Bipedal Balance Control with Whole-body Musculoskeletal Standing and Falling Simulations

Chengtian Ma, Yunyue Wei, Chenhui Zuo, Chen Zhang, Yanan Sui

Tsinghua University

Abstract:
Balance control is important for human and bipedal robotic systems. While dynamic balance during locomotion has received considerable attention, quantitative understanding of static balance and falling remains limited. This work presents a hierarchical control pipeline for simulating human balance via a comprehensive whole-body musculoskeletal system. We identified spatiotemporal dynamics of balancing during stable standing, revealed the impact of muscle injury on balancing behavior, and generated fall contact patterns that aligned with clinical data. Furthermore, our simulated hip exoskeleton assistance demonstrated improvement in balance maintenance and reduced muscle effort under perturbation. This work offers unique muscle-level insights into human balance dynamics that are challenging to capture experimentally. It could provide a foundation for developing targeted interventions for individuals with balance impairments and support the advancement of humanoid robotic systems. Project page: `https://lnsgroup.cc/research/bipedal_balance`.

Keywords: balance control, bipedal standing and falling, musculoskeletal system

1 Introduction

Bipedal locomotion and balance have been extensively studied in robotics and control, with research demonstrating the inherent challenges of stabilizing underactuated, high-degree-of-freedom systems [1, 2]. Robotic bipedal stability during locomotion has been analyzed through criteria like limit cycles and gait periodicity [3, 4]. While bipedal standing seems naturally achievable, stable standing remains a challenging problem, particularly when accounting for double-support phases and area contact [5]. However, the foundational behavior enabling bipedalism, human static standing, remains understudied due to the lack of models capable of capturing its musculoskeletal complexity. Humans rely on an obligate bipedal stance, requiring precise coordination between neuromuscular control and biomechanical dynamics. Unlike occasional bipeds seen in other animals, human standing is characterized by a vertical spine, the absence of auxiliary balancing structures (e.g., prehensile tails), and an upper body for multitasking while maintaining balance [6]. This skill is acquired over years of development, relying on advanced neural integration and reflexes [7]. Mastering static balance could lag behind walking in children [8], highlighting its difficulty. The process integrates multisensory inputs and fine-grained muscle coordination [9], posing a control challenge distinct from robot locomotion.

Humans achieve remarkable balance adaptability, but the principles underlying this robustness remain poorly understood. The human musculoskeletal system, with its high degrees of freedom and nonlinearity, requires more sophisticated control mechanisms to perform the dynamical control

9th Conference on Robot Learning (CoRL 2025), Seoul, Korea.
A short version won the Best Paper Award at the How Do Robots Care Workshop, ICRA 2025. Correspondence to Yanan Sui: ysui@tsinghua.edu.cn

of stable standing. A critical gap persists: *no prior work has systematically investigated bipedal standing control using a full-body musculoskeletal model*, which is essential to decode fine-grained, muscle-level postural dynamics [10].

In contrast to maintaining balance, falling is a leading cause of injury among aging populations [11], resulting in both psychological trauma and physical harm [12]. Despite its significance, falling dynamics remain poorly understood due to data scarcity and ethical constraints in real-world experiments [13].

In this work, we introduce a hierarchical control and analysis method for full-body bipedal balance control. Our framework, deployed on a full-body human musculoskeletal model with 700 muscle-tendon units [14], enables biomechanically plausible simulations of standing balance and falls that would be limited in real experiments. Through extensive simulations, we identify specific dynamical balance behavior during stable standing, reveal the adaptive balancing strategies under injury conditions, and demonstrate how hip exoskeleton assistance can enhance balance while reducing muscular effort. We provide muscle-level insights during balance that remain inaccessible through conventional experimental methods. To the best of our knowledge, our approach is the first to leverage a high-dimensional, whole-body musculoskeletal model to systematically study human balance dynamics at muscle-level resolution, providing insights for both biomechanical understanding of bipedal balance and assistive technology development.

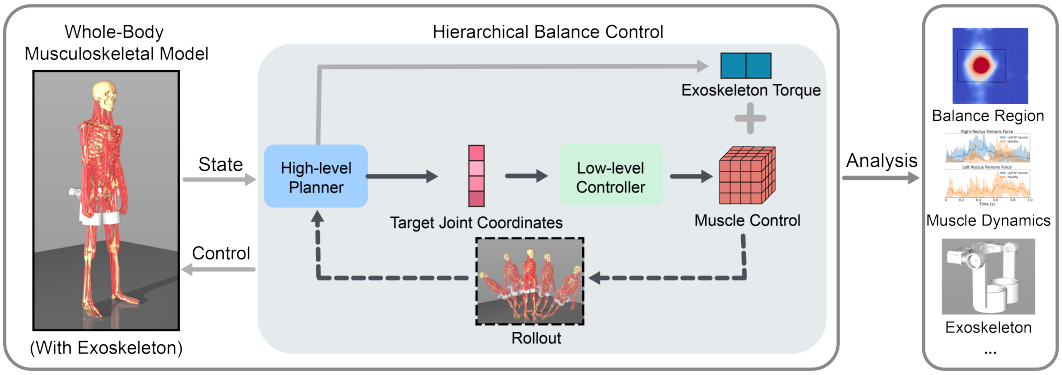


Figure 1: Control and analysis of human musculoskeletal standing and falling. HBC enables training-free balance control and efficient collection of dynamical behavior for balance analysis. Our method supports concurrent control planning for exoskeleton-assisted scenarios, facilitating integrated evaluation and optimization of human-exoskeleton interaction.

2 Related Work

2.1 Bipedal Balance from Robotics to Human Musculoskeletal Models

The balance control of bipedal robots has been extensively studied, with numerous approaches demonstrating the inherent challenges of stabilizing underactuated systems. Traditional research has employed simplified models such as single inverted pendulum [15] and double inverted pendulum models [16, 17] to represent bipedal balance dynamics by focusing on the Center of Mass (CoM) motion and the Zero Moment Point (ZMP).

However, human balance control is far more complex than robot models can capture, involving intricate sensorimotor integration executed through high-dimensional muscular systems [18] rather than simplified bipedal standing frameworks like ZMP strategies [19, 20]. A gap between theoretical models and biological reality remains, necessitating the use of musculoskeletal models [21, 22]. Recent work introduced a full-body human musculoskeletal system that simulates whole-body dynamics [14]. Control methods using deep reinforcement learning (DRL) have been explored for these high-dimensional systems [23, 24, 25, 26, 27]. While progress has been made in improving locomotion capabilities, the control of full-body standing balance remains unexplored.

2.2 Investigation of Human Balance and Fall

While robotic fall-prediction systems have advanced through multi-sensor fusion [28] and machine learning [29], human fall dynamics lack physiologically accurate understanding. Existing datasets [30, 13] primarily consist of voluntary falls, which differ substantially from real-world falls in terms of impact dynamics and protective responses [31].

Exoskeletons have been used to assist balance for various populations [32, 33], but experimental studies still have significant limitations. Most exoskeleton research has focused on locomotion tasks [34, 35], with limited investigation into balance scenarios. Studies that target at assisting balance with exoskeletons typically involve small participant numbers and controlled lab conditions [33, 36]. These limitations in experimental scope and participant diversity hinder a comprehensive understanding of how exoskeletons might benefit the balance control. Musculoskeletal simulations have aided exoskeleton design [37] by estimating joint torques [38] or serving as control testbeds [39]. To the best of our knowledge, there are no existing works that use musculoskeletal simulation to test and validate the performance of balance-targeted exoskeletons.

3 Problem Setting

3.1 Musculoskeletal Model Dynamics

We use the full-body MS-Human-700 model [14], which comprises of 90 rigid body segments, 206 joints and 700 muscle-tendon units. It is implemented in the MuJoCo physics engine [40]. The actuators of the model are 700 Hill-type [41] muscles. The force generated by each actuator, and the temporal relation between muscle activation act and the input control signal u can be described by the following equations:

$$f_m(act) = f_{max} \cdot [F_l(l_m) \cdot F_v(v_m) \cdot act + F_p(l_m)], \quad \frac{\partial act}{\partial t} = \frac{u - act}{\tau(u, act)}, \tag{1}$$

where F_l and F_v are active force-length and force-velocity functions, F_p is the passive force-length function, and l_m, v_m are normalized muscle length and velocity. f_{max} is the maximum isometric muscle force. Muscle activation act is calculated with control u in Eq. (19) [42].

3.2 Standing Task Design and Balance-Related Scenarios

We treat the human standing control problem as a finite horizon Markov decision process with state $s \in \mathcal{S}$, control $u \in \mathcal{U}$, dynamics $s_{t+1} = f(s_t, u_t)$, and time step t. For a given initial state s_0 and a desired standing horizon T, we aim to find a control sequence $U_T^\star = (u_0, ..., u_{T-1})$ that successfully maintains the standing posture by minimizing a pre-defined cost function C:

$$U_T^\star = \text{argmin}_{U_T} \sum_{t=0}^{T-1} C(s_t, u_t) \tag{2}$$

The control goal for standing is: (1) Near-zero CoM horizontal velocity and (2) CoM within the support region, ensuring no flipping tendency. The model may step for balance, but its initial position remains constant across experiments.

4 The Control Algorithm and Analysis for Static Balancing

In this section, we present our control and analysis method for whole-body static balance in humans. We begin by introducing a hierarchical control framework that enables training-free control of high-dimensional balance dynamics. We then detail our analysis for evaluating balance performance under varying physiological conditions and with assistance.

Algorithm 1: Hierarchical Balance Control (HBC)

Input: Model dynamics f, total time steps T, execution length t_e, rollout horizon h, particle number n, MPPI iteration number r, initial distribution parameter μ, σ

Output: Action sequence U^*

```
s_0 ~ S_0, U* ← ∅
for t = 0, ··· , T − 1 do
    if t mod t_e = 0 then
        for i = 1, ··· , r do
            z_1, ..., z_N ~ N(μ, σ)
            c_1, ··· , c_n ← Rollout(z_1, ··· , z_n)
            Update μ, σ using e.q. (4)
        end
        z* ~ N(μ, σ)
    end
    u_t ← π(s_t, z*)
    U* = U* ∪ u_t
    s_{t+1} ← f(s_t, u_t)
end
```

4.1 Hierarchical Balance Control Algorithm

We propose Hierarchical Balance Control (HBC), a hierarchical algorithm for the balance control of human body. As shown in Figure 1, a high-level planner first proposes a set of major joint coordinates $z^* \in \mathcal{Z}$ as the target of the low-level controller, $\pi(u|s, z)$, which coordinates muscle controls to achieve the target joint coordinates. HBC uses model predictive control (MPC) to optimize the target joint coordinates according to the cumulative cost function over a short horizon:

$$z^\star = \text{argmin}_z \sum_{t=t_p}^{t_p+H-1} C(s_t, u_t), u_t = \pi(s_t, z), \tag{3}$$

where $t_p = \{t \mid t \mod t_e = 0\}$ is the planning timestep and $H \ll T$ is the planning horizon. Compared to vanilla MPC methods, HBC reduces the parameter space from the scale of muscle number to joint number, which is a feasible dimension for effective planning and control.

For the high-level controller, we integrate Model Predictive Path Integral (MPPI, [43]) as a sampling-based planner. At the beginning of each rollout process, N sets of target joint angles $z_1, z_2, ..., z_N$ are sampled from the current target distribution $\mathcal{N}(\mu, \sigma)$ to generate N trajectories. The cumulative cost $c_n = \sum_{t=t_p}^{t_p+H-1} C(s_t, \pi(s_t, z_n))$ is calculated from the parallel rollout trajectory of target z_n. The new mean and covariance are calculated based on the weighted average of top k targets $(z_1^*, z_2^*, ..., z_k^*)$ whose corresponding trajectories return the minimal costs. The update is implemented as follows:

$$\mu = \frac{\sum_{j=1}^{k} w_j \cdot z_j^*}{\sum_{j=1}^{k} w_j}, \sigma = \sqrt{\frac{\sum_{j=1}^{k} w_j \cdot (z_j^* - \mu)^2}{\sum_{j=1}^{k} w_j}}, \tag{4}$$

where $w_j = e^{-\frac{1}{\lambda}} c_j$ is the aggregate weight for z_j, and λ is the softmax temperature. The full method of getting control sequences is shown in Algorithm 1.

For the low-level controller, we use a normalized PD control-like formulation over muscle lengths to derive desired muscle forces [25]:

$$f_m = min(0, k_p \cdot (l_m^* - l_m)/l_{range} + k_d \cdot (0 - \dot{l}_m)), \tag{5}$$

where f_m stands for muscle forces, k_p and k_d are PD control gains, l_m stands for actual muscle lengths, $\dot{l}_m$ stands for muscle velocities and l_m^* is target muscle lengths. The difference between the maximum muscle length and the minimum muscle length of each muscle, l_{range}, is used to normalize muscle lengths and stabilize control effect. Given desired muscle forces, the desired

muscle control can be derived according to the muscle actuator dynamics in e.q. (19) to facilitate full-body control.

Compared to DRL-based methods, HBC is a training-free method which generates effective controls in minutes. This advantage enables the collection of large dynamical dataset during static balance control under different conditions for statistical analysis.

4.2 Standing and Fall under Different Physiological Conditions

We apply HBC over the MS-Human-700 model to collect dynamical balance behavior during stand control. The model was initialized in a natural standing posture and then stood for 5 seconds. We consider the following criteria for balance and fall monitoring:

Balance behavior monitoring. During each simulation, the model is considered to have achieved stable standing if two conditions were met: (1) No body parts other than the feet contacted the ground. (2) The CoM didn't leave the real-time support polygon when 5s is reached. We collected CoM positions, muscle force of the model and body posture data at a 500 Hz frequency for analysis.

Fall detection and recording. When we apply HBC from an up-straight posture for 5 seconds, the human model has small possibility to fall from standing. We track the CoM position of the MS-Human-700 model and the support polygon formed by foot contact points (heels and toes). We define fall duration as the duration between two events: The initialization event and the contact event. The initialization event occurs when the CoM crosses from inside to outside the support polygon, marking the start of falling. The contact event is defined as the moment when the body-ground contact force reaches its peak, indicating impact. During the 5-second trial, these events determine the model's balance status: the initialization event signals the start of a fall, while the contact event marks ground collision. Beyond healthy conditions, we model impaired balance resulting from muscle injury by simulating reduced muscular function. Specifically, we restrict the force-generating capacity of the left rectus femoris (RF) muscle to replicate the effects of injury, as reported in [44].

4.3 Exoskeleton-Assisted Standing

As illustrated in Figure 1, the high-level planner in HBC concurrently optimizes control strategies for balance-assistive devices such as a hip exoskeleton, enabling efficient simulation and design of exoskeletal systems. Assisted balance is simulated by applying torque actuation at hip joints—a widely accepted approach for modeling exoskeleton effects via externally applied forces [45]. To enhance assisted balance performance, we optimize the exoskeleton design parameters p (e.g., control gains). Formulated as a black-box optimization problem, we apply Bayesian optimization to improve the exoskeleton's posture correction policy.

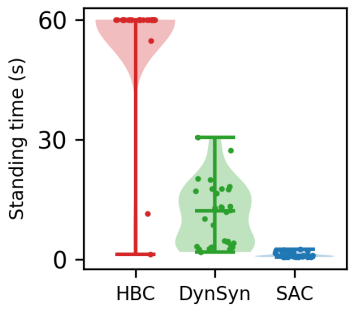


Figure 2: Standing durations achieved by three algorithms. The evaluations for balancing were terminated at 60 seconds.

5 Experiments

In the experiments, we first evaluate the control performance of the HBC algorithm for a high-dimensional whole-body standing task, with biomechanical fidelity check of the simulation results. We then generate balance and fall behaviors and conduct comprehensive analysis on healthy, injured, and exoskeleton-assisted conditions.

5.1 Performance Evaluation of the HBC Algorithm

We compare HBC against competitive DRL-based control baselines on the full-body balance control task. Our selected baselines include Dynamic Synergy Representation (DynSyn), the leading DRL-

based control method for the control musculoskeletal systems [27], and Soft Actor Critic (SAC) [46]. As shown in Figure 2, our approach successfully achieves static balance control on the full-body musculoskeletal system, significantly outperforming both baseline methods and attaining the maximum standing duration in most trials. We carried out a fidelity check on our simulated control results by comparing the muscle activation calculated by HBC with the human experimental results reported in [47, 48]. Figure 3 shows that the simulation results demonstrate consistency with real-world experiments.

5.2 Study of Balance Dynamics

In this setion, we study the dynamical behavior during balance control over a healthy model. A total of 2,800 falling trajectories were collected, where similar data collection in real world will be very difficult to achieve. For all plots on the X-Y coordinate plane, the initial support polygon is demonstrated by connecting the starting positions of heels and toes.

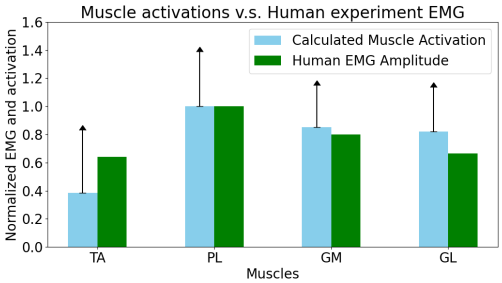
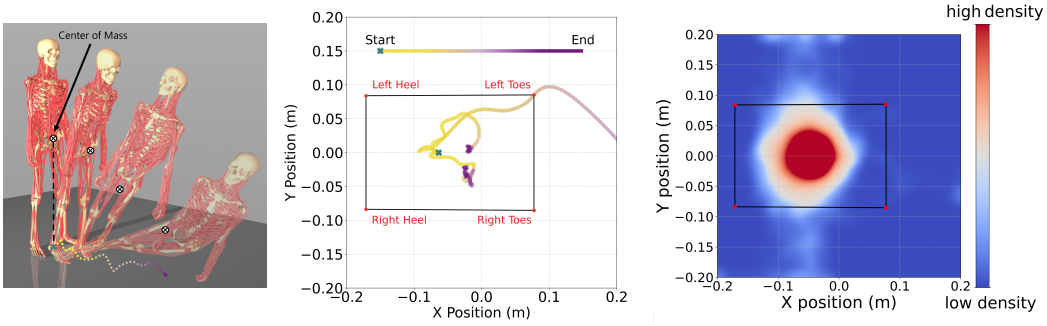


Figure 3: Comparisons between calculated muscle activations and real human EMG. Simulated muscles activity (blue) of the standing task in normalized root mean square values. Experimental electromyography (EMG) signals of a human subject (green). Four experimentally measured muscles: TA, tibialis anterior; PL, peroneus longus; GM, gastrocnemius medialis; GL, gastrocnemius lateralis.

CoM trajectories during balancing. As discussed in Section 2, CoM position and movement are important metrics to investigate bipedal balance. We observe horizontal CoM position to track the subtle, dynamic movements of human body during balance as demonstrated in Figure 4a. Representational CoM trajectories are visualized in Figure 4b, where two of them are from successful balancing trajectories with the end converging near the starting point, and one from a falling trajectory. The CoM trajectories of balancing are usually irregular at the beginning, and converge to a smaller area, showing that balancing of a high-dimensional, muscle -controlled human body is dynamical and influenced by instant control and responses.

Balance region achieved during stable standing. Although CoM series data are very noisy and diverse due to the dynamic nature of balance, it is possible to draw statistical insights from the



(a) (b) (c)

Figure 4: Human static balancing behavior dynamics. (a) Visualization of a human fall with the center of CoM projection trajectory. (b) CoM trajectories during balance and falling, with each trajectory color-coded over time from yellow to purple. (c) Density plot of CoM coordinates during successful balance trials. The red area highlights the balance region, representing the highest density of CoM positions associated with stable balancing dynamics.

overall distribution of CoM positions arrived during balance. We visualized the spatial density of CoM coordinates in 640,000 frames of successful trials during standing in Figure 4c. Similar data collection in real world will be very difficult to achieve.

The distribution of collision positions following falls. We recorded a total of 2,895 falls during the experiments and present statistical analyses based on the resulting falling behaviors. We present statistical results from the collected falling behavior. Post-fall collision locations are closely associated with injury risk and offer insights that are difficult to obtain in real-world experimentation. As shown in Figure 5, the most frequent collision regions during falls include the upper extremities (wrist and forearm), lower leg, and pelvis. This findings align with frequently injured sites in clinical data [49, 50]. Notably, 37% of fall-related injuries involve broken or fractured bones [49], highlighting the biomechanical vulnerability of these impact sites. Our simulation framework facilitates detailed analysis of such injury-prone collision patterns, providing a biomechanically accurate testbed for the development of fall prevention and mitigation strategies.

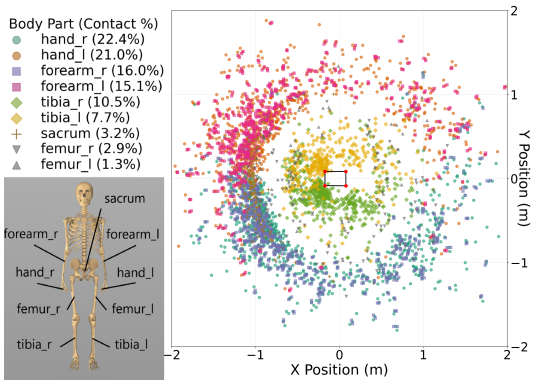


Figure 5: Distribution of 2,895 collision positions in falling. We plot the body parts and their corresponding collision positions when the contact forces reach maximum value.

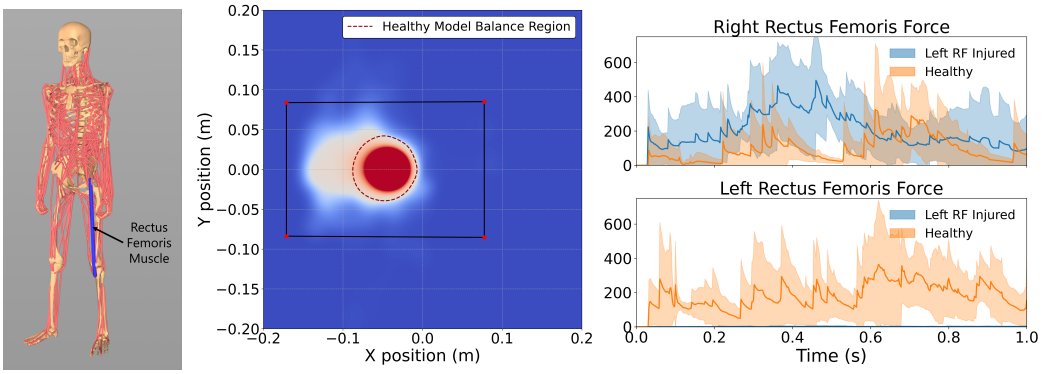
5.3 Study of Injury Effect on Balance

With the injured model as introduced in Section 4.2, a total of 1280 falling trajectories were collected. We investigated the CoM behavior and muscle force responses under the injury model. Figure 6b shows that the balance region shrinks with unilateral RF injury. The comparative analysis of the density distributions reveals two notable phenomena. First, the balance region exhibits a more concentrated pattern in the injured model, suggesting enhanced predictability and more conservative balancing strategy under injury. Second, there's an asymmetric expansion of the low-density region (white area) in the posterior direction, indicating an increased tendency of forward-leaning postural adjustments. Figure 6c shows a significant increase in right RF muscle force during balance maintenance, suggesting a weight shift to the right leg for stability.

5.4 Simulation of Assisted Balance

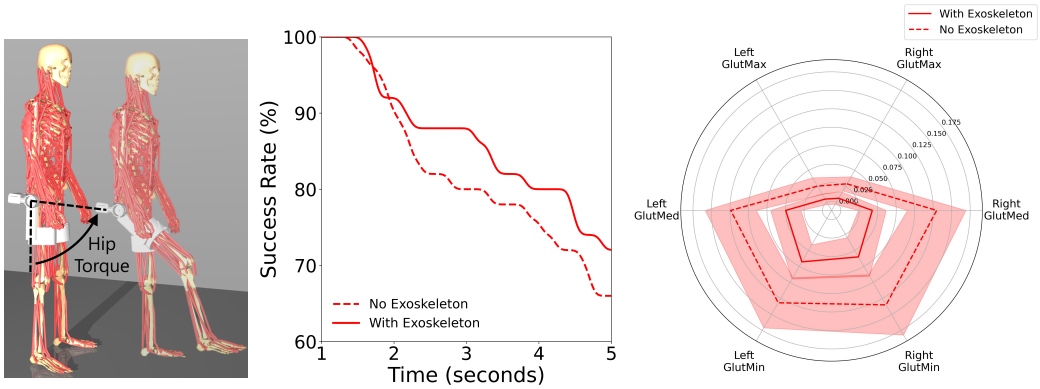
We simulated a assisted balance scenario with a hip exoskeleton as shown in Figure 7a. We employed optimized exoskeleton parameters and validated the effectiveness of this exoskeleton control policy in a perturbation test: Models with and without exoskeleton assistance are pushed in random directions for 3 times with intervals of 1 second. As in Figure 7b, exoskeleton-assisted balance achieved a higher success rate in maintaining balance under perturbation in the 5-second simulation.

In Figure 7c, we carried out a ablation study over the assisted balance by recording the muscle activations of the gluteus maximus, gluteus medius, and gluteus minimus, which play critical roles in lower-limb movement and postural stability. We observe that muscle activation levels reduce with exoskeleton assistance, showing assistive devices' potential to save muscular effort and reduce metabolic cost. Such muscle-level data are difficult to obtain in traditional experiments due to the limitations of surface EMG and the inaccessibility of deep muscles. Our simulation pipeline enables validation of exoskeleton effects on balance maintenance. These findings suggest the utility of musculoskeletal simulations in evaluating and optimizing assistive device performance prior to costly physical prototyping and human subject testing.



(a) (b) (c)

Figure 6: Balancing behavior under muscle injury conditions. (a) Illustration of the injured muscle (left rectus femoris). (b) Density plot of CoM coordinates under muscle injury. The red dashed contour represents the balance region of the healthy model, highlighting the reduction in balance region due to impaired muscle function. (c) Muscle forces during standing; increased force in the right rectus femoris indicates compensatory activation in response to left side injury.



(a) (b) (c)

Figure 7: Balancing behavior under exoskeleton assistance. (a) Visualization of the hip exoskeleton device, using joint torques as actuators. (b) Success rate of maintaining standing posture, where the exoskeleton assistance helps better static balance. (c) Activation levels of gluteal muscle activations (GlutMax: gluteus maximus, GlutMed: gluteus medius, GlutMin: gluteus minimus) during the standing simulation, where muscle activation levels are reduced with exoskeleton assistance.

6 Conclusion

We present a training-free hierarchical control method for studying full-body human bipedal balance using musculoskeletal simulation. Our Hierarchical Balance Control (HBC) approach enables effective planning and control by reducing the parameter space from muscle to joint level, and outperforms leading DRL-based methods.

This work contributes to the understanding of human balance control by providing access to muscle-level insights that are difficult to obtain experimentally. Future work could explore more complex scenarios and develop targeted interventions for populations with balance limitations. While simulation has inherent limitations compared to real-world studies, our approach offers a complementary tool for investigating questions in human movement science that face practical experimental constraints.

7 Limitation

This work demonstrates the challenge of maintaining human-like balance on a high-dimensional whole-body musculoskeletal model and proposes a hierarchical control method to achieve stable balance. While we achieves training-free balance control carried out comprehensive analysis under various conditions, several limitations remain.

First, although HBC outperforms DRL baselines, we used an intuitive cost function design without extensive tuning. In contrast to DRL methods that often rely on carefully engineered reward functions, our approach might benefit from more sophisticated cost function designs to further improve performance and biological fidelity. Second, our simulations primarily focus on a standard healthy adult model. We did not extensively simulate vulnerable populations such as elderly individuals or those with specific neurological conditions that affect balance. These populations exhibit distinct biomechanical and neuromuscular characteristics that may lead to different balance control strategies and falling patterns, limiting the generalizability of our conclusions to these important user groups. Finally, our exoskeleton simulation uses simplified torque actuation rather than modeling the complete mechanical interface between the device and human tissue. This approximation, while computationally efficient, may not fully capture the complex effects that occur in real human-exoskeleton systems.

8 Acknowledgment

This work is funded by STI2030-Major Projects-2022ZD0209400, NSFC-62206152, and Tsinghua Dushi Fund.

References

[1] Y. Fujimoto, S. Obata, and A. Kawamura. Robust biped walking with active interaction control between foot and ground. In *Proceedings. 1998 IEEE international conference on robotics and automation (cat. no. 98ch36146)*, volume 3, pages 2030–2035. IEEE, 1998.

[2] T. Sugihara. Standing stabilizability and stepping maneuver in planar bipedalism based on the best com-zmp regulator. In *2009 IEEE International Conference on Robotics and Automation*, pages 1966–1971. IEEE, 2009.

[3] F. Gubina, H. Hemami, and R. B. McGhee. On the dynamic stability of biped locomotion. *IEEE Transactions on Biomedical Engineering*, (2):102–108, 1974.

[4] J. Pratt and G. Pratt. Intuitive control of a planar bipedal walking robot. In *Proceedings. 1998 IEEE international conference on robotics and automation (Cat. No. 98CH36146)*, volume 3, pages 2014–2021. IEEE, 1998.

[5] C. Mummolo, W. Z. Peng, C. Gonzalez, and J. H. Kim. Contact-dependent balance stability of biped robots. *Journal of Mechanisms and Robotics*, 10(2):021009, 02 2018. doi:10.1115/1.4038978. URL `https://doi.org/10.1115/1.4038978`.

[6] J. R. Skoyles. Human balance, the evolution of bipedalism and dysequilibrium syndrome. *Medical Hypotheses*, 66(6):1060–1068, 2006. ISSN 0306-9877. doi:https://doi.org/10.1016/j.mehy.2006.01.042.

[7] H. Diener, F. Bootz, J. Dichgans, and W. Bruzek. Variability of postural "reflexes" in humans. *Experimental brain research*, 52(3):423–428, 1983.

[8] Y. Breniére and B. Bril. Development of postural control of gravity forces in children during the first 5 years of walking. *Experimental brain research*, 121:255–262, 1998.

[9] J. Massion. Postural control systems in developmental perspective. *Neuroscience & Biobehavioral Reviews*, 22(4):465–472, 1998.

[10] F. J. Valero-Cuevas, H. Hoffmann, M. U. Kurse, J. J. Kutch, and E. A. Theodorou. Computational models for neuromuscular function. *IEEE reviews in biomedical engineering*, 2:110–135, 2009.

[11] R. S. Kakara, R. Lee, and E. N. Eckstrom. Cause-specific mortality among adults aged 65 years in the united states, 1999 through 2020. *Public Health Reports*, 139(1):54–58, 2024.

[12] R. Luque, E. Casilari, M.-J. Morón, and G. Redondo. Comparison and characterization of android-based fall detection systems. *Sensors*, 14(10):18543–18574, 2014.

[13] W. et al. Kinematic analysis of video-captured falls experienced by older adults in long-term care. *Journal of Biomechanics*, 48(6), 2015.

[14] 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)*, pages 13062–13069, 2024. doi:10.1109/ICRA57147.2024.10610081.

[15] P. Morasso, A. Cherif, and J. Zenzeri. Quiet standing: The single inverted pendulum model is not so bad after all. *PloS one*, 14(3):e0213870, 2019.

[16] Y. Suzuki, T. Nomura, M. Casadio, and P. Morasso. Intermittent control with ankle, hip, and mixed strategies during quiet standing: A theoretical proposal based on a double inverted pendulum model. *Journal of Theoretical Biology*, 310:55–79, 2012. ISSN 0022-5193. doi:https://doi.org/10.1016/j.jtbi.2012.06.019.

[17] M. Günther, O. Müller, and R. Blickhan. Watching quiet human stance to shake off its straitjacket. *Archive of Applied Mechanics*, 81:283–302, 2011.

[18] F. B. Horak. Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls? *Age and ageing*, 35(suppl_2):ii7–ii11, 2006.

[19] J. H. Kim, J. Lee, and Y. Oh. Stability regions for standing balance of biped humanoid robots. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4735–4740, 2017. doi:10.1109/ICRA.2017.7989549.

[20] M. Vukobratović and B. Borovac. Zero-moment point—thirty five years of its life. *International journal of humanoid robotics*, 1(01):157–173, 2004.

[21] 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*, 54(11):1940–1950, 2007.

[22] C. Vittorio, W. Huawei, D. Guillaume, S. Massimo, and K. Vikash. Myosuite – a contact-rich simulation suite for musculoskeletal motor control. `https://github.com/myohub/myosuite`, 2022. URL `https://arxiv.org/abs/2205.13600`.

[23] S. Lee, M. Park, K. Lee, and J. Lee. Scalable muscle-actuated human simulation and control. 38(4), 2019. ISSN 0730-0301. doi:10.1145/3306346.3322972. URL `https://doi.org/10.1145/3306346.3322972`.

[24] P. Schumacher, D. Häufle, D. Büchler, S. Schmitt, and G. Martius. Dep-rl: Embodied exploration for reinforcement learning in overactuated and musculoskeletal systems, 2022. URL `https://arxiv.org/abs/2206.00484`.

[25] Y. Feng, X. Xu, and L. Liu. Musclevae: Model-based controllers of muscle-actuated characters, 2023. URL `https://arxiv.org/abs/2312.07340`.

[26] C. Berg, V. Caggiano, and V. Kumar. Sar: Generalization of physiological agility and dexterity via synergistic action representation. *Autonomous Robots*, 48(8):28, 2024.

[27] K. He, C. Zuo, C. Ma, and Y. Sui. Dynsyn: Dynamical synergistic representation for efficient learning and control in overactuated embodied systems, 2024. URL `https://arxiv.org/abs/2407.11472`.

[28] R. Subburaman, D. Kanoulas, L. Muratore, N. G. Tsagarakis, and J. Lee. Human inspired fall prediction method for humanoid robots. *Robotics and Autonomous Systems*, 121:103257, 2019.

[29] M. E. Mungai, G. Prabhakaran, and J. W. Grizzle. Fall prediction for bipedal robots: The standing phase. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13135–13141. IEEE, 2024.

[30] E. Hsiao and S. Robinovitch. Common protective movements govern unexpected falls from standing height. *Journal of Biomechanics*, 31(1), 1997.

[31] e. a. J. Klenk. Comparison of acceleration signals of simulated and real-world backward falls. *Medical Engineering Physics*, 33(3), 2011.

[32] I. Farkhatdinov, J. Ebert, G. Van Oort, M. Vlutters, E. Van Asseldonk, and E. Burdet. Assisting human balance in standing with a robotic exoskeleton. *IEEE Robotics and automation letters*, 4(2):414–421, 2019.

[33] A. Emmens, E. van Asseldonk, M. Masciullo, M. Arquilla, I. Pisotta, N. L. Tagliamonte, F. Tamburella, M. Molinari, and H. van der Kooij. Improving the standing balance of paraplegics through the use of a wearable exoskeleton. In *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, pages 707–712, 2018. doi: 10.1109/BIOROB.2018.8488066.

[34] S. Luo, M. Jiang, S. Zhang, J. Zhu, S. Yu, I. Dominguez Silva, T. Wang, E. Rouse, B. Zhou, H. Yuk, et al. Experiment-free exoskeleton assistance via learning in simulation. *Nature*, 630 (8016):353–359, 2024.

[35] Y. Lu, Y. Huang, R. Yang, Y. Wang, Y. Ikegami, Y. Nakamura, and Q. Wang. A human-prosthesis coupled musculoskeletal model for transtibial amputees. *IEEE Transactions on Biomedical Engineering*, 2025.

[36] V. Monaco, P. Tropea, F. Aprigliano, D. Martelli, A. Parri, M. Cortese, R. Molino-Lova, N. Vitiello, and S. Micera. An ecologically-controlled exoskeleton can improve balance recovery after slippage. *Scientific reports*, 7(1):46721, 2017.

[37] D. D. Molinaro, K. L. Scherpereel, E. B. Schonhaut, G. Evangelopoulos, M. K. Shepherd, and A. J. Young. Task-agnostic exoskeleton control via biological joint moment estimation. *Nature*, 635(8038):337–344, 2024.

[38] S. K. Lam and I. Vujaklija. Joint torque prediction via hybrid neuromusculoskeletal modelling during gait using statistical ground reaction estimates: an exploratory study. *Sensors*, 21(19): 6597, 2021.

[39] F. Mo, Q. Zhang, H. Zhang, J. Long, Y. Wang, G. Chen, and J. Ye. A simulation-based framework with a proprioceptive musculoskeletal model for evaluating the rehabilitation exoskeleton system. *Computer Methods and Programs in Biomedicine*, 208:106270, 2021.

[40] 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*, pages 5026–5033, 2012. doi:10.1109/IROS.2012.6386109.

[41] F. E. Zajac. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Critical reviews in biomedical engineering*, 17(4):359–411, 1989.

[42] M. Millard, T. Uchida, A. Seth, and S. L. Delp. Flexing computational muscle: modeling and simulation of musculotendon dynamics. *Journal of biomechanical engineering*, 135(2): 021005, 2013.

[43] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou. Aggressive driving with model predictive path integral control. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1433–1440, 2016. doi:10.1109/ICRA.2016.7487277.

[44] B. S. Unuvar, M. S. Torlak, H. Gercek, O. Tufekci, K. Erdagi, and B. Işik. Comparison and relationship of quadriceps femoris angle, muscle strength, and balance in athletes and non-athletes. *Indian Journal of Orthopaedics*, 57(8):1243–1250, 2023.

[45] C. L. Dembia, A. Silder, T. K. Uchida, J. L. Hicks, and S. L. Delp. Simulating ideal assistive devices to reduce the metabolic cost of walking with heavy loads. *PloS one*, 12(7):e0180320, 2017.

[46] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor, 2018. URL `https://arxiv.org/abs/1801.01290`.

[47] M. Blaszczyszyn, A. Szczkesna, and K. Piechota. semg activation of the flexor muscles in the foot during balance tasks by young and older women: A pilot study. *International journal of environmental research and public health*, 16(22):4307, 2019.

[48] E. Yamanaka, Y. Horiuchi, and I. Nojima. Emg-emg coherence during voluntary control of human standing tasks: a systematic scoping review. *Frontiers in Neuroscience*, 17:1145751, 2023.

[49] M. Do, V. Chang, N. Kuran, and W. Thompson. Fall-related injuries among canadian seniors, 2005–2013: an analysis of the canadian community health survey. *Health promotion and chronic disease prevention in Canada: research, policy and practice*, 35(7):99, 2015.

[50] K. M. DeGoede, J. A. Ashton-Miller, and A. Schultz. Fall-related upper body injuries in the older adult: a review of the biomechanical issues. *Journal of biomechanics*, 36(7):1043–1053, 2003.

[51] T. Howell, N. Gileadi, S. Tunyasuvunakool, K. Zakka, T. Erez, and Y. Tassa. Predictive sampling: Real-time behaviour synthesis with mujoco, 2022. URL `https://arxiv.org/abs/2212.00541`.

[52] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13:455–492, 1998.

[53] M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy. Botorch: A framework for efficient monte-carlo bayesian optimization. *Advances in neural information processing systems*, 33:21524–21538, 2020.

9 Supplementary material

9.1 Implementation detail

MJPC environment. We used MuJoCo MPC [51] in our experiments to deploy HBC in simulation time. With the vey high dimensionality of the MS-Human-700 model, the simulation speed was set to 10%. A 5-second experiment will cost 50 seconds, which is much shorter than the training time of reinforcement learning methods. HBC experiments used 16-core CPU + RTX 3080 GPU (600ms/32-step, 16-particle planning).

DRL experiments set up. RL training used DynSyn [27] on 224-core CPU + RTX 4070 GPU (CPU simulation was the bottleneck; adding more GPUs doesn't accelerate training). RL polices trained for >1 day still fails at stable standing, while HBC can generate 2000 trajectories within 1 day, enabling significant speedup.

9.2 Intuitive cost function design

Standing balance could be described by the dynamics of CoM and ZMP. We set the control goal of human standing simulated by a musculoskeletal model as follow: (1) the horizontal velocity of the CoM should be near 0; (2) the ZMP should fall inside the support region, represented by a posture that has no trend of flipping over. The model may take a step to maintain balance, but the initial position of the model is kept the same for all experiments.

Considering these basic mechanical requirements of balancing, and hints of a desirable standing posture, we arrange components of the cost function C as follows. The weights were tuned to $(w_H, w_R, w_{P_c}, w_{V_c}, w_I) = (300, 300, 300, 10, 1)$.

$$C = w_H \cdot C_H + w_R \cdot C_R + w_{P_c} \cdot C_{P_c} + w_{V_c} \cdot C_{V_c} + w_I \cdot C_I \tag{6}$$

(1) **Height,** C_H: This term penalizes the difference between body height and the initial height, which represents the status of the upright, not-flipping-over posture. It reaches 0 when the difference between head and feet of the musculoskeletal model is the same as its initial value. The height cost represents the status of the upright, not-flipping-over posture. H_{head} is the height of the head, and H_{feet} is the average height of left and right feet.

$$C_H = z_{head} - z_{feet} \tag{7}$$

(2) **Rotation,** C_R: This term penalizes large rotational angle of the model's upper body. $\vec{t}$ represents the unit vector pointing from the pelvis to the head, and $\vec{z}$ is the vertical unit vector.

$$C_R = 1 - \vec{t} \cdot \vec{z} \tag{8}$$

(3) **CoM position** C_{P_c}: This term penalizes the mismatch between the horizontal CoM and the center of the support region, which is the convex hull formed by the heels and toes. (x_{CoM}, Y_{CoM}) refers to the horizontal position of the center of mass, while (x_{feet}, y_{feet}) is the centroid of the polygon formed by the left and right heels and toes.

$$C_{P_c} = |(x_{CoM}, y_{CoM}) - (x_{feet}, y_{feet})| \tag{9}$$

(4) **CoM velocity,** C_{V_c}: This term penalizes large horizontal velocity of CoM.

$$C_{V_c} = |(\dot{x}_{CoM}, \dot{y}_{CoM})| \tag{10}$$

(5) **Imitation,** C_I: This term penalizes the mismatch between the joint positions and joint angles of a desirable, natural standing posture. q is the joint angles of the model, and q_{ref} is the full-body joint angles of a natural standing posture.

$$C_I = |q - q_{ref}| \tag{11}$$

9.3 Exoskeleton control policy and optimization

We implemented a weighted postural PD control over the joint torque actuation placed at the left and right hip flexion joints. The high-level planner (as discussed in Section 4.1) is adapted to plan an extra target posture indicating the overall leaning direction of the body, represented by the tilt angle of the pelvis. The control policy of the exoskeleton torque actuation is separated into two parts mixed by a weight: (1) Hip flexion joint angle PD control. (2) Postural PD control. The control is formulated as follow:

$$\tau_e^i = (1-w) \cdot (k_{p_e} \cdot (q_i^* - q_i) + k_{d_e} \cdot (0 - \dot{q}_i)) + w \cdot (k_{p_t} \cdot (q_t^* - q_t) + k_{d_t} \cdot (0 - \dot{q}_t). \tag{12}$$

$i = 1, 2$ represents the left or right side respectively. τ_e^i is the torque actuation value. k_{p_e} and k_{p_e} are the joint angle PD control constants, while k_{p_t} and k_{p_t} are the postural PD control constants over the tilt angle of the pelvis. q_i^* and q_t^* are the target values of the hip joint angles and the pelvis tilt angle. w is the weight between the two PD control policies.

We found that the assistive effect of the exoskeleton was very sensitive to the k_{p_e}, k_{p_e}, k_{p_t} and w values. Therefore, we carried out Bayesian optimization (BO) to determine a set of parameter to ensure performance across trials. We define $\boldsymbol{x} = (k_{p_e}, k_{p_e}, k_{p_t}, w)$, and formulate the parameter search as a black-box optimization problem:

$$\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}} \left[-\sum_{t=0}^{T-1} C(s_t, u_t) \right] \tag{13}$$

where the objective $f(\boldsymbol{x})$ is the negative cumulative cost function under parameter $\boldsymbol{x}$, averaged over 5 independent trials. We assume the observation noise is i.i.d Gaussian: $y = f(\boldsymbol{x}) + \eta, \eta \sim \mathcal{N}(0, \sigma^2)$. Given sampled data $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), \cdots, (\boldsymbol{x}_n, y_n)\}$ We use Gaussian process to model the objective function with posterior mean and covariance estimation under kernel function k:

$$\begin{aligned} \mu_n(\boldsymbol{x}) &= \boldsymbol{k}_n(\boldsymbol{x})^T (\boldsymbol{K}_n + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{y}_n \\ k_n(\boldsymbol{x}, \boldsymbol{x}') &= k(\boldsymbol{x}, \boldsymbol{x}') - \boldsymbol{k}_n(\boldsymbol{x})^T (\boldsymbol{K}_n + \sigma^2 \boldsymbol{I})^{-1} \mathbf{k}_n(\boldsymbol{x}') \\ \sigma_n^2(\boldsymbol{x}) &= k_n(\boldsymbol{x}, \boldsymbol{x}), \end{aligned} \tag{14}$$

where $\boldsymbol{k}_n(\boldsymbol{x}) = [k(\boldsymbol{x}_1, \boldsymbol{x}), ..., k(\boldsymbol{x}_n, \boldsymbol{x})]$ is the covariance between $\boldsymbol{x}$ and sampled points, $\boldsymbol{K}_n$ is the covariance of sampled positions: $[k(\boldsymbol{x}, \boldsymbol{x}')]_{\boldsymbol{x}, \boldsymbol{x}' \in X_n}$. Given the GP posterior, we optimize the

Algorithm 2: Exoskeletal parameter search via Bayeisan optimization

Input: Initial dataset $\mathcal{D}_0$, Gaussian process $\mathcal{M}_0$

for $n = 1, 2, \cdots$ **do**

  Update $\mathcal{M}_t$ based on $\mathcal{D}_{n-1}$

  $\boldsymbol{x}_n \leftarrow \text{argmax}_{\boldsymbol{x} \in \mathcal{X}} \text{EI}_n(\boldsymbol{x}))$

  $y_n \leftarrow f(\boldsymbol{x}_n) + \eta$

  $\mathcal{D}_n \leftarrow \mathcal{D}_{n-1} \cup (\boldsymbol{x}_n, y_n)$

end

Expected Improvement acquisition function to sample the next parameter [52]:

$$\text{EI}_n(\boldsymbol{x}) = \mathbb{E}[[f(\boldsymbol{x}) - f_n^*, 0]^+] \tag{15}$$

$$= (\mu_n(\boldsymbol{x}) - f_n^*)\Phi(\frac{\mu_n(\boldsymbol{x}) - f_n^*}{\sigma_t(\boldsymbol{x})}) + \sigma_n(\boldsymbol{x})\Phi(\frac{\mu_n(\boldsymbol{x}) - f_n^*}{\sigma_n(\boldsymbol{x})}), \tag{16}$$

where $f_n^* = \max_{i \in n} y_i$. Our sequentual optimization procedure is illustrated in Algorithm 2. We implemented the overall BO procedure based on BoTorch [53]. Figure 8 shows the optimization performance over 600 iterations. The effectiveness of this exoskeleton control policy is validated in a perturbation test: Models with and without the exoskeleton assistance are pushed in random directions for 3 times with intervals of 1 second. We observe that under the optimize parameter, the model with the assistance achieves a higher success rate in maintaining balance under perturbation through the 5-second simulation.

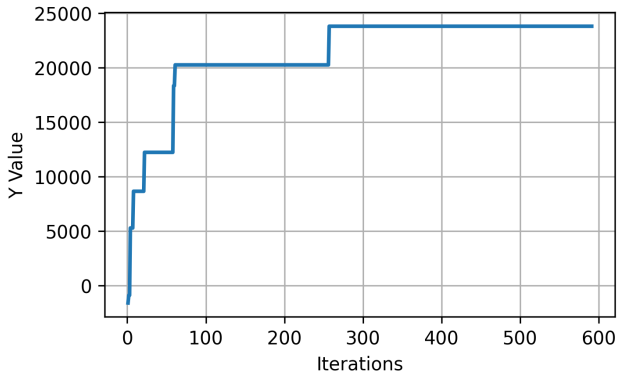


Figure 8: Exoskeleton parameter optimization results. The curve shows the best objective value during optimization.

9.4 Musculoskeletal Model Dynamics

Rigid-body dynamics

The musculoskeletal model used in this work is the MS-Human-700 model [14]. It comprises of 90 rigid body segments, 206 joints and 700 muscle-tendon units. By actuating its 700 muscle-tendon units, the model can be controlled and perform human-like tasks. The dynamics of the model can be formulated as follow:

$$M(q)\ddot{q} + c(q, \dot{q}) = J_m^T f_m(act) + J_c^T f_c + \tau_{ext}. \tag{17}$$

On the left side of the equation, q stands for generalized coordinates of joints, $M(q)$ stands for the mass distribution matrix, and $c(q, \dot{q})$ stands for Coriolis and the gravitational force. On the right side, J_m and J_c stand for Jacobian matrices that map forces to the generalized coordinates, f_c is the constraint force, $f_m(act)$ stands for actuator forces generated by muscle-tendon units determined by muscle activations (act), and τ_{ext} stands for all external torque when interacting with environments.

MS-Human-700 model is implemented in the MuJoCo physics engine [40]. The actuators of the model in this work are 700 Hill-type [41] muscles. The actuator force generated by each muscle-tendon unit, and the temporal relation between muscle activation act and the input control signal of the musculoskeletal model u can be decided by the following equations:

$$f_m(act) = f_{max} \cdot [F_l(l_m) \cdot F_v(v_m) \cdot act + F_p(l_m)]. \tag{18}$$

$$\frac{\partial act}{\partial t} = \frac{u - act}{\tau(u, act)}, \tag{19}$$

In Eq. (18), F_l and F_v represent force-length and force-velocity functions which are actuator gains, F_p is the passive force that works as actuator bias, and l_m, v_m are normalized length and normalized velocity of the muscle. f_{max} is the maximum isometric muscle force as specified in the model. In the first-order nonlinear system described by Eq. (19), muscle activation act is calculated. The time parameter τ is computed following Millard et al. [42]. τ is the time constant related to the latency in activation and deactivation.

Muscular Inverse Dynamics

To tackle the challenge of musculoskeletal model postural control posed by its very high dimensionality and non-trivial nature, we try to avoid directly manipulating the control of 700 muscles, but approach a proper target pose defined by joint angles of the model. The length of muscles is one characteristic of the muscles that can be determined by a given set of joint angles on MS-Human-700 model. A PD control-like formulation over muscle lengths can be used to derive desired muscle forces. A normalized PD control is implemented as

$$f_m = min(0, k_p \cdot (l_m^* - l_m)/l_{range} + kd \cdot (0 - \dot{l}_m)), \tag{20}$$

where f_m stands for muscle forces, k_p and k_d are PD control gains, l_m stands for actual muscle lengths, $\dot{l}_m$ stands for muscle velocities and l_m^* is target muscle lengths. The difference between the maximum muscle length and the minimum muscle length of each muscle, l_{range}, is used to normalize muscle lengths and stablize control effect.

To fully control the muscle-actuated musculoskeletal model after obtaining desired muscle forces, inverse dynamics of the muscles are applied. We assume that a target muscle activation, determined by desired muscle forces, should be approached in the next timestep. Eq. (18) and Eq. (19) can therefore be reformulated as follows:

$$act^* = \frac{f^*/f_{max} - F_p}{F_l \cdot F_m}, \tag{21}$$

$$act^* = act + ts \cdot \frac{u - act}{\tau}, \tag{22}$$

where act^* is the target muscle activation and ts is the simulation timestep. According to MyoSuite [22], the discrete time constant can be approximated, making it possible to obtain the desired input control in closed form $u = \mathcal{PD}(s, z^*)$, where z^* denotes the desired target joint angles.