MASTER THESIS

INSIGHTS ON HUMAN STANDING BALANCE BASED ON DEEP LEARNING-DRIVEN MUSCULOSKELETAL SIMULATIONS

Carlota Trigo La Blanca

FACULTY OF ENGINEERING TECHNOLOGY
DEPARTMENT OF BIOMECHANICAL ENGINEERING

EXAMINATION COMMITTEE
Dr. Huawei Wang
Dr. Guillaume Durandau
Dr. Edwin van Asseldonk
Prof. Dr. Ir. Massimo Sartori

DOCUMENT NUMBER
BE - 935

UNIVERSITY OF TWENTE.

27/06/2023

Insights on human standing balance based on deep learning driven musculoskeletal simulations

Carlota Trigo La Blanca[1]

During the first year of human life, the ability to stand and maintain balance is acquired effortlessly, yet there are still gaps in the understanding of how humans perform those tasks. This thesis aims to enhance the comprehension of human movement control by training a reinforcement learning policy that controls a musculoskeletal model of the lower limb to perform standing balance with and without perturbation tasks. A reward function was designed based on the pelvis position error, the stability margin, and the metabolic cost. The observation space was composed of joint position, velocity, pelvis position, pelvis error, and muscle activations. Training for standing balance entailed 15 million steps. For the standing balance with perturbation task, a randomized force between 1 and 50 N was applied to the pelvis for 0.1 seconds. Two separate 20-million-step training were conducted, with the perturbation applied in the anterioposterior and mediolateral directions, respectively. The results align with existing literature, although the lower limb model does not replicate the expected human response. The model withstands higher anterioposterior perturbations compared to mediolateral perturbations, using the ankle strategy for both types of perturbations. Overall, the myoLeg model is able to stand and withstand perturbations using solely information on joint position, velocity, pelvis position, error, and muscle activation information.

***Index Terms*—human balance control, deep reinforcement learning, musculoskeletal simulations**

I. Introduction

As humans, the ability to stand is essential. Standing does not only allow one to initiate other movements, such as walking or running; it also allows one to perform daily-life activities that involve using the upper body, such as reaching for an object. Understanding how standing is performed and controlled is vital for the development of new solutions for motor impairment or wearable rehabilitation robotics.

Standing balance can be defined as the strategies used to maintain an upright posture and to keep the center of mass (COM) within the base of support (BoS) [1]. This is an inherently difficult task, given the constant challenge provided by the effect of gravity and sensory noise [2].

Research has shown that the body controller receives information from active and passive contributors. The passive ones involve the intrinsic mechanical properties of muscles and tendons [3]. The active contributor is formed by sensory information, coming from the visual, vestibular, and proprioceptive systems. The visual system signals translation and rotation by means of direction-sensitive ganglion cells [4]. The somatosensory system relies on information provided by receptors in muscles, joints, and skin, which gives an estimate of joint position and movement. The main receptors providing this information are muscle spindles and Golgi tendon organs. The first are sensors for length and velocity in the muscle, while the latter encode muscle force production. Although it has been studied that they are mostly sensitive to transient inputs, studies show that their low-frequency sensitivity could encode the muscle action being used to maintain balance [5] [6]. Last, the vestibular system encodes the motion of the head using sensors present in the otoliths and semicircular canals. They detect linear and rotational motion, respectively [7] [8].

Using the information from all sensory sources, the body estimates the position of the COM. It then uses different strategies to maintain it within the BoS and thus, keep balance. There are two main strategies: the ankle and the hip strategy. The first is performed by applying a torque on the ankle joint that aims to stabilize the body. It is normally used in situations of low and slow perturbations, as it is limited by the ability of the ankle to exert a torque and keep the center of pressure (CoP) within the BoS. On the other hand, the hip strategy involves the movement of the trunk forward or backward by rotating on the hip axis. It is used in higher amplitude perturbations, and it is limited by the ability to generate horizontal forces against the ground [9].

So far, there are no noninvasive methods that allow obtaining insight on how balance is achieved in terms of internal forces and neural controls in the body. This issue, combined with the extremely limited possibilities of experimenting on humans lead to the development of modeling approaches. Musculoskeletal modeling and simulations can be used as a non-invasive approach, that allows gaining insight into the internal forces and controllers in the body [10]. A musculoskeletal model allows to represent the dynamics of the human body as if they were controlled by the neural system [11]. It uses a representation of the body as rigid elements (bones) connected by joints and actuated by tensile and contractile elements, muscles [10].

Regarding standing balance, the field has been explored both through experimental and simulation studies [12]. The optimal control simulation framework has been thoroughly explored for this task. In 1991, He et al. investigated the regulation of posture using a dynamical model of a cat's leg [13]. Later on, the work of Kuo et al. explored coordination and the selection of control strategies in humans [14] [15]. Further, Atkeson et al. demonstrated how a single optimization criterion can lead to diverse balance strategies [16]. Shen et al. used nonlinear predictive control simulations to explore the function of the arms [17] [18]. Later on, they extended their work by publishing a study in which they used deep reinforcement learning (deepRL) driven simulations to reproduce human

[1]Department of Biomechanical Engineering, University of Twente, The Netherlands.

balance using a 2D model of the human body [19].

There are different frameworks in which musculoskeletal modeling can be implemented. Physics-based engines present a very accurate description of the musculoskeletal system. However, they have limited contact-rich interaction and are also highly computationally expensive (e.g. OpenSim [20]). On the other hand, physics engines are more efficient computationally and they provide contact-rich interactions, however, their musculoskeletal modeling support is extremely limited. An example of this engine is MuJoCo [21]. MyoSuite is a novel simulation framework that merges the most effective features of two frameworks: It combines the accurate musculoskeletal representation of OpenSim with the efficiency and contact dynamics of MuJoCo [22]. To achieve a physiologically accurate representation, the models used in Myosuite are adapted from OpenSim models to MuJoco models. This adaptation is performed via the MyoSim pipeline [23], which simplifies the model but still results in fast, stable, and realistic simulations.

The computational power of MuJoCo, and thus MyoSuite, allows for the simulations to be driven by deepRL. In deepRL, neural networks are used to tackle reinforcement learning problems, providing accurate neural representations that mimic the brain [24] [25]. This approach allows to develop controllers with high-dimensional inputs and outputs, making them more similar to humans [26]. The integration of deepRL with neuromechanical simulations in MyoSuite offers the potential to achieve human-like behavior as has already been demonstrated in other studies [27] [28].

So far, the MyoSuite models have been successfully trained for the upper body, being able to perform extremely complex tasks [22]. However, the lower body model is still under development. The aim of this thesis is to train the MyoLeg model to perform standing balance with and without perturbations.

II. Methodology

The Myosuite framework allows tasks to be solved using deepRL. DeepRL is the computational application of learning from interaction, which is achieved by using neural networks (NN). In these types of problems, the model learns what to do by performing an action and obtaining a reward [29]. For this thesis, the tasks are defined to be standing balance and standing balance with perturbation.

There are four main elements in a reinforcement learning problem: a policy, a reward function, an agent, and an environment. The environment is the "world" in which the agent operates. The agent receives constant feedback on its actions through the reward function and the state of the environment. That is achieved through the policy, which maps states to actions, and the reward received the agent can decide which action to take next. [29]. Fig 1 shows a diagram of the basic functioning of deepRL problems.

During training, the optimal policy is found by using Proximal Policy Optimization (PPO) algorithm. PPO has been demonstrated to be stable and effective in high-dimensional and continuous tasks, such as the ones in hand. An Actor-Critic architecture is used, thus two NNs are trained, one for the policy (actor-network) and one for the value function (critic network). Both NN share an architecture formed by 4 layers: one input, one output, and two hidden layers with 64 neurons each. The activation function being used is a sigmoid function.

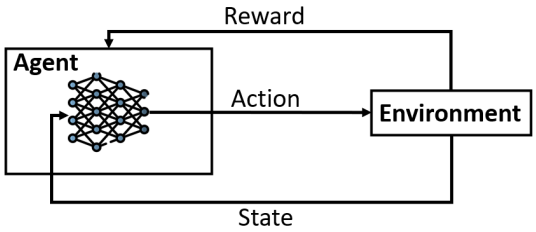

Fig. 1: Deep Reinforcement Learning Diagram. Figure adapted from [30]

The agent is the myoLeg model (see Fig. 2). This model was adapted from Rajagopal's 2015 OpenSim model [31]. The torso was removed in order to simplify the standing balance task. The model consists of 20 degrees of freedom (DoF), which are detailed in Table I. The model includes 80 actuators, which represent 60 biological muscles.

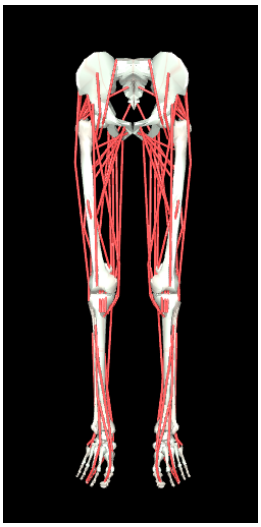

Fig. 2: MyoLeg model, without the torso.

TABLE I: Degrees of Freedom of the lower limb model

Joint	DoF	Description
Pelvis	6	Translations in 3D Rotation Tilt List
Hip	3	Flexion/Extension Adduction/Abduction Internal/External Rotation
Knee	1	Flexion/Extension
Ankle	2	Dorsiflexion/Plantarflexion Inversion/Eversion
Toes	1	Flexion/Extension

After taking an action, the agent receives information about the state of the environment. In deepRL, the environment can be fully or partially observable. In this work, the environment was kept partially observable. For the tasks performed in this work, the information that the model was receiving was: joint position, joint velocities, muscle activations, the position of the pelvis, and error of the pelvis with respect to its target position. The total size of the observation space was 155, from which 35 observations correspond to joint position, 34 to joint velocity, 80 to muscle activations, 3 to the position of the pelvis and 3 to the error of the pelvis with respect to its target position.

The aim of the agent throughout the training is to maximize a reward. In this case, the reward is based on the position error, the metabolic cost of the muscles, and the position of the center of mass with respect to the base of support. The mathematical formulation can be seen in Eqn. 1. In this equation, w_e, w_m, and w_c are the weights used for each of the reward terms. Term r_ε refers to the Frobenius norm of the pelvis position error, r_{met} is the metabolic cost of the muscles, computed as the sum of activations squared normalized by the number of muscles. Last, r_{com} is a binary term, which represents whether the center of mass is within the base of support or not.

$$R = w_e \cdot r_\varepsilon + w_m \cdot r_{met} + w_c \cdot r_{COM} \tag{1}$$

where:

$$r_\varepsilon = -1 \cdot \|X_{target} - X_{current}\|_F$$
$$r_{met} = \frac{-1}{N_{mus}} \sum a_i^2$$
$$r_{com} = \begin{cases} +1 & \text{if } COM \in BoS \\ -1 & \text{if } COM \notin BoS \end{cases}$$

The base of support was defined as the polygon formed by the x and y position of the toes and the calcaneus of both feet. The position of the center of mass was computed as shown in Eqn. 2, where X and M represent the position and mass of the bodies respectively.

$$X_{COM} = \frac{\sum_{i=0}^{N_{bodies}} X_i \cdot M_i}{\sum_{i=0}^{N_{bodies}} M_i} \tag{2}$$

The horizon represents the extent to which the agent considers the reward of its actions. For this task, it was set at 100 steps (1 step is 0.01 seconds) and the reward weights w_e, w_c, and w_m were set to 1. The initial position was prescribed for all joints and the target position was a range of values for the pelvis. Both the initial and the target position were kept constant during training and evaluation. With this configuration, a 15M step training was run to obtain the standing balance results.

For the standing balance with perturbation task, the perturbation is a force applied at the pelvis. The force was modeled to have a random magnitude between 1 and 50 Newtons (N), applied in between the first 10 and 20 % of the episode. The perturbation duration is 10 simulation steps (0.1 seconds). The horizon was set to 200 steps, as the response to a perturbation in humans lasts approximately 2 seconds [32]. Two 20M steps training were run, the first with the perturbation applied in the anterior-posterior (AP) direction and the second with the perturbation applied in the mediolateral (ML) direction. A transfer learning approach was applied; therefore, the training was run on top of the most successful standing balance policy.

The standing balance policy was evaluated over 100 episodes. The standing balance with perturbation policies were evaluated by testing the behavior over 200 episodes. The success rate of the policies was calculated as shown in Eqn. 3. In this case, an episode was considered to be successful if the model didn't fall in the timespan of 700 steps.

$$\%Success = \frac{successfulEpisodes}{totalEpisodes} \cdot 100 \tag{3}$$

III. Results

This section presents the results obtained from the most successful policies after training for standing balance and standing balance with perturbation.

A. Standing Balance Task

A 15-million-step training was performed with the reward function presented in Eqn. 1. Fig. 3 shows the mean reward per episode. As can be seen, the reward increases until it stabilizes at a value of 99. The success rate over 100 episodes was computed for this policy, and the result was 100%.

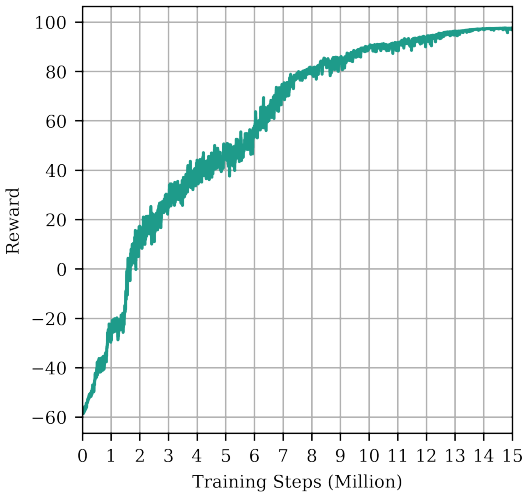

Fig. 3: Evolution of the reward over the training steps of the standing balance policy training.

When performing standing balance, the model is asked to minimize the error in the pelvis with respect to its target position. The mean error and 2 SD in the X, Y, and Z axis are shown in Fig. 4. After the first 100 steps in which the model changes position, it reaches a stable pose in which the error oscillates around 0 cm.

In terms of muscle activity, most muscles show no or very low activation. The activations driving the movement come from the left vastus intermedius, left extensor digitorium longus, right gluteus medius, right and left rectus femoris,

and right and left tibialis anterior. These activations are non-constant, and submaximal, with neither of them exceeding 0.5. See Fig. 5 for an overview of all muscle activations.

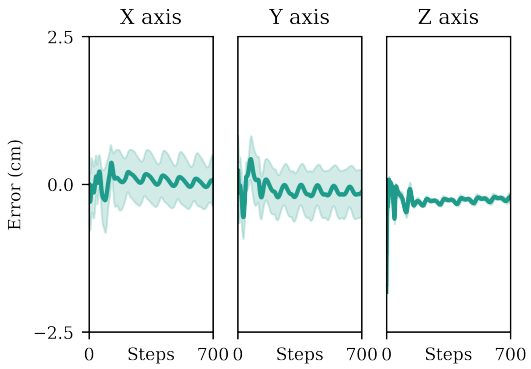

Fig. 4: Mean pelvis error with 2 SDs for the standing balance policy. From left to right: error in the x-axis, error in the y-axis and error in the z-axis.

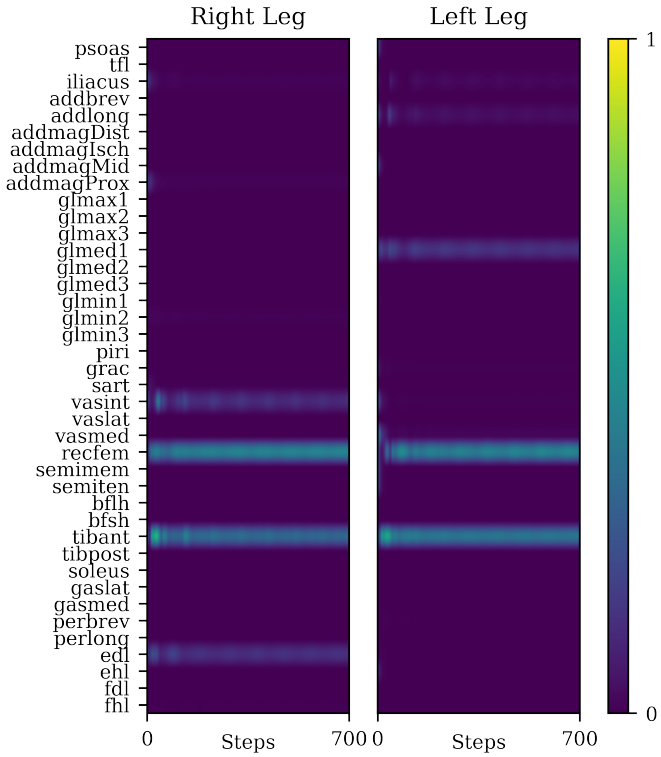

Fig. 5: Mean muscle activations for the standing balance task. The left leg is presented on the right, the right leg is presented on the left. The color map indicates muscle activation levels, with dark blue representing no activation and yellow representing maximum activation.

Joint angles and joint torques are displayed in Fig. 7. The joint angles present a stable pose, which is non-symmetric in the model. Regarding the torques, those coming from the hip flexion and rotation, subtalar flexion, and MTP flexion show small amplitude oscillations around 0 after stabilization of the pose. Hip adduction, knee flexion, and left ankle flexion torque also show oscillations, however, they are not centered around 0 meaning that there is some almost constant moment being applied at those joints.

Last, the position of the center of mass with respect to the base of support was extracted, this is shown in Fig. 6. With the exception of the initial change in the position of the feet, the base of support remains in the same configuration throughout the whole simulation, with a small movement of the center of mass.

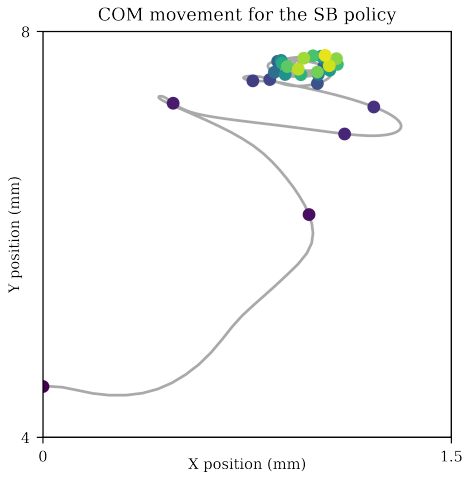

Fig. 6: Zoomed in view of the movement of the COM within the BoS for the standing balance policy. The dots represent the position of the center of mass, turning lighter as the simulation advances.

B. Standing Balance with Perturbation

Similarly to the standing balance task, the standing balance with perturbation policies were evaluated over 200 episodes of 700 steps. In this evaluation, the success rate was computed. For the AP perturbation, it is 96%, while for ML perturbation, the success rate is 71%. Fig. 8 shows the distribution of the perturbation magnitude and the step at which the perturbation is applied. For the AP perturbation, the maximum that can be sustained without falling is 50.0N, although there is a small region of magnitudes for which the model still falls. In the ML direction, the maximum magnitude borderline can be drawn at 33.5N.

As a general note, to ease visualization and analysis, the figures for the standing balance with perturbation tasks are displayed in bins per perturbation magnitude. That is, from 0 to 50 N in increments of 10 N. Note that for the ML perturbation, plots will only reach the range 30 - 40 N, since the maximum perturbation that can be withstood is 33.5N. In the upcoming figures, only the successful episodes are being considered

Figs. 9 and 10 show the mean error in the pelvis and mean center of mass displacement for both perturbation types. Both the displacement and the error begin with an oscillation that then stabilizes, this oscillation is caused by the perturbation applied. Perturbations up to 40N in the AP direction and up to 30N in the ML direction have a very small effect on the

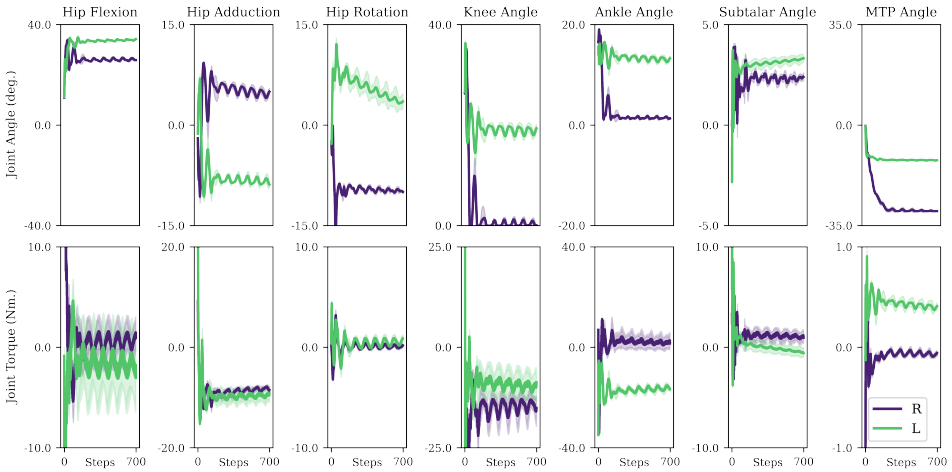

Fig. 7: Mean joint angles and joint torques with 2SD for the standing balance (without perturbation) task. Right leg is shown in purple and left leg is shown in green.

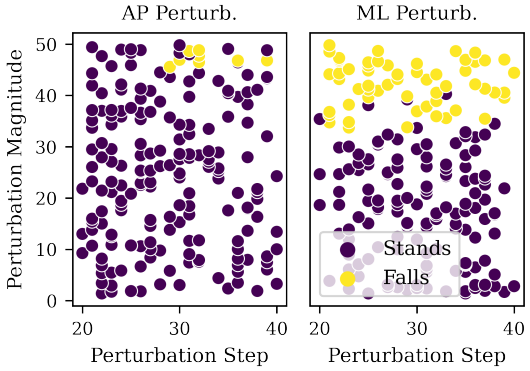

Fig. 8: Distribution of the perturbation magnitude and step in which it was applied. On the left are the results for the AP perturbation, on the right are those of the ML perturbation. Colored in yellow are the combinations for which the model falls.

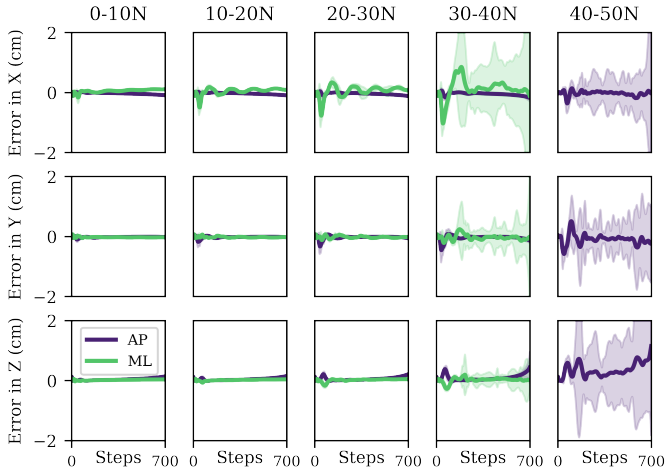

Fig. 9: Mean error in the pelvis with 2SD for the standing balance with perturbation task. AP perturbation is shown in purple and the ML perturbation in green. From top to bottom: errors in the x, y, and z-axis. Each column represents a range of perturbation magnitudes.

error in the pelvis. In the last perturbation range, the effect of the perturbation is larger, although it always stabilizes near 0 cm of error. Regarding the displacement of the COM, up to 10N, it is almost negligible. Perturbations higher than 10N cause a larger displacement of the com, although the model can stabilize near the starting position.

The effect of the perturbations on the muscle activations was analyzed. This is shown in Fig. 12 for the AP perturbation and Fig. 14 for the ML perturbation. For simplicity, only the active muscle pairs are shown, as all the other muscles showed minimal muscle activation. For the AP perturbation, the muscle activations are kept low throughout the whole timespan. For perturbations up to 40 N, the muscle activations remain mostly at the same activation level. However, for the last perturbation range, there is more variability in the activation of the muscles. Specifically, the left adductor longus, right gluteus maximus, and right and left rectus femoris and tibialis anterior show significant activation changes. Similarly, in the ML direction, there is not much change in the activity up to 30N. However, from that point, there is an increase in the variability level of the activation, while always keeping the activation submaximal. The most relevant muscles for this perturbation direction are the left adductor longus, right gluteus maximus, and right and left rectus femoris and tibialis anterior.

Figs. 13 and 15 collect the joint angles and torques for both perturbation directions. In both cases, the joint angles are more symmetric than in the no-perturbation case, however, for the ML perturbation the pose remains more symmetric than for the AP one. For the AP perturbation, similarly to the muscle activations, perturbations up to 40N have very small effects on the model, while the last perturbation range shows big variability both for the angles and torques. Besides,

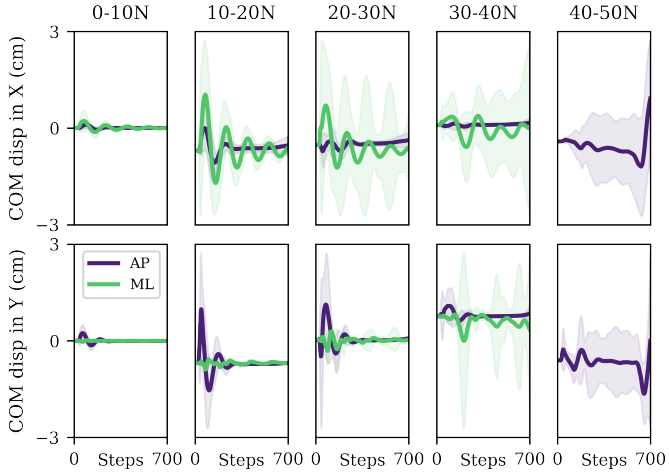

Fig. 10: Mean center of mass displacement with 2SD for the standing balance with perturbation task. AP perturbation is shown in purple and the ML perturbation in green. From top to bottom: displacement on the x and y-axis. Each column represents a range of perturbation magnitudes.

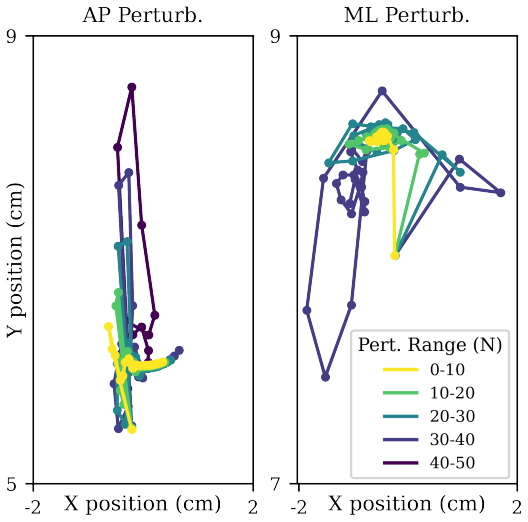

Fig. 11: Zoomed-in view of the movement of the COM for different perturbation ranges. From left to right, the AP and ML perturbation respectively

the oscillatory behavior exhibited in the joint angles and torques of the no-perturbation case is greatly diminished. Most of the torques converge toward 0 or very small values, suggesting that those that do not (knee extension and ankle plantarflexion torques) are the ones driving the movement. In the ML perturbation, both joint angles and torques show oscillations, although it is more remarkable for the joint torques. Both angles and torques show a trend, increasing the amplitude of the oscillations with the perturbation range. Besides, similarly to the AP perturbation case, the variability in the last perturbation range increases both for the angles and the torques. In terms of joint angles, it is interesting to see that both hips are externally rotated, which means that the toes are pointed towards the plane in which the perturbation is applied. Regarding the torques, as in the AP perturbation, most of them converge towards 0. The ones that do not do so are the hip flexion, knee extension, ankle plantarflexion, and right subtalar torque.

Last, the effect of the perturbation magnitude on the COM displacement was also studied. Fig. 11 shows the movement of the center of mass for the AP and ML directions. In both cases, the behavior is similar, with the displacement of the COM being higher in the direction of the perturbation. That is, for the AP perturbation, the displacement is greater in the Y direction and for the ML perturbation, it is higher in the X axis. Although it cannot be seen, the center of mass does not go out of the base of support for any of the successful episodes.

In order to be able to analyze the possible reasons that might be making the model fall, separate plots were extracted and included in Appendix A. Figs. 16 and 17 show the muscle activation pattern and joint angles and torques for the AP perturbation direction while Figs. 18 and 19 do so for the ML perturbation. In terms of muscle activation, in both perturbation directions, the muscles activated for the unsuccessful trials are the same as in the successful ones. The activation patterns are similar for successful and unsuccessful trials until the 200-step mark, time at which the muscles begin to activate more. It is relevant to note that for the AP direction, the mean activations for the unsuccessful episodes are mostly submaximal (except for the adductor longus), while for the ML direction, the right tibialis anterior, right gluteus maximus, left adductor longus, and left gluteus medius reach maximal activations.

In terms of joint torques, similarly to the muscle activations, the torques are stable and similar to the successful episodes approximately until the 200-step mark for both perturbation types. In the AP direction, after that mark, the joint torques show that the hip flexion torque, as well as the left hip adduction and rotation torques, are no longer able to converge to the stable values. In the ML direction, something similar happens, but in this case, it is the left hip flexion, right hip adduction, left hip rotation, and knee flexion torques that are not able to converge. This lack of stability in the torques, has an effect on the joint angles for both perturbation directions, causing them to be highly variable and become asymmetric.

IV. Discussion

The goals of this thesis involve training the myoLeg model to perform standing balance with and without perturbation tasks and analyze its performance.

A. Standing Balance Task

As already mentioned, during training, a deepRL agent aims to maximize the reward obtained. Fig. 3, shows the reward obtained during the training of the standing balance policy. This reward keeps increasing during training until it reaches a maximum value of 99. The stagnation in the reward generally means that the policy has been fully optimized, and thus, the agent has finished learning. The reward reaching this

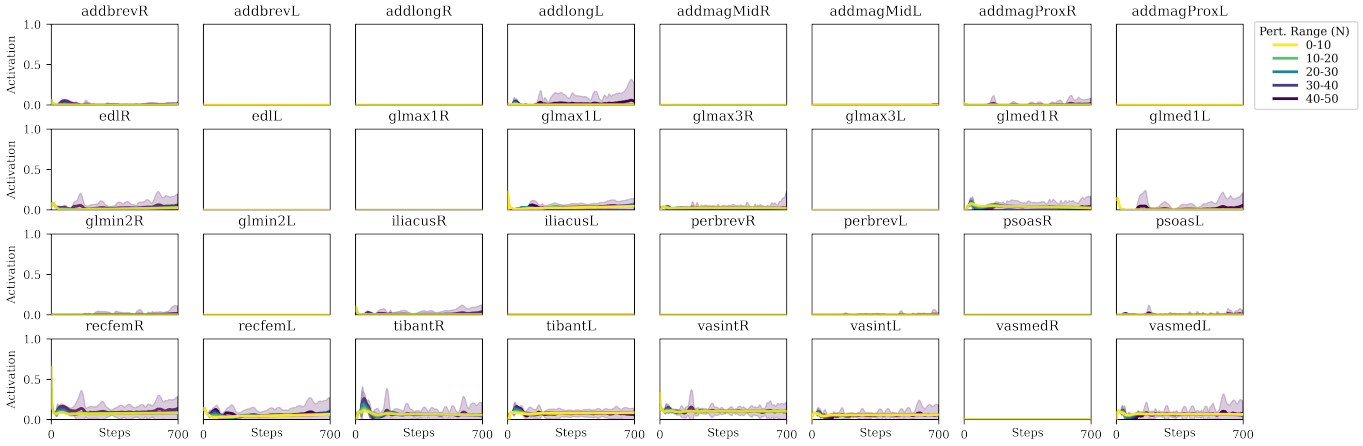

Fig. 12: Mean muscle activation with 2SD per perturbation range for the AP perturbation. For simplicity, only the active muscle pairs are presented, all the others showed minimal muscle activation.

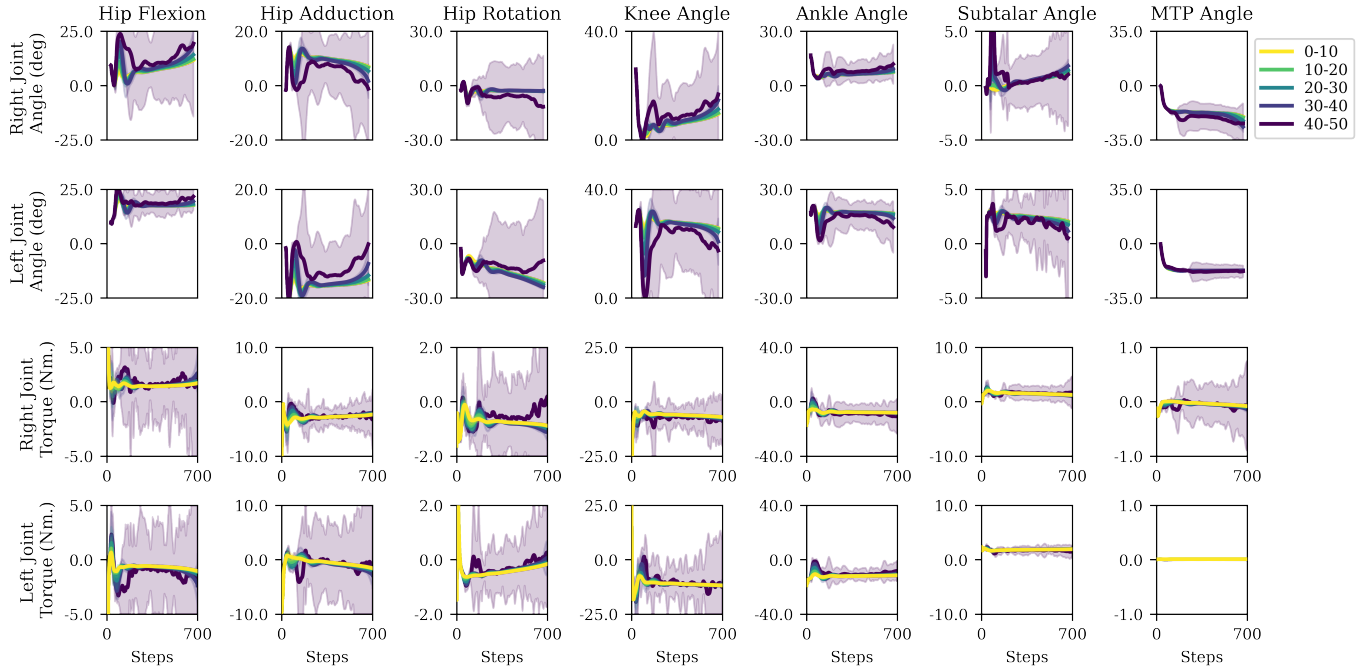

Fig. 13: Mean joint angles and torques with 2SD for the standing balance with AP perturbation task. The top two row shows the joint angles (the first row is the right side, the second row is the left side) and the last two rows show the joint torques (the third row is the right side, the fourth row is the left side).

threshold indicates that the training was sufficiently long to converge on a behavior; whether it is successful or not cannot be exclusively determined by the training reward.

The standing balance policy was evaluated over 100 episodes, giving a success rate of 100%. This means that the agent successfully kept balance for 600 steps on the episodes, which was the minimum that was required for the standing balance task to be deemed accomplished.

The policy is successful in reducing the error in the pelvis with respect to the target position, especially in terms of height, in which the standard deviation is almost null and the error has a mean value of 0.2 mm upon stabilization. There is, however, an oscillatory behavior present in all three axes, which is especially significant in the X and Y directions. A possible explanation for them could be the agent trying to minimize the error in the pelvis in order to obtain a higher reward. A possible solution to prevent them could be a term in the reward function that reduces the movement of the pelvis, such as a penalty for pelvis acceleration.

The muscle activations obtained from the training are low, which is consistent with results obtained in other studies in which low muscle activations are shown to be sufficient to stand [33]. In humans, the central nervous system keeps most of the muscles inactive to reduce the energy expenditure of the task while using a few muscles to modulate the position of the COM. The two main muscles used for this task in humans are

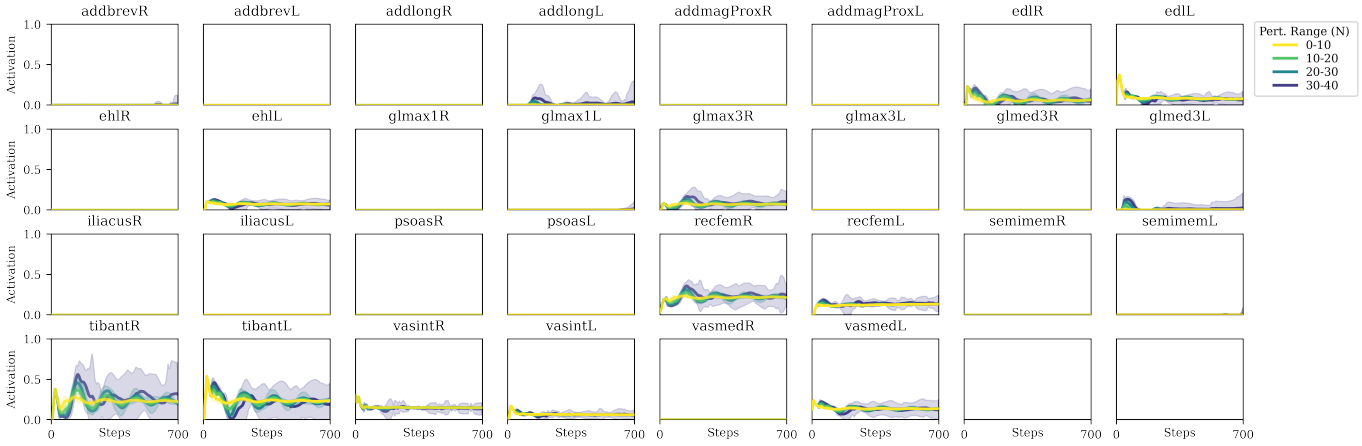

Fig. 14: Mean muscle activation with 2SD per perturbation range for the ML perturbation. For simplicity, only the active muscle pairs are presented, all the others showed minimal muscle activation.

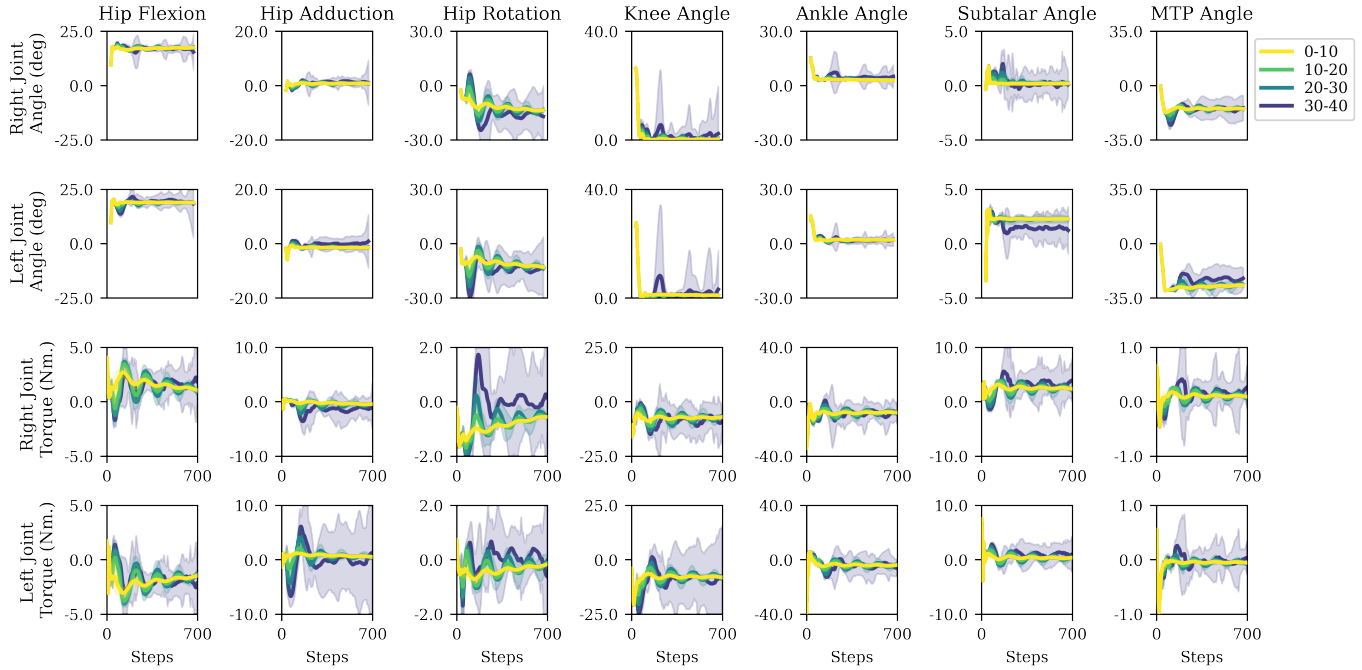

Fig. 15: Mean joint angles and torques with 2SD for the standing balance with ML perturbation task. The top two row shows the joint angles (the first row is the right side, the second row is the left side) and the last two rows show the joint torques (the third row is the right side, the fourth row is the left side).

the rectus femoris and tibialis anterior [33]. Given Fig. 5, the policy shows coherent results, as the activation of the rectus femoris and tibialis anterior are the most prominent ones. The other muscles that are active show a smaller activation when compared to the aforementioned, but they are still influencing the behavior. It is believed that they remain active to prevent the model from collapsing, although they may also be having an effect on the already-mentioned oscillations in the pelvis.

In humans, it is the coordinated action of the ankle, knee, and hip torque that modulates the position of the center of mass in order to maintain balance, although, for a quiet stance, the highest contribution is expected to come from the ankle moment. The ankle torque obtained as a result of the simulation shows significantly different results for the right and left leg, suggesting independent control of each of the limbs. This could be arising from the non-symmetrical pose in which the model is standing, as can be seen in the joint angles. All joint torque shows a constant oscillatory behavior, meaning that the torque is constantly being modulated. Most of them oscillate around 0 Nm., however, the left ankle torque, the knee flexion torque, and the hip adduction torque do not show that behavior. It is believed that these torques are the ones that keep the model from collapsing.

In humans, for balance to be achieved the main condition is to keep the center of mass within the base of support. The result of this policy shows reasonable human-like behavior in

these terms, as apart from an initial change of position, the center of mass remains at the same position with very small variations, as shown in Fig. 6.

B. Standing Balance with Perturbation Task

For standing balance with perturbation tasks, the success rate of the policy decreases with respect to the standing balance without perturbation one. This allowed to conclude the maximal perturbation that the agent can withstand with this training. The agent is able to withstand higher perturbations in the AP than in the ML direction, which is counterintuitive, given that the base of support is larger in the frontal than in the sagittal plane.

In the cases in which the policy is successful, the error in the pelvis position remains small and stable. It is especially remarkable that the oscillations mentioned in the stable position for the standing balance task are greatly reduced, and the mean error remains centered around 0, both for the AP and the ML perturbations. The displacement of the center of mass shows a high deviation in the time region in which the perturbation is applied. However, apart from that, the center of mass displacement remains centered around 0, showing that after perturbation, the model is able to return to the stable position. When analyzing the movement of the center of mass with respect to the perturbation applied, consistent results are obtained. For the AP direction, the larger the perturbation, the larger the movement in the medial plane. For the ML perturbation, the opposite happens. The larger the perturbation, the higher the movement in the frontal plane, as it is the direction in which it is applied.

Response to AP perturbations is believed to be modulated by ankle dorsiflexors and plantarflexors [34] thus, the response seen in Fig. 12 in which the tibialis anterior increases its activation with the perturbation magnitude, is consistent with the expected behavior. The activation of the rectus femoris, vastus intermedius, and the vastus medialis muscles can also be explained, as they are muscles that act in the same plane as the movement generated by the perturbation. Their role in this case is the stabilization of the movement, as well as assisting to keep the center of mass within the base of support. The other active muscles show very low activation levels, with small variations with the perturbation applied. The activation of these muscles is necessary to prevent the model from collapsing.

In the ML direction, the response is normally regulated by changing the hip loading and activating adductors and abductors [34], depending on the direction of the perturbation. In this case, since the perturbation is applied as a push from right to left, it would be expected to have activation of the left abductors and right adductors [35]. The model doesn't behave as expected, given that it is activating the left adductors and the right gluteus. However, it is interesting to see that the activation in the tibialis anterior and rectus femoris is higher and more variable with perturbation force than in the aforementioned muscles. These activations are not surprising, given that it was found in other studies that in humans there are muscle activations of muscles not belonging to the perturbation plane [36]. Given the pose of the model, with both hips in external rotation, it is believed that the tibialis anterior and rectus femoris govern the response to the ML perturbation while keeping the other active muscles as stabilizers.

Regarding the episodes for which the model falls, it depends on the perturbation direction. For the AP direction, the model only falls in 8 out of the 200 episodes that were run for evaluation. There is not a clear limit in perturbation magnitude that makes it fall, as the model can withstand perturbations above and below these. In the ML direction, the number of unsuccessful trials is larger (58 out of 200) and in this case, there is a clear limit from which the model cannot withstand perturbations.

For both perturbation directions, a pattern was found in the unsuccessful trials before and after the 200-step mark. For the unsuccessful ones, the mean muscle activations are similar in pattern to those of the successful ones, although higher in magnitude as the perturbation increases. A similar pattern was found for the joint angles and torques. After 200 steps, the activation levels increase and so do the joint torques, which diverge greatly from the torques in the successful trials. This pattern suggests that the model is able to withstand the perturbation in those 200 steps and that it may be falling from trying to keep balance from a non-stable position.

To rule out the possibility of the model falling because of the influence of the metabolic cost term, a separate training and evaluation step was performed, setting the metabolic cost term to 0. The result of the evaluation showed that the lack of minimization in the metabolic cost made the model fall in 100% of the cases. The response showed maximal muscle activation in 22 muscles, out of which 13 were on the right leg. These activations caused extremely high and asymmetrical torques, that made the model collapse. Overall, this confirms that the metabolic cost term is necessary for the model to be able to withstand perturbations.

In order to find an explanation for why the model falls, we must recall that a transfer learning approach was used, running the training of standing balance with perturbation on top of the standing balance without perturbation policy. Although this policy was 100% successful in keeping balance, it was trained to be able to perform this task from a stable position. It is possible that, after applying the perturbation, the model was not able to return to that stable position with low acceleration of the center of mass that would keep it from deviating highly from it. In that case, the model would not be able to stand, given that it was not trained to learn how to do so.

Limitations to this work arise from the lack of the torso in the simulations, which may introduce discrepancies with the results in the literature. Furthermore, the oscillatory behavior in the pelvis error for the standing balance policy may indicate that an extra term is needed in the reward. This may also affect the standing balance with perturbation policies, given that they were trained on top of the standing balance one.

In terms of future work, addressing the aforementioned limitations is extremely relevant to the performance of the myoLeg model in standing balance. The oscillatory behavior observed in the pelvis can be reduced by adding a corrective term in the reward that penalizes the pelvis acceleration. Furthermore, in order to prevent "early" falls in the standing

balance with perturbation policies, it is suggested to retrain the standing balance policy, with more random starting points in order for it to learn how to return to balance in a non-stable position. As a suggestion, an additional term could be added to the reward function, that gives a bonus for the ground reaction forces being higher than 0, in order to ensure that both feet stay on the ground and prevent changes of pose. Additionally, incorporating the torso in the model will allow for more realistic conclusions as to whether the myoLeg is able to stand as humans do. The addition of the torso will very likely require tuning the reward function, as now the model will have different inertia and mass to balance. Obtaining balance control for the myoLeg model with the torso will allow for more accurate and realistic insight into the human control of balance.

V. Conclusion

The goal of this thesis is to train the myoLeg model to perform standing balance and standing balance with perturbation tasks. This is done with the aim of getting more insight into human balance control.

The myoLeg model was successfully trained to perform standing balance tasks, with a success rate of 100% over 100 episodes. The error in the pelvis and the displacement of the center of mass were minimized, though they both show oscillatory behavior, and the center of mass was continually kept within the base of support. The muscle activations obtained are low in magnitude and consistent with other studies. This indicates that minimal muscle activation is sufficient for standing. The joint torques were not as expected, but they can be explained given the non-symmetrical pose in which the model is standing.

Regarding the standing balance with perturbation tasks, the results were achieved with a success rate of 96% for the AP direction and 71% for the ML direction. This decrease was expected and allowed us to conclude the maximum perturbation magnitude that can be withstood in both directions. The model is able to withstand higher perturbations in the AP than in the ML direction, this is counter-intuitive, given that the base of support is larger in the frontal than in the medial plane. The observed muscle activations were consistent with the literature for the AP perturbations, with the ankle dorsiflexors governing. For the ML perturbation case, the muscle activations were not as expected, and it is the ankle dorsiflexors along with the rectus femoris that are governing the response to those perturbations.

Note from the author

All information related to the code used for this work can be found in its GitHub repository.

References

[1] C. Rival, H. Ceyte, and I. Olivier, "Developmental changes of static standing balance in children," *Neuroscience Letters*, vol. 376, no. 2, pp. 133–136, 3 2005.

[2] P. A. Forbes, A. Chen, and J. S. Blouin, "Sensorimotor control of standing balance," in *Handbook of Clinical Neurology*. Elsevier B.V., 1 2018, vol. 159, pp. 61–83.

[3] R. J. Peterka, "Model-Based Interpretations of Experimental Data Related to the Control of Balance During Stance and Gait in Humans," 2016, pp. 245–270.

[4] S. Sabbah, J. A. Gemmer, A. Bhatia-Lin, G. Manoff, G. Castro, J. K. Siegel, N. Jeffery, and D. M. Berson, "A retinal code for motion along the gravitational and body axes," *Nature*, vol. 546, no. 7659, pp. 492–497, 6 2017.

[5] J. Day, L. R. Bent, I. Birznieks, V. G. Macefield, and A. G. Cresswell, "Muscle spindles in human tibialis anterior encode muscle fascicle length changes," *J Neurophysiol*, vol. 117, pp. 1489–1498, 2017. [Online]. Available: www.jn.org

[6] J. A. Stephens, R. M. Reinking, and D. G. Stuart, "Tendon organs of cat medial gastrocnemius: responses to active and passive forces as a function of muscle length." *Journal of neurophysiology*, vol. 38, no. 5, pp. 1217–31, 9 1975.

[7] Löwenstein. O and A. Sand, "The mechanism of the semicircular canal. A study of the responses of single-fibre preparations to angular accelerations and to rotation at constant speed," *Proceedings of the Royal Society of London. Series B - Biological Sciences*, vol. 129, no. 855, pp. 256–275, 8 1940. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rspb.1940.0039

[8] Löwenstein. O and R. Saunders, "Otolith-controlled response from the first-order neurons of the labyrinth of the bullfrog (Rana catesbeiana) to changes in linear acceleration," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 191, no. 1105, pp. 475–505, 12 1975. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rspb.1975.0140

[9] G. M. Blenkinsop, M. T. G. Pain, and M. J. Hiley, "Balance control strategies during perturbed and unperturbed balance in standing and handstand," 2017. [Online]. Available: http://dx.doi.org/10.1098/rsos.161018Electronicsupplementarymaterialisavailableonlineathttps://dx.doi.org/10.6084/m9.figshare.c.3825463.

[10] T. Bassani and F. Galbusera, "Musculoskeletal modeling," in *Biomechanics of the Spine: Basic Concepts, Spinal Disorders and Treatments*. Elsevier, 4 2018, pp. 257–277.

[11] M. Sartori and D. Farina, "Neuromusculoskeletal modeling for neurorehabilitation technologies," *Biosystems and Biorobotics*, vol. 1, pp. 1217–1220, 2013. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-34546-3_201

[12] D. Tokur, M. Grimmer, and A. Seyfarth, "Review of balance recovery in response to external perturbations during daily activities," *Human Movement Science*, vol. 69, 2 2020.

[13] J. He, W. S. Levine, and G. E. Loeb, "Feedback Gains for Correcting Small Perturbations to Standing Posture," Tech. Rep. 3, 1991.

[14] J. H. J. Allum, D. J. Allum-Mecklenburg, F. P. Harris, R. Probst, A. D. Kuo, and F. E. Zajac, "Human standing posture: multi-joint movement strategies based on biomechanical constraints," Tech. Rep. 153.

[15] A. D. Kuo, "An Optimal Control Model for Analyzing Human Postural Balance," Tech. Rep. 1, 1995.

[16] Institute of Electrical and Electronics Engineers., *Humanoids 07 : proceedings of the 2007 7th IEEE-RAS International Conference on Humanoid Robots : November 29-December 1, 2007, Pittsburgh, Pennsylvania, USA*.

[17] K. Shen, A. Chemori, and M. Hayashibe, "Reproducing Human Arm Strategy and Its Contribution to Balance Recovery Through Model Predictive Control," *Frontiers in Neurorobotics*, vol. 15, 5 2021.

[18] ——, "Effectiveness Evaluation of Arm Usage for Human Quiet Standing Balance Recovery through Nonlinear Model Predictive Control," in *2020 3rd International Conference on Control and Robots, ICCR 2020*. Institute of Electrical and Electronics Engineers Inc., 12 2020, pp. 150–153.

[19] K. Shen, G. Li, A. Chemori, and M. Hayashibe, "Self-organizing neural network for reproducing human postural mode alternation through deep reinforcement learning," *Scientific Reports*, vol. 13, no. 1, 12 2023.

[20] A. Seth, J. L. Hicks, T. K. Uchida, A. Habib, C. L. Dembia, J. J. Dunne, C. F. Ong, M. S. Demers, A. Rajagopal, M. Millard, S. R. Hamner, E. M. Arnold, J. R. Yong, S. K. Lakshmikanth, M. A. Sherman, J. P. Ku, and S. L. Delp, "OpenSim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement," 2018. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1006223.g001

[21] E. Todorov, T. Erez, and Y. Tassa, "MuJoCo: A physics engine for model-based control," *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012.

[22] V. Caggiano, H. Wang, G. Durandau GVDURANDAU, M. Sartori MSARTORI, V. Kumar VIKASHPLUS, R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer, V. Caggiano, H. Wang, G. Durandau, M. Sartori, and V. Kumar,

"MyoSuite: A contact-rich simulation suite for musculoskeletal motor control," *Proceedings of Machine Learning Research*, vol. 168, pp. 1–16, 2022. [Online]. Available: https://sites.google.com/view/myosuite

[23] H. Wang, V. Caggiano, G. Durandau, M. Sartori, and V. Kumar, "MyoSim: Fast and physiologically realistic MuJoCo models for musculoskeletal and exoskeletal studies," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 8104–8111, 2022.

[24] D. L. Yamins, H. Hong, C. F. Cadieu, E. A. Solomon, D. Seibert, and J. J. DiCarlo, "Performance-optimized hierarchical models predict neural responses in higher visual cortex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 23, pp. 8619–8624, 6 2014.

[25] B. A. Richards, T. P. Lillicrap, P. Beaudoin, Y. Bengio, R. Bogacz, A. Christensen, C. Clopath, R. P. Costa, A. de Berker, S. Ganguli, C. J. Gillon, D. Hafner, A. Kepecs, N. Kriegeskorte, P. Latham, G. W. Lindsay, K. D. Miller, R. Naud, C. C. Pack, P. Poirazi, P. Roelfsema, J. Sacramento, A. Saxe, B. Scellier, A. C. Schapiro, W. Senn, G. Wayne, D. Yamins, F. Zenke, J. Zylberberg, D. Therien, and K. P. Kording, "A deep learning framework for neuroscience," pp. 1761–1770, 11 2019.

[26] S. Song, 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," 12 2021.

[27] X. B. Peng and M. van de Panne, "Learning locomotion skills using DeepRL: Does the choice of action space matter?" in *Proceedings - SCA 2017: ACM SIGGRAPH / Eurographics Symposium on Computer Animation*. Association for Computing Machinery, Inc, 7 2017.

[28] J. Merel, Y. Tassa, D. TB, S. Srinivasan, J. Lemmon, Z. Wang, G. Wayne, and N. Heess, "Learning human behaviors from motion capture by adversarial imitation," 7 2017. [Online]. Available: http://arxiv.org/abs/1707.02201

[29] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction Second edition, in progress," Tech. Rep., 2015.

[30] V. Vijayan, "Deep Reinforcement Learning: Value Functions, DQN, Actor-Critic method, Back-propagation through stochastic functions," 6 2021. [Online]. Available: https://shorturl.at/yQWX3

[31] 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, 10 2016.

[32] M. Mihelj, Z. Matjačic, M. Matjačic´b, and T. Bajd, "Postural activity of constrained subject in response to disturbance in sagittal plane," Tech. Rep., 2000. [Online]. Available: www.elsevier.com/locate/gaitpost

[33] S. A. Chvatal and L. H. Ting, "Common muscle synergies for balance and walking," *Frontiers in Computational Neuroscience*, no. APR 2013, 4 2013.

[34] D. A. Winter, F. Prince, J. S. Frank, C. Powell, and K. F. Zabjek, "Unified theory regarding A/P and M/L balance in quiet stance," *Journal of Neurophysiology*, vol. 75, no. 6, pp. 2334–2343, 1996.

[35] S. Rietdyk, A. E. Patla, D. A. Winter, M. G. Ishac, and C. E. Little, "Balance recovery from medio-lateral perturbations of the upper body during standing," Tech. Rep., 1999.

[36] S. M. Henry, J. Fung, F. B. Horak, and R. S. Dow, "EMG Responses to Maintain Stance During Multidirectional Surface Translations," Tech. Rep., 1998.

APPENDIX A
FIGURES INCLUDING THE FALLEN TRIALS

This appendix includes the mean muscle activations, mean joint angles, and mean joint torques figures used for the analysis of possible reasons that might be making the model fall. For comparison, the figures include the mean per perturbation range (only four ranges are shown) and the mean of the unsuccessful episodes. As in the other muscle activation figures, only the active muscles are shown, as all the other show minimal muscle activations.

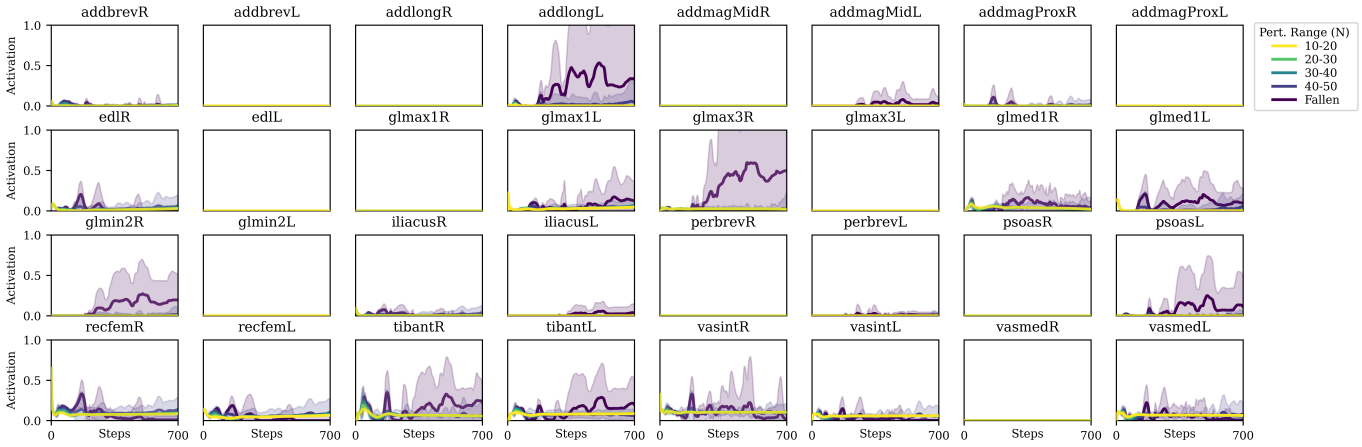

Fig. 16: Mean muscle activation with 2SD for AP perturbation including the unsuccessful trials.

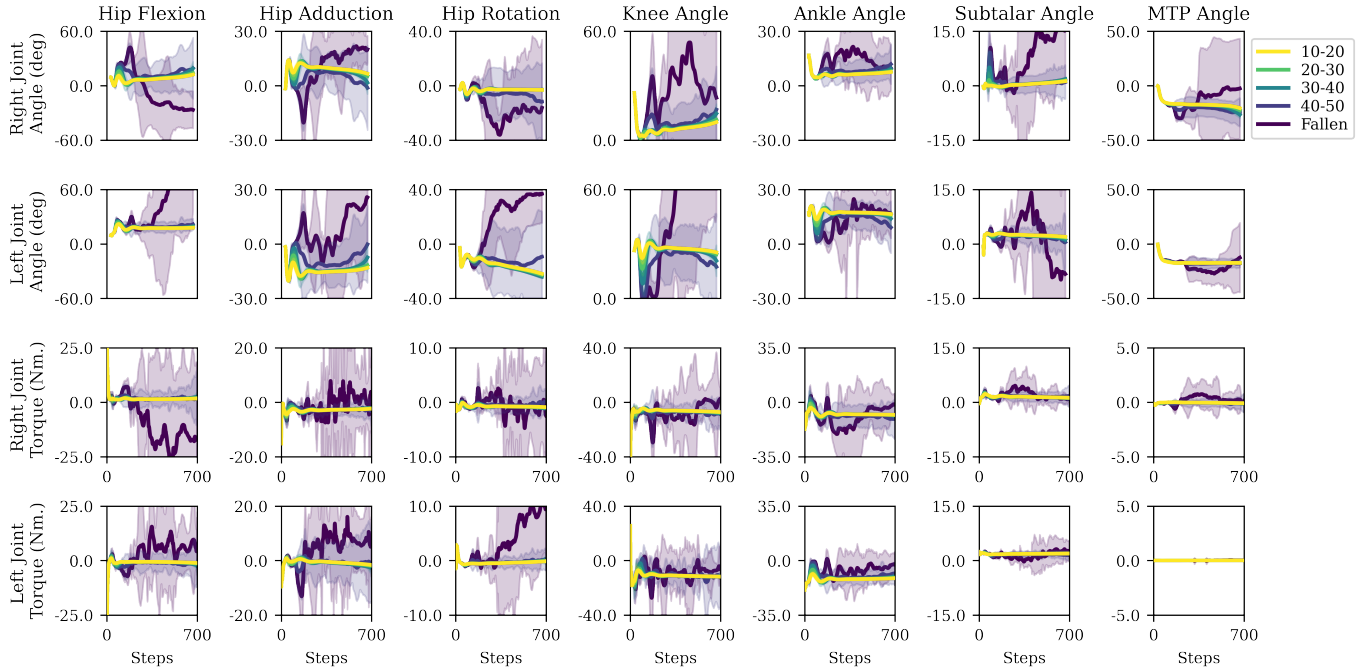

Fig. 17: Mean joint angles and torques with 2SD for the standing balance with AP perturbation task including the unsuccessful trials. The top two row shows the joint angles (the first row is the right side, the second row is the left side) and the last two rows show the joint torques (the third row is the right side, the fourth row is the left side).

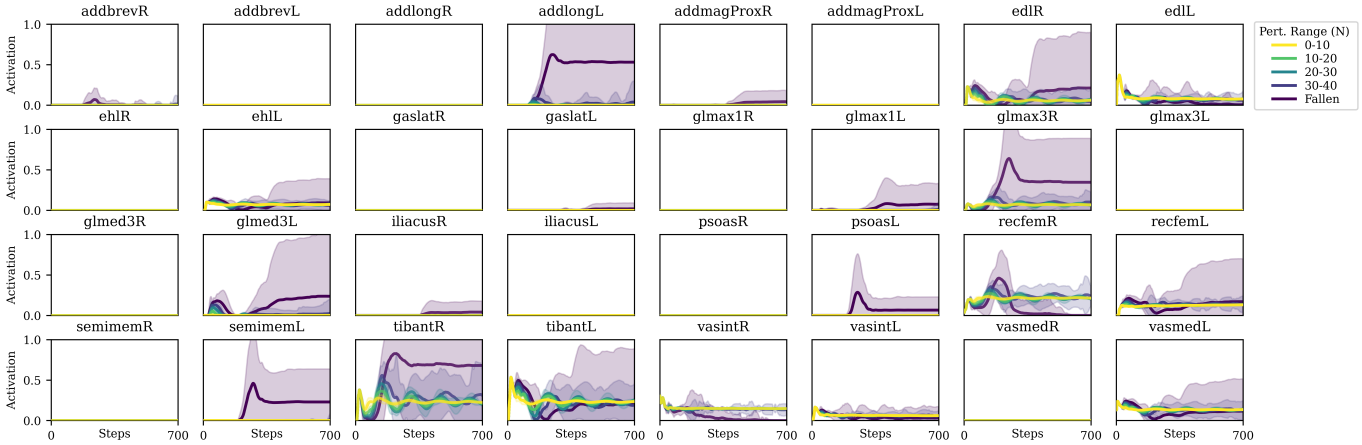

Fig. 18: Mean muscle activation with 2SD for ML perturbation including the unsuccessful trials.

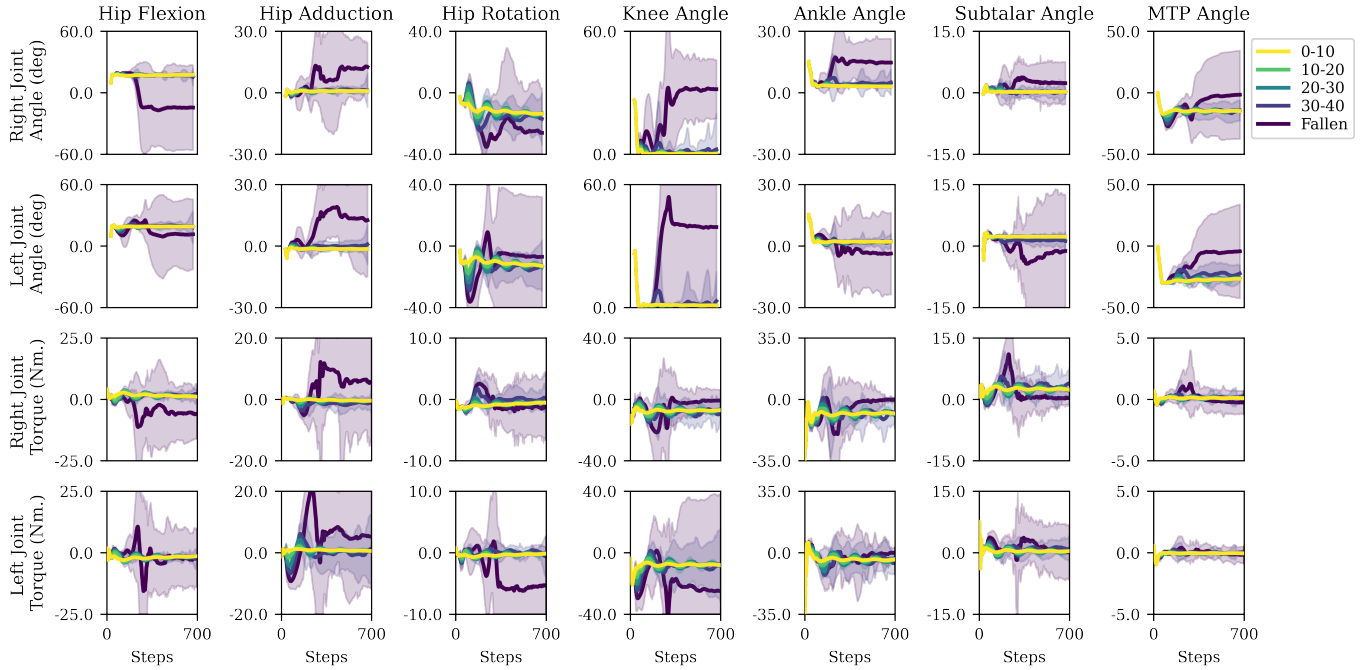

Fig. 19: Mean joint angles and torques with 2SD for the standing balance with ML perturbation task including the unsuccessful trials. The top two row shows the joint angles (the first row is the right side, the second row is the left side) and the last two rows show the joint torques (the third row is the right side, the fourth row is the left side).