IMPEDANCEMIMIC: LEARNED IMPEDANCE AS ACTION SPACE FOR FORCE-ACCURATE MOTION IMITATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Recovering the forces behind a motion is ill-posed: different controllers can reproduce the same movement while exerting different contact forces. Most motion-imitation methods anchor the humanoid in a physics simulator that grounds the contact dynamics, and drive it with a fixed-gain proportional–derivative (PD) controller. We show that fixed gains enforce a compromise between tracking the motion and recovering the resulting forces, where contact and free swing place opposing demands on the gains. To overcome this limitation, we present ImpedanceMimic, a motion-imitation reinforcement-learning policy that outputs per-joint stiffness, damping and position setpoints, supervised by force measurements withheld at inference. Across 129 participants, nine datasets and six diverse activities from lunges to jumping, ImpedanceMimic outperforms fixed-gain PD, direct torque and PHC on kinematic and force tracking, and approaches a dedicated motion-to-force regressor. With gains learned as multiples of inertia-scaled priors, our policy is the first to drive subject-scaled biomechanical models for every participant, without anthropometric input, and transfers zero-shot to new bodies. The learned impedance forms an activity-specific, morphology-invariant signature of the underlying motor strategy, allowing the policy to generalize to unseen activities. This signature occupies a low-dimensional, interpretable space where stiffness scales with contact force and damping anticipates contact, consistent with human perturbation studies. We release our code, model checkpoints and benchmarks at impedancemimic.github.io/ImpedanceMimic.

1 INTRODUCTION

Physical intelligence requires knowing not only how a body moves, but also the latent control that produces the motion and the forces it exerts on the world. Contact forces determine a humanoid's footstep stability, a patient's knee loading, and an athlete's performance. While easy-to-deploy cameras have made motion datasets abundant (Ionescu et al., 2013; Mahmood et al., 2019; Zhang et al., 2026), force data remains scarce: instrumented force plates confine recordings to the laboratory, and mobile sensorized insoles resolve only normal foot pressure and are drift-prone.

Recovering the control that drives motion is therefore of interest, since a control learned where force was measured can be applied where it was not. Recent efforts have standardized large corpora of synchronized motion and force-plate recordings (Werling et al., 2024) and extended their coverage past steady gait to cutting, lunging and jumping (Scherpereel et al., 2023). These corpora have so far been used to fit the map from kinematics to ground reaction forces (GRF) directly (Mourot et al., 2022; Tan et al., 2026). Such a map returns the expected force for a motion, with no physics at inference to keep its output dynamically consistent, and it never represents the control that produced the force, which remains buried and uninterpretable in its hidden layers.

Physics-based motion imitation restores both the physics and the control. Tracking a recorded motion inside a simulator yields contact forces from the rollout itself (Peng et al., 2018; Luo et al., 2023), making them dynamically consistent by construction, and the control is an explicit output rather than an internal encoding, and GPU-parallel engines now train such policies at scale (Google DeepMind & NVIDIA, 2025). Driving the simulated body requires choosing an action space, and nearly all imitation tracks the reference with a proportional–derivative (PD) controller whose stiffness and damping are fixed by hand (Peng & Van De Panne, 2017). Yet different controllers can produce the same motion with different contact forces, so kinematic fidelity does not guarantee force accuracy. Recent work has begun to bring measured force into the loop, as an observation (Hori et al., 2026) or as a training reward (Liu et al., 2026), but keeps the fixed-gain PD controller. Other action spaces have been explored, such as learned gains (Yuan et al., 2021), muscles (Li et al., 2026) and direct torque (Cotton, 2025), but with limited scrutiny of measured forces.

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.

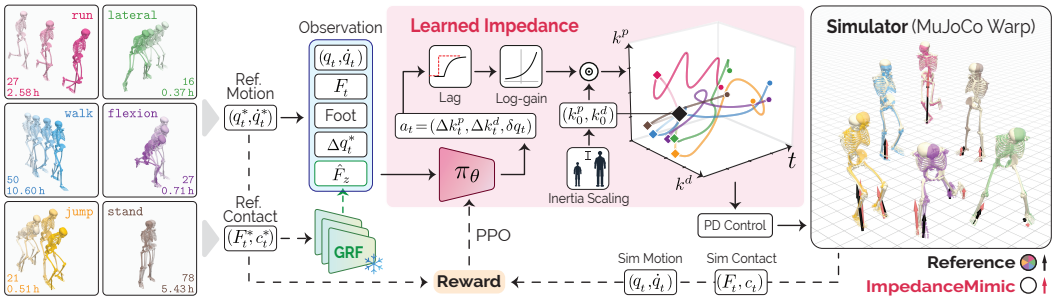


Figure 1: **ImpedanceMimic learns joint impedance to reproduce measured contact forces.** From a reference motion and a frozen kinematics-to-GRF estimate, the policy modulates inertia-scaled nominal gains into time-varying stiffness and damping per joint, tracing a trajectory characteristic of each task. These gains set a PD law whose torques drive a subject-scaled model in a physics simulator, producing contact forces. Dashed lines mark training-only signals, withheld at inference.

We present **ImpedanceMimic**, a reinforcement-learning imitation policy that outputs per-joint stiffness, damping and position setpoints, supervised by force-plate GRF and center of pressure (CoP). On 129 participants, nine datasets and six diverse activities, ImpedanceMimic outperforms fixed-PD, direct-torque and state-of-the-art imitation controllers on kinematic and force accuracy. To our knowledge, no prior controller reproduces measured per-foot contact forces across subjects and activities, including unseen ones, with no force measurement at inference. Collectively, we contribute:

1. **Action-space bounds on force recovery.** We derive a lower bound on stiffness, set by the torque the reference force demands, and an upper bound on damping, set by discrete-time stability, which oppose each other between contact and free-swing. A fixed-gain policy forfeits either the pose or the force, whereas a learned impedance satisfies each bound.
2. **A morphology-agnostic controller.** Expressing impedance as a multiple of an inertia-derived prior enables our policy, trained on 107 subject-scaled models, to transfer zero-shot to 22 unseen subjects without any anthropometric input. To our knowledge, this is the first physics-based imitator to transfer across subject-specific biomechanical embodiments.
3. **Learned impedance encodes a physiological signature of the task.** The learned gains occupy a low-dimensional, interpretable space where stiffness scales with contact force and damping anticipates contact state. This signature allows our policy to generalize to activities withheld from its training, where fixed action spaces fail.
4. **An open benchmark.** We release code, checkpoints and the split manifest of a benchmark over 129 participants and six activities, with fixed-PD, direct-torque, PHC (Luo et al., 2023) and GaitDynamics (Tan et al., 2026) baselines and extensive ablations.

2 Related Work

Physics-based motion imitation. Executing a recorded motion inside a simulator makes the resulting forces dynamically consistent. Reinforcement learning has made this routine at scale, from single-clip tracking (Peng et al., 2018) through adversarial objectives (Peng et al., 2021) to controllers covering large motion corpora (Luo et al., 2023; 2024; Tessler et al., 2024), and has been used to generate kinetic labels (Liu et al., 2025). The objective throughout is kinematic. These methods originate in animation, where success means the motion looks right, and physics is often used to repair generated motion rather than to measure it (Yuan et al., 2023). On the same criterion, Peng & Van De Panne (2017) compared torque, PD-target, velocity and muscle parameterizations and found fixed-gain PD tracking superior, the choice the field has adopted since. A parallel line in biomechanics solves the problem by trajectory optimization over musculoskeletal models (Dembia et al., 2020), physically rigorous but one trial at a time. Recent work carries RL imitation into anatomically grounded models with torque or muscle actuation (Cotton, 2025; Simos et al., 2025; Li et al., 2026; Choi et al., 2026), with limited validation against measured contact forces. The physical accuracy of these simulated motions has rarely been examined. Hori et al. (2026) uses measured force as an observation, driving an SMPL humanoid from inertial and insole sensors that are needed at inference too. Liu et al. (2026) adds GRF and CoP to the reward, but validates on a single walking participant and applies an external residual wrench to the pelvis (Yuan & Kitani, 2020).

Learned impedance control. Letting a policy set its own impedance originates in robotics, where variable stiffness and damping were learned to make contact-rich manipulation and legged hopping robust to contact uncertainty (Buchli et al., 2011; Martín-Martín et al., 2019; Bogdanovic et al., 2020), the contact force being a quantity to minimize rather than to reproduce. The idea has since entered physics-based character control and motion capture. SimPoE (Yuan et al., 2021) scales nominal gains per joint at every PD step and adds a residual root wrench to the action, ReGAIL (Boursin et al., 2025) puts servo gains and damping in the action so that a controller can alternate tension and relaxation, and Neural PhysCap (Shimada et al., 2021) and OSDCap (Le et al., 2024) regress the gains from video, the latter together with an external force inside a neural Kalman filter. In each case the gains serve a kinematic objective or the fidelity of a pose. None ties the gains to the body it controls, and none evaluates them against measured force. Le et al. (2026) comes closest, learning the gains of a center-of-mass PD controller against force plates, though the body it controls is a single point mass. No prior work has learned the impedance of an articulated human model against measured contact forces, nor shown its transfer across subjects and activities. We first motivate our approach on a single joint, which we then scale to human motion imitation in ImpedanceMimic.

3 METHOD

Ill-posed contact forces. We model the body as a floating-base rigid-body system with generalized coordinates $\boldsymbol{q} \in \mathbb{R}^n$, $n = 31$, six of which describe the pelvis and $m = 15$ of which are actuated joints. Its dynamics obey the Newton–Euler equations of motion,

$$\mathbf{M}(\boldsymbol{q})\,\ddot{\boldsymbol{q}} \;+\; \boldsymbol{h}(\boldsymbol{q},\dot{\boldsymbol{q}}) \;=\; \mathbf{S}^{\top}\boldsymbol{\tau} \;+\; \mathbf{J}_c(\boldsymbol{q})^{\top}\boldsymbol{\lambda}, \tag{1}$$

with mass matrix $\mathbf{M}$, bias forces $\boldsymbol{h}$ (Coriolis, centrifugal, gravitational), joint torques $\boldsymbol{\tau}$ mapped by the selection matrix $\mathbf{S}$, and contact forces $\boldsymbol{\lambda}$ through their Jacobian $\mathbf{J}_c$ (dimensions in App. A.2.2). Each GRF is the resultant of its contact forces applied at its CoP. A recorded motion $\boldsymbol{q}^*(t)$ fixes the left-hand side of Eq. 1, and its torque-free pelvis rows fix the net contact wrench to the rate of change of whole-body momentum, but not how that wrench is realized: in double support, its split between the feet and each CoP remains undetermined. Simulation resolves this through a contact model and a joint-torque trajectory $\boldsymbol{\tau}(t)$, so the accuracy of the recovered forces is that of $\boldsymbol{\tau}(t)$.

3.1 ACTION-SPACE IMPEDANCE BOUNDS

A policy $\pi_\theta(\boldsymbol{a}_t \mid \boldsymbol{o}_t)$ acts every $\Delta t = 10\,\text{ms}$ while the simulator steps every $h = 2\,\text{ms}$. The action space determines how the joint torque $\hat{\boldsymbol{\tau}}$ in Eq. 1 evolves over the $N = 5$ substeps: a direct-torque policy holds $\hat{\boldsymbol{\tau}} = \tau_{\text{scale}}\,\boldsymbol{a}_t$ fixed, so it cannot react between queries. Peng & Van De Panne (2017) wrap it in a proportional–derivative controller recomputed every substep, to smooth learning:

$$\hat{\boldsymbol{\tau}} \;=\; \boldsymbol{k}^p \circ \big(\boldsymbol{q}^*_{t+1} + \alpha\,\delta\boldsymbol{q} - \boldsymbol{q}\big) \;+\; \boldsymbol{k}^d \circ \big(\dot{\boldsymbol{q}}^*_{t+1} - \dot{\boldsymbol{q}}\big), \tag{2}$$

clipped to $\pm\boldsymbol{\tau}^{\lim}$, where $\delta\boldsymbol{q} \in [-1,1]^m$ are per-joint position offsets scaled by α, and $\circ$ the element-wise product. We extend this controller by emitting per-joint gain residuals $\Delta\boldsymbol{k}^p, \Delta\boldsymbol{k}^d \in [-1,1]^m$,

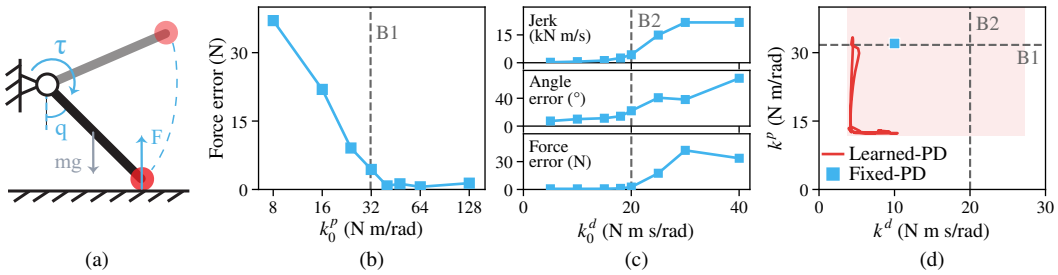
$$\boldsymbol{k}^p = \boldsymbol{k}^p_0 \circ \exp\big(s_p\,\Delta\boldsymbol{k}^p\big), \qquad \boldsymbol{k}^d = \boldsymbol{k}^d_0 \circ \exp\big(s_d\,\Delta\boldsymbol{k}^d\big), \tag{3}$$

about base gains $\boldsymbol{k}^p_0, \boldsymbol{k}^d_0$, with log-scales $s_p, s_d > 0$ setting the gain range. The exponential keeps gains strictly positive to maintain stability at the joint. We define two action spaces: *Fixed-PD* ($\boldsymbol{a} = \delta\boldsymbol{q}$, gains held at $(\boldsymbol{k}^p_0, \boldsymbol{k}^d_0)$), and *Learned-PD* ($\boldsymbol{a} = (\Delta\boldsymbol{k}^p, \Delta\boldsymbol{k}^d, \delta\boldsymbol{q})$). As observed empirically by Bogdanovic et al. (2020), the contact the environment holds the joint, so k^d has little effect and k^p sets the force the PD can produce. Out of contact the joint drives only the light distal segment, so k^d governs the closed-loop dynamics. We formalize this as two bounds.

Proposition 1 (Force authority, Bound 1)**.** *At the reference state $(\boldsymbol{q},\dot{\boldsymbol{q}}) = (\boldsymbol{q}^*,\dot{\boldsymbol{q}}^*)$ the velocity-error term of Eq. 2 vanishes and the displacement is $\mathcal{O}(\Delta t)$-small, so the PD torque reduces to $\boldsymbol{\tau} \approx \alpha\,\boldsymbol{k}^p \circ \delta\boldsymbol{q}$: the offset $\delta\boldsymbol{q}$ is the policy's sole lever for torque production at the reference pose, giving an achievable interval of half-width $\alpha\boldsymbol{k}^p$. Producing the reference contact force therefore requires the inverse-dynamics torque $\boldsymbol{\tau}^*$ of Eq. 1, which is attainable at the reference pose only if*

$$k^p \;\geq\; \frac{\max_t |\tau^*(t)|}{\alpha}. \tag{4}$$

Below this bound, the lever $\alpha\boldsymbol{k}^p$ is too small to supply the torque required by the reference force. The offset $\delta\boldsymbol{q}$ saturates, leaving the tracking term $\boldsymbol{k}^p \circ (\boldsymbol{q}^* - \boldsymbol{q})$ as the only source of torque. A fixed-gain policy below the bound must therefore choose between forfeiting kinematic or force tracking.



(a) (b) (c) (d)

Figure 2: **Action-space bounds on a single joint. (a)** A hinged rod tracks a trajectory and presses against a compliant floor. **(b)** Fixed-PD force error against the base stiffness. **(c)** Fixed-PD jerk, angle and force error against the base damping. **(d)** The k^p–k^d plane with both bounds, the nominal gains and Learned-PD's trajectory over one cycle.

Proposition 2 (Feedback stability, Bound 2)**.** *In continuous time any positive k^d is stabilizing. In simulation the PD of Eq. 2 is applied as an explicit actuator force at every physics step h (Peng & Van De Panne, 2017). On a free joint of inertia I, discretizing the error dynamics yields a 2×2 linear map whose eigenvalues must lie inside the unit circle. The Jury criterion gives (App. A.2.1)*

$$0 < k^d < \frac{2I}{h} \qquad \textit{and} \qquad 0 < k^p < \frac{4I}{h^2} - \frac{2k^d}{h}. \tag{5}$$

The damping condition is the more restrictive case, and depends on I, the changing inertia that the joint drives at a given moment. During contact, the joint acts on the kinematic chain of the whole body above it. During swing, it drives only the distal segment, whose inertia is orders of magnitude smaller, meaning the admissible k^d also decreases. A damping gain that is stable in contact can therefore be unstable in flight, where the feedback amplifies rather than corrects the error between policy steps and the torque alternates at $\pm\boldsymbol{\tau}^{\lim}$ every substep. This chatter is the classical limit on the impedance a sampled-data controller can render (Colgate & Brown, 1994). Stable PD avoids it by evaluating the controller at the next state (Tan et al., 2011), but the explicit PD used in imitation does not, so each joint inherits a damping ceiling set by its distal inertia.

Remark 1 (Phase-dependent feasibility)**.** *Bound 1 demands a stiff joint in contact and higher damping to keep it from oscillating, whereas Bound 2 needs a lightly damped joint in swing. Fixed gains must therefore favor one phase over the other, whereas state-dependent gains can match each bound.*

Validation. We test both bounds on the rod of Fig. 2a, where they evaluate to $k^p \geq 32\,\text{Nm/rad}$ and $k^d < 20\,\text{Nm s/rad}$ (App. A.2.1). Trained with identical base gains, reward and PPO (Schulman et al., 2017), Fixed-PD forfeits force tracking when its stiffness is below Bound 1 (Fig. 2b) and loses kinematic and force tracking for damping above Bound 2 (Fig. 2c). Learned-PD instead reaches the same force accuracy by raising k^p to Bound 1 in contact and holding k^d under Bound 2 throughout (Fig. 2d). Section 4.3 verifies these bounds on the humanoid (Figs.3 and 9, Table14).

3.2 ImpedanceMimic

Action space. The policy outputs $\boldsymbol{a}_t = (\Delta\boldsymbol{k}_t^p, \Delta\boldsymbol{k}_t^d, \delta\boldsymbol{q}_t) \in [-1,1]^{3m}$ with log-scales $s_p = 1.25$ and $s_d = 1.5$. A first-order lag mimics activation dynamics by smoothing gain changes,

$$\tilde{\Delta\boldsymbol{k}}_t = \tilde{\Delta\boldsymbol{k}}_{t-1} + \beta\big(\Delta\boldsymbol{k}_t - \tilde{\Delta\boldsymbol{k}}_{t-1}\big), \qquad \beta = 1 - e^{-\Delta t/T_{\text{act}}}, \quad T_{\text{act}} = 50\ \text{ms}, \tag{6}$$

that would otherwise appear as force transients, and lets the policy stiffen before touchdown. Base gains scale with each subject i's diagonal joint-space inertia $\boldsymbol{I}_{jj}(i)$ at the rest pose,

$$\boldsymbol{k}_0^p = \boldsymbol{I}_{jj} \circ \boldsymbol{\omega}_n^2, \qquad \boldsymbol{k}_0^d = 2\,\boldsymbol{\zeta} \circ \boldsymbol{I}_{jj} \circ \boldsymbol{\omega}_n, \tag{7}$$

where the natural frequency $\boldsymbol{\omega}_n$ and damping ratio $\boldsymbol{\zeta}$ are set per joint by nominal gains (Table 8), such that a multiple of the nominal results in the same closed-loop behavior across bodies.

Kinematics-to-GRF estimator. Setting the impedance before touchdown requires anticipating the contact force, which the policy cannot read from a force plate at inference. The actor therefore receives an estimate $\hat{\boldsymbol{F}}_t$ computed from the reference kinematics. A center-of-mass physics prior, corrected by a residual network ensemble, is fitted once per split on the training subjects' plate-valid frames and then frozen (App. A.2.2). It reaches a GRF correlation of 0.979 on seen and 0.928 on unseen subjects, on par with the dedicated GaitDynamics regressor (Tan et al., 2026).

Observations. The actor observes $\boldsymbol{o}_t \in \mathbb{R}^{215}$ (Table 7): proprioception and contact state from the simulator, and three lookaheads at $\{10, 50, 250\}$ ms, namely the kinematic offset $\boldsymbol{q}^*_{t+k} - \boldsymbol{q}_t$, the estimated GRF $\hat{\boldsymbol{F}}_{t+k}$, and foot positioning. Every entry is a simulator state or a function of the reference kinematics, so no force-plate signal reaches the actor and the policy runs plate-free at inference (Fig. 3). The critic receives the current-frame GRF to lower advantage variance (Pinto et al., 2017), and a linear head on the actor, discarded at inference, regresses the measured GRF at three lookahead offsets with a weighted Huber loss ($w_{\text{aux}} = 0.25$) in the PPO objective,

$$\mathcal{L} = \mathcal{L}_{\text{PPO}} + w_{\text{aux}} \,\text{Huber}\big(\hat{\boldsymbol{F}}^{\text{aux}}(\boldsymbol{o}_t),\, \mathcal{N}_{\text{priv}}(\boldsymbol{F}^*_{t+\{1,5,25\}})\big), \tag{8}$$

Reward. The reward combines kinematic imitation with force-plate supervision,

$$r_t = r_t^{\text{kin}} + w_{\text{grf}}\, r_t^{\text{grf}} + 0.25\, r_t^{\text{shear}} + 0.15\, r_t^{\text{ee}}, \tag{9}$$

with w_{grf} annealed linearly, such that the policy first learns to track and then to produce faithful forces. The kinematic term combines pose, velocity and pelvis tracking (App. A.2.2). Each joint's pose error is normalized by its amplitude over the clip, and the root term tracks pelvis velocity, such that it is independent of the subject's leg length. With $\Delta\boldsymbol{F}_f = \boldsymbol{F}_f - \boldsymbol{F}^*_f$ and $\Delta\boldsymbol{c}_f = \boldsymbol{c}_f - \boldsymbol{c}^*_f$ the force and CoP errors of foot f, and $\bar{F}_f = \max(F^*_{z,f},\, 0.5)$ the load they are scaled by,

$$r^{\text{grf}} = \exp\Big(-\sum_f g_f \big[\alpha_{\text{grf}} \tfrac{\|\Delta\boldsymbol{F}_f\|^2}{\bar{F}_f^2} + 5\,\|\Delta\boldsymbol{c}_f\|^2\big]\Big), \quad r^{\text{shear}} = \exp\Big(-5\sum_f g_f\, \|\Delta\boldsymbol{F}_{f,\parallel}\|^2\Big), \tag{10}$$

with $\alpha_{\text{grf}} = 4$. The contact gate g_f is active when either the simulated or the reference vertical force exceeds 0.1 BW, to penalize mistimed contacts. r^{ee} tracks heel and toe positions.

Datasets. We assemble 20 h of motion from 129 participants across 9 public datasets with synchronized force plates (Fig. 1, App. A.1.1), distributed through AddBiomechanics (Werling et al., 2024) with joint kinematics, per-foot GRF and centers of pressure. Recordings cover six activity classes, walking, running, standing, lateral movements, flexion and jumping, with body masses ranging from 44 to 126 kg and heights from 1.46 to 1.95 m. We resample to 100 Hz, low-pass filter, cut event-aligned clips of 1–10 s and normalize forces to body weight (App. A.1).

Simulation model. Every subject is simulated with the no-arm Rajagopal model, scaled by AddBiomechanics and converted to MuJoCo (Todorov et al., 2012) with the joints driven by torques. All subjects share the topology of 14 bodies and the $n = 31$ coordinates of Eq. 1, while masses, inertias, segment lengths, contact geometry and joint ranges are subject-specific. Each foot carries eight contact spheres building on Falisse et al. (2019) (App. A.2.2). The policy runs at 100 Hz, the rate of the reference data, and the physics at $h = 2$ ms.

Training and evaluation. We train actor and critic ($[512, 256, 128]$, SiLU) with PPO for 2 billion steps on 8192 parallel worlds in one batched MuJoCo Warp environment, each world carrying its own subject-scaled model (App. A.2.2). Episodes start at a uniformly sampled frame with noise on pose and velocity and terminate when the clip ends or the pelvis drops below 0.4 m (Peng et al., 2018). Our policy receives no subject identifier or anthropometric descriptor: morphology enters only through the priors of Eq. 7. We report three benchmarks (App. A.1.3): *seen* (train on all 129 subjects, evaluate on the 20% of each subject's held-out clips), *unseen* (train on 107 subjects, evaluate zero-shot on 22 held out) and *leave-one-activity-out* (train on five activities, evaluate on the sixth). Evaluation rollouts report survival, RMSE and Pearson correlation for joint angles and per-foot forces, center-of-pressure error and torque consistency (App. A.3).

Baselines. We compare ImpedanceMimic to two other action spaces, extending the analysis of Peng & Van De Panne (2017) to contact forces, and to two external methods. The action-space baselines share ImpedanceMimic's simulator, data, observations, reward, network and PPO settings. *Fixed-PD* outputs the position offset $\delta\boldsymbol{q}_t$ with fixed inertia-scaled base gains (App. A.4.1). *Torque* outputs $\hat{\boldsymbol{\tau}}_t = \boldsymbol{\tau}_{\text{scale}} \circ \boldsymbol{a}_t$ with the per-joint, per-subject scale $\boldsymbol{\tau}_{\text{scale}} = \alpha\, \boldsymbol{k}_0^p$, such that its authority is inertia-normalized like the PD baselines (App. A.2.2). *GaitDynamics* (Tan et al., 2026) is a generative diffusion model that maps joint kinematics to bilateral GRF and CoP (App. A.2.3). Given it has no physics and no kinematics to imitate, it serves as an upper bound for force accuracy. *PHC* (Luo et al., 2023) is the state-of-the-art physics-based imitation controller and we use it to test whether kinematic fidelity alone yields accurate forces (App. A.2.3).

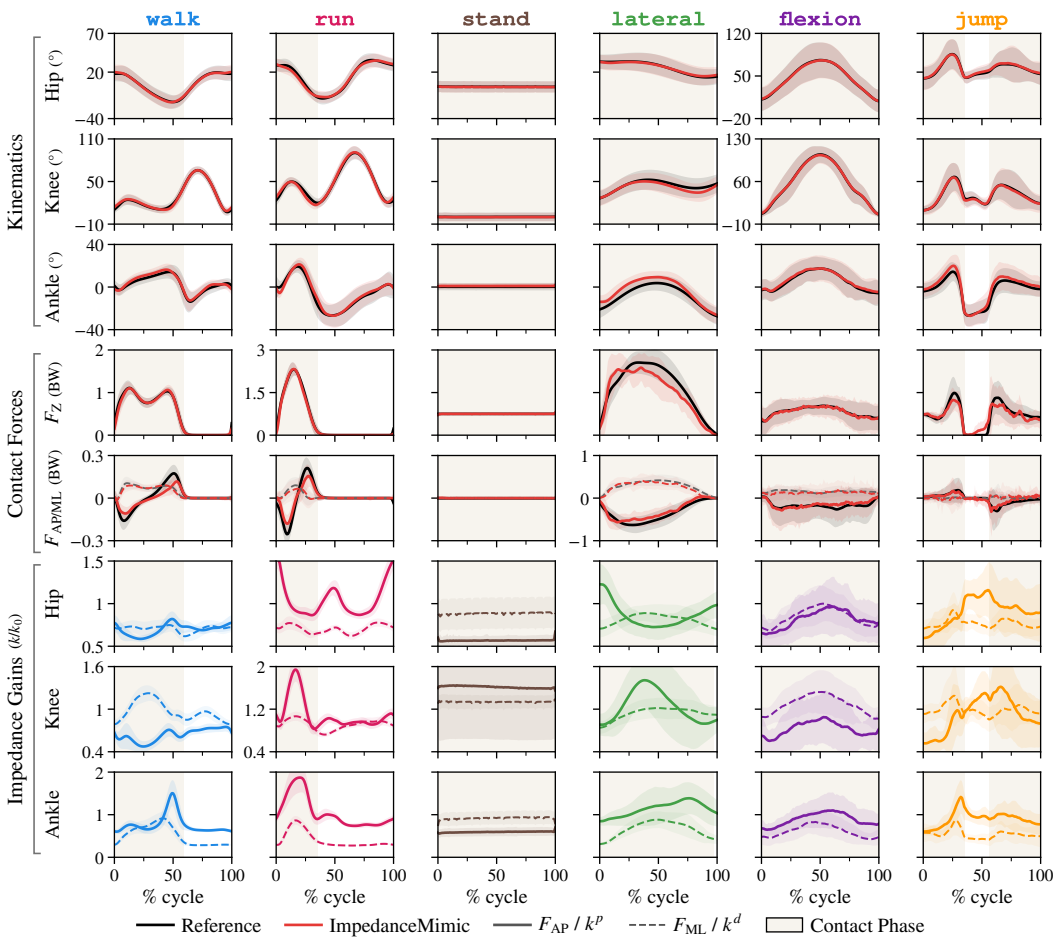


Figure 3: **Reference and simulated trajectories across the six activities.** Cycle-normalized hip, knee and ankle angles, vertical and horizontal GRF, and the learned hip, knee and ankle gains k^p/k_0^p (solid) and k^d/k_0^d (dashed). Curves show the mean $\pm$ 1 s.d. over all cycles of the held-out clips.

4 Results

4.1 Learned Impedance Tracks Motion and Forces Simultaneously

ImpedanceMimic outperforms all baselines for combined kinematic and force tracking (Table 1, Fig. 3). On held-out clips from seen subjects, it holds the lowest joint-angle error in every activity and matches or leads every controller on the contact-force metrics of Table 10, aside from the Pearson correlation on vertical GRF during standing, where a near-constant reference depresses every method, GaitDynamics included. Fixed-PD and Torque land at opposite ends of the trade-off of Remark 1, trading force for kinematic accuracy, and vice-versa. When removing the force reward in our ablations (Table 19), the joint-angle error improves at the cost of force accuracy, confirming that optimizing for motion trades force away.

Activities further separate into those whose forces follow from the motion and those whose forces do not. On cyclic walking and running, where the leg acts as a spring, Fixed-PD is highly accurate. In comparison, on the four non-cyclic activities, it gives up over a third of its force correlation while ImpedanceMimic retains most of it. Torque exhibits the mirror pattern in kinematics, its error growing from under one and a half times that of ImpedanceMimic on walking to over double on jumping. Fixed-PD pays the excess in pose or force frame by frame, and the per-clip violation share predicts its running errors ($\rho = +0.33$) while leaving ImpedanceMimic's untouched (Fig. 10).

We observe that Bound 2 replicates its behaviour from the single-joint example to the humanoid. With nominal gains, Fixed-PD's torque jerk sits an order of magnitude above ImpedanceMimic's in

Table 1: **ImpedanceMimic produces force-accurate motion imitation across six diverse activities and transfers zero-shot to unseen subject morphologies.** Each baseline trains one policy on all six activities across 1) *Seen subjects*: held-out clips from training subjects (same morphologies, different trials), 2) *Unseen subjects*: all clips from a disjoint set H of held-out subjects (new morphologies, zero-shot). **Bold** / underline = best / second best within each activity (ties share a rank). Full metrics, including per-seed (3) standard deviations, in Appendix A.3.1.

Activity	walk			run			stand			lateral			flexion			jump			**Mean**		
Method	Ep.↑	q↓	F_{Z_r}↑	Ep.↑	q↓	F_{Z_r}↑	Ep.↑	q↓	F_{Z_r}↑	Ep.↑	q↓	F_{Z_r}↑	Ep.↑	q↓	F_{Z_r}↑	Ep.↑	q↓	F_{Z_r}↑	Ep.↑	q↓	F_{Z_r}↑
Seen subjects, held-out clips																					
*GaitDynamics**	–	–	*1.00*	–	–	*1.00*	–	–	*.78*	–	–	*.90*	–	–	*.92*	–	–	*.90*	–	–	*.92*
Torque	99.4	2.40	**.98**	99.8	3.67	.96	99.2	2.30	**.69**	96.1	6.44	.89	93.8	5.19	.69	76.9	7.91	**.74**	94.2	4.65	**.83**
Fixed-PD	99.9	2.51	.96	99.8	3.49	.97	97.8	1.34	.30	96.0	3.85	.86	91.6	2.81	.64	71.7	3.88	.63	92.8	2.98	.73
PHC	98.9	9.78	.33	96.3	11.8	.53	97.6	7.74	.19	94.5	12.6	.53	83.0	13.3	.50	68.7	15.5	.44	89.9	11.8	.42
ImpedanceMimic	**100**	**1.70**	.98	**100**	**2.40**	**.99**	**99.6**	**1.10**	.51	**99.8**	**3.18**	**.90**	**97.6**	**2.40**	**.71**	**82.9**	**3.26**	.69	**96.6**	**2.34**	.80
Unseen subjects, zero-shot																					
*GaitDynamics**	–	–	*.99*	–	–	*1.00*	–	–	*.65*	–	–	*.90*	–	–	*.91*	–	–	*.87*	–	–	*.89*
Torque	**100**	2.85	**.98**	96.4	3.79	.97	**99.9**	2.26	**.68**	81.8	8.11	.80	74.8	6.24	.59	54.5	8.42	**.66**	84.6	5.28	**.78**
Fixed-PD	99.6	2.89	.94	98.6	3.38	.96	99.2	1.32	.46	94.1	4.24	.77	84.4	3.24	.51	71.9	4.21	.47	91.3	3.21	.69
PHC	95.2	10.2	.40	96.8	12.7	.50	93.1	7.30	.40	88.3	14.4	.50	75.7	14.1	.46	53.8	17.8	.35	83.8	12.7	.44
ImpedanceMimic	**100**	**2.04**	.97	**100**	**2.47**	**.99**	99.7	**.99**	.53	**99.7**	**3.63**	**.84**	**99.7**	**2.54**	**.66**	**89.6**	**3.42**	.61	**98.1**	**2.52**	.76

*GaitDynamics is shown in italics for reference and excluded from the ranking, since it is not a controller and serves as a reference for GRF.
Ep. = normalized episode survival before termination. q = joint-angle RMSE (°). F_{Z_r} = Pearson correlation w.r.t. reference vertical GRF.

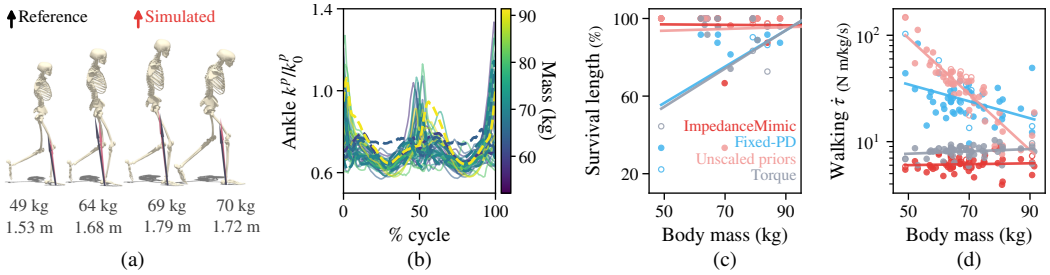


Figure 4: **One controller for every body. (a)** Unseen walkers at the same gait event under one policy with no anthropometric input, with their measured and simulated GRF. **(b)** Median ankle k^p/k_0^p over the walking cycle per subject (39 seen solid, 6 unseen dashed), colored by mass. **(c)** Survival on lateral movements and **(d)** torque jerk on walking clips against mass (seen filled, unseen open).

every activity, including standing, due to the nominal ankle damping exceeding the free-foot bound by a third. When capping damping below the bound, torque jerk significantly reduced and cyclic force tracking was recovered (Table 15, App. A.4.1). Even capped, Fixed-PD dispays almost three times ImpedanceMimic's jerk, whereas the learned policy holds the ankle at a third its nominal value in swing and meets Bound 1 in contact by raising k^p (Fig. 3).

PHC, which optimizes kinematic imitation alone, tracks the kinematics five times worse than ImpedanceMimic and recovers half of its force correlation, leading only on root position, which follows from our velocity-tracking root term. The learned impedance also trains fastest, exceeding PHC's final reward at its first evaluation, Fixed-PD's within a sixth of the budget and Torque's shortly past half (Fig. 8). Against GaitDynamics, the regression upper bound, ImpedanceMimic reaches 87% of its vertical force correlation. Although GaitDynamics demonstrates the strongest performance for predicting contact forces, it produces negative GRFs and places the CoP outside its the foot area (App. A.4.2). In the contrary, ImpedanceMimic does not violate any physics, highlighting the advantage of predicting contact forces physics simulators.

4.2 One Controller for Every Body

Evaluated on 22 held-out participants, ImpedanceMimic maintains its kinematic and force accuracy and survives nearly every clip (Table 1, Fig. 4a). This is attributed to the base gains that scale with the subject's joint-space inertia (Eq. 7), the policy only has to learn multiples of the same closed-loop natural frequency and damping ratio on every body. Dividing each subject's ankle stiffness trajectory by its own k_0^p collapses 45 walking subjects from 52 to 91 kg onto one band, a contact-phase multiple of 0.71 ± 0.03 uncorrelated with body mass, with the unseen walkers inside it (Fig. 4b).

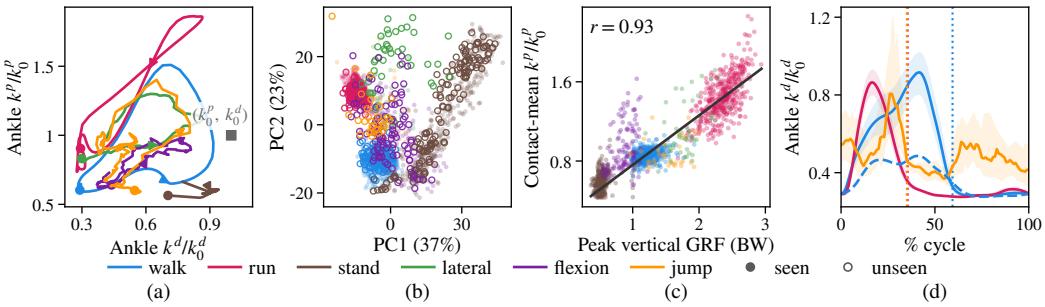


Figure 5: **The learned impedance encodes the activity.** Ankle gains of the seen-split policy divided by the subject's nominal k_0. **(a)** Median k^d–k^p loop per activity. **(b)** Principal components of the gain cycles fitted on seen subjects, with the unseen subjects projected onto them. **(c)** Contact-mean k^p/k_0^p against peak vertical GRF. **(d)** k^d/k_0^d over the cycle (median, interquartile band), toe-off and take-off dotted; dashed: the no-walk policy on walking.

Table 2: **Leave-one-activity-out generalization.** One policy per fold, trained on five activities and evaluated zero-shot on all clips of the held-out activity. **Bold** / underline = best / second best within each activity. Full metrics, including per-seed (3) standard deviations, in Appendix A.3.2.

Held-out activity	walk			run			stand			lateral			flexion			jump			**Mean**		
Method	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$	Ep.↑	q↓	$F^{\uparrow}_{Z_r}$
*GaitDynamics**	–	–	*.96*	–	–	*.99*	–	–	*.85*	–	–	*.85*	–	–	*.86*	–	–	*.87*	–	–	*.90*
Torque	54.9	8.00	**.72**	62.6	10.6	.84	**96.2**	4.21	.59	86.3	7.71	.82	66.1	8.88	.64	70.3	9.98	**.65**	72.7	8.23	.71
Fixed-PD	37.9	4.47	.45	68.3	5.58	.73	68.6	3.12	.40	93.6	4.19	.77	71.1	3.53	.60	69.6	4.10	.59	68.2	4.17	.59
PHC	38.2	19.5	.35	64.8	15.4	.43	44.4	15.8	.29	89.7	13.5	.50	73.8	15.8	.44	60.7	16.3	.36	61.9	16.1	.40
ImpedanceMimic	**85.4**	**3.44**	.67	**99.0**	**3.96**	**.91**	92.3	**2.40**	**.64**	**99.1**	**3.45**	**.83**	**80.3**	**3.14**	**.65**	**82.3**	**3.69**	.63	**89.7**	**3.35**	**.72**

*GaitDynamics is shown in italics for reference and excluded from the ranking, since it is not a controller and serves as a reference for GRF. Ep. = normalized episode survival before termination. q = joint-angle RMSE (°). F_{Z_r} = Pearson correlation w.r.t. reference vertical GRF.

ImpedanceMimic's survival, force correlation and torque jerk are flat in body mass (Fig. 4c,d). Fixed-PD, which has the scaling but not the modulation, degrades toward light bodies: on lateral movements its survival falls with mass, from near-full on the heaviest subjects to a fifth on the lightest, and on walking it carries an order of magnitude more torque jerk than ImpedanceMimic on the lightest walker. The ablation with learned gains but unscaled priors survives the lateral trials but pays the steepest mass-graded walking jerk, since the bound moves with the body and its nominal damping does not. Torque, whose per-joint scale carries the same inertia prior but no feedback between queries, transfers walking to every body, yet loses survival in proportion to body size on lateral movements alongside Fixed-PD, and remains the action space that loses the most survival on new bodies (Table 1): the prior transfers authority, not control. Inertia scaling buys transfer, learned modulation buys force accuracy, and ImpedanceMimic is the only action space with both.

4.3 Impedance Encodes a Physiological Signature of Task

Withholding an activity tests whether the learned impedance transfers across tasks as it does across bodies. ImpedanceMimic survives nine in ten held-out clips where every baseline survives under three quarters, its joint-angle error barely moves where Torque's nearly doubles, and on force it edges past Torque at well under half the kinematic error (Table 2); root drift on unseen locomotion is the exception. The two controllers fail differently: Fixed-PD by instability, surviving just over a third of clips without walking with its torque jerk doubled, and ImpedanceMimic by force, still walking on the walk fold but with its force correlation down from 0.98 to 0.67. One impedance chosen for five activities is unstable on the sixth; a force profile never supervised on is what a controller cannot reproduce.

Each activity has its own impedance signature, and it transfers (Fig. 5a,b). Over the 16,862 held-out cycles of 117 subjects, each divided by its own nominal k_0, the ankle traces one closed loop per activity: walking and running raise stiffness to one and a half and nearly twice nominal through stance and release damping to a third in swing, while standing stays below nominal in both (Table 13). Activity explains half the variance of the gain cycles, and a logistic regression recovers it at

$92.3 \pm 2.3\%$ on held-out subjects and $87.6 \pm 2.7\%$ on the 22 unseen ones projected onto the seen basis (App. A.4.1): the encoding belongs to the action space and the task, not to one network or one body.

Two rules generate the signatures, one per bound. Stiffness scales with load (Fig. 5c): contact-mean ankle k^p/k_0^p correlates with the cycle's peak vertical GRF at $r = 0.93$, and the learned k^p meets Bound 1 on nearly every running contact frame where Fixed-PD's nominal gain meets it on half to three quarters (Fig. 9). Damping is released before unloading (Fig. 5d): ankle k^d peaks $17.5 \pm 0.3\%$ of the cycle before toe-off in every one of $11{,}533$ walking cycles and settles at a third of nominal, under Bound 2, which the nominal damping itself violates. Where the comparison is valid the values are human: over the stance window of Rouse et al. (2014) the learned stiffness rises from 2.3 to 3.0 N m/rad/kg against their 1.5 to 6.5, quiet standing holds 0.4 of the toppling stiffness (Loram & Lakie, 2002), the running peak exceeds the walking one as for a leg spring stiffening with speed (McMahon & Cheng, 1990; Farley & Gonzalez, 1996), and damping leads unloading as Lee & Hogan (2015) report, though its magnitude follows Bound 2 (App. A.4.1).

The no-walk policy makes the test direct: driven on walking it never trained on, its ankle damping keeps the shared timing at half amplitude while its push-off stiffness rises only half as far (Fig. 5d, dashed). The contact-gated damping transfers and keeps it walking; the load-scaled stiffness must be learned per activity, and its absence is the force it loses.

5 Limitations and Future Work

ImpedanceMimic recovers forces inside a simulation, and is where limitations stem from. Our huamnoid has no arms, and penalizes our results for jumping and deep flexion, where whole-body angular momentum matters most and the corpus is thinnest. Physical consistency means consistency with the fitted contact model (Fig. 6) and with reference kinematics already reconciled with the measured forces. The root position is not tracked, only its velocity, and it drifts by a quarter of a meter on seen subjects (Tables 10 and 11). A residual wrench would remove the drift (Yuan & Kitani, 2020; Yuan et al., 2021; Cotton, 2025), however we do not implement it precisely because it would absorb the force error we set out to measure. The estimator is noncausal, reading the reference 400 ms ahead, but the oracle rows of Tables 19 and 20 cap what any better force signal buys at 0.02 in force correlation: a causal replacement need only match the estimate, not the plates. Future work follows: arms are additive, and muscle actuation would supply state-dependent impedance intrinsically, at six times the simulation cost (App. A.4).

6 Conclusion

We introduced ImpedanceMimic, a motion-imitation policy that emits per-joint, state-dependent impedance, supervised by measured ground reaction forces withheld at inference. Across 129 participants and six activities, ImpedanceMimic outperforms all other baselines in tracking kinematic and force targets, and approaches the GaitDynamics regressor on force correlation while producing physically consistent forces. We show that fixed gains cannot satisfy force authority and feedback stability within one stride, where learned impedance unites the kinematic fidelity of Fixed-PD with the force authority of Torque. Expressed as multiples of a body-specific inertial prior, the learned gains are invariant to morphology, enabling zero-shot transfer across bodies. Finally, the learned impedance gains constitute an interpretable, low-dimensional and physiological encoding of the task, one that extends to activities withheld from training.

AI USE STATEMENT

In this work, we used generative AI tools for literature research, for polishing author-written text, and as a coding assistant, including for LaTeX table formatting. All research ideas, the experimental design, and the analysis are our own. We therefore did not use generative AI for ideation or for the generation of research artifacts such as figures, data, or proofs.

We have reviewed all AI-assisted work: we read every paper we cite in its original published form, and we reviewed every line of AI-assisted code. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Code is released openly at https://impedancemimic.github.io/ImpedanceMimic/. Experimental data used in this work is open-sourced at https://addbiomechanics.org/.

REFERENCES

Karl J Åström and Björn Wittenmark. *Computer-controlled systems: theory and design*. Courier Corporation, 2013.

Miroslav Bogdanovic, Majid Khadiv, and Ludovic Righetti. Learning variable impedance control for contact sensitive tasks. *IEEE Robotics and Automation Letters*, 5(4):6129–6136, 2020.

Paul Boursin, Yannis Kedadry, Tony Chevalier, Victor Zordan, Paul Kry, Sophie Grégoire, and Marie-Paule Cani. Imitation in relative terms using regail: Making motion controllers agile and transferrable. *Computers & Graphics*, pp. 104457, 2025.

Jonas Buchli, Freek Stulp, Evangelos Theodorou, and Stefan Schaal. Learning variable impedance control. *The International Journal of Robotics Research*, 30(7):820–833, 2011.

Jonathan Camargo, Aditya Ramanathan, Will Flanagan, and Aaron Young. A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions. *Journal of biomechanics*, 119:110320, 2021.

Josh Carter, Xi Chen, Dario Cazzola, Grant Trewartha, and Ezio Preatoni. Full body kinematics and ground reaction forces of fifty heterogeneous runners completing treadmill running at various speeds and gradients. 2024.

Itak Choi, Ilseung Park, Eni Halilaj, and Inseung Kang. Musculoskeletal motion imitation for learning personalized exoskeleton control policy in impaired gait. *arXiv preprint arXiv:2604.09431*, 2026.

J Edward Colgate and J Michael Brown. Factors affecting the z-width of a haptic display. In *Proceedings of the 1994 IEEE international conference on robotics and automation*, pp. 3205–3210. IEEE, 1994.

R James Cotton. Kintwin: imitation learning with torque and muscle driven biomechanical models enables precise replication of able-bodied and impaired movement from markerless motion capture. *arXiv preprint arXiv:2505.13436*, 2025.

Christopher L Dembia, Nicholas A Bianco, Antoine Falisse, Jennifer L Hicks, and Scott L Delp. Opensim moco: Musculoskeletal optimal control. *PLoS computational biology*, 16(12): e1008493, 2020.

Damiana A Dos Santos, Claudiane A Fukuchi, Reginaldo K Fukuchi, and Marcos Duarte. A data set with kinematic and ground reaction forces of human balance. *PeerJ*, 5:e3626, 2017.

Antoine Falisse, Gil Serrancolí, Christopher L Dembia, Joris Gillis, Ilse Jonkers, and Friedl De Groote. Rapid predictive simulations with complex musculoskeletal models suggest that diverse healthy and pathological human gaits can emerge from similar control strategies. *Journal of The Royal Society Interface*, 16(157):20190402, 2019.

Claire T Farley and Octavio Gonzalez. Leg stiffness and stride frequency in human running. *Journal of biomechanics*, 29(2):181–186, 1996.

Benjamin J Fregly, Thor F Besier, David G Lloyd, Scott L Delp, Scott A Banks, Marcus G Pandy, and Darryl D D'lima. Grand challenge competition to predict in vivo knee loads. *Journal of orthopaedic research*, 30(4):503–513, 2012.

Google DeepMind and NVIDIA. MuJoCo Warp: GPU-optimized version of the MuJoCo physics simulator. `https://github.com/google-deepmind/mujoco_warp`, 2025.

Samuel R Hamner and Scott L Delp. Muscle contributions to fore-aft and vertical body mass center accelerations over a range of running speeds. *Journal of biomechanics*, 46(4):780–787, 2013.

Xingjian Han, Ben Senderling, Stanley To, Deepak Kumar, Emily Whiting, and Jun Saito. Groundlink: A dataset unifying human body movement and ground reaction dynamics. In *SIGGRAPH Asia 2023 Conference Papers*, pp. 1–10, 2023.

Ryosuke Hori, Jyun-Ting Song, Zhengyi Luo, Jinkun Cao, Soyong Shin, Hideo Saito, and Kris Kitani. Ground reaction inertial poser: Physics-based human motion capture from sparse imus and insole pressure sensors. *arXiv preprint arXiv:2603.16233*, 2026.

Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE transactions on pattern analysis and machine intelligence*, 36(7):1325–1339, 2013.

Cuong Le, Viktor Johansson, Manon Kok, and Bastian Wandt. Optimal-state dynamics estimation for physics-based human motion capture from videos. *Advances in Neural Information Processing Systems*, 37:43609–43631, 2024.

Cuong Le, Urs Waldmann, Bastian Wandt, and Mårten Wadenbäck. Gravity-guided contact dynamics estimation from 3d human motions. *arXiv preprint arXiv:2606.08133*, 2026.

Hyunglae Lee and Neville Hogan. Time-varying ankle mechanical impedance during human locomotion. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(5):755–764, 2015.

Chengkun Li, Cheryl Wang, Bianca Ziliotto, Merkourios Simos, Jozsef Kovecses, Guillaume Durandau, and Alexander Mathis. Towards embodied ai with musclemimic: Unlocking full-body musculoskeletal motor learning at scale. *arXiv preprint arXiv:2603.25544*, 2026.

Xinpeng Liu, Junxuan Liang, Zili Lin, Haowen Hou, Yong-Lu Li, and Cewu Lu. Imdy: Human inverse dynamics from imitated observations. In *International Conference on Learning Representations*, volume 2025, pp. 1209–1229, 2025.

Xinyi Liu, Jangwhan Ahn, Edgar Lobaton, Jennie Si, and He Huang. Imitation learning from human motion alone does not guarantee biomechanically plausible gait kinetics. *arXiv preprint arXiv:2603.12408*, 2026.

Ian D Loram and Martin Lakie. Direct measurement of human ankle stiffness during quiet standing: the intrinsic mechanical stiffness is insufficient for stability. *The journal of physiology*, 545(3): 1041–1053, 2002.

Zhengyi Luo, Jinkun Cao, Kris Kitani, Weipeng Xu, et al. Perpetual humanoid control for real-time simulated avatars. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10895–10904, 2023.

Zhengyi Luo, Jinkun Cao, Josh Merel, Alexander Winkler, Jing Huang, Kris Kitani, and Weipeng Xu. Universal humanoid motion representations for physics-based control. In *International Conference on Learning Representations*, volume 2024, pp. 56766–56782, 2024.

Naureen Mahmood, Nima Ghorbani, Nikolaus F Troje, Gerard Pons-Moll, and Michael J Black. Amass: Archive of motion capture as surface shapes. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 5442–5451, 2019.

Roberto Martín-Martín, Michelle A Lee, Rachel Gardner, Silvio Savarese, Jeannette Bohg, and Animesh Garg. Variable impedance control in end-effector space: An action space for reinforcement learning in contact-rich tasks. In *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 1010–1017. IEEE, 2019.

Thomas A McMahon and George C Cheng. The mechanics of running: how does stiffness couple with speed? *Journal of biomechanics*, 23:65–78, 1990.

Jason K Moore, Sandra K Hnat, and Antonie J van den Bogert. An elaborate data set on human gait and the effect of mechanical perturbations. *PeerJ*, 3:e918, 2015.

Lucas Mourot, Ludovic Hoyet, François Le Clerc, and Pierre Hellier. Underpressure: Deep learning for foot contact detection, ground reaction force estimation and footskate cleanup. In *Computer Graphics Forum*, volume 41, pp. 195–206. Wiley Online Library, 2022.

Xue Bin Peng and Michiel Van De Panne. Learning locomotion skills using deeprl: Does the choice of action space matter? In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 1–13, 2017.

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

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

Lerrel Pinto, Marcin Andrychowicz, Peter Welinder, Wojciech Zaremba, and Pieter Abbeel. Asymmetric actor critic for image-based robot learning. *arXiv preprint arXiv:1710.06542*, 2017.

Apoorva Rajagopal, Christopher L Dembia, Matthew S DeMers, Denny D Delp, Jennifer L Hicks, and Scott L Delp. Full-body musculoskeletal model for muscle-driven simulation of human gait. *IEEE transactions on biomedical engineering*, 63(10):2068–2079, 2016.

Elliott J Rouse, Levi J Hargrove, Eric J Perreault, and Todd A Kuiken. Estimation of human ankle impedance during the stance phase of walking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):870–878, 2014.

Keaton Scherpereel, Dean Molinaro, Omer Inan, Max Shepherd, and Aaron Young. A human lower-limb biomechanics and wearable sensors dataset during cyclic and non-cyclic activities. *Scientific Data*, 10(1):924, 2023.

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

Soshi Shimada, Vladislav Golyanik, Weipeng Xu, Patrick Pérez, and Christian Theobalt. Neural monocular 3d human motion capture with physical awareness. *ACM Transactions on Graphics (ToG)*, 40(4):1–15, 2021.

Merkourios Simos, Alberto Silvio Chiappa, and Alexander Mathis. Kinesis: Motion imitation for human musculoskeletal locomotion. *arXiv preprint arXiv:2503.14637*, 2025.

Jie Tan, Karen Liu, and Greg Turk. Stable proportional-derivative controllers. *IEEE Computer Graphics and Applications*, 31(4):34–44, 2011.

Tian Tan, Zachary A Strout, Roy TH Cheung, and Peter B Shull. Strike index estimation using a convolutional neural network with a single, shoe-mounted inertial sensor. *Journal of Biomechanics*, 139:111145, 2022a.

Tian Tan, Dianxin Wang, Peter B Shull, and Eni Halilaj. Imu and smartphone camera fusion for knee adduction and knee flexion moment estimation during walking. *IEEE Transactions on Industrial Informatics*, 19(2):1445–1455, 2022b.

Tian Tan, Tom Van Wouwe, Keenon F Werling, C Karen Liu, Scott L Delp, Jennifer L Hicks, and Akshay S Chaudhari. Gaitdynamics: a generative foundation model for analyzing human walking and running. *Nature biomedical engineering*, pp. 1–13, 2026.

Chen Tessler, Yunrong Guo, Ofir Nabati, Gal Chechik, and Xue Bin Peng. Maskedmimic: Unified physics-based character control through masked motion inpainting. *ACM Transactions On Graphics (TOG)*, 43(6):1–21, 2024.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems*, pp. 5026–5033. IEEE, 2012.

Huawei Wang, Vittorio Caggiano, Guillaume Durandau, Massimo Sartori, and Vikash Kumar. Myosim: Fast and physiologically realistic mujoco models for musculoskeletal and exoskeletal studies. In *2022 International Conference on Robotics and Automation (ICRA)*, pp. 8104–8111. IEEE, 2022.

Huawei Wang, Akash Basu, Guillaume Durandau, and Massimo Sartori. A wearable real-time kinetic measurement sensor setup for human locomotion. *Wearable technologies*, 4:e11, 2023.

Keenon Werling, Nicholas A Bianco, Michael Raitor, Jon Stingel, Jennifer L Hicks, Steven H Collins, Scott L Delp, and C Karen Liu. Addbiomechanics: Automating model scaling, inverse kinematics, and inverse dynamics from human motion data through sequential optimization. *Plos one*, 18(11):e0295152, 2023.

Keenon Werling, Janelle Kaneda, Tian Tan, Rishi Agarwal, Six Skov, Tom Van Wouwe, Scott Uhlrich, Nicholas Bianco, Carmichael Ong, Antoine Falisse, et al. Addbiomechanics dataset: Capturing the physics of human motion at scale. In *European Conference on Computer Vision*, pp. 490–508. Springer, 2024.

Ye Yuan and Kris Kitani. Residual force control for agile human behavior imitation and extended motion synthesis. *Advances in Neural Information Processing Systems*, 33:21763–21774, 2020.

Ye Yuan, Shih-En Wei, Tomas Simon, Kris Kitani, and Jason Saragih. Simpoe: Simulated character control for 3d human pose estimation. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7155–7165. IEEE, 2021.

Ye Yuan, Jiaming Song, Umar Iqbal, Arash Vahdat, and Jan Kautz. Physdiff: Physics-guided human motion diffusion model. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 15964–15975. IEEE, 2023.

Jiahao Zhang, Joseph Liu, Young-Yoon Lee, Seonghyeon Moon, Victor Zordan, Guy Tevet, C Karen Liu, Stephen Gould, Oren Jacob, Haomiao Jiang, et al. Romo: A large-scale, richly organized dataset and semantic taxonomy for human motion generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16408–16419, 2026.

A Appendix

A.1 Datasets and Preprocessing

A.1.1 Dataset Descriptions

Our corpus pools nine public motion-capture datasets that record joint kinematics together with ground reaction forces (GRF). Eight are taken from the AddBiomechanics open-source release (Werling et al., 2024), which provides for every subject a Rajagopal lower-body model (Rajagopal et al., 2016) scaled to that subject and, for every trial, joint angles and force-plate data in a unified format. We add a ninth dataset, Scherpereel et al. (2023), for non-cyclic activity coverage. It is not part of the AddBiomechanics release, so we processed its raw motion-capture files through the AddBiomechanics tool (Werling et al., 2023) (Appendix A.1.2). After curation, the corpus contains 129 participants, 972 source trials, 8,493 clips and 20.2 hours of motion at 100 Hz (Table 3). We exclude AddBiomechanics' largest paired motion–force datasets: Carter et al. (2024) lacks CoP data, and Camargo et al. (2021) has excessive marker error.

Table 3: **Dataset summary.** Some subjects from the source dataset were dropped due to data quality issues. H is the number of held-out participants (Section A.1.3). Trials are segmented into 1–10 s clips. The original rate was resampled to 100 Hz across datasets. TM = instrumented treadmill, OG = overground force plates. Activities: `W`alk, `R`un, `S`tand, `F`lexion, `L`ateral, `J`ump.

Dataset	Subj.	H	Trials	Clips	Frames	Hours	Rate (Hz)	Setting	Activities
Dos Santos et al. (2017)	49/49	9	294	1764	1,762,236	4.90	100	OG	S
Fregly et al. (2012)	6/6	1	113	115	19,061	0.05	120–200	OG	W
Hamner & Delp (2013)	10/10	2	36	36	5,226	0.01	100–120	TM	R
Han et al. (2023)	7/7	1	117	144	90,479	0.25	250	OG	W S F L J
Moore et al. (2015)	12/15	2	35	2047	1,967,364	5.46	100	TM	W S
Scherpereel et al. (2023)	11/12	2	116	1001	464,459	1.29	200	OG	S F L J
Tan et al. (2022a)	8/16	1	79	818	771,555	2.14	200	TM	R
Tan et al. (2022b)	17/17	3	68	1648	1,577,293	4.38	100	TM	W
Wang et al. (2023)	9/12	1	114	920	614,376	1.71	100	TM/OG	W R F J
Total	129	22	972	8493	7,272,049	20.20			

Our preprocessing and quality checks exclude data at three levels: whole subjects, individual trials, and single frames. *Subjects*: we retain 12/15 subjects of Moore et al. (2015), 8/16 of Tan et al. (2022a) and 9/12 of Wang et al. (2023), dropping those that failed AddBiomechanics processing, and 11/12 of Scherpereel et al. (2023), excluding AB02 whose markers were reattached across four sessions yielding an inconsistent scaled model. *Trials*: every trial-level exclusion follows one of three criteria: (i) unmeasured external forces, removing the walking-pole trials of Fregly et al. (2012) and the ball-toss, tug-of-war and weight-lifting trials of Scherpereel et al. (2023); (ii) incomplete force-plate coverage, removing the treadmill activities of Scherpereel et al. (2023) and the stepping trials of Han et al. (2023); and (iii) environment geometry we do not simulate, removing the sit-to-stand, stair and ramp trials of Scherpereel et al. (2023). *Frames*: frames in which both feet load a single plate, or plate coverage is lost, are marked invalid (Han et al., 2023; Tan et al., 2022a).

A.1.2 Data Preprocessing

Subject models and trial extraction. Each AddBiomechanics subject file (`.b3d`) holds a scaled OpenSim model and its trials. We read the final dynamics pass and export one HDF5 per trial with the 23 generalized coordinates of the no-arm Rajagopal model (Rajagopal et al., 2016), their velocities, and the two-foot GRF and center of pressure (CoP), rejoining consecutive 20 s segments of the same recording. The MuJoCo model adds the eight auxiliary knee coordinates tied to the knee angles by coupling constraints, giving the $n = 31$ coordinates of Eq. 1: six pelvis, 15 actuated and ten passive. Each scaled `.osim` model is converted to MuJoCo with MyoConverter (Wang et al., 2022), keeping skeleton, inertia and coupling constraints and discarding muscles and tendons (App. A.2.2). Joint torques are recomputed with MuJoCo inverse dynamics on each subject's model, with GRF normalized to a body weight estimated from standing trials or, failing that, from the pelvis residual force.

Signal processing. Joint angles and GRFs are resampled to 100 Hz with a polyphase FIR filter. Isolated GRF dropouts of up to three frames are interpolated, the first and last two frames of every trial are truncated, and frames with vertical force below 0.02 BW are flagged invalid. GRF, CoP and pelvis velocity are rotated into a heading-local frame so that observations are invariant to laboratory orientation. For the five treadmill datasets (Hamner & Delp, 2013; Moore et al., 2015; Tan et al., 2022a;b; Wang et al., 2023), whose kinematics AddBiomechanics expresses in the belt frame, we recover a laboratory frame from the planted-foot constraint: a loaded foot cannot translate, so its forward-kinematic displacement gives the belt travel, which is spline-interpolated across strides and added to the pelvis translation, velocity and CoP, leaving a residual foot slide below 2 cm in walking. Loaded frames whose corrected CoP falls outside an oriented bounding box of the foot are interpolated from neighboring in-box frames or zeroed.

Clip segmentation. Valid trials are segmented into clips of 1–10 s at biomechanically meaningful boundaries: walking clips start and end at double-support onsets and running clips at foot strikes, greedily packed up to the 10 s ceiling so that a clip covers whole strides, standing is cut into 10 s windows, and non-cyclic overground activities are cut at the midpoints of bilateral standing phases. Cyclic clips reach a median length of 9.7–10 s, lateral, flexion and jump clips 2.1, 3.4 and 4.1 s.

A.1.3 Data Splits and Evaluation Protocol

Splits. All splits are drawn once with a fixed seed and frozen for every experiment (Table 4). *Seen subjects, held-out clips* withholds a random 20% of each participant's clips per activity, training on 129 participants and 6,796 clips. *Unseen subjects, zero-shot* withholds $\lfloor 0.2N_d \rfloor$ participants per dataset, training on the 7,128 clips of the remaining 107 and evaluating the 22 unseen participants zero-shot, 3–14 per activity. *Leave-one-activity-out* trains on five activities and evaluates zero-shot on the sixth, excluding datasets that contain only the left-out activity so that transfer to a new activity is not confounded with transfer to a new morphology: the stand fold drops Dos Santos et al. (2017), the walk fold Fregly et al. (2012) and Tan et al. (2022b), and the run fold Hamner & Delp (2013) and Tan et al. (2022a), leaving 4,112 of the 8,493 clips for evaluation.

Table 4: **Evaluation clips per split axis.** Seen = Seen subjects, held-out clips. Unseen = Unseen subjects, zero-shot. LOAO reports the participants available for training in each fold.

	Seen	*Unseen*	*LOAO*	
`Activity`	Eval clips	Eval clips	Train subj.	Eval clips
`walk`	816	670	106	2351
`run`	207	128	111	159
`stand`	387	360	80	214
`lateral`	93	80	129	462
`flexion`	125	79	129	589
`jump`	69	48	129	337
Total	1697	1365		4112

Evaluation protocol. Each evaluation clip is split into valid segments of at least 25 frames: every segment is one episode, initialized at its first reference frame and rolled out with the policy's mean action at 100 Hz until the segment ends or the episode terminates (pelvis below 0.4 m, or numerically invalid state). The first 0.1 s is excluded from all fidelity metrics, since the reset places the model in the reference pose without settling. We report the metrics of Tables 10 and 11 per clip. Activity scores average over clips, and the overall score is the mean of the six activity means, each reported as mean and standard deviation over three training seeds.

A.1.4 SUBJECT CHARACTERISTICS

Table 5 lists every participant with sex, age, mass, height, the activities they contribute with trial and clip counts, and the number of held-out clips and participants. The 129 participants span 44–126 kg and 1.46–1.95 m (mass 69.5 ± 12.5 kg, height 1.71 ± 0.10 m). Ages are reported for 102 participants (34 ± 18 yr) and range from 19 to 84. Age is unreported for Fregly et al. (2012), Hamner & Delp (2013) and Scherpereel et al. (2023). Sex is 45 female and 84 male.

Table 5: **Subject characteristics across all datasets.** Activities are coded as W = walk, R = run, S = stand, F = flexion, L = lateral, J = jump, with (#trials, #clips) for each activity. Trials were segmented into clips ∈ [1,10] s. *Held-out clips (H-O)*: per activity, $\mathrm{round}(0.2 \times n_{\mathrm{clips}})$ clips are held out, randomly sampled (seed 42). *Held-out participants (**bold** rows)*: per dataset, $\lfloor 0.2 \times N \rfloor$ participants are held out, randomly sampled (seed 42). Since all datasets have $N \geq 6$, at least one participant is always withheld, yielding 3-14 unseen participants per activity.

Subj.	Sex	Age	Mass	Ht.	Activities	H-O
Dos Santos et al. (2017)						
s0	M	25	74.3	1.72	S(6,36)	7
s1	M	26	126.3	1.92	S(6,36)	7
s2	F	22	52.9	1.67	S(6,36)	7
s3	M	33	75.8	1.79	S(6,36)	7
s4	M	24	61.0	1.84	S(6,36)	7
s5	**M**	**25**	**83.2**	**1.74**	**S(6,36)**	**7**
s6	**F**	**24**	**71.8**	**1.58**	**S(6,36)**	**7**
s7	M	36	64.0	1.82	S(6,36)	7
s8	**F**	**25**	**61.2**	**1.69**	**S(6,36)**	**7**
s9	F	31	61.7	1.62	S(6,36)	7
s10	M	32	77.6	1.92	S(6,36)	7
s11	F	24	48.8	1.58	S(6,36)	7
s12	M	70	65.3	1.75	S(6,36)	7
s13	M	30	95.4	1.71	S(6,36)	7
s14	**F**	**31**	**64.9**	**1.53**	**S(6,36)**	**7**
s15	**M**	**23**	**89.3**	**1.81**	**S(6,36)**	**7**
s16	M	31	77.9	1.73	S(6,36)	7
s17	**F**	**61**	**70.6**	**1.50**	**S(6,36)**	**7**
s18	M	63	72.2	1.72	S(6,36)	7
s19	M	28	79.1	1.86	S(6,36)	7
s20	F	28	59.3	1.68	S(6,36)	7
s21	M	62	70.5	1.64	S(6,36)	7
s22	M	29	68.2	1.76	S(6,36)	7
s23	M	68	70.2	1.67	S(6,36)	7
s24	F	21	62.5	1.70	S(6,36)	7
s25	M	22	61.5	1.81	S(6,36)	7
s26	F	25	61.2	1.70	S(6,36)	7
s27	M	62	73.1	1.74	S(6,36)	7
s28	M	84	66.3	1.56	S(6,36)	7
s29	F	68	49.2	1.47	S(6,36)	7
s30	M	61	74.2	1.62	S(6,36)	7
s31	F	72	65.7	1.46	S(6,36)	7
s32	F	72	62.0	1.54	S(6,36)	7
s33	F	63	59.8	1.61	S(6,36)	7
s34	**F**	**67**	**66.3**	**1.60**	**S(6,36)**	**7**
s35	M	69	73.2	1.60	S(6,36)	7
s36	F	80	52.7	1.51	S(6,36)	7
s37	**M**	**72**	**92.9**	**1.72**	**S(6,36)**	**7**
s38	F	69	66.1	1.60	S(6,36)	7
s39	F	28	44.3	1.54	S(6,36)	7
s40	F	68	65.3	1.51	S(6,36)	7
s41	F	63	55.6	1.61	S(6,36)	7
s42	M	63	96.0	1.78	S(6,36)	7
s43	**M**	**61**	**82.2**	**1.70**	**S(6,36)**	**7**
s44	F	24	51.7	1.61	S(6,36)	7
s45	M	27	101.8	1.72	S(6,36)	7
s46	F	34	53.5	1.56	S(6,36)	7
s47	M	37	69.6	1.55	S(6,36)	7
s48	F	64	60.8	1.58	S(6,36)	7
Fregly et al. (2012)						
1GC	M		64.6	1.66	W(14,14)	3
2GC	M		67.0	1.72	W(15,15)	3
3GC	F		78.4	1.67	W(31,31)	6
4GC	M		66.7	1.68	W(31,31)	6
5GC	M		75.0	1.80	W(1,3)	1
6GC	**M**		**70.0**	**1.72**	**W(21,21)**	**4**
Hamner & Delp (2013)						
s01	**M**		**72.8**	**1.77**	**R(4,4)**	**1**
s02	**M**		**76.5**	**1.82**	**R(3,3)**	**1**
s03	M		64.0	1.78	R(4,4)	1
s04	M		80.5	1.77	R(4,4)	1
s08	M		82.4	1.79	R(4,4)	1
s10	M		69.3	1.70	R(3,3)	1
s11	M		69.3	1.74	R(4,4)	1
s17	M		68.4	1.68	R(3,3)	1
s19	M		65.1	1.72	R(4,4)	1
s20	M		67.2	1.67	R(3,3)	1

Subj.	Sex	Age	Mass	Ht.	Activities	H-O
Han et al. (2023)						
s001	F	22	69.9	1.68	W(3,3),S(11,14),F(8,13),L(3,3)	8
s002	M	21	66.7	1.70	W(2,2),S(8,9),F(9,11)	4
s003	F	23	53.1	1.59	W(3,3),F(10,14),L(2,2)	4
s004	M	20	71.7	1.82	W(2,2),S(4,4),F(6,10),L(1,1)	3
s005	M	24	90.7	1.95	W(3,3),S(5,8),F(8,11),L(1,1)	5
s006	**F**	**35**	**49.0**	**1.53**	**W(3,3),S(4,4),F(9,11),L(3,3),J(2,2)**	**5**
s007	F	22	64.0	1.75	S(4,4),F(3,3)	2
Moore et al. (2015)						
s3	**F**	**32**	**60.0**	**1.62**	**W(3,129),S(3,6)**	**27**
s5	M	23	65.0	1.73	W(3,186),S(3,6)	38
s6	M	26	80.0	1.77	W(2,122),S(2,4)	25
s7	F	29	63.0	1.72	W(3,185),S(3,6)	38
s8	M	20	70.0	1.57	W(3,134),S(3,6)	28
s9	M	20	64.0	1.69	W(3,138),S(3,6)	29
s10	**M**	**19**	**90.7**	**1.77**	**W(3,186),S(3,6)**	**38**
s12	M	22	81.0	1.85	W(3,179),S(3,6)	37
s13	F	21	64.0	1.70	W(3,186),S(3,6)	38
s15	M	22	79.4	1.83	W(3,183),S(2,4)	38
s16	F	28	52.2	1.69	W(3,180),S(3,6)	37
s17	M	23	87.0	1.86	W(3,171),S(3,6)	35
Scherpereel et al. (2023)						
AB01	**M**		**79.1**	**1.75**	**S(1,11),F(2,20),L(4,38),J(2,15)**	**17**
AB03	M		113.9	1.79	S(1,11),F(2,18),L(5,36),J(2,11)	14
AB05	F		71.8	1.69	S(1,6),F(2,20),L(5,43),J(2,10)	18
AB06	M		80.7	1.83	S(1,8),F(2,26),L(5,44),J(2,18)	19
AB07	F		62.0	1.65	S(1,13),F(2,24),L(5,42),J(3,14)	17
AB08	M		88.1	1.71	S(1,9),F(2,28),L(5,41),J(3,13)	19
AB09	**M**		**83.9**	**1.77**	**S(1,9),F(2,18),L(5,39),J(3,13)**	**18**
AB10	M		67.5	1.73	S(1,10),F(2,25),L(5,41),J(3,10)	19
AB11	F		64.2	1.57	S(1,8),F(2,36),L(5,39),J(3,18)	23
AB12	F		63.3	1.65	S(1,9),F(2,31),L(5,48),J(3,21)	22
AB13	F		67.7	1.65	S(1,9),F(2,37),L(5,41),J(3,20)	24
Tan et al. (2022a)						
s1	**F**	**25**	**56.1**	**1.58**	**R(11,109)**	**22**
s3	F	24	57.4	1.69	R(5,48)	10
s4	M	25	69.0	1.78	R(7,82)	16
s5	F	22	52.1	1.58	R(12,122)	24
s6	M	24	68.0	1.75	R(11,114)	23
s7	M	23	72.3	1.78	R(11,113)	23
s8	F	24	53.0	1.59	R(10,104)	21
s9	M	23	66.0	1.75	R(12,126)	25
Tan et al. (2022b)						
02	M	25	71.4	1.83	W(4,97)	19
04	M	23	57.0	1.62	W(4,97)	19
05	**M**	**23**	**64.0**	**1.68**	**W(4,97)**	**19**
06	**M**	**24**	**67.2**	**1.75**	**W(4,97)**	**19**
07	M	22	74.4	1.71	W(4,97)	19
08	M	22	59.0	1.79	W(4,97)	19
09	**M**	**24**	**70.7**	**1.77**	**W(4,96)**	**19**
10	M	22	58.1	1.76	W(4,97)	19
11	M	23	66.9	1.80	W(4,97)	19
12	M	24	63.8	1.73	W(4,97)	19
13	M	22	73.2	1.76	W(4,97)	19
14	M	22	61.8	1.81	W(4,97)	19
15	M	24	64.6	1.76	W(4,97)	19
17	M	26	74.8	1.87	W(4,97)	19
18	M	23	68.9	1.82	W(4,97)	19
19	M	23	91.4	1.85	W(4,97)	19
20	M	22	57.4	1.70	W(4,97)	19
Wang et al. (2023)						
s04	M	23	75.2	1.80	W(6,42),R(3,21),F(2,21),J(2,18)	20
s05	F	23	65.1	1.71	W(6,44),R(3,19),F(2,28),J(2,19)	23
s06	F	23	56.0	1.68	W(6,42),R(3,14),F(2,34),J(2,21)	22
s07	M	23	80.8	1.82	W(6,40),R(3,19),F(2,18),J(2,21)	20
s08	F	28	61.5	1.62	W(6,40),R(3,20),F(2,30),J(2,19)	22
s09	M	26	81.0	1.82	W(6,40),R(3,18),F(2,26),J(2,17)	20
s10	M	28	71.6	1.66	W(6,41),R(3,18),F(2,22),J(2,19)	20
s11	F	25	62.4	1.68	W(4,26),R(3,18),F(2,24),J(2,20)	18
s12	**M**	**30**	**69.0**	**1.79**	**W(6,41),R(2,12),F(2,30),J(2,18)**	**20**

A.2 METHOD

A.2.1 ACTION-SPACE IMPEDANCE BOUNDS

Proof of Proposition 2.

Proof. Between policy queries the command $q^* + \alpha\,\delta q$ and gains (k^p, k^d) are held fixed. Define $e = q - (q^* + \alpha\,\delta q)$ and $v = \dot{q} - \dot{q}^*$. The PD torque (Eq. 2) reduces to $\hat{\tau} = -k^p e - k^d v$. On a free joint of inertia I, Newton's law gives $I\dot{v} = -k^p e - k^d v$. Discretizing with semi-implicit Euler ($v_+ = v + h\dot{v}$, $e_+ = e + hv_+$):

$$\begin{pmatrix} e_+ \\ v_+ \end{pmatrix} = A \begin{pmatrix} e \\ v \end{pmatrix}, \qquad A = \begin{pmatrix} 1 - h^2 k^p / I & h(1 - hk^d/I) \\ -hk^p/I & 1 - hk^d/I \end{pmatrix}. \tag{11}$$

Asymptotic stability requires all eigenvalues of A inside the unit circle. For a real 2×2 matrix the Jury criterion (Åström & Wittenmark, 2013) gives necessary and sufficient conditions: $|\det A| < 1$ and $|\operatorname{tr} A| < 1 + \det A$. Computing:

$$\det A = 1 - \frac{hk^d}{I}, \qquad \operatorname{tr} A = 2 - \frac{hk^d}{I} - \frac{h^2 k^p}{I}. \tag{12}$$

The condition $|\det A| < 1$ gives $0 < k^d < 2I/h$. Substituting into $|\operatorname{tr} A| < 1 + \det A$ and simplifying gives $0 < k^p < 4I/h^2 - 2k^d/h$. □

Single-joint model. The model is a uniform rod on a hinge whose tip contacts a compliant floor, so that one joint alternates between a grounded force task and a free kinematic task within one cycle. Every constant follows from the rod itself (Table 6). All results are over 3 seeds.

Table 6: One-joint model, controller and protocol.

Rod	$L = 0.20\,\text{m}$, $m = 1.5\,\text{kg}$, uniform; $I_{jj} = mL^2/3 = 0.020\,\text{kg}\,\text{m}^2$; hinge height $z = 60\,\text{mm}$
Contact	sphere $r = 20\,\text{mm}$, contact at $q_c = 11.5°$; solref $= (0.01, 1)$, i.e. time constant $10\,\text{ms}$ at damping ratio 1
Rates	physics $h = 2\,\text{ms}$, policy $100\,\text{Hz}$ ($N = 5$ substeps)
Actuator	$\lvert\hat{\tau}\rvert \le \tau^{\lim} = 40\,\text{Nm}$, Torque: $\hat{\tau} = \tau_{\text{scale}}\,a$ where $\tau_{\text{scale}} = 40\,\text{Nm}$, held over Δt
Gain priors	$k_0^p = I_{jj}\omega_n^2 = 32\,\text{Nm/rad}$ ($\omega_n = 40\,\text{rad/s}$), $k_0^d = I_{jj}/h = 10\,\text{Nm s/rad} = \frac{1}{2}\cdot 2I_{jj}/h$
Learned-PD	$k = k_0\, e^{\overline{\Delta k}}$, $\overline{\Delta k} \in [-1, 1]$ (log-scale $s = 1$), $\alpha = 0.5\,\text{rad}$
Reference	$1\,\text{s}$ cycle, contact 53%; F^*: $\sin^2$ bell, peak $mg = 14.7\,\text{N}$; free phase: $20°$
Reward	$r^{\text{kin}} + 0.25\, r^{\text{force}}$; $r^{\text{kin}} = \exp(-2\,q_e^2) + \exp(-0.1\,\dot{q}_e^2)$; $r^{\text{force}} = \exp(-2\,\text{gate}\cdot F_e^2)$
Randomization	per episode, unobserved: floor time constant $\times\, e^{\mathcal{U}(-\ln 1.4,\, \ln 1.4)}$, rod inertia $\times\,\mathcal{U}(0.85, 1.15)$; reference-state initialization; termination at $\lvert q - q^*\rvert > 20°$
Sweeps	$\xi_F \in \{0.25, 0.5, 1, 2, 3, 4, 6, 8, 10\}$ at $\xi_I = 1$
Gain cuts	$k_0^d \in \{0.25, \ldots, 2\} \times 2I_{jj}/h$ at nominal, $k_0^p \in \{0.25, \ldots, 4\} \times 32$ at $\xi_F = 6$
Policies	16 conditions $\times$ 4 action spaces $\times$ 2 reward modes $\times$ 3 seeds $+$ 39 gain cuts
PPO	128 envs $\times$ 32 steps $\times$ 2000 updates (8.2 M steps per policy), network $(64, 64)$
Evaluation	$\bar{e}_q = \text{mean}_t \lvert q - q^*\rvert$, $\varepsilon_F = \text{RMS}_{\text{stance}}(F - F^*)/F^*_{\max}$

Sweep protocol. Two multipliers move the model relative to the bounds. The force multiplier ξ_F scales the reference force F^*, and hence τ^*, leaving q^* unchanged, so it raises Bound 1 against a fixed k_0^p. The inertia multiplier ξ_I scales the rod's mass and inertia and, through Eq. 7, both gain priors with it. The gain cuts instead hold the reference fixed and retrain Fixed-PD with one base gain set to a specified fraction of its bound, isolating each bound in turn.

Bound values for this model. Bound 1 is the net joint torque the contact demands, divided by the offset authority. With the contact force acting at the rod tip and the rod's own weight at its center,

$$\frac{\max_t |\tau^*(t)|}{\alpha} = \frac{mgL\cos q_c}{\alpha}\left(\xi_F - \tfrac{1}{2}\right) = 5.77\,\xi_F - 2.88\ \ \text{Nm/rad} \qquad (\xi_F \ge 1.3), \tag{13}$$

and below $\xi_F = 1.3$ the binding requirement is the free-phase lift instead ($4.5\,\text{Nm/rad}$). Bound 2 is $2I_{jj}/h = 20\,\text{Nm s/rad}$, which the prior $k_0^d = 10\,\text{Nm s/rad}$ meets with a factor of two in hand. The exponential parameterization of Eq. 3 gives Learned-PD a searchable range of $k^p \in [k_0^p e^{-s},\, k_0^p e^{s}] =$

$[11.8, 87.0]$ Nm/rad and $k^d \in [3.7, 27.2]$ Nm s/rad at $s = 1$. The damping ceiling sits above Bound 2, so staying stable in flight is something the policy has to learn rather than something the parameterization enforces.

Force sweep. Raising ξ_F drives Bound 1 through the fixed k_0^p, which happens when the peak torque exhausts the offset authority, $\max_t |\tau^*(t)| = \alpha k_0^p$. Substituting $k_0^p = I_{jj}\omega_n^2 = mL^2\omega_n^2/3$ into Eq. 13, the mass cancels and the crossing

$$\xi_F = \frac{1}{2} + \frac{\alpha L \omega_n^2}{3\, g \cos q_c} = 6.05 \tag{14}$$

is fixed by the controller bandwidth, the action range, the rod geometry and gravity alone. The failure appears there: Fixed-PD's force error rises from < 1 N below the crossing to 31 N at $\xi_F = 10$, while Learned-PD stays below 1 N throughout at comparable angle error. At the crossing itself, $\xi_F = 6$, Bound 1 evaluates to 31.7 Nm/rad and the nominal $k_0^p = 32$ Nm/rad sits on it. This is the operating point for the gain trajectories and gain cuts of Fig. 2.

Gain cuts. Sweeping k_0^p at $\xi_F = 6$ tests Bound 1 in isolation. The force error drops from 37 N at $k_0^p = 8$ Nm/rad to under 1 N above $k_0^p = 40$ Nm/rad, bracketing the bound at 31.7 Nm/rad, while the angle error stays between 9 and 13°: below the bound the controller holds the pose and gives up the force. Sweeping k_0^d at the nominal reference tests Bound 2. The torque jerk jumps from 1.2 kNm/s at $k_0^d = 15$ Nm s/rad to 15 kNm/s at $k_0^d = 25$ Nm s/rad, bracketing the bound at 20 Nm s/rad, with angle and force error rising sharply alongside it.

A.2.2 IMPEDANCEMIMIC

Algorithm 1 ImpedanceMimic Training.

Require: Reference motions $\mathcal{D} = \{(\boldsymbol{q}^*, \dot{\boldsymbol{q}}^*, \boldsymbol{F}^*, \boldsymbol{c}^*)\}$, subject-specific models $\{m_s\}_{s=1}^S$
Ensure: Policy π_θ, GRF estimator $\hat{\boldsymbol{F}}$
Stage 1: Kinematics-only GRF estimator
$\boldsymbol{F}^{\text{prior}} \leftarrow \textsc{PhysicsPrior}(\boldsymbol{q}^*, \dot{\boldsymbol{q}}^*)$ ▷ center-of-mass dynamics
Train ensemble $\{f_{\phi_i}\}_{i=1}^K$ to regress the residual $\boldsymbol{F}^* - \boldsymbol{F}^{\text{prior}}$
$\hat{\boldsymbol{F}} \leftarrow \boldsymbol{F}^{\text{prior}} + \frac{1}{K}\sum_i f_{\phi_i}$; $\quad \hat{c} \leftarrow \mathbf{1}[\hat{F}_z > \epsilon_c]$ ▷ frozen during RL
Stage 2: Impedance policy learning
Scale nominal gains $\boldsymbol{k}_0^p, \boldsymbol{k}_0^d$ per subject by joint inertia
Initialize actor π_θ with auxiliary head h_θ, and critic V_ψ
repeat
  for $t = 1$ **to** T, in parallel over environments **do**
    $\boldsymbol{o}_t \leftarrow \textsc{Obs}\big(\boldsymbol{q}_t, \dot{\boldsymbol{q}}_t,\ \boldsymbol{q}^*_{t+\Delta},\ \hat{\boldsymbol{F}}_{t+\Delta}, \hat{c}_{t+\Delta}\big)$ ▷ no force-plate input
    $\boldsymbol{a}_t \sim \pi_\theta(\cdot \mid \boldsymbol{o}_t)$
    $(\boldsymbol{k}_t^p, \boldsymbol{k}_t^d, \boldsymbol{q}_t^{\text{cmd}}) \leftarrow \textsc{Decode}(\boldsymbol{a}_t;\ \boldsymbol{k}_0^p, \boldsymbol{k}_0^d, \boldsymbol{q}^*_{t+1})$ ▷ gain lag
    Simulate with $\boldsymbol{\tau} = \text{clip}\big(\boldsymbol{k}_t^p \circ (\boldsymbol{q}_t^{\text{cmd}} - \boldsymbol{q}) + \boldsymbol{k}_t^d \circ (\dot{\boldsymbol{q}}^*_{t+1} - \dot{\boldsymbol{q}}),\ \boldsymbol{\tau}^{\text{lim}}\big)$
    $r_t \leftarrow r(\boldsymbol{q}_{t+1}, \boldsymbol{F}_{t+1};\ \boldsymbol{q}^*_{t+1}, \boldsymbol{F}^*_{t+1})$ ▷ kinematic + GRF tracking
    Reset terminated episodes via reference-state initialization
  end for
  Estimate advantages $\hat{A}_t$ with GAE using privileged critic $V_\psi(\boldsymbol{o}_t, \boldsymbol{F}_t^*)$
  Update θ, ψ by minimizing $\mathcal{L}_{\text{PPO}} + w_{\text{aux}}\, \ell\big(h_\theta(\boldsymbol{o}_t), \boldsymbol{F}^*_{t+\Delta}\big)$
until converged

Kinematics-to-GRF estimator. The estimator maps reference kinematics to per-foot GRF in two stages, a closed-form physics prior and a learned residual, fitted once per split on the training subjects, shared by the three seeds and then frozen. On the evaluation sets it reaches a vertical-GRF correlation of 0.979 on seen and 0.928 on unseen subjects (anterior–posterior 0.883 / 0.806, mediolateral 0.978 / 0.977). The prior alone reaches 0.94 on the vertical component of the training frames. A contact flag $\hat{F}_{z,f} > 0.1$ BW replaces the plate-derived flag wherever the actor's observations need one.

Physics prior. The whole-body CoM of every reference frame is computed with the subject's model and smoothed per clip with the five-tap binomial kernel $[1, 4, 6, 4, 1]/16$; central differences give its acceleration. The vertical total $\max\big((\ddot{\boldsymbol{c}}+g)/g \cdot \boldsymbol{e}_{\mathrm{up}},\, 0\big)$ is split with $w_f = \mathrm{ramp}(h_{\mathrm{top}} - h_f;\, h_{\mathrm{ramp}}) \cdot \mathrm{ramp}(v_{\mathrm{top}} - v_f;\, v_{\mathrm{ramp}})$, where h_f and v_f are the lowest foot-body height and speed of foot f ($h_{\mathrm{top}} = 0.10$ m, $h_{\mathrm{ramp}} = 0.04$ m, $v_{\mathrm{top}} = 1.0$ m/s, $v_{\mathrm{ramp}} = 0.6$ m/s), normalized across feet with inverse CoM-to-heel horizontal distances. The horizontal components are the vertical force times the horizontal-over-vertical components of the vector from the lowest foot point to the CoM (vertical leg length floored at 0.3 m).

Residual network. Per-frame features, all in the pelvis heading frame with the root translation removed: pelvis height, tilt and list; the actuated joint angles and velocities; root angular velocity; root and CoM linear velocity and CoM height; the four foot-body heights and speeds; the heel offsets from the CoM; the prior (rotated into the heading frame) and its contact weight. A window of twelve offsets spanning ± 40 frames (Table 8), clamped inside the clip, is concatenated and standardized. The network is an MLP with a zero-initialized output layer, so the fit starts at the prior. Targets are $\boldsymbol{F}^* - \boldsymbol{F}^{\mathrm{prior}}$ in the heading frame, trained with a Huber loss and AdamW on the plate-valid frames of the run's training subjects plus their left/right mirror images, with one clip in twenty withheld as a fit diagnostic. The fit takes about two minutes at environment initialization and is saved with the run, so evaluation applies the same frozen estimator to held-out subjects. The ensemble mean is smoothed per clip with the binomial kernel, rotated back to the world frame, and its vertical components are clamped at zero.

Table 7: Actor observation (215 dims). Lookahead offsets are in reference frames at 100 Hz.

Component	Contents	Dims
Proprioception	pelvis height, the three pelvis rotations and the 25 remaining coordinates, actuated and passive (29); all generalized velocities (31)	60
Contact state	per foot: vertical contact force (BW, clamped at 0) and heel height	4
Kinematic lookahead	$\boldsymbol{q}^*_{t+k} - \boldsymbol{q}_t$ for $k \in \{1, 5, 25\}$, clamped at the clip end	93
GRF lookahead	$\hat{\boldsymbol{F}}_{t+k}$, six components, $k \in \{1, 5, 25\}$	18
Foot lookahead	per foot at $k \in \{0, 1, 5, 25\}$: heel height, toe height, heel vertical velocity and acceleration, estimated contact flag	40

Observations. Raw observations are clipped to ± 100 and normalized with running mean and variance (clip ± 5). The privileged features, the measured current-frame GRF (6 dims) and the measured GRF at the three lookahead offsets (18 dims), have their own running normalizer $\mathcal{N}_{\mathrm{priv}}$. The critic input is the concatenation of the normalized observation and the normalized current-frame GRF.

Reward. The kinematic term is $r^{\mathrm{kin}} = 0.65\, r_{\mathrm{pose}} + 0.10\, r_{\mathrm{vel}} + 0.25\, r_{\mathrm{root}}$, with

$$r_{\mathrm{pose}} = \exp\Big(-0.04 \sum_j \frac{(q_j - q_j^*)^2}{\sigma_j^2} - 2 \sum_k \Delta\theta_k^2\Big),$$
$$r_{\mathrm{vel}} = e^{-0.02\,\|\dot{\boldsymbol{q}} - \dot{\boldsymbol{q}}^*\|^2}, \qquad r_{\mathrm{root}} = e^{-5\,\|\boldsymbol{v}_{\mathrm{pelvis}} - \boldsymbol{v}^*_{\mathrm{pelvis}}\|^2}, \tag{15}$$

where $\Delta\theta$ is the pelvis tilt/list/rotation error, the joint sums run over the actuated joints, $\boldsymbol{v}_{\mathrm{pelvis}}$ is the three-dimensional pelvis velocity in world axes, and σ_j is the standard deviation of q_j^* over the clip, clipped to $[0.05, 0.141]$ rad. The end-effector term is $r^{\mathrm{ee}} = \exp\big(-40 \sum_b \|\Delta(\boldsymbol{p}_b - \boldsymbol{p}_{\mathrm{pelvis}})\|^2\big)$ over the heel and toe bodies of both feet, with positions taken relative to the pelvis in world axes: invariant to pelvis drift but sensitive to yaw. CoP errors in Eq. 10 are taken relative to each foot's calcaneus in world axes, and are additionally gated on a valid measured CoP and a loaded simulated foot ($F_{z,f} > 0.05$ BW). The Torque baseline uses the unnormalized kernel $\exp(-2 \sum_j (q_j - q_j^*)^2 - 2 \sum_k \Delta\theta_k^2)$ in place of r_{pose}, since under the amplitude-normalized kernel it reached a joint-angle RMSE of $\approx 42^\circ$.

Humanoid model and contact. The topology of the humanoid carries ten passive coordinates alongside the $m = 15$ actuated joints. The coupled knee translations and rotations are tied to the

knee angle by equality constraints, and one metatarsophalangeal joint per foot is passive, with a linear spring–damper ($k = 25$ N m/rad, $d = 2$ N m s/rad) over the range $[-60°, 0]$. In Eq. 1, $\mathbf{M}(\boldsymbol{q}) \in \mathbb{R}^{n\times n}$, $\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \in \mathbb{R}^n$, $\boldsymbol{\tau} \in \mathbb{R}^m$, $\mathbf{S} \in \mathbb{R}^{m\times n}$, $\boldsymbol{\lambda} \in \mathbb{R}^{3n_c}$ at the $n_c \leq 16$ active contacts and $\mathbf{J}_c(\boldsymbol{q}) \in \mathbb{R}^{3n_c\times n}$; the per-foot ground reaction forces $\boldsymbol{F}_r, \boldsymbol{F}_l \in \mathbb{R}^3$ are the resultants of the contact forces on each foot. The metatarsophalangeal spring–damper torques are absorbed into $\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and the knee coupling constraint forces are omitted. The eight contact spheres per foot (radius 32 mm) build on the six-sphere foot of Falisse et al. (2019) (heel, midfoot, first and fifth metatarsal heads, hallux and lateral toe), adding a posterolateral heel sphere and a third-metatarsal sphere; friction is $\mu = 1.0$, the contact time constant $t_c = 0.04$ s, and a common vertical offset of 5 mm was fitted on standing frames.

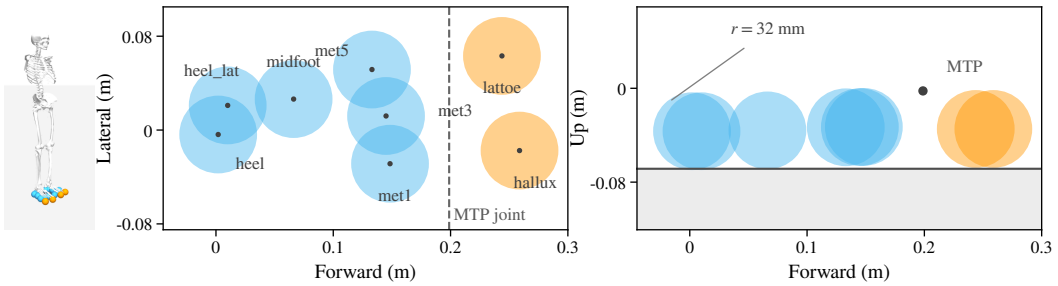


Figure 6: **The simulated body and contact model. (a)** The simulated body on its contact spheres (bone meshes for visualization only). **(b)** Top and **(c)** side view of the eight contact spheres of the right foot: six on the calcaneus body (blue), two on the toes body (orange), the passive MTP joint dashed. Per-foot GRF and CoP are the resultants of these spheres' contacts.

Training. Every reset draws a start frame uniformly from the world's subject's pool, then adds Gaussian noise of 0.05 rad to all non-root coordinates and the pelvis orientation and 0.2 rad/s to all velocities. The lagged gain state of Eq. 6 is carried across these resets and zeroed only at full resets and at evaluation. Episodes terminate when the clip ends, when the pelvis falls below 0.4 m, on a non-finite state, or at $\max_j |\dot{q}_j| > 100$ rad/s. A single Adam optimizer updates actor, critic and auxiliary head. Per minibatch the loss is the clipped surrogate plus $0.5 \cdot \frac{1}{2}(V - R)^2$, minus 0.01 times the policy entropy, plus the auxiliary term of Eq. 8 (Huber $\delta = 1$ in normalized units). Advantages are normalized per minibatch and the gradient norm is clipped at 0.5. Time-outs at the clip end bootstrap the value. The actor's mean head is initialized uniformly in $\pm 3 \times 10^{-3}$ and its log standard deviation at zero.

Implementation. The environment runs all worlds in one MuJoCo Warp data object with zero-copy GPU views of $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$ and the contact wrenches. Rollouts, advantage estimation and PPO updates run on the GPU without host synchronization, and the physics substep block is captured as a CUDA graph. Per-foot forces and CoP are read from the contact wrenches of the heel and toe bodies, $\boldsymbol{F}_f = \sum_b \boldsymbol{f}_b$ and $\boldsymbol{c}_f = (-\tau_y, \tau_x)/\max(F_{z,f}, 1\text{ N})$, where $\boldsymbol{\tau} = \sum_b (\boldsymbol{\tau}_b + \boldsymbol{c}_{\text{root}} \times \boldsymbol{f}_b)$ is the total moment shifted from the root subtree CoM, about which MuJoCo expresses the wrenches, to the world origin. Each production run (2 billion samples) takes 6.6 h of wall-clock time on one NVIDIA H200 with four runs sharing the GPU under MPS, i.e. $\approx$84k environment steps per second per process. Simulation and training hyperparameters are listed in Table 8.

A.2.3 Baseline Implementation Details

GaitDynamics (Tan et al., 2026). We use the force-refinement transformer from github.com/stanfordnmbl/GaitDynamics (6-layer encoder, $d_{\text{model}} = 192$, 8 attention heads, $\sim$2M parameters), which regresses bilateral GRF and CoP from a 150-frame (1.5 s) window of kinematics. The architecture, optimizer (Adan, lr 4×10^{-4}), EMA inference, and training budget (7 680 epochs) are kept verbatim from upstream. The `einops` calls inside the rotary-embedding module are replaced by plain-PyTorch reshapes (numerically identical, parity-tested). We retrain on our 9-dataset corpus covering all six activities rather than AddBiomechanics, using the same per-split subject/clip assignments as our other baselines. The one substantive adaptation is heading

Table 8: ImpedanceMimic hyperparameters.

Category	Parameter	Value
Simulation	Sim timestep h	0.002 s
	Control rate	100 Hz (5 substeps)
	Contact time constant t_c	0.04 s
	Foot model	Falisse 8-sphere (per foot)
	Parallel worlds N	8192
Action space (impedance)	Gain log-scale s_p (k^p)	1.25
	Gain log-scale s_d (k^d)	1.5
	Gain lag T_{act}	50 ms
	Position offset scale α	0.5 rad
	k^p_{nom} (hip / knee / ankle / sub / lum)	400 / 400 / 300 / 100 / 500 N m/rad
	k^d_{nom} (hip / knee / ankle / sub / lum)	40 / 40 / 30 / 10 / 50 N m s/rad
	τ^{lim} (hip / knee / ankle / sub / lum)	275 / 300 / 150 / 50 / 160 N m
Observation	Kinematic lookahead offsets	$\{1, 5, 25\}$ frames
	Foot lookahead offsets	$\{0, 1, 5, 25\}$ frames
	GRF lookahead offsets	$\{1, 5, 25\}$ frames
GRF estimator	Window offsets	{-40,-25,-15,-10,-5, 0, 1, 5, 10, 15, 25, 40}
	Hidden size	1024
	Layers	2
	Ensemble size K	3
	Loss	Huber ($\delta = 0.1$)
	Gradient steps (per member)	4000
Reward weights	w_{kin}	1.0
	w_{grf}	$0.25 \rightarrow 0.5$ (annealed)
	w_{shear}	0.25
	w_{ee}	0.15
Reward scales	α_{pose}	0.04
	α_{vel}	0.02
	α_{root}	5.0
	α_{grf}	4.0
	α_{cop}	5.0
	α_{shear}	5.0
	α_{ee}	40.0
	GRF relative error floor	0.5
	Pose amplitude floor	0.05
Network	Actor	$[512, 256, 128]$, SiLU
	Mean head	Linear$(128 \rightarrow 45)$
	Aux head	Linear$(128 \rightarrow 18)$
	Critic	$[512, 256, 128, 1]$, SiLU
	Action distribution	tanh-Normal, global $\log \sigma$
PPO	Total timesteps	2×10^9
	Unroll length T	20
	Minibatches	32 (5 120 samples each)
	Epochs per collect	4
	Discount γ	0.98
	GAE λ	0.95
	Clip ϵ	0.3
	Learning rate	$3 \times 10^{-4} \rightarrow 3 \times 10^{-6}$ (linear)
	Entropy coeff	0.01
	Value loss coeff	0.5
	Aux loss weight w_{aux}	0.25
	Grad norm clip	0.5
	Optimizer	Adam
RSI and termination	Fall height	0.4 m
	$\dot{q}$ limit	100 rad/s
	RSI noise σ_q	0.05 rad
	RSI noise $\sigma_{\dot{q}}$	0.2 rad/s
	Start-frame sampling	uniform over the subject's frames

alignment: upstream feeds clips with $> 45°$ pelvis heading range unaligned, though the model is trained only on aligned data. We align every clip at inference, which raises lateral GRF tracking from 0.65 to 0.89.

PHC (Luo et al., 2023). We re-implement PHC (github.com/ZhengyiLuo/PHC) on our MuJoCo Warp simulator and Rajagopal character so that it trains and scores on the exact same data, physics, subject models, and evaluation protocol. We use the same observations (544 dims), reward (exponential imitation kernel and AMP discriminator), PPO constants ($\gamma = 0.99$, horizon 32, minibatch 16 384, 6 mini-epochs, lr 2×10^{-5}, fixed $\sigma = \exp(-2.9)$), and the PMCP curriculum (cumulative-failure sampling every 245.76M samples; training episodes terminate at 0.25 m of tracking error, the curriculum's pass/fail evaluation at 0.5 m). The action space is absolute PD targets at mid $\pm\, 0.7\times$range, with PD gains taken from our inertia-scaled priors, since PHC's SMPL gains do not transfer to the Rajagopal skeleton. Three benign divergences are documented: we bootstrap on timeouts (upstream does not), use a single-stage per-subject motion sampler (upstream uses a two-stage subset), and compute reference velocities by finite differences (upstream Gaussian-smooths at $\sigma = 2$ frames). The main systematic difference is the control rate: our 100 Hz policy against upstream's 30 Hz makes the 32-step rollout span 0.32 s instead of 1.07 s, compresses the γ-horizon from 3.3 s to 1 s, and shrinks the AMP window from 0.33 s to 0.1 s. The one term we drop is upstream's joint-power penalty, $-5{\times}10^{-4} \sum_j |\tau_j \dot{q}_j|$ per step. Its coefficient was calibrated on AMASS, a kinematics-only corpus of mostly moderate-speed locomotion. On our corpus the reference joint power $\sum_j |\tau_j \dot{q}_j|$ is $1{,}004$ W in running against 276 W in walking and 2 W in standing, so the penalty reaches -0.5 per step on running and -1.05 at jump take-off against an imitation reward capped at 1. The per-step reward turns negative on running clips, early termination becomes optimal, and training stalls: with the penalty, PHC reaches 57.2% survival, $15.5°$ and $F_{\mathrm{Z}_r} = 0.23$ on seen subjects at 2 B steps, with a 39% curriculum pass rate and an episode length still rising (Table 9). Continuing those runs to 6 B steps raises survival to 69.7% and F_{Z_r} to 0.30, and retraining at the native 33 Hz rate gives $17.6°$ against $15.5°$, so the control rate is not the bottleneck. Without the penalty, the configuration of every PHC row in the paper, the curriculum passes 91–92% of training clips, the episode length is flat after 1 B steps, and the policy reaches 89.9% survival, $11.8°$ and $F_{\mathrm{Z}_r} = 0.42$ at the same 2 B budget.

Table 9: **PHC with the upstream joint-power penalty, at the paper's 2 B budget and continued to 6 B steps** (three seeds, mean$_{\pm\text{std}}$). The paper's PHC rows use the same implementation without the penalty (App. A.2.3). Columns as in Table 10.

		Stability	Kinematics				Ground Reaction Forces						Torques†		
Budget	Activity	Ep.↑ (%)	$q^{\downarrow}$ (°)	$\dot{q}^{\downarrow}$ (°/s)	$p_x{}^{\downarrow}$ (m)	$\mathrm{kin}_r{}^{\uparrow}$	$F_{\mathrm{Z}}{}^{\downarrow}$ (BW)	$F_{\mathrm{Z}_r}^{\uparrow}$	$F_{\mathrm{ML}_r}^{\uparrow}$	$F_{\mathrm{AP}_r}^{\uparrow}$	$\mathrm{CoP}_{\mathrm{ML}}^{\downarrow}$ (mm)	$\mathrm{CoP}_{\mathrm{AP}}^{\downarrow}$ (mm)	$\tau^{\downarrow}$ (Nm/kg)	$\tau_r{}^{\uparrow}$	$\dot{\tau}^{\downarrow}$ (Nm/kg/s)
Seen subjects, held-out clips															
	walk	40.4±7.8	15.71±1.17	123.6±3.1	.12±.01	.40±.03	.47±.01	.27±.03	.17±.03	.04±.02	46±8	82±10	1.01±.03	.07±.02	41.6±6.0
	run	10.5±0.3	19.43±.23	213.9±1.9	.18±.01	.18±.00	.84±.01	.16±.02	.03±.02	.03±.03	146±30	231±18	1.40±.03	.06±.01	63.2±5.9
	stand	96.6±0.3	8.91±.34	10.0±0.3	.03±.00	.07±.00	.08±.01	.12±.06	.54±.18	.03±.02	18±1	53±1	.59±.03	.06±.01	1.5±0.3
2 B	lateral	73.8±2.7	16.67±.44	130.8±0.5	.13±.00	.24±.02	.62±.01	.19±.03	.15±.05	.00±.03	61±35	118±20	1.08±.02	.07±.02	57.6±2.4
	flexion	67.8±0.8	16.35±.83	64.1±1.9	.13±.02	.37±.00	.26±.01	.36±.03	.21±.10	.18±.03	25±1	67±2	.69±.03	.19±.02	16.8±1.4
	jump	54.3±2.1	16.11±.36	96.7±1.5	.15±.01	.33±.01	.41±.01	.30±.02	.24±.09	.07±.01	34±11	91±30	.75±.02	.15±.00	23.5±1.9
	Mean	57.2±2.0	15.53±.45	106.5±0.3	.12±.01	.27±.01	.45±.00	.23±.02	.21±.07	.05±.01	55±9	107±7	.92±.02	.10±.01	34.1±1.2
	walk	92.1±3.0	10.94±.37	131.3±0.5	.05±.00	.63±.02	.45±.00	.37±.01	.27±.02	.12±.01	29±5	66±3	.93±.01	.15±.01	51.2±0.5
	run	15.8±0.4	19.21±.24	212.1±2.2	.17±.01	.18±.03	.82±.00	.19±.02	.00±.03	.05±.01	179±75	209±55	1.34±.02	.10±.01	51.5±4.4
	stand	97.5±0.4	9.25±.54	9.4±0.2	.03±.00	.08±.00	.08±.01	.18±.05	.61±.19	.03±.05	17±2	48±1	.59±.03	.06±.01	1.2±0.2
6 B	lateral	81.1±0.6	15.48±.24	126.0±1.9	.14±.01	.34±.01	.59±.00	.27±.01	.23±.02	.03±.00	97±44	132±24	1.02±.02	.12±.02	49.9±4.4
	flexion	76.0±0.4	14.88±.27	59.7±1.2	.11±.00	.45±.01	.23±.00	.47±.02	.25±.03	.22±.01	32±5	74±7	.65±.02	.25±.03	15.6±0.5
	jump	56.0±0.6	16.55±.34	97.6±3.9	.15±.02	.34±.01	.40±.01	.33±.01	.22±.02	.06±.02	38±17	87±17	.74±.05	.17±.02	22.3±5.2
	Mean	69.7±0.5	14.39±.15	106.0±1.5	.11±.01	.34±.01	.43±.00	.30±.01	.26±.03	.05±.01	65±10	103±11	.88±.01	.14±.01	32.0±2.1
Unseen subjects (H), zero-shot															
	walk	29.6±18.2	18.22±1.70	124.8±4.7	.16±.04	.33±.11	.47±.02	.28±.08	.15±.06	.08±.05	36±8	72±8	.96±.06	.07±.05	33.4±4.7
	run	11.2±0.2	20.04±.50	231.1±3.9	.20±.01	.17±.01	.81±.01	.16±.02	.00±.01	.05±.03	201±49	255±66	1.60±.01	.06±.00	89.9±4.1
	stand	95.6±0.7	8.01±1.45	9.9±0.6	.03±.00	.09±.01	.09±.00	.12±.06	.57±.39	.02±.17	24±4	54±6	.62±.02	.07±.01	2.7±0.6
2 B	lateral	72.7±2.3	18.40±.81	127.2±3.8	.17±.01	.20±.02	.67±.01	.11±.02	.08±.05	.01±.02	135±55	155±10	1.30±.03	.06±.01	81.0±9.4
	flexion	55.7±6.7	17.24±.08	74.5±5.1	.18±.02	.33±.01	.30±.01	.29±.08	.03±.06	.15±.04	47±32	94±19	.84±.04	.11±.03	29.1±2.0
	jump	43.1±4.5	17.75±.92	107.7±2.8	.22±.02	.25±.02	.46±.02	.20±.05	.06±.04	.04±.03	72±25	107±9	.96±.04	.09±.01	51.7±5.7
	Mean	51.3±1.6	16.61±.53	112.5±2.3	.16±.01	.23±.02	.47±.01	.19±.01	.15±.06	.06±.01	86±23	123±12	1.05±.01	.08±.01	48.0±1.5
	walk	80.9±1.1	13.00±.46	134.1±1.7	.08±.01	.58±.01	.45±.01	.38±.01	.26±.02	.05±.01	28±3	63±3	.90±.00	.16±.00	49.4±0.9
	run	13.8±0.9	19.56±.07	220.7±5.4	.16±.02	.19±.01	.78±.01	.19±.02	.02±.09	.07±.01	230±140	292±137	1.48±.07	.09±.02	63.8±14.5
	stand	97.2±0.8	8.45±.78	8.7±1.0	.02±.00	.10±.01	.08±.02	.22±.22	.66±.07	.05±.06	16±1	49±4	.55±.06	.07±.00	1.5±0.3
6 B	lateral	78.4±1.3	17.87±.53	122.6±0.9	.17±.01	.24±.02	.64±.01	.18±.04	.12±.04	.02±.01	64±3	136±15	1.16±.04	.09±.02	59.0±9.2
	flexion	66.0±2.6	15.95±.92	67.8±2.0	.14±.02	.40±.03	.26±.02	.42±.05	.12±.05	.21±.03	33±17	81±5	.75±.02	.19±.03	26.7±1.9
	jump	49.9±1.2	18.25±.66	104.7±3.5	.21±.00	.27±.02	.44±.01	.25±.02	.07±.08	.06±.01	51±19	106±18	.87±.02	.12±.02	38.1±1.4
	Mean	64.4±0.3	15.51±.27	109.8±0.7	.13±.01	.30±.01	.44±.01	.27±.03	.20±.04	.06±.02	71±25	121±26	.95±.01	.12±.01	39.8±2.0

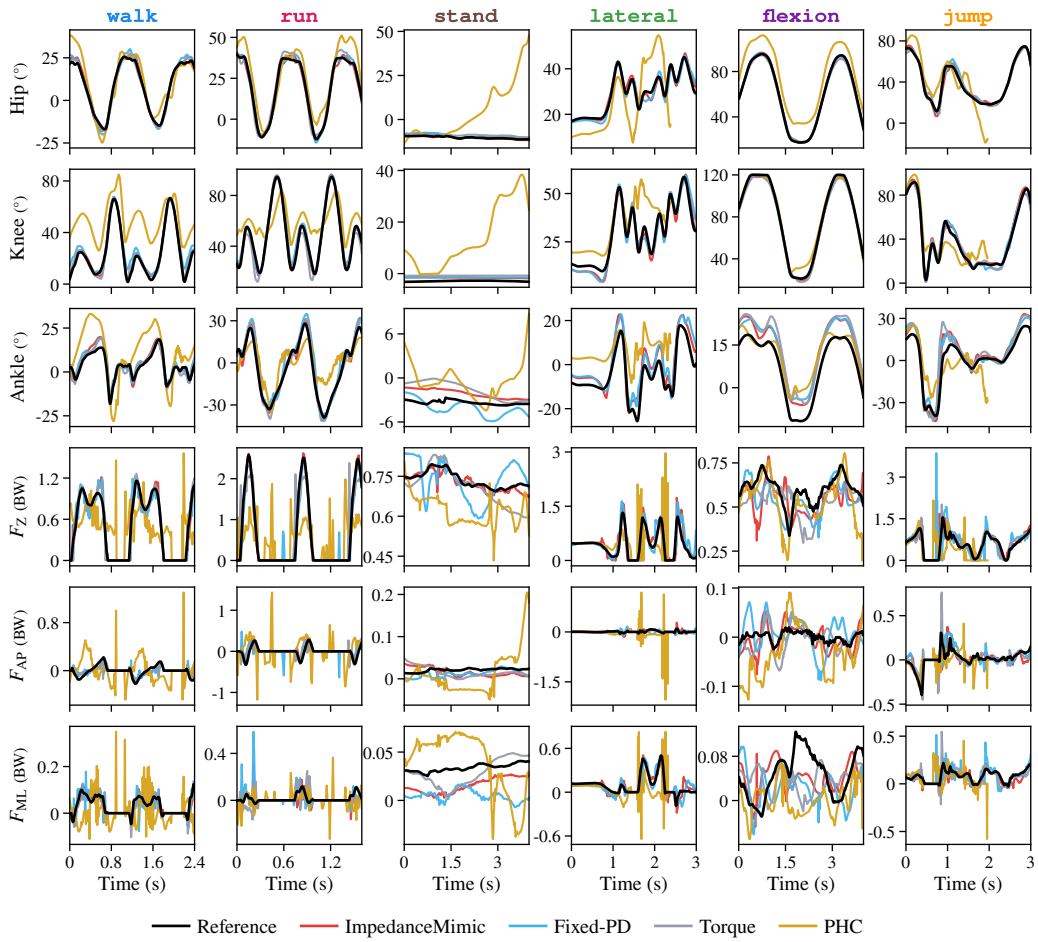


Figure 7: **Reference and simulated trajectories of the four controllers.** One representative clip per activity (columns): right hip, knee and ankle angles, then vertical, anterior–posterior and medio-lateral GRF of the right foot (reference black; a line that stops early is a terminated episode). ImpedanceMimic tracks the reference in every row, Fixed-PD holds the pose but its forces spike at contact transients and wander in standing, Torque recovers the forces but rounds off the pose where contact changes (squat ankle, landing), and PHC drifts in both.

A.3 Results

A.3.1 A Single Policy for Force-Accurate Motion Imitation

The results for seen, held-out clips and zero-shot morphology are presented in Table 10.

A.3.2 Cross-Activity Generalization

The results for cross-activity generalization are presented in Table 11.

A.4 Training Details and Compute

ImpedanceMimic, Fixed-PD, Torque. All three action-space variants train for 2 B environment steps with 8 192 parallel environments in MJWarp (Google DeepMind & NVIDIA, 2025), using PPO with $\gamma = 0.98$, $\lambda_{\text{GAE}} = 0.95$, horizon length 20, 32 minibatches, 4 update epochs per batch, and learning rate 3×10^{-4}. The actor and critic are both [512, 256, 128] MLPs. Each configuration is trained with 3 random seeds across 8 data splits (seen, unseen, and 6 leave-one-activity-out), for

Table 10: **Full metrics for seen- and unseen-subject held-out clips.** *seen* = held-out clips of training subjects, *unseen* = all clips of held-out subjects H. Columns are RMSE (lower better, ↓) unless marked ↑. Subscript r denotes Pearson correlation with the reference signal. Cells report mean$_{\pm\text{std}}$ over 3 seeds (leading zeros are omitted). **Bold** / underline = best / second best within each activity. Ranking is performed on the unrounded values; ties share a rank.

		Stability	Kinematics				Ground Reaction Forces						Torques†		
Activity	Method	Ep.↑ (%)	q↓ (°)	$\dot{q}$↓ (°/s)	p_x↓ (m)	kin_r↑	F_Z↓ (BW)	F_{Z_r}↑	F_{ML_r}↑	F_{AP_r}↑	CoP_{ML}↓ (mm)	CoP_{AP}↓ (mm)	τ↓ (Nm/kg)	τ_r↑	$\dot{\tau}$↓ (Nm/kg/s)
Seen subjects, held-out clips															
	*GaitDynamics**	–	–	–	–	–	*.04±.00*	*1.00±.00*	*.99±.00*	*.99±.00*	*7±1*	*14±0*	–	–	–
	Torque	99.4±.5	2.40±.01	52.9±.9	.29±.02	.93±.00	**.08**±.00	**.98**±.00	.89±.00	.83±.01	**19**±0	**41**±0	.43±.00	.59±.01	8.4±.2
walk	Fixed-PD	99.9±.0	2.51±.01	55.9±.6	.29±.02	.96±.00	.13±.00	.96±.00	.77±.01	.70±.01	21±0	45±0	.55±.01	.48±.01	24.2±1.4
	PHC	98.9±.5	9.78±.24	114±2	**.04**±.00	.70±.01	.47±.00	.33±.01	.15±.00	.23±.01	24±1	73±1	1.49±.01	.10±.00	96.4±2.9
	ImpedanceMimic	**100**±0	**1.70**±.01	**29.6**±.0	.16±.00	**.98**±.00	.09±.00	.98±.00	**.89**±.00	**.84**±.00	20±1	42±0	**.40**±.00	**.62**±.00	**6.3**±.1
	*GaitDynamics**	–	–	–	–	–	*.04±.00*	*1.00±.00*	*.97±.00*	*.99±.00*	*5±0*	*11±0*	–	–	–
	Torque	99.8±.3	3.67±.03	106±2	.76±.07	.93±.00	.23±.00	.96±.00	.42±.02	.57±.03	**21**±0	47±1	.89±.01	.64±.00	**18.1**±.7
run	Fixed-PD	99.8±.2	3.49±.01	104±1	.88±.04	.96±.00	.20±.00	.97±.00	.33±.01	.45±.01	22±1	45±1	1.18±.01	.48±.01	86.9±4.0
	PHC	96.3±.8	11.8±.1	180±1	**.08**±.00	.56±.03	.68±.02	.53±.04	.12±.01	.00±.02	27±2	68±1	1.77±.01	.19±.01	120±8
	ImpedanceMimic	**100**±0	**2.40**±.02	**63.3**±.3	.82±.06	**.98**±.00	**.12**±.00	**.99**±.00	**.61**±.03	**.84**±.01	22±1	**44**±1	**.89**±.01	**.64**±.00	18.9±.9
	*GaitDynamics**	–	–	–	–	–	*.03±.01*	*.78±.04*	*.88±.03*	*.53±.07*	*4±1*	*8±1*	–	–	–
	Torque	99.2±.3	2.30±.03	10.8±.5	**.02**±.00	.38±.01	**.04**±.00	**.69**±.03	.64±.07	**.29**±.02	**13**±0	41±7	**.32**±.01	**.15**±.01	1.0±.1
stand	Fixed-PD	97.8±.3	1.34±.01	10.8±.6	.04±.00	.62±.00	.11±.00	.30±.02	.45±.01	.18±.03	16±3	54±6	.55±.01	.13±.01	6.3±.9
	PHC	97.6±.3	7.74±.60	10.2±.2	.03±.00	.09±.02	.08±.00	.19±.07	.65±.26	.03±.07	17±1	46±2	.60±.04	.08±.00	2.4±.2
	ImpedanceMimic	**99.6**±.3	**1.10**±.02	**4.3**±.2	.02±.00	**.65**±.01	.07±.01	.51±.04	**.82**±.03	.17±.02	**13**±0	**33**±1	.84±.01	.14±.01	**.9**±.1
	*GaitDynamics**	–	–	–	–	–	*.21±.01*	*.90±.01*	*.87±.01*	*.86±.01*	*29±2*	*31±1*	–	–	–
	Torque	96.1±1.6	6.44±.24	139±2	.18±.01	.78±.01	.25±.01	.89±.01	.74±.00	.65±.01	**29**±1	54±2	.61±.00	.50±.00	27.3±1.2
lateral	Fixed-PD	96.0±.5	3.85±.05	71.2±.8	.16±.01	.91±.00	.29±.01	.86±.00	.73±.01	.62±.02	36±14	**54**±5	.69±.00	.51±.01	42.7±1.0
	PHC	94.5±1.3	12.6±.3	123±2	**.08**±.00	.53±.01	.53±.01	.53±.02	.42±.02	.15±.01	31±1	68±4	1.26±.02	.23±.01	110±2
	ImpedanceMimic	**99.8**±.2	**3.18**±.03	**58.6**±1.1	.15±.01	**.94**±.00	**.24**±.01	**.90**±.01	**.79**±.00	**.71**±.01	32±4	57±9	**.57**±.01	**.57**±.01	**21.0**±1.6
	*GaitDynamics**	–	–	–	–	–	*.05±.00*	*.92±.01*	*.95±.01*	*.83±.00*	*9±1*	*20±1*	–	–	–
	Torque	93.8±1.6	5.19±.22	109±7	.13±.00	.79±.01	**.17**±.00	.69±.01	.56±.02	.52±.04	**23**±1	60±2	**.42**±.01	.47±.01	13.0±.9
flexion	Fixed-PD	91.6±.2	2.81±.01	42.0±1.9	.12±.01	.92±.00	.21±.01	.64±.00	.50±.00	.47±.01	25±1	62±2	.48±.01	.45±.01	17.6±.6
	PHC	83.0±2.7	13.3±.1	68.7±2.9	**.08**±.00	.52±.00	.24±.00	.50±.02	.34±.02	.23±.02	34±16	74±15	.76±.02	.27±.00	38.6±2.4
	ImpedanceMimic	**97.6**±1.8	**2.40**±.01	**39.8**±1.4	.13±.01	**.94**±.00	.18±.00	**.71**±.02	**.60**±.02	**.53**±.00	**23**±1	**59**±1	.43±.01	**.50**±.00	**10.6**±.4
	*GaitDynamics**	–	–	–	–	–	*.13±.01*	*.90±.02*	*.85±.01*	*.75±.01*	*35±2*	*51±1*	–	–	–
	Torque	76.9±.6	7.91±.38	168±2	.29±.05	.71±.01	**.28**±.00	**.74**±.01	.54±.01	.39±.03	33±9	71±7	**.56**±.00	**.40**±.01	**24.9**±.3
jump	Fixed-PD	71.7±1.9	3.88±.07	76.4±1.2	.23±.03	.89±.01	.37±.01	.63±.02	.45±.02	.31±.01	27±2	72±4	.87±.02	.30±.01	64.4±1.4
	PHC	68.7±2.0	15.5±.4	110±1	**.11**±.01	.42±.02	.41±.01	.44±.02	.33±.01	.14±.03	36±6	79±8	.96±.01	.21±.01	65.1±2.8
	ImpedanceMimic	**82.9**±1.4	**3.26**±.10	**67.0**±1.4	.24±.02	**.92**±.00	.32±.00	.69±.00	**.57**±.01	**.42**±.01	**27**±1	**66**±2	.66±.02	.38±.01	31.5±2.3
	*GaitDynamics**	–	–	–	–	–	*.08±.00*	*.92±.01*	*.92±.01*	*.82±.01*	*15±1*	*23±0*	–	–	–
	Torque	94.2±.4	4.65±.10	97.5±1.9	.28±.01	.75±.00	.18±.00	**.83**±.00	.63±.01	.54±.01	23±1	52±2	**.54**±.00	.46±.00	15.5±.1
Mean	Fixed-PD	92.8±.4	2.98±.02	60.0±.3	.29±.00	.88±.00	.22±.00	.73±.01	.54±.01	.45±.00	25±3	55±2	.72±.01	.39±.00	40.4±1.3
	PHC	89.9±.7	11.8±.2	101±1	**.07**±.00	.47±.01	.40±.01	.42±.01	.34±.04	.04±.02	28±2	68±3	1.14±.01	.18±.00	71.9±.7
	ImpedanceMimic	**96.6**±.6	**2.34**±.01	**43.8**±.1	.25±.01	**.90**±.00	**.17**±.00	.80±.01	**.71**±.01	**.58**±.01	**23**±1	**50**±1	.63±.01	**.48**±.00	**14.9**±.6
Unseen subjects (H), zero-shot															
	*GaitDynamics**	–	–	–	–	–	*.06±.00*	*.99±.00*	*.98±.00*	*.97±.00*	*11±0*	*18±0*	–	–	–
	Torque	**100**±0	2.85±.07	62.1±1.0	.29±.01	.93±.00	**.09**±.00	**.98**±.00	.87±.01	**.85**±.01	**21**±1	**39**±1	.42±.00	.59±.01	8.9±.2
walk	Fixed-PD	99.6±.1	2.89±.02	61.7±.7	.40±.06	.96±.00	.16±.00	.94±.00	.71±.00	.68±.02	24±0	43±2	.58±.01	.47±.01	29.6±1.6
	PHC	95.2±1.8	10.2±.5	116±3	**.05**±.01	.69±.01	.46±.00	.40±.01	.15±.01	.19±.02	24±1	66±1	1.41±.00	.10±.00	94.9±4.3
	ImpedanceMimic	**100**±0	**2.04**±.02	**33.7**±.3	.19±.01	**.98**±.00	.11±.00	.97±.00	**.87**±.01	.82±.00	23±1	40±1	**.40**±.00	**.62**±.00	**7.0**±.1
	*GaitDynamics**	–	–	–	–	–	*.06±.00*	*1.00±.00*	*.96±.00*	*.98±.00*	*7±0*	*15±1*	–	–	–
	Torque	96.4±2.3	3.79±.03	110±3	.61±.10	.91±.01	.19±.01	.97±.00	.32±.05	.64±.02	**21**±0	**42**±2	**1.02**±.01	**.64**±.01	**17.9**±.2
run	Fixed-PD	98.6±.0	3.38±.05	111±4	.60±.03	.95±.00	.19±.01	.96±.00	.20±.04	.51±.03	21±1	42±1	1.26±.01	.47±.01	80.3±5.4
	PHC	96.8±1.4	12.7±.4	181±3	**.08**±.01	.54±.01	.66±.01	.50±.02	.11±.04	.00±.03	26±1	72±6	1.84±.01	.19±.01	118±3
	ImpedanceMimic	**100**±0	**2.47**±.02	**67.6**±.4	.78±.06	**.98**±.00	**.13**±.00	**.99**±.00	**.44**±.07	**.85**±.01	**21**±0	43±1	1.03±.00	.62±.01	20.1±.7
	*GaitDynamics**	–	–	–	–	–	*.04±.00*	*.65±.03*	*.95±.01*	*.35±.01*	*6±0*	*15±1*	–	–	–
	Torque	**99.9**±.1	2.26±.08	10.4±1.5	**.03**±.00	.39±.02	**.05**±.00	**.68**±.05	.72±.05	**.31**±.06	**12**±0	35±6	**.34**±.00	.16±.00	**1.2**±.1
stand	Fixed-PD	99.2±.2	1.32±.02	6.2±.3	.03±.01	.67±.01	.08±.00	.46±.01	.59±.04	.14±.02	13±1	40±3	.54±.01	**.18**±.00	3.3±.3
	PHC	93.1±4.4	7.30±.45	14.3±4.7	.04±.01	.09±.01	.08±.01	.40±.05	.73±.15	.05±.12	18±1	47±3	.61±.03	.07±.02	4.2±1.0
	ImpedanceMimic	99.7±.2	**.99**±.01	**4.5**±.7	.03±.01	**.67**±.01	.06±.00	.53±.03	**.83**±.01	.20±.04	15±1	**29**±3	.88±.02	.18±.02	1.3±.3
	*GaitDynamics**	–	–	–	–	–	*.23±.02*	*.90±.03*	*.85±.01*	*.84±.03*	*33±2*	*40±3*	–	–	–
	Torque	81.8±1.6	8.11±.68	194±14	.27±.02	.69±.04	.35±.02	.80±.03	.62±.02	.56±.07	35±7	66±4	.84±.03	.39±.02	49.8±3.9
lateral	Fixed-PD	94.1±.5	4.24±.08	85.2±.3	.26±.01	.89±.01	.39±.01	.77±.02	.60±.03	.54±.03	**34**±3	70±9	.96±.02	.41±.02	70.2±2.4
	PHC	88.3±1.1	14.4±.8	122±2	**.13**±.00	.45±.03	.55±.02	.50±.04	.39±.04	.13±.05	42±11	79±12	1.31±.02	.25±.03	114±2
	ImpedanceMimic	**99.7**±.4	**3.63**±.13	**74.3**±5.1	.29±.02	**.92**±.00	**.32**±.03	**.84**±.03	**.69**±.03	**.62**±.03	35±4	**64**±2	**.75**±.03	**.49**±.01	**34.7**±3.7
	*GaitDynamics**	–	–	–	–	–	*.07±.00*	*.91±.01*	*.91±.01*	*.76±.02*	*15±0*	*33±1*	–	–	–
	Torque	74.8±10.0	6.24±.60	158±30	.26±.02	.71±.03	.24±.02	.59±.02	.42±.01	.50±.04	33±14	75±7	.55±.03	.36±.02	24.7±3.1
flexion	Fixed-PD	84.4±4.7	3.24±.13	64.1±1.3	.25±.02	.86±.01	.31±.00	.51±.02	.31±.03	.38±.01	24±2	74±2	.79±.04	.32±.01	54.6±7.1
	PHC	75.7±2.0	14.1±.4	72.0±2.4	**.09**±.00	.46±.02	.27±.01	.46±.04	.27±.03	.28±.02	23±1	72±1	.86±.01	.23±.01	49.7±2.3
	ImpedanceMimic	**99.7**±.2	**2.54**±.02	**47.1**±1.6	.29±.02	**.90**±.01	**.23**±.01	**.66**±.01	**.44**±.02	**.53**±.01	**22**±1	**67**±3	**.53**±.01	**.41**±.01	**14.5**±.7
	*GaitDynamics**	–	–	–	–	–	*.16±.00*	*.87±.01*	*.85±.01*	*.67±.03*	*38±3*	*54±1*	–	–	–
	Torque	54.5±6.2	8.42±1.05	202±6	.54±.05	.68±.03	**.36**±.04	**.66**±.06	.42±.03	**.34**±.05	33±5	**78**±2	**.68**±.04	**.33**±.05	**34.7**±3.8
jump	Fixed-PD	71.9±1.5	4.21±.07	88.0±1.3	.37±.02	.86±.00	.48±.01	.47±.02	.34±.00	.22±.02	**31**±5	80±2	1.05±.01	.20±.01	84.2±3.2
	PHC	53.8±2.6	17.8±1.2	120±4	**.16**±.01	.34±.04	.47±.01	.35±.01	.25±.03	.13±.04	37±8	87±10	1.10±.04	.15±.01	84.5±5.6
	ImpedanceMimic	**89.6**±1.3	**3.42**±.05	**76.4**±2.1	.43±.05	**.91**±.01	.40±.01	.61±.01	**.49**±.01	.33±.01	32±7	79±9	.78±.02	.31±.01	41.0±1.1
	*GaitDynamics**	–	–	–	–	–	*.10±.00*	*.89±.01*	*.92±.00*	*.76±.00*	*18±1*	*29±1*	–	–	–
	Torque	84.6±2.8	5.28±.21	123±4	.33±.01	.72±.01	.21±.01	**.78**±.02	.56±.02	.53±.02	26±1	56±1	**.64**±.01	.41±.01	22.8±1.7
Mean	Fixed-PD	91.3±.7	3.21±.04	69.3±.9	.32±.02	.87±.00	.27±.00	.69±.00	.46±.00	.41±.00	**25**±1	58±2	.86±.00	.34±.00	53.7±1.2
	PHC	83.8±.9	12.7±.4	104±1	**.09**±.00	.43±.00	.42±.00	.44±.02	.32±.04	.05±.03	28±3	71±3	1.19±.01	.17±.00	77.6±.5
	ImpedanceMimic	**98.1**±.2	**2.52**±.01	**50.6**±.8	.34±.02	**.89**±.00	**.21**±.00	.76±.01	**.62**±.02	**.56**±.01	25±0	**54**±2	.73±.00	**.44**±.01	**19.8**±.4

*GaitDynamics is shown in italics for reference and excluded from the ranking, since it is not a controller and serves as a reference for GRF.
†Metrics are computed over 9 main joints: {hip flexion, hip adduction, knee, ankle} left/right, lumbar extension against inverse dynamics.
Ep. = mean survival episode length before termination. q, $\dot{q}$ = joint angles / velocities. p_x = root forward-position error. {kin_r, F_{Z_r}, F_{AP_r}, F_{ML_r}, τ_r} = Pearson correlation w.r.t. reference {17 joint angles, vertical GRF, anterior–posterior GRF, medio-lateral GRF, inverse-dynamics torques}. $\text{CoP}_{AP/ML}$ = center-of-pressure in the anterior–posterior / medio-lateral direction.

Table 11: **Full metrics for leave-one-activity-out benchmark.** Columns are RMSE (lower better, ↓) unless marked ↑. Subscript r denotes Pearson correlation with the reference signal. Cells report mean$_{\pm\text{std}}$ over 3 seeds (leading zeros are omitted). **Bold** / underline = best / second best within each activity. Ranking is performed on the unrounded values; ties share a rank.

		Stability	Kinematics				Ground Reaction Forces						Torques†		
Held-out Activity	Method	Ep.↑ (%)	q↓ (°)	$\dot{q}$↓ (°/s)	p_x↓ (m)	kin$_r$↑	F_{Z}↓ (BW)	$F^{\uparrow}_{\text{Z}_r}$	$F^{\uparrow}_{\text{ML}_r}$	$F^{\uparrow}_{\text{AP}_r}$	CoP$^{\downarrow}_{\text{ML}}$ (mm)	CoP$^{\downarrow}_{\text{AP}}$ (mm)	τ↓ (Nm/kg)	τ_r↑	$\dot{\tau}$↓ (Nm/kg/s)
walk	*GaitDynamics**	–	–	–	–	–	*.12*±.00	*.96*±.00	*.93*±.01	*.91*±.01	*15*±1	*40*±2	–	–	–
	Torque	54.9±1.4	8.00±.13	152±17	1.54±.47	.70±.00	**.34**±.03	**.72**±.03	**.51**±.01	.35±.04	**35**±7	**89**±9	**.68**±.03	.29±.02	**24.1**±2.2
	Fixed-PD	37.9±4.9	4.47±.06	109±2	1.00±.13	.85±.01	.59±.01	.45±.02	.13±.02	.13±.02	89±5	149±5	1.37±.03	.14±.01	149±7
	PHC	38.2±4.3	19.5±.5	148±5	**.28**±.06	.37±.01	.50±.01	.35±.02	.15±.02	.07±.01	37±9	97±9	1.42±.01	.11±.01	117±4
	ImpedanceMimic	**85.4**±1.7	**3.44**±.25	**80.2**±7.2	1.96±.61	**.91**±.01	.40±.03	.67±.04	.48±.02	**.42**±.03	43±6	102±12	.82±.08	**.31**±.03	29.9±5.6
run	*GaitDynamics**	–	–	–	–	–	*.11*±.00	*.99*±.00	*.97*±.00	*.92*±.02	*13*±1	*26*±2	–	–	–
	Torque	62.6±1.3	10.6±.5	186±16	1.23±.17	.65±.01	.40±.01	.84±.01	.26±.11	.30±.06	25±2	**67**±9	.90±.01	.44±.02	24.2±2.3
	Fixed-PD	68.3±13.5	5.58±.40	115±3	1.76±.28	.87±.01	.52±.07	.73±.06	.26±.07	.16±.19	24±3	74±4	1.20±.08	.33±.04	91.3±6.4
	PHC	64.8±4.2	15.4±.8	186±5	**.13**±.02	.38±.03	.70±.02	.43±.03	.24±.03	.08±.04	35±4	82±13	1.71±.02	.12±.02	123±6
	ImpedanceMimic	**99.0**±.5	**3.96**±.36	**72.1**±3.2	2.22±.22	**.94**±.01	**.30**±.04	**.91**±.02	**.45**±.21	**.45**±.08	**23**±2	69±8	**.89**±.03	**.50**±.03	**23.6**±2.7
stand	*GaitDynamics**	–	–	–	–	–	*.07*±.00	*.85*±.02	*.95*±.00	*.40*±.06	*15*±1	*25*±1	–	–	–
	Torque	**96.2**±1.4	4.21±.36	109±26	.69±.28	.68±.05	.21±.06	.59±.07	.52±.02	.10±.01	18±1	60±9	**.44**±.05	.25±.03	17.9±4.3
	Fixed-PD	68.6±2.7	3.12±.03	71.4±2.0	.30±.02	.65±.01	.37±.02	.40±.01	.26±.00	.08±.01	26±4	84±5	.91±.02	.13±.01	100±4
	PHC	44.4±2.2	15.8±.7	81.0±1.9	**.16**±.01	.17±.00	.30±.00	.29±.02	.01±.16	.04±.02	44±34	89±30	.82±.03	.07±.00	49.7±5.9
	ImpedanceMimic	92.3±.6	**2.40**±.02	**36.4**±1.6	.25±.03	**.80**±.01	**.18**±.00	**.64**±.00	**.68**±.02	**.19**±.02	**17**±0	**55**±2	.44±.01	**.26**±.02	**12.9**±.9
lateral	*GaitDynamics**	–	–	–	–	–	*.27*±.01	*.85*±.02	*.80*±.01	*.67*±.02	*31*±2	*42*±4	–	–	–
	Torque	86.3±2.3	7.71±.12	170±2	.22±.01	.73±.01	.30±.01	.82±.01	.69±.02	.50±.01	**28**±0	**57**±3	.66±.00	.44±.01	32.8±1.0
	Fixed-PD	93.6±.3	4.19±.07	75.8±1.4	.19±.02	.90±.00	.35±.01	.77±.02	.66±.02	.47±.02	33±2	69±8	.79±.02	.42±.01	48.6±1.5
	PHC	89.7±.9	13.5±.5	124±2	**.13**±.00	.47±.02	.52±.01	.50±.01	.38±.02	.06±.02	36±6	72±8	1.26±.02	.22±.01	106±3
	ImpedanceMimic	**99.1**±.2	**3.45**±.06	**66.7**±4.1	.17±.00	**.93**±.00	**.30**±.01	**.83**±.01	**.73**±.01	**.55**±.02	34±7	59±7	**.64**±.03	**.48**±.02	**24.6**±1.2
flexion	*GaitDynamics**	–	–	–	–	–	*.11*±.00	*.86*±.02	*.89*±.01	*.70*±.01	*25*±3	*42*±2	–	–	–
	Torque	66.1±3.3	8.88±.25	137±11	.13±.01	.64±.01	**.21**±.01	.64±.01	.47±.04	**.47**±.02	27±1	73±2	**.53**±.02	.36±.02	23.3±4.2
	Fixed-PD	71.1±.8	3.53±.07	**38.8**±1.7	**.10**±.01	.85±.00	.23±.01	.60±.02	.39±.03	.44±.01	**25**±1	71±2	.57±.01	.37±.01	26.1±2.2
	PHC	73.8±1.1	15.8±.5	83.6±4.5	.13±.00	.40±.01	.28±.01	.44±.02	.32±.07	.16±.04	30±7	76±5	.84±.03	.22±.02	47.6±4.5
	ImpedanceMimic	**80.3**±.4	**3.14**±.11	49.5±.8	.16±.01	**.87**±.01	.22±.00	**.65**±.01	**.49**±.02	.44±.02	25±1	**67**±2	.55±.02	**.39**±.02	**18.5**±2.3
jump	*GaitDynamics**	–	–	–	–	–	*.16*±.00	*.87*±.01	*.85*±.01	*.64*±.00	*34*±1	*57*±1	–	–	–
	Torque	70.3±3.6	9.98±.54	203±8	.30±.01	.65±.01	**.33**±.02	**.65**±.02	.49±.03	.31±.02	28±3	69±1	**.61**±.01	**.36**±.01	**26.5**±.5
	Fixed-PD	69.6±1.1	4.10±.02	**78.1**±.5	.23±.01	.88±.00	.38±.01	.59±.01	.42±.04	.27±.01	**25**±0	69±2	.88±.01	.27±.00	61.9±1.0
	PHC	60.7±1.3	16.3±.2	116±2	**.13**±.01	.36±.01	.43±.00	.36±.01	.27±.02	.09±.02	40±8	83±9	1.02±.03	.17±.01	73.9±3.4
	ImpedanceMimic	**82.3**±3.4	**3.69**±.18	78.3±.6	.29±.03	**.90**±.01	.36±.01	.63±.01	**.50**±.02	**.35**±.02	26±1	**68**±2	.71±.04	.35±.03	32.5±3.7
Mean	*GaitDynamics**	–	–	–	–	–	*.14*±.00	*.90*±.01	*.90*±.00	*.71*±.01	*22*±1	*39*±0	–	–	–
	Torque	72.7±.5	8.23±.05	160±6	.68±.07	.67±.00	.30±.01	.71±.01	.49±.01	.34±.01	**27**±1	**69**±3	**.64**±.01	.36±.01	24.8±.6
	Fixed-PD	68.2±2.2	4.17±.06	81.4±.6	.60±.03	.83±.00	.41±.01	.59±.01	.35±.02	.26±.03	37±0	86±2	.95±.01	.28±.01	79.5±.7
	PHC	61.9±.2	16.1±.4	123±1	**.16**±.01	.36±.00	.45±.00	.40±.01	.23±.02	.06±.02	37±4	83±6	1.18±.01	.15±.00	86.1±2.0
	ImpedanceMimic	**89.7**±.6	**3.35**±.09	**63.9**±1.5	.84±.13	**.89**±.00	**.29**±.01	**.72**±.01	**.55**±.04	**.40**±.02	28±2	70±2	.68±.01	**.38**±.00	**23.7**±1.0

*GaitDynamics is shown in italics for reference and excluded from the ranking, since it is not a controller and serves as a reference for GRF.
†Metrics are computed over 9 main joints: {hip flexion, hip adduction, knee, ankle} left/right, lumbar extension against inverse dynamics.
Ep. = mean survival episode length before termination. q, $\dot{q}$ = joint angles / velocities. p_x = root forward-position error. {kin$_r$, F_{Z_r}, F_{AP_r}, F_{ML_r}, τ_r} = Pearson correlation w.r.t. reference {17 joint angles, vertical GRF, anterior–posterior GRF, medio-lateral GRF, inverse-dynamics torques}. CoP$_{\text{AP/ML}}$ = center-of-pressure in the anterior–posterior / medio-lateral direction.

24 runs per method. ImpedanceMimic was trained on NVIDIA H200 GPUs; Fixed-PD, Torque and PHC on NVIDIA RTX PRO 6000 GPUs.

PHC (Luo et al., 2023). PHC uses its own PPO hyperparameters ($\gamma = 0.99$, horizon 32, minibatch 16 384, 6 mini-epochs, lr 2×10^{-5}, fixed $\sigma = \exp(-2.9)$) with a [1024, 512] actor and its AMP discriminator, trained for 2 B steps on the same 8 192 environments (details in Appendix A.2.3).

GaitDynamics (Tan et al., 2026). The force-refinement transformer (2 M parameters) trains for 7 680 epochs ($\sim$24.6 M windows of 150 frames each) using the Adan optimizer (lr 4×10^{-4}) and converges in $\sim$13 min on a single GPU. Full implementation details are in Appendix A.2.3.

Throughput comparison. Table 12 reports training throughput on both GPU types. With 8 192 parallel environments on one GPU, ImpedanceMimic reaches 84 k steps/s (H200) and Fixed-PD / Torque 96–98 k, against the 13 k steps/s that MuscleMimic (Li et al., 2026) reports on an H100 at the same environment count, simulator and control rate. The gap is the cost of the actuator model: our character has 15 hinge joints, whereas MuscleMimic integrates 416 Hill-type muscles at each substep and needs a 50 M-parameter policy (vs. 0.56 M for our actor and critic together), for $\sim$20 h per billion steps against our $\sim$3.3 h.

Training curves. Figure 8 shows training convergence on the seen split. ImpedanceMimic converges fastest, reaching near-plateau reward within 0.5 B steps. Torque starts far below the PD action spaces but closes most of the gap, and Fixed-PD finishes lowest of the three, all still rising gently at 2 B steps. PHC reaches its plateau by 1 B steps with a curriculum pass rate of 91–92%, and its reward per step is not comparable to the other three, which share one reward definition (App. A.2.3). GaitDynamics (supervised) trains for 7,680 epochs.

Table 12: **Training throughput and compute budget.** RL methods train for 2 B steps with 8 192 parallel environments. Throughput is per process, with four (RL rows) or three (PHC) processes sharing one GPU under MPS. GPU-hours cover all 24 runs of a method on the GPU type used for the paper's runs, divided by the processes per GPU. The nine gain-variant runs of Table 15 add ∼11 GPU-h, counted in the total. †20 M-step benchmark.

Method	Steps/s		Hours/run		Total
	H200	RTX 6000	H200	RTX 6000	GPU-h
GaitDynamics	*(supervised)*		*0.2*		*2*
Torque	98 k	116 k	5.7	4.8	29
Fixed-PD	96 k	117 k	5.8	4.7	28
PHC	57 k	74 k	9.8	7.5	60
ImpedanceMimic	84 k	111 k†	6.6	5.0†	40
Total					**∼170**

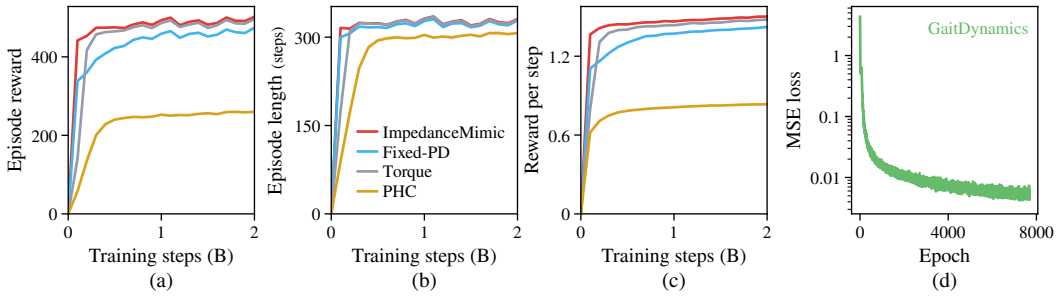


Figure 8: **Training curves** (seen split, mean ± sd over 3 seeds, PHC in its penalty-free configuration). **(a)** Episode reward. **(b)** Episode length. **(c)** Reward per step (ImpedanceMimic, Fixed-PD and Torque share one reward definition, PHC uses its own). **(d)** GaitDynamics MSE loss (log scale).

A.4.1 Learned Impedance as an Encoder of Motor Control

Analyzed set. Gains are the realized k_t^p, k_t^d of the seen-split policy (seed 0) on the 1,697 held-out clips of the 129 training subjects, and of the unseen-split policy (seed 0) on the 1,365 clips of the 22 held-out subjects. Walking and running cycles run from heel strike to the next heel strike of the same leg (contact = ipsilateral vertical GRF above 0.1 BW, released below 0.05 BW), keeping cycles of 0.6–3 s with a stance fraction of 50–75% for walking and 0.5–1 s with 25–55% for running. Standing uses the quiet-stance clips as windows, lateral cycles are the cutting stances, flexion cycles run from descent onset through the deepest point (50%) to the return, and jump cycles are centered on flight (take-off at 35%, landing at 55%). Of 17,067 seen cycles, 16,862 pass: 11,533 walking from 39 subjects, 4,772 running from 27, 355 standing from 61, 48 lateral from 10, 120 flexion from 20 and 34 jumping from 11, plus 13,283 unseen cycles from 20 of the 22 held-out subjects. Every cycle is resampled to 101 samples and its gains are divided by the subject's nominal k_0; Table 13 lists them in physical units.

Per-activity signatures. Standing holds the hip and ankle at 0.6× nominal stiffness and 0.9× nominal damping, with the knee at 1.1× both. Walking modulates the ankle most: its stiffness rises from 0.6× at heel strike to 1.5× at push-off (49% of the cycle) and its damping falls to 0.3× in swing. Running raises knee and ankle stiffness to 1.9× nominal within the first 20% of the cycle, the impact-absorption phase, with ankle damping again at 0.3× in swing. Flexion ramps stiffness with depth, from 0.6× at onset to 0.8–1.0× at the deepest point, and ankle damping rises with it (0.4× to 0.8×). Jumping concentrates stiffness at push-off (ankle 1.4×, just before take-off) and at landing (hip and knee 1.1–1.2×), then returns to sub-nominal values. Lateral cutting loads the hip at foot strike (1.2×) and the knee and ankle through mid- and late stance (1.3–1.4×).

Activity versus subject. A logistic regression on 10 PCs with leave-subject-out cross-validation on the seen held-out clips reaches 92.2% macro accuracy on the balanced subsample of 34 cycles per activity (94.6 and 90.0% for seeds 1 and 2). Walk, run and stand are at 100% recall, lateral at 91%, flexion at 82% and jump at 79% (Table 17), the jump–flexion confusion reflecting deep knee

Table 13: **Learned gains per activity and joint** (seen split, median over cycles of the contact-phase mean, swing-phase mean, and peak of the median waveform). Normalized values are multiples of the subject-specific nominal k_0 of Eq. 7, physical values absolute and per kilogram of body mass.

		k^p/k_0^p			k^d/k_0^d		k^p (N m/rad)		k^p (N m/rad/kg)		k^d (N m s/rad)		k^d (N m s/rad/kg)	
Activity	Joint	contact	swing	peak	contact	swing	contact	peak	contact	peak	contact	swing	contact	swing
walk	hip	0.68	0.72	0.82	0.73	0.71	192	243	2.8	3.5	20.3	20.1	0.29	0.29
	knee	0.58	0.72	0.74	1.03	0.89	163	233	2.4	3.4	28.1	24.0	0.42	0.36
	ankle	0.85	0.65	1.51	0.66	0.29	173	338	2.5	5.0	13.4	6.0	0.20	0.09
run	hip	1.05	1.05	1.53	0.71	0.70	265	402	4.0	6.0	18.5	18.7	0.27	0.27
	knee	1.34	0.97	1.95	0.97	0.88	336	502	5.1	7.6	25.4	23.0	0.37	0.33
	ankle	1.45	0.83	1.86	0.60	0.29	269	370	4.3	5.9	11.7	5.9	0.18	0.09
stand	hip	0.56	0.90	0.59	0.89	0.65	135	153	2.0	2.3	20.9	19.7	0.31	0.27
	knee	1.06	0.75	1.14	1.09	0.85	283	349	4.0	5.1	27.5	26.4	0.42	0.37
	ankle	0.59	0.76	0.66	0.91	0.41	112	134	1.7	2.0	17.3	10.0	0.25	0.13
lateral	hip	0.83	0.93	1.20	0.81	0.70	240	348	3.7	5.4	23.1	19.2	0.35	0.30
	knee	1.02	0.81	1.39	0.95	0.90	305	452	4.7	7.0	32.0	29.8	0.47	0.45
	ankle	1.06	0.95	1.30	0.64	0.39	223	325	3.5	4.7	14.7	8.7	0.22	0.13
flexion	hip	0.83	0.77	0.94	0.87	0.74	235	347	3.3	5.2	25.4	23.1	0.35	0.31
	knee	0.69	0.72	0.79	1.07	0.82	214	385	3.0	5.5	30.5	25.0	0.45	0.33
	ankle	0.86	0.73	0.99	0.64	0.30	184	285	2.7	4.2	13.3	6.7	0.20	0.09
jump	hip	0.77	1.03	1.12	0.75	0.74	233	397	3.4	5.9	21.4	20.9	0.34	0.31
	knee	0.83	1.10	1.22	0.98	0.94	263	533	4.0	8.2	32.2	26.5	0.49	0.43
	ankle	0.82	0.98	1.40	0.54	0.42	179	321	2.9	5.1	12.5	8.8	0.20	0.14

flexion and high ground forces in both. Activity explains $\eta^2 = 0.53$ of the variance in the 10-PC space (Cohen's "large" threshold is 0.14), and a nested decomposition on all 16,862 cycles finds PC1 to be 91% activity, 3% subject and $<2\%$ speed (Table 18). Subject identity within walking is recoverable at 65.9% ($25.7\times$ the 2.6% chance level), so body morphology leaves a trace, but far below the cross-activity accuracy.

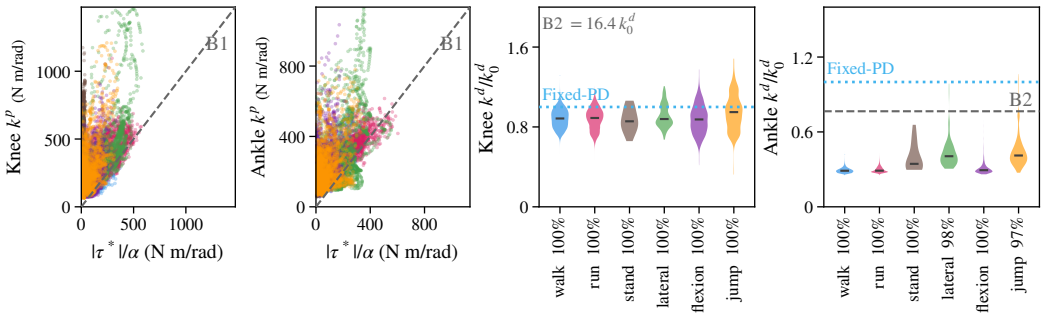
Damping lead. Ankle k^d/k_0^d peaks at 41% of the walking cycle against a mean toe-off at 59.4%: a lead of 18% of the cycle, 186 ms at the median cycle duration, with the peak preceding toe-off in every one of the 11,533 cycles. Running shows the same 18% lead (129 ms) and jumping 8% (120–145 ms before take-off). At toe-off the ankle damping has fallen to $0.34\times$ nominal in walking and $0.37\times$ in running. The realized gains include the 50 ms activation lag of Eq. 6, so the commanded lead is about 50 ms longer.

Bounds on the humanoid. Table 14 and Fig. 9 evaluate both bounds of Section 3.1 frame by frame. Read as predictions, the single-joint bounds say where they should bind on the humanoid, only at the ankle and subtalar joints in swing and at the running knee and ankle in stance, and the counts confirm exactly that pattern: the nominal fixed gains fail precisely there and nowhere else, while the learned policy satisfies both nearly everywhere. Bound 1 uses the inverse-dynamics torque of the reference clip, $|\tau_t^*|/\alpha$, and counts the contact frames on which the realized k_t^p exceeds it. Bound 2 uses the diagonal joint-space inertia I_{jj} of the rest pose and counts the swing frames on which $k_t^d < 2I_{jj}/h$. Because $k_0^d \propto I_{jj}$, the second bound is a fixed multiple of the nominal gain on every subject: $100\times$ at the hip, $16\times$ at the knee, $1.16\times$ at the subtalar joint and $0.77\times$ at the ankle. The nominal ankle damping (30 N m s/rad on the reference subject, whose $2I_{jj}/h = 23$ N m s/rad) therefore violates Bound 2 in every swing phase, the chatter of the nominal variant in Table 15. The learned policy holds the ankle at a median 0.29–$0.41\times$ nominal in swing. Bound 1 binds where the demand exceeds the nominal gain: the median per-cycle demand $\max_t |\tau_t^*|/(\alpha k_0^p)$ is $1.0\times$ at the walking ankle, $1.5\times$ at the running knee and $1.7\times$ at the running ankle, and the learned gain meets it on 92–96% of those contact frames against the nominal gains' 93, 72 and 53%. The strict per-cycle form, $\min_t k_t^p \geq \max_t |\tau_t^*|/\alpha$ over the whole stance, holds for only 1–59% of gait cycles: the policy raises k^p when the demand arrives rather than holding it high throughout. Table 15 turns the bounds into an intervention on these gains. Capping the ankle and subtalar damping at I_{jj}/h — removing the Bound-2 violation and changing nothing else — collapses torque jerk from 130 to 40 N m/kg/s (118 to 6 in standing), cuts the standing force error from 0.71 to 0.11 BW, and raises survival and force correlation on every seen and unseen activity, with survival up on five of the six held-out-activity folds (no-walk $31 \rightarrow 38\%$, no-stand $46 \rightarrow 69\%$). What the cap leaves open is exactly Bound 1's territory (Table 14): doubling every stiffness prior on top closes the running force

gap (F_{Z_r} .97 → .98, F_Z .20 → .14 BW) but not jumping, and torque jerk climbs back from 40 to 76 N m/kg/s, since with the damping pinned at the swing ceiling the stiffer spring is underdamped in contact. No fixed pair meets both phases (Remark 1). The learned schedule does, at close to a third of even the capped variant's jerk.

Table 14: **Action-space bounds on the humanoid** (seen split, learned / nominal fixed gains). Bound 1: share of contact frames with $k^p > |\tau^*|/\alpha$, and the median per-cycle demand $\max_t |\tau^*|/(\alpha k_0^p)$ in units of the nominal gain. Bound 2: share of swing frames with $k^d < 2I_{jj}/h$, a subject-independent multiple of k_0^d since $k_0^d \propto I_{jj}$.

	Bound 2	walk			run			stand			lateral			flexion			jump		
Joint	($\times k_0^d$)	B1 %	$\lvert\tau^*\rvert/\alpha k_0^p$	B2 %	B1 %	$\lvert\tau^*\rvert/\alpha k_0^p$	B2 %	B1 %	$\lvert\tau^*\rvert/\alpha k_0^p$	B2 %	B1 %	$\lvert\tau^*\rvert/\alpha k_0^p$	B2 %	B1 %	$\lvert\tau^*\rvert/\alpha k_0^p$	B2 %	B1 %	$\lvert\tau^*\rvert/\alpha k_0^p$	B2 %
hip	100.27	98 / 100	0.76	100 / 100	99 / 99	0.79	100 / 100	100 / 100	0.04	100 / 100	95 / 97	0.84	100 / 100	94 / 98	0.73	100 / 100	99 / 99	0.64	100 / 100
knee	16.43	97 / 100	0.48	100 / 100	96 / 72	1.50	100 / 100	100 / 100	0.07	100 / 100	97 / 89	0.99	100 / 100	97 / 100	0.56	100 / 100	99 / 100	0.62	100 / 100
ankle	0.77	92 / 93	1.02	100 / 0	93 / 53	1.72	100 / 0	100 / 100	0.19	100 / 0	79 / 76	1.05	98 / 0	99 / 100	0.49	100 / 0	94 / 94	0.88	97 / 0
subtalar	1.16	100 / 100	0.40	100 / 100	66 / 57	1.74	100 / 100	100 / 100	0.15	100 / 100	81 / 82	0.56	100 / 100	99 / 100	0.37	100 / 100	95 / 95	0.95	100 / 100



(a) (b) (c) (d)

Figure 9: **Action-space bounds on the humanoid** (seen split, learned policy; activity colors as in Fig. 3). **(a, b)** Learned k^p against the per-frame demand $|\tau^*|/\alpha$ on contact frames (2,500 per activity), with Bound 1 dashed; the per-activity share of frames above it, learned and nominal, is in Table 14. **(c, d)** Distribution of k^d/k_0^d over swing frames, with Bound 2 dashed, the nominal Fixed-PD gain dotted and the share of swing frames under the bound on the ticks. The nominal ankle damping sits above the bound and never satisfies it in swing.

Table 15: **Fixed-PD gain variants** (mean of the six activity means, 3 seeds, mean ± sd; columns and conventions of Table 10). *Nominal* holds the inertia-scaled priors k_0^p, k_0^d of Eq. 7. The *capped* variant sets $k_0^d \leftarrow \min(k_0^d, I_{jj}/h)$, half of Bound 2 (ankle 30 → 12, subtalar 10 → 6 N m s/rad on the reference subject), changes nothing else, and is the Fixed-PD of Tables 1–11. The third variant doubles every stiffness prior on top of the cap (seen split only).

	Fixed-PD variant	Ep.↑ (%)	$q^{\downarrow}$ (°)	kin_r↑	F_Z↓ (BW)	$F_{Z_r}^{\uparrow}$	$\dot{\tau}^{\downarrow}$ (Nm/kg/s)
Seen	nominal (k_0^p, k_0^d)	89.1±.5	4.24±.03	.76±.00	.44±.00	.58±.00	130±1
	k^d capped at I_{jj}/h	92.8±.4	2.98±.02	.88±.00	.22±.00	.73±.01	40.4±1.3
	capped k^d, 2× stiffness	96.2±.1	2.59±.03	.87±.00	.22±.00	.72±.00	75.8±1.7
Unseen	nominal (k_0^p, k_0^d)	88.3±1.0	4.29±.04	.75±.00	.47±.01	.54±.01	139±3
	k^d capped at I_{jj}/h	91.3±.7	3.21±.04	.87±.00	.27±.00	.69±.00	53.7±1.2

Trade-off accounting. On every contact frame, Eq. 2 makes Proposition 1 an accounting identity: with $\Delta\tau = \max(0, |\tau^*| - |\tau|)$ and damping neglected, $|q^* - q| + \Delta\tau/k^p \geq \max(0, |\tau^*|/k^p - \alpha)$, so any demand the lever cannot supply is paid in pose error, in force shortfall, or in both. Figure 10 evaluates it on the 3.1M sagittal contact frame-joints of the walking and running clips (seed 0): the rollouts satisfy it on 98.9% (Fixed-PD) and 99.98% (ImpedanceMimic) of frames at a 0.02 rad tolerance. The demand exceeds the fixed nominal lever on 2.2% of walking and 26.7% of running frame-joints, so Bound 1 is a running constraint, and on binding frames Fixed-PD pays along both axes while ImpedanceMimic, by raising k^p, keeps the excess small and rare. Per clip, the share of reference frames whose demand exceeds the nominal lever, computable from the reference alone,

predicts Fixed-PD's errors within running (Spearman $\rho = +0.33$ for both q and F_Z RMSE) and not ImpedanceMimic's (-0.24 and -0.17). Bound 2's entry in the ledger is the damping intervention of Table 15: capping k^d alone — the gain Bound 1 does not constrain — collapses the torque jerk and halves the vertical-force error, costs incurred on frames that carry no Bound-1 demand. The two mechanisms separate exactly as the theory says.

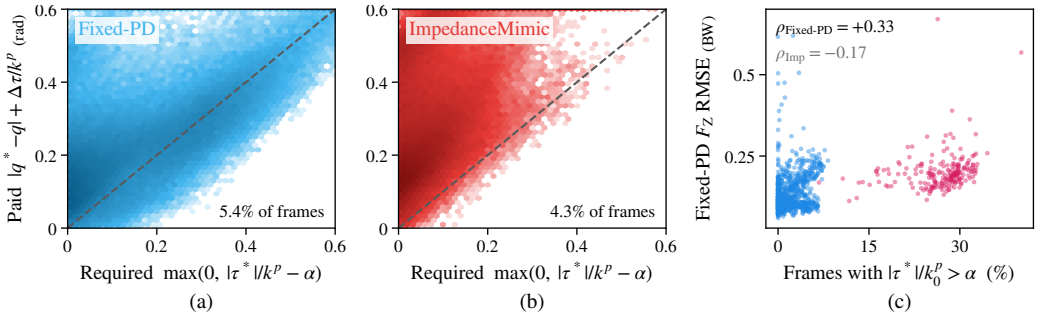


Figure 10: **Bound-1 trade-off accounting** (seen split, seed 0; sagittal joints, contact frames of walking and running). **(a, b)** Paid cost $|q^*-q|+\Delta\tau/k^p$ against the required excess $\max(0, |\tau^*|/k^p-\alpha)$ on the frames where the demand exceeds the lever (share in the corner), identity line dashed: Fixed-PD's binding frames reach large demands and pay accordingly, ImpedanceMimic's stay small. **(c)** Per-clip share of reference frames with $|\tau^*|/k_0^p > \alpha$ against Fixed-PD's vertical-force RMSE, colored by activity (walk blue, run magenta), with within-running Spearman correlations for both controllers.

Cross-projection. Features are standardized on the seen cycles and projected on the first 10 PCs of the seen set (59.8% of variance in two, 75.0% in five); unseen cycles are transformed with the same scaler and basis. The nearest seen-activity centroid in that space recovers the activity of 99.9% of unseen walking cycles, 100% of running, 84.5% of standing, 86.0% of lateral, 38.4% of flexion and 44.8% of jumping cycles. A logistic regression trained on all seen cycles reaches 99.5% overall and 89.1% macro accuracy on the unseen cycles (walk 99.8, run 100, stand 97.9, lateral 88.4, flexion 72.7, jump 75.9% recall). Trained on balanced draws of 34 seen cycles per activity, macro accuracy is $90.5 \pm 1.4\%$ over 20 draws. Over the three seen-split seeds the balanced macro accuracy is $87.6 \pm 2.7\%$ and the overall accuracy $99.4 \pm 0.1\%$.

Gain substitution. At evaluation we replace the policy's gain actions by a frozen schedule and keep its own position offsets: the schedule is the population mean of the seen-split normalized gain cycles above, one 101-point waveform per activity and joint, indexed at runtime by the activity and the movement-cycle phase and rescaled by each subject's nominal k_0. Table 16 compares three conditions on the seen split (seed 0, activity means): the network's own gains, the frozen schedule, and gains pinned at nominal. The schedule is an oracle in two inputs only, the activity label and the phase, yet it retains most of the network's performance, and its losses concentrate on flexion and jumping, where a cycle phase is least well defined. Pinning the gains at nominal collapses gait (22% and 27% survival on walking and running): the modulation, not the nominal level, is what stands the cyclic activities. The gap from nominal to schedule is therefore the information carried by the schedule, and the small remaining gap to the network is the value of per-frame feedback.

Table 16: **Gain substitution** (seen split, seed 0). Cells are survival (%) / F_{Z_r}; q is the activity-mean joint-angle RMSE. Schedule = frozen population-mean gain cycles (activity + phase in, gains out), offsets from the policy; Nominal = gains pinned at k_0.

Gains	walk	run	stand	lateral	flexion	jump	Mean	q (°)
Network	100 / .98	100 / .99	99.3 / .47	99.6 / .90	95.6 / .69	81.4 / .69	96.0 / .79	2.32
Schedule	97.1 / .94	99.6 / .98	92.9 / .44	97.0 / .83	76.5 / .62	70.7 / .65	89.0 / .74	2.90
Nominal	22.2 / .28	27.2 / .58	85.4 / .36	89.4 / .64	62.1 / .34	45.5 / .39	55.3 / .43	5.64

Identifiability of the push-off gain. The PD torque decomposes as $\tau = k^p(q^* - q) + k^p\alpha\,\delta q + k^d(\dot{q}^* - \dot{q})$: a spring about the reference, a feedforward lever set by the position offset, and damping.

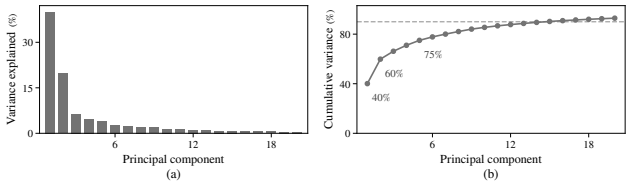


Figure 11: PCA of normalized gain profiles at hip, knee, ankle.

Table 17: Gain-only activity classification (LOSO logistic regression on 10 PCs)

Activity	Recall	Precision	F1
walk	100%	97%	.99
run	100%	94%	.97
stand	100%	100%	1.00
lateral	91%	89%	.90
flexion	82%	82%	.82
jump	79%	90%	.84
Macro	**92.2%**	**92.1%**	**.92**

We reconstruct the three terms from the ankle traces of the 11,533 walking cycles (the offset follows exactly from the recorded torque, gains and states wherever the torque is below its 150 N m limit). In early stance the torque is small and the terms cancel. At mid-stance (25–35% of the cycle, -43 N m) the feedforward term supplies 86%, the spring 15% and damping -1%, with $\alpha\,\delta q = 0.25$ rad against a tracking error of 2.7°. At push-off (44–54%, -95 N m) the feedforward term supplies 109%, the spring 1% and damping -9%, with the offset at 70% of its authority and the torque at its limit on 4% of frames. The gain is a spring constant about the commanded setpoint at all times, but a spring about the reference trajectory only insofar as the spring term carries the torque, which it does not at push-off.

Toppling stiffness. The 61 standing subjects have a mean stature of 1.68 m; placing the center of mass at 0.56 of stature puts it 0.86 m above the ankle, so the gravitational toppling stiffness mgh is 8.4 N m/rad per kilogram. The learned contact-phase ankle stiffness of 1.7 N m/rad/kg per ankle, 3.3 for both, is 0.4 of it.

The no-walk policy on walking. The leave-walking-out policy (seed 0), driven on the walking clips it never trained on, yields 25,327 valid cycles from 21 subjects of the Moore2015 and Wang2023 walking sets. Against the seen-split policy's 11,533 cycles, its median ankle k^d/k_0^d waveform correlates at $r = 0.96$, peaks at 40% of the cycle against 42%, precedes toe-off in 96% of cycles against 100% and settles at the same 0.29× in swing, but reaches 0.49× at its peak against 0.92×. Its k^p/k_0^p waveform correlates at 0.89 and rises from 0.78× in early stance to 1.24× at push-off, against 0.71× to 1.51×, i.e. 58% of the seen policy's push-off rise. The within-walking correlation of contact stiffness with peak GRF is 0.23 against 0.34.

Table 18: Variance decomposition of the gain manifold (nested ANOVA, seen held-out clips).

PC	Activity	Dataset	Subject	Speed	Residual
1	**91.2%**	0.7%	3.3%	1.9%	2.9%
2	**59.7%**	10.1%	14.9%	0.9%	14.5%
3	22.0%	**41.0%**	13.9%	0.1%	23.0%
4	3.6%	3.6%	**38.3%**	14.8%	39.7%
5	**46.4%**	2.3%	13.6%	0.9%	36.8%

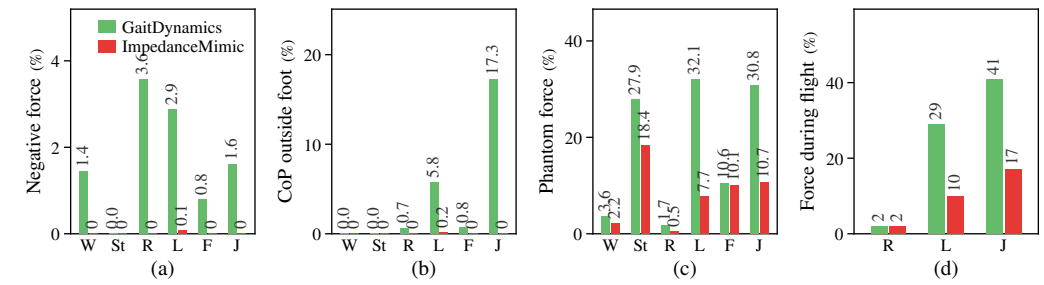


Figure 12: **Physical consistency audit on the unseen split** (1 365 clips, 22 held-out subjects, bars = 3-seed means). **(a, b)** Structurally impossible violations: negative vertical force and center of pressure outside the foot. GaitDynamics violates both on every activity, ImpedanceMimic stays at or below 0.2% of frames. **(c, d)** Contact-timing violations: phantom force on an unloaded foot, and force during measured flight (activities with >100 flight frames). Both methods show nonzero rates, but for the simulator these are timing errors, not force on an airborne body.

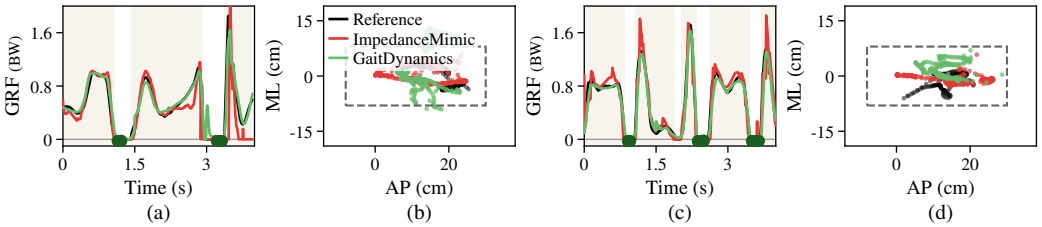


Figure 13: **Per-clip vertical GRF and foot-frame center of pressure** for a held-out jump **(a, b)** and lateral-shuffle **(c, d)** trial, right foot; measured contact phases shaded. GaitDynamics (green) predicts negative vertical force during flight (dark markers) and places the center of pressure well outside the foot box (dashed, -8 to 30 cm AP, ± 8 cm ML). ImpedanceMimic (red) can do neither by construction.

A.4.2 Physical Consistency: GaitDynamics vs. ImpedanceMimic

GaitDynamics is a supervised kinematics-to-GRF regressor: given measured joint angles over a 1.5 s window, it predicts the 12 force/CoP channels that the GRF columns score. It has no body, no contact model, and no notion of momentum, as it fits the force-plate distribution of the training clips. ImpedanceMimic produces the motion in a physical simulation, and the ground-reaction forces are determined by the contact physics. The regressor achieves higher Pearson correlation on the vertical GRF F_{Z_r} on every activity (Table 1). This section asks whether its predicted forces are *physically possible*, and how ImpedanceMimic compares.

Consistency with the contact physics is distinct from accuracy against the measured plates (Tables 10 and 11): a force can satisfy every constraint of the simulated body and still differ from the measurement, so we report the two separately. We score both methods on the unseen split with identical per-frame metrics, reporting mean $\pm$ sd over the three seeds of each method. For GaitDynamics the forces are the model's predictions, and for ImpedanceMimic they are the per-frame contact forces recorded during the evaluation rollouts. We establish two physical-consistency metrics:

1. **Negative vertical force** (unilateral contact): GaitDynamics pulls on the ground on $1.2 \pm 0.3\%$ of foot-frames overall ($3.6 \pm 1.4\%$ in running); ImpedanceMimic on 0.00% of frames on every activity except lateral ($0.08 \pm 0.04\%$) and jump (0.01%), numerical contact transients (Fig. 12a).
2. **CoP outside the foot**: GaitDynamics places the center of pressure beyond the foot box on $5.8 \pm 1.3\%$ of lateral and $17.3 \pm 3.0\%$ of jump loaded frames; ImpedanceMimic on $0.2 \pm 0.1\%$ of lateral frames and none elsewhere (Fig. 12b).

We also evaluate contact timing errors: *phantom force* (force >0.05 BW on a foot the plates show unloaded) and *force during flight* (total force >0.1 BW while both feet are measured as airborne).

- **Phantom force**: GaitDynamics $27.9 \pm 1.0\%$ on stand, $32.1 \pm 1.4\%$ on lateral, $30.8 \pm 2.1\%$ on jump and $3.6 \pm 0.5\%$ on walk, while ImpedanceMimic scores $18.4 \pm 1.5\%$ / $7.7 \pm 0.9\%$ / $10.7 \pm 0.9\%$ / $2.2 \pm 0.3\%$ (Fig. 12c).
- **Force during flight** (activities with >100 flight frames): GaitDynamics predicts $>$ 0.1 BW during $28.7 \pm 5.5\%$ of lateral and $40.7 \pm 3.1\%$ of jump airborne frames, against ImpedanceMimic's $10.3 \pm 1.5\%$ / $17.3 \pm 1.2\%$ (Fig. 12d). On running, both are at 2%.

On steady gait the regressor is both accurate and nearly consistent (walk phantom 3.6%, 2% of running flight frames loaded). Away from gait it produces forces no body could exert: force on the ground during 41% of jump flight frames, center of pressure off the foot on a sixth of loaded jump frames, and vertical force pulling on the ground on 1–4% of frames. Such behaviors are illustrated in Figure 13. The simulated rows satisfy the two structural constraints by construction, up to 0.2% of frames of contact transients. Where they go wrong is *when* contact happens, not *whether* the contact is physical.

A.4.3 Ablations

We ablate ImpedanceMimic by removing each design element individually, training 11 variants on the same data for 2B steps each (3 seeds). Removing the estimator observation or its learned residual also replaces the estimator's contact flag by the kinematic prior's, which agrees with the plates on 96% of frames instead of 99%. Numbers below are means of the six activity means, the convention of Table 1 and of the tables' **Mean** block.

Seen subjects (Table 19). Three elements drive force fidelity on held-out clips of training subjects. Removing the GRF reward causes the largest drop: F_{Z_r} falls from .80 to .69, F_{AP_r} from .58 to .30 and F_{ML_r} from .71 to .48, while kinematics improve (q 2.34→1.94°, kin_r .90→.94): the reward trades kinematic precision for force accuracy. Removing inertia scaling on (k_0^p, k_0^d) raises the joint-velocity error from 44 to 79°/s and the torque jerk from 14.9 to 56.6 Nm/kg/s, concentrated in fast activities (run $\dot{\tau}$: 19→119, lateral: 21→75), and costs .04 in F_{Z_r}, mostly on flexion and jumping ($-.06$ and $-.05$). Removing amplitude normalization reveals the inverse trade-off: q degrades by 0.6° and kin_r drops from .90 to .84, but force tracking improves (F_{Z_r} .80→.85, F_{AP_r} .58→.64), with the gain concentrated on standing ($+.24$). The normalization steers optimization toward low-range joints at a force cost. The end-effector term anchors the pelvis: without it survival falls from 96.6 to 94.7% (flexion -6.8, jump -3.9 points), root drift grows from 0.25 to 0.32 m and the $\mathrm{CoP}_{\mathrm{AP}}$ error by 7 mm. The remaining elements change F_{Z_r} by at most .02 on seen subjects: the privileged training channels, the gain lag, velocity-error damping and the foot lookahead are individually dispensable when the GRF reward is present, and the estimator observation and its learned residual cost .02 each, concentrated on lateral, flexion and jumping.

Unseen subjects (Table 20). The same hierarchy holds on zero-shot morphologies, and two elements matter more there. Without the GRF reward, F_{Z_r} drops from .76 to .66 and F_{AP_r} from .56 to .27. Without inertia scaling, $\dot{\tau}$ rises from 19.8 to 48.0 Nm/kg/s and F_{Z_r} falls by .02, with flexion the most affected ($-.07$). Without amplitude normalization the kinematics–force trade-off persists (q 2.52→3.18°; F_{Z_r} .76→.79). The kinematics-to-GRF estimator matters on new bodies: removing its observation costs .04 in F_{Z_r} (standing $-.10$, lateral $-.05$, flexion and jumping $-.04$) and removing its learned residual .03 (standing $-.11$), while walking and running are unchanged. On unseen subjects the estimator carries the force information that the motion alone does not determine. The end-effector term again holds survival (98.1→95.6%, jump -9.5 points) and root drift (0.34→0.45 m). The privileged channels, velocity damping and foot lookahead remain null ($|\Delta F_{\mathrm{Z}_r}| \leq .01$).

Oracle ceiling. The oracle rows replace the actor's estimated-GRF observation by the measured reference GRF at the same lookahead offsets, with the training recipe otherwise identical. Conditioning on the measured force moves the activity-mean force correlations by at most 0.02, on F_{AP_r} ($0.58 \rightarrow 0.60$ seen, $0.56 \rightarrow 0.58$ unseen), within two seed standard deviations, and leaves kinematics unchanged, while the motion-only policy survives slightly more on unseen subjects (98.1 against 97.4%). The motion-only policy therefore operates at the oracle ceiling of its architecture: force accuracy is limited by the controller and the contact model, not by the force signal available to the actor.

Table 19: **Ablation study — seen subjects, held-out clips.** Each row removes one design element from ImpedanceMimic. Columns are RMSE (lower better, ↓) unless marked ↑. Cells report mean$_{\pm\text{std}}$ over 3 seeds (leading zeros are omitted). **Bold** / underline = best / second best within each activity.

		Stability	Kinematics				Ground Reaction Forces						Torques†		
Activity	Method	Ep.↑ (%)	q↓ (°)	$\dot{q}$↓ (°/s)	p_x↓ (m)	kin$_r$↑	F_{Z}↓ (BW)	F_{Z_r}↑	F_{ML_r}↑	F_{AP_r}↑	CoP$_{\mathrm{ML}}$↓ (mm)	CoP$_{\mathrm{AP}}$↓ (mm)	τ↓ (Nm/kg)	τ_r↑	$\dot{\tau}$↓ (Nm/kg/s)
walk	ImpedanceMimic	**100**±0	1.70±.01	29.6±0.0	.15±.00	.98±.00	.09±.00	**.98**±.00	.89±.00	.83±.00	20±1	42±0	.40±.00	.62±.00	6.3±0.1
	Oracle GRF obs	*100*±0	*1.72*±.01	*29.4*±0.2	*.17*±.01	*.98*±.00	*.09*±.00	*.98*±.00	*.89*±.00	*.84*±.00	*20*±0	*41*±0	*.40*±.00	*.62*±.00	*6.2*±0.1
	- GRF reward	**100**±0	**1.33**±.02	28.5±0.5	**.10**±.00	**.99**±.00	.25±.01	.84±.01	.67±.01	.46±.02	21±1	44±0	.50±.01	.48±.01	10.9±0.4
	- Aux head	**100**±0	1.71±.02	29.3±0.3	.15±.01	.98±.00	.09±.00	**.98**±.00	.89±.00	.84±.00	20±0	41±0	.39±.00	.62±.01	6.3±0.1
	- Aux & critic	**100**±0	1.70±.04	29.3±0.4	.16±.00	.98±.00	.09±.00	**.98**±.00	.89±.00	.83±.01	20±0	41±0	.40±.00	.62±.00	6.4±0.1
	- Gain lag	**100**±0.1	1.86±.02	30.2±0.3	.14±.00	.98±.00	.11±.00	.97±.00	.87±.00	.82±.00	20±1	40±1	**.38**±.00	**.64**±.00	**5.7**±0.2
	- Inertia scaling	**100**±0	1.86±.03	59.2±1.6	.18±.01	.97±.00	.12±.00	.96±.00	.83±.00	.75±.00	21±0	41±1	.58±.01	.51±.00	32.0±1.0
	- Amp norm	**100**±0	1.95±.02	30.5±0.3	.14±.00	.95±.00	**.08**±.00	**.98**±.00	**.90**±.00	**.85**±.00	20±0	42±1	.40±.00	.62±.00	6.2±0.0
	- Estimator obs	**100**±0.1	1.68±.02	29.4±0.3	.16±.01	.98±.00	.09±.00	.97±.00	.88±.01	.84±.00	20±1	41±1	.39±.00	.62±.00	6.0±0.1
	- Learned residual	**100**±0	1.70±.01	29.6±0.1	.15±.01	.98±.00	.09±.00	**.98**±.00	.89±.00	.83±.00	21±1	40±0	.40±.00	.62±.00	6.1±0.1
	- EE reward	**100**±0.1	1.73±.03	**27.7**±0.4	.26±.01	.98±.00	**.08**±.00	**.98**±.00	**.90**±.00	.84±.00	**19**±0	53±1	.42±.00	.59±.00	6.1±0.0
	- Vel damping	**100**±0.1	1.83±.04	33.1±0.4	.15±.00	.98±.00	.09±.00	**.98**±.00	.89±.01	.84±.00	20±0	**39**±1	.40±.00	.60±.01	6.4±0.1
	- Foot LA	**100**±0	1.72±.01	28.8±0.3	.17±.01	.98±.00	.09±.00	**.98**±.00	.89±.00	.83±.01	20±0	42±1	.40±.01	.62±.01	6.0±0.1
stand	ImpedanceMimic	99.6±0.3	1.10±.02	4.3±0.2	**.02**±.00	.65±.01	.07±.00	.51±.04	.82±.03	.17±.02	13±0	33±1	.84±.01	.14±.01	0.9±0.1
	Oracle GRF obs	*99.6*±0.3	*1.07*±.01	*4.4*±0.4	*.03*±.00	*.65*±.00	*.06*±.00	*.56*±.01	*.81*±.03	*.24*±.02	*13*±0	*32*±2	*.84*±.01	*.16*±.00	*0.9*±0.1
	- GRF reward	**99.8**±0	**0.72**±.01	4.1±0.2	.03±.00	**.82**±.00	.09±.00	.33±.01	.37±.04	.00±.01	16±0	29±1	.85±.01	.17±.00	1.2±0.1
	- Aux head	**99.8**±0.1	1.07±.01	4.3±0.3	**.02**±.00	.65±.00	.06±.00	.54±.03	.81±.02	.18±.01	13±1	32±1	.81±.02	.15±.01	0.9±0.1
	- Aux & critic	99.6±0.1	1.06±.03	4.1±0.1	**.02**±.00	.66±.00	.06±.00	.52±.00	.81±.02	.18±.02	13±0	33±1	.83±.00	.15±.01	0.8±0.1
	- Gain lag	99.7±0.1	1.11±.01	4.2±0.1	**.02**±.00	.64±.00	.06±.00	.59±.03	.82±.01	.18±.01	**12**±0	**26**±1	.81±.02	**.19**±.01	0.9±0.1
	- Inertia scaling	**99.8**±0.1	1.10±.02	8.1±0.2	.03±.00	.64±.01	.08±.00	.44±.02	.68±.02	.19±.01	13±1	38±3	.95±.00	.12±.01	3.1±0.2
	- Amp norm	**99.8**±0.1	2.12±.03	4.3±0.1	**.02**±.00	.50±.00	**.04**±.00	**.74**±.02	**.84**±.02	**.38**±.01	**12**±1	42±2	**.58**±.02	.17±.01	**0.6**±0.1
	- Estimator obs	99.7±0.4	1.06±.04	4.1±0.4	.03±.00	.66±.02	.06±.00	.51±.03	.80±.01	.20±.02	13±0	32±1	.84±.01	.15±.01	0.8±0.1
	- Learned residual	99.6±0.1	1.07±.02	4.2±0.2	.03±.00	.66±.01	.07±.00	.50±.01	.80±.02	.19±.02	13±0	33±1	.81±.03	.14±.00	0.9±0.1
	- EE reward	98.6±0.2	1.08±.02	6.7±0.3	.09±.01	.61±.01	.08±.00	.47±.01	.79±.03	.21±.01	13±1	42±1	.83±.03	.13±.00	1.8±0.1
	- Vel damping	**99.8**±0.1	1.06±.01	5.3±0.2	**.02**±.00	.64±.00	.06±.00	.54±.03	.82±.01	.22±.01	13±0	32±1	.83±.01	.15±.01	0.8±0.1
	- Foot LA	**99.8**±0.1	1.07±.02	**3.8**±0.2	.03±.00	.67±.01	.07±.00	.51±.02	.81±.02	.18±.02	13±0	33±1	.83±.02	.17±.01	0.9±0.1
run	ImpedanceMimic	**100**±0	2.40±.02	63.3±0.3	.82±.06	**.98**±.00	.12±.00	**.99**±.00	.61±.03	.84±.01	22±1	44±1	.89±.01	.64±.00	18.9±0.9
	Oracle GRF obs	*100*±0	*2.38*±.04	*62.8*±0.7	*.72*±.08	*.98*±.00	*.12*±.00	*.99*±.00	*.57*±.04	*.85*±.00	*22*±1	*43*±1	*.87*±.01	*.65*±.01	*18.5*±0.7
	- GRF reward	**100**±0	**2.20**±.03	**59.0**±1.8	**.26**±.01	**.98**±.00	.23±.01	.96±.01	.40±.02	.13±.05	**21**±0	**43**±2	.99±.01	.59±.01	19.9±0.3
	- Aux head	**100**±0	2.38±.05	62.9±0.9	.76±.05	**.98**±.00	.12±.00	**.99**±.00	.59±.04	.84±.00	22±0	**43**±1	.88±.00	.64±.00	19.1±0.9
	- Aux & critic	**100**±0.1	2.40±.02	62.7±2.5	.83±.05	**.98**±.00	.12±.00	**.99**±.00	.61±.00	.85±.01	22±1	44±1	.89±.02	.64±.01	20.6±3.1
	- Gain lag	**100**±0	2.30±.02	64.1±1.9	1.17±.07	**.98**±.00	.12±.00	**.99**±.00	.57±.04	.84±.01	22±1	44±1	.88±.01	.64±.01	22.0±1.5
	- Inertia scaling	**100**±0.1	2.59±.01	141.0±3.3	.86±.06	.97±.00	.16±.00	.98±.00	.41±.01	.78±.00	22±1	45±2	1.33±.02	.45±.01	119.2±2.6
	- Amp norm	**100**±0	2.59±.03	64.2±0.5	.80±.05	.96±.00	**.11**±.00	**.99**±.00	.63±.02	.84±.00	**21**±0	44±1	.90±.02	.63±.01	19.2±0.8
	- Estimator obs	**100**±0	2.37±.01	62.6±1.5	.82±.10	**.98**±.00	.12±.00	**.99**±.00	.56±.04	.83±.00	**21**±1	**43**±1	.88±.00	.64±.01	18.4±1.1
	- Learned residual	**100**±0	2.38±.04	63.6±1.6	.82±.07	**.98**±.00	.12±.00	**.99**±.00	.59±.05	.84±.01	**21**±0	**43**±1	.89±.01	.64±.01	19.3±1.3
	- EE reward	**100**±0	2.43±.02	60.5±0.4	.85±.04	**.98**±.00	**.11**±.00	**.99**±.00	**.64**±.01	**.87**±.01	22±1	46±1	.90±.02	.64±.01	**18.1**±0.5
	- Vel damping	**100**±0	2.41±.02	67.4±1.5	.70±.02	**.98**±.00	.12±.00	**.99**±.00	.55±.04	.84±.01	**21**±0	**43**±2	.91±.01	.63±.00	20.0±1.3
	- Foot LA	**100**±0	2.41±.03	62.7±1.0	.86±.06	**.98**±.00	.12±.00	**.99**±.00	.63±.01	.84±.01	**21**±1	44±1	**.87**±.01	**.65**±.01	18.4±1.0
lateral	ImpedanceMimic	99.8±0.2	3.18±.03	58.6±1.1	.15±.01	.94±.00	.24±.01	.90±.01	.79±.00	.71±.01	32±4	57±9	.57±.01	**.57**±.01	21.0±1.6
	Oracle GRF obs	*99.9*±0.1	*3.18*±.03	*60.5*±1.0	*.16*±.00	*.94*±.00	*.23*±.00	*.90*±.00	*.79*±.00	*.71*±.01	*29*±1	*52*±5	*.57*±.00	*.57*±.00	*20.4*±0.9
	- GRF reward	99.9±0.1	**2.75**±.03	65.8±2.0	**.12**±.01	**.95**±.00	.36±.00	.79±.00	.66±.02	.50±.01	**29**±1	59±3	.72±.01	.46±.01	29.9±1.2
	- Aux head	99.9±0.2	3.15±.06	60.3±2.4	.15±.01	.94±.00	.25±.00	.89±.00	.78±.01	.69±.02	35±4	59±3	.59±.02	.55±.01	22.1±0.2
	- Aux & critic	99.8±0.3	3.12±.06	59.8±1.1	.14±.00	.94±.00	.24±.01	.90±.00	.79±.02	.70±.02	**29**±1	53±2	.56±.01	**.57**±.01	20.9±1.3
	- Gain lag	99.6±0.5	3.24±.02	60.9±1.2	.16±.00	.93±.00	.26±.01	.88±.01	.78±.00	.68±.01	**29**±2	**52**±2	.59±.01	.55±.00	23.5±1.2
	- Inertia scaling	99.7±0.5	3.41±.04	102.3±1.5	.15±.01	.93±.00	.29±.01	.87±.01	.72±.01	.63±.02	33±8	56±7	.90±.03	.43±.01	74.9±4.4
	- Amp norm	99.8±0.2	3.80±.07	63.5±1.0	.15±.01	.90±.00	**.23**±.00	**.91**±.01	**.80**±.00	**.72**±.01	**29**±2	**52**±3	**.55**±.02	**.57**±.02	**20.3**±2.0
	- Estimator obs	99.8±0.2	3.38±.06	60.8±1.8	.16±.01	.93±.00	.28±.01	.86±.01	.75±.01	.64±.01	31±4	56±2	.60±.00	.53±.00	21.1±1.0
	- Learned residual	99.9±0.2	3.35±.04	62.0±1.7	.15±.01	.93±.00	.27±.01	.87±.01	.76±.01	.66±.00	30±3	56±5	.59±.01	.54±.01	22.5±2.6
	- EE reward	99.7±0.1	3.27±.03	**58.1**±1.9	.18±.02	.94±.00	.24±.00	.89±.00	.78±.01	.70±.02	**29**±3	57±6	.59±.01	.54±.01	20.6±1.7
	- Vel damping	99.9±0.1	3.35±.03	65.8±1.7	.16±.01	.93±.00	.24±.01	.90±.01	.78±.03	.69±.01	31±3	58±14	.58±.01	.56±.00	21.5±0.5
	- Foot LA	**100**±0.1	3.14±.06	58.5±1.9	.14±.01	.94±.00	.26±.01	.89±.01	.79±.01	.69±.02	31±5	59±11	.56±.01	**.57**±.01	21.1±1.0
flexion	ImpedanceMimic	97.6±1.8	2.40±.01	39.8±1.4	.13±.01	.94±.00	.18±.00	.71±.02	.60±.02	.53±.00	**23**±1	59±1	.43±.01	.49±.00	10.6±0.4
	Oracle GRF obs	*97.8*±1.5	*2.38*±.03	*39.6*±2.9	*.13*±.01	*.93*±.00	*.18*±.01	*.71*±.01	*.60*±.00	*.54*±.02	*23*±1	*60*±0	*.43*±.01	*.50*±.01	*10.5*±1.0
	- GRF reward	**99.8**±0.3	**1.83**±.05	38.2±1.4	.14±.00	**.96**±.00	.23±.00	.62±.01	.36±.01	.38±.01	**23**±0	61±1	.51±.01	.45±.01	14.4±0.8
	- Aux head	98.4±2.2	2.38±.07	39.9±1.8	.12±.00	.94±.00	.18±.00	.70±.00	.58±.00	.53±.01	**23**±0	**58**±1	.42±.01	.51±.01	10.2±0.6
	- Aux & critic	97.9±1.4	2.42±.03	42.7±1.6	.14±.03	.94±.01	.18±.01	.70±.01	.57±.01	.53±.00	24±1	59±1	.43±.01	.49±.00	10.9±0.2
	- Gain lag	96.6±0.9	2.66±.07	46.0±3.0	.17±.02	.93±.00	.19±.01	.69±.02	.56±.01	.48±.01	**23**±0	**58**±1	.47±.02	.48±.02	13.2±1.2
	- Inertia scaling	98.1±1.0	2.64±.09	60.7±3.1	.23±.01	.92±.00	.23±.01	.65±.01	.49±.01	.39±.02	24±1	63±2	.62±.02	.40±.01	28.6±2.2
	- Amp norm	99.2±0.2	3.28±.07	39.5±0.9	**.09**±.01	.89±.00	**.16**±.00	**.73**±.00	**.66**±.01	**.58**±.01	**23**±0	**58**±1	**.39**±.01	**.52**±.02	**7.9**±0.5
	- Estimator obs	97.6±1.1	2.44±.04	41.6±2.3	.13±.00	.93±.00	.19±.00	.68±.01	.55±.01	.51±.00	24±0	59±1	.44±.01	.49±.01	10.8±0.1
	- Learned residual	96.7±1.1	2.47±.06	40.0±1.7	.12±.02	.93±.00	.18±.00	.69±.01	.57±.03	.53±.00	**23**±1	60±0	.43±.01	.49±.00	10.3±0.6
	- EE reward	90.8±4.9	2.67±.11	**35.7**±3.0	.23±.02	.91±.01	.20±.01	.68±.01	.57±.00	.51±.02	**23**±0	65±2	.45±.01	.47±.02	9.8±1.1
	- Vel damping	97.9±0.5	2.46±.07	44.5±2.4	.13±.01	.93±.00	.18±.01	.70±.00	.57±.03	.50±.00	**23**±1	**58**±1	.45±.01	.49±.01	11.3±0.6
	- Foot LA	99.0±0.6	2.41±.07	40.9±1.1	.14±.02	.94±.00	.18±.00	.70±.01	.59±.01	.52±.02	**23**±0	59±2	.43±.02	.51±.02	10.2±0.4
jump	ImpedanceMimic	82.9±1.4	3.26±.10	67.0±1.4	.24±.02	.92±.00	.32±.00	.69±.00	.57±.01	.42±.01	27±1	66±2	.66±.02	.38±.01	31.5±2.3
	Oracle GRF obs	*84.0*±1.7	*3.24*±.08	*66.9*±2.6	*.26*±.02	*.92*±.01	*.32*±.01	*.70*±.01	*.59*±.04	*.43*±.01	*26*±1	*66*±2	*.64*±.02	*.40*±.01	*30.3*±1.1
	- GRF reward	83.7±0.9	**2.79**±.04	66.2±1.8	.25±.03	**.94**±.00	.39±.00	.61±.01	.39±.03	.32±.01	27±2	67±3	.72±.02	.36±.01	35.0±2.4
	- Aux head	83.8±2.8	3.30±.10	68.0±1.5	.24±.01	.91±.01	.33±.01	.69±.02	.55±.03	.41±.01	27±3	68±3	.67±.02	.39±.01	33.5±3.2
	- Aux & critic	83.5±0.4	3.27±.15	70.6±2.1	.25±.01	.91±.00	.33±.01	.68±.01	.55±.00	.40±.01	26±1	68±1	.67±.02	.37±.01	32.6±2.5
	- Gain lag	**86.0**±1.2	3.18±.05	**66.0**±1.5	.25±.02	.92±.00	.32±.01	.70±.01	.55±.02	.41±.01	26±0	**65**±1	.65±.01	.39±.01	30.7±1.3
	- Inertia scaling	81.8±0.5	3.55±.04	99.5±3.1	.28±.00	.91±.00	.37±.01	.64±.02	.47±.02	.32±.01	27±1	70±2	1.03±.01	.28±.01	81.8±2.2
	- Amp norm	84.4±0.6	4.07±.10	69.8±6.5	**.21**±.03	.86±.01	**.29**±.01	**.72**±.02	**.63**±.01	**.43**±.01	27±2	69±4	**.63**±.03	**.40**±.01	31.1±3.5
	- Estimator obs	82.4±2.8	3.41±.10	68.9±2.3	.25±.01	.92±.00	.36±.01	.65±.01	.53±.01	.38±.02	29±5	77±15	.69±.04	.37±.02	33.9±4.9
	- Learned residual	82.8±1.8	3.35±.08	67.7±1.3	.27±.01	.91±.00	.35±.01	.66±.01	.54±.03	.40±.01	26±0	68±1	.68±.02	.37±.01	33.6±1.7
	- EE reward	79.0±0.7	3.36±.06	66.2±2.3	.32±.06	.91±.00	.36±.00	.66±.00	.51±.01	.39±.03	**25**±1	77±1	.70±.02	.34±.01	32.4±3.0
	- Vel damping	84.8±3.1	3.30±.01	71.1±3.2	.24±.04	.91±.00	.32±.02	.71±.01	.56±.02	.41±.01	26±1	**65**±2	**.63**±.01	**.40**±.00	**29.2**±1.4
	- Foot LA	85.0±0.4	3.20±.08	67.5±1.1	.24±.03	.92±.00	.32±.01	.71±.02	.59±.01	.42±.01	26±1	67±1	.64±.03	**.40**±.02	30.8±3.7
Mean	ImpedanceMimic	96.6±0.6	2.34±.01	43.8±0.1	.25±.01	.90±.00	.17±.00	.80±.01	.71±.01	.58±.01	23±1	50±1	.63±.01	.48±.00	14.9±0.6
	Oracle GRF obs	*96.9*±0.6	*2.33*±.02	*43.9*±0.2	*.24*±.02	*.90*±.00	*.17*±.00	*.81*±.00	*.71*±.01	*.60*±.01	*22*±0	*49*±1	*.63*±.00	*.48*±.00	*14.4*±0.2
	- GRF reward	97.2±0.1	**1.94**±.00	43.6±0.3	**.15**±.00	**.94**±.00	.26±.00	.69±.01	.48±.02	.30±.01	23±0	51±1	.72±.00	.42±.00	18.6±0.4
	- Aux head	97.0±0.8	2.33±.03	44.1±0.6	.24±.01	.90±.00	.17±.00	.80±.01	.70±.00	.58±.01	23±1	50±1	.63±.00	.48±.00	15.3±0.7
	- Aux & critic	96.8±0.2	2.33±.03	44.9±1.0	.26±.01	.90±.00	.17±.00	.79±.00	.70±.00	.58±.01	22±0	50±0	.63±.01	.47±.01	15.4±1.2
	- Gain lag	97.0±0.2	2.39±.01	45.2±0.1	.32±.01	.90±.00	.18±.00	.80±.01	.69±.01	.57±.00	22±1	**48**±0	.63±.01	.48±.00	16.0±0.5
	- Inertia scaling	96.5±0.2	2.53±.02	78.5±1.8	.29±.01	.89±.00	.21±.01	.76±.01	.60±.01	.51±.01	23±1	52±2	.90±.00	.37±.00	56.6±0.5
	- Amp norm	97.2±0.1	2.97±.02	45.3±1.2	.24±.01	.84±.00	**.15**±.00	**.85**±.01	**.74**±.00	**.64**±.00	22±1	51±1	**.58**±.00	.49±.00	**14.2**±0.9
	- Estimator obs	96.6±0.6	2.39±.01	44.6±0.8	.26±.02	.90±.00	.18±.00	.78±.01	.68±.00	.57±.00	23±2	51±3	.64±.01	.47±.00	15.2±1.1
	- Learned residual	96.5±0.4	2.39±.02	44.5±0.7	.26±.02	.90±.00	.18±.00	.78±.00	.69±.02	.57±.00	22±0	50±1	.63±.01	.47±.00	15.5±0.5
	- EE reward	94.7±0.9	2.42±.03	**42.5**±0.4	.32±.01	.89±.00	.18±.00	.78±.00	.70±.01	.59±.00	**22**±0	57±1	.65±.01	.45±.01	14.8±0.4
	- Vel damping	97.1±0.4	2.40±.02	47.9±0.8	.23±.01	.89±.00	.17±.00	.80±.01	.70±.02	.59±.00	22±1	50±3	.63±.01	.47±.00	14.9±0.2
	- Foot LA	**97.3**±0.1	2.32±.02	43.7±0.6	.26±.01	.90±.00	.17±.00	.79±.01	.72±.01	.58±.01	22±1	51±2	.62±.01	**.49**±.01	14.5±0.9

†Metrics are computed over 9 main joints: {hip flexion, hip adduction, knee, ankle} left/right, lumbar extension against inverse dynamics. Ep. = mean survival episode length before termination. q, $\dot{q}$ = joint angles / velocities. p_x = root forward-position error. {kin$_r$, F_{Z_r}, F_{AP_r}, F_{ML_r}, τ_r} = Pearson correlation w.r.t. reference {17 joint angles, vertical GRF, anterior–posterior GRF, medio-lateral GRF, inverse-dynamics torques}. CoP$_{\mathrm{AP/ML}}$ = center-of-pressure in the anterior–posterior / medio-lateral direction.

Table 20: **Ablation study — unseen subjects, zero-shot.** Each row removes one design element from ImpedanceMimic. Columns are RMSE (lower better, ↓) unless marked ↑. Cells report mean$_{\pm\text{std}}$ over 3 seeds (leading zeros are omitted). **Bold** / underline = best / second best within each activity.

		Stability	Kinematics				Ground Reaction Forces						Torques†		
Activity	Method	Ep.↑ (%)	q↓ (°)	$\dot{q}$↓ (°/s)	p_x↓ (m)	kin$_r$↑	F_{Z}↓ (BW)	F_{Z_r}↑	F_{ML_r}↑	F_{AP_r}↑	CoP$^{\downarrow}_{\mathrm{ML}}$ (mm)	CoP$^{\downarrow}_{\mathrm{AP}}$ (mm)	τ↓ (Nm/kg)	τ_r↑	$\dot{\tau}$↓ (Nm/kg/s)
walk	ImpedanceMimic	**100**±0	2.04±.02	33.7±0.3	.19±.01	**.98**±.00	.11±.00	**.97**±.00	.87±.01	.82±.00	23±1	40±1	.40±.00	.62±.00	7.0±0.1
	Oracle GRF obs	*100*±0	*1.97*±.03	*32.8*±0.5	*.19*±.01	*.98*±.00	*.11*±.01	*.97*±.01	*.88*±.00	*.84*±.01	22±0	*43*±1	*.40*±.00	*.62*±.00	*6.6*±0.2
	- GRF reward	**100**±0	**1.55**±.02	32.6±0.1	**.16**±.02	**.98**±.00	.28±.01	.81±.01	.63±.01	.49±.01	23±0	45±1	.52±.00	.47±.00	12.4±0.0
	- Aux head	**100**±0	2.07±.03	33.9±0.7	.20±.01	**.98**±.00	.11±.00	**.97**±.00	.87±.01	.82±.01	**22**±0	42±2	.40±.01	.62±.01	6.8±0.3
	- Aux & critic	**100**±0	2.01±.02	33.5±0.2	.19±.01	**.98**±.00	.11±.01	**.97**±.00	.87±.01	.83±.01	23±1	41±2	.40±.00	.62±.00	6.8±0.2
	- Gain lag	**100**±0	2.25±.03	35.9±1.5	.24±.03	.97±.00	.13±.01	.95±.01	.85±.01	.79±.02	23±1	40±1	**.39**±.01	**.63**±.01	6.6±0.4
	- Inertia scaling	**100**±0	2.28±.03	73.0±2.9	.24±.04	.97±.00	.15±.01	.94±.01	.80±.01	.72±.03	23±0	40±2	.61±.01	.50±.02	35.2±0.7
	- Amp norm	**100**±0	2.36±.07	34.5±0.6	.18±.02	.95±.01	**.10**±.00	**.97**±.00	.88±.01	**.85**±.00	**22**±0	41±1	.40±.01	.62±.01	6.5±0.2
	- Estimator obs	**100**±0	2.05±.02	34.3±0.1	.21±.01	**.98**±.00	.12±.00	.96±.00	.86±.00	.82±.02	23±1	**39**±2	.40±.01	.61±.01	6.6±0.2
	- Learned residual	**100**±0	2.05±.03	34.0±0.3	.20±.00	**.98**±.00	.11±.00	**.97**±.00	.86±.01	.82±.01	23±0	40±1	.40±.01	.62±.01	6.8±0.2
	- EE reward	**100**±0	2.06±.03	**31.7**±1.1	.37±.04	**.98**±.00	**.10**±.00	**.97**±.00	**.89**±.01	.84±.01	**22**±1	51±2	.42±.01	.59±.01	**6.4**±0.2
	- Vel damping	**100**±0	2.09±.04	38.2±1.0	.21±.02	**.98**±.00	.11±.01	**.97**±.01	.87±.01	.83±.01	**22**±1	40±0	.41±.00	.60±.00	7.0±0.3
	- Foot LA	**100**±0	2.01±.03	33.2±0.3	.23±.01	**.98**±.00	.12±.00	.96±.00	.88±.00	.82±.01	**22**±1	42±1	.40±.00	.61±.00	6.6±0.1
stand	ImpedanceMimic	99.7±0.2	0.99±.01	4.5±0.7	.03±.01	.67±.01	.06±.00	.53±.03	.83±.01	.20±.04	15±1	29±3	.88±.02	.18±.02	1.3±0.3
	Oracle GRF obs	*99.8*±0.1	*.98*±.02	*3.5*±0.1	*.03*±.00	*.70*±.01	*.06*±.00	*.55*±.02	*.79*±.02	*.27*±.03	*14*±0	*30*±1	*.86*±.03	*.21*±.00	*1.0*±0.1
	- GRF reward	99.7±0.1	**0.64**±.01	3.6±0.3	.03±.00	**.84**±.00	.08±.00	.40±.01	.46±.02	.01±.02	16±0	27±1	.92±.01	.21±.00	1.6±0.1
	- Aux head	99.8±0.1	0.97±.01	4.1±0.7	.04±.01	.67±.01	.06±.00	.50±.02	.84±.02	.21±.07	14±1	30±2	.88±.02	.18±.01	1.2±0.1
	- Aux & critic	99.6±0.2	0.98±.01	4.2±0.4	.03±.00	.68±.01	.06±.00	.52±.04	.80±.02	.21±.08	14±0	29±2	.88±.03	.19±.01	1.2±0.1
	- Gain lag	99.7±0.1	1.05±.05	3.7±0.5	.03±.00	.65±.00	.06±.00	.54±.01	.86±.00	.22±.02	13±0	**25**±1	.86±.02	**.22**±.01	1.1±0.2
	- Inertia scaling	99.7±0.2	1.00±.01	5.4±0.4	.03±.01	.69±.01	.07±.00	.51±.03	.83±.05	.12±.11	14±1	34±4	.97±.02	.18±.01	2.3±0.2
	- Amp norm	**100**±0.1	1.92±.07	4.3±1.0	**.02**±.00	50±.02	**.04**±.01	**.61**±.03	.87±.06	**.35**±.09	**11**±0	40±7	**.60**±.01	.21±.02	**0.8**±0.1
	- Estimator obs	99.8±0.1	0.95±.04	4.0±0.3	.03±.00	.69±.02	.06±.00	.43±.02	.76±.01	.22±.05	13±1	33±1	.92±.01	.19±.01	1.0±0.1
	- Learned residual	99.8±0.2	0.97±.02	3.8±0.2	.03±.01	.69±.00	.06±.00	.42±.02	.82±.06	.21±.09	13±1	31±1	.90±.00	.19±.00	1.1±0.1
	- EE reward	99.2±0	1.02±.03	6.4±1.1	.08±.01	.62±.01	.08±.00	.55±.02	.84±.02	.30±.02	13±1	40±4	.86±.01	.15±.01	1.5±0.2
	- Vel damping	99.6±0.2	0.97±.02	5.7±0.7	.03±.01	.66±.01	.06±.00	.53±.01	.85±.05	.21±.05	14±1	28±2	.87±.01	.18±.01	1.2±0.1
	- Foot LA	99.6±0.1	0.96±.03	**3.3**±0.2	**.02**±.00	.69±.00	.06±.00	.52±.03	**.88**±.02	.22±.06	13±0	30±2	.88±.03	.21±.01	1.0±0.1
run	ImpedanceMimic	**100**±0	2.47±.02	67.6±0.4	.78±.06	**.98**±.00	.13±.00	**.99**±.00	.44±.06	.85±.01	21±0	43±1	**1.03**±.00	.62±.00	20.1±0.7
	Oracle GRF obs	*100*±0	*2.44*±.02	*67.8*±1.3	*.70*±.14	*.98*±.00	*.13*±.00	*.99*±.00	*.42*±.04	*.88*±.00	*27*±1	*42*±1	*1.04*±.00	*.61*±.01	*19.6*±0.8
	- GRF reward	**100**±0	**2.24**±.05	64.5±2.0	**.26**±.07	**.98**±.00	.23±.01	.95±.00	.25±.05	.11±.08	**20**±0	**38**±2	1.17±.02	.55±.01	21.0±0.3
	- Aux head	**100**±0	2.44±.06	66.2±1.1	.71±.06	**.98**±.00	.14±.00	.98±.00	.46±.06	.85±.01	21±1	42±1	1.04±.01	.62±.01	20.6±0.2
	- Aux & critic	**100**±0	2.42±.08	65.5±1.7	.73±.15	**.98**±.00	.13±.01	.98±.00	.49±.05	.87±.01	21±1	43±2	**1.03**±.01	**.63**±.01	**18.4**±1.0
	- Gain lag	**100**±0.1	2.28±.04	66.4±3.4	.93±.17	**.98**±.00	.14±.01	.98±.00	.43±.10	.84±.01	21±0	44±1	1.05±.01	.60±.02	23.5±3.6
	- Inertia scaling	**100**±0	2.59±.04	148.3±1.7	.75±.05	.97±.00	.18±.01	.97±.00	.25±.05	.80±.01	21±1	43±2	1.47±.01	.40±.01	127.3±1.1
	- Amp norm	**100**±0	2.69±.05	68.4±0.7	.74±.04	.95±.00	.13±.00	**.99**±.00	.49±.07	.86±.02	21±1	42±1	1.08±.01	.60±.00	19.7±0.6
	- Estimator obs	99.9±0.2	2.48±.03	68.3±2.6	.73±.06	**.98**±.00	.14±.00	.98±.00	**.52**±.04	.87±.01	21±1	43±1	1.04±.01	.61±.00	21.1±1.9
	- Learned residual	**100**±0	2.44±.11	65.1±2.3	.77±.19	**.98**±.00	.14±.01	.98±.00	.46±.08	.85±.01	21±1	41±1	**1.03**±.03	.62±.01	19.4±1.5
	- EE reward	**100**±0	2.46±.04	**62.8**±0.4	.79±.10	**.98**±.00	**.12**±.01	**.99**±.00	.46±.06	**.89**±.00	22±1	46±2	1.07±.01	.60±.00	18.9±0.3
	- Vel damping	99.9±0.1	2.42±.03	71.0±1.1	.62±.09	.97±.00	.15±.01	.98±.00	.37±.05	.83±.01	22±1	44±2	1.08±.01	.59±.01	19.7±0.8
	- Foot LA	99.9±0.1	2.52±.10	67.5±1.5	.92±.18	**.98**±.00	.14±.01	.98±.00	.45±.09	.85±.00	21±1	43±1	1.04±.01	.61±.01	19.9±1.2
lateral	ImpedanceMimic	99.7±0.4	3.63±.13	74.3±5.1	.29±.02	.92±.00	.32±.03	.84±.03	.69±.02	.62±.03	35±4	64±2	.75±.03	.49±.01	34.7±3.7
	Oracle GRF obs	*98.3*±1.3	*3.50*±.13	*72.3*±1.0	*.28*±.04	*.92*±.00	*.30*±.01	*.86*±.01	*.72*±.01	*.63*±.01	*34*±4	*65*±12	*.73*±.02	*.50*±.02	*32.7*±0.6
	- GRF reward	99.7±0.1	**3.37**±.06	84.0±4.0	**.22**±.03	**.94**±.00	.45±.00	.69±.01	.55±.01	.36±.03	37±8	68±7	.95±.04	.39±.02	44.9±3.5
	- Aux head	99.4±0.5	3.57±.07	73.2±0.5	.29±.02	.93±.01	.32±.01	.84±.01	.72±.00	.61±.00	34±5	61±5	.74±.01	.49±.00	32.8±0.7
	- Aux & critic	99.5±0.2	3.53±.07	75.3±1.8	.26±.01	.92±.00	.32±.02	.83±.02	.70±.03	.59±.03	33±2	62±2	.75±.02	.49±.02	34.1±1.5
	- Gain lag	99.6±0.3	3.59±.03	**72.0**±0.4	.24±.02	.92±.00	.32±.01	.83±.01	.71±.02	.59±.02	33±4	63±6	.76±.02	.47±.01	34.2±4.7
	- Inertia scaling	99.2±1.0	3.78±.08	80.6±3.6	.25±.02	.92±.00	.30±.01	.85±.01	.70±.03	.60±.02	33±1	**56**±1	.76±.01	.48±.00	35.9±1.1
	- Amp norm	99.7±0.5	4.37±.10	80.0±2.7	.26±.02	.87±.01	**.29**±.02	**.86**±.01	**.74**±.02	**.64**±.02	32±3	62±4	**.71**±.01	**.50**±.00	**30.8**±1.8
	- Estimator obs	99.7±0.6	3.75±.08	75.8±1.9	.30±.03	.92±.00	.37±.02	.78±.02	.66±.02	.54±.01	36±1	67±2	.79±.02	.46±.02	34.3±1.4
	- Learned residual	**100**±0	3.71±.05	74.9±3.1	.30±.02	.92±.00	.35±.03	.80±.03	.67±.03	.56±.03	36±8	61±4	.76±.03	.48±.01	31.7±3.6
	- EE reward	98.7±0.3	3.80±.08	73.0±0.8	.34±.03	.91±.00	.34±.01	.81±.01	.70±.00	.61±.03	**31**±2	62±4	.76±.01	.46±.01	33.3±2.6
	- Vel damping	99.7±0.3	3.75±.17	78.6±2.1	.29±.04	.91±.01	.31±.02	.84±.02	.69±.02	.62±.02	37±4	64±4	.75±.03	.48±.02	32.7±1.2
	- Foot LA	**100**±0	3.63±.16	74.5±1.5	.28±.02	.92±.01	.34±.02	.82±.03	.67±.04	.57±.04	33±4	63±0	.77±.03	.47±.02	32.9±1.9
flexion	ImpedanceMimic	99.7±0.2	2.54±.02	47.1±1.6	.29±.02	.90±.01	.23±.01	**.66**±.01	.44±.02	.53±.01	**22**±1	67±3	.53±.01	.41±.01	14.5±0.7
	Oracle GRF obs	*99.4*±0.7	*2.56*±.03	*47.1*±3.2	*.26*±.03	*.90*±.01	*.23*±.01	*.66*±.02	*.47*±.04	*.52*±.03	*22*±1	*66*±1	*.54*±.01	*.41*±.01	*14.6*±0.7
	- GRF reward	**99.8**±0.4	**2.06**±.05	**42.4**±2.1	.38±.05	**.93**±.00	.30±.02	.56±.01	.33±.02	.38±.03	23±1	70±2	.64±.01	.36±.01	21.3±1.7
	- Aux head	99.5±0.4	2.50±.02	43.8±1.3	.26±.02	.91±.01	.23±.01	.65±.01	.47±.03	.52±.04	23±1	67±2	.53±.02	.41±.01	13.7±0.7
	- Aux & critic	99.6±0.3	2.50±.01	45.5±2.0	.25±.03	.91±.01	.22±.00	.64±.00	.47±.06	.54±.03	**22**±1	67±2	.52±.01	.41±.01	14.0±0.6
	- Gain lag	98.3±0.9	2.70±.04	49.9±2.5	.33±.03	.89±.01	.24±.00	.64±.01	.47±.02	.47±.04	**22**±0	**65**±1	.55±.01	.39±.01	16.5±0.5
	- Inertia scaling	98.0±0.5	2.88±.02	65.9±1.5	.54±.05	.87±.01	.28±.01	.58±.01	.44±.02	.44±.02	23±1	68±1	.68±.02	.35±.02	27.5±1.1
	- Amp norm	99.6±0.4	3.35±.07	47.9±2.0	**.18**±.00	.85±.01	**.21**±.01	**.66**±.02	**.60**±.04	**.55**±.01	23±1	67±4	**.49**±.01	**.43**±.02	**12.0**±0.3
	- Estimator obs	98.5±0.3	2.63±.04	47.5±3.4	.27±.04	.90±.00	.25±.02	.62±.03	.42±.04	.47±.02	**22**±1	67±2	.56±.02	.39±.01	16.2±1.7
	- Learned residual	99.3±0.2	2.56±.04	49.4±0.8	.31±.02	.90±.01	.25±.01	.61±.02	.41±.03	.52±.02	**22**±1	68±2	.56±.01	.39±.01	16.3±0.5
	- EE reward	95.9±1.7	2.75±.06	54.5±1.0	.56±.08	.88±.01	.29±.02	.56±.02	.42±.02	.48±.01	**22**±1	72±3	.62±.02	.35±.01	18.2±1.3
	- Vel damping	99.0±0.8	2.55±.06	47.7±1.5	.26±.05	.90±.01	.23±.01	.65±.01	.47±.05	.53±.01	23±1	67±1	.54±.02	.40±.01	14.9±0.5
	- Foot LA	99.5±0.5	2.58±.02	47.2±0.6	.32±.07	.90±.01	.24±.00	.63±.01	.44±.05	.52±.04	**22**±1	67±1	.55±.01	.40±.01	15.2±0.1
jump	ImpedanceMimic	89.6±1.3	3.42±.05	76.4±2.1	.43±.05	.90±.01	.40±.01	.61±.01	.49±.01	.33±.00	32±7	79±9	.78±.02	.30±.01	41.0±1.1
	Oracle GRF obs	*86.6*±1.6	*3.51*±.14	*78.0*±3.3	*.40*±.10	*.90*±.00	*.39*±.00	*.61*±.00	*.51*±.03	*.33*±.02	*27*±1	*74*±3	*.77*±.02	*.31*±.00	*39.7*±1.8
	- GRF reward	90.5±0.7	**3.11**±.12	78.0±1.0	.43±.07	**.93**±.00	.47±.01	.53±.02	.35±.01	.27±.02	29±3	73±1	.87±.02	.28±.00	47.0±2.2
	- Aux head	90.8±1.4	3.48±.17	79.7±1.8	.42±.04	.91±.00	.41±.01	.60±.01	.49±.03	.33±.02	28±0	75±3	.80±.02	.29±.01	43.0±3.2
	- Aux & critic	89.4±1.7	3.50±.09	77.3±2.3	.40±.02	.90±.00	.40±.01	.60±.01	.48±.04	.34±.02	31±4	77±8	.78±.03	**.31**±.01	40.1±2.0
	- Gain lag	91.0±1.5	3.41±.04	**71.7**±2.0	.38±.04	.91±.00	.39±.01	.62±.01	.48±.03	.35±.02	28±1	74±2	.76±.04	**.31**±.01	39.3±4.4
	- Inertia scaling	89.0±0.3	3.83±.13	91.6±1.1	.41±.05	.89±.01	.41±.01	.60±.01	.48±.03	.31±.02	29±1	**71**±1	.91±.03	.27±.01	59.9±4.7
	- Amp norm	**92.1**±1.4	4.40±.02	81.8±3.6	**.36**±.02	.85±.00	**.37**±.01	**.65**±.00	**.54**±.01	**.36**±.01	29±2	74±4	**.75**±.03	**.31**±.01	**38.5**±3.7
	- Estimator obs	89.7±2.2	3.64±.11	80.7±3.9	.41±.10	.90±.01	.43±.00	.57±.02	.48±.02	.31±.02	28±1	74±3	.81±.02	.29±.01	43.7±2.7
	- Learned residual	90.0±1.1	3.53±.13	77.8±7.2	.43±.10	.90±.01	.42±.01	.57±.01	.44±.03	.34±.02	30±3	77±2	.80±.06	.29±.02	41.3±5.9
	- EE reward	80.0±5.2	3.70±.26	80.4±3.5	.56±.03	.89±.02	.44±.01	.56±.02	.41±.03	.28±.01	34±9	81±5	.84±.03	.26±.03	45.0±1.9
	- Vel damping	89.6±2.5	3.54±.02	79.6±2.0	.41±.05	.90±.00	.40±.01	.60±.02	.51±.01	.33±.02	**27**±0	74±2	.77±.01	.29±.01	38.9±1.4
	- Foot LA	89.9±2.1	3.45±.08	75.0±2.2	.45±.10	.91±.00	.40±.01	.61±.00	.49±.03	.33±.01	31±6	77±7	.78±.03	.30±.01	39.6±4.3
Mean	ImpedanceMimic	98.1±0.2	2.52±.01	50.6±0.8	.34±.02	.89±.00	.21±.00	.76±.01	.62±.02	.56±.01	25±0	54±2	.73±.00	.44±.00	19.8±0.4
	Oracle GRF obs	*97.4*±0.6	*2.49*±.05	*50.3*±0.5	*.31*±.02	*.90*±.00	*.20*±.00	*.77*±.01	*.63*±.01	*.58*±.00	*24*±1	*53*±2	*.72*±.01	*.45*±.00	*19.0*±0.3
	- GRF reward	98.3±0.1	**2.16**±.02	50.8±0.3	**.25**±.01	**.93**±.00	.30±.00	.66±.00	.43±.01	.27±.02	25±2	53±1	.84±.01	.37±.01	24.7±0.6
	- Aux head	98.3±0.2	2.50±.04	50.1±0.5	.32±.01	.89±.00	.21±.00	.76±.00	.64±.01	.56±.01	24±1	53±1	.73±.00	.44±.00	19.7±0.5
	- Aux & critic	98.0±0.3	2.49±.02	50.2±1.1	.31±.03	.89±.00	.21±.00	.76±.00	.63±.01	.56±.01	24±0	53±1	.73±.00	.44±.01	19.1±0.4
	- Gain lag	98.1±0.3	2.55±.01	**49.9**±0.3	.36±.04	.89±.00	.21±.00	.76±.00	.63±.01	.54±.01	23±1	**52**±1	.73±.01	.44±.00	20.2±1.2
	- Inertia scaling	97.7±0.2	2.72±.01	77.5±0.8	.37±.01	.88±.00	.23±.00	.74±.00	.58±.03	.50±.02	24±0	52±1	.90±.01	.36±.00	48.0±1.3
	- Amp norm	**98.6**±0.3	3.18±.05	52.8±0.7	.29±.01	.83±.00	**.19**±.00	**.79**±.01	**.69**±.01	**.60**±.02	**23**±0	54±2	**.67**±.00	**.45**±.00	**18.1**±0.7
	- Estimator obs	97.9±0.3	2.58±.03	51.8±0.5	.33±.01	.89±.00	.23±.01	.72±.01	.62±.01	.54±.01	24±0	54±1	.75±.01	.42±.00	20.5±0.6
	- Learned residual	98.2±0.2	2.55±.01	50.8±1.3	.34±.03	.89±.00	.22±.01	.73±.01	.61±.01	.55±.02	24±1	53±1	.74±.01	.43±.00	19.4±1.4
	- EE reward	95.6±0.6	2.63±.04	51.5±0.7	.45±.03	.88±.00	.23±.01	.74±.00	.62±.01	.57±.01	24±2	59±1	.76±.00	.40±.00	20.6±0.3
	- Vel damping	98.0±0.4	2.56±.03	53.5±0.5	.30±.01	.89±.00	.21±.00	.76±.00	.63±.01	.56±.01	24±0	53±1	.74±.00	.42±.00	19.0±0.3
	- Foot LA	98.2±0.4	2.52±.04	50.1±0.7	.37±.04	.89±.00	.22±.00	.75±.00	.63±.01	.55±.02	24±1	54±1	.74±.00	.43±.00	19.2±0.6

†Metrics are computed over 9 main joints: {hip flexion, hip adduction, knee, ankle} left/right, lumbar extension against inverse dynamics. Ep. = mean survival episode length before termination. q, $\dot{q}$ = joint angles / velocities. p_x = root forward-position error. {kin$_r$, F_{Z_r}, F_{AP_r}, F_{ML_r}, τ_r} = Pearson correlation w.r.t. reference {17 joint angles, vertical GRF, anterior–posterior GRF, medio-lateral GRF, inverse-dynamics torques}. CoP$_{\mathrm{AP/ML}}$ = center-of-pressure in the anterior–posterior / medio-lateral direction.