

KINESIS: Motion Imitation for Human Musculoskeletal Locomotion

Merkourios Simos¹, Alberto Chiappa¹, Alexander Mathis¹

Abstract—How do humans move? Advances in reinforcement learning (RL) have produced impressive results in capturing human motion using physics-based humanoid control. However, torque-controlled humanoids fail to model key aspects of human motor control such as biomechanical joint constraints and non-linear, overactuated musculotendon control. We present KINESIS, a model-free motion imitation framework that tackles these challenges. KINESIS is trained on 1.8 hours of locomotion data and achieves strong motion imitation performance on unseen trajectories. Through a negative mining approach, KINESIS learns robust locomotion priors that we leverage to deploy the policy on several downstream tasks such as text-to-control, target point reaching, and football penalty kicks. Importantly, KINESIS learns to generate muscle activity patterns that correlate well with human EMG activity. We show that these results scale seamlessly across biomechanical model complexity, demonstrating control of up to 290 muscles. Overall, the physiological plausibility makes KINESIS a promising model for tackling challenging problems in human motor control. Code, videos, and benchmarks are available at <https://github.com/amathislab/Kinesis>.

I. INTRODUCTION

A hallmark of embodied intelligence is the ability to efficiently navigate the world. In humans and other animals, this capacity to *locomote* translates to a diverse skillset including walking, turning, and running, movements that are driven by hundreds of musculotendon organs, constrained by the intricate joint morphology of the skeletal system [1]. How biological organisms master the complexity of the musculoskeletal system is a crucial question for neuroscience, and it could hold the key to unlocking robust and efficient high-dimensional motor control in robots.

Deep reinforcement learning (DRL) has shown success in the field of physics-based motor control, enabling the control of humanoids that run, jump, and navigate complex terrain [2], [3]. However, due to different scientific goals, these approaches use unrealistic models of the human body, which consist of a simplified kinematic tree and are notably controlled by torque controllers instead of muscles. Realistic musculoskeletal models, which are necessary to capture the true complexity of human locomotion, are receiving increasing attention across scientific domains [4]–[8]. Yet, despite the significant progress of the last years [6], [8], current works on muscle-driven locomotion are still limited compared to their torque-controlled counterparts, primarily

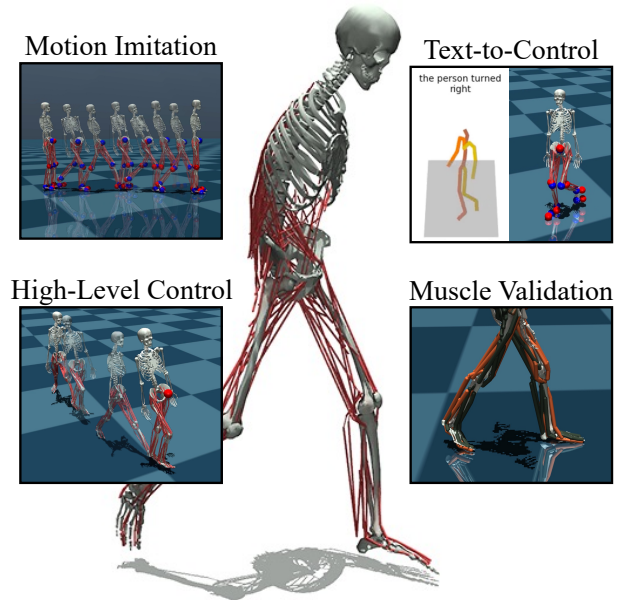


Fig. 1. KINESIS is a model-free RL policy for the musculoskeletal control of locomotion. **Top left:** Our policy is trained on a curated set of MoCap data focusing on locomotion, and can successfully imitate unseen motion clips of the same locomotion types. **Top right:** KINESIS can be deployed zero-shot on synthetic text-conditioned motion sequences. **Bottom left:** KINESIS is fine-tuned to reach goal locations, and produces human-like motion without reference motion guidance. **Bottom right:** Muscle activity patterns produced during motion imitation (active muscles shown in orange, inactive muscles shown in black) correlate well with human electrophysiology data (EMG).

due to their narrow focus on forward gait [9]–[11] or lack of quantitative comparisons to human data [12]–[14].

We propose **KINESIS**, a model-free motion imitation framework that facilitates the development of effective and scalable muscle-based control policies of locomotion (Fig. 1). We trained KINESIS on 1.8 hours of motion capture data that we curated for whole-body locomotion. We implemented this framework on three biomechanically validated musculoskeletal models of increasing complexity, and demonstrate that the learned policies achieve quantifiably high motion tracking performance and high-level controllability through natural language and target-reaching commands. Furthermore, we demonstrate that KINESIS serves as a strong foundation for downstream tasks by fine-tuning the policy to perform football penalty kicks. Finally, we find that the learned policies produce physiologically plausible muscle activity patterns, which we validate by comparisons with human electromyography (EMG) recordings across multiple subjects.

¹EPFL

Contact: alexander.mathis@epfl.ch

We thank the reviewers and members of the Mathis Group for helpful feedback. In particular, we thank Bianca Ziliotto, Chengkun Li, Alessandro Marin Vargas, and Johann Brea for feedback on earlier versions. This project was funded by Swiss SNF grant (310030 212516).

II. RELATED WORK

a) Musculoskeletal Models of the Human Body: Efforts to model the human musculoskeletal system trace back to the Hill-type model, which provided a mathematical formulation of the dynamical properties of muscles [15], and much later of musculotendon units [16]. Since then, many works have incorporated these models into numerical simulations of muscle-based systems [8]. In particular, the OpenSim platform provided an open-source tool for the design of accurate musculoskeletal models of the human body [4], [17]. Unfortunately, it is computationally costly compared to physics engines that are used in robotics, machine learning and RL [18]. To this end, Lee et al. [12] implemented a full musculoskeletal model in DART [12], [19], and incorporated it into a motor skill task. Similarly, Caggiano et al. [20] developed MyoS uite, a suite of physiologically accurate biomechanical models implemented in Mujoco by porting models from OpenSim. Contemporary RL work uses alternative musculoskeletal models with more muscles [21]; however, those are not open-source. In this work, we use three variants of the MyoLeg model from MyoS uite, with the most complex variant featuring 290 muscles, more than any other open-source model.

b) Control of Musculoskeletal Agents: For high-dimensional musculoskeletal control, some works have utilized reference motion data [12]–[14], [22], whereas others focused on task-oriented learning frameworks [10], [11], [23]–[31]. MuscleVAE combined motion imitation of various locomotion skills with a muscle fatigue model using a model-based policy [13]. More recently, MAGNET introduced a three-stage learning approach for imitation learning of a large MoCap dataset and developed a distilled model for real-time muscle activity generation [14]. While the current lack of available training code for either of these works hinders direct comparison, KINESIS presents several comparative advantages. First, KINESIS is based on a single-stage, model-free algorithm, demonstrating seamless application to three musculoskeletal models with virtually no hyperparameter tuning. Second, we do not rely on model distillation for real-time muscle activity generation; our full model can run at 3.8-times the real speed on the CPU of a commercial laptop. Third, we present comprehensive quantitative metrics for our imitation and EMG comparison results as a baseline for future work.

c) Comparison with human muscle activity: Electromyography (EMG) signals are a common way to validate the physiological plausibility of biomechanical models fit to motion capture data [8], [32]–[34]. In particular, the simulated muscle activity patterns are verified against physiological sequences measured by EMG. This matching is non-trivial, due to the non-linear, high dimensional, and over-actuated nature of biomechanical systems [4], [5]. Some studies have investigated the correspondence between RL-based muscle control and human muscle activity, but for control policies generating a single motion skill at a time [6], [22]. MAGNET focused on real-time EMG prediction, but

the quantitative correspondence of the generated activity patterns to multi-subject ground truth data remains unclear [14]. Here, we propose a validation benchmark based on EMG data from human subjects during locomotion [35], [36] that can be applied to any muscle-based control policy modeling leg muscles, and we show that KINESIS significantly outperforms state-of-the-art control policies that use the MyoLeg model [9], [10].

III. METHODS

In this section, we describe the three musculoskeletal models from the MyoS uite library [20], the curated KIT-Lo comotion reference dataset [37], and the motion imitation learning framework.

A. Musculoskeletal Model (MyoLeg)

MyoS uite [20] is a library built on top of the Mujoco physics simulator [18], featuring several biologically-accurate models of different human body parts. Among them, MyoLeg is a locomotion-oriented musculoskeletal simulation of the human body that is biologically validated against an OpenSim [4] model of the full human body by [17]. In this work, we apply the same training method to three variants of the MyoLeg with increasing complexity:

- *Legs:* The simplest version is actuated at the legs by 80 musculotendon units, while the upper body is modeled as a rigid volume with mass.
- *Legs+Abs:* This model includes six additional abdominal muscles (86 in total) that enable basic postural control of the upper body.
- *Legs+Back:* The most advanced version is a combination of the basic model of the legs with a comprehensive model of the back [38], enabling full lumbar control. It consists of 290 musculotendon actuators (Fig. 1).

In all three models, the body dimensions closely match the mean human body shape of the SMPL model [39], enabling the seamless use of MoCap datasets.

B. Human Motion Dataset

a) KIT-Lo comotion: We utilize human motion clips from the KIT Motion Dataset [37]. KIT is a collection of MoCap datasets under a unified representation framework, containing a variety of full-body motor skills. Because of our focus on the motion of the lower body, we extracted a locomotion-specific subset of 1054 reference motion sequences (KIT-Lo comotion), representing 1.8 hours of MoCap data. We selected motion sequences based on their motion description, using key terms such as 'walk', 'turn', or 'run'. We filtered this selection further by evaluating motion categories on feasibility, diversity, and length. In the curated form, *KIT-Lo comotion* consists of five broad locomotion skills: (1) *walk*, in which the subject walks forward at varying speeds, (2) *gradual turn*, in which the subject turns to the left or right while walking, (3) *turn in place*, where the subject walks in a straight line, then performs a U-turn and resumes walking, (4) *walk backwards*, where the subject takes backward steps, and (5) *run*, where the subject breaks into a short running stride. We extracted all motions from

the AMASS dataset [40], which use the SMPL body format [39].

b) Data pre-processing: We downsampled the motions to 30 frames per second before aligning them with the reference frame used in Mujoco. We computed joint rotations in quaternion space, which, along with the global translation of the pelvis, fully describe the motion. For motion imitation, we considered the reference body set $\{head, pelvis, knees, ankles, toes\}$, removing head tracking for the basic model since it lacks upper postural control.

c) Inverse kinematics: Although motion imitation is performed in Cartesian space, we still need to define the initial position of the model in joint angles, and the joint rotation space of MyoLeg is different compared to the skeletal model that describes motions in the SMPL kinematic structure. Therefore, we performed inverse kinematics to align the model with the reference motion frame. We repeated this process for multiple frames per reference trajectory, thus enabling the initialization at random starting points during the motion.

C. Motion Imitation

a) Preliminaries: The motion imitation problem can be cast as a Partially-Observable Markov Decision Process (POMDP) $\mathcal{M} = \langle \mathcal{S}, \mathcal{O}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma \rangle$, characterized by the state space \mathcal{S} , the observation function \mathcal{O} , the action space \mathcal{A} , the transition function \mathcal{T} , the reward function \mathcal{R} and the discount factor γ . At each time step, an agent interacts with the simulated environment in Mujoco to control the MyoLeg model and track a reference motion. Through the observation function $\mathcal{O} : \mathcal{S} \rightarrow \mathbb{R}^N$, the environment provides an observation vector $\mathbf{o}_t \in \mathbb{R}^N$, describing the proprioceptive state of the body. Specifically, the observation vector includes: the height, the tilt and the velocity of the pelvis, the joint kinematics (position, velocity, angular position and angular velocity for each joint), and the feet contact forces at time t , as well as the target reference pose (absolute and relative) at time $t + 1$. The agent controls the body by specifying a control signal for each of the M muscles, represented by a vector $\mathbf{a}_t \in \mathcal{A} \subset \mathbb{R}^M$. The next state $\mathbf{s}_{t+1} \in \mathcal{S}$ is determined by the transition function $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$, defined by the physics simulator and the muscle dynamics. N is equal to 309, 363, and 453 for the basic, medium, and advanced models respectively.

b) Reward structure: The objective of the agent is to maximize the discounted cumulative reward $R = \sum_{t=0}^T \gamma^t r_t$, where T denotes the termination point of the POMDP and γ denotes the discount factor. The reward $r_t \in \mathbb{R}$ is the output of the function $\mathcal{R} : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$, which maps the current state \mathbf{s}_t and action \mathbf{a}_t into a reward signal. In a motion imitation task, the goal is to reproduce the target movement as faithfully as possible. As in previous motion imitation work [3], [41], [42], we defined the imitation component of the reward function as the sum of the negative Euclidean distance between the body's joint position vector \mathbf{p}_t^i and the position vector $\hat{\mathbf{p}}_{t+1}^i$ of the target pose:

$$r_t^{pos} = \exp \left(-k_{pos} \sum_i \|\mathbf{p}_t^i - \hat{\mathbf{p}}_{t+1}^i\|^2 \right). \quad (1)$$

The joints included in the sum are the pelvis, knees, ankles, toes, as well as the head for the Legs+Abs and L+Back models. We apply an exponential kernel to make the rewards positive and avoid encouraging early termination. The scaling factor k_{pos} controls the rate at which the reward decreases with higher distance from the target motion frame.

Similarly, we defined a velocity reward:

$$r_t^{vel} = \exp \left(-k_{vel} \sum_i \|\mathbf{v}_t^i - \hat{\mathbf{v}}_{t+1}^i\|^2 \right), \quad (2)$$

where \mathbf{v}^i and $\hat{\mathbf{v}}^i$ denote the joint velocities of the same keypoints for the skeleton and target pose.

Since the musculoskeletal model is overactuated, with antagonist muscles controlling the same joints, there are infinite sequences of muscle activation that produce the same motion. To distinguish between these solutions and discourage simultaneous co-activation of antagonist muscles, we add L1 and L2 regularization in terms of the expended energy at each time step through the reward component:

$$r_t^e = -\|\mathbf{m}_t\| - \|\mathbf{m}_t\|_2, \quad (3)$$

where \mathbf{m}_t is the vector containing the muscle activation signals for time step t . Finally, we optionally include an upright reward component:

$$r_t^{up} = \exp \left(-k_{up} (\theta_{f,t}^2 + \theta_{s,t}^2) \right), \quad (4)$$

where $\theta_{f,t}$ and $\theta_{s,t}$ denote the forward and sideways tilt of the pelvis in radians, at time step t .

The total reward at each time step is a weighted sum of the reward components, with the individual weights chosen as hyperparameters.

c) Muscle actuation: We tested two widely-used motor control schemes to map the policy output into the activation signal. First, we followed an indirect control scheme, in which the policy output $\mathbf{a}_t \in \mathbb{R}^M$ represents the desired next-step muscle lengths [12], [13], [42]. A proportional-derivative (PD) controller is responsible for finding the force which a muscle should apply for the joint to reach the desired position, according to the equation $F_i = F_i^{\text{peak}} (k_p (a_i - l_i) - k_d v_i)$, where i denotes the muscle index, F_i^{peak} denotes the maximum force that the muscle can exert, l denotes the current muscle length, v denotes current muscle velocity, while k_p and k_d are scaling factors that are empirically determined and remain fixed.

To convert muscle force into muscle activity $\mathbf{m}_t \in \mathbb{R}^M$, we use the force-length-velocity relationship:

$$m_i = \frac{F_i - F_P(l_i)}{F_L(l_i) F_V(v_i)}, \quad (5)$$

where F_P denotes the force exerted by the muscle's passive element as a function of its length, F_L denotes the

TABLE I
QUANTITATIVE MOTION IMITATION RESULTS. MEAN AND SEM FOR N=3 EVALUATION RUNS.

Model	KIT-Loocomotion Train (946 motions)			KIT-Loocomotion Test (108 motions)		
	Frame Coverage \uparrow	Success \uparrow	$E_{MP,JPE} \downarrow$	Frame Coverage \uparrow	Success \uparrow	$E_{MP,JPE} \downarrow$
Legs-PD	96.98 \pm 0.06 %	90.63 \pm 0.21 %	48.8 \pm 0.16	96.38 \pm 0.79 %	91.36 \pm 1.76 %	48.76 \pm 0.08
Legs-Direct	98.10 \pm 0.05 %	93.27 \pm 0.26 %	42.72 \pm 0.22	97.99 \pm 0.35 %	94.14 \pm 0.91 %	43.04 \pm 0.42
Legs+Abs	98.99 \pm 0.06 %	96.48 \pm 0.12 %	44.01 \pm 0.22	98.63 \pm 0.41 %	95.68 \pm 1.10 %	44.17 \pm 0.61
Legs+Back	99.22 \pm 0.06 %	97.04 \pm 0.38 %	42.26 \pm 0.04	99.43 \pm 0.03 %	96.91 \pm 0.50 %	42.84 \pm 0.29

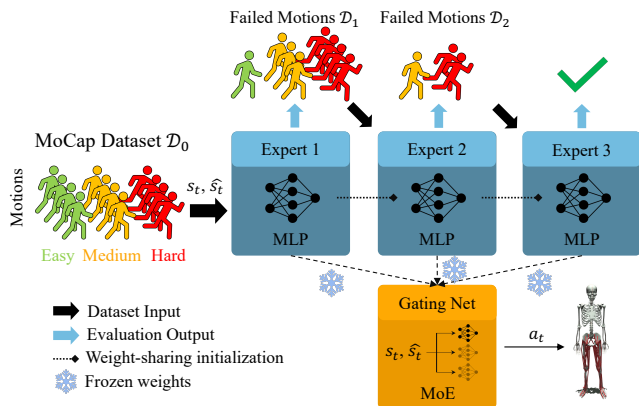


Fig. 2. An illustration of the hard negative mining strategy. The first policy network (expert) is trained on the entire dataset, containing five locomotion skills of varying difficulty (here denoted as easy, medium, and hard). After a given number of training epochs, the motions that the first expert successfully imitates are removed from the dataset, and the reduced dataset is used to train a new copy of the expert. The process repeats until the dataset is empty. Finally, a MoE gating network is trained to select the appropriate expert for a specific time step, given the current proprioceptive state and target pose.

active isometric force exerted by the muscle as a function of its length, and F_V denotes the active force exerted by a maximally activated muscle at its resting length, as a function of its velocity. These parameters are provided in MyoLeg.

Alternatively, we implemented a direct mapping $\mathbf{m}_t = (\mathbf{a}_t + 1)/2$, which simply rescales the policy output into the muscle activation range. We compared these two methods in the Legs model, and we found that the direct control scheme reduced training and simulation time significantly, while achieving better results. Therefore, we used the direct control scheme for the Legs+Abs and Legs+Back models, as well as for the downstream tasks.

D. Reinforcement Learning

In RL, the agent interacting with the environment is defined by the policy $\pi(\mathbf{a}_t|\mathbf{o}_t)$, which specifies the conditional distribution over actions $a_t \in \mathbb{R}^M$ given an observation $\mathbf{o}_t \in \mathbb{R}^N$. We built on the motion imitation approach by Luo et al. [42], using a training set \mathcal{D} with reference motion sequences, and following a Mixture-of-Experts (MoE) approach.

Training episodes begin by sampling a random reference motion from \mathcal{D} and a random starting frame, setting the musculoskeletal model to the reference state, and initiating environment dynamics. The distance between the reference frame and the model is computed as the episode progresses. If any tracked keypoint deviates more than 0.15 meters from

its reference, the episode ends early; otherwise, it concludes at the end of the sequence without further reward, and a new motion is sampled.

We modeled each expert policy with a 6-hidden layer [2048, 1536, 1024, 1024, 512, 512] multilayer perceptron (MLP) [42]. The control policy $\pi_\theta(\mathbf{a}_t|\mathbf{o}_t) = \mathcal{N}(\mu(\mathbf{o}_t, \theta), \Sigma)$ is a multivariate Gaussian distribution with a covariance matrix Σ and mean $\mu(\mathbf{o}_t, \theta)$, where θ represents the learned policy network parameters. The policy is trained using proximal policy optimization (PPO) and Lattice, an exploration method that is adapted for high-dimensional control [28], [43]. The same MLP architecture models the value function.

a) *Hard negative mining*: The agent must generate a diverse array of control plans to successfully imitate the reference motions. However, the policy required for one motion sequence may be incompatible with policies learned for other motions. To address this, we followed [42] and we employed an ensemble policy based on negative hard mining (Fig. 2). In more detail, the training phase began by sampling motions from the entire training dataset. When the imitation success rate reached a plateau, we identified and isolated the motion sequences that the policy failed on. We then initialized a copy of the policy with the currently-learned weights, and we initiated a new round of training using only the isolated motion sequences. We continued this process until there were no failed motion sequences left.

b) *Mixture of experts*: Ultimately, we obtained a set of different experts that covered the entire training dataset. To combine these policy networks into a single, universal policy, we froze their parameters and trained an MoE model on the entire training set [44]. The gating network is a three-layer [1024, 512, 256] MLP that takes as input the current observable state and outputs a vector $\hat{\mathbf{a}}_t \in \mathbb{R}^n$ that determines which expert's output will be used to actuate the model. We used a softmax transformation so that \hat{a}_t^i denotes the probability of using the i 'th expert's output as the selected action. Essentially, the MoE network acts as a gating mechanism that stochastically selects an expert based on the current observations (Fig. 2). We found that three experts were sufficient to achieve high imitation performance for all three musculoskeletal models.

E. High-level Control Tasks

Beyond motion imitation, a truly intelligent agent should be able to apply its learned skills to novel, high-level control problems. We selected three downstream tasks—text-to-control, target goal reaching, and football penalty kick—that progressively test the policy's ability to generalize its learned

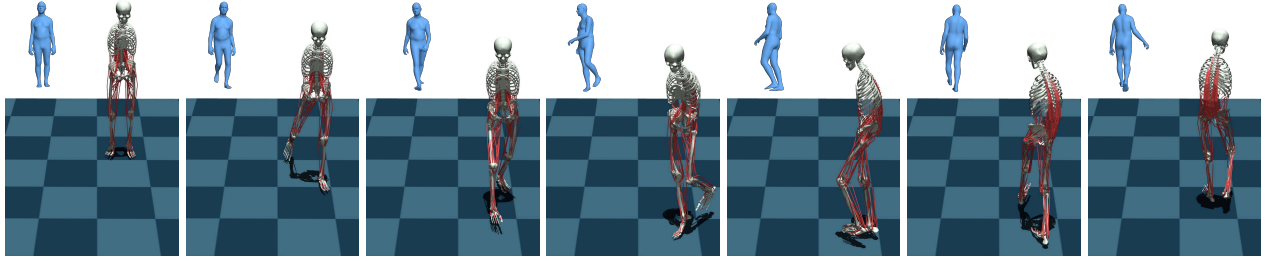


Fig. 3. The Legs+Back musculoskeletal model is imitating a “Turn in Place” motion; a rendering of the reference motion is shown in blue.

TABLE II

KINESIS HYPERPARAMETERS. DIFFERENT VALUES FOR INDIVIDUAL MODELS ARE DENOTED AS LEGS/LEGS+ABS/LEGS+BACK

Hyperparameter	Value	Hyperparameter	Value
Sim. timestep	150^{-1}	Reward w_{vel}	0.2
Control interval	5	Reward w_e	0.1/0.2/0.2
PD controller k_p	1	Reward w_{up}	0.1/0.0/0.0
PD controller k_d	1	Reward $w_{tracking}$	0.8
Reward k_{pos}	200	Reward $w_{success}$	20
Reward k_{vel}	5	PPO γ	0.99
Reward k_{up}	3	PPO ϵ	0.2
Reward w_{pos}	0.6	PPO τ	0.95
PPO learning rate	5×10^{-5}	PPO optimizer	Adam
Activation function	SiLU		

locomotion priors to increasingly abstract and interactive challenges.

a) Text-to-control: Natural language is an ideal medium to convey high-level behavioral goals. To this end, we leveraged Human Motion Diffusion Model (MDM) to transform the motion imitation task into a text-conditioned motor control task. MDM generates motion sequences as keypoints that follow the SMPL model [45], which is perfectly aligned with the dimensions of our model’s skeleton. We extracted the kinematics from the generated motions and initiated the motion imitation task without additional training, evaluating the musculoskeletal model in a zero-shot manner.

b) Target goal reaching: Although the abstraction of natural language is powerful, sometimes a more concrete goal formulation is favorable; for example, when we need to specify an exact location to locomote to. To test whether the locomotion skills acquired by KINESIS during motion imitation training can be transferred to other tasks without relying on reference data, we designed a target goal-reaching task. The objective was for the model to reach a designated position within a 2×2 m range from a given initial pose and facing direction. We provided the policy with proprioceptive input and target keypoints, but we set all “targets” to the current position of the respective body part, apart from the target root position, which was set to the target goal, following [42]. We updated the reward function to better match the high-level task by replacing the position and tracking velocity rewards with two new components, $r_t^{track} = \delta_{t-1} - \delta_t$, and $r_t^{success} = \mathbf{1}_{\{\delta_t < 0.1\}}$, where δ_t denotes the distance between the root and the target position at time t [46]. We set the duration of each episode to 5 seconds, and terminated the episode early if the height of the root dropped below 0.7 m. To leverage the locomotion priors of KINESIS, we used the second expert policy as the starting point for fine-tuning.

c) Penalty kick: Locomotion is often combined with other goal-directed movements that involve interactions with objects, such as a football. To test the capacity of KINESIS in this challenging context, we deployed the Legs+Back goal-reaching policy on a penalty-kick task, where the model must step to the ball and kick it towards the net while beating the goalkeeper. Similar to the goal reaching task, the observation space remains the same, and the position of the ball is given to the policy as the target root position. To incentivize the policy to score goals, we simply provide the velocity of the ball towards the net as large positive reward. The environment was adapted from the 2025 MyoChallenge [47].

IV. EXPERIMENTS

A. Implementation Details

a) Training and inference: All models were trained on a single NVIDIA A100 GPU and 128 CPU threads. We used 128 parallel environment instances. Each training step consisted of experience accumulation for 51 200 steps, gathered in total from all environments, followed by 10 rounds of policy optimization using PPO [43]. All hyperparameters are shown in Table II.

Training was performed at a simulation frequency of 150 Hz and a control frequency of 30 Hz. At those settings, inference on the CPU of a modern laptop is 3.8 times faster than real-time. The simulation and control frequencies can be increased significantly, without loss of performance.

b) Datasets: We trained KINESIS on the training split of the curated KIT-LoComotion dataset, consisting of 946 motions and five locomotion skills (walking, walking while turning, turning around in place, walking backwards, running). We evaluated the learned policy on the test split, containing 108 unseen motions from the same skill categories. For text-to-motion validation, we used MDM to generate synthetic motion trajectories for each skill category. Text prompts were chosen to ensure that generated motions matched the learned skills in style and velocity. The datasets were the same across all three musculoskeletal models.

c) Evaluation Metrics: We report the motion imitation frame coverage as the percentage of the dataset (in frames) that the model successfully imitates. We also report the trial success rate, considering a trial successful when the tracked body markers of the model remain within 0.5 meters of the reference motion on average, throughout the trial trajectory. Furthermore, we report the global mean per-joint position error (MPJPE) in millimeters, averaged across all frames. All

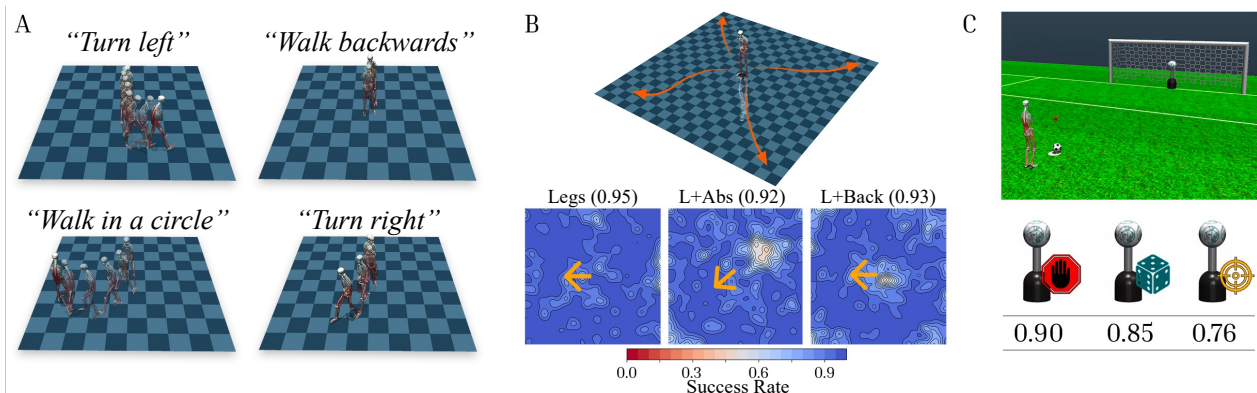


Fig. 4. Examples of musculoskeletal motor control during high-level tasks. **A:** *Text to motion*. Four examples of natural language prompts that KINESIS executes zero-shot. **B:** *Target goal reaching*. The agent is tasked with reaching random positions in \mathbb{R}^2 space starting from an initial position and orientation (orange arrows); success rate heatmaps are shown for the three musculoskeletal models (average success rate in parentheses). **C:** *Penalty kick*. The agent is tasked with beating the goalkeeper in a penalty kick, starting from a random position behind the ball. The policy consistently beats three goalkeeper strategies of increasing difficulty: a static strategy (red), a random walk strategy (blue), and an active blocking strategy (yellow); success rate after 100 kicks.

motion imitation experiments were performed three times. In the target goal-reaching task, we simulated 1000 random trials, and considered successful those trials that ended with the model within 0.1 meters of the target. In the penalty kick task, we simulated 1000 kicks with the initial position chosen randomly behind the ball, and considered successful the penalty kicks where a goal was scored against the goalkeeper while the model remained standing.

B. Results: Motion Imitation

a) Performance on KIT-locomotion Dataset.: A control policy trained using negative mining and three motor experts combined with a MoE module achieved high performance on all locomotion skills (Table I; example in Fig. 3). Imitation performance was higher with direct control vs. PD-control (Legs-PD vs. Legs-Direct); model training and inference was also faster. We thus focus on direct control from now on. The two more complex models (Legs+Abs, Legs+Back) achieved slightly higher imitation performance, highlighting the importance of the upper body for locomotion and the good scaling properties of our method.

C. Results: High-level Control Tasks

a) Text-to-Control: As a first task, we considered text-to-motion control, which we evaluated qualitatively. We prompted MDM with textual descriptions of simple locomotion (e.g., “the person walked in a circle”). KINESIS could successfully imitate the generated motions in a zero-shot manner, across all three musculoskeletal models (see example trajectories for the Legs+Back model in Fig. 4A).

b) Target Goal Reaching: The fine-tuned policy achieved strong performance on the target reaching task (Fig. 4B). KINESIS consistently navigated to the target and remained there for the duration of the episode. Notably, the trained agents utilized not only the locomotion skills acquired during the original imitation learning, but also additional skills that had not been previously encountered, such as side steps and diagonal backward motion. Only the Legs+Abs model struggled somewhat for targets right behind the initialization (success rate around 60%). Importantly, the agents retained a realistic gait after fine-tuning on this task.

c) Football penalty kick: The Legs+Back model is placed in a penalty kick task against a simulated goalkeeper that follows three adversarial strategies of increasing difficulty. More specifically, the goalkeeper can either remain static at the center of the goal, move randomly under Brownian noise, or track the lateral position of the ball with a bounded velocity. We found that adding a simple reward component for ball velocity enabled the fine-tuned KINESIS policy to consistently beat the goalkeeper (Fig. 4C).

D. Results: EMG Analysis

Imitation performance and downstream controllability are great challenges to test the capacity of models for human motion. We hypothesized that pretraining with motion imitation also leads to better alignment with muscle patterns. To assess how closely muscle-based control policies mimic biological muscle control, we developed a benchmark that measures the correlation of generated muscle activity with recorded human electromyography (EMG) measurements. For this evaluation, we adapted two datasets that include recordings from leg muscles in different subjects during various forms of locomotion [35], [36]. We compared the musculoskeletal models trained with KINESIS with two recent locomotion baselines, DynSyn and DEP-RL [9], [10], that perform forward locomotion using the Legs model.

First, we quantified the inherent variability across human subjects by averaging the individual muscle activity patterns across the periodic gait cycle [35], and calculating the pair-wise Pearson correlation (Fig. 5A). Human-to-human variability represents a strong upper bound for model performance, since a model within that range would be just as good as another human’s muscle dynamics for predicting muscle activity. Computing the correlation of generated average activity patterns across motion types and musculoskeletal models, we found that KINESIS achieves positive correlation and significantly outperforms both baselines in the forward locomotion test. KINESIS was able to capture the qualitative characteristics of the muscle activity patterns; the muscles shown in Fig. 5A are especially relevant for identifying gait disorders [48]. In addition, KINESIS matches the human

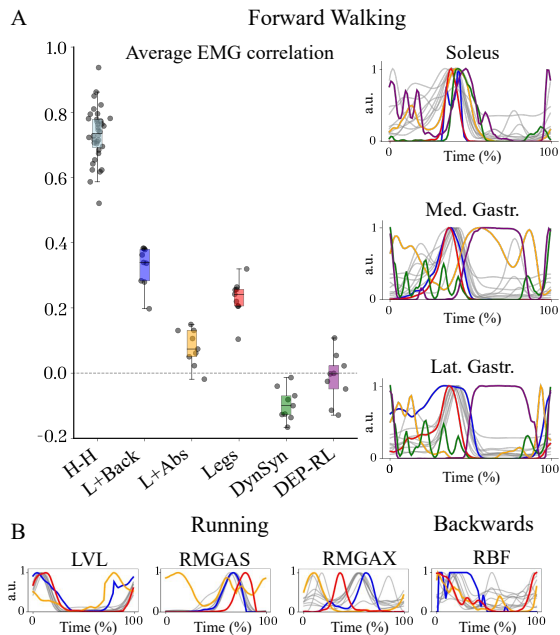


Fig. 5. Correlation of generated gait-averaged muscle activity patterns with recorded human EMG data. **Left, forward walking:** When averaging over gait cycles and muscles, the correlation of muscle activity patterns produced by KINESIS with EMG data, across embodiments, is significantly higher compared to alternative methods using the same musculoskeletal model. Between different KINESIS models, the Legs+Back model has the highest correlation. Individual points denote the average muscle activity correlation between the model and one subject ($N = 9$), and pairwise correlation between human subjects (H-H, $N = 36$). **Right:** Example gait-averaged muscle activity patterns during forward walking. **Bottom:** Example gait-averaged muscle activity patterns during running (LVL: left vastus lateralis, RMGAS: right medial gastrocnemius, RMGAX: right gluteus maximus) and backward walking (RBF: right biceps femoris).

EMG patterns for several other muscles during running and backward walking (Fig. 5B).

Interestingly, KINESIS Legs+Abs produced less human-like muscle activity compared to Legs, despite the added postural control afforded by the abdominal muscles. We later found an error in this particular embodiment that led to a significant and unrealistic increase in upper body weight. This discrepancy might explain the drop in EMG correlation and highlights the value of this analysis in evaluating both musculoskeletal models and policies.

RL from scratch often does not necessarily produce biologically plausible behavior (e.g., [11], [29]), whereas imitation learning explicitly ensures alignment with human motion. We therefore evaluated whether the advantage of KINESIS in terms of EMG correlation was strictly due to this gait similarity (Fig. 6). We calculated the mean joint angle error (MPJAE) between generated and human gaits by using the available kinematic data, considering four joints of the left leg that showed high correlation across human subjects (hip flexion, hip adduction, knee angle, ankle angle). As expected, we found that DynSyn and DEP-RL had high MPJAE. However, the MPJAE of KINESIS was also significantly higher compared to inter-subject values. Despite achieving only minor improvements in kinematic fidelity over alternative methods, KINESIS learned substantially

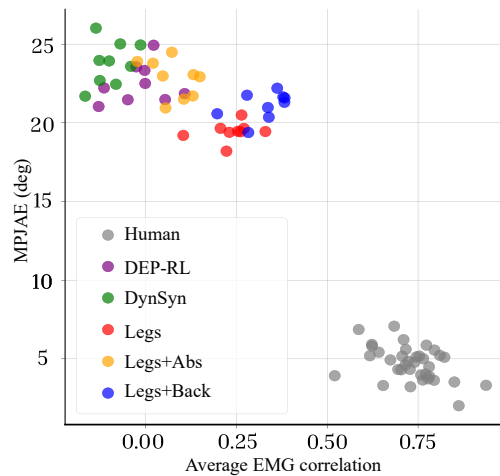


Fig. 6. Average EMG correlation between models and human data versus mean joint angle error (MPJAE) between the movement generated by the model and the human trajectory.

more accurate EMG patterns. We believe this is a promising result in support of our initial hypothesis, and suggests that motion imitation methods can close the gap to the human ceiling further as future musculoskeletal models evolve.

V. CONCLUSION

We presented KINESIS, an RL-based motion imitation framework for musculoskeletal locomotion that demonstrates high imitation performance on 1.8h of MoCap data and effective deployment across downstream tasks. To our knowledge, KINESIS is the first model-free policy combining diverse locomotion skills while generating human-like EMG patterns without being explicitly trained to do so [8]. KINESIS formed the basis of our winning solution to Myochallenge 2025 [47]; more importantly, it could serve as an *in silico* model of the human sensorimotor system to tackle fundamental questions about motor control (see Appendix in arXiv version for more details). Significant challenges remain towards a better understanding of human sensorimotor control, including more comprehensive training datasets, and refined musculoskeletal models — nevertheless, our results suggest a promising path.

REFERENCES

- [1] P. Ramdya and A. J. Ijspeert, “The neuromechanics of animal locomotion: From biology to robotics and back,” *Science Robotics*, vol. 8, no. 78, p. eadg0279, 2023.
- [2] C. Tessler, Y. Guo, O. Nabati, G. Chechik, and X. B. Peng, “Masked-mimic: Unified physics-based character control through masked motion inpainting,” *arXiv preprint arXiv:2409.14393*, 2024.
- [3] Z. Luo, J. Cao, J. Merel, A. Winkler, J. Huang, K. M. Kitani, and W. Xu, “Universal humanoid motion representations for physics-based control,” in *The Twelfth International Conference on Learning Representations*, 2024.
- [4] S. L. Delp, F. C. Anderson, A. S. Arnold, P. Loan, A. Habib, C. T. John, E. Guendelman, and D. G. Thelen, “Opensim: open-source software to create and analyze dynamic simulations of movement,” *IEEE transactions on biomedical engineering*, vol. 54, no. 11, pp. 1940–1950, 2007.
- [5] G. E. Loeb, “Learning to use muscles,” *Journal of human kinetics*, vol. 76, p. 9, 2021.
- [6] S. Song, Ł. Kidziński, X. B. Peng, C. Ong, J. Hicks, S. Levine, C. G. Atkeson, and S. L. Delp, “Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation,” *Journal of neuroengineering and rehabilitation*, vol. 18, no. 1, p. 126, 2021.

- [7] M. W. Mathis, A. P. Rotondo, E. F. Chang, A. S. Tolia, and A. Mathis, "Decoding the brain: From neural representations to mechanistic models," *Cell*, vol. 187, no. 21, pp. 5814–5832, 2024.
- [8] M. Denayer, E. Alfio, M. A. Díaz, M. Sartori, F. De Groot, K. De Pauw, and T. Verstraten, "A prisma systematic review through time on predictive musculoskeletal simulations," *Journal of Neuro-Engineering and Rehabilitation*, vol. 22, no. 1, pp. 1–24, 2025.
- [9] K. He, C. Zuo, C. Ma, and Y. Sui, "DySyn: dynamical synergistic representation for efficient learning and control in overactuated embodied systems," *arXiv preprint arXiv:2407.11472*, 2024.
- [10] P. Schumacher, D. F. Haeufle, D. Büchler, S. Schmitt, and G. Martius, "Dep-rl: Embodied exploration for reinforcement learning in overactuated and musculoskeletal systems," *International Conference on Learning Representations (ICLR)*, 2023.
- [11] P. Schumacher, T. Geijtenbeek, V. Caggiano, V. Kumar, S. Schmitt, G. Martius, and D. F. Haeufle, "Emergence of natural and robust bipedal walking by learning from biologically plausible objectives," *iScience*, 2025.
- [12] S. Lee, M. Park, K. Lee, and J. Lee, "Scalable muscle-actuated human simulation and control," *ACM Transactions On Graphics (TOG)*, vol. 38, no. 4, pp. 1–13, 2019.
- [13] Y. Feng, X. Xu, and L. Liu, "Musclevae: Model-based controllers of muscle-actuated characters," in *SIGGRAPH Asia 2023 Conference Papers*, 2023, pp. 1–11.
- [14] J. Park, E. Jung, J. Lee, and J. Won, "Magnet: Muscle activation generation networks for diverse human movement," in *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, 2025, pp. 1–11.
- [15] A. V. Hill, "The heat of shortening and the dynamic constants of muscle," *Proceedings of the Royal Society of London. Series B-Biological Sciences*, vol. 126, no. 843, pp. 136–195, 1938.
- [16] F. E. Zajac, "Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control," *Critical reviews in biomedical engineering*, vol. 17, no. 4, pp. 359–411, 1989.
- [17] A. Rajagopal, C. L. Dembia, M. S. DeMers, D. D. Delp, J. L. Hicks, and S. L. Delp, "Full-body musculoskeletal model for muscle-driven simulation of human gait," *IEEE transactions on biomedical engineering*, vol. 63, no. 10, pp. 2068–2079, 2016.
- [18] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2012, pp. 5026–5033.
- [19] J. Lee, M. X. Grey, S. Ha, T. Kunz, S. Jain, Y. Ye, S. S. Srinivasa, M. Stilman, and C. Karen Liu, "Dart: Dynamic animation and robotics toolkit," *The Journal of Open Source Software*, vol. 3, no. 22, p. 500, 2018.
- [20] V. Caggiano, H. Wang, G. Durandau, M. Sartori, and V. Kumar, "MyoSuite – A contact-rich simulation suite for musculoskeletal motor control," *arXiv preprint arXiv:2205.13600*, 2022.
- [21] C. Zuo, K. He, J. Shao, and Y. Sui, "Self model for embodied intelligence: Modeling full-body human musculoskeletal system and locomotion control with hierarchical low-dimensional representation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 13 062–13 069.
- [22] W. Qin, R. Tao, L. Sun, and K. Dong, "Muscle-driven virtual human motion generation approach based on deep reinforcement learning," *Computer Animation and Virtual Worlds*, vol. 33, no. 3-4, p. e2092, 2022.
- [23] Ł. Kidziński, C. Ong, S. P. Mohanty, J. Hicks, S. Carroll, B. Zhou, H. Zeng, F. Wang, R. Lian, H. Tian *et al.*, "Artificial intelligence for prosthetics: Challenge solutions," in *The NeurIPS'18 Competition: From Machine Learning to Intelligent Conversations*. Springer, 2020, pp. 69–128.
- [24] V. Caggiano, H. Wang, G. Durandau, S. Song, Y. Tassa, M. Sartori, and V. Kumar, "MyoChallenge: Learning contact-rich manipulation using a musculoskeletal hand," 2022.
- [25] J. Park, S. Min, P. S. Chang, J. Lee, M. S. Park, and J. Lee, "Generative gaitnet," in *ACM SIGGRAPH 2022 Conference Proceedings*, 2022, pp. 1–9.
- [26] V. Caggiano, G. Durandau, H. Wang, A. Chiappa, A. Mathis, P. Tano, N. Patel, A. Pouget, P. Schumacher, G. Martius *et al.*, "MyoChallenge 2022: Learning contact-rich manipulation using a musculoskeletal hand," in *NeurIPS 2022 Competition Track*. PMLR, 2023, pp. 233–250.
- [27] C. Berg, V. Caggiano, and V. Kumar, "SAR: generalization of physiological agility and dexterity via synergistic action representation," *Autonomous Robots*, vol. 48, no. 8, p. 28, 2024.
- [28] A. S. Chiappa, A. Marin Vargas, A. Huang, and A. Mathis, "Latent exploration for Reinforcement Learning," *Advances in Neural Information Processing Systems*, vol. 36, 2024.
- [29] A. S. Chiappa, P. Tano, N. Patel, A. Ingster, A. Pouget, and A. Mathis, "Acquiring musculoskeletal skills with curriculum-based reinforcement learning," *Neuron*, vol. 112, no. 23, pp. 3969–3983, 2024.
- [30] V. Caggiano, G. Durandau, S. Song, C. K. Tan, H. Wang, B. Hodossy, P. Schumacher, L. Gionfrida, M. Sartori, and V. Kumar, "MyoChallenge 2024: Physiological dexterity and agility in bionic humans," in *NeurIPS 2024 Competition Track*, 2024.
- [31] Y. Wei, S. Zhuang, V. Zhuang, and Y. Sui, "Motion control of high-dimensional musculoskeletal systems with hierarchical model-based planning," *arXiv preprint arXiv:2505.08238*, 2025.
- [32] J. L. Hicks, T. K. Uchida, A. Seth, A. Rajagopal, and S. L. Delp, "Is my model good enough? best practices for verification and validation of musculoskeletal models and simulations of movement," *Journal of biomechanical engineering*, vol. 137, no. 2, p. 020905, 2015.
- [33] T. DeWolf, S. Schneider, P. Soubiran, A. Roggenbach, and M. Mathis, "Neuro-musculoskeletal modeling reveals muscle-level neural dynamics of adaptive learning in sensorimotor cortex," *bioRxiv*, pp. 2024–09, 2024.
- [34] D. Schneider, S. Reiß, M. Kugler, A. Jaus, K. Peng, S. Sutschet, M. S. Sarfraz, S. Matthiesen, and R. Stiefelhagen, "Muscles in time: Learning to understand human motion in-depth by simulating muscle activations," *Advances in Neural Information Processing Systems*, vol. 37, pp. 67 251–67 281, 2024.
- [35] H. Wang, A. Basu, G. Durandau, and M. Sartori, "A wearable real-time kinetic measurement sensor setup for human locomotion," *Wearable technologies*, vol. 4, p. e11, 2023.
- [36] K. Scherpereel, D. Molinaro, O. Inan, M. Shepherd, and A. Young, "A human lower-limb biomechanics and wearable sensors dataset during cyclic and non-cyclic activities," *Scientific Data*, vol. 10, no. 1, p. 924, 2023.
- [37] C. Mandery, O. Terlemez, M. Do, N. Vahrenkamp, and T. Asfour, "The KIT whole-body human motion database," in *International Conference on Advanced Robotics (ICAR)*, 2015, pp. 329–336.
- [38] R. Walia, M. Billot, K. Garzon-Aguirre, S. Subramanian, H. Wang, M. I. Refai, and G. Durandau, "Myoback: A musculoskeletal model of the human back with integrated exoskeleton," *bioRxiv*, pp. 2025–03, 2025.
- [39] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, *SMPL: A Skinned Multi-Person Linear Model*. Association for Computing Machinery, 2023, vol. 2, pp. 851–866.
- [40] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black, "AMASS: Archive of motion capture as surface shapes," in *International Conference on Computer Vision*, Oct. 2019, pp. 5442–5451.
- [41] X. B. Peng, P. Abbeel, S. Levine, and M. Van de Panne, "Deepmimic: Example-guided deep reinforcement learning of physics-based character skills," *ACM Transactions On Graphics (TOG)*, vol. 37, no. 4, pp. 1–14, 2018.
- [42] Z. Luo, J. Cao, K. Kitani, W. Xu *et al.*, "Perpetual humanoid control for real-time simulated avatars," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 10 895–10 904.
- [43] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [44] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixtures of local experts," *Neural computation*, vol. 3, no. 1, pp. 79–87, 1991.
- [45] G. Tevet, S. Raab, B. Gordon, Y. Shafir, D. Cohen-or, and A. H. Bermano, "Human motion diffusion model," in *The Eleventh International Conference on Learning Representations*, 2023.
- [46] J. Won, D. Gopinath, and J. Hodgins, "Physics-based character controllers using conditional vaes," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 4, pp. 1–12, 2022.
- [47] "Myochallenge 2025." [Online]. Available: sites.google.com/view/myosuite/myochallenge/myochallenge-2025
- [48] R. L. Lenhart, C. A. Francis, A. L. Lenz, and D. G. Thelen, "Empirical evaluation of gastrocnemius and soleus function during walking," *Journal of biomechanics*, vol. 47, no. 12, pp. 2969–2974, 2014.