot as the human partner in the hardware experiments. The authors also thank Li-Yu Lo, Yuhao Huang, and Yinsen Jia for insightful discussions, and Weizhe Ni, Xinyuan Luo, Jiaxun Liu, Boyuan Wang, and Zijiang Yang for their assistance with the hardware experiments.

REFERENCES

- [1] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang, “Expressive whole-body control for humanoid robots,” *arXiv preprint arXiv:2402.16796*, 2024. 1, 2
- [2] M. Ji, X. Peng, F. Liu, J. Li, G. Yang, X. Cheng, and X. Wang, “ExBody2: Advanced expressive humanoid whole-body control,” *arXiv preprint arXiv:2412.13196*, 2024. 1, 2
- [3] X. B. Peng, Z. Ma, P. Abbeel, S. Levine, and A. Kanazawa, “AMP: Adversarial motion priors for stylized physics-based character control,” *arXiv preprint arXiv:2104.02180*, 2021. 1, 2, 4, 5
- [4] G. Margolis, M. Wang, N. Fey, and P. Agrawal, “SoftMimic: Learning compliant whole-body control from examples,” *arXiv preprint arXiv:2510.17792*, 2025. 2, 3, 4, 5
- [5] K. Kosuge, T. Hayashi, Y. Hirata, and R. Tobiya, “Dance partner robot – Ms DanceR,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, 2003, pp. 3459–3464. 2
- [6] T. Kobayashi, E. Dean-Leon, J. R. Guadarrama-Olvera, F. Bergner, and G. Cheng, “Whole-body multicontact haptic human-humanoid interaction based on leader-follower switching: A robot dance of the ‘box step,’” *Advanced Intelligent Systems*, vol. 4, no. 2, p. 2100038, 2022. 2
- [7] M. Charbonneau, F. J. Andrade Chavez, and K. Mombaur, “Dancing with REEM-C: A robot-to-human physical-social communication study,” *arXiv preprint arXiv:2408.05301*, 2024. 2
- [8] R. Watanabe, C. Li, and M. Hutter, “DFM: Deep fourier mimic for expressive dance motion learning,” *arXiv preprint arXiv:2502.10980*, 2025. 2
- [9] H. Chen, W. Zhang, P. Li, S. Ma, K. Ma, Y. Jin, Z. Xu, X. Wang, Y. Zheng, Z. Wang, J. Zhao, Y. Chen, and W. Ding, “Rhythm: Learning interactive whole-body control for dual humanoids,” *arXiv preprint arXiv:2603.02856*, 2026. 2
- [10] Y. Du, Y. Li, B. Jia, Y. Lin, P. Zhou, W. Liang, Y. Yang, and S. Huang, “Learning human-humanoid coordination for collaborative object carrying,” *arXiv preprint arXiv:2510.14293*, 2025. 2
- [11] G. C. R. Bethala, H. Huang, N. Pudasaini, A. M. Ali, S. Yuan, C. Wen, A. Tzes, and Y. Fang, “H²-COMPACT: Human-humanoid co-manipulation via adaptive contact trajectory policies,” *arXiv preprint arXiv:2505.17627*, 2025. 2
- [12] Y. Wang, Q. Zhao, Y. F. Lau, R. Yu, H. W. Tsui, Q. Chen, J. Wang, J. Pang, and P. Tan, “HumanX: Toward agile and generalizable humanoid interaction skills from human videos,” *arXiv preprint arXiv:2602.02473*, 2026. 2
- [13] W.-J. Huang, Y.-Y. Zhang, Y.-L. Wei, Z.-W. Xia, J. Tan, Y.-M. Li, Z. Zhao, and W.-S. Zheng, “Learning whole-body human-humanoid interaction from human-human demonstrations,” *arXiv preprint arXiv:2601.09518*, 2026. 2
- [14] X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne, “DeepMimic: Example-guided deep reinforcement learning of physics-based character skills,” *arXiv preprint arXiv:1804.02717*, 2018. 2, 4, 5
- [15] X. B. Peng, Y. Guo, L. Halper, S. Levine, and S. Fidler, “ASE: Large-scale reusable adversarial skill embeddings for physically simulated characters,” *arXiv preprint arXiv:2205.01906*, 2022. 2

- [16] T. He, Z. Luo, W. Xiao, C. Zhang, K. Kitani, C. Liu, and G. Shi, "Learning human-to-humanoid real-time whole-body teleoperation," *arXiv preprint arXiv:2403.04436*, 2024. [2](#)
- [17] T. He, Z. Luo, X. He, W. Xiao, C. Zhang, W. Zhang, K. Kitani, C. Liu, and G. Shi, "OmniH2O: Universal and dexterous human-to-humanoid whole-body teleoperation and learning," *arXiv preprint arXiv:2406.08858*, 2024. [2](#)
- [18] Q. Liao, T. E. Truong, X. Huang, Y. Gao, G. Tevet, K. Sreenath, and C. K. Liu, "BeyondMimic: From motion tracking to versatile humanoid control via guided diffusion," *arXiv preprint arXiv:2508.08241*, 2025. [2](#), [4](#), [5](#)
- [19] N. Hogan, "Impedance control: An approach to manipulation: Parts I–III," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 1–24, 1985. [2](#)
- [20] B. Xu, H. Weng, Q. Lu, Y. Gao, and H. Xu, "FACET: Force-adaptive control via impedance reference tracking for legged robots," *arXiv preprint arXiv:2505.06883*, 2025. [2](#)
- [21] P. Zhi, P. Li, J. Yin, B. Jia, and S. Huang, "Learning a unified policy for position and force control in legged loco-manipulation," *arXiv preprint arXiv:2505.20829*, 2025. [2](#)
- [22] X. Luo, X. Chen, X. Yin, H. Wu, B. Xia, Z. Chen, J. Li, B. Chen, and X. Cheng, "CEER: Compliant end-effector and root control as a unified interface for hierarchical humanoid loco-manipulation," *arXiv preprint arXiv:2605.19981*, 2026. [2](#)
- [23] Q. Lu, Y. Feng, B. Shi, M. Pisen, Z. Bao, and C. K. Liu, "GentleHumanoid: Learning upper-body compliance for contact-rich human and object interaction," *arXiv preprint arXiv:2511.04679*, 2025. [2](#)
- [24] Y. Wang, Y. Sun, P. Patel, K. Daniilidis, M. J. Black, and M. Kocabas, "PromptHMR: Promptable human mesh recovery," *arXiv preprint arXiv:2504.06397*, 2025. [3](#)
- [25] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas, and M. J. Black, "Expressive body capture: 3D hands, face, and body from a single image," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 10975–10985. [3](#)
- [26] J. P. Araujo, Y. Ze, P. Xu, J. Wu, and C. K. Liu, "Retargeting matters: General motion retargeting for humanoid motion tracking," *arXiv preprint arXiv:2510.02252*, 2025. [3](#)
- [27] K. Zakka, Q. Liao, B. Yi, L. Le Lay, K. Sreenath, and P. Abbeel, "mjlab: A lightweight framework for GPU-accelerated robot learning," *arXiv preprint arXiv:2601.22074*, 2026. [5](#)

APPENDIX

Tables V to VII list the reward terms, the observations, and the randomization, force replay and episode settings of the two trained policies.

TABLE V: Reward terms of the two trained policies. Tracking terms are exponential kernels $\exp(-e^2/\sigma^2)$ of the error e with the listed σ ; weights multiply the per-step term. The two policies share every term except the locomotion objective: the decoupled policy draws its locomotion style from the adversarial term, the whole-body tracker (in the style of SoftMimic) from reference tracking of the lower body of the adapted reference.

Term	Decoupled	Whole-body	Parameters
<i>Partner following</i>			
Right foot follows the partner's left foot	2	2	offset 0.5 m behind, $\sigma = 0.25$ m, deadband 0.05 m, height tracked
Left foot follows the partner's right foot	2	2	same
Feet air time	0.5	0.5	air time between 0.2 s and 0.6 s
<i>Upper-body compliance</i>			
Arm positions, anchor-relative	1	1	$\sigma = 0.3$ m; shoulder, elbow, wrist links
Arm orientations, anchor-relative	2	2	$\sigma = 0.4$ rad
Force matching at contacted wrists	2	2	$\sigma = 15$ N, mean over active contacts
Torque matching at contacted wrists	2	2	$\sigma = 2$ N m
Contacted-link position emphasis	2.5	2.5	$\sigma = 0.2$ m
Contacted-link orientation emphasis	2.5	2.5	$\sigma = 0.3$ rad
<i>Lower body</i>			
Locomotion objective: adversarial term on the lower body	0.1 (coef.)		7 bodies, positions and 6D orientations per frame, 2 frames, pelvis anchor; discriminator 256, 256, gradient penalty 5
Anchor position, global		0.5	$\sigma = 0.3$ m
Anchor orientation, global		0.5	$\sigma = 0.4$ rad
Lower-body positions, anchor-relative		1	$\sigma = 0.3$ m; pelvis, hip, knee, ankle links
Lower-body orientations, anchor-relative		1	$\sigma = 0.4$ rad
Lower-body linear velocities, global		1	$\sigma = 1$ m s ⁻¹
Lower-body angular velocities, global		1	$\sigma = 3.14$ rad s ⁻¹
<i>Regularization and safety</i>			
Action rate	-0.1	-0.1	squared
Joint limit violation	-10	-10	all joints
Self collision	-10	-10	contact force above 10 N
Contact with the partner's body	-25	-25	fraction of steps in contact
Joint acceleration	-10 ⁻⁵	-10 ⁻⁵	hip and knee joints, squared

TABLE VI: Observations of the two trained policies, identical for both. Per-step sizes are stacked over the listed history. Noise is uniform in the listed range and applied to the actor only; the actor's joint positions also carry the per-joint encoder bias of Table VII. Both networks are MLPs with hidden layers 512, 256, 128 and ELU activations. The AMP discriminator of the decoupled policy reads 2 frames of 63 features (positions and 6D orientations of 7 lower-body and torso links in the pelvis frame), which the policy itself does not see.

Term	Per step	History	Actor noise	Actor	Critic
Base linear velocity (IMU)	3	5			✓
Base angular velocity (IMU)	3	5	±0.2	✓	✓
Joint positions	29	5	±0.01	✓	✓
Joint velocities	29	5	±0.5	✓	✓
Previous actions	29	5		✓	✓
Projected gravity	3	5	±0.05	✓	✓
Reference anchor orientation, free reference	6	5	±0.05	✓	✓
Reference anchor position, free reference	3	5			✓
Reference anchor pose, served reference	9	5			✓
Reference joint positions and velocities, free reference	58	5		✓	✓
Reference joint positions, served reference	29	5			✓
Partner pelvis and knees, robot pelvis frame	9	50	±0.25	✓	✓
Commanded stiffness per wrist (log), linear and rotational	4	5		✓	✓
Robot body positions, base frame (14 bodies)	42	5			✓
Robot body orientations, base frame (14 bodies)	84	5			✓
Applied wrench per wrist	12	5			✓
Scheduled wrench per wrist	12	5			✓
Environment spring stiffness per wrist (log)	4	5			✓

TABLE VII: Randomization, force replay and episode settings, identical for both policies.

Setting	Value
<i>Randomization (at episode start)</i>	
Torso centre of mass offset	uniform, $x \pm 0.025$ m, $y \pm 0.05$ m, $z \pm 0.05$ m
Foot friction coefficient	uniform, 0.3 to 1.2, one draw for all foot geoms
Joint encoder bias	uniform, ± 0.01 rad per joint, added to the actor's joint positions
Actor observation noise	uniform, per observation term, at the ranges the policy trained with
Reset root pose	$x, y \pm 0.15$ m, yaw ± 0.3 rad about the clip pose
<i>Force replay</i>	
Environment spring	stiffness and setpoint from the force schedule, force scale 1
Setpoint frame	torso frame frozen at the rising edge of each event, yaw only
Pushes	none besides the scheduled events
<i>Episode</i>	
Length, control rate	20 s, 50 Hz
Clip end	clip-chaining to another clip of the training data
Termination	anchor height below 0.3 m, anchor orientation off by more than 1.2 rad, a foot more than 1.2 m from its target, time out
Parallel environments	16384