Reproducing Synergetic Human Balance Recovery Modalities through Predictive and Learning Control Strategies

Doctoral Thesis

Thesis Title

Reproducing Synergetic Human Balance Recovery Modalities

through Predictive and Learning Control Strategies

Department of Robotics

Graduate School of Engineering

TOHOKU UNIVERSITY

KELI SHEN

Advising Professor at Tohoku Univ.	Professor Mitsuhiro Hayashibe
Research Advisor at Tohoku Univ.	–
Dissertation Committee Members Name marked with "○" is the Chief Examiner	○ Prof. Mitsuhiro Hayashibe 1 Prof. Kazuya Yoshida 2 Prof. Yasuhisa Hirata 3 Dr. Ahmed Chemori (*LIRMM, University of Montpellier, CNRS*)

TOHOKU UNIVERSITY

Graduate School of Engineering

Reproducing Synergetic Human Balance Recovery Modalities through Predictive and Learning Control Strategies

(協調的な人間のバランス回復モダリティ再現のための予測と学習制御戦略)

A dissertation submitted for the degree of

Doctor of Philosophy (Engineering)

Department of Robotics

by

Keli SHEN

July 9, 2021

Reproducing Synergetic Human Balance Recovery Modalities through Predictive and Learning Control Strategies

Keli SHEN

Abstract

Postural balance control for standing is a primary function required by humans in daily life. However, with the emergence of a super-aging society, the postural balance control ability of humans is reducing owing to the sensory-motor impairments caused by aging. An efficient solution to this problem is to understand the balance mechanism. The adaptation ability of human postural control is a key research topic in this field. The adaptability of human postural control has not yet been reproduced via computational modeling. This motivated us to conduct the work presented in this thesis. Numerous studies have investigated balance control, and different strategies have been proposed. In our previous work, we have made the following contributions to this research topic: the reproduction of human balance recovery modalities through model predictive control (MPC) and deep reinforcement learning (DRL) and motion mechanism analysis through the recruitment of multisegment postural coordination. There are three main strategies for balance control, i.e., the (*i*) ankle strategy, (*ii*) hip-ankle strategy, and (*iii*) stepping strategy. However, in recent years, it has been confirmed that the arm strategy improves the balance control ability for standing and walking. These strategies are considered to be efficient for balance control.

This thesis describes our work on the reproduction of synergetic human balance recovery modalities through predictive and learning control strategies. Predictive and learning control are considered efficient and necessary methods for controlling human motor behavior. Predictive control is consistent with the human motor control determined by the central nervous system (CNS); it is based on the predictive models of the human body and the environment with which humans interact. In an unknown environment, learning control can be applied for motor behavior adaptation. There are four contributions of this work toward understanding the mechanisms of balance control through computational modeling and simulations and postural coordination analysis.

The first contribution was the reproduction of the hip–ankle strategy for quiet standing balance recovery through MPC. A three-link model, including a foot with unilateral constraints, a lower body, and an upper body, was built. Then, we derived and linearized a dynamical equation. The bound constraints of the model were set on the basis of the

human balance capability; these included the angles and balance torques of the ankle and hip. Unilateral constraints were applied to the foot; this made the model similar to a human quiet standing case. Numerical MPC was proposed to predict the best method of maintaining balance under different disturbance forces. Finally, we reproduced the proposed ankle and hip–ankle strategies via simulation and analyzed the results using kinematic and dynamic indices. Additionally, we examined the energy consumption. The robustness of the controller was verified. This study may help obtain a better understanding of human quiet standing balance control using a simplified model.

The second contribution was the reproduction of the arm strategy for quiet standing balance recovery through nonlinear MPC (NMPC). A literature review showed that several studies have investigated the improvement of the balance control ability during quiet standing and walking motions. In imperial literature, arm usage (arm strategy) has been proposed to control balance during walking motion. However, limited research exists on the contributions of the arm strategy to balance recovery during quiet standing through predictive control along with the ankle and hip strategies. Therefore, in this study, we built a simplified model with arms and proposed a controller based on NMPC to achieve human-like balance control using the arm strategy. Three arm states of the model, namely, active arms, passive arms, and fixed arms, were considered to examine the contributions of arm usage to human balance recovery during quiet standing. Furthermore, various indices, such as the root mean square deviation of joint angles, CoM phase cycles, and recovery energy consumption, were used to reveal the mechanism of the arm strategy. We computationally reproduced human-like balance recovery with and without arm rotation during quiet standing while applying different magnitudes of perturbing forces on the upper body. Furthermore, we compared the results of simulations and human balance recovery experiments to verify the contributions of the arm strategy. Finally, the arm strategy with active arm usage was confirmed as an efficient balance recovery approach by comparing the motion intensity, energy consumption, and synergistic levels of the motor command.

The third contribution was that the synergy analysis of multisegment balance control was executed in human experiments and simulations to analyze postural balance control features, which were classified as the corresponding ankle, hip, and arm strategies for static postural balance recovery. The human body behaves as a coordinated unit rather than a collection of decoupled degrees of freedom (DoFs) for completing motor tasks. Muscle synergy could be a method of controlling several DoFs simplified by the CNS. Two types of synergy were applied to understand the mechanisms of balance recovery motion on the basis of different simplified models through human experiments and simulations. Different balance control coordination modes were produced according to the limited human body joint recruitment combined with the constraints of the angle and torque of joints. These modes were classified as corresponding balance recovery strategies. The relationship between the type of synergy and control strategy can help understand which

control strategy is being employed by a subject through motion observation. This scheme would be beneficial for understanding human motion during rehabilitation training.

The fourth contribution of this work was that postural coordination modes were reproduced in head-target tracking tasks through model-free DRL. Kinematic and dynamic constraints were applied to the model and DRL reward function to explore the random behavior for different tracking frequencies during the training process. After training, different postural behaviors were produced for different frequencies using model-free DRL. The transitions between postural coordination types, i.e., in-phase and antiphase modes, existed in dynamic postural control tasks. These modes were considered as emergent phenomena that were influenced by the target tracking frequencies. First, the transition between the in-phase and antiphase modes was distinguished by the joint correlation of the ankle and hip. In addition, the temporal joint angles and phase portraits of the hip and ankle were compared to classify the motion transition depending on frequencies. Finally, the CoM deviations were compared for different frequencies to analyze the motion features. It is worth noting that the coordination mode transitions contributed to a decrease in the CoM deviation. It is reasonable that for high-speed tracking tasks, the agent should reduce the CoM deviations to achieve a stable tracking performance. In this study, postural mode emergence was generated without assuming a prior model regarding body dynamics and using only a learning scheme. The learning process is essential for postural control, particularly for conditions in which body characteristics may change because of growth or aging. This is important for introducing hyperadaptivity in postural control.

Acknowledgments

I wish to express my sincere gratitude to my Professor Mitsuhiro Hayashibe, for providing me with the opportunity to work at the Neuro-Robotics Lab, and for his continuous support and guidance during my doctor's degree. I want to make appreciation to my co-supervisor Professor Ahmed Chemori, for his dedication and contribution to my research.

I would also like to thank my thesis committee members, Professor Mitsuhiro Hayashibe, Professor Kazuya Yoshida, and Professor Yasuhisa Hirata, and my co-supervisor Professor Ahmed Chemori, for their thoughtful questions and valuable comments regarding this research project.

I am grateful to all my colleagues and members of the Neuro-Robotics Lab, especially Mr. Guanda Li, Jiazheng Chai for his kind support, interesting discussions and advice regarding my research. Besides, thank Prof. Dai Owaki for his help for letting me to access to motion capture system for motion analysis.

Finally, to my parents, wife (Fang Chen) and daughter (Jihe Shen). Thank you for the love and support over the years and letting me follow my heart.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

World Population Prospects 2019: Summary [1], published by the Population Division of the United Nations Department of Economic and Social Affairs, provides a comprehensive overview of global demographic patterns and prospects. The report confirms that the world's population is ageing and the number of countries with declining population sizes is increasing due to increasing life expectancy and declining fertility levels (see Figure 1.1). The ensuing changes in the size, composition and distribution of the world's population have important implications for achieving the Sustainable Development Goals (SDGs), which are globally recognized goals for improving economic prosperity and social well-being while protecting the environment. The world's population is getting older, with age groups of 65 years for both men and women, and is growing fastest (see Figure 1.2).

In order to tackle the problems faced by an ageing society, it is important to develop rehabilitation programs to improve the standard of living and extend the life span of the elderly. Rehabilitation is the care needed when daily functions are or may be limited due to ageing or health conditions such as chronic diseases, disabilities, injuries, and trauma. Examples of limitations in functioning are difficulty thinking, seeing, hearing, communicating, moving around, having relationships, or keeping a job. When provided after other interventions such as surgical care, rehabilitation not only contributes to realizing the full benefits of the intervention but also has preventive features in terms of avoiding complications and future difficulties in daily functioning. Health and demographic changes in the increasing prevalence of non-communicable diseases and the ageing of the population

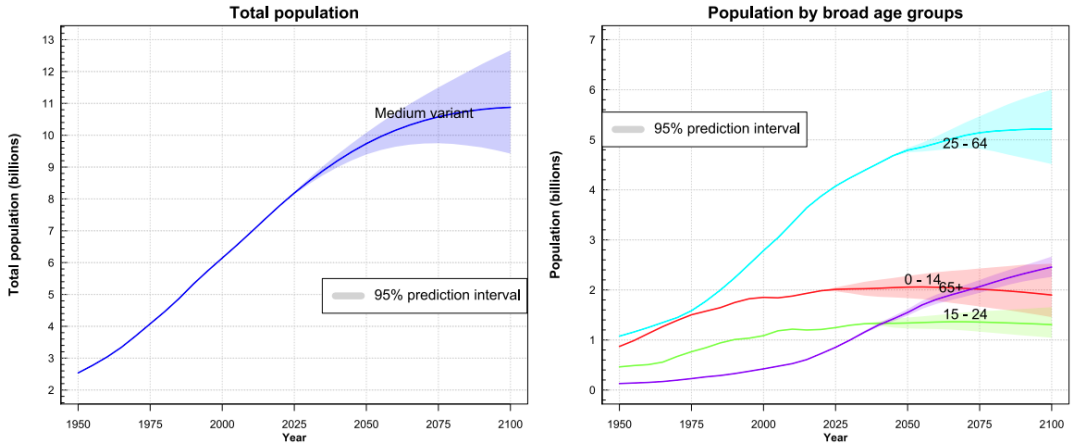


Figure 1.1: The line charts of total population and population by board group [1]. The ratio of the population above 65 will increase from 7 per cent now to 26 per cent by 2100. While the population ratio of children will maintain low increment.

have already contributed to the rapid global increase in the number of people experiencing a functional decline. Thus, the already very high need for rehabilitation will increase further in the coming years. This is a major challenge for health and social systems in addressing these needs.

Rehabilitation addresses the impact of health conditions on a person's daily life by optimizing function and reducing the experience of disability. Rehabilitation expands the focus of health beyond preventive and curative care, enabling people with health conditions to be as independent as possible and to participate in education, work, and meaningful life roles. Everyone may need rehabilitation at some point in their lives, whether they have experienced an injury, illness, or disease, or whether their functioning has declined with age.

Globally, it is estimated that one in three people today live with a health condition that would benefit from rehabilitation. This global need for rehabilitation is projected to increase in the coming years due to changes in the health and characteristics of the population. For example, people are living longer, but with more chronic diseases and disabilities. In addition, many people are living with medium- to long-term consequences of COVID-19 and need rehabilitation to recover from their illnesses. Many countries are not prepared to meet the existing rehabilitation needs. In some low- and middle-income countries, more than 50 per cent of people do not receive the rehabilitation services they need.

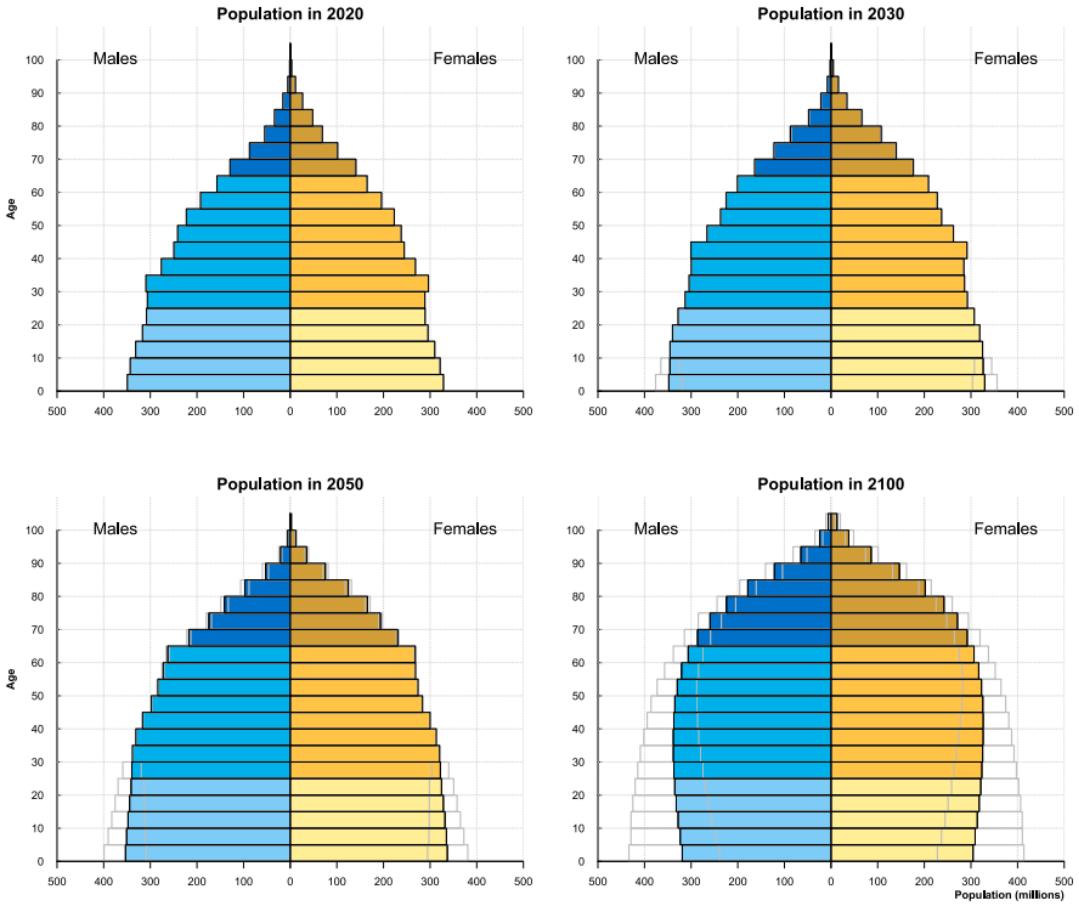


Figure 1.2: The population pyramids of the population of males and females [1]. The population above 60 will increase a lot by 2100. While the population ratio of the young will decrease.

In 2017, WHO launched the Rehabilitation 2030 initiative [2]. It emphasizes the need for coordinated and concerted action by all stakeholders to provide quality and timely rehabilitation through actions such as Develop a strong interdisciplinary rehabilitation workforce. Expand funding for rehabilitation. Improve data collection and research on rehabilitation. WHO will continue to support countries to implement these actions through increased provision of technical support, development of guidance and practical tools, and development of robust evidence for rehabilitation.

In this thesis, we make more effort to study the mechanism of balance recovery through numerical systems. Different balance recovery strategies will be introduced and the related work will be mentioned as the background of our study.

1.1 Balance Strategies

1.1.1 Hip-ankle Strategy

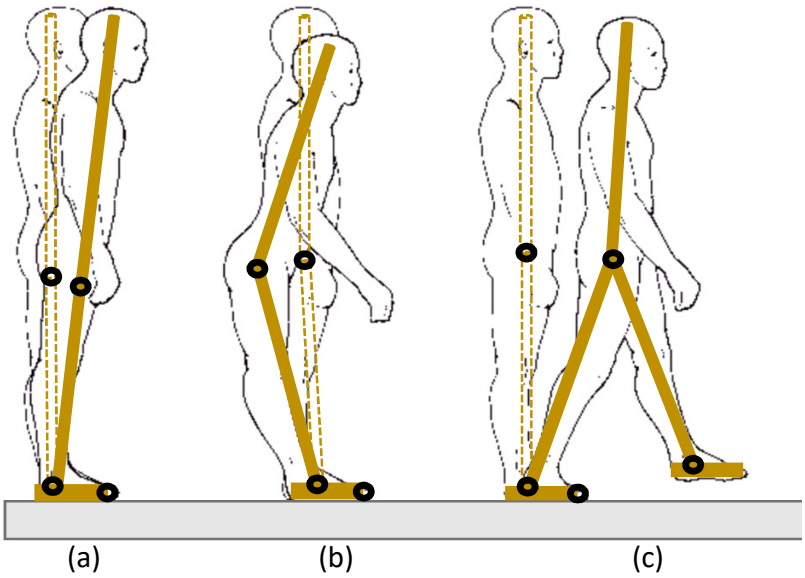


Figure 1.3: The three balance strategies with different external disturbing forces in A/P direction: (a) ankle strategy, (b) hip strategy and (c) stepping strategy in [3]

.

Human balance recovery is an important topic in human rehabilitation area. Human balance has been extensively studied for many years. First, three main balance strategies were observed in human push-recovery experiments with different extents of perturbation in the anterior-posterior (A/P) direction. For a small pushing force on the back, the human tries to move the ankles (”ankle strategy”) to keep balance while keeping the knees, hips and neck straight. For a larger pushing force on the back, the human tends to rotate the hips (”hip strategy”) when the ankle movement is not large enough to keep balance. Besides, in case the human cannot keep quiet standing, the human does need to step forward (”stepping strategy”) for balance recovery. Three balance strategies are illustrated in Figure 1.3. Nevertheless, in some special environments, the human cannot step. In quiet standing cases, ankle strategy and hip strategy are two choices for human balance control. After obtaining rules from human balance experiment, many researchers tried to model this kind of human posture behavior for further motor learning and control.

In the literature one can find many related works about human balance control, both in clinical human experiments and numerical simulations.

Vukobratovic *et al.* stressed the importance of artificial locomotion systems for rehabilitation equipment design [4] and proposed the concept of Zero Monment Point (ZMP) as biped locomotion stability criteria [5]. Hemami *et al.* used nonlinear feedback to linearize a compound inverted pendulum system for postural stability analysis [6]. Goddard *et al.* studied single-support postural stabilization of a three-link planar model of a biped while considering the system state and input constraints in the frontal plane [7].

Gatev *et al.* proposed to use a feedforward to assess strategies of balance during quiet standing and evaluate the influence of narrow stance width and absence of sight. Feedforward controller predicts an external input or behaves using higher-order processing rather than simple negative feedback of a variable [8]. The postural responses to unexpected small and slow or large and faster disturbances during quiet standing are defined as 'ankle strategy' and 'hip strategy' [9], respectively. The choice of a balance strategy to disturbance is based on the available suitable sensory data [10]. However, Kuo et al. proposed a mixed hip-ankle strategy in the anteroposterior (A/P) direction; implemented on biomechinical optimization models instead of a pure ankle strategy to recover postural balance from different disturbances. Considering the big moment of inertia of the whole body and the difficulty of separating the control of ankle and hip joints, the objective was to minimize the neural effort and the prediction was based on the limited effort of the ankle joint torque to recover balance from unstable postures [11], [12]. The studies of quiet standing [13]–[15] declared the strategies of postural balance recovery should be divided into two coordinations: anteroposterior (A/P) and mediolateral (M/L). A/P balance studies include ankle strategy, hip strategy and stepping strategy. M/L balance studies only concentrate on sway motion control [15]–[18].

Pai *et al.* [19] predicted the Center of mass (CoM) velocity and position of an inverted pendulum with a foot segment within the size of the base of support (BOS) for balance recovery from forward and backward falls. Kajita *et al.* [20] used a preview control of zero-moment point (ZMP) to generate biped walking pattern. A ZMP tacking servo controller based on the future reference preview control was designed to compensate the eventual ZMP error caused by the difference between the amplified model and the precise multibody model. Azevedo *et al.* [21] studied the human walking and balance strategies and proposed 'Trajectory-Free Nonliear Model Predictive Control' to simulate various

walking and also the stable standing situations. Hofmann [22] studied humanoid walking and balance control in his thesis and highlighted the importance of horizontal motion control of the CoM for balance recovery. Stephens [23]–[26] studied humanoid push balance recovery during walking and quiet stance from multiple aspects: balance indices, control methods, etc. However, he did not consider the balance recovery with long-time disturbances and under-actuated humanoid feet in simulations. Liu *et al.* [27] proposed a balance controller based on a trajectory library used for nonlinear systems control with constraints, such as a humanoid standing balance control. Kiemel [28] demonstrated that a humanoid robot can keep balance for larger disturbances using the hip-ankle strategy than using the ankle strategy from the first experimental evidence and used a bang-bang control to implement the proposed hip-ankle strategy. Nenchev [29] studied the decision of ankle and hip strategies for balance recovery depending on acceleration data measured during the impact. Aftab *et al.* [30]–[33] proposed a multiple step balance recovery scheme based on Linear Model Predictive Control (LMPC) by minimizing the horizontal CoM velocity and angular velocity of the flywheel, including the use of a hip strategy and a variable step duration to correct large perturbations on the inverted pendulum or inverted pendulum plus flywheel, so the typical kinematics of human hip strategy cannot be observed. Choi *et al.* [34] studied a trajectory-free reactive stepping controller using momentum control. The proposed controller is able to make a humanoid model moving passively without following a planned trajectory in the direction of the disturbance and achieve a natural stepping for different pushes. Ashtiani *et al.* [35] developed a control scheme based on Model Predictive Control (MPC) and Capture Point (CP) for balance recovery with pushes. The proposed MPC was used to guide the CP to a desired position by regulating the ZMP and the Centroidal Moment Pivot (CMP). Penco *et al.* [36] developed a retargeting framework to make the iCub robot repeating the human operator's motion for keeping whole-body balance. Yamamoto [37] studied the maximal output admissible (MOA) set of a CP feedback controller for adaptive humanoid balance with external disturbances in both M/L and A/P directions.

Comparing the previous work, we found that they did not mention about long-time disturbing forces, robustness of MPC, energy consumption of hip and ankle joints, the evolution of Ground Reaction Force and feet with unilateral constraints. The contributions of our study are as follows.

1) We built a three-link simplified human model where foot is unilaterally constrained to remain in contact with the ground. This makes our model balance more human-like.
2) A Numerical MPC with system states and control constraints is proposed to implement a human-like balance strategy and autonomous switch between ankle strategy and hip-ankle strategy during quiet standing balance for the different disturbing forces. This illustrates that N-MPC is similar to the behavior employed by the human brain and nervous system from neuroscience viewpoint. N-MPC is also endowed with prediction aspect that enables to predict future behavior and select control balance strategy through minimization of the energy consumption of the whole body. N-MPC can handle state and input constraints simultaneously. This is crucial for realistic requirements since body limitations including joint ranges and input torques can be considered. N-MPC is also a robust controller able to optimize the balance strategy and to deal with different external disturbances, such as small and large, short-time and long-time disturbing forces.
3) The Center of Mass (CoM) and the Center of Pressure (CoP) are used as evaluating indices and constraints for both upright. Different disturbing forces are considered: namely small and large disturbing forces, short-time and long-time disturbing forces.
4) From the obtained simulation results, we analyzed the mixed hip-ankle strategy regarding two aspects: kinematic and dynamical indices, energy consumption aspect. Also, we tested the robustness of the proposed controller. And we verified our model and control approach can implement much more human-like balance behavior. In this way, motor control on ankle and hip may become more clear and the understanding provides a very efficient guidance for the elderly's quiet standing balance.

1.1.2 Arm Strategy

Balance control mechanism of human has been researched to enhance balance ability of human and humanoid robots [13]. In specific, principal balance recovery strategies, namely, ankle, hip, and stepping strategies have been studied based on human experiments [9], [38], [39] and artificial systems [3], [11], [12]. These strategies have been considered as efficient means to help preventing falls and analyze the mechanism of balance control during standing and walking motions in human rehabilitation and humanoid robot control. For instance, human upright posture (UP) dynamic stability with a simplified inverted

model or hip-ankle model has been studied based on bifurcation analyses to improve balance ability relared to fall prevention and rehabilitation [40]–[42]. Additionally, arm strategy has been considered as an efficient means to contribute to balance control and reduce the effects of a fall [3], [33], [43], [44].

Many studies related to the arm strategy have been conducted through human experiments and simulations. Nashner *et al.* [45] studied rapid postural adjustment associated with a class of voluntary movements, including arm rotation, that disturb the postural balance. Ledebt [46] concluded that arm postures help stabilize the body to maintain the upright position and that balance control improves because of the arm movement. Furthermore, he considered maximization of the gait efficiency based on an organism's propensity for convergence toward a stable coordination between the arms and legs. Atkeson *et al.* [23] studied optimal control with boundary constraints from one optimization criterion to realize a multi-link model balance control and observed the movement of shoulder joints for the different perturbations. Aoustin *et al.* showed that arm swinging can help minimize the energy consumption during walking [47]. Nakada *et al.* [48] reviewed the mechanism of arm strategy for balance recovery and proposed Q-learning to produce appropriate arm control torques for humanoid. They concluded that the arm rotation strategy can widen the range of perturbation impulses. Bruijn *et al.* [49] studied the influence of arm swinging on balance control for a perturbation as well as the local and global stability of the steady-state gait and concluded that arm movements contributed to the overall stability of human gait. Milosevic *et al.* [50] estimated the effectiveness of arm motions in clinical balance and mobility. Wagner [51] verified that in a dynamic balance task during challenged locomotion, the contribution of the upper body motions, particularly the one of arm movements, to human balance regulation increases with the difficulty of the task. The considered balance recovery tasks are in anteroposterior (A/P) direction. Objero [52] showed that arm movements are important for the control of mediolateral (M/L) postural sway, based on human experimental data.

It is worth noting that all the previous works did not cover the verification of the arm strategy with multiple cases, e.g., active arms, passive arms, and fixed arms, in their human experiments to discuss the usefulness of arm rotations. To our best knowledge, these arm strategies are relevant for stability improvement and energy efficiency in human and

humanoid/bipedal walking and standing. Furthermore, the previous work did not leverage nonlinear model predictive control (NMPC) for addressing multiple constraints of the ankle, hip, arm joint angles and torques and reproducing human-like balance recovery controller in their artificial systems. The features of NMPC consistent with the capacity of the human body and brain such as constraints handling, predictive horizon, optimization, and robustness are not considered very well in all the previous work.

Therefore, we further developed the mechanism of arm strategy for balance recovery based on previous works and compared the results with human balance recovery experimental results. The contributions of our study are summarized as follows.

1) A three-joint, five-link model is built to represent the human body structure for studying quiet standing balance recovery in the A/P direction. This model includes the foot, the lower body, the upper body, and the arms.

2) With the system states and the input constraints, an NMPC is proposed from a neuroscience perspective to reproduce human-like balanced behavior evoked by the human central nervous system. The proposed NMPC also has a predictive aspect that allows predicting the future behavior and computer an optimal control balance strategy by minimizing systemic energy consumption of the whole body. Furthermore, the NMPC technique can handle simultaneously the state and input constraints, which is important to meet realistic requirements due to physical limitations of the human body such as joint ranges and torques satuation. All the previously proposed control techniques can not take into account constraints naturally. In this research, we proposed NMPC which can naturally take into account constraints. Different magnitudes of disturbing forces are applied to the model to observe the autonomous switch between the ankle, hip, and arm strategies and to examine the robustness of the proposed solution.

3) Various indexes are verified to evaluate the capability of balance recovery. The root mean square (RMS) deviation and energy consumption are compared for different cases, namely, active arms, passive arms, and fixed arms. These three cases of arm usages are recruited for balance recovery. The obtained data indicate that balance recovery with active arms is the most effective strategy, and balance control with arm usage is better than that without arm usage.

4) Phase portraits of joint angles are considered to analyze the control pattern of balance recovery motion.

5) Ankle torque boundary constraints are set with different values. Besides, the relationship between ankle capacity and active arm usage is discussed since in our daily life ankle is easy to be injured, we want to observe how arm usages contribute to balance in this case.

6) By comparing the results of the numerical simulation and human experiments, human-like balance recovery with arm strategy is implemented and arm movements are found to enhance the capability of balance recovery.

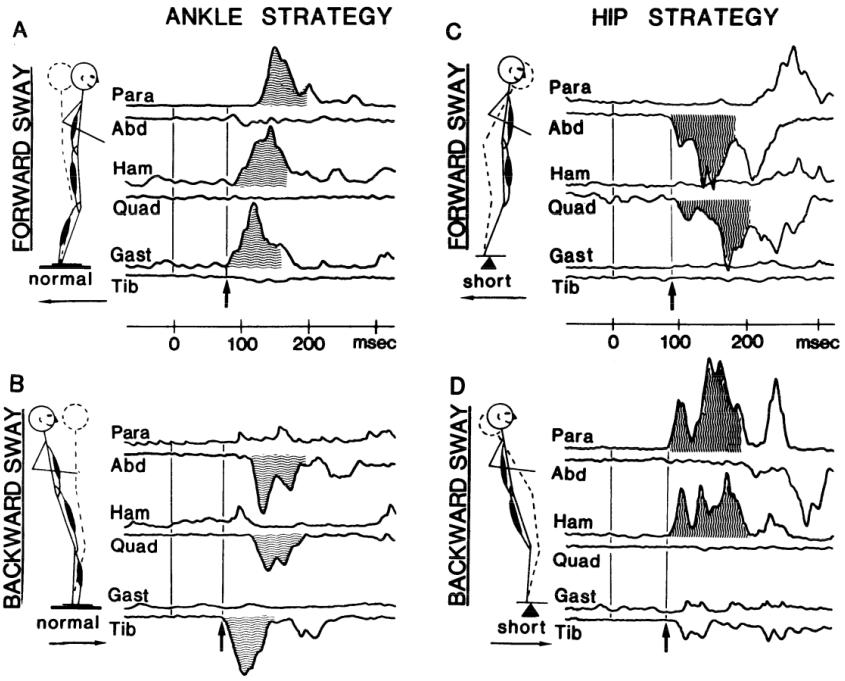


Figure 1.4: Muscle response patterns to forward (top) and backward (bottom) sway perturbations [9]: A and B: on normal surface, ankle strategy; C and D: on short surface, hip strategy. EMGs for 6 muscles are arranged in antagonist pairs with the dorsal muscle EMGs directed up and the ventral EMGs directed down. First 100 *ms* EMG response shaded in this and subsequent figures. Dashed stick figures represent corrected body position. Average of last 5 of 20 trials for subject A triggered 150 *ms* prior to platform perturbation (solid line, 0 *ms*). Arrows indicate latency of first EMG. Abbreviations for this and subsequent figures: Para, lumbar para spinal muscles; Abd, rectus abdominis; Ham, hamstrings; Quad, rectus femoris; Gast, medial gastrocnemius; Tib, tibialis anterior.

1.2 Synergy analysis

The mechanism of human balance recovery has been studied in the literature where various balance control approaches such as static and stepping strategies have been proposed [11], [13]. Understanding balance strategy is crucial for the elderly population due to a decrease in standing ability (motor impairment) after losing muscle strength.

Many works related to this research topic have been conducted regarding various aspects. Hip-ankle strategy for balance maintenance was simulated with a double inverted pendulum based on numerical model predictive control [3]. An improved balance strategy with arm usage based on nonlinear model predictive control and its effectiveness was studied in [53]. Besides, another multi-link model including ankle, hip, and neck was proposed in [54] to analyze standing postural control and its motor coordination. However, the aforementioned works were more focused on the human balance control strategy. Indeed, the relation between balance and coordination is very important for understanding balance control mechanism [55]. Then, the study in [56] indicates the ability of balance control can be improved with strong intended coordination over multiple joints.

Synergy was defined as an effective tool to explain the co-work of muscles and joint coordination [57]. A reasonable assumption is that the movement control is simplified by the central nervous system (CNS) through a layered and modular structure. At the lowest level of its hierarchy, muscle employment may be controlled by several functional units, thereby reducing the dimensionality of the output space. Higher levels in the hierarchy may employ and combine these output modules flexibly to control various movements. Consequently, organizing muscle synergy is one way through which the task of controlling a number of degrees of freedom (DoF) is simplified by the CNS. Indeed, two types of synergies were introduced in [58] to describe different structures of muscle patterns that can be shared in motions: (*i*) muscle-related spatial synergy, (*ii*) muscle, and time-related spatiotemporal synergy. In our study, since we only focus on joint coordination across the simplified human model, spatial synergy is employed.

Besides, patterns of coordinated multi-joint motion have been studied in various tasks such as reaching, walking, sit to stand and standing. Coordinating multiple joints in [59] for producing a straight hand path saves a definite computational resource through the multi-

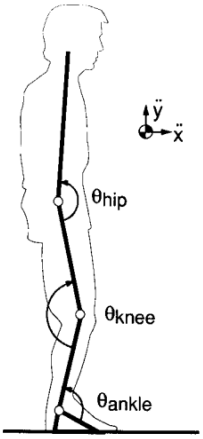


Figure 1.5: Four-segment sagittal plane model of the human body in [11].

joint coordination. The motor synergy of walking motion generated with deep learning has been quantified in Chai's work [60], in which it was concluded that a good synergy level contributes for improving energy efficiency. In sit to stand studies, two aspects have been studied. Yang *et al.* [61] hava distinguished different standing-up strategies based on recruitment of muscle group and Yamasaki *et al.* [62] studied joint torques organization and their correlation. The standing coordination has been studied in [54] with a multi-link model including neck, hip, and ankle. However, this work ignored the joint rotation of arms and knees which are definitely useful in our daily life for improving balance ability of quiet standing while applying a pushing force on the back of a human. The main contributions of this paper are summarized as follows.

- An eight-joint model for human balance recovery is proposed and four-level of pushing forces are considered to distinguish different balance strategies.
- Spatial motor synergy is extracted to classify the organization of joints.
- The transition of balance strategies is discussed according to the space of the extracted synergies.

1.3 Related Work

The development of balance strategies has been conducted from numerical simulations and human experiments. These are without a doubt a huge leap towards understanding

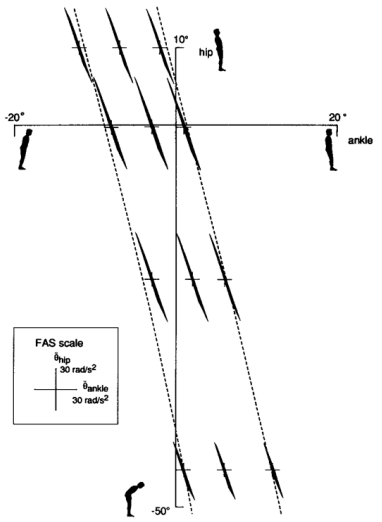


Figure 1.6: Ankle-hip FAS plotted for 12 separate body positions in [11].

the balance recovery movements, however, the relation of limbs and brain for balance recovery is still not clear for previous research. Some of them focus on the switch of body joints, or optimal motion selection. The whole body cooperation might be necessary for correct balance mechanism understanding based on our best knowledge. In this section, I will introduce some of the representative balance strategy researches both in human experiments and numerical simulations and their applications.

1.3.1 Central programming of postural movements: adaptation to altered support-surface configurations

The ankle strategy and the hip strategy were first established from experimental studies in humans. In this section, one of the classical studies is presented in [9]. Two different perturbations are applied in order to observe the configuration of the postural movements in relation to the properties of the support surface. These are the forward and backward sway shown in Figure 1.4. The activation patterns of the ankle, thigh, and trunk muscles were stereotyped in all subjects after repeated exposure to the perturbations on a normal support surface. Contraction began in the muscles of the ankle joint that opposed the swaying perturbation and then radiated to the thighs, then trunk muscles in a distal to proximal order on the same dorsal or ventral aspect of the body.

Figure 1.4, A and B show the ensemble mean responses elicited by perturbing the subject

in the forward and backward directions, respectively. EMG traces are placed in segment groups where the functional tensor response is upward and the flexor response is downward. After a sufficient number of exercises on a short 9.0 *cm* surface, subjects used different but equally stereotyped muscle activation patterns to maintain stability (Figure 1.4, C and D). The perturbations activated thigh and trunk muscles that were antagonistic to those activated by the normal support surface. The abdominal and quadriceps muscles responded to the perturbation of the forward swing. The hamstrings and paraspinal muscles responded to backward rocking perturbations.

These results confirm the fact that the choice between the two different postural movement strategies is influenced not only by current feedback information, but also by previous experience. Specifically, when the change in support configuration is triggered by an external event, the change in the combination of ankle and hip strategies is incremental rather than instantaneous. An important caveat regarding this study is the application of EMS response patterns to distinguish between ankle and hip strategies of perturbation from the support surface.

1.3.2 Human standing posture: multi−joint movement strategies based on bio-mechanical constraints

The ankle strategy and hip strategy can also be distinguished from the joint movements except muscle patterns mentioned in section 1.3.1. A four-segment sagittal plane model of the human body was built in [11] (see Figure 1.5). In joint coordinates, angles are defined to be positive in extension, with zero degrees corresponding to the full upright anatomical position. Center of mass movements in x and y directions are defined to be positive in the anterior and upward directions, respectively.

To evaluate the transition between the ankle and hip strategies, a set of all feasible accelerations that the muscles could induce in a nearly upright position (i.e., the ”feasible acceleration set” or FAS) was calculated. The ankle-hip FAS plotted for 12 separate body positions are shown in Figure 1.6. The origin indicates that the body position in each FAS plot is indicated by a cross-hair. The body position is indicated by the angle of the ankle and the angle of the hip as defined in Figure 1.5. The dashed lines indicate the

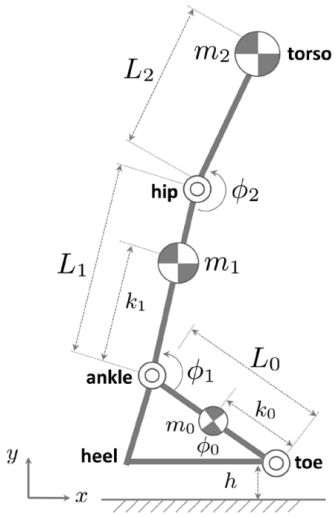


Figure 1.7: Three link sagittal biped is composed of three rigid links in [27].

position boundaries beyond which stable balance cannot be achieved. The line on the left is the boundary for falling forward, and the line on the right is the boundary for falling backwards. Note that the FAS is similar in size and orientation throughout the body's position space. The inset shows the scale of the axes in the FAS plot. At the extremes of the stable flat-footed posture (the positions near the boundary shown in Figure 1.6), use the ankle strategy (i.e., accelerate along the ankle axis toward the upright posture) to bring the body back upright. The position easily violates the constraints of flat feet. The only recourse is to use the hip strategy to move the body away from the boundary before using the ankle strategy to restore the nominal position. The contribution of this study is that FAS was proposed to analyze insights into the control of standing multi-joint upright postures based on biomechanical constraints.

1.3.3 A Two-Joint Human Posture Control Model With Realistic Neural Delays

In this section, A simplified human model in [27] represented with a double inverted pendulum (see Figure 1.7) is proposed for the numerical simulation study of posture control. a nonlinear optimal control model for posture regulation illustrated in Figure 1.8 was developed in this study. The proposed model consists of two main components: body dynamics and performance measure.

Body dynamics are the dynamics of a double inverted pendulum in the sagittal plane, controlled by the ankle and hip torques. Performance measures are nonlinear quadratic in the centre of pressure and quadratic in the control. Realistic values of both sensory and motor delays are included in the dynamic model. The problem of this nonlinear fourth-order regulator is largely solved by the model predictive control method. The resulting feedback control replicates both the experimentally observed perturbations and the tuned nonlinear response. The responses to different amplitudes of the perturbations are qualitatively and quantitatively different. Small perturbations produce mainly ankle motion, while large perturbations produce more hip motion than ankle motion. This can be compared in Figures 1.9 and 1.10.

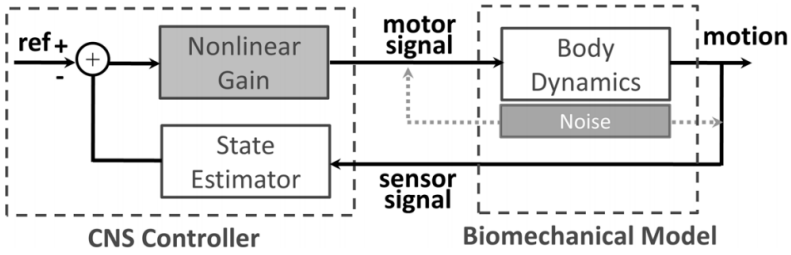


Figure 1.8: System diagram of balance control model with delays and noise/perturbations shown. State estimator is assumed to be well approximated by a Kalman filter/predictor in [27].

1.3.4 Multiple balance strategies from one optimization criterion

Multiple strategies for standing balance [23] have been observed in humans, including using the ankles to apply torque to the ground, using the hips and/or arms to generate horizontal ground forces, and using the knees and hips to squat. This study shows that multiple strategies can arise from the same optimization criterion. It is likely that humanoid robots will exhibit the same balance strategies as humans.

Figure 1.11 shows a one-link inverted pendulum responding to perturbations in the sagittal plane (forward and backward motion only). In this simple model, all joints except the ankle are held in fixed positions. This model of standing balance has only one strategy: applying torque to the ankles. The height of one link model is 2 $[m]$ and the total weight is 70 $[kg]$. The one-step optimization criterion is a combination of the deviation of the joint angle θ_a from zero (directly above) and a quadratic penalty on the joint torque

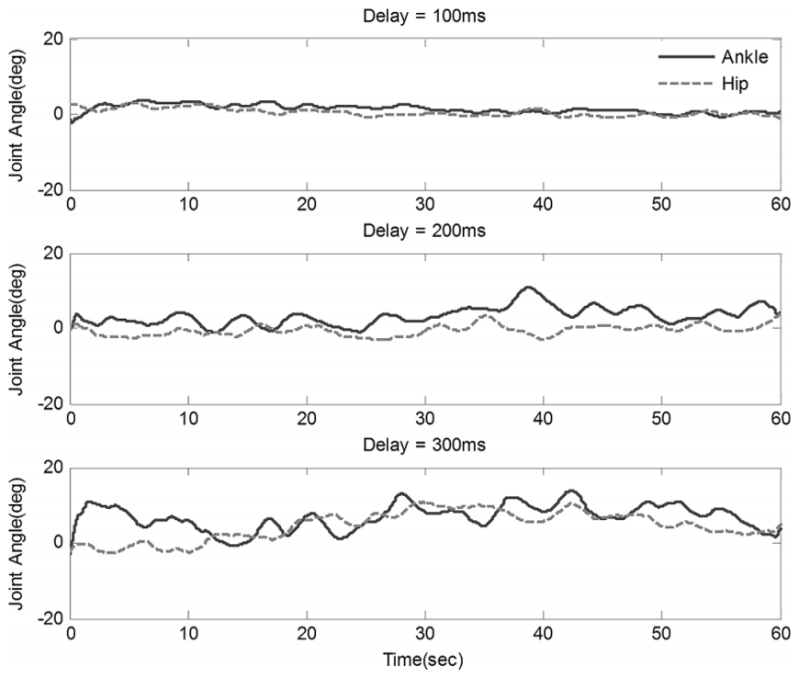


Figure 1.9: Joint angle trajectories of ankle joint (solid line) and hip joint (dotted line) with different time delays under fixed perturbation $\|\varepsilon\| = 0.01$ in [27].

$L(\mathbf{x},\mathbf{u}) = \mathrm{T} * \theta_a^2 + 0.002 * \mathrm{T} * \tau_a^2$ where 0.002 weights the torque penalty for position error. T is the simulation time step (0.01 $[s]$). There is no cost associated with joint velocity. The total orbit cost is obtained by summing the criteria for one step at each step of the orbit. The optimal trajectory minimizes the total trajectory cost.

Figure 1.12 shows a two link inverted pendulum responding to a perturbation in the sagittal plane. This multi-link pendulum is a useful model for standing balance of a human and a humanoid robot. The bottom link is the leg, and the top link is the torso. Each link is 1 meter long, and is modeled as a thin rod of 35 $[kg]$ for a total body weight of 70 $[kg]$. The one step optimization criterion is a straightforward extension of the one link optimization criterion: a combination of quadratic penalties on the deviations of the joint angles from zero (straight up) and on the joint torques: $L(\mathbf{x},\mathbf{u}) = \mathrm{T} * \left(\theta_a^2 + \theta_h^2\right) + 0.002 * \mathrm{T} * \left(\tau_a^2 + \tau_h^2\right)$ where 0.002 weights the torque penalty relative to the position error. And T is the time step of the simulation (0.01 $[s]$). There are no costs associated with joint velocities.

To explore more complex and human-like strategies, we subdivided the two link models into four link models, including knees, shoulders, and arms (Figure 1.13). Each link is modeled as a thin rod, with calf and thigh lengths of 0.5 $[m]$ and 17.5 $[kg]$, respectively.

The torso is 1m long and weighs 26.25 $[kg]$, and the arms are 1m long and weigh 8.75 $[kg]$. The symmetrical legs of 0.2 meter remained the same. Impulse perturbations were applied horizontally to the center of the torso. The one-step optimization criterion is a combination of quadratic penalties for joint angle deviation from zero (straight up), joint velocity, and joint torque: $L(\mathbf{x},\mathbf{u}) = \mathrm{T} * \left(\theta_a^2 + \theta_h^2\right) + 0.002 * \mathrm{T} * \left(\tau_a^2 + \tau_h^2\right)$ where 0.002 weights the torque penalty for errors in position and velocity. Also, T is the simulation time step $(0.01[s])$. For the four-link case, we introduced a penalty on the joint velocity to reduce knee and shoulder oscillations. Figures 1.13 and 1.14 show the response to the largest perturbation that can be handled in each direction.

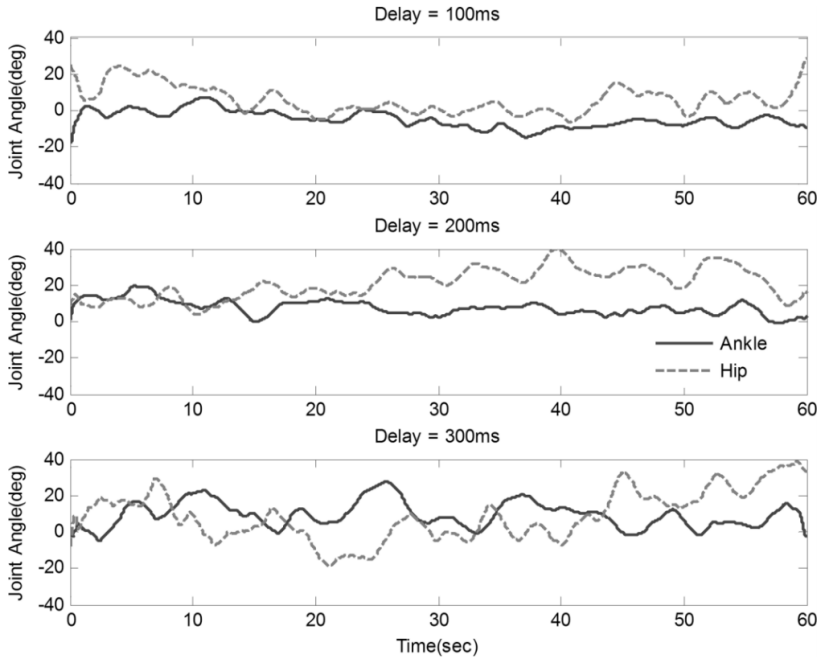


Figure 1.10: Joint angle trajectories of ankle joint (solid line) and hip joint (dotted line) with different time delays under fixed perturbation $\|\varepsilon\| = 0.1$ in [27].

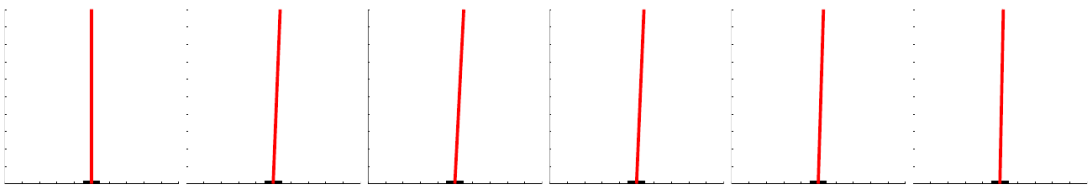


Figure 1.11: Configurations every quarter of a second of a simulated one link inverted pendulum response to an impulse of 15 Newton-seconds forward (to the right). The black rectangle indicates the extent of the symmetric foot in [23].

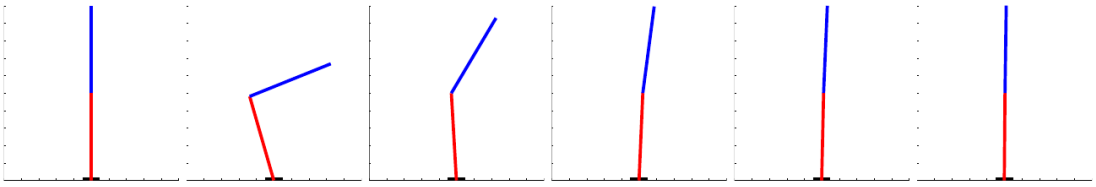


Figure 1.12: Configurations every half second of a simulated two link inverted pendulum response to an impulse of 25 Newton-seconds forward (to the right). The black rectangle indicates the extent of the symmetric foot [23].

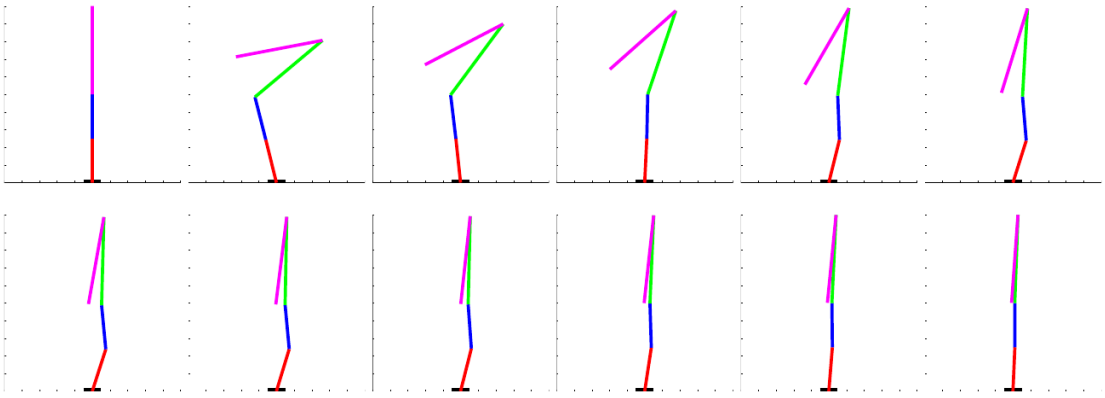


Figure 1.13: Configurations every quarter second from a simulated four link inverted pendulum response to a forward impulse (to the right) of 22.5 Newton-seconds. The black rectangle indicates the extent of the symmetric foot [23].

1.3.5 Balance Recovery Prediction with Multiple Strategies for Standing Humans

Another interesting related work is on balance control of flywheel inverted pendulum by model predictive control [32]. The human body is represented using the much-vaunted inverted pendulum model, in which the entire mass of the system is considered to be concentrated at the center of gravity (COM). The model is further supplemented with feet to represent the base of support, and flywheel segments are included to illustrate the inertial effects of rotation of the segments shown in Figure 1.15. Three independent recovery actions are included in this model. They correspond to the mechanical effects of the three classical recovery strategies described in the literature (ankle, hip, and recovery step).

This model is placed in a closed loop using a model prediction controller (Figure 1.16). This type of controller iteratively solves a series of optimization problems and computes an optimal control strategy based on a simplified model of the system. In this work, a

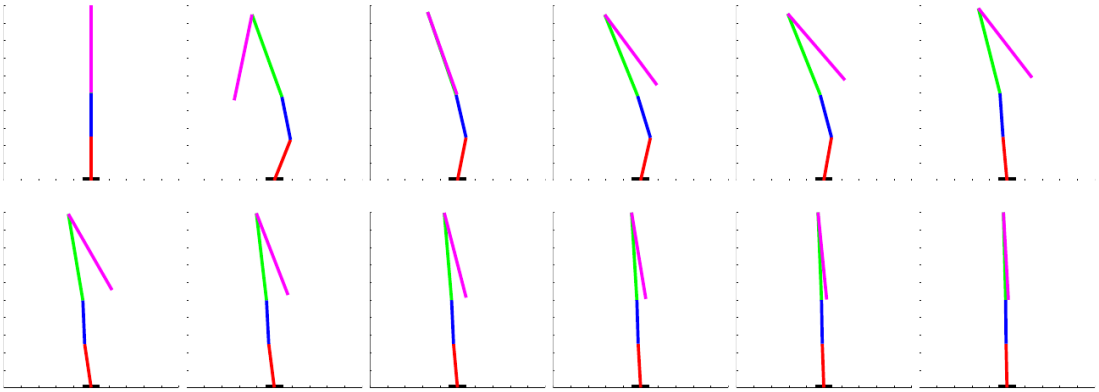


Figure 1.14: Configurations every quarter second from a simulated four link inverted pendulum response to a backward impulse (to the left) of 22.5 Newton-seconds [23].

simplified representation of the mechanical system, called the internal model, is used to estimate the consequences of the control strategy for future system states. Each time the controller is called, an optimal control strategy is selected, i.e., the control action (position of the center of pressure (CoP), position of the foot, and rotation of the inertia wheel) at each time step in the time range. Put the internal model into the desired state. Since the internal model is different from the mechanical (or real) system, applying the control strategy as a whole to the mechanical system may result in divergent motions from the expected trajectory. Therefore, only the initiation of this control (the first step) is applied to the mechanical system, and the strategy is updated via a new call to the controller using this latest system state.

The optimal control strategy is chosen primarily by minimizing a cost function defined to quickly return the mechanical system to a steady state attitude. This consists of minimizing the horizontal CoM velocity and angular velocity of the flywheel segment over the entire prediction range. In addition, a cost associated with the leg swing is introduced by minimizing the swing foot acceleration. Three additional terms are also included to smooth the motion trajectory and converge the CoP under the ankle position at the end of the recovery. These latter are less penalizing and do not affect the overall recovery motion as much. A limitation of this study, however, is that it omits upper body movements where human-like balance recovery movements can be observed.

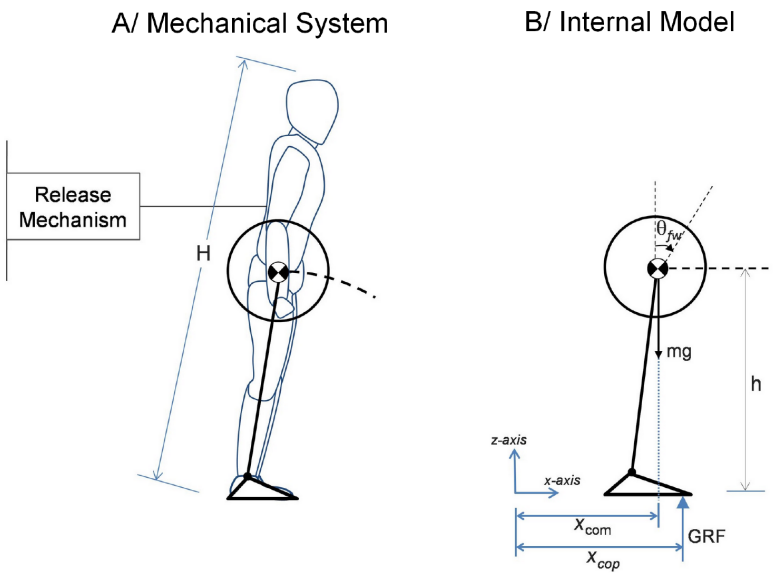


Figure 1.15: The two representations used of the human body. Left: Mechanical system: simple inverted pendulum plus flywheel model, i.e. the CoM follows a circular arc. Right: Internal model: linearized inverted pendulum plus flywheel model, i.e. the CoM travels at a constant height h [32].

1.3.6 Wearable Devices for Balance Rehabilitation

Wearable devices such as the exoskeleton [63] and other wearable devices [64] have been developed at a fast pace over the last few decades to improve the walking and standing abilities of patients after stroke or spinal cord injury. However, suitable balance control strategies for wearable devices remain challenging because wearable devices need to support human balance with the same movement patterns. In this section, we describe some of the latest wearable devices for rehabilitation and point out some of their limitations in [65]. The reasons for studying the movement and control mechanisms of human balance will be very well understood after an overview of their applications. This may lead to the development of balance control methods for wearable devices.

Many commercially available wearable devices, such as the typical device shown in Figure 1.17, have only knee and hip flexion-extension actuation combined with passive or fixed ankle motors [66]. While hip and knee joint actuation can generate a gait pattern, it is difficult to generate human-like motion and control self-balance. Plantar/dorsiflexion actuation of the ankle is necessary to stabilize the body in the sagittal plane by reducing the effects of gravitational torque in the standing position. In addition, uncontrolled hip abduction/internal rotation does not allow the body to step out to the side. This is necessary

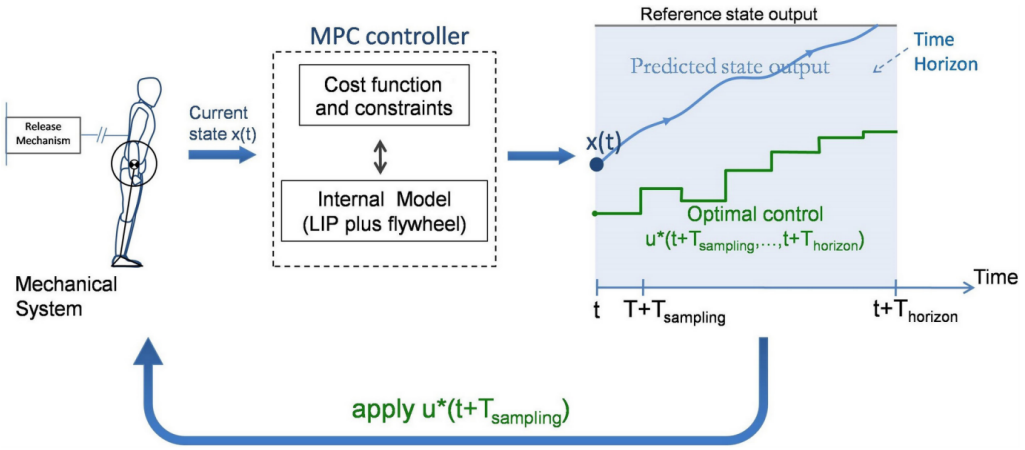


Figure 1.16: The basic principle of Model Predictive Control. At any time instant t the most recent mechanical system state $x(t)$ is sampled and fed to the controller. The control is computed minimizing the cost function 2 to bring the internal model to a standstill posture. The computed control vector u, consisting of the CoM and flywheel jerks and the future step positions, is then applied back to the mechanical system [32].

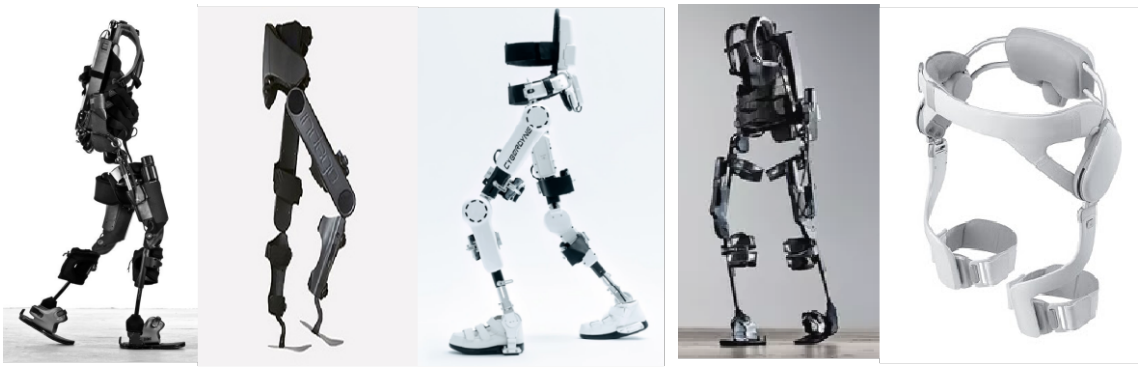


Figure 1.17: Commercial wearable devices.

to maintain balance in the lateral plane. Note that many wearable devices are not equipped with these degrees of operating freedom. Therefore, these limitations are due to their mechanical design. Also, most control strategies of wearable devices are set up with predefined motor trajectories. This indicates that these control strategies lack the adaptive capabilities needed to maintain balance in different situations.

An efficient approach to develop a suitable balance control strategy for wearable standing support devices is to consider the human balance mechanism. First, humans have a better ability to maintain balance than wearable devices. Second, a wearable device may be more predictable and more comfortable if the control strategy is consistent with the actions that the user might naturally perform. Thus, human-like control strategies may improve the coordination between the user and the wearable device and increase user acceptance of the

device. The wearable device should be adaptively controlled to produce human-like quiet standing balance movements through ankle and hip strategies.

1.3.7 Muscle Synergy Analysis of Balance Postural Responses

Postural balance control is a natural behavior that requires the spatial and temporal coordination of multiple muscles. Complex muscle activation patterns characterize the postural response. This illustrates the need for independent muscle control. This work [67] discussed whether general neural strategies could be used to control postural balance in humans. And they thought that the synergy of several muscles could explain the trial-to-trial variability of automatic postural responses from different perturbation directions, and different postural strategies.

Postural balance responses to multi-directional support-surface translations in 16 leg and lower-back muscles of the subject's right side, including: rectus abdominalis (REAB), tensor fascia lata (TFL), biceps femoris long head (BFLH), tibialis anterior (TA), semitendinosus (SEMT), semimembranosus (SEMB), rectus femoris (RFEM), peroneus (PERO), medial gastrocnemius (MGAS), lateral gastrocnemius (LGAS), erector spinae (ERSP), external oblique (EXOB), gluteus medius (GMED), vastus lateralis (VLAT), vastus medialis (VMED), and soleus (SOL), were analyzed in 9 healthy subjects. Six or fewer muscle synergies were recruited to reconstruct the postural balance responses in each subject, illustrated in Figure 1.18.

The composition and temporal activation of several muscle synergies identified in all subjects was consistent with the "ankle" and "lower back" strategies previously identified in the human postural balance response in Figure 1.19. The average peak activation coefficient for the synergistic Synergy W_1 of the "ankle" strategy was substantially greater during APR1 than in the other time bins. The average peak activity of the inner and outer "hip" strategy, Synergy W_5, was highest during the early bins APR 1 and APR 2. For synergy W_{2-3}, the average peak activity was greater in APR 2 and APR 3 than in the other time bins, whereas the average peak activity for synergy W_4, the "hip" strategy synergy. The average peak activity of synergy W_4 was significantly higher during the "hip" strategy, APR 3. Interestingly, the activation of muscle synergy 6, the "knee" strategy synergy, was

the only muscle synergy that had a higher average peak activity at APR2 compared to all other time bins. Furthermore, the trial-to-trial variation in muscle activation patterns was successfully replicated by adjusting the synergy of the various muscles. This indicates that the changes in the activation of each muscle are related and represent changes in the amplitude of the descending nerve commands that activate each muscle synergy. Thus, they concluded that muscle synergy may represent a general neural strategy for muscle coordination in postural tasks.

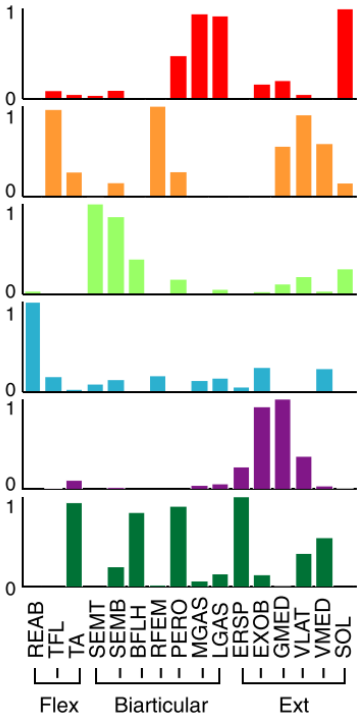


Figure 1.18: Muscle synergy vectors for a representative subject in [67]. Muscle synergy vectors, W_i, extracted from EMG data during quiet stance and 3 APR time bins. Each bar represents the relative level of activation of each muscle within the synergy.

1.4 Research Objective

The purpose of this research is to understand the mechanism of human balance recovery control through reproducing human-like balance strategies from static and dynamic balance motions and analyzing the motion features according to postural coordination modes. The further descriptions of research objective are as follows.

- Reproduce human-like hip-ankle strategy for balance recovery with double inverted pendulum through numerical model predictive control;

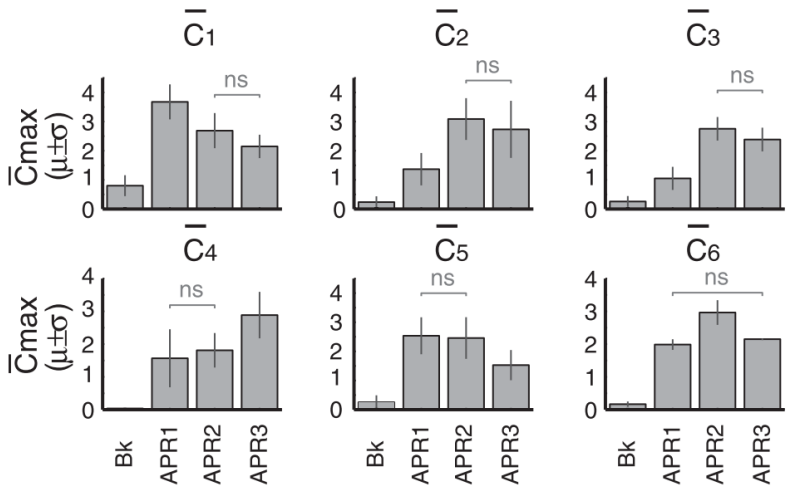


Figure 1.19: Mean peak activation of muscle synergies [67] across subjects in all 4 time bins. Each muscle synergy activation level is statistically different in magnitude (P is less than 0.05) from the others in the sub-panel, except where specified by the designation *ns*. Mean peak activation of muscle synergy W_1, an "ankle" strategy synergy, is largest in APR1, whereas mean peak activation of muscle synergy W_4, a "hip" strategy synergy is larger in APR3. "Knee" strategy synergy W6 has its peak in APR2

- Reproduce three types of arm strategies through nonlinear model predictive control and evaluate their contributions to balance recovery;

- Balance strategy classification and feature analysis through the extraction of synergies both in human experiments and simulations;

- Reproduce postural coordination modes in head-target tracking tasks through model-free deep reinforcement learning.

1.5 Structure of thesis

This dissertation is organized in six thematic chapters dealing with all aspects relevant to reproduced human balance recovery strategies from static and dynamic balance motions and their mechanism analysis both in simulation and human experiment studies. The current chapter, Chapter 1, presents the introduction to this study. It summarizes the background, types of the simplified human model, control methods and their associated challenges, related works including the state of art in extended wearable device application, and the objectives of this research. The subsequent chapters can be summarized as follows.

1.5.0.1 Chapter 2: Human-like Balance Recovery Based on Numerical Model Predictive Control Strategy

This chapter introduces the simulation software MuJoco and the basic concept of MPC. Then, we describe the proposed MPC to predict the best way to maintain balance for different disturbance forces. We discuss in detail, to simulate balance recovery, a three-link model including afoot with unilateral constraints, lower body, and upper body was built. Then, we derived the dynamical equation and made its linearization. By learning human balance capability, we set bound constraints on our model including angles and balance torques of ankle and hip. Unilateral constraints are set on the foot, which makes our model more like a human quiet standing case. Finally, we reproduced the proposed ankle and hip-ankle strategy in simulation and analyzed the obtained results from kinematic and dynamic indices such as the evolution of joint angles, torques, and the Center of Mass (CoM) and Center of Pressure (CoP) as well as the energy consumption aspect. The robustness of the controller was verified in this study as well. This chapter can be considered the first part of contribution of this dissertation.

1.5.0.2 Chapter 3: Reproducing Human Arm Strategy Through Model Predictive Control and its Contribution to Human Balance Recovery

This chapter describes another balance strategy (arm strategy) and its implementation in our study. First, the softwares and hard-wares used in our research such as Opensim, Optitack, Forceplate are introduced. And, data processing methodologies in OpenSim including musculoskeletal model scaling, inverse dynamics and inverse kinematics in OpenSim are given in details. Then, the A simplified model with arms is built and a controller based on NMPC are proposed to implement human-like balance control. Three-arm states of the model, namely, active arms, passive arms, and fixed arms, were considered to discuss the contributions of arm usage to human balance recovery during quiet standing. Various indices are recruited to investigate the contributions of the arm strategy. This chapter can be considered the second part of contribution of this dissertation.

1.5.0.3 Chapter 4: Postural Synergy Extraction and Analysis

This chapter describes the synergy extraction and analysis of balance motion features of the corresponding ankle strategy, hip strategy and arm strategy discussed in the first and second studies about static postural balance recovery. We discuss in detail, the synergy extraction methodology such as PCA, experimental protocol, data processing, synergy patterns used to analyze motion features, joint coordination both in human experiments and simulations, and possible applications. This chapter can be considered the third part of contribution of this dissertation.

1.5.0.4 Chapter 5: Reproducing Postural Coordination Modes Through Model-free DRL

This chapter described dynamic balance motion named head-target tracking tasks through model-free learning methodology DRL. The kinematic and dynamic constraints are adapted in the model and DRL reward function to explore the random behavior for different tracking frequencies during the training process. postural coordination transition for different tracking frequencies are analyzed with various indices. First, the transition between in-phase and anti-phase is distinguished through the joint correlation of ankle and hip. Then, the temporal joint angles and phase portraits of hip and ankle are compared to classify the motion transition depending on frequencies as well. Finally, the CoM deviations are compared for different frequencies to analyze the motion features. This chapter can be considered the final part of contribution of this dissertation.

1.5.0.5 Chapter 6: Conclusions and Future Work

This chapter concerns the final remarks and summarizes the present study. We talk about the results, outcomes, and the future work of this research.

Chapter 2

Human-like Balance Recovery Based on MPC Strategy

This chapter is about human-like hip-ankle strategy reproduction for balance recovery through numerical model predictive control(N-MPC). After reviewing previous works, we found that they did not cover long-time disturbing forces, the robustness of MPC, the energy consumption of the hip and ankle joints, and the evolution of the ground reaction force and feet with unilateral constraints. The contributions of this paper are as follows.

1) Numerical MPC with system states and control constraints is proposed to implement a human-like balance strategy and autonomous switching between ankle strategy and hip-ankle strategies during quiet standing balance for the different disturbing forces. This model illustrates that N-MPC is similar to the predictive behavior elicited by the central nervous system from a neuroscience viewpoint. N-MPC is also endowed with a predictive aspect that enables it to predict future behavior and select a control balance strategy through the minimization of the energy consumption of the whole body. N-MPC can handle state and input constraints simultaneously. This is crucial to fulfilling realistic requirements because body limitations, including joint ranges and input torques, can be considered.

2) The CoM and the center of pressure (CoP) are used as evaluation indices and constraints

to maintain an upright orientation. Different disturbing forces are considered, namely small and large disturbing forces and short-time and long-time disturbing forces.

3) From the obtained simulation results, we analyzed the mixed hip-ankle strategy regarding two aspects: kinematic and dynamical indices and energy consumption. In addition, we tested the robustness of the proposed controller and verified that our model and control approach can implement a much more human-like balance behavior. Thus, the proposed controller could shed light on human motor control on the ankle and hip may become an efficient guide to understand the elderly's quiet standing balance.

The rest of this chapter is organized as follows. In Section 2.1, the software MuJoCo and the concept of MPC are introduced. In Section 2.2, the three-link model is described and its dynamic equation is derived and linearized. The proposed N-MPC is described particularly in Section 2.3. In Section 2.4, simulation settings and results are presented and discussed. Our conclusions are presented in Section 2.5. The structure of this work is illustrated in Figure 2.1.

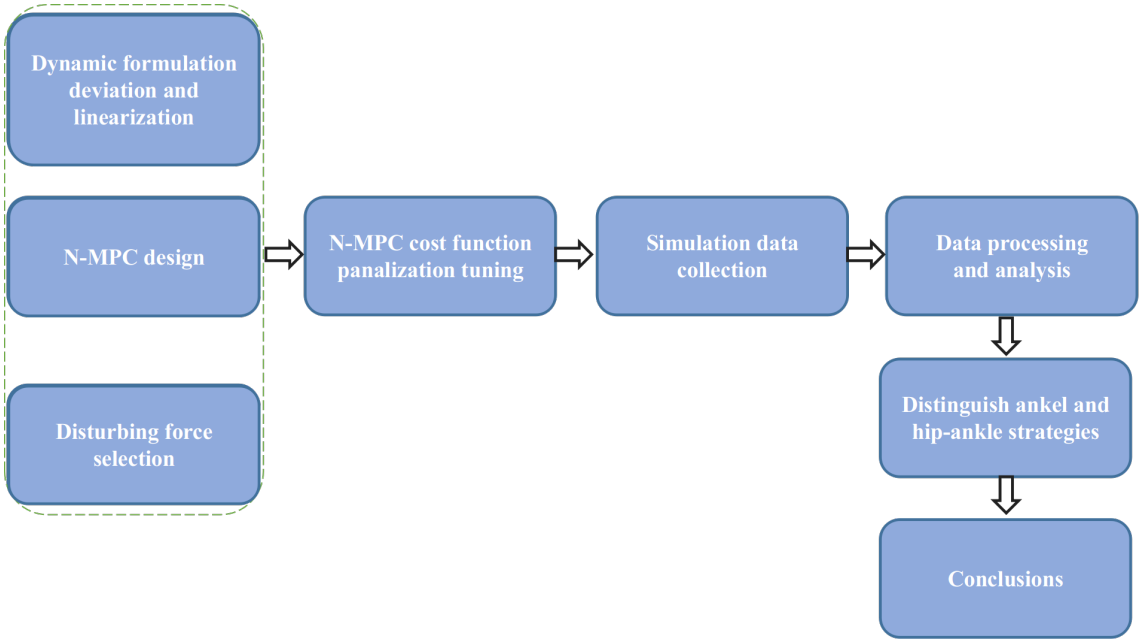


Figure 2.1: Sturcture of this work.

2.1 Methodology Introduction

In this section, the physics simulation software Mujoco and the concept of MPC used in this thesis are introduced as follows.

2.1.1 MuJoCo

MuJoCo [68] stands for Multi-Joint dynamics with Contact. It is a physics engine aiming to facilitate research and development in robotics, biomechanics, graphics and animation, machine learning, and other areas where fast and accurate simulation of complex dynamical systems is needed. MuJoCo is a dynamic library with C API, compatible with Windows, Linux and maxOS. It is intended for researchers and developers with computational background. It includes the XML parser, model compiler, simulator and interactive OpenGL visualizer. It further exposes a large number of functions for computing physics-related quantities, not necessarily in a simulation loop. MuJoCo can be used to implement model-based computations such as control synthesis, state estimation, system identification, mechanism design, data analysis through inverse dynamics, parallel sampling for machine learning applications. It can also be used as a more traditional simulator, including applications to gaming and interactive virtual environments. The MuJoCo environment is illustrated in Figure 2.2

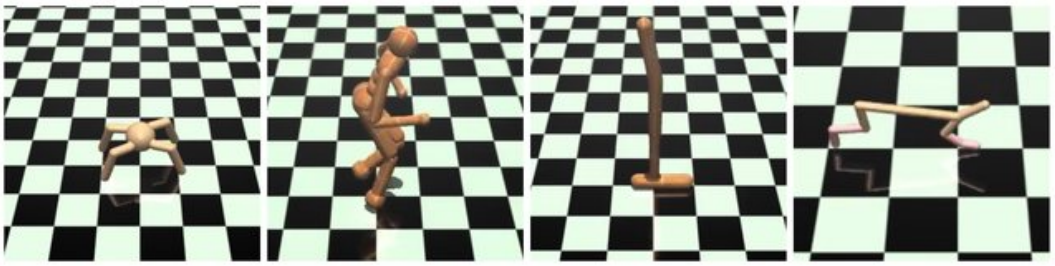


Figure 2.2: Illustration of the MuJoCo environment [69]. From left to right: Ant, Humanoid, Hopper, HalfCheetah, respectively.

The notable features of MuJoCo are outlined briefly as follows.

- Generalized coordinates combined with modern contact dynamics.
- Soft, convex and analytically-invertible contact dynamics.

- Tendon geometry.
- General actuation model.
- Re-configurable computation pipeline.
- Model compilation.
- Separation of model and data.
- Interactive simulation and visualization.
- Powerful yet intuitive modeling language.

In our study, we take advantage of MuJoCo to simulate the human-like balance recovery movements through optimal control algorithms such as Model Predictive Control and Deep Reinforcement Learning. In this chapter, the next chapter and chapter 5, we will show the simulation results and scenarios in detail.

2.1.2 Model Predictive Control

In this section, we will introduce model predictive control (MPC) based on [70] from the following four aspects briefly:

- The basic idea of MPC.
- The Algorithm of MPC.
- The advantages and drawbacks of MPC.
- The implementation tool.

2.1.2.1 Basic Idea of MPC

The idea of MPC for linear or nonlinear systems is to utilize a model of the process in order to predict and optimize future system behavior. Here, the formulation of models is described

$$x^{+} = f(x,u) \tag{2.1.1}$$

where $f : X \times U \to X$ is a nonlinear map which derives the next state x^+ through a current state x and a current control value u. Starting from the current state $x(n)$, for any given control sequence $u(0),...,u(N-1)$ with horizon length $N \geq 2$, Eq. (2.1.1) can be iterated to construct a prediction trajectory x_u, which is defined by

$$x_u(0) = x(n), \quad x_u(k+1) = f(x_u(k), u(k)), \quad k = 0,\ldots,N-1 \tag{2.1.2}$$

predictions $x_u(k)$ for the state of the system $x(n+k)$ at time t_{n+k} in the future can be obtained through this way. Thus, a prediction of the behavior of the system on the discrete interval $t_n,...,t_{n+N}$ depends on the chosen control sequence $u(0),...,u(N-1)$.

Then, Optimal control is used to determine $u(0),...,u(N-1)$ such that the next state x_u is as close as possible to the reference $x_* = 0$. After that, the distance between $x_u(k)$ and $x_* = 0$ for $k = 0,...,N-1$ is measured through the function $\ell(x_u(k), u(k))$. Here, the deviation of state from the reference and the distance of the control values $u(k)$ to a reference control $u_* = 0$ are penalized to reach the purpose. A common representation of the function $\ell(x_u(k), u(k))$ based on the quadratic function

$$\ell(x_u(k), u(k)) = \|x_u(k)\|^2 + \lambda \|u(k)\|^2, \tag{2.1.3}$$

where $\|\cdot\|$ denotes the usual Euclidean norm and $\lambda \geq 0$ is a weighting parameter to penalize the control.

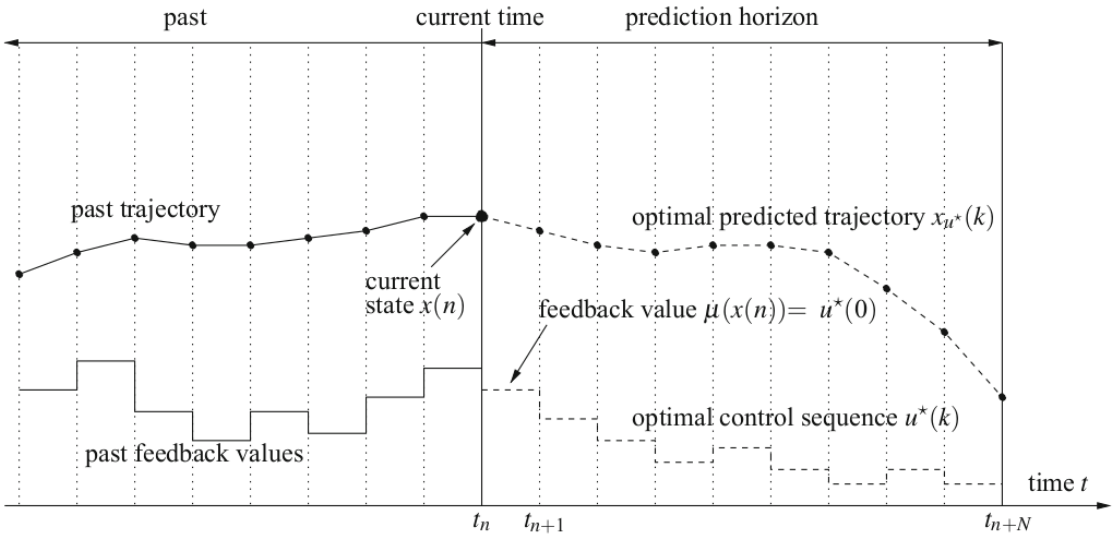


Figure 2.3: Illustration of the MPC step at time t_n [70].

Then, the optimal control problem reads

$$\text{minimize} \quad J(x(n),u(\cdot)) := \sum_{k=0}^{N-1} \ell\left(x_u(k),u(k)\right) \tag{2.1.4}$$

w.r.t all admissible control sequences $u(0),...,u(N-1)$ with x_u generated by Eq. (2.1.2).

Finally, the solution of optimal control problem can be obtained through minimizing control sequence $u^*(0),...,u^*(N-1)$,

$$\min_{u(0),...,u(N-1)} J(x(n),u(\cdot)) = \sum_{k=0}^{N-1} \ell\left(x_{u^*}(k),u^\star(k)\right) \tag{2.1.5}$$

Finally, the solution of optimal control problem can be obtained by minimizing the control sequence.

2.1.2.2 Algorithm of MPC

The basic algorithm of MPC for the constant reference $x^{ref} = x_*$ is described as follows. At each sampling time $t_n,\ n = 0,1,2,...$:

1) Measure the state $x(n) \in X$ of the system.

2) Set the initial state $x_0 := x(n)$, solve the optimal control problem:

$$\begin{aligned} &\text{minimize} \quad J_N\left(x_0,u(\cdot)\right) := \sum_{k=0}^{N-1} \ell\left(x_u\left(k,x_0\right),u(k)\right) \\ &\text{with respect to } u(\cdot) \in \mathbb{U}^N\left(x_0\right), \text{ subject to} \\ &x_u\left(0,x_0\right) = x_0, \quad x_u\left(k+1,x_0\right) = f\left(x_u\left(k,x_0\right),u(k)\right), \end{aligned} \tag{2.1.6}$$

and denote the obtained optimal control sequence by $u^\star(\cdot) \in \mathbb{U}^N\left(x_0\right)$.

3) Use the feedback control value $\mu_N(x(n)) := u^*(0) \in U$ in the next sampling period.

4) Repeat the above process from 1).

2.1.2.3 Advantages and drawbacks of MPC

MPC shows many advantages compared with other control methods:

- It is particularly attractive to staff with only a limited knowledge of control because

the concepts are very intuitive and at the same time the tuning is relatively easy.

- It can be used to control a great variety of processes, from those with relatively simple dynamics to more complex ones, including systems with long delay times or non-minimum phase or unstable ones.
- The multi-variable case can easily be dealt with.
- It intrinsically has compensation for dead times.
- It introduces feed forward control in a natural way to compensate for measurable disturbances.
- The resulting controller is an easy-to-implement control law.
- Its extension to the treatment of constraints is conceptually simple, and these can be systematically included during the design process.
- It is very useful when future references (robotics or batch processes) are known.
- It is a totally open methodology based on certain basic principles which allows for future extensions.

However, There are still some drawbacks in MPC:

- Although the resulting control law is easy to implement and requires little computation, its derivation is more complex than that of the classical PID controllers. If the process dynamic does not change, the derivation of the controller can be done beforehand, but in the adaptive control case all the computation has to be carried out at every sampling time.
- When constraints are considered, the amount of computation required is even higher.
- The greatest drawback is the need for an appropriate model of the process to be available. The design algorithm is based on prior knowledge of the model and is independent of it, but it is obvious that the benefits obtained will be affected by the discrepancies existing between the real process and the model used.

2.1.2.4 Implementation tool of MPC

MPC Algorithm in our study is implemented with MPCTools. A brief introduction is given here. MPCTools is a control and estimation tool for linear and nonlinear dynamic models. Model predictive control hierarchy consists of three components: the estimator, the target calculator, and the regulator. MPCTools provides an interface to CasADi solvers for each of these components to facilitate simulation of MPC on any system of interest. The regulator determines an optimal open-loop control trajectory from the current state estimate (from the state estimator) to the targets (from the target calculation). The state estimator reconstructs the most probable state from input and output data from the plant. The target calculation finds a state and input target to yield the desired output steady states of the model, given integrated disturbance estimates and setpoint values.

In the regulator, we use model predictive control (MPC), which uses a mathematical model to forecast and optimize the predicted future behavior of the plant. Stage costs, terminal cost, and hard/soft constraints on states and inputs can all be specified by the user.

The estimator uses moving horizon estimation (MHE). In this approach, the model is used to best reconstruct the state trajectories by solving a nonlinear optimization problem. Stage cost, arrival cost, and system variable bounds can be specified by the user.

The computational algorithm options for both the state estimator and regulator are described in CasADi Documentation. MPCTools allows solver selection from the options in CasADi for each part of the MPC problem.

The target calculation problem uses CasADi solvers as well. It can have an objective function specified by the user as well as constraints.

2.2 Dynamic equation of the three-link model

To implement quiet standing balance recovery, we consider the human body as a three-link simplified model comprising a unilaterally constrained foot, an ankle joint, a lower body, a hip joint, and an upper body, as illustrated in Figure 4.1. The physical parameters of our model are summarized in Tables 2.1. Based on an existing anthropometric database

[71], the total body height is 1.6 $[m]$ and the total body mass is 66.3 $[kg]$. m_0, m_1, and m_2 represent the masses of the foot, the lower body, and the upper body respectively. L_0, L_1, and L_2 represent the lengths of the foot, the lower body, and the upper body respectively. q_0, q_1, and q_2 represent toe, ankle, and hip angles, respectively. It is worth noting that we ignore the body segments between the ankle joint and the hip joint, and between the hip joint and the head. This is consistent with the case of human quiet standing balance because humans maintain their knee joint angle within a certain range of disturbing forces acting on their body. However, if these disturbing forces become too large, they need to bend their knee and step forward to avoid falling.

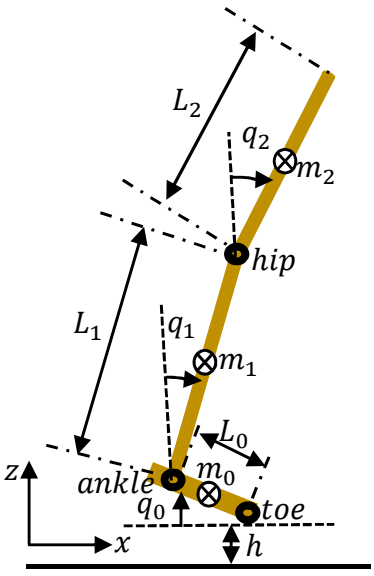


Figure 2.4: Illustration of the three-link, two-joint model for balance recovery strategy. m_0, m_1, m_2 represent the masses of foot, lower body and upper body respectively. L_0 represents the length from the toe to the ankle. L_1, L_2 represent the lengths of lower body and upper body respectively. q_0, q_1, q_2 represent the toe angle, the ankle angle and hip angle respectively.

First, we use Lagrange formalism [72]–[74] to derive the dynamic equation of motion for this two-joint, three-link model controlled by the ankle and hip torques. The Lagrange equations are as follows:

Table 2.1: Summary of the physical parameters of the three-link, two-joint model inspired by human.

Link	Mass [kg]	Length [m]
Foot	1.3	0.3
Lower body	30	1.0
Upper body	35	0.6

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) - \frac{\partial L}{\partial q_1} = \tau_a, \tag{2.2.1}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_2}\right) - \frac{\partial L}{\partial q_2} = \tau_h, \tag{2.2.2}$$

$$L = T - V. \tag{2.2.3}$$

where T is the total kinetic energy, V is the total potential energy, τ_a is the ankle torque and τ_h is the hip torque. In this mechanical system, we assume that the toe rotates with q_0 and h is the vertical displacement of the toe with respect to its original point on the ground. Then, we set unilateral constraints on the foot to make the toe and the heel of the foot remain in contact with the ground. There are two degrees of freedom (DOFs) for the foot constraints: the toe angle q_0 and the displacement of the toe with respect to the ground h. The constraint forces can be solved via the Lagrange method for q_0 and h.

Let us now express the complete dynamic equation of motion. First, the kinetic energies of the foot T_0, of the lower body T_1, and of the upper body T_2 are computed separately as follows

$$T_0 = \frac{1}{2} m_0 \left(\frac{1}{3} {L_0}^2 \dot{q}_0^2 + L_0 \dot{q}_0 \dot{h} + \dot{h}^2\right), \tag{2.2.4}$$

$$T_1 = \frac{1}{2}m_1[{L_0}^2\dot{q}_0^2 + \frac{1}{3}{L_1}^2\dot{q}_1^2 + \dot{h}^2 + L_0L_1\dot{q}_0\dot{q}_1 \sin(q_0 - q_1) + 2\dot{h}L_0\dot{q}_0\cos q_0 - \dot{h}L_1\dot{q}_1\sin q_1], \tag{2.2.5}$$

$$T_2 = \frac{1}{2}m_2[{L_0}^2\dot{q}_0^2 + {L_1}^2\dot{q}_1^2 + \frac{1}{3}{L_2}^2\dot{q}_2^2 + \dot{h}^2 + L_0L_2\dot{q}_0\dot{q}_2 \sin(q_0 - q_2) + L_1L_2\dot{q}_1\dot{q}_2\cos(q_2 - q_1) - \dot{h}L_2\dot{q}_2 \sin q_2 + 2\dot{h}L_0\dot{q}_0\cos q_0 - 2\dot{h}L_1\dot{q}_1\sin q_1]. \tag{2.2.6}$$

The total kinetic energy of the whole body is given by:

$$T = T_0 + T_1 + T_2. \tag{2.2.7}$$

Second, the potential energies of the foot V_0, the lower body V_1 and the upper body V_2 are derived as follows,

$$V_0 = m_0g(\frac{1}{2}L_0\sin q_0 + h), \tag{2.2.8}$$

$$V_1 = m_1g(\frac{1}{2}L_1\cos q_1 + L_0\sin q_0 + h), \tag{2.2.9}$$

$$V_2 = m_2g(\frac{1}{2}L_2\cos q_2 + L_1\cos q_1 + L_0\sin q_0 + h). \tag{2.2.10}$$

The total potential energy of the whole body is given by:

$$V = V_0 + V_1 + V_2. \tag{2.2.11}$$

Here, g represents the gravity coefficient. Because our balance implementation is for quiet standing balance, our model foot cannot leave the ground, which means that stepping and rotation of the foot are not allowed. We set the angle, angular velocity, and acceleration of the toe equal to zero, i.e., $q_0 = \dot{q}_0 = \ddot{q}_0 = 0$. The vertical displacement, velocity, and acceleration of the toe are also set to zero, i.e., $h = \dot{h} = \ddot{h} = 0$. These unilateral constraints are required to meet with the condition of quiet standing. Based on the above equations of unilateral constraints, the complete dynamic equation of motion can be expressed as follows:

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} \tau_a \\ \tau_h \end{bmatrix}. \tag{2.2.12}$$

Where,

$$\begin{aligned} M_{11} &= \frac{1}{3} m_1 {L_1}^2 + m_2 {L_1}^2, \\ M_{12} &= M_{21} = \frac{1}{2} m_2 L_1 L_2 \cos(q_2 - q_1), \\ M_{22} &= \frac{1}{3} m_2 {L_2}^2, \\ C_1 &= -\frac{1}{2} m_2 L_1 L_2 \dot{q}_2^2 \sin(q_1 - q_2) - \frac{1}{2} m_1 g L_1 \sin q_1 \\ &\quad - m_2 g L_1 sinq_1, \\ C_2 &= -\frac{1}{2} m_2 L_1 L_2 \dot{q}_1^2 sin(q_1 - q_2) - \frac{1}{2} m_2 g L_2 \sin q_2. \end{aligned}$$

In (2.2.12), M_{11} and M_{22} are the effective inertia terms, M_{12} and M_{21} are the coupling inertia terms. C_1 and C_2 are the total of centrifugal, Coriolis, and gravity forces.

2.3 Numerical Model Predictive Control (N-MPC)

In this section, we propose a linear MPC scheme for our balance recovery problem. First, we propose to linearize the dynamic model around the unstable vertical equilibrium point, which is $q_1 = q_2 = 0$ in our system.Because the disturbance of the upright standing body is considered to be small enough, it is possible to linearize the dynamic equation of motion. The state vector $\boldsymbol{X}$ and the control input vector $\boldsymbol{\tau}$ are defined as follows:

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \dot{q}_1 \\ \dot{q}_2 \end{bmatrix}, \boldsymbol{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} \tau_a \\ \tau_h \end{bmatrix}.$$

Then, the dynamic equation of motion can be converted via linearization to

$$\dot{\boldsymbol{X}} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}\boldsymbol{\tau}, \tag{2.3.1}$$

where $\boldsymbol{A}$ is the 4×4 state matrix and $\boldsymbol{B}$ is the 4×2 control matrix. After obtaining the state-space representation of our model, we introduce the concept of MPC and propose our N-MPC approach with boundary conditions.

MPC is also referred to as receding horizon predictive control [75]. It should be noted that we use discrete-time MPC because the proposed MPC is implemented in discrete time through a combination of discrete-time state space functions. Here, the discrete-time state space equation is given by:

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{\tau}(k), \tag{2.3.2}$$

where $\boldsymbol{x}(k+1)$ represents the 4×1 vector of the angles and angular velocities of the ankle and hip joints at time $k+1$. $\boldsymbol{x}(k)$ represents the 4×1 vector of the angles and angular velocities of the ankle and hip joints at time k and $\boldsymbol{\tau}(k)$ represents the 2×1 vector of the

ankle and hip torques at time k. To implement N-MPC with boundary conditions, the cost function and constraints need to be defined as follows in a finite time N as follows.

The cost function is

$$J\left(\boldsymbol{x}(0), \boldsymbol{\tau}_{(0,N-1)}\right) = \sum_{0}^{N-1} l(\boldsymbol{x}, k, \boldsymbol{\tau}) + V_f \tag{2.3.3}$$

$$l(\boldsymbol{x}, k, \boldsymbol{\tau}) = \frac{1}{2}\left(\boldsymbol{x}^T(k)\boldsymbol{Q}\boldsymbol{x}(k) + \boldsymbol{\tau}^T(k)\boldsymbol{R}\boldsymbol{\tau}(k)\right),$$
$$V_f = \frac{1}{2}\boldsymbol{x}^T(N)Q_f\boldsymbol{x}(N),$$

where $\boldsymbol{Q} > 0$ and $\boldsymbol{Q}_f > 0$ are 4×4 real symmetric matrices and $\boldsymbol{R} > 0$ is a 2×2 real symmetric matrix. $\boldsymbol{Q}$ and $\boldsymbol{R}$ can be used as tuning parameters to penalize the states and the control inputs. The terminal weighting $\boldsymbol{Q}_f$ is defined to be equal to the solution of the algebraic Riccatti equation (ARE) [76]. This makes V_f become a Lyapunov function to achieve stable MPC performance. Then, by tuning $\boldsymbol{Q}$ and $\boldsymbol{R}$ into suitable values, the MPC controller can be improved.

The objective is to minimize $J\left(\boldsymbol{x}(0), \boldsymbol{\tau}_{(0,N-1)}\right)$ subject to the following constraints:

1) The discrete time state space function:

$$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{\tau}(k).$$

2) For all $i = 1, 2$ and k = $0, 1, 2, ..., N-1$, the torques should satisfy

$$\tau_{min}(i) \leq \tau_i(k) \leq \tau_{max}(i),$$

where $\tau_{min}(1) = -20\ [Nm]$, $\tau_{min}(2) = -100\ [Nm]$, $\tau_{max}(1) = 20\ [Nm]$, and $\tau_{max}(2) = 100\ [Nm]$.

3) For all $i = 1,...,4$ and k = $0,...,N$, the system states satisfy:

$$x_{min}(i) \leq x_i(k) \leq x_{max}(i),$$

where $x_{min}(1) = -0.26\ [rad]$, $x_{min}(2) = -0.35\ [rad]$, $x_{min}(3) = -\infty\ [rad/s]$, $x_{min}(4) = -\infty\ [rad/s]$, $x_{max}(1) = 0.5\ [rad]$, $x_{max}(2) = 1.4\ [rad]$, $x_{max}(3) = \infty\ [rad/s]$, and $x_{max}(4) = \infty\ [rad/s]$.

4) For k = $0,...,N$, the CoP should satisfy:

$$CoP_{min} \leq CoP_i(k) \leq CoP_{max}(i),$$

where $CoP_{min} = -0.15\ [m]$ and $CoP_{max} = 0.15\ [m]$.

5) For k = $0,...,N$, the CoM should satisfy:

$$CoM_{min} \leq CoM_i(k) \leq CoM_{max}(i).$$

where $CoM_{min} = -0.15\ [m]$ and $CoM_{max} = 0.15\ [m]$.

The N-MPC problem described above can be solved as an iterative open-loop optimal control problem with a finite horizon and an observable initial state for each sampling time. For instance, let N-MPC starts at $k = 0$ with the observed initial states $\boldsymbol{x}(0) = \boldsymbol{x}$ and a prediction horizon $k = N$ (here $N = 20$). Then, the prediction-based optimal control sequence for the whole horizon can be obtained as

$$\boldsymbol{\tau}_{opt} = [\boldsymbol{\tau}_{opt}(0), \boldsymbol{\tau}_{opt}(1), \boldsymbol{\tau}_{opt}(2)...\boldsymbol{\tau}_{opt}(N-1)] \tag{2.3.4}$$

The sequence of the predicted states is given by

$$\boldsymbol{x}_{opt} = [\boldsymbol{x}_{opt}(1), \boldsymbol{x}_{opt}(2)...\boldsymbol{x}_{opt}(N)] \tag{2.3.5}$$

Then, the first sample of the obtained optimal control sequence $\boldsymbol{\tau}_{opt}(0)$ is applied to the system and produces the states $\boldsymbol{x}(1)$. Here, $\boldsymbol{x}(1)$ are the observed states, which can be

identical or different from the predicted states $\boldsymbol{x}_{opt}(1)$. In the next sampling time, $\boldsymbol{x}(1)$ becomes the new initial variables for the new optimal control problem at the sampling time $k = 1$. Then, the N-MPC repeats the above described optimal process and obtains the new optimal control inputs for the current system. Afterward, the new initial state variables can be observed for the forthcoming optimal process. Thus, N-MPC is an iterative optimal control algorithm. The control loop of N-MPC in this research is illustrated in Figure 2.5.

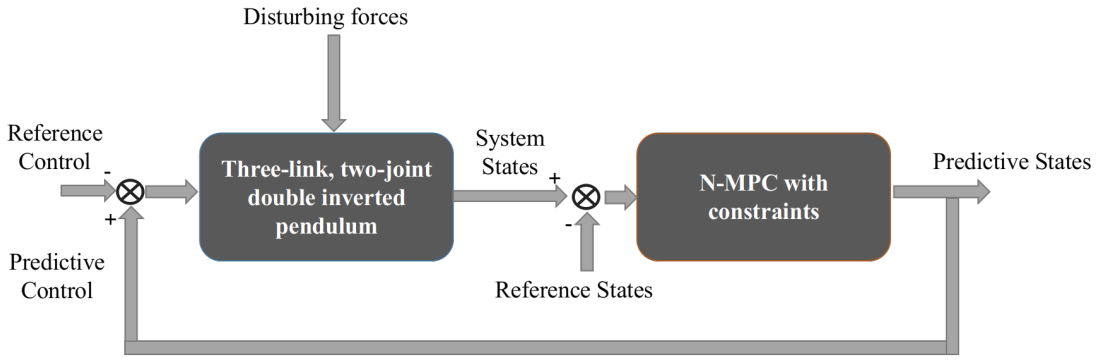


Figure 2.5: The control loop of N-MPC in this work.

Stability analyses of MPC have been discussed from different perspectives in the literature [77]–[80]. Here, the stability of the proposed N-MPC is analyzed in a concise form. Here, the sufficient conditions here that ensure closed-loop asymptotic stability are obtained from a previous work [77]:

A1: state constraint satisfied in the terminal constraint set.

A2: control constraint satisfied in the terminal constraint set.

A3: the terminal constraint set is positively invariant under the control law.

A4: $V_f(x(k+1)) - V_f(x(k)) + l(\boldsymbol{x}, k, \boldsymbol{\tau}) \leq 0$, where $V_f(\cdot)$ is a local Lyapunov function.

With the constraints set for the proposed in N-MPC, conditions A1 through A3 are satisfied. Let $\boldsymbol{Q_f} > 0$ satisfy the Lyapunov equation

$$\boldsymbol{A}^T \boldsymbol{Q}_f \boldsymbol{A} + \boldsymbol{Q} = 0.$$

Then, $V_f = (1/2)\boldsymbol{x}^T \boldsymbol{Q}_f \boldsymbol{x}$ satisfies A4 with equality. Thus, the closed-loop system with N-MPC is asymptotically stable, which means that all the states converge to the origin. This is also verified by the obtained results, which are presented in the next section.

To implement the control scheme N-MPC scheme in the hip-ankle balance recovery simulations, we used CasADi 3.4.5 to solve the numerical optimization problems [81].

CasADi is an open-source tool that implements algorithmic differentiation (AD) on user-defined symbolic expressions. CasADi also provides standardized interfaces to a variety of numerical routines, such as the simulation, optimization, and solution of linear and nonlinear equations. The IPOPT solver, which is based on the primal-dual nonlinear interior-point (IP) method, was used in the proposed N-MPC scheme. IPOPT can solve optimization problems with boundary constraints for all variables. A multiple-shooting technique was applied for faster numerical integration and optimization. A fourth-order Runge-Kutta method (RK4) was used for the numerical integration of ordinary differential equations (ODE) [82].

The control method described in [72] is effective for three kinds of constraints, namely actuated state constraints, underactuated state constraints, and constraints on some specific composite variables. However, in this study, we need to simultaneously consider the state and input constraints, to meet the requirements of a human-like balance behavior. This is also one of the advantages of the proposed controller. To the best of our knowledge, the method proposed in the aforementioned study is not straightforwardly applicable to solve our problem.

2.4 Numerical Simulation of the Proposed recovery strategy

In this section, we describe our implementation of the proposed human-like hip-ankle strategy for balance recovery and the control scheme based on N-MPC in the Mujoco simulation environment [68]. The sampling period used was 0.01 $[s]$. The obtained simulation results are described and analyzed from kinematic, dynamic, energy consumption perspectives.

2.4.1 Kinematic and Dynamic Analysis

In this section, we analyze the kinematic and dynamic aspects of the proposed hip-ankle strategy, such as joint angles and velocities, CoM, CoP, and control inputs. We pushed the position of the CoM of the upper body with different disturbing forces along the same

direction within 0.1 $[s]$. The disturbing forces were set as follows: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, and 120 $[N]$. The model can recover balance after a perturbation within a recovery time of 12 $[s]$. The state weight $\boldsymbol{Q}$ and the input weight $\boldsymbol{R}$ are unchangeable, $\boldsymbol{I}$ is a 4×4 identity matrix.

$$\boldsymbol{Q} = 10^{3} * \boldsymbol{I},$$
$$\boldsymbol{R} = 10^{-4} * \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix}.$$

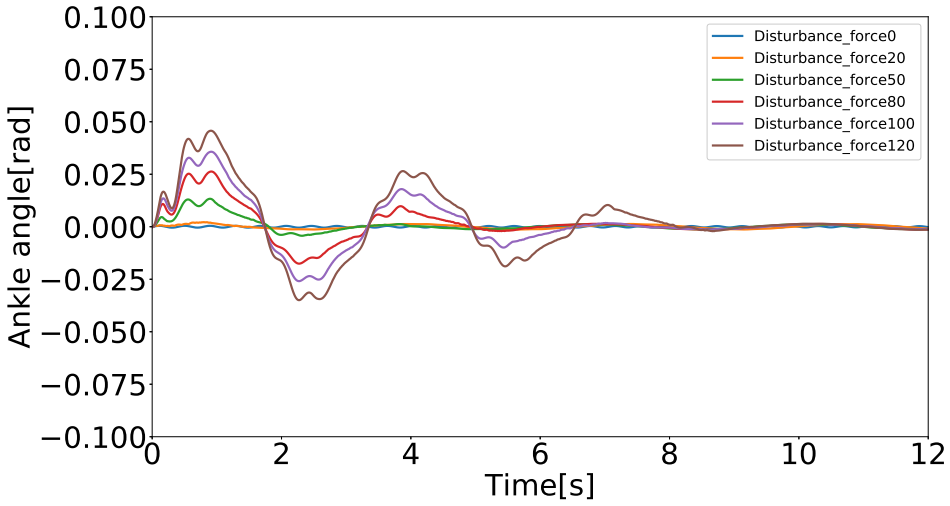


Figure 2.6: Evolution of the ankle angle versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

The evolution of the ankle angle for different disturbing forces is shown in Figure 2.6. Based on these results, we compare the influence of different external forces on the amplitude of the ankle angle and balance recovery time. For small disturbing forces, the ankle joint amplitude changes slightly, and the recovery time is also short. This is similar to human-like balance because, for a small pushing force, our body sways a little and maintains balance easily. For larger disturbing forces, the ankle joint changes considerably, and thus balance recovery takes a longer time. Figure 2.7 illustrates the relationship between the ankle joint position and velocity in a phase portrait representation. It can be seen that no matter how large the disturbance force is, the cycle finally converges to the origin $(0,0)$.

Figure 2.8 shows the evolution of the hip joint angle for different disturbing forces. From

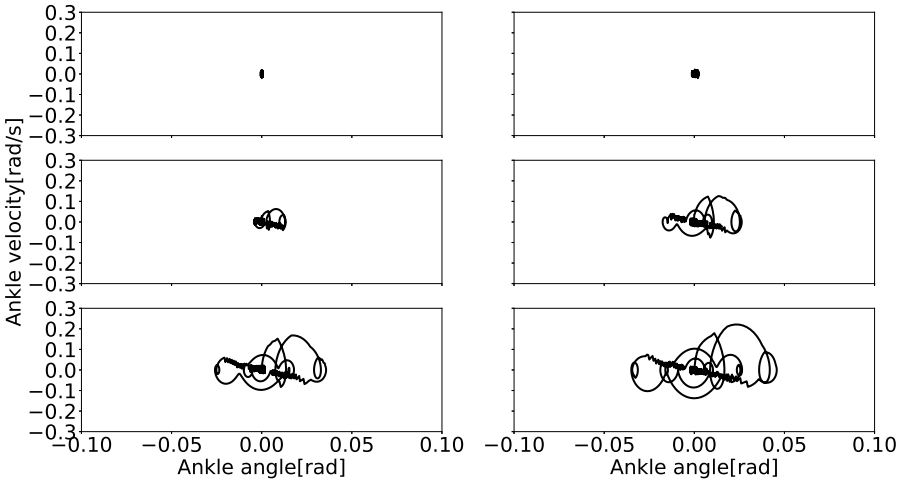


Figure 2.7: Evolution of the ankle phase versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

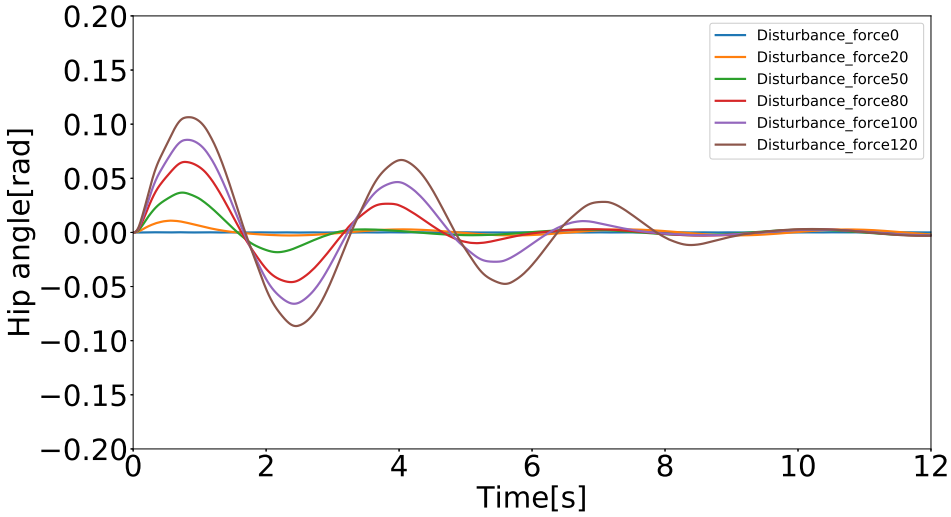


Figure 2.8: Evolution of the hip angle versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

the changes in amplitude of the hip joint angle, it is worth noting that for a disturbing force of 20 $[N]$, the hip starts to react to achieve balance recovery, which indicates that the ankle strategy contains a small amount of hip rotation. Similar results have been reported in human postural balance experiments by Nashner et al.[38] and Horak *et al.* [39]. Moreover, for a disturbing force of 50 $[N]$, hip rotation plays an important role, as evidenced by comparing the ankle joint amplitudes shown in Fig. 2.6 and the hip joint angle amplitudes shown in Fig. 2.8. For disturbing forces of 100 $[N]$ and 120 $[N]$, the sway of the ankle is not enough to maintain balance and therefore the hip sways as well for balance control. These results indicate that the hip-ankle strategy (not a pure hip strategy) is used to maintain balance against a certain range of disturbing forces, which is similiar to

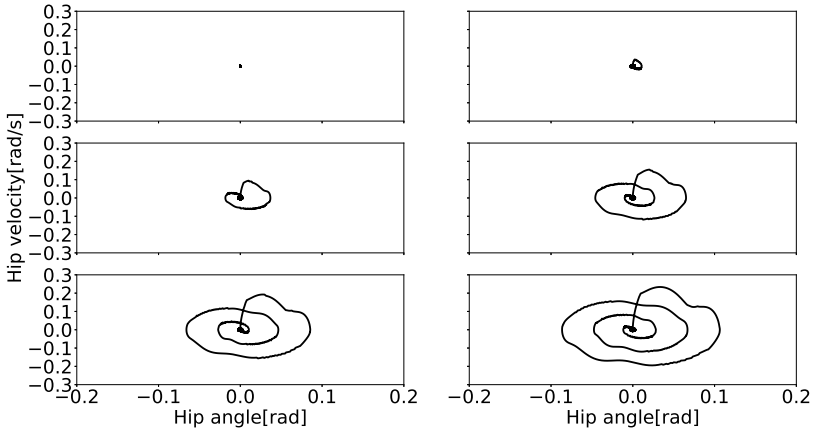


Figure 2.9: Evolution of the hip phase portrait versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

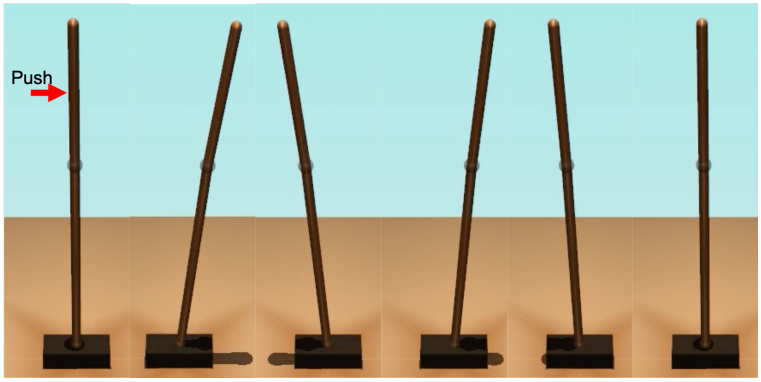


Figure 2.10: Screenshot of simulation animation for the disturbing force: 120 $[N]$.

the published results of human movement experiments published by Runge *et al.* [83]. It is also worth noting that the hip joint angle amplitude is larger than the ankle joint amplitude, which is similar to the results of human experiments published by Colobert *et al.* [84]. The hip joint angle also converges to zero but takes a longer time than the ankle joint angle. The relationship between hip velocity and hip angle is illustrated through a phase portrait in Fig. 2.9. It can be seen that the cycle shown in Fig. 2.9 is larger than that shown in Fig. 2.7. This comparison also shows that for larger disturbances, our model favors using more hip movement because ankle movement is not sufficient to maintain balance. In our simulations, the ankle and hip movements observed were different from the results of Aftab *et al.* [30]–[33] where the upper body was not included. A screenshot of a simulation animation for a disturbing force of 120 $[N]$ is shown in Fig. 2.10. We note that the deviation of the hip angle is within '$\pm$ 0.15 $[rad]$' owning to our linearization assumption. This is an important limitation of the proposed model.

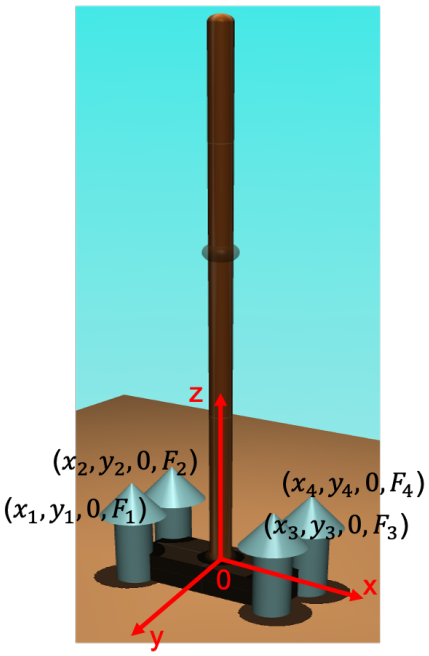


Figure 2.11: The schematic diagram for calculating the CoP location. The origin of world coordinate system is point O at the foot botterm center. The positions and ground reaction forces of four load cells under pressure are defined by $(x_1,\ y_1,\ 0,\ F_1)$, $(x_2,\ y_2,\ 0,\ F_2)$, $(x_3,\ y_3,\ 0,\ F_3)$, $(x_4,\ y_4,\ 0,\ F_4)$.

The definitions of the CoP and the CoM are given as follows. A schematic diagram for calculating the location of the CoP is shown in Figure 2.11. The origin of the world coordinate system is point O at the center of the foot's bottom. The positions and ground reaction forces of four load cells under pressure are defined by $(x_1,\ y_1,\ 0,\ F_1)$, $(x_2,\ y_2,\ 0,\ F_2)$, $(x_3,\ y_3,\ 0,\ F_3)$, and $(x_4,\ y_4,\ 0,\ F_4)$. Because the model sways in the x-axis direction, the CoP and CoM in the y-axis direction are always zeros and are omitted in the study. The formula for calculating the location of the CoP in the x-axis direction is as follows:

$$CoP = \frac{F_1x_1 + F_2x_2 + F_3x_3 + F_4x_4}{F_1 + F_2 + F_3 + F_4}. \tag{2.4.1}$$

The CoM is calculated with the CoMs in the x-axis direction and the masses of the foot, the lower body, and the upper body. The CoMs in each part of the model can be obtained online during the simulation using the Mujoco API and are represented by x_f, x_l, and x_u. In addition, m_0, m_1, and m_2 represent the masses of the foot, the lower body, and the upper body, respectively, as mentioned in Section 2. The formula for calculating the location of

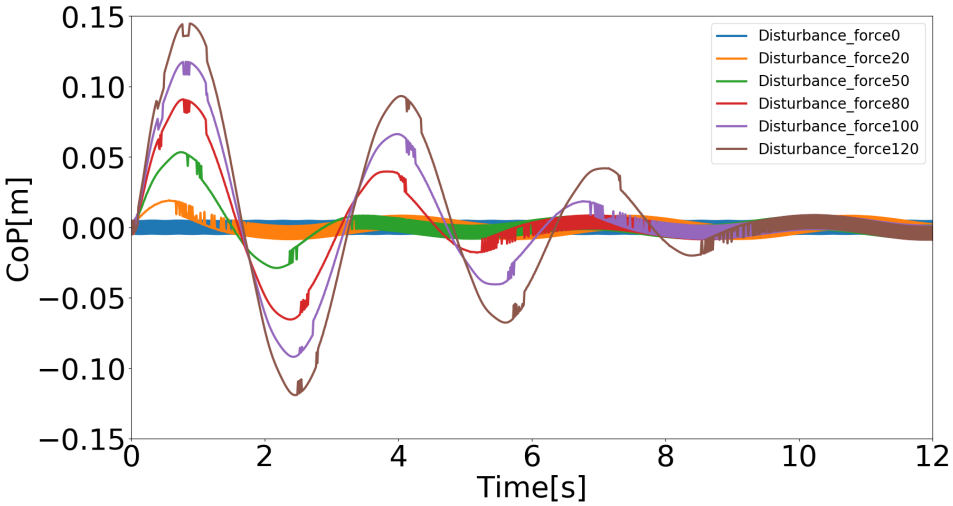


Figure 2.12: Evolution of the CoP versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

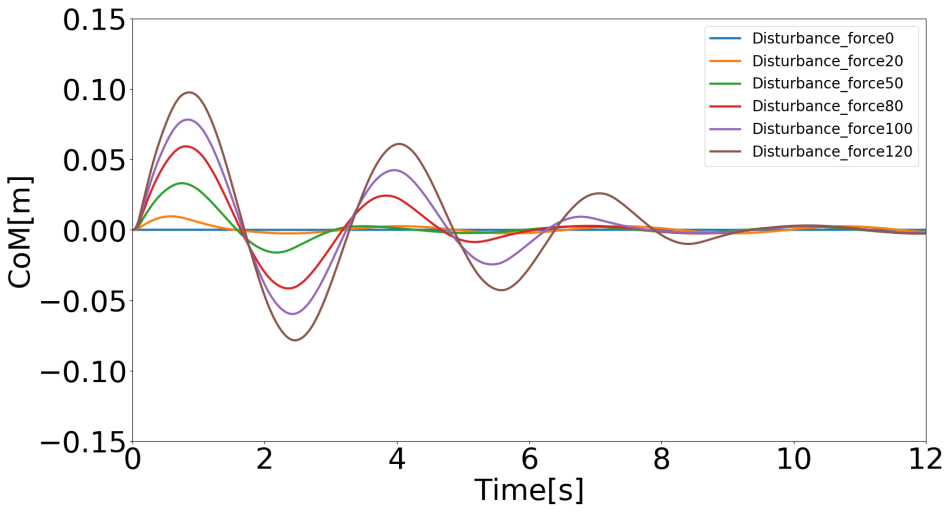


Figure 2.13: Evolution of the CoM versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

the CoM in the x-axis direction is as follows:

$$CoM = \frac{m_0 x_f + m_1 x_l + m_2 x_u}{m_0 + m_1 + m_2}. \tag{2.4.2}$$

The evolution of the CoP and the CoM is represented in Figures 2.12 and 2.13, respectively. One can observe that the CoP and CoM amplitudes become progressively higher as the disturbing force becomes greater. In addition, it is worth noting that the CoP amplitude is

generally larger than the CoM amplitude. In our simulations, We make the foot model not leave the ground via unilateral constraints. We also take the CoP as a criterion to evaluate the dynamic stability of the body. The CoP remains permanently inside the footprint. If the CoP were outside the footprint, our optimization problem would become infeasible, and maintaining the balance behavior could not be guaranteed.

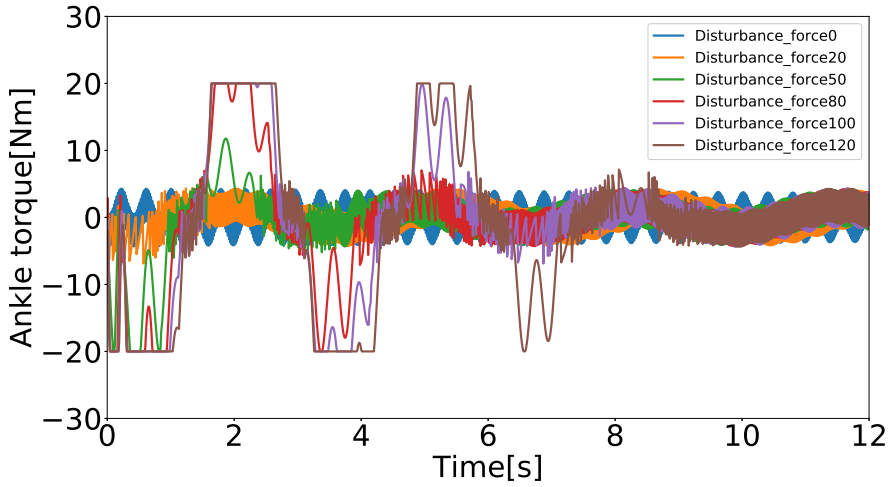


Figure 2.14: Evolution of the ankle balance recovery torque versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

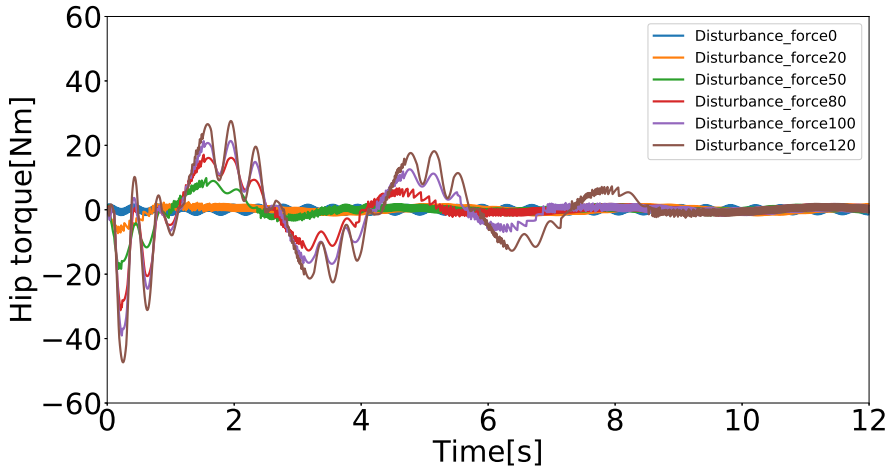


Figure 2.15: Evolution of the hip balance recovery torque versus time for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

The changes of the ankle and hip reaction torques for different disturbing forces are depicted in Figures 2.14 and 2.15, respectively. Here, we consider boundary constraints of the input torques (ie. saturation) for the ankle and hip joints. This makes the behavior of our balance model more similar to the hip-ankle strategy. From Figure 2.14, it can be

seen that the ankle input torque starts to be saturated at -20 $[N]$ for a disturbing force of 50 $[N]$. This indicates that when the pushing force becomes larger enough, the ankle joint torque produces a maximum torque of -20 $[N]$ to maintain balance. However, the ankle input torque seems insufficient to maintain balance, and thus the hip input torque is produced to help the body maintain balance. By observing the ankle and hip input torques for pushing forces of 100 $[N]$ and 120 $[N]$, we find that larger hip input torques are used and the balance recovery duration becomes longer.

In this subsection, we analyzed the evolution of the angles, phase portraits, CoP, CoM, and input torques of the ankle and hip joints. As shown through our simulation results, we implemented a hip-ankle strategy with a unilaterally constrained foot and analyzed the resulting balance behavior against the different pushing forces.

2.4.2 Energy Consumption Viewpoint

In this subsection, we analyze our implementation of the hip-ankle strategy and from a new analysis perspectives: energy consumption. The joint energy is calculated as follows, where, τ is the joint input torque and $\dot{q}$ is the angular velocity, t is time.

$$W = \int \tau \dot{q} dt. \tag{2.4.3}$$

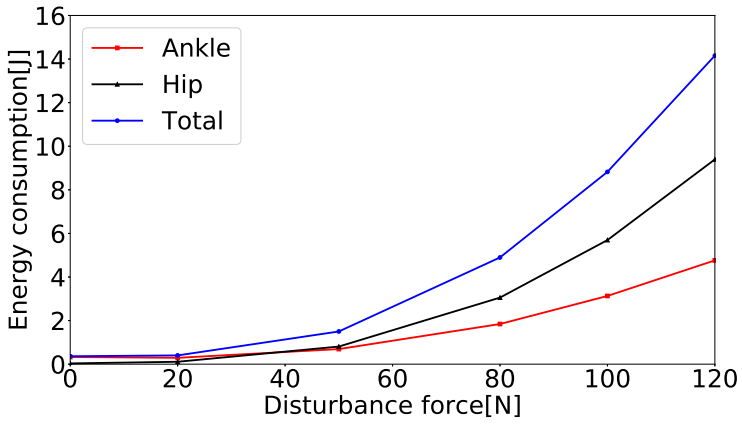


Figure 2.16: Evolution of the ankle, hip and total energy respectively for different disturbing forces: 0 $[N]$, 20 $[N]$, 50 $[N]$, 80 $[N]$, 100 $[N]$, 120 $[N]$.

Figure 2.16 clearly shows that as the disturbing force increases, the energy consumption increases to maintain balance. We compared the energy consumption of the ankle joint

with that of the hip joint and note three main concluding observations. First, the energy consumption of the hip is larger than that of the ankle, for disturbing forces greater than 50 $[N]$. This indicates that when the pushing force is considerable, the hip joint needs to make a higher effort to maintain balance. Secondly, for pushing forces lower than 50 $[N]$, the energy consumption of the ankle is larger than that of the hip. This indicates that the ankle a higher effort for balance control than the hip joint for small disturbing forces. Thirdly, as pushing forces increase, the energy consumption of our model for balance control increases.

The state weight $\boldsymbol{Q}$ and input weight $\boldsymbol{R}$ in the cost function (2.3.3) were adjusted to find the minimum energy consumption for balance recovery. The condition for adjusting the weights is to recover balance within 12 $[s]$. First, $\boldsymbol{Q}$ is varies over five cases while $\boldsymbol{R}$ is not changed as follows. $\boldsymbol{I}$ is a 4×4 identity matrix. The energy consumption of the ankle and hip joints, as well as the total energy consumed, are depicted in Figure 2.17.

$$
\begin{aligned}
\boldsymbol{Q}_1 &= 10^3 \boldsymbol{I}, \\
\boldsymbol{Q}_2 &= 10^4 \boldsymbol{I}, \\
\boldsymbol{Q}_3 &= 10^5 \boldsymbol{I}, \\
\boldsymbol{Q}_4 &= 10^6 \boldsymbol{I}, \\
\boldsymbol{Q}_5 &= 10^7 \boldsymbol{I}, \\
\boldsymbol{R} &= 10^{-4} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}
$$

Then, $\boldsymbol{Q}$ is kept unchanged and $\boldsymbol{R}$ is adjusted over five cases as follows. Figure 2.18 shows

the energy consumption of the ankle and hip joints and the total energy consumed.

$$\boldsymbol{Q} = 10^3 \boldsymbol{I},$$

$$\boldsymbol{R}_1 = 10^{-4} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\boldsymbol{R}_2 = 10^{-5} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\boldsymbol{R}_3 = 10^{-6} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\boldsymbol{R}_4 = 10^{-7} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\boldsymbol{R}_5 = 10^{-8} \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix}.$$

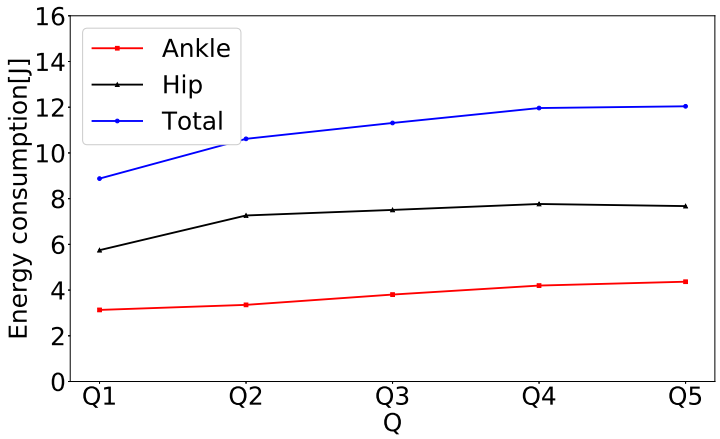


Figure 2.17: Evolution of the ankle energy, hip energy and total energy of both joints for different state weight $\boldsymbol{Q}$ settings and a constant control weight $\boldsymbol{R}$: $\boldsymbol{Q}_1$, $\boldsymbol{Q}_2$, $\boldsymbol{Q}_3$, $\boldsymbol{Q}_4$, $\boldsymbol{Q}_5$.

In these calculations, the disturbing force was set to 100 $[N]$ and its acting time was 0.1 $[s]$. By comparing the energy consumption values shown in Figures 2.17 and 2.18, which were obtained by respectively changing the settings of $\boldsymbol{Q}$ and $\boldsymbol{R}$, we found that the weight settings $\boldsymbol{Q}_1$ and $\boldsymbol{R}_1$ yielded the minimum energy consumption. Thus, we choose weight settings $\boldsymbol{Q}_1$ and $\boldsymbol{R}_1$ as the initial settings in our balance simulations, which is representative of human behavior. When a human is pushed, their body tries to predict a way to maintain balance with low energy consumption.

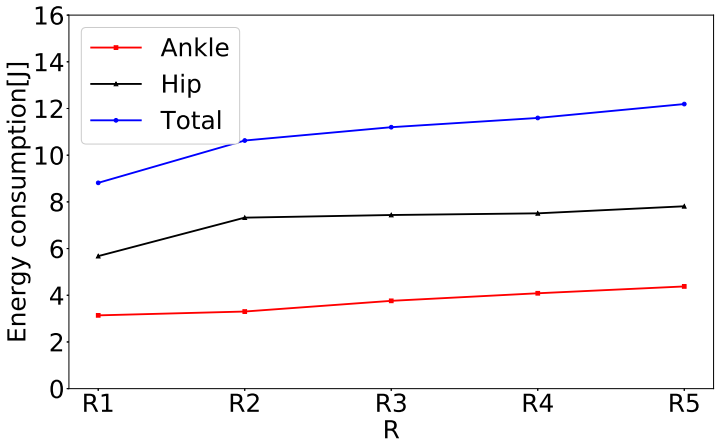


Figure 2.18: Evolution of the ankle energy, hip energy and total energy of both joints for different control weight $\boldsymbol{R}$ settings and a constant state weight $\boldsymbol{Q}$: $\boldsymbol{R}_1$, $\boldsymbol{R}_2$, $\boldsymbol{R}_3$, $\boldsymbol{R}_4$, $\boldsymbol{R}_5$.

Table 2.2: The ankle joint energy, hip joint energy and total energy of both joints for different disturbing forces and their application time under consideration of balance recovery within 12 $[s]$.

Force [N]	Force time [s]	Ankle joint energy [J]	Hip joint energy [J]	Total energy [J]
5	0.10	0.286	0.042	0.356
5	8.00	0.340	0.227	0.567
10	0.10	0.290	0.066	0.356
10	7.00	1.851	2.554	4.405
15	0.10	0.287	0.083	0.370
15	6.00	2.386	2.408	4.794
20	0.10	0.291	0.110	0.401
20	0.70	4.018	8.016	12.034
30	0.10	0.354	0.245	0.600
30	0.45	4.351	8.818	13.169
40	0.10	0.487	0.455	0.942
40	0.30	3.630	7.236	10.866
50	0.10	0.724	0.822	1.546
50	0.25	4.196	8.491	12.687
60	0.10	1.074	1.358	2.432
60	0.20	4.131	8.162	12.193
80	0.10	1.842	3.054	4.896
100	0.10	3.129	5.693	8.822
120	0.10	4.762	9.395	14.157

We then analyzed the robustness of the proposed controller. We tested the longer time acting for different disturbances and found the maximum acting time that allows for balance recovery within 12 $[s]$. The results obtained are shown in Table. 2.2. For instance, when the disturbing force is 5 $[N]$, the maximum acting time is 8.00 $[s]$. When the disturbing force is 10 $[N]$, the maximum acting time is 7.00 $[s]$. For a disturbing force of 15 $[N]$, the maximum acting time is 6.00 $[s]$. In contrast, for a disturbing force of 20 $[N]$, the maximum acting time is 0.70 $[s]$, etc. These results indicate that the proposed controller has good robustness. For an acting time of 8.00 $[s]$ and a force of 5 $[N]$, the ankle joint energy $(0.340$ [J]$)$ is larger than the hip joint energy $(0.227$ [J]$)$. Therefore, the ankle strategy plays a crucial role in this case. However, for an acting time of 7 $[s]$ and a force of 10 $[N]$, the ankle joint energy $(1.851$ [J]$)$ is lower than the hip joint energy $(2.554$ [J]$)$. The hip strategy thus plays a major role when the disturbing force and acting time are increased.

2.4.3 Discussion

The stability of the closed-loop system with the proposed control strategy, which was previously analyzed, was confirmed through the results presented in this section.

In the cost function (2.3.3), the states and input torques are adjusted to desired values. In our case, the desired values are zeros. This makes the model remain upright around the vertical equilibrium point. Figures 2.6, 2.14, 2.8, and 2.15 respectively show that the angles and torques of the ankle and hip joints converge to zeros after several seconds. The limit cycles indicate that the evolution of the angles and angular velocities of the ankle and hip versus time in Figures 2.7 and 2.9 converge to the origin. The CoM and CoP are determined by the model postures and input torques. Thus, similar results are shown in Figures 2.13 and 2.12, respectively. In the proposed N-MPC, the ankle input torque constraints are given, and Figure 2.14 shows the input saturation of the ankle. However, the hip input torque is not saturated simultaneously, which indicates that the hip makes a much higher effort to maintain balance in our model. The energy consumption of the model is determined by the balance recovery time and the angular velocities and input torques of the ankle and hip. As shown in Figures 2.17 and 2.18, changing $\boldsymbol{Q}$ and $\boldsymbol{R}$ in the cost function (2.3.3) can affect the resulting energy consumption. The robustness of the

proposed N-MPC is evident from the results shown in Table. 2.2, where it can be seen that the proposed model can achieve balance recovery within 12 $[s]$ for different disturbing forces and acting times.

2.5 Conclusions

In this paper, we proposed a new model with a unilaterally constrained foot and derived its dynamic equation of motion. Subsequently, we proposed an N-MPC scheme for our model and provided a detailed explanation of our implementation of N-MPC. We implemented the hip-ankle strategy based on the proposed model and controller in a simulated physical environment. Finally, we analyzed the obtained simulation results, which were found to be similar to those of previous human balance experiments in two perspectives: kinematic and dynamic aspects and energy consumption. This helped us gain a better understanding of the hip-ankle strategy from new perspectives. Notably, this study may also be meaningful for the control of exoskeleton devices because N-MPC is very practical for bio-mechanical optimization control.

Chapter 3

Reproducing Arm Strategy Through MPC and its Contribution

This chapter is about arm strategy for balance recovery during quiet standing through nonlinear model predictive control (NMPC). It is worth noting that all the previous works did not cover the verification of the arm strategy with multiple cases, e.g., active arms, passive arms, and fixed arms, in their human experiments to discuss the usefulness of arm rotations. To our best knowledge, these arm strategies are relevant for stability improvement and energy efficiency in human and humanoid/bipedal walking and standing. Furthermore, they did not leverage nonlinear model predictive control (NMPC) for addressing multiple constraints of the ankle, hip, arm joint angles and torques and reproducing human-like balance recovery controller in their artificial systems. The features of NMPC consistent with the capacity of the human body and central nervous system such as constraints handling, predictive horizon, optimization, and robustness are not considered very well in all the previous works.

Therefore, we further studied the mechanism of arm strategy for balance recovery based on previous works and compared the results with human balance recovery experimental results. The contributions of our study are summarized as follows.

1) A three-joint, five-link model is built to represent the human body structure for studying quiet standing balance recovery in the A/P direction. This model includes the foot, the

lower body, the upper body, and the arms.

2) With the system states and the input constraints, an NMPC is proposed from a neuroscience perspective to reproduce human-like balanced behavior evoked by the human central nervous system. The proposed NMPC also has a predictive aspect that allows predicting the future behavior and compute an optimal control balance strategy by minimizing systemic energy consumption of the whole body. Furthermore, the NMPC technique can handle simultaneously the state and input constraints, which is important to meet realistic requirements due to physical limitations of the human body such as joint ranges and torques satuation. All the previously proposed control techniques can not take into account constraints naturally. In this research, we proposed NMPC which can naturally take into account constraints. Different magnitudes of disturbing forces are applied to the model to observe the autonomous switch between the ankle, hip, and arm strategies and to examine the robustness of the proposed solution.

3) Various indexes are verified to evaluate the capability of balance recovery. The root mean square (RMS) deviation and energy consumption are compared for different cases, namely, active arms, passive arms, and fixed arms. These three cases of arm usages are recruited for balance recovery. The obtained data indicate that balance recovery with active arms is the most effective strategy, and balance control with arm usage is better than that without arm usage.

4) Phase portraits of joint angles are considered to analyze the control pattern of balance recovery motion.

5) Ankle torque boundary constraints are set with different values. Besides, the relationship between ankle capacity and active arm usage is discussed since in our daily life ankle is easy to be injured, we want to observe how arm usages contribute to balance in this case.

6) By comparing the results of the numerical simulation and human experiments, human-like balance recovery with arm strategy is implemented and arm movements are found to enhance the capability of balance recovery.

The paper is organized as follows. The hardware, software, and related methodology are introduced in section 3.1. In section 3.2, the simplified model with three different arm

usages and their dynamic equation are introduced. The balance recovery controller based on NMPC is proposed in section 3.3. In section 3.4, the results of simulation and human experiments are discussed to verify if actuated arm usage contributes to balance control. The conclusions of this study are summarized as well as future work in section 3.5.

3.1 Methodology Used in Human Experiment

In this section, we omitted the simulation-related control methods NMPC and the simulation platform MuJoco because we have introduced them in the previous chapter. The hardware for the motion and force data collection such as OptiTrack system, AMTI force plate, the software Opensim used for the computation of the inverse kinematics and the inverse dynamics, and data processing methods about human experiments are described in this section.

In this study, the OptiTrack system shown is an optical motion capture system consisting of 8 cameras, which is used for tracking the human translations through tracking the three-dimensional positions of markers attached on the skin of the subject. These marker trajectories are captured at 100 $[Hz]$ and filtered with a second-order Butter-worth low-pass filter with a cut-off frequency of 6 $[Hz]$. Besides, Motive shown is the software platform for the marker position recording of the human experiment, data processing, and data export. The procedure of Motive used in human experiments is demonstrated in Figure 3.1.

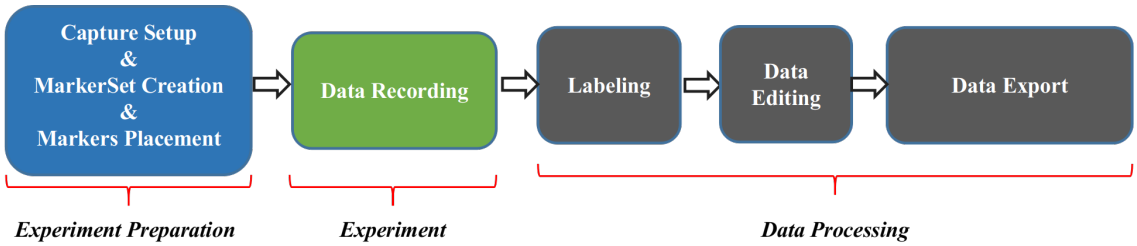


Figure 3.1: The procedure of Motive used in human experiments.

1) **Capture Setup**: The first step is to get Motive ready for actual capture. There are three main steps: Calibration of the capture volume, placing retro-reflective (or active) markers onto subjects, and creating associated assets or models in Motive. Besides, Two AMTI force plates are connected to the system, synchronization settings need to be configured

accordingly.

2) **MarkerSet Creation**: once Motive is prepared, the next step is to place markers on the subject and create a corresponding MarkerSet. For each Take, involved MarkerSet is displayed in the Assets pane, and the related properties show up at the Properties pane when an asset is selected within Motive. The MarkerSet is a list of marker labels that are used to annotate reconstructed markers. After the experiment, we will manually label markers in post-processing. When doing so, having a defined set of labels (MarkerSet) makes this process much easier. MarkerSets within a Take will be listed in the Labeling Pane, and each label can be assigned through the Labeling process.

3) **Data Recording**: then, we start the recording and let Motive automatically generate new names on the fly.

4) **Labeling**: after experiments, we label the maker trajectory manually using the Marker-Set we created.

5) **Data Editing**: after labeling, we observe the labeled marker trajectory in Channels and Graphics. We check tracking errors from recorded capture data and post-process them through the Edit Tools in Motive. Since our experiments exhibit simple balance motions, there are few maker labeling errors. However, Missing marker position sometimes can be found through the trajectory graphs on the Graph View pane. we use interpolation options including Constant, Cubic to fill the trajectory gaps. Then, we apply a noise filter to trajectory data, which modifies the marker trajectory smoother.

6) **Data Export**: After data processing, we export the data including the maker time-series position and ground reaction forces to 'CSV' documents.

3.1.1 AMTI Force Plate

The standard approach to obtain the human joint torques used in the experiment is to compute inverse dynamics. To do this, it is necessary to get the contacting forces with the environment. And, the accuracy of force information is important for achieving reliable data as well. Thus, we choose the AMTI's AccuGait Optimized (c) multi-axis force

platform shown in Figure 3.2 to record the ground reaction force in Figure 3.3 between feet and force plates in the experiment.

Figure 3.2: AMTI's AccuGait Optimized multi-axis force platform is a portable solution for quantifying human gait and balance.

Each type of force platform has its way to compute the ground reaction measures. The set of equations for an AMTI force plate are derived from the output signals labeled F_x, F_y, F_z, M_x, M_y, and M_z shown in Figure 3.4. The six AMTI equations are

$$\begin{aligned} F'_x = F_x f_x \quad F'_y = F_y f_y \quad F'_z = F_z f_z \\ COP_x = -\left(M g_y + F'_x z\right)/F'_z \\ CoP_y = \left(M_x g_x - F'_y z\right)/F'_z \\ M'_z = M_z g_z + F'_x y - F'_y x \end{aligned} \tag{3.1.1}$$

where $F'_x,\ F'_y,\ F'_z$ are the three components of ground reaction force, $(CoP_x,\ CoP_y, 0)$ denote the three coordinates of COP, M'_z is the free moment of force, $f_x,\ f_y,\ f_z$ represent scale factors that convert the forces from voltages $[V]$ to newtons $[N]$, and $g_x,\ g_y,\ g_z$ are scale factors that transform the moments from voltages $[V]$ to newton meters $[Nm]$. The six out signals from an AMTI force plate are fed to an amplifier unit with select-able gain levels that determine the exact values of $f_x,\ f_y,\ f_z$ and $g_x,\ g_y,\ g_z$.

3.1.2 OpenSim

OpenSim is a freely available, open-source software system that lets users develop models of musculoskeletal structures and create dynamic simulations of a wide variety of movements such as grasping, touching, walking, and balance. OpenSim provides a platform where the biomechanics community can build a library of simulations that can be developed through multi-institutional collaboration. In this study, we use the Scale Tool, Inverse

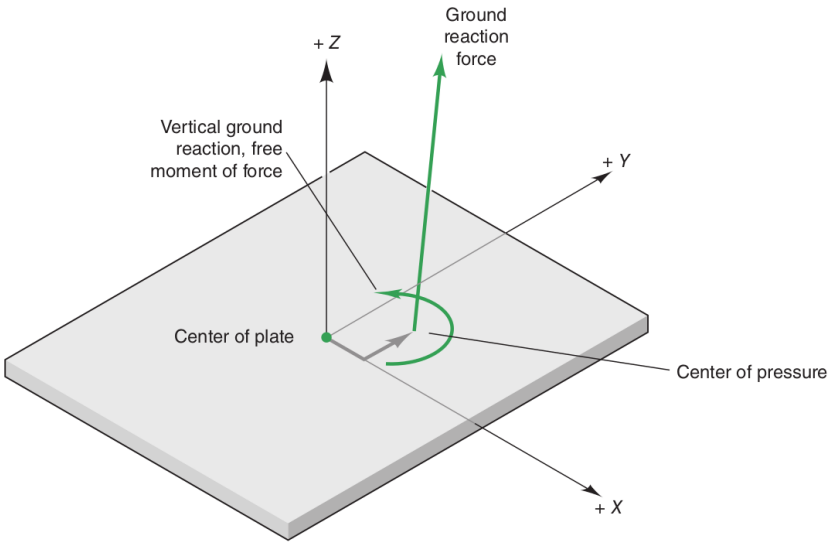


Figure 3.3: AMTI's AccuGait Optimized multi-axis force platform with its reaction to an applied force and vertical moment of force (reprinted from [85]).

Kinematics Tool, and Inverse Dynamics Tool to compute the joint angles and torques. The overview of the computation procedure in OpenSim is shown in Figure 3.5. The detailed processing protocol is described as follows.

1) **Scaling**: a generic musculoskeletal model (Rajagopal2015) [86] is scaled through minimizing the optimal cost function containing the distance between experimental marker data and existing virtual maker settings using the Opensim Scaling Model Tool. Then, a subject-specific model could be obtained. Accordingly, the length, mass, and inertia of each body segment are scaled as well. The comparison of the generic model and the subject-specific model is illustrated in Figure 3.6. This model includes bony geometry for the full-body, 37 DoFs to define joint kinematics, Hill-type models of 80 muscle-tendon units actuating the lower limbs, and 17 ideal torque actuators driving the upper body.

2) **Inverse Kinematics**: applying the Inverse Kinematics (IK) tool in OpenSim, The simulation of the subject's motions could be implemented with the best pose at each time frame computed through minimizing the sum of squared errors between model makers and experimental makers. Furthermore, the IK Tool is a solver of the weighted least squares equation involving minimizing both marker and coordinate errors. Marker error is defined as the deviation between a scaled virtual marker attached to the model and an experimental marker. Each marker has its weight for specifying the impact of the marker's error to

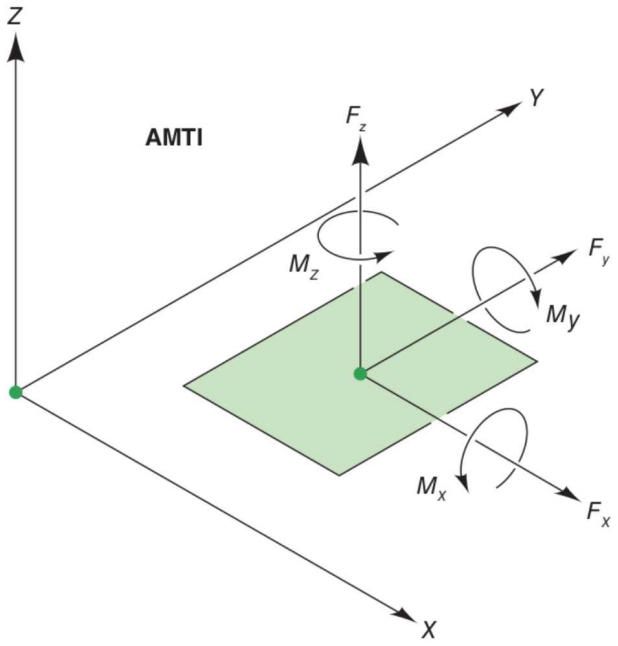


Figure 3.4: Outputs from an AMTI force plate include three forces F_x, F_y, F_z and three moments M_x, M_y, M_z (reprinted from [85]).

derive the best pose. And, the deviation between a computed coordinate value and the one directly obtained from the motion capture system is defined as a coordinate error. In our study, we omit the coordinate error and only consider the marker error. we could compute joint angles in this step. The rough scenario of balance recovery motion and its evolution of joint angles are shown in Figure 3.7 and 3.8, respectively.

3) **Inverse Dynamics**: in this step, joint torque can be computed through the Inverse Dynamics Tool in Opensim based on the joint angles obtained in the previous step and the ground reaction forces of both sides feet from experiments that are applied to the model at the calcaneus bones. The evolution of the joint angles and the torques in human balance recovery motion is demonstrated in Figure 3.9.

3.2 Dynamic Equation of Simplified Models

To achieve quiet standing balance control, we regard the human body structure as a three-joint and five-link model consisting of left-right arm joint, hip joint, ankle joint, and right arm, left arm, upper body, lower body, fixed foot (e.g. Figure 3.10). Table 3.1 summarizes the physical parameters of our model. Based on an existing anthropometric database [71] and the previous work [23] dealing of the optimization-based balance recovery strategy,

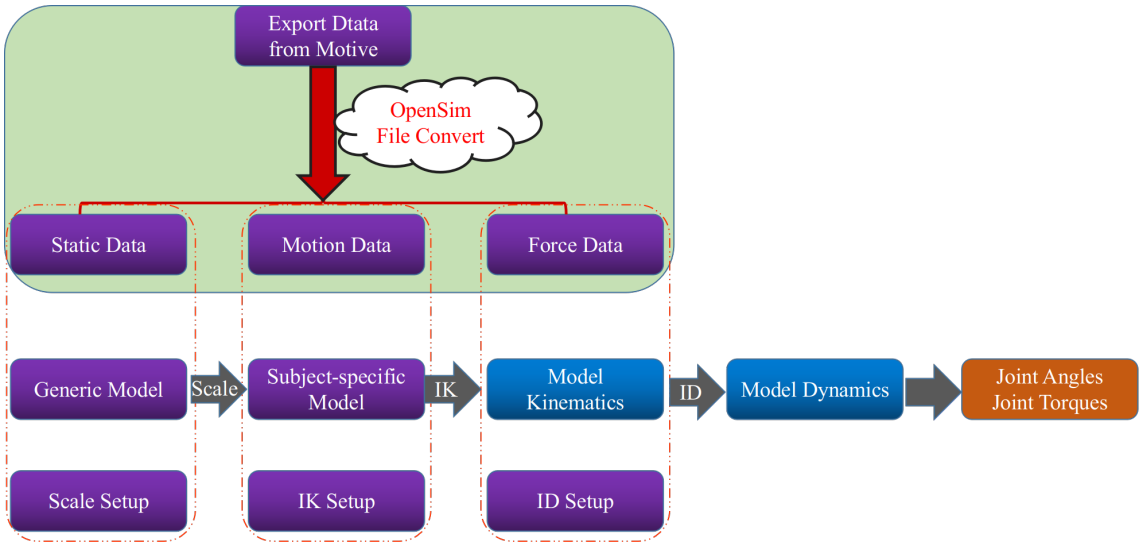


Figure 3.5: Overview of computation procedure in OpenSim.

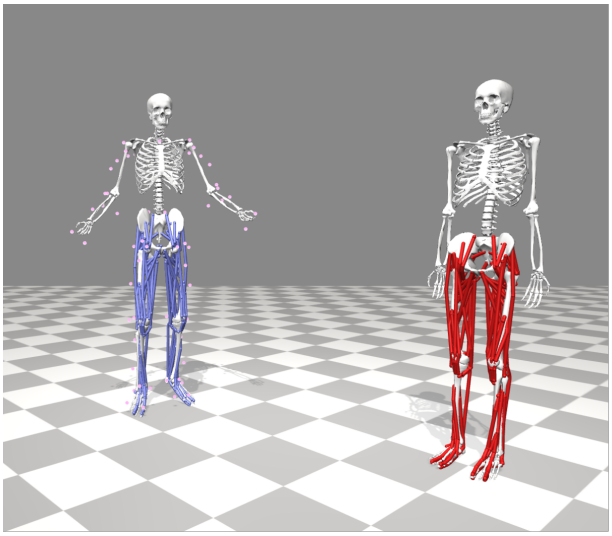


Figure 3.6: The subject-specific model (left) after scaling versus the generic model (right).

the height and mass of the whole-body are 1.7 $[m]$ and 69.3 $[kg]$, respectively. Further, m_4, m_3, m_2, m_1, and m_0 represent the masses of the left arm, right arm, upper body, lower body, and foot, respectively; L_4, L_3, L_2, L_1, and L_0 represent the lengths of left arm, right arm, upper body, lower body, and foot, respectively; and q_3, q_2, and q_1 represent the left-right arm angle, hip angle, and ankle angle, respectively. Note that the body segments between the head and the left-right arm joint, between the left-right arm joint and the hip joint, and between the hip joint and the ankle joint are ignored.

First, the dynamic equations of motion for this three-joint, five-link model controlled by the arm, hip, and ankle joint torques are computed based on Lagrange mechanics [87]. The Lagrange equations and dynamic equation of motions are derived for the model with

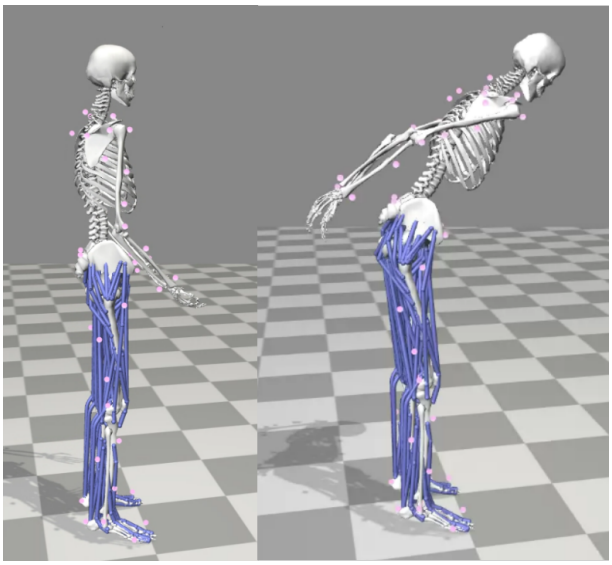


Figure 3.7: The rough scenario of balance recovery motion with arm strategy after IK processing.

Table 3.1: Physical parameters of the three-joint, five-link model.

Link	Mass [kg]	Length [m]	Height [m]
Foot	1.3	0.3	0.1
Lower body	35	1.0	–
Upper body	25	0.6	–
Right arm	4	0.6	–
Left arm	4	0.6	–
Total mass [kg]	69.3	–	–
Total height [m]	–	–	1.7

three different arm states separately: active arms, passive arms, and fixed arms, as shown in Table 3.2. In that table, T and V represent the kinetic and potential energy, respectively; τ_{arm}, τ_{hip}, and τ_{ankle} represents the arm torque, hip torque, and ankle torque, respectively; M_{11}, M_{12}, M_{13}, M_{21}, M_{22}, M_{23}, M_{31}, M_{32} and M_{33} are the inertia terms; and C_1, C_2, and C_3 denote the total centrifugal, Coriolis, and gravity forces.

3.3 Proposed NMPC for Balance Recovery

In this section, an NMPC scheme [70] is proposed to resolve the balance recovery problem. The control loop of NMPC for the implementation of reproducing arm strategy for balance

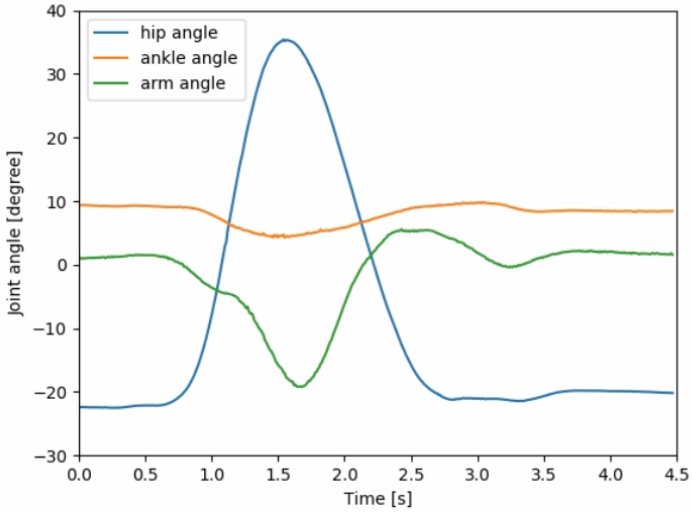


Figure 3.8: The computed joint angles of balance recovery motion with arm strategy through IK tool.

Table 3.2: Lagrange equations and dynamic equation of motions for the model with three different cases: (1) active arms; (2) passive arms; (3) fixed arms.

Case	Lagrange equations	Dynamic equation of motion
(1)	$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right)-\frac{\partial L}{\partial q_1}=\tau_{ankle}$, $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_2}\right)-\frac{\partial L}{\partial q_2}=\tau_{hip}$, $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_3}\right)-\frac{\partial L}{\partial q_3}=\tau_{arm}$, $L=T-V$.	$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} \tau_{ankle} \\ \tau_{hip} \\ \tau_{arm} \end{bmatrix}$
(2)	$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right)-\frac{\partial L}{\partial q_1}=\tau_{ankle}$, $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_2}\right)-\frac{\partial L}{\partial q_2}=\tau_{hip}$, $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_3}\right)-\frac{\partial L}{\partial q_3}=0$, $L=T-V$.	$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} \tau_{ankle} \\ \tau_{hip} \\ 0 \end{bmatrix}$
(3)	$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right)-\frac{\partial L}{\partial q_1}=\tau_{ankle}$, $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_2}\right)-\frac{\partial L}{\partial q_2}=\tau_{hip}$, $L=T-V$.	$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} \tau_{ankle} \\ \tau_{hip} \end{bmatrix}$

recovery in simulation is illustrated in Figure3.11. This problem can be solved as an iterative open-loop optimal control problem with a finite horizon and an observable initial states for each sampling time. The procedure of NMPC with constraints is illustrated in Figure 3.12 to strengthen the NMPC concept explanation. For example, let NMPC starts at

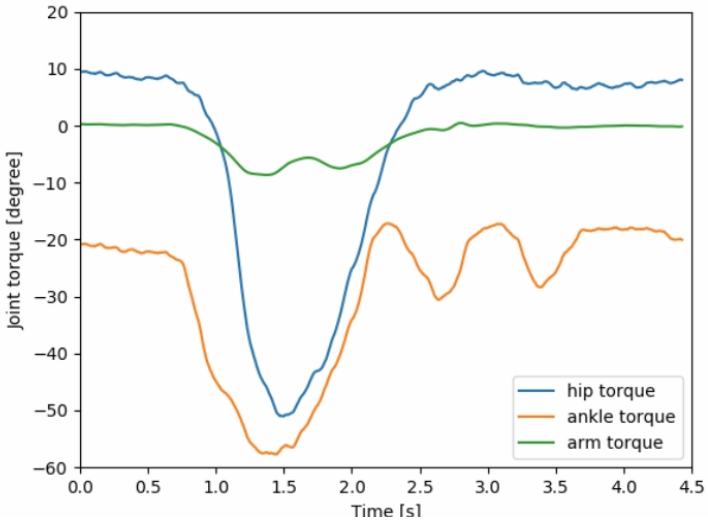


Figure 3.9: The computed joint torques of balance recovery motion with arm strategy through ID tool.

$k = 0$ with a prediction horizon $k = N_t$ (here $N_t = 5$) and the initial states $\boldsymbol{x}(0) = \boldsymbol{x}$. The predictive optimal control sequence for the entire horizon can be obtained as follows,

$$\boldsymbol{\tau}_{opt} = [\boldsymbol{\tau}_{opt}(0), \boldsymbol{\tau}_{opt}(1), \boldsymbol{\tau}_{opt}(2)...\boldsymbol{\tau}_{opt}(N_t - 1)] \tag{3.3.1}$$

The sequence of the predicted states is denoted by,

$$\boldsymbol{x}_{opt} = [\boldsymbol{x}_{opt}(1), \boldsymbol{x}_{opt}(2)...\boldsymbol{x}_{opt}(N_t)] \tag{3.3.2}$$

Then, the first sample of the optimal control sequence $\boldsymbol{\tau}_{opt}(0)$ is applied to the system to produce the state $\boldsymbol{x}(1)$. And, the initial state is updated by $\boldsymbol{x}(1)$ for the new optimal control problem at the sampling time $k = 1$. Then, the above-described optimization process is repeated with the concept of receding horizon to obtain a new optimal control sequence for the current system. Subsequently, the new initial states can be computed for the next optimal process. Therefore, NMPC is considered as a receding horizon iterative optimal control algorithm.

The cost function considered in the optimal control problem of the NMPC is given by

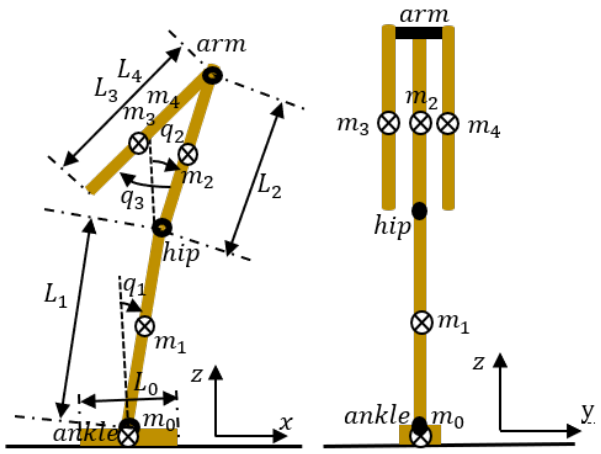


Figure 3.10: Structure of the three-joint and five-link model. m_4, m_3, m_2, m_1, and m_0 represent the masses of the left arm, right arm, upper body, lower body, and foot, respectively. L_4, L_3, L_2, L_1, and L_0 represent the lengths of the left arm, right arm, upper body, lower body, and foot, respectively. q_3, q_2, and q_1 represent the left-right arm angle, hip angle, and ankle angle, respectively. Here, the right arm and left arm share the same joint motor.

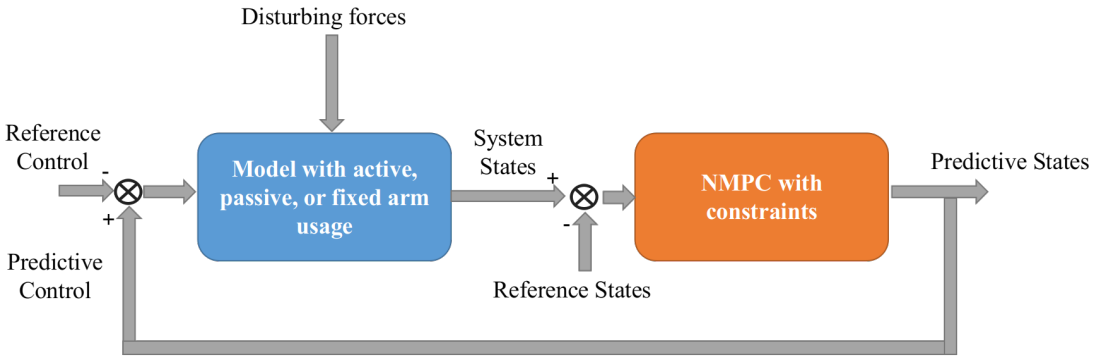


Figure 3.11: The control loop of NMPC for the implementation of reproducing arm strategy for balance recovery in simulation.

$$J\left(\boldsymbol{x}(0),\boldsymbol{\tau}_{(0,N-1)}\right)=\sum_{0}^{N-1} l(\boldsymbol{x},k,\boldsymbol{\tau})+V_f, \tag{3.3.3}$$

$$l(\boldsymbol{x},k,\boldsymbol{\tau})=\frac{1}{2}\left(\boldsymbol{x}^T(k)\boldsymbol{Q}\boldsymbol{x}(k)+\boldsymbol{\tau}^T(k)\boldsymbol{R}\boldsymbol{\tau}(k)\right), \tag{3.3.4}$$

$$V_f=\frac{1}{2}\boldsymbol{x}^T(N)Q_f\boldsymbol{x}(N). \tag{3.3.5}$$

The penalty weighting dimension and constraints of the NMPC differ for the model with the three different arm states, including active arms, passive arms, and fixed arms.

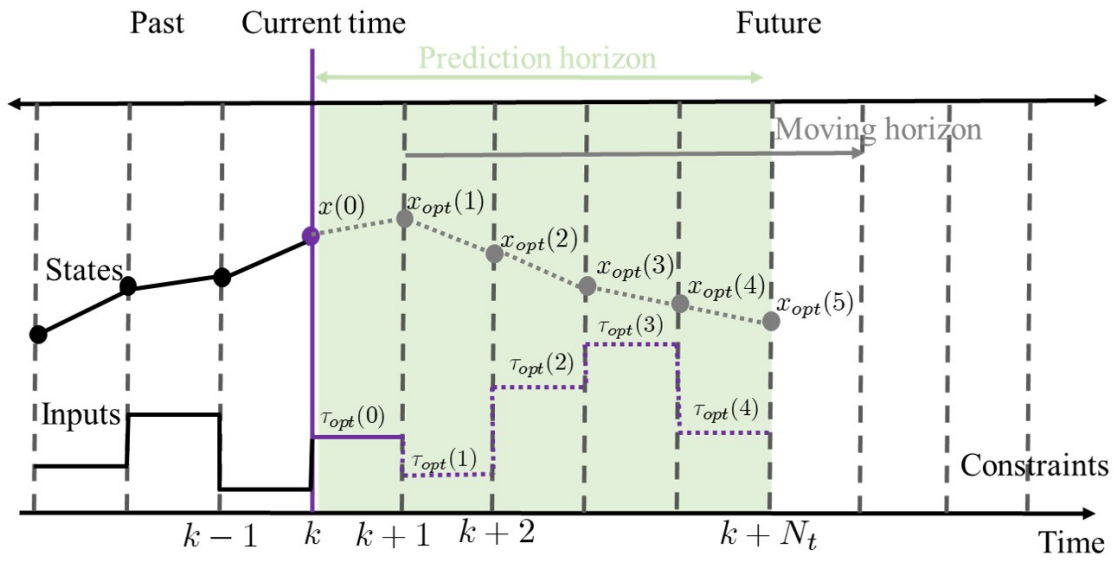


Figure 3.12: Schematic description of the NMPC at time k. The proposed dynamic model is recruited to predict the future motion state sequence $\boldsymbol{x}_{opt}$ of the model system and compute optimal control input sequence $\boldsymbol{\tau}_{opt}$ of balance recovery based on the current state through solving an optimization problem. For instance, let NMPC starts at $k = 0$ with a prediction horizon N_t (here $N_t = 5$) and the initial states $\boldsymbol{x}(0)$.

The cost function (3.3.3) is considered such that $\boldsymbol{Q}$, $\boldsymbol{Q}_f$ and $\boldsymbol{R}$ are positive definite symmetric matrices. The states and the control torques can be penalized by tuning $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively. Increasing $\boldsymbol{Q}$ is aimed to minimize the state tracking error while increasing $\boldsymbol{R}$ means a reduction of energy consumption. In this research, the ratio between $\boldsymbol{Q}$ and $\boldsymbol{R}$ for three cases of arm usages is set as the same value 10^3 named one optimization criterion [23]. Further, terminal weighting $\boldsymbol{Q}_f = 10^5$ can be used as a tuning parameter to penalize the terminal states to achieve stable NMPC performance.

The objective is to minimize the cost $J\left(\boldsymbol{x}(0), \boldsymbol{\tau}_{(0,N_t-1)}\right)$ subject to the following control input and state boundary for the model with three different arm strategies:

1) NMPC for Model with Active Arms

For $i_a = 1,2,3$, which represent ankle, hip, and arm joints respectively, and $k_a = 0,\ldots,N_t - 1$, boundary settings of the control inputs have been selected based on the work of [23] where a constrained-based optimization is proposed for a multi-balance recovery strategy:

$$\tau_{min}(i_a) \leq \tau_{i_a}(k_a) \leq \tau_{max}(i_a),$$

where $\tau_{min}(1) = -120\ [Nm]$, $\tau_{min}(2) = -500\ [Nm]$, $\tau_{min}(3) = -200\ [Nm]$, $\tau_{max}(1) = 120\ [Nm]$, $\tau_{max}(2) = 500\ [Nm]$, and $\tau_{max}(3) = 200\ [Nm]$.

For all $i_a = 1, ..., 6$ and $k_a = 0, ..., N_t$, the states including angles and angular velocities of ankle, hip and arm joints are bounded by

$$x_{min}(i_a) \leq x_{i_a}(k_a) \leq x_{max}(i_a),$$

where $x_{min}(1) = -0.2\ [rad]$, $x_{min}(2) = -0.35\ [rad]$, $x_{min}(3) = -2.5\ [rad]$, $x_{min}(4) = -\infty\ [rad/s]$, $x_{min}(5) = -\infty\ [rad/s]$, $x_{min}(6) = -\infty\ [rad/s]$, $x_{max}(1) = 0.4\ [rad]$, $x_{max}(2) = 1.3\ [rad]$, $x_{max}(3) = 0.5\ [rad]$, $x_{max}(4) = \infty\ [rad/s]$, and $x_{max}(5) = \infty\ [rad/s]$, $x_{max}(6) = \infty\ [rad/s]$. It is necessary to point out that the three first elements of x denote joint angles, and the three last elements represent angular velocities; this is why the unit changes from $[rad]$ to $[rad/s]$. We just put negative infinity in boundary settings for implementation purposes to keep a wide range of velocity values. However, based on the obtained results, the evolution of the velocities remains very reasonable, i.e. within the interval $[-1.2,\ 1.2]$ as it can be observed from Figures 3.26, 3.27, 3.28.

2) NMPC for Model with Passive Arms

For all $i_p = 1, 2$ representing the notation of ankle and hip joints respectively and $k_p = 0, ..., N_t - 1$, the control inputs are bounded by

$$\tau_{min}(i_p) \leq \tau_{i_p}(k_p) \leq \tau_{max}(i_p),$$

where $\tau_{min}(1) = -120\ [Nm]$, $\tau_{min}(2) = -500\ [Nm]$, $\tau_{max}(1) = 120\ [Nm]$, and $\tau_{max}(2) = 500\ [Nm]$.

For all $i_p = 1, ..., 6$ representing joint angles and angular velocities of ankle and hip, arm and prediction horizon $k_p = 0, ..., N_t$, the states are bounded by the same constraint settings as the case with active arms.

3) NMPC for Model with Fixed Arms

For all $i_f = 1, 2$ representing ankle and hip joints respectively and $k_f = 0, ..., N_t - 1$, the control inputs are bounded by the same constraint settings as the case with passive arm.

For all $i_f = 1, ..., 4$ representing joint angles and angular velocities of ankle and hip, and prediction horizon $k_f = 0, ..., N_t$, the system states are bounded by

$$x_{min}(i_f) \leq x_{i_f}(k_f) \leq x_{max}(i_f),$$

where $x_{min}(1) = -0.2\ [rad]$, $x_{min}(2) = -0.35\ [rad]$, $x_{min}(3) = -\infty\ [rad/s]$, $x_{min}(4) = -\infty\ [rad/s]$, $x_{max}(1) = 0.4\ [rad]$, $x_{max}(2) = 1.3\ [rad]$, $x_{max}(3) = \infty\ [rad/s]$, and $x_{max}(4) = \infty\ [rad/s]$.

With the system states and the input constraints, an NMPC is proposed from a neuro-science perspective to reproduce human-like balanced behavior evoked by the human central nervous system. The proposed NMPC also has a predictive aspect that allows predicting the future behavior and computes an optimal control balance strategy by minimizing systemic energy consumption of the whole body. Furthermore, the NMPC technique can handle simultaneously the state and input constraints, which is important to meet realistic requirements due to physical limitations of the human body such as joint ranges and torques satuation. All the previously proposed control techniques can not take into account constraints naturally. In this research, we proposed NMPC which can naturally take into account constraints. Different magnitudes of disturbing forces are applied to the model to observe the autonomous switch between the ankle, hip, and arm strategies and to examine the robustness of the proposed solution.

3.4 Results of Simulation and Discussion Compared to Human Experiments

3.4.1 Simulation Results and Discussion

The preparation and procedure of this work is illustrated in Figure 3.13. In this section, we analyze the model motion intensity using the total RMS deviation of the joint angles to verify the effectiveness of the arm strategy. The simulation settings are demonstrated in Figure 3.14. We pushed the position of the center of mass of the upper body with different disturbing forces backward and forward for 1 $[s]$which could be different with the previous

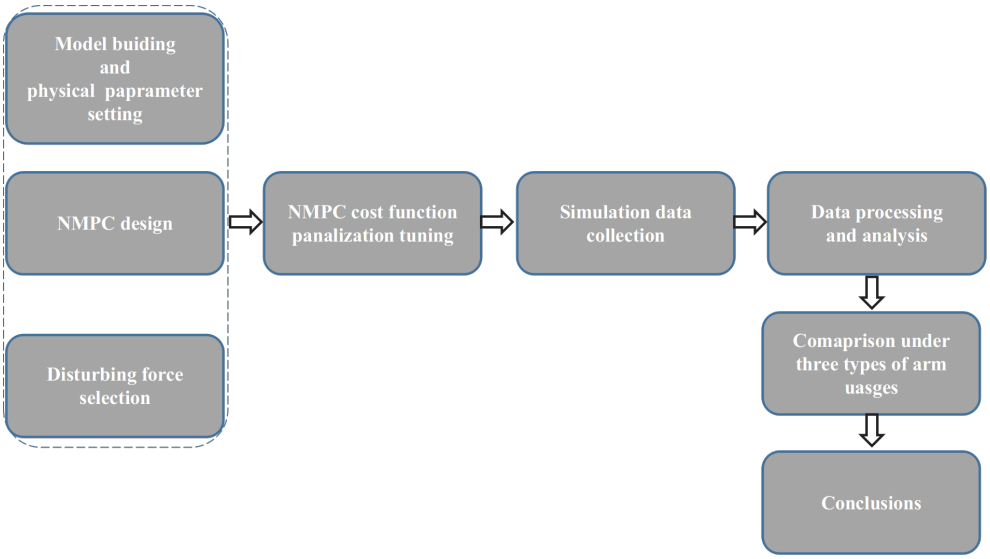


Figure 3.13: The preparation and procedure of the simulation parts.

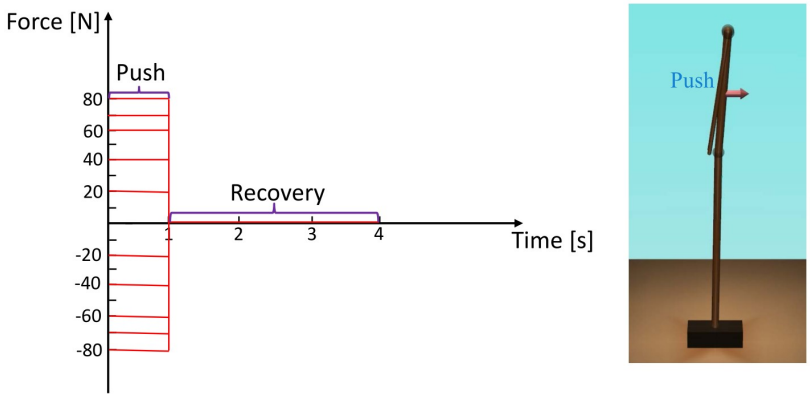


Figure 3.14: The common simulation settings for the three different arm states.

study on perturbation setting with a balance board [88], [89]. The maximum simulation time is set as 4 $[s]$ that can make the model finish the process of balance recovery. The disturbing forces were as follows [23]:

1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, -70 $[N]$, and -80 $[N]$.

2) No force: 0 $[N]$.

3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$.

NMPC controller produces predictive ankle-hip strategy after a perturbation while arm strategy can be employed only for the model of active arm setting. For the disturbing forces -80 $[N]$ and 80$[N]$, only the model with the active arm can realize balance recovery from the unstable states. The models with passive arm and fixed arm are unable to obtain a

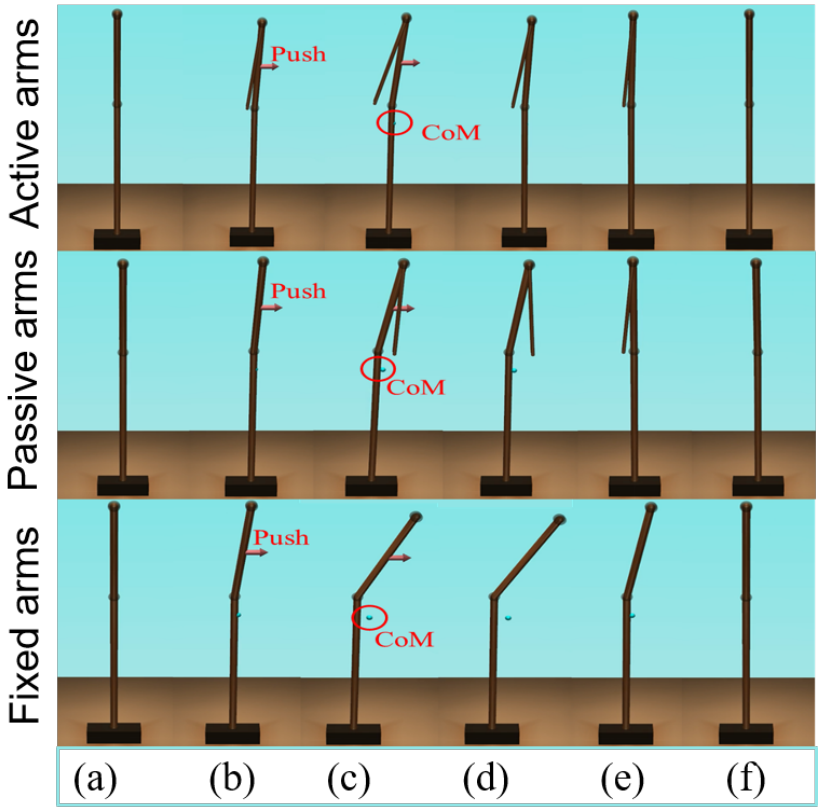


Figure 3.15: Schematic of the balance behavior for the three different arm states ('top to bottom: active arms, passive arms, and fixed arms'), for a disturbing force of 70 $[N]$. (a) and (f) represent the equilibrium states; (b) represents the pushing forward process; (c)-(e) represents the balance recovery behavior.

solution for balance control under the same disturbing force. This indicates that the active arm rotation strategy widens the range of the disturbing forces; this result is similar to the conclusions derived in [48], [90].

The schematic of the movements of the models with active, passive, and fixed arms for a disturbing force of 70 $[N]$ is illustrated in Figure 3.15. The figure shows that the model with active arms has a better ability to realize balance recovery than the other two models. This is because the deviation of the center of mass of the model with active arm usage in the x-axis direction (e.g. Figure 3.16) is less than the other with passive and fixed arm usages. Figure 3.17 shows the center of mass for three different arm states is located within the stable region according to the evolution of the whole body center of mass (CoM) velocity versus its position. Based on the obtained results from this Figure, we concluded that it is located within the stable region. This indicates that active arm usage can maintain the center of mass of the body to remain close to the origin (equilibrium point). From

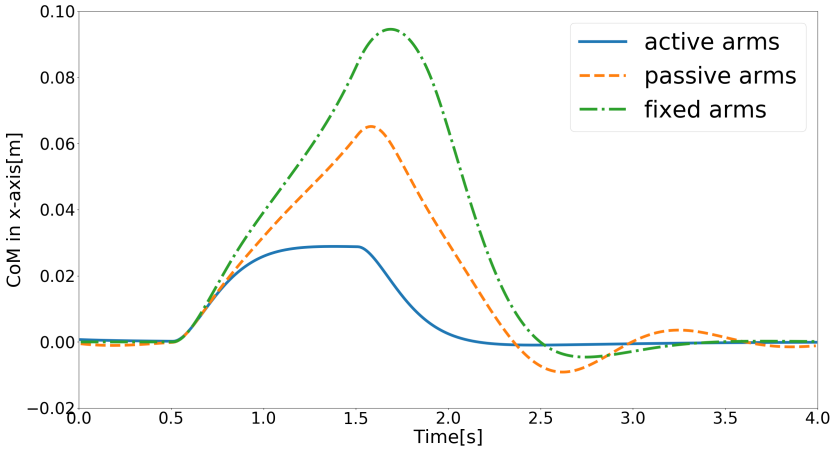


Figure 3.16: Evolution of center of mass of the model in the x-axis direction for the three different arm states (active arms, passive arms, and fixed arms) for a disturbing force of 70 $[N]$.

the stability aspect, active arm usage shows more advantage in balance recovery tasks by comparing the deviation of the center of mass. Even though the ankle angles almost show the same performance for the different disturbing forces with three types of arm usages in two directions in Figure 3.18, the active arm usage reduces the effort of the ankle joint and helps the model recover balance in the fast speed, which saves the recovery time and improves the robustness of balance recovery movements. Similarly, the ankle torque shows the least time effort with active arm usage demonstrated in Figure 3.21 compared with the other two types of arm usages. It is worth noting that the ankle torque shows the different capacity recruitment w.r.t the different disturbing forces in general.

The hip angle exhibits the least performance from the recovery time and the maximum deviation of joint angles for the model with active arm usages shown in Figure 3.19. And for the disturbing force $-70[N]$, Only the model with fixed arm usages reaches the hip angle constraint boundary. Besides, the hip angle with fixed arm usages makes the maximum effort for all different disturbing forces due to its limited ankle efforts without arm rotation usages. As the disturbing force increases in Figure 3.22, more hip torque is used. This indicates the more energy consumption of hip joints might be applied to compensate the other joint limited efforts for balance recovery motion.

What's more, it is worth noting that the active arm joint angle converges asymptotically while the passive arm joint angle oscillates a lot in Figure 3.20. This indicates active arm usages might make the balance maintenance more robust than the case with passive arm

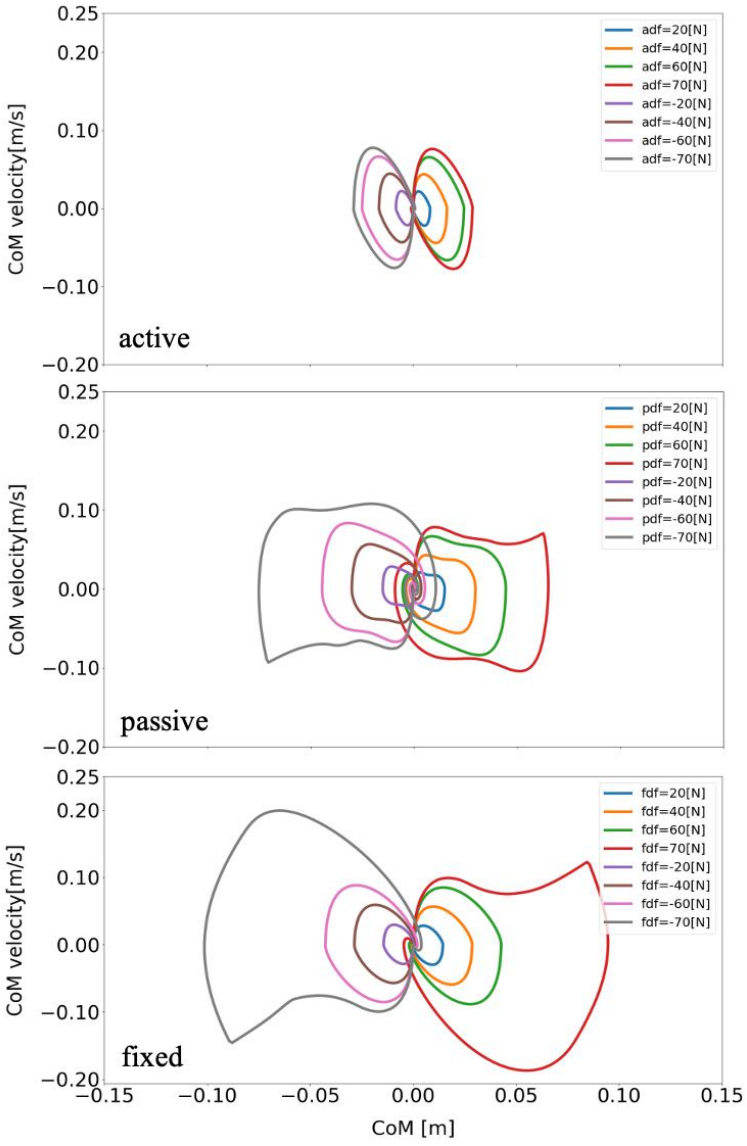


Figure 3.17: Evolution of the center of mass phase portrait of the model for three different arm states (active arm, passive arm, fixed arm) for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, and -70 $[N]$. 2) No force: 0 $[N]$. 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$. "a," "p," and "f" in the labels "adf," "pdf," and "fdf" represent the cases with active arms, passive arms, fixed arms, respectively, and "df" represents the disturbing forces.

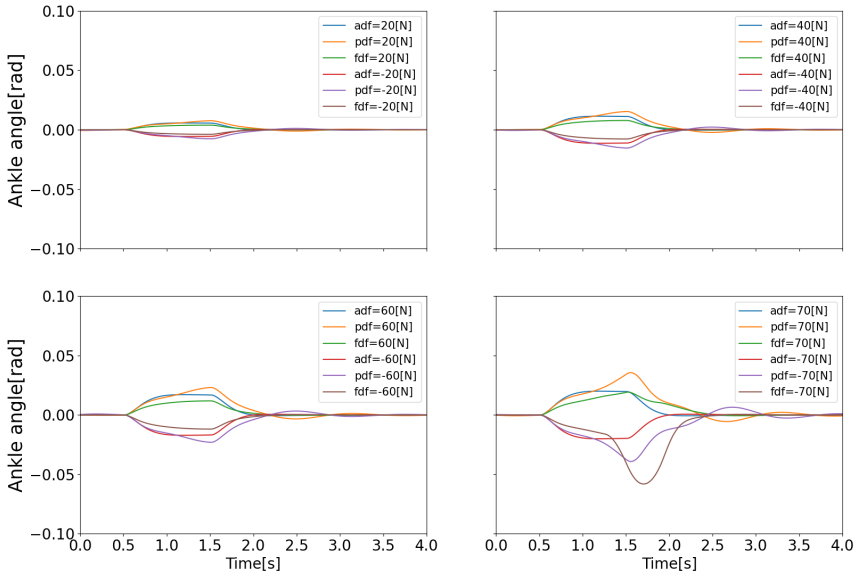


Figure 3.18: Evolution of ankle angles for the three different arm states (active arms, passive arms, and fixed arms) for different disturbing forces.

usages. And as the disturbing force increases, the active arm joint torque increase as well in Figure 3.23. Thus, the more active arm joint energy consumption is applied to help balance recovery control.

The total RMS deviation can be calculated by

$$Total\ RMS\ deviation = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(q_1(t)^2 + q_2(t)^2)}\ ,$$

where, N denotes the total samples number, which can be computed from the recovery time and the sampling period, $q_1(t)$ and $q_2(t)$ represent the ankle and hip angles at each sampling point, respectively.

The evolution of the total RMS deviation of the model for the three different arm states (active arm, passive arm, and fixed arm) for different disturbing forces is illustrated in Figure 3.24. Here, the total RMS deviation is defined to represent the body motion intensity. Figure 3.24 shows that the total RMS deviation of the balance recovery motion with active arms is less than that with passive arms. Furthermore, the total RMS deviation of the balance recovery motion with passive arms is less than that with fixed arms for the following disturbing forces: -20 $[N]$, -40 $[N]$, -60 $[N]$, -70 $[N]$, 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$. This indicates that arm movements contribute to human body balance control

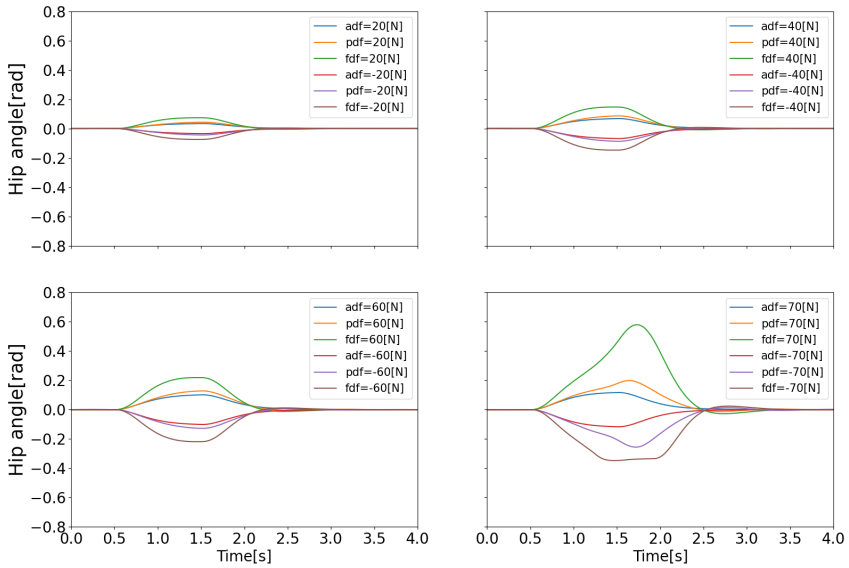


Figure 3.19: Evolution of hip angles for the three different arm states (active arms, passive arms, and fixed arms) for different disturbing forces.

Table 3.3: Contributions of active arm usage to balance recovery under different ankle capacities. Five cases are defined as follows:(1) tc = [-80, 80] [Nm], df = 56 [N]; (2) tc = [-100, 100] [Nm], df = 56 [N]; (3) tc = [-120, 120] [Nm], df = 56 [N]; (4) tc = [-100, 100] [Nm], df = 70 [N]; (5) tc = [-120, 120] [Nm], df = 70 [N]. AEC represents arm energy consumption, and TEC represents total energy consumption.

Case	Recovery time [s]	Arm RMS deviation	AEC [J]	TEC [J]
(1)	5.53	0.505	31.681	74.284
(2)	3.2	0.071	0.81	3.11
(3)	3.2	0.071	0.796	3.097
(4)	5.31	0.523	29.914	85.602
(5)	3.32	0.087	1.248	4.824

and reduce the motion intensity of the hip joint. This conclusion in accordance with the one obtained from a human experiment [51]. Besides, it is worth to note that the proposed model based on NMPC can recover after a wide range of perturbations; therefore, the robustness of the NMPC is verified as well. This is one of the advantages of the proposed controller with active arm usage.

The joint energy is calculated as follows, where, τ is the joint input torque matrix of ankle,

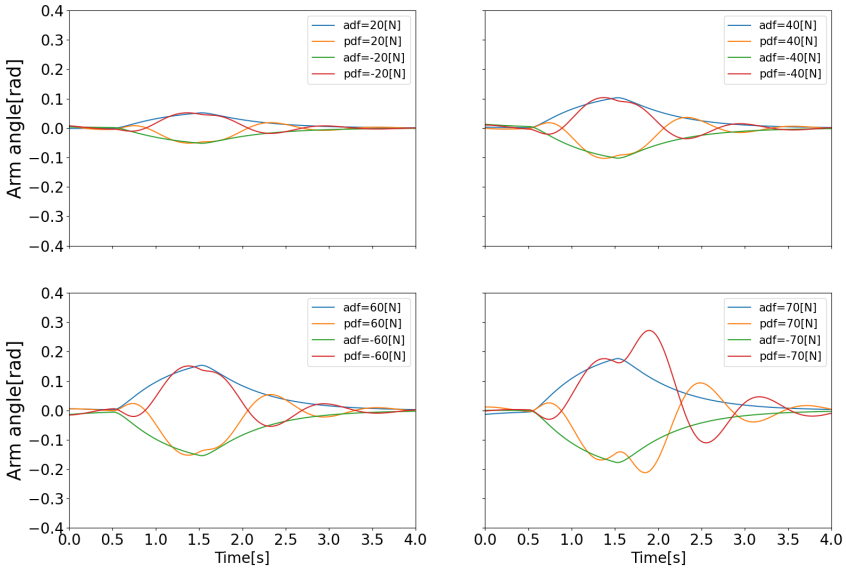


Figure 3.20: Evolution of arm angles for the three different arm states (active arms, passive arms, and fixed arms) for different disturbing forces.

Table 3.4: Joint correlation of balance recovery motion for the model with active arm usage.

Disturbing forces [N]	Ankle-hip correlation	Hip-arm correlation
-70	0.900	0.967
-60	0.900	0.970
-40	0.896	0.962
-20	0.896	0.964
20	0.896	0.966
40	0.898	0.968
60	0.899	0.968
70	0.900	0.964
Mean joint correlation	0.898	0.966

hip and arm, and $\dot{q}$ is their corresponding angular velocity matrix, t is time.

$$W = \int \tau \dot{q} dt. \tag{3.4.1}$$

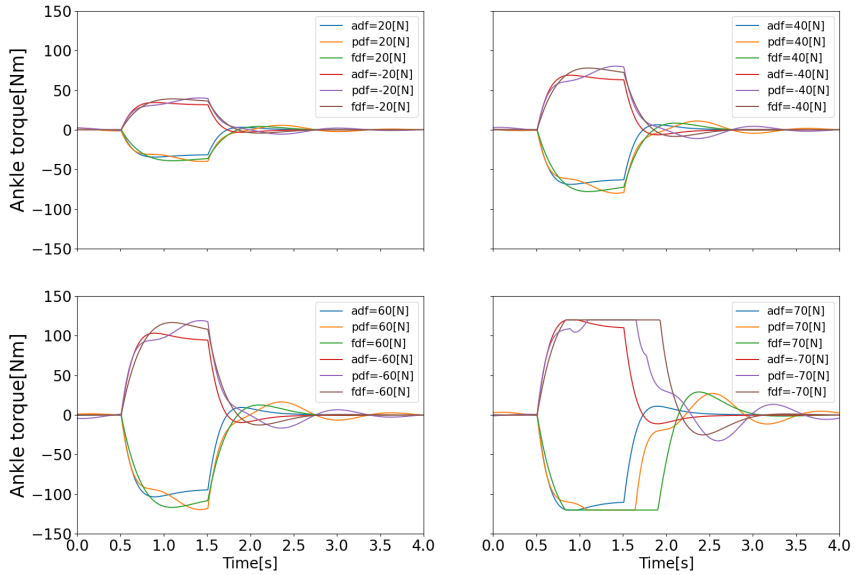


Figure 3.21: Evolution of ankle torques for the three different arm states (active arms, passive arms, and fixed arms) for different disturbing forces.

Figure 3.25 compares the energy consumption of the model for three different arm states and for different disturbing forces. The energy consumption in this research is metabolic energy, which can be computed through the total joint actuator energy consumption of ankle, hip, arms in Eq. 3.4.1. First, we observe that as the disturbing force intensifies, the balance recovery motion consumes more energy for each case. Most importantly, for the same amount of push, the energy consumption for the balance recovery of the model with active arm rotation is the least followed by passive arm rotation. It is the biggest in the case without arm rotation. This indicates clearly that balance recovery with arm strategy can reduce energy consumption, which is human-like energy-efficient. Humans also optimize the motion behavior for balance recovery to save energy. Thus, the contribution of arm usage to human balance recovery can also be acknowledged from the perspective of energy-efficiency.

Furthermore, there are consistent limit cycles of the balance recovery for the model with active arm usage over the different disturbing forces, indicating natural temporal regulation on the coordination of ankle, hip, and arm joints, respectively, in Figures 3.26 , 3.27, and 3.28. It means that there is temporal pattern to make compensation against disturbing forces, which can be viewed as there is control strategy since it forms similar form of phase portrait. It implies there is consistent ankle-hip-arm control strategy for active arm usage. However, it is noticeable that portrait form is largely deformed for the passive and fixed

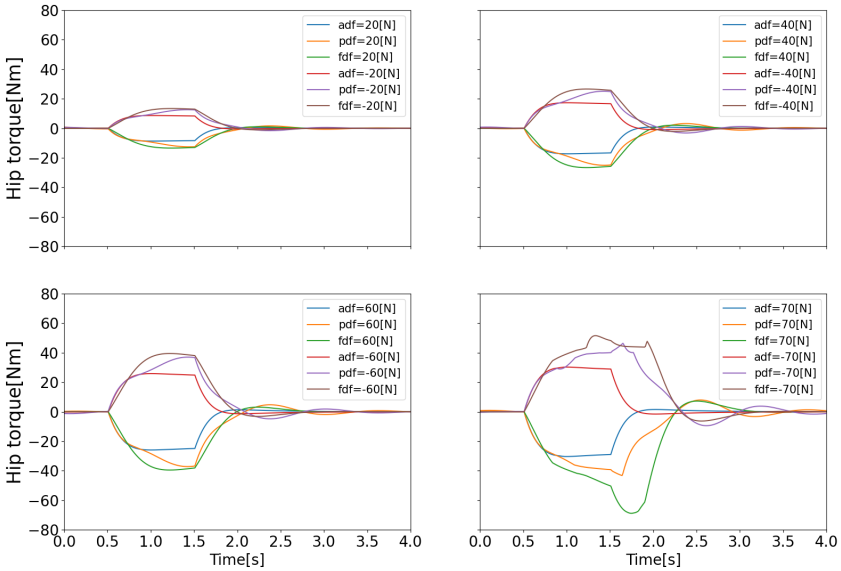


Figure 3.22: Evolution of hip torques for the three different arm states (active arms, passive arms, and fixed arms) for different disturbing forces.

arm cases.

Similarly, the relationship of the ankle, hip, arm angles shows aligned spatial pattern over the joints, which represents synergetic joint coordination, and the maximum deviations exhibit linear approximations in Figure 3.29. The joint correlation of neighboring joints, such as ankle and hip, hip and arm, under the different disturbing forces are computed for synergy existence confirmation [91]. The mean joint correlation in 3.4 between ankle and hip is 0.898 and the one between hip and arm is 0.966. Thus, The balance motion with active arm rotation is highly coordinated, which means there is a good synergy performance. The synergy pattern here can represent the ability of task sharing and balance stabilization. However, the motions of balance recovery for the cases of the model with passive and fixed arms did not exhibit similar synergy performance because of the absence of certain patterns. From the synergy analysis perspective, the balance recovery for the model with active arm usage is better than that for the other two cases.

Table 3.3 shows the contribution of active arm usage to the balance recovery under different ankle capacities. Since ankle is most common injured body site [92], we want to observe how arm usages improve the ability of balance maintenance. Here, five cases are considered through different ankle torque constraints (tc) and disturbing forces (df) [93]:
(1) tc $= [-80,\ 80]\ [Nm]$, df $=\ 56\ [N]$,

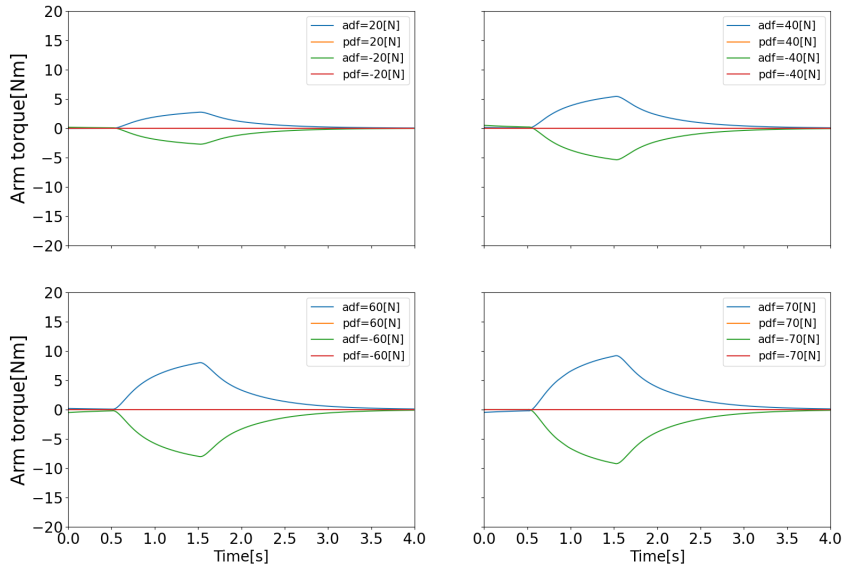


Figure 3.23: Evolution of arm torques for the three different arm states (active arms, passive arms, and fixed arms) for different disturbing forces.

Table 3.5: Mean of the peak-to-peak value of the joint angles (Unit: $[degree]$) and torques (Unit: $[Nm]$) of the ankle, hip, and arm for different pushing forces.

Push magnitude	Ankle		Hip		Arm	
	Angle	Torque	Angle	Torque	Angle	Torque
Small Push	2.35	33.70	14.10	24.03	12.45	4.35
Middle Push	4.58	37.82	24.53	47.19	19.54	6.95
Large Push	5.51	46.60	44.91	61.80	24.86	8.32

5

(2) tc $= [-100,\ 100]\ [Nm]$, df $=\ 56\ [N]$,

(3) tc $= [-120,\ 120]\ [Nm]$, df $=\ 56\ [N]$,

(4) tc $= [-100,\ 100]\ [Nm]$, df $=\ 70\ [N]$,

(5)tc $= [-120,\ 120]\ [Nm]$, df $=\ 70\ [N]$.

Comparing the above ankle boundary constraint settings, we can note that the ankle capacity of case (1) is weaker than that of cases (2) and (3). For the same disturbing force df $=\ 56\ [N]$ applied on the center of mass of the upper body, in cases (1), (2), (3), the ankle capacity of case (1) reaches the maximum limit. Therefore, this model needs more efforts for balance recovery. Besides, a longer recovery time and a higher energy consumption are required for case (1), compared with those for cases (2) and (3). For a limited ankle capacity, such as in case (1), the active arm RMS deviation is seven times those in cases (2) and (3), and arm energy consumption is approximately 39 times that in

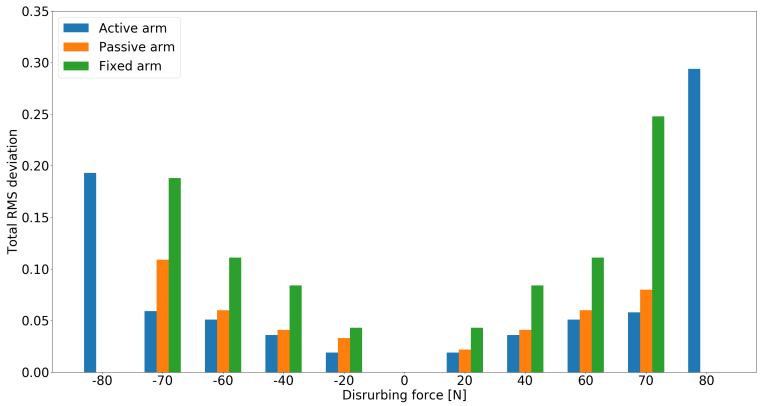


Figure 3.24: Evolution of the total RMS deviation of the model for three different arm states (active arm, passive arm, and fixed arm) for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, -70 $[N]$, and -80 $[N]$; 2) No force: 0 $[N]$; 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. There was no solution for the cases of passive and fixed arms under the disturbing forces -80 $[N]$ and 80 $[N]$.

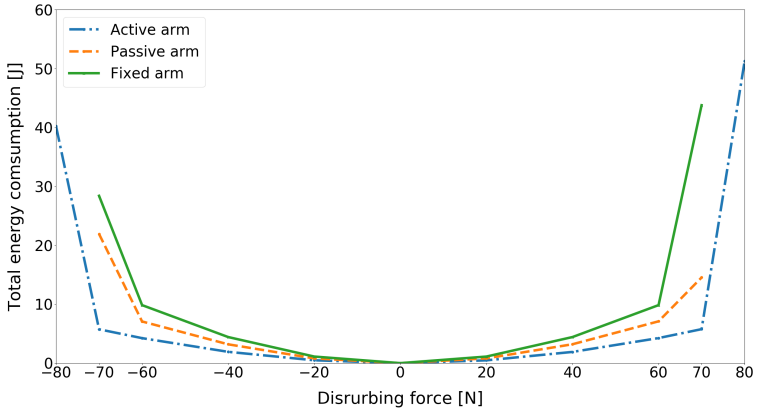


Figure 3.25: Comparison of total energy consumption of the model for three different arm states (active arm, passive arm, fixed arm) for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, -70 $[N]$, and -80 $[N]$; 2) No force: 0 $[N]$; 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$.

cases (2) and (3). Similarly, for a same disturbing force of 70 $[N]$, the active arm RMS deviation in case (4) is six times that in cases (2) and (3), and the arm energy consumption is approximately 24 times that in cases (2) and (3). These observations show that for a limited ankle capacity, arm rotation makes more effort for balance recovery. Furthermore, the ankle capacity in cases (2) and (3) for a disturbing force df $=$ 56 $[N]$ does not reach the maximum limit, and the movements of balance recovery are almost the same. For the same ankle capacity in cases (2) and (4), the disturbing force df $=$ 70 $[N]$ in case (4) makes the ankle capacity reaching the maximum limit, and active arms need more effort for balance recovery than that in case (2). By comparing cases (3) and (5), although the ankle capacity does not reach the maximum limit in both cases, more efforts are required

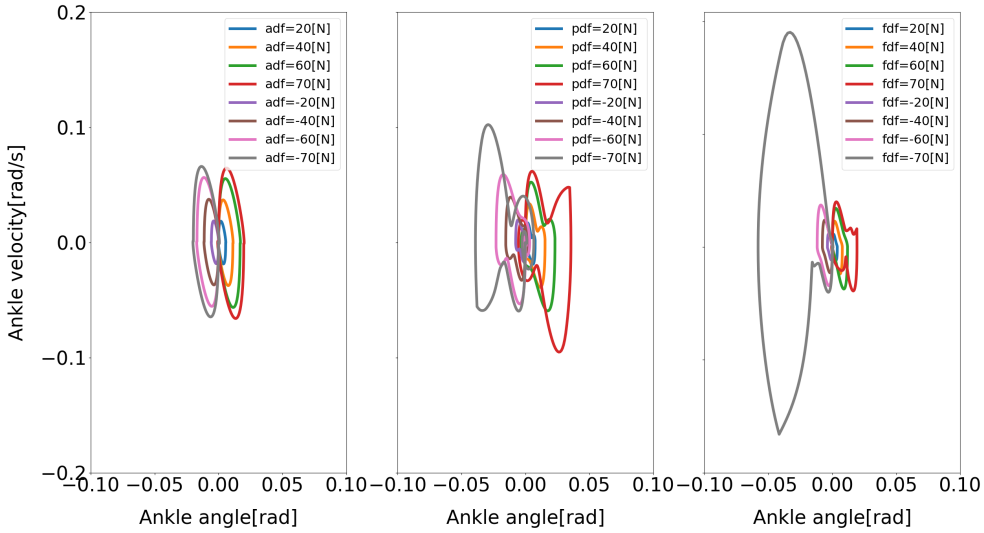


Figure 3.26: Evolution of the ankle phase portrait of the model for three different arm states (active arm, passive arm, fixed arm) for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, and -70 $[N]$; 2) No force: 0 $[N]$; 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$. "a," "p," and "f" in the labels "adf," "pdf," and "fdf" represent the cases with active arms, passive arms, fixed arms, respectively, and "df" represents the disturbing forces.

for a bigger disturbing force.

In the simulation study, various indexes are verified to evaluate the capability of balance recovery. The obtained data indicate that balance recovery with active arms is the most effective strategy, and balance control with arm usage is better than that without arm usage. Besides, Phase portraits of joint angles are considered to analyze the control pattern of balance recovery motion. Furthermore, Ankle torque boundary constraints are set with different values. The relationship between ankle capacity and active arm usage is discussed since in our daily life ankle is easy to be injured, we want to observe how arm usages contribute to balance in this case. Regarding the comparison of our work with previous studies dealing mainly with human balance control without arm strategy, it is worth to point out that in our study we considered three cases including (i) active, (ii) passive and (iii) fixed arms. This last one corresponds to the case without arm strategy from the literature. Indeed, the obtained results show clearly that the balance model with active arm strategy leads to a less energy consumption, a more robust control, a more synergetic motion, and an improved balance ability, compared to the case without arm strategy.

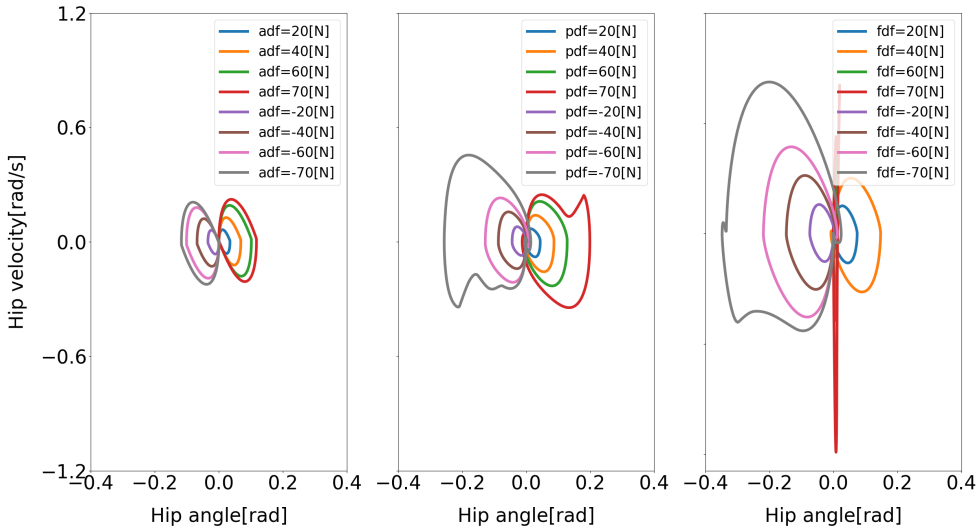


Figure 3.27: Evolution of the hip phase portrait of the model for three different arm states (active arm, passive arm, fixed arm) for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, and -70 $[N]$. 2) No force: 0 $[N]$. 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$. "a," "p," and "f" in the labels "adf," "pdf," and "fdf" represent the cases with active arms, passive arms, fixed arms, respectively, and "df" represents the disturbing forces.

3.4.2 Comparison with Human Experimental Results

Now, we apply three different magnitudes of a pushing force, namely, small push, medium push, and large push, to the backs of the subjects to observe the contributions of arm rotation to human quiet standing balance recovery. The magnitudes of the pushing forces are distinguished by the maximum position deviation of the marker on the subject's neck, and ground reaction force measured by two AMTI force plates. Furthermore, it is important to point out that even though the same pushing force is applied to all the subjects, there is no guarantee that the balance behaviour of the subjects would be exactly the same. Consequently, we decided to quantify the levels of this pushing force and classify them into three levels (small, medium and large). The key point behind this is to distinguish the subjects' balance recovery behavior based on these different levels of the pushing force. The subjects were 5 healthy men(mean age (25 $\pm$ 5) years, mean height (175 $\pm$ 10) cm, mean weight (70 $\pm$ 10) kg) without any known motor or neurological impairment. The protocols of human experiments were designed according to the Declaration of Helsinki and approved by the Tohoku University ethics committee. In fact, the human experiments have been conducted in two main stages. During the first one, dealing with a pre-training, the subjects are pushed with different forces (according to the three levels explained above,

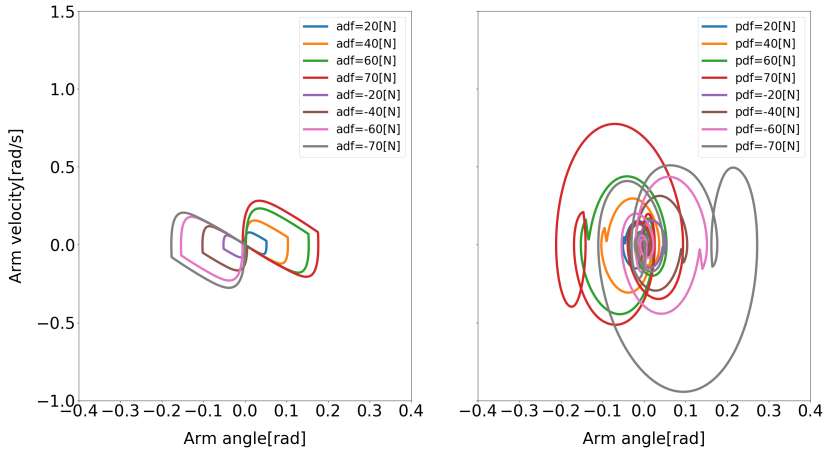


Figure 3.28: Evolution of the arm phase portrait of the model for two different arm states (active arm, passive arm) for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, and -70 $[N]$. 2) No force: 0 $[N]$. 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$. "a" and "p" in the labels "adf" and "pdf" represent the cases with active arms and passive arms, respectively, and "df" represents the disturbing forces.

respecting the order : small, then medium, then large for security purposes) to learn how to maintain their balance. During the second stage, dealing with the final experimental tests, the previous different pushing force levels are considered, while disturbing the subjects in stand-up positions, and their behaviour data are recorded. The exact spot of the push force is the upper back of the subject. For each level push force, five repetitions are performed for a single subject. The motion of the subject is tracked using 42 markers in the Optitrack system with eight cameras and the ground reaction forces are measured using two force-plates. Then, we export the tracking data of the motion and ground reaction forces and convert them to a standard data format, which can be used in OpenSim [86]. Then, we obtain the joint angles and torques for each subject through model scaling, inverse kinematics and dynamics in OpenSim. Here, the inverse dynamics could be solved by using the top-down method. These results can be used to analyze the balance recovery motion and the functions of the ankle, hip, and arm for different magnitudes of disturbing forces.

Our discussion in this paragraph focuses on the representative movements on the Subject 1 since the trends discussed for this subject are consistent across all the subjects. The balance recovery motion of the representative subject 1 for a large pushing force is shown in Figure 3.30. Here, the active arm usage of subject 1 is in a good accordance with the one obtained in our simulation shown in Figure 3.15 reproduced by the proposed NMPC,

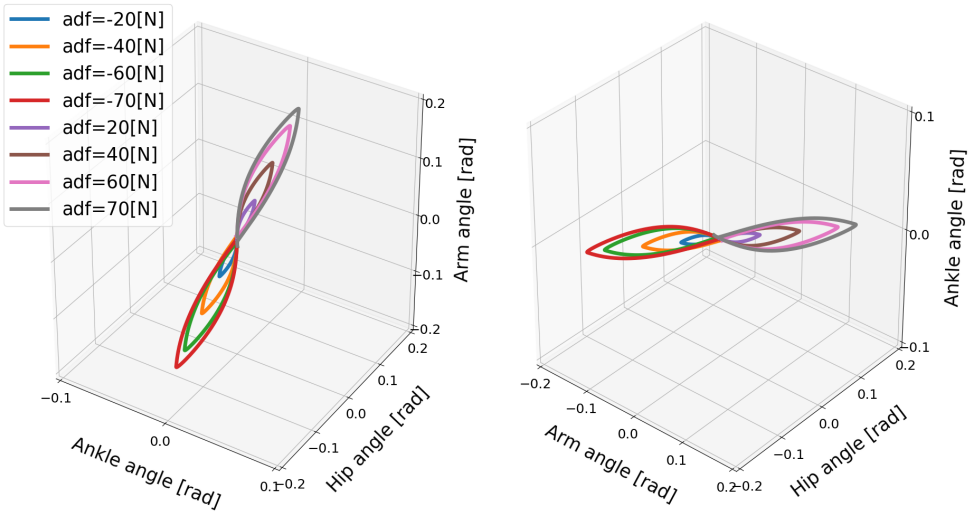
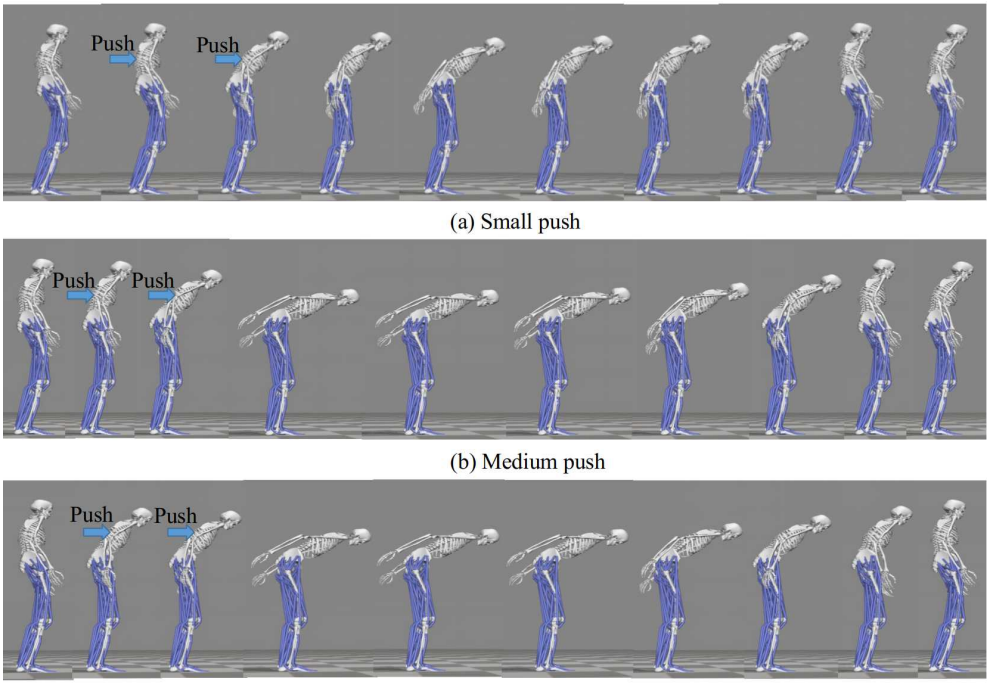


Figure 3.29: Evolution of the ankle, hip, and arm angles of the model for active arms for different disturbing forces: 1) Push backward: -20 $[N]$, -40 $[N]$, -60 $[N]$, -70 $[N]$. 2) No force: 0 $[N]$. 3) Push forward: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$. "a" in the label "adf" represents the case with active arms, and "df" represents the disturbing force. The mean joint correlation between ankle and hip is 0.898 and the one between hip and arm is 0.966.

where anti-phase between arm and hip joint angles. Figure 3.31 shows the evolution of the joint angles and torques of the ankle, hip, and arm for the three different magnitudes of push. The ankle joint angles change slightly, which is similar with the simulation results illustrated in Figure 3.26, because of the structural limitation of the ankle joint compared to other joints. It is important to note that ankle joint angle and torque don't change from middle push to large push. It means ankle usage meets saturation due to its limited capacity. We have observed this phenomenon also in the simulation study. The hip joint rotates by a larger degree as the magnitude of the pushing force increases. This illustrates that hip joints play a major role in balance recovery. Furthermore, the deviation of the arm joint angles and torques increases. This is because when the magnitude of the pushing force increases, subject 1 attempts to recover balance through more efforts of the arm rotation. From Figure 3.31, we note that subject 1 spends a longer time in recovering balance for the large push. Table 3.5 presents the mean of the peak-to-peak values of the joint angles and torques of the ankle, hip, and arm of five subjects for different magnitudes of pushing force. Here, the deviation of the arm joint angles and torques is positively correlated with the magnitude of push force. For large push, we can notice that ankle joint angle increases only 1 degree from middle push case. It implies that ankle usage is already near the saturation due to mechanical constraints, thus the arm strategy to compensate disturbance is essential for large push. This process is consistent to the behavior we have



(a) Small push

(b) Medium push

(c) Large push

Figure 3.30: Balance recovery motion of Subject 1 for large push, illustrated through the human motions at different time instant.

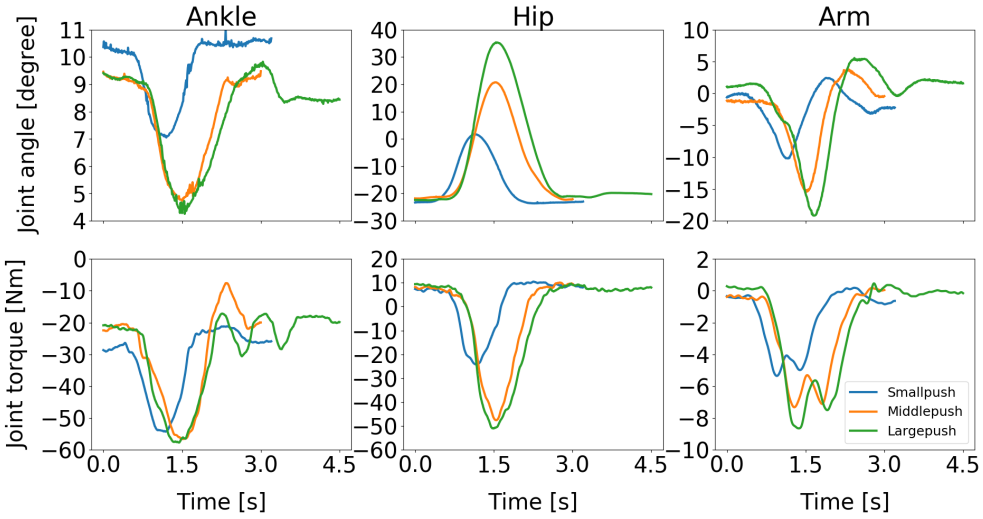


Figure 3.31: Evolution of the joint angles and torques of the ankle, hip, and arm for three different magnitudes of push: 1) Small push, 2) Medium push, and 3) Large push. It is important to note that ankle joint angle and torque don't change from middle push to large push. It implies necessity of arm usage after ankle usage saturation.

observed in the proposed NMPC controller. Consequently, we conclude that active arm usage contributes to balance recovery in human experiment and the consistent behavior between the predictive controller study and human experiment.

3.5 Conclusions

In this study, we built a simplified human model with arms and proposed an NMPC scheme to reproduce human balance behavior with arm usages. Three-arm states, active arms, passive arms, and fixed arms, were considered to study the contributions of the arm movements to balance recovery with different magnitudes of a disturbing force during quiet standing. The contribution of arm usage to human balance control was verified by comparing the total RMS deviation of joint angles, and balance control with active arms was found to be the most effective in terms of energy consumption and the disturbance effect minimization. Furthermore, the synergetic motion pattern was observed with kinematics during balance recovery with active arms while it was confirmed with joint correlation along with the steady smooth limit cycle pattern, and the total energy consumption was compared. Finally, the results of human experiments were compared with simulation to verify that active arm usage contributes to balance recovery. This will help us to gain a better understanding of the mechanism of quiet standing balance with arm strategy and to develop an effective balance controller for rehabilitation.

Chapter 4

Postural Synergy Extraction and Analysis

4.1 Synergy Extraction Methods

During the last few decades, many studies have identified motor primitives (i.e. synergies) at electromyographic, kinetic and kinematic levels. Typically, unsupervised learning algorithms have been applied for the identification of such components, including independent component analysis (ICA), principal component analysis (PCA) and non-negative matrix factorization (NMF) etc. The calculation for extracting various types of synergies have been formalized. The basic approach for synergy extraction can be simplified into following steps:

- Acquisition and pre-processing of data during a complex behavior. It can be a recording of EMG activity or joint angles or joint torques etc., depending on whether it is muscle or kinematic or kinetic synergies.
- Use computational factorization analysis like ICA, PCA, or NMF, etc., to extract the set of synergies from the recorded data.
- Evaluate if a combination of these synergies can well represent and reproduce the observed data.

In our study, we have used spatial and spatiotemporal synergies extracted from the joints angles and torques of the simplified human-body models to analyze the mechanism of human balance recovery motion.

4.1.1 Spatial Motor Synergy Computation

The definition of spatial motor synergy has been introduced in [58]. The basic idea is that instantaneous co-variations can be shown by groups of degrees of freedom, indicating adaptive coordination of multiple joints or muscles. This means the hypothesis that the proportion of different degrees of freedom signals should be unchanged over time. This fact is explained by the classical assumption that a set of multiple joints or muscles of the body can be controlled synchronously without specific time-delays by the CNS through reducing the dimension of the DoFs. This can be crucial for saving computational resources of the multiple joint or muscle control, which implies reduced command that simplifies the complexity of body control. The formulation of the spatial motor synergy is as follows:

$$\mathbf{x}_n^r(t) = \sum_{s=1}^{S} \mathbf{w}_s \cdot c_s^r(t) + \text{ residuals} \tag{4.1.1}$$

Where $\mathbf{x}_n^r(t)$ denotes the values of n joint angles (For our simplified body model, $1 \leq n \leq 8$) as source signals at time instant t in trial number. The number of spatial motor synergies is denoted by S. The spatial motor synergy patterns $\mathbf{w}_s$ are extracted as invariant column vectors over trials. The time-dependent mixing weights of the synergies $c_s^r(t)$ and the residuals vary in each trial. It is worth noting that the spatial decomposition indicates the contribution of source signals in space.

4.1.2 Spatiotemporal Motor Synergy Computation

In the previous section, spatial synergy has been introduced as an efficient way to analyze the recorded data features. However, spatial synergy only includes an amplitude parameter, the temporal dependence in the recorded data derives from the activation weights $c_s^r(t)$ in 4.1.3. Thus, spatiotemporal synergy including both an amplitude parameter and a temporal parameter is introduced necessarily to analyze the recorded data from spatial and

time-varying aspects. The spatiotemporal synergy is formulated by

$$\mathbf{x}(t) = \sum_{i=1}^{N} c_i \cdot \mathbf{w}_i(t - t_i) + \text{ residuals } . \tag{4.1.2}$$

Where $\mathbf{x}(t)$ denotes the recorded source signals, $\mathbf{w}_i(t - t_i)$ represents ith synergy after the onset delay t_i. c_i is an activation weight parameter.

The simplified matrix form of Eqs. (4.1.1) or (4.1.2) is derived in Eq. (4.1.3), where the term of *residuals* is omitted. X, W, and C represent the matrix of the source signals, the synergies, and the corresponding activation weights, respectively.

$$X = W \cdot C \tag{4.1.3}$$

As one of the classical decomposition algorithms, principal component analysis (PCA) [94] is used to solve Eq. (4.1.3). The key role of PCA in this study is to minimize the nonlinear least square problem defined as the reconstruction error in Eq. (4.1.4) w. r. t W and C with the Frobenius Norm $\|\cdot\|_F$. All the necessary computation can be implemented in PYTHON using the "*scikit-learn*" library.

$$E^2 = \|X - W \cdot C\|_F^2 \tag{4.1.4}$$

The reconstruction accuracy metric R^2 defined in Eq. (4.1.5) is used to quantify how well source signals can be represented by the inner multiplication of the extracted W and the corresponding C term. It is worth noting that R^2 ranges from zero to one. Besides, $R^2 = 1$ indicates the perfect reconstruction of the source signals.

$$R^2 = 1 - \frac{\|X - W \cdot C\|_F^2}{\|X - \bar{X}\|_F^2} \tag{4.1.5}$$

In this study, the approach to evaluating joint angles is more synergetic through observing the corresponding joint angles can be reconstructed with fewer spatial synergy and the higher reconstruction accuracy R^2.

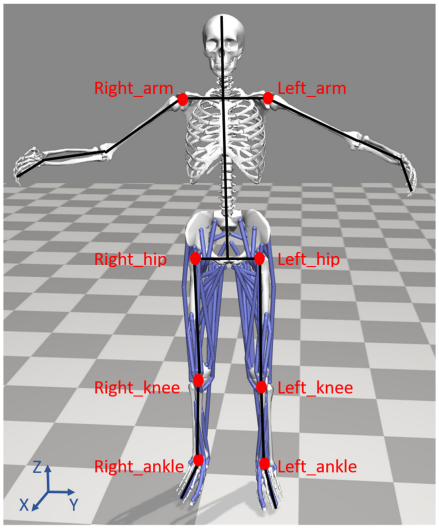


Figure 4.1: Structure of the simplified eight-joint model including the joints of right arm, left arm, right hip, left hip, right knee, left knee, right ankle, left ankle. The proposed model setting is consistent with our protocols of human experiments.

4.2 Motor Synergy Analysis for Human Balance Recovery Strategies

The main contributions of this section are summarized as follows.

- An eight-joint model for human balance recovery is proposed and four-level pushing forces are considered to distinguish different balance strategies.
- Spatial motor synergy is extracted to classify the organization of joints.
- The transition of balance strategies is discussed according to the space of the extracted synergies.

The rest of the section is organized as follows. In subsection 4.2.1, the methods of the human experiments are described. Data processing is shown in subsection 4.2.2 and spatial motor synergy is demonstrated in subsection 4.2.3.

4.2.1 Methods of Human Experiments

The main purpose of the present study is to explore the relationship between balance recovery strategies and the characteristics of motor synergy for a simplified human body

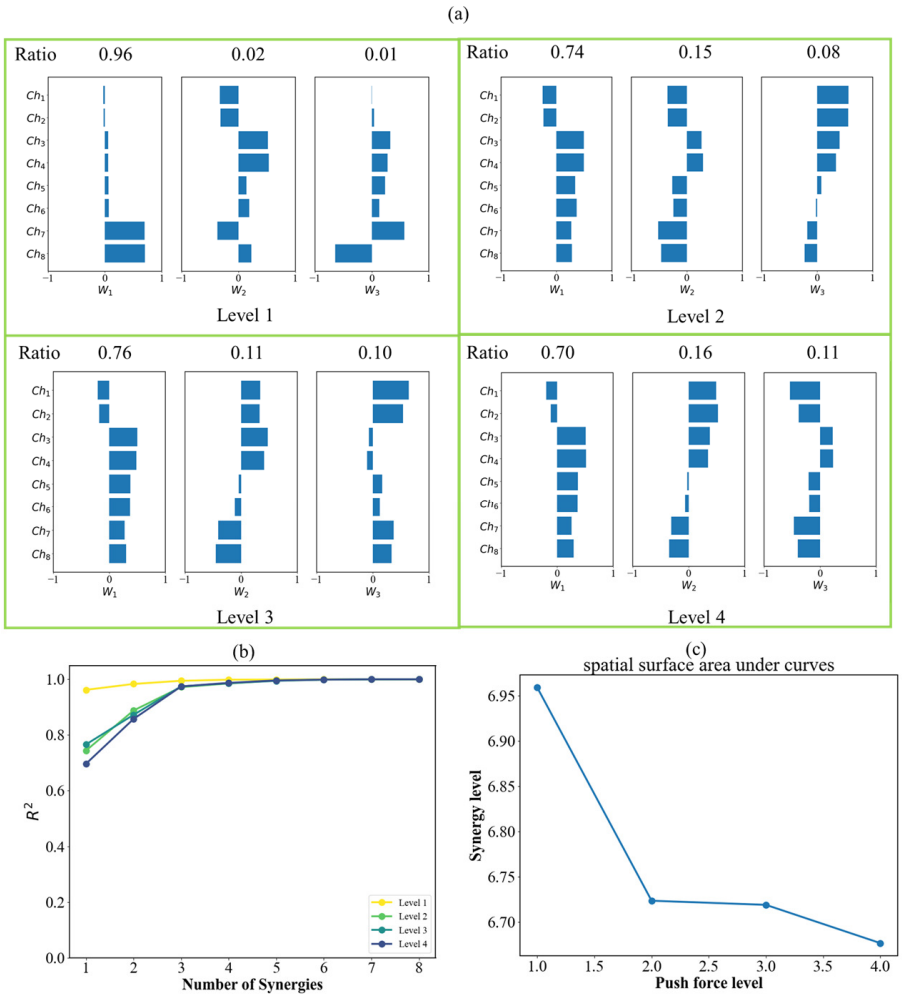


Figure 4.2: (*a*) Comparison of the extracted spatial synergies of joint angles for the four-level pushing forces, Ch_1, Ch_2, Ch_3, Ch_4, Ch_5, Ch_6, Ch_7, and Ch_8 represent the angles of the right arm, left arm, right hip, left hip, right knee, left knee, right ankle and left ankle, respectively. (*b*) Reconstruction accuracy for the four-level pushing forces. (*c*) Synergy level computed by spatial surface area under curves of the reconstruction accuracy R^2.

structure as illustrated in Figure 4.1. The participants were 6 healthy male subjects (mean age (25 $\pm$ 5) years, mean height (175 $\pm$ 10) cm, mean weight (70 $\pm$ 10) kg) without any known motor or neurological impairment. Following the Declaration of Helsinki, the protocols of human experiments were designed for measuring human postural coordination.

We instructed the subjects to stand upright keeping the two feet on the force-plates (AMTI) [95] to measure the ground reaction force (GrF). First, a static pose with 64 makers was captured for model scaling in Opensim [96]. The maker set of the running model in [86] without head makers was used in our experiments. After completing the static pose, the 22 less important markers for inverse kinematics computation were removed to simplify the

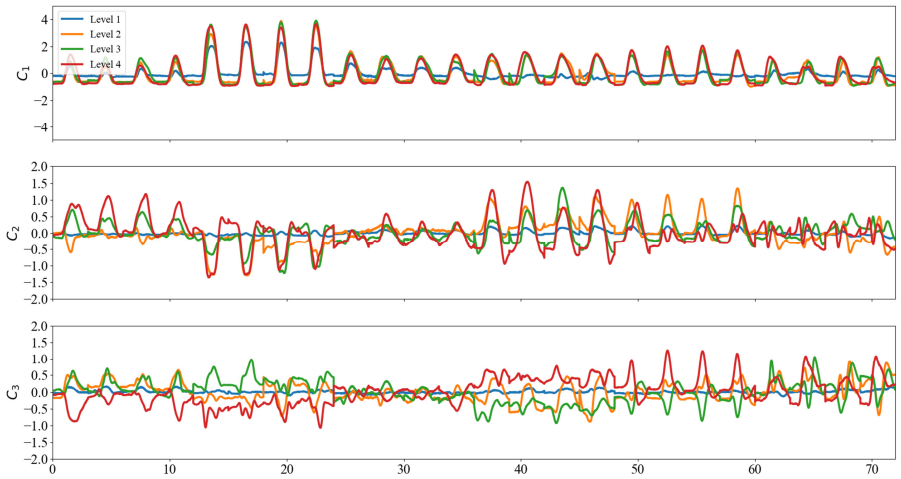


Figure 4.3: Corresponding activation weights of the extracted spatial synergies of joint angles for the four-level pushing forces, C_1, C_2, C_3, represent time-related activation where the balance recovery process can be observed.

process of marker labeling. Then, we asked the subjects to maintain balance through the rotation of the arm, hip, knee, and ankle joints after a disturbing push applied on their back with a stick. Forty-two marker positions were captured at a frequency of 100 Hz through OptiTrack system to measure body kinematics. Each subject was disturbed with four-level pushing forces: weak, small, median, and large forces for 4 trials separately and these distinguished forces were checked by measuring the position deviations of the marker named 'C7' attached on the neck of each subject. To make each time reaction behavior complete, subjects were pushed the next time after maintaining a balanced position without having the obvious body swaying movements.

4.2.2 Data Processing

The conducted experiments include one trial for static pose and 12 trials of balance recovery motion recorded for each subject. Then, we labeled the makers with their names at the beginning of the recording and checked if they were labeled or not frame by frame. After finishing maker labeling, the motion trajectory of each maker was checked, and smoothed by a filter. Then, the recording of maker positions and two force-plates were exported in 'CSV' files. In fact, the current extracted data form can not be directly used for scaling, inverse kinematics (IK), and inverse dynamics (ID). Therefore, we rearranged data forms of static and dynamic motion, filtered ground reaction forces, and removed force noises.

The converted forms of maker positions and GrF were recruited in Opensim. The data processing in the software Opensim was divided into three steps:

1) **Scaling**: the maker positions of static pose were imported and the height and weight of the subject were necessary for scaling process as well. after scaling, we get a scaling model for the subject. This model can be used for the computation of IK and ID.

2) **Inverse kinematics**: the maker positions of balance recovery motion were imported and the temporal joint angles of the whole body were computed and extracted.

3) **Inverse dynamics**: the converted forces from the two force-plates were imported and filtered at 6 Hz. Based on the obtained inverse kinematics results and GrF, the temporal joint torques were computed and extracted.

Then, the extracted joint angles and torques of the both-side arm, hip, knee, ankle were normalized between -1 to $+1$ according to the time of the lowest position at z-axis of the maker 'C7'. The time range 1.5 $[s]$ before and 1.5 $[s]$ after the push was employed for the synergy analysis.

4.2.3 Result Analysis and Discussion

To explore the mechanism of balance recovery after pushing with four-level forces, the kinematic characteristics obtained from the subjects are analyzed and discussed in this section. We also study on the synergy structure alternation regarding W depending on the different levels of push.

From the extraction of spatial synergies shown in Figure 4.2, the ratios of reconstruction accuracy, R^2, and synergy level are varying for the four-level pushing forces. The high ratio of synergy W_1 illustrates common balance recovery spatial feature for all levels in Figure 4.2 (a). For the pushing force level 1, its ratio of W_1 is 0.96, which is almost equal to the total ratio from W_1 to W_3 for the pushing force from level 2 to 4. The eight joint angles can be reconstructed by the three synergies with the high $R^2 = (0.965 \pm 0.005)$ shown in Figure 4.2 (b), which simplifies the kinematic motion modes. This illustrates that the high synergy level of balance recovery movements for the four-level pushing forces.

Even though the reconstruction accuracy by the recruitment of three synergies are almost the same for level $2-4$ pushing forces, the combinations of joint angles and the ratio of R^2 in Figure 4.2 (a) are different respects to the space of synergies W_1, W_2 and W_3. The synergy level of balance in Figure 4.2 (c) is computed by the spatial surface area under curves of the reconstruction accuracy R^2. Although the high reconstruction accuracy exists with the recruitment of W_1, W_2 and W_3 for all pushing forces in this study, the synergy level seems depending on pushing force level. The synergy level decreases as the pushing force increases. This is reasonable as the limitation of joint capability caused by external disturbance leads to the reduction of synergy level.

We can also notice that there is synergy structure alternation regarding W depending on the different levels of push. Ankle angle variation in the space of W_1 is the strongest for the pushing force level 1 as Ch_7 and Ch_8 show the strongest intensity. This indicates that the ankle strategy exists for the pushing force level 1 for small disturbance. The relative joint angle phase between ankle and hip in the space of W_1 is in-phase. Additionally, there is only 1 motion mode for this level.

Secondly, from level 2, we observe there are 2 motion modes combined. The strong coordination of hip joints can be observed through the space of synergy W_1 for the pushing forces from level 2 to 4 as Ch_3 and Ch_4 show the strongest intensity. This should indicate the employment of hip strategy from small level of push. The $1st$ mode W_1 reflects hip strategy. Then, we still observe the ankle strategy at the $2nd$ motion mode at W_2, as as Ch_7 and Ch_8 show the strongest intensity for W_2. Between level 2 and level 3, it demonstrates similar motion mode for W_1 and W_2. However, we can notice the anti-phase between ankle and hip joint angles in synergies W_2 is observed more clearly for level 3. The larger anti-phase of ankle and hip exists for balance recovery reflecting ankle-hip strategy. This opposite flexion of ankle joints relative to hip joints is a well-known balance strategy, for higher disturbing forces.

Then, for further strong push for level 4, arm usage can be noticed as as Ch_1 and Ch_2 show stronger intensity for W_2, and efforts of active arm swing increase as the pushing force increases, where arm strategy is applied for improving the ability of balance recovery. This result is in accordance with our previously obtained conclusion in reproduced balance recovery movements through nonlinear model predictive control (NMPC) [53] that arm

usage helps balance control. Second, the in-phase and anti-phase of the ankle and hip joint illustrates the transition of ankle strategy and hip-ankle strategy. This kind of joint recruitment improves the ability of balance recovery based on the previous conclusion stated in [11].

The evolution of the corresponding activation weights of the extracted spatial synergies of joint angles for the four-level pushing forces is illustrated in Figure 4.3. C_1, C_2, C_3 represent time-related activation where the balance recovery process can be observed. As described in our human experiment protocols, for each level pushing force, each subject performs four trials. For each trial, the time of motion is limited to 3 seconds. There are twenty-four trials that can be observed for each level of pushing force. In C_1, the balance behavior can be well understood. Before the pick value of activation, subjects are pushed, and then they behave in the balance recovery process. Even though the pushing forces are divided into four levels, the balance recovery time is almost the same. This is because humans have the predictive ability and can adjust their standing postures very fast under the cooperation of CNS and body muscles.

4.3 Synergy Analysis for Reproducing Balance Recovery with Active Arm Usage

4.3.1 Models and Methods

4.3.1.1 Description of the Simplified Human Model

To implement the quiet standing balance recovery, we consider the human body as a five-link model comprising a fixed foot, an ankle joint, a lower body, a hip joint, an upper body, a left-right arm joint, a right arm, and a left arm, as illustrated in Figure 4.4. Based on natural body parts weights studied in [97] (Whole trunk: 55.1 percent, two arms: 11.4 percent, two legs (without feet): 33.4 percent, physical parameters are described in Table. 4.1.

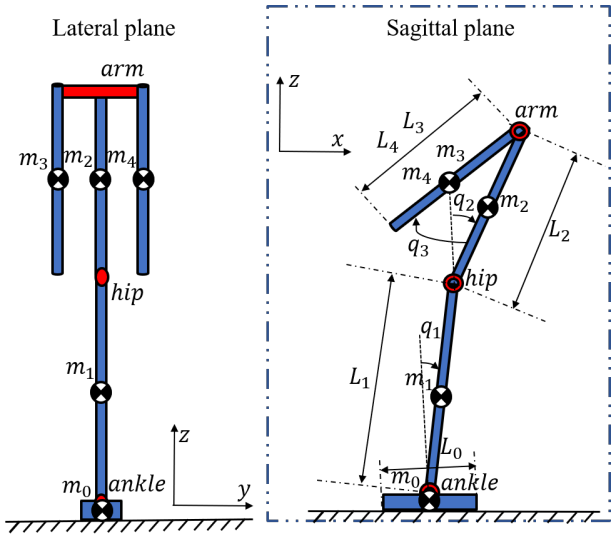


Figure 4.4: Structure of the five-link, three-joint model. m_0, m_1, m_2, m_3, m_4 represent the masses of foot, lower body, upper body, right arm and left arm, respectively. L_0, L_1, L_2, L_3, L_4 represent the lengths of foot, lower body, upper body, right arm, and left arm, respectively. q_1, q_2, q_3 represent the ankle joint angle, hip joint angle, left-right arm joint angle, respectively. The right and left arms share the same joint motor.

4.3.1.2 Dynamic Equations of the Simplified Model

Lagrange formalism is applied to derive the dynamic equation of motion for this five-link, three-joint model controlled by the ankle, hip, and arm joint-torques. The Lagrange equations and dynamic equation of motion are derived for the model with active arm usage.

The dynamic equations of the model with active arms should consider the rotation of the arm joint in the sagittal plane under the control torques generated by the joint actuator. The Lagrange equations for this case are as follows:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) - \frac{\partial L}{\partial q_1} = \tau_{ankle}, \tag{4.3.1}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_2}\right) - \frac{\partial L}{\partial q_2} = \tau_{hip}, \tag{4.3.2}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_3}\right) - \frac{\partial L}{\partial q_3} = \tau_{arm}, \tag{4.3.3}$$

$$L = T - V. \tag{4.3.4}$$

Where, L is the Lagrangian function, T is the total kinetic energy, V is the total potential

Table 4.1: Summary of the physical parameters of the proposed five-link, three-joint model.

Link	Mass [kg]	Length [m]	Height [m]
Foot	1.3	0.3	0.1
Lower body	22.712	1.0	–
Upper body	37.468	0.6	–
Right arm	3.876	0.6	–
Left arm	3.876	0.6	–
Total weight [kg]	69.232	–	–
Total height [m]	–	–	1.7

energy, τ_{ankle} is the ankle torque, τ_{hip} is the hip torque, and τ_{arm} is the arm torque.

The resulting dynamic equation of motion can be expressed in a matrix form as follows:

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} \tau_{ankle} \\ \tau_{hip} \\ \tau_{arm} \end{bmatrix}. \tag{4.3.5}$$

In 4.3.5, M_{11}, M_{22} and M_{33} are the effective inertia terms, M_{12}, M_{13}, M_{21}, M_{23}, M_{31}, and M_{32} are the coupling inertia terms. C_1 C_2, and C_3 terms include centrifugal, Coriolis, and gravity forces.

4.3.1.3 Proposed NMPC For Balance Recovery

In this section, we propose an NMPC scheme [70]to resolve the balance recovery problem. The NMPC problem described above can be solved as an iterative open-loop optimal control problem with a finite horizon and an observable initial state for each sampling time.

The cost function considered in the optimal control problem of NMPC is

$$J\left(\boldsymbol{x}(0), \boldsymbol{\tau}_{(0,N-1)}\right) = \sum_{0}^{N-1} l(\boldsymbol{x}, k, \boldsymbol{\tau}) + V_f,$$

$$l(\boldsymbol{x},k,\boldsymbol{\tau}) = \frac{1}{2}\left(\boldsymbol{x}^T(k)\boldsymbol{Q}\boldsymbol{x}(k) + \boldsymbol{\tau}^T(k)\boldsymbol{R}\boldsymbol{\tau}(k)\right),$$
$$V_f = \frac{1}{2}\boldsymbol{x}^T(N)Q_f\boldsymbol{x}(N).$$

The objective is to minimize the cost $J\left(\boldsymbol{x}(0), \boldsymbol{\tau}_{(0,N-1)}\right)$ subject to the state and control constraints:

1) The joint torques should satisfy the input constraints:

$$-120\ [Nm] \leq \tau_{ankle} \leq 120\ [Nm],$$
$$-500\ [Nm] \leq \tau_{hip} \leq 500\ [Nm],$$
$$-200\ [Nm] \leq \tau_{arm} \leq 200\ [Nm].$$

2) The joint angles satisfy the state constraints:

$$-0.2\ [rad] \leq x_{ankle} \leq 0.4\ [rad],$$
$$-0.35\ [rad] \leq x_{hip} \leq 1.3\ [rad],$$
$$-2.5\ [rad] \leq x_{arm} \leq 0.5\ [rad].$$

4.3.2 Balance Experimental Results Considering Mass Estimation Errors

In this section, balance recovery through active arm usage is implemented considering the mass estimation error. The normal distribution shown in Figure 4.5 is taken into account to reproduce the mass estimation errors, where the mean or expectation of the distribution is equal to zero ($\mu = 0$) and standard deviation σ is set with 0.02. The general form of its probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \tag{4.3.6}$$

And, the mass of model considering Gaussian noise M_{err} can be computed with the real mass of model M_{ori} and normal distribution $f(x)$ in Eq. 4.3.6,

$$M_{err} = M_{ori}(1 + f(x)) \tag{4.3.7}$$

The correlation coefficient $\rho_{X,Y}$ in this study is computed by

$$\rho_{X,Y} = \operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathrm{E}\left[(X - \mu_X)(Y - \mu_Y)\right]}{\sigma_X \sigma_Y} \tag{4.3.8}$$

Where X and Y represent two random variables (in this study, they can be selected from the set of joint angles or the set of joint torques), σ_X and σ_Y represent standard deviations, μ_X and μ_Y represent expected values, E means the expected value operator,cov represents covariance, and corr represents the correlation coefficient. The Pearson correlation [98] is defined only if both standard deviations are finite and positive. An alternative formula purely in terms of moments is

$$\rho_{X,Y} = \frac{\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)}{\sqrt{\mathrm{E}(X^2) - \mathrm{E}(X)^2} \cdot \sqrt{\mathrm{E}(Y^2) - \mathrm{E}(Y)^2}} \tag{4.3.9}$$

In the simulation procedure, the CoM of the upper body is chosen as a push position with different disturbing forces for a duration of 1 $[s]$. The disturbing forces are determined in sagittal plane: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. And, ten trials of simulation are implemented for each disturbing force.

The corresponding results of joint angles are demonstrated in Figures 4.6. And, the heat-map plots in Figures 4.7 are applied to illustrate the correlation of the joint angles and torques. In Figure 4.6, the joint angle deviations increase as the disturbing forces increase. It is worth noting that the evolution of the joint angles in 10 trials does not change a lot for each disturbing force even though the mass errors are introduced in the process of

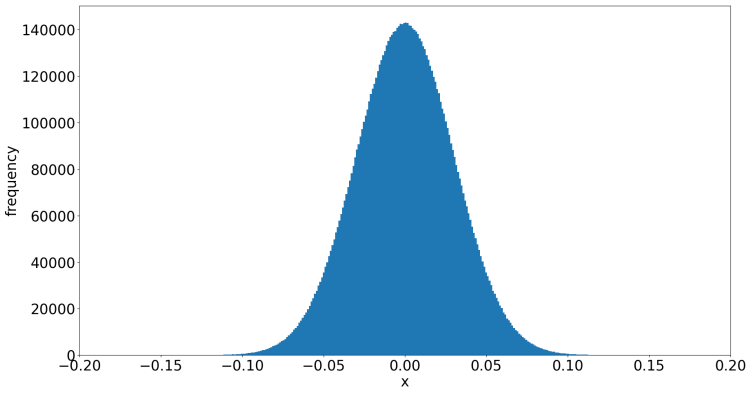


Figure 4.5: Normal distribution applied for mass estimation errors existing in simulation

simulations. This indicates the robustness characteristics of NMPC. And, for the push forces 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$, the model employs the similar balance recovery motion phase with the increase of joint angle deviations while the motion phase switches for the disturbing force 80$[N]$. The correlations of joint angles indicate a similar relation in Figure 4.7. For the disturbing forces 20 $[N]$, 40 $[N]$, 60 $[N]$, and 70 $[N]$, the mean correlation between ankle and hip joint angles is 0.8525, the mean correlation between ankle and arm joint angles is 0.7875, and the mean correlation between hip and arm joint angles is 0.955, which indicate the high correlation among ankle, hip and arm joint angles during the balance recovery motions. However, for the disturbing force 80 $[N]$, the correlation between ankle and hip joint angles is 0.7875, the correlation between ankle and arm joint angles is 0.21, and the correlation between hip and arm joint angles is -0.44, which indicate the low correlation between arm and ankle as well as hip and arm joint angles during the balance recovery motions.

4.3.3 Normalization of the Obtained Results

The purpose of this study is to analyze the balance recovery motion modes through the extracted synergies from joint angles obtained from the simulation. However, the obtained joint angles and torques can not be directly applied to extract synergy due to their different amplitude ranges. Thus, it is necessary to normalize the obtained joint angles. In this study, each feature is scaled by its maximum absolute value. It does not shift/center the data and thus does not destroy any sparsity. For its implementation, **sklearn.preprocessing.MaxAbsScaler** [99] is chosen to scale joint angles and torques

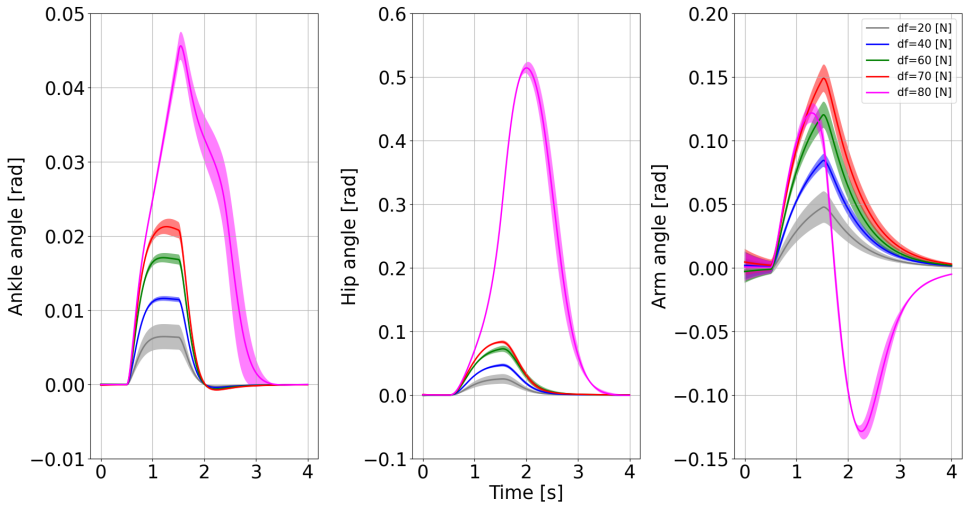


Figure 4.6: Evolution of the joint angles with active arm strategy for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. The "df" represents the disturbing forces.

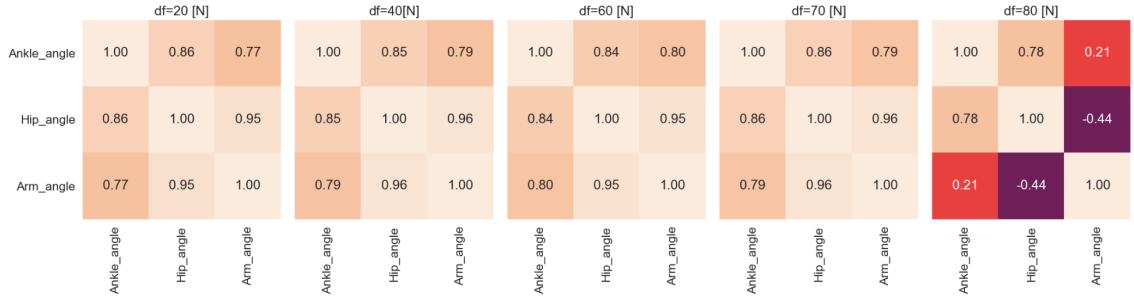


Figure 4.7: Correlation of the joint angles with active arm strategy for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. The "df" represents the disturbing forces.

separately. This estimator scales and translates each feature individually such that the maximal absolute value of each feature in the training set will be 1.0. Evolution plots of the normalized joint angles and torques with active arm strategy for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ are exhibited in Figures 4.8, respectively. These normalized evolution of joint angles is proportional to their original signals exhibited in Figures 4.6. Thus, all the features of balance recovery motion results are kept after normalization.

4.3.4 Spatial Synergy Extraction of Joint Angles

In this section, spatial synergy extracted from joint angles is analyzed to understand the mechanism of balance recovery movements reproduced through NMPC.

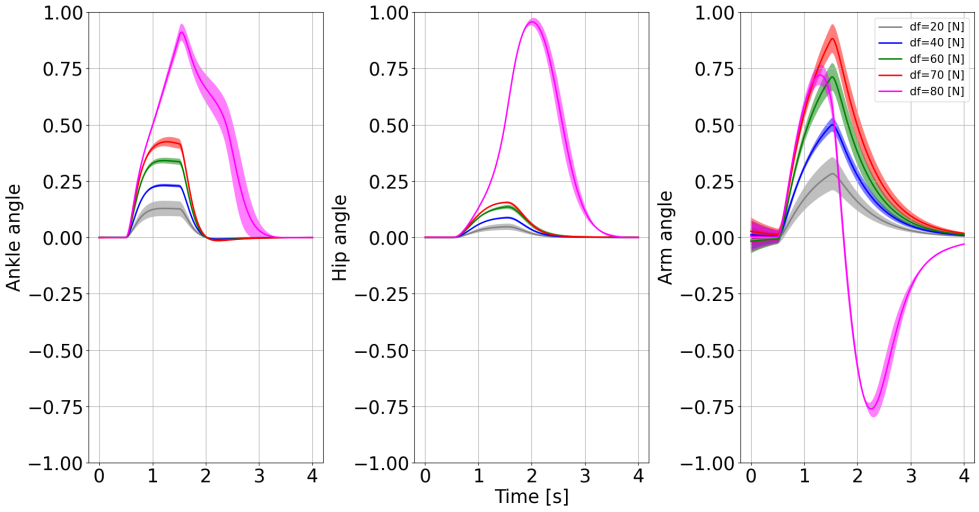


Figure 4.8: Evolution of the normalized joint angles with active arm strategy for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. The "df" represents the disturbing forces.

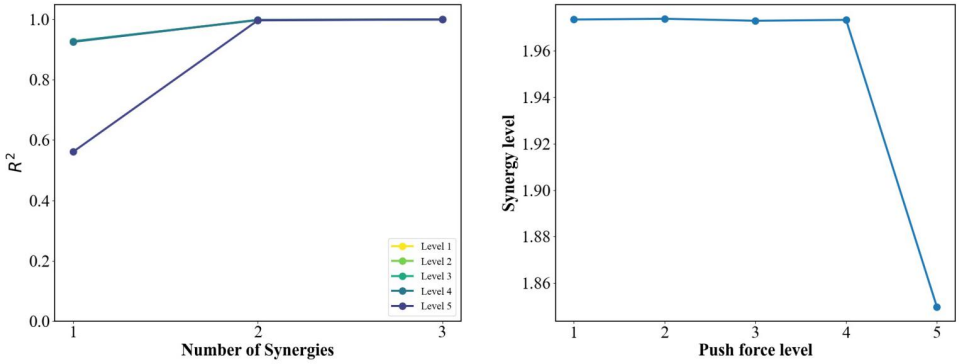


Figure 4.9: The left plot describes spatial reconstruction accuracy of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ w.r.t levels from 1 to 5. Synergy level computed by spatial surface area under curves of the reconstruction accuracy R^2 shown in the right plot.

Spatial synergies of joint angles are extracted to analyze the balance recovery mechanism. spatial reconstruction accuracy of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ w.r.t levels from 1 to 5 and the corresponding synergy level computed by the spatial surface area under curves of the reconstruction accuracy R^2 shown in Figure 4.9. From the reconstruction accuracy of joint angles, only two synergies can reconstruct the original joint angles. This indicates the kinematic joint angles have high synergetic performance, especially, for the disturbing forces from level 1 to 4, the kinematic signals with the accuracy above 90 percent can be reconstructed through one synergy. However, the largest disturbing force can reduce the synergetic performance.

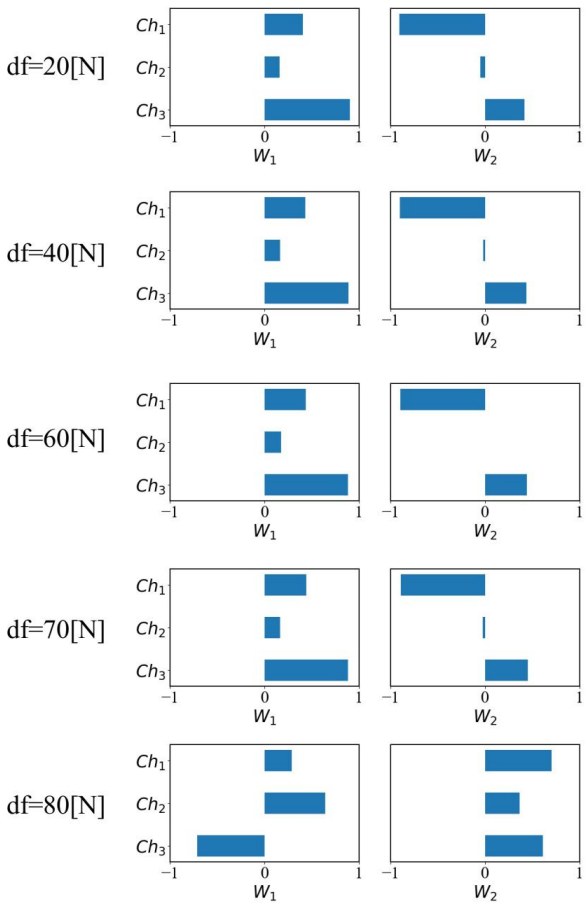


Figure 4.10: Corresponding spatial synergies of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. Ch_1, Ch_2 Ch_3 represent joint angles of ankle, hip, and arm respectively. W_1 and W_2 represent the extracted kinematic synergies of joint angles. The "df" represents the disturbing forces.

For instance, the R^2 for the disturbing force $80[N]$ (Level 5) is less than 0.6. This can be illustrated in the right plot of Figure 4.9, where for the disturbing force $80[N]$ (Level 5), the synergy level decreases a lot.

The corresponding spatial synergies of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ are shown in Figure 4.10. Here, the Ch_1, Ch_2 Ch_3 represent joint angles of ankle, hip, and arm respectively. The spatial synergies of joint angles for $df = 20,\ 40,\ 60,\ 70\ [N]$ exhibit similar motion modes, and arm joint angle modes are recruited much more than the modes of ankle and hip joint angles in W_1. However, for $df = 80\ [N]$, arm joint angle mode switches as anti-phase w.r.t the modes of ankle and hip joint angles in synergy W_1. Besides, the mode of the hip joint angle is recruited much more than the cases for $df = 20,\ 40,\ 60,\ 70\ [N]$, due to the impact of the boundary limitations of ankle and arm joint angles on the NMPC optimal solution

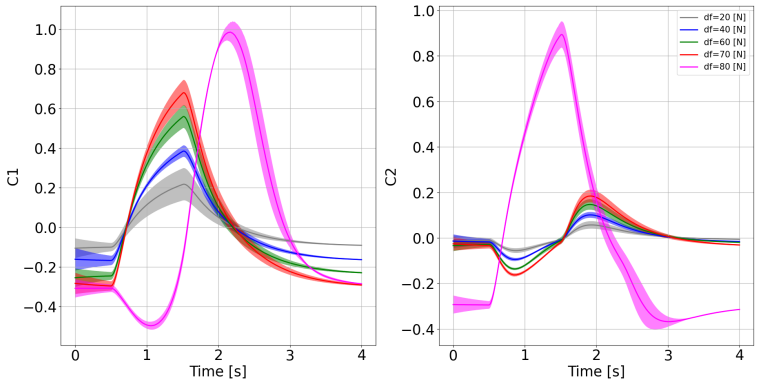


Figure 4.11: Corresponding activation weights of the extracted spatial synergies of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ where the balance recovery process can be observed. The "df" represents the disturbing forces.

decisions. Thus, the extracted synergies represent the motion features of the obtained joint angles with fewer elements. This indicates the high synergetic movements between joint angles exist in balance recovery tasks based on active arm strategy reproduced through NMPC.

Even though the synergies of joint angles for $df = 20,\ 40,\ 60,\ 70\ [N]$ are similar, the corresponding evolution of joint angles is different, which has been illustrated through the corresponding activation weights in Figure 4.11. For the different disturbing forces, the corresponding activation weights are different. It is worth noting that for the same disturbing force, the corresponding activation weights in each trial are different due to the introduced mass estimation errors. However, for $df = 80\ [N]$, there is the phase transition in the evolution of activation weight C_1. And, the deviations of C_1 and C_2 for $df = 80\ [N]$ remain acceptable ranges despite the introduced mass estimation errors. This indicates the robustness of NMPC for dealing with the existing system noise or uncertainty, which is similar to our human ability to deal with the uncertainty information.

4.3.5 Spatiotemporal Synergy Extraction of Joint Angles

In previous section, the spatial synergies are extracted to analyze the spatial joint coordination from kinematic aspects. However, spatial synergies only contain the spatial information of motion, where time-related information can not be analyzed from spatial synergy. The corresponding activation weights are necessary for analyzing the time-related

evolution of motion.

In this section, spatiotemporal synergies of motion are extracted to understand their spatial and temporal coordination existing in balance recovery based on active arm strategy.

In this section, Spatiotemporal synergies of joint angles are extracted to analyze the motion mechanism of balance recovery based on active arm strategy reproduced through NMPC. Figure 4.12 shows the spatiotemporal synergies keep all the motion features with one synergy, where reconstruction accuracy R^2 for each disturbing force is almost near 1.0. This indicates balance recovery motion has high spatiotemporal synergy. Spatiotemporal synergy can represent more features of balance recovery motion than spatial synergy through comparing the reconstruction accuracy R^2 and synergy level shown in Figures 4.9 and 4.12. And in Figure 4.12, the synergy level decreases since the boundary constraints of ankle joint torques influences the NMPC optimal solutions, which further affects the coordination of joint angles in Figure 4.6.

The corresponding spatiotemporal synergies of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ are exhibited in Figure 4.13. Ch_1, Ch_2 Ch_3 represent joint angles of ankle, hip, and arm respectively. W_1 represents the extracted kinematic synergy of joint angles. The "df" represents the disturbing forces. The spatiotemporal synergies of the ankle, hip, arm joint angles for $df = 20,\ 40,\ 60,\ 70\ [N]$ are similar because of the similar motion behavior except for the amplitudes of swinging. And, the amplitude difference of joint angles is illustrated in Figure 4.14 where the activation weight C_1 for the disturbing force from 20 $[N]$ to 70 $[N]$ increases. However, the spatiotemporal synergy for the disturbing force 80 $[N]$ is totally different from those for the smaller disturbing forces. This is because the optimal solution of NMPC for balance recovery is changed facing the boundary constraints of ankle joint torques. It is worth noting that the evolution waves of spatiotemporal synergy are ratio with those of joint angles, which illustrates the spatial and temporal features of the original information are kept in the extracted spatiotemporal synergies. This is the main advantage of using the spatiotemporal synergy for motion feature analysis.

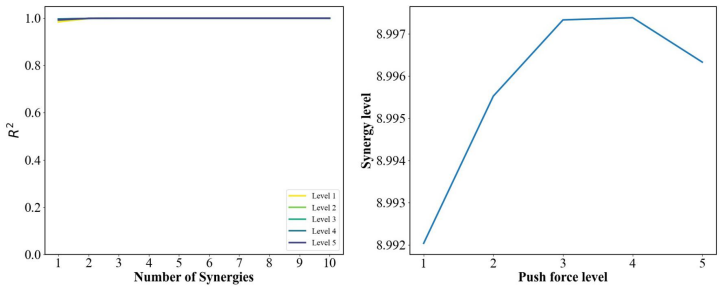


Figure 4.12: The left plot describes spatiotemporal reconstruction accuracy of joint angles for for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ w.r.t levels from 1 to 5. Synergy level computed by spatiotemporal surface area under curves of the reconstruction accuracy R^2 shown in the right plot.

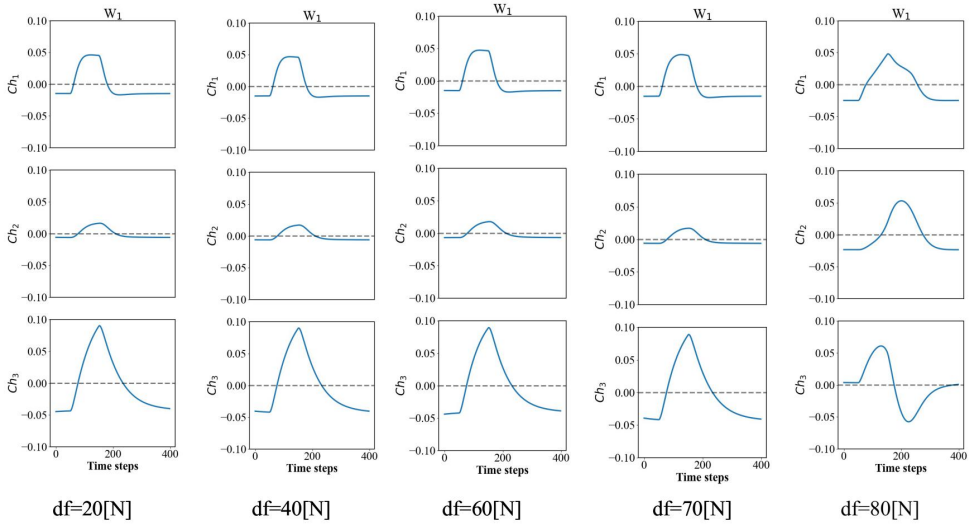


Figure 4.13: Corresponding spatiotemporal synergies of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$. Ch_1, Ch_2 Ch_3 represent joint angles of ankle, hip, and arm respectively. W_1 represent the extracted kinematic synergy of joint angles. The "df" represents the disturbing forces.

4.4 Conclusions

In this chapter, the extraction methods of spatial synergy and spatiotemporal synergy are introduced first in section 4.1.

Then, in section 4.2, various balance recovery motions in human trials for four different levels of pushing forces are classified as balance recovery approaches like arm strategy, ankle strategy, and hip-ankle strategy based on the extracted spatial synergy. The common patterns and the high synergy level of balance recovery motions are confirmed as well, which indicates the strong coordination between joint angles based on the switch of balance

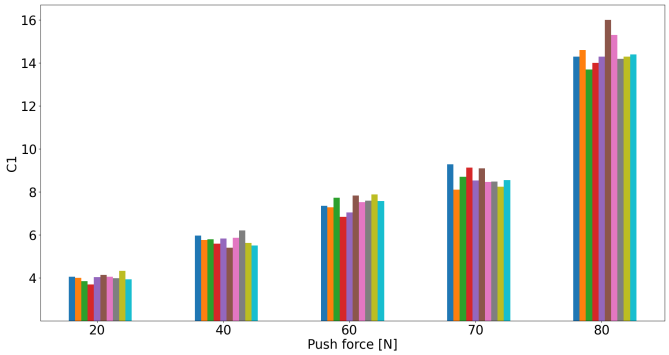


Figure 4.14: Corresponding activation weights of the extracted spatiotemporal synergies of joint angles for different disturbing forces: 20 $[N]$, 40 $[N]$, 60 $[N]$, 70 $[N]$, and 80 $[N]$ where the balance recovery process can be observed. The "df" represents the disturbing forces.

recovery strategies. The synergy level is computed to further study the performance of human joint capability for the different pushing forces.

In section 4.3, the models and methods using for reproducing active arm strategy are explained first. Then, the experimental joint angles are normalized and analyzed directly for understanding the balance recovery mechanism. Furthermore, spatial synergy and spatiotemporal synergy of joint angles are extracted respectively to analyze the motion and control features of balance recovery. And, the reproduced balance recovery behavior displays the high coordination ability of motion and control.

Chapter 5

Reproducing Postural Coordination Modes Through Model-free DRL

After studying static balance strategies in previous chapters, we will discuss dynamic balance strategy through postural coordination modes reproduced through model-free Deep Reinforcement Learning (DRL) in head-target tracking tasks. In this chapter, the basic DRL and soft actor-critic (SAC) used in our study are first introduced. Then, the model and specific setting of the algorithm are further described. Trajectory tracking swing motions by tracking speeds were trained through model-free DRL. After training, the postural coordination modes can be observed w.r.t. different tracking frequencies. Finally, the corresponding hip ankle joint correlation was computed to classify the transition of human posture including in-phase and anti-phase between joint angles of ankle and hip. The temporal joint angles and phase portraits of hip and ankle were compared to classify the motion transition depending on frequencies as well. And, the CoM deviations were compared for different frequencies to analyze the motion features. It is worth noting that the coordination mode transitions contribute to the decrease of CoM deviation. The emergent phenomena of postural coordination mode change from in-phase to antiphase decrease the deviation of CoM. It is reasonable that for high-speed tracking tasks, the agent should reduce the CoM deviations for stable tracking performance.

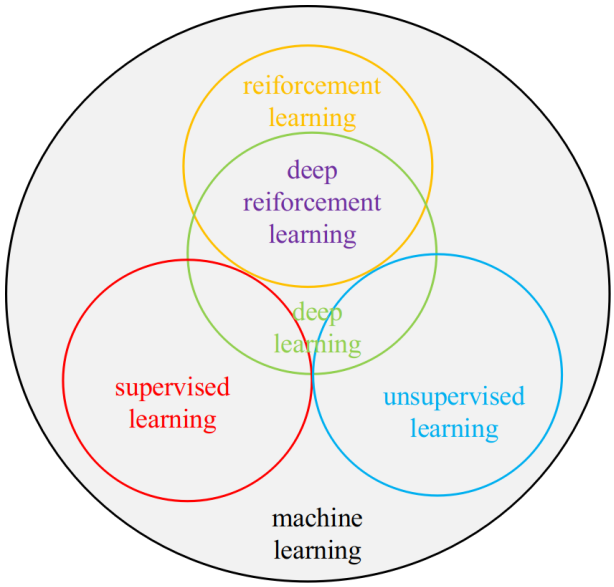


Figure 5.1: Illustration on the relation of deep reinforcement learning (DRL) with respect to machine learning (ML).

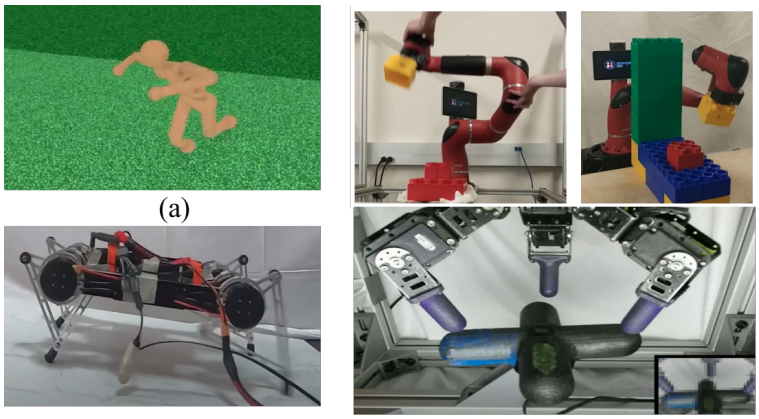


Figure 5.2: (a) Locomotion for a humanoid agent in OpenAI Gym [100]. (b) Minitaur robot walking in [101]. (c) Robotic manipulation learning in [101].

5.1 Overview of DRL Implementation with SAC

5.1.1 Basic Concepts of DRL

Reinforcement learning (RL) [105], [106] is a machine learning technique (see illustration of Figure 5.1) concerning how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward [107]. In a nutshell [108], RL is the study of agents and how they learn by trial and error. The idea is that rewarding or punishing an agent for its behavior makes it more likely to repeat or forego that behavior

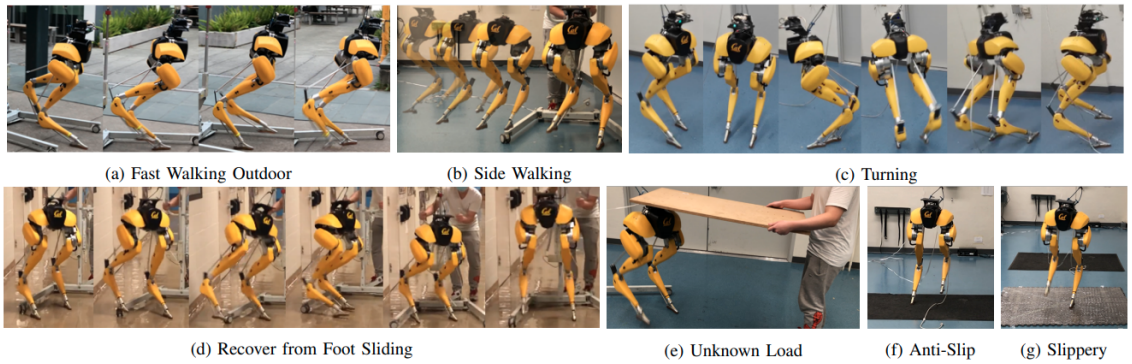
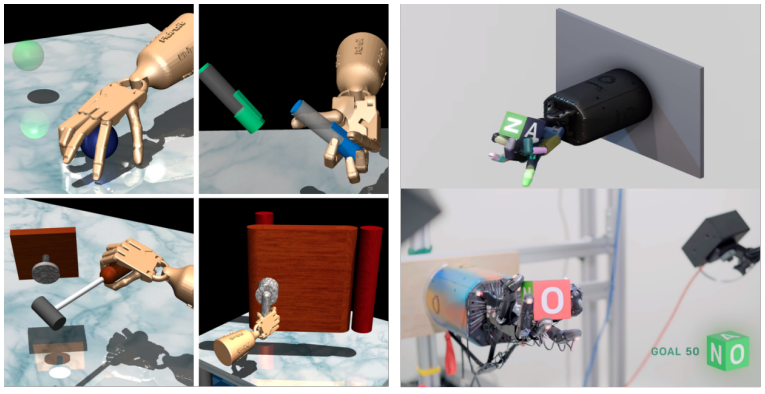


Figure 5.3: The proposed learned walking policy extensively on Cassie in the real-world with different scenarios [102]. In the experiments, the policy enables Cassie to perform various agile behaviors such as fast forward and backward walking, sideways walking, changing walking height, and turning around. Moreover, empowered by the proposed policy, the robot can recover from random perturbation and adapt to change different ground friction changes as well as unknown loads.



(a) (b)

Figure 5.4: A wide range of dexterous manipulation skills using DRL methods in simulation (a) [103] and real world (b) [104].

in the future. DRL [109], [110] represents the combination of RL with deep learning. DRL has succeed in a wide variety of fields including robotics [111]–[116], control in simulation and other real world applications [101]–[103], [117], operation research, economics, games [118]. For instance, through combining deep learning in RL (DRL), real-world learning for challenging tasks such as locomotion for a humanoid agent [100] (see Figure 5.2(a)) and real biped robots [102] (see Figure 5.3) and Minitaur robot (see Figure 5.2(b)), robotic manipulation [101] (see Figure 5.2(c)), dexterous robotic hand motion [103] and real-world dexterous manipulation [104] (see Figure 5.4) were successfully achieved.

The agent and the environment represent the main characters of RL. The interaction loop between agent and environment is illustrated in Figure 5.5. The environment means the

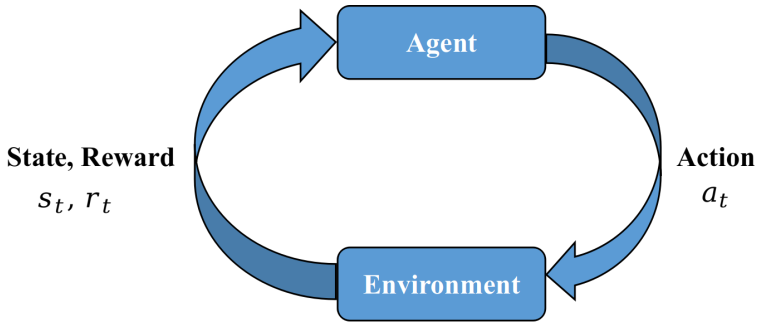


Figure 5.5: Agent-environment interaction loop in RL [108].

world that the agent evolves in and interacts with. At every step of interaction, the agent sees a (possibly partial) observation of the state of the world and then decides on an appropriate action to take. The environment changes when the agent acts on it, but may change by its own as well. The agent perceives a reward signal from the environment as well, a number that estimates how good or bad the current world state is. The purpose of the agent is to maximize its cumulative reward, called return. Reinforcement learning methods are approaches that the agent can learn behaviors to implement its tasks. Additional terminology is introduced to understand more specifically what RL does according to [108].

5.1.1.1 States and Observations

A state *s* is a variable describing completely the state of the world. There is no information about the world which is hidden from the state. An observation *o* describes a state partially. This may omit information. In DRL, states and observations are generally represented by a real-valued vector, matrix, or higher-order tensor. For example, the RGB matrix of its pixel values could be used to represent a visual observation; joint angles and velocities might be used to represent the state of a robot. When the agent can observe the complete state of the environment, this means that the environment is fully observed. When the agent can only see a partial observation, this means that the environment is partially observed.

5.1.1.2 Action Spaces

Different environments allow different types of actions to be used. The action space is a set of all valid actions in a given environment. Some environments (e.g. Go and Atari) have discrete action spaces in which the agent can only use a limited number of actions.

Agents have continuous action spaces where they can control robots in a real or simulated world; actions are real-valued vectors in a continuous space. This distinction has a quite profound effect on the approach in DRL. Certain families of algorithms can only be used directly in one case and must substantially be redesigned for other cases.

5.1.1.3 Policies

A policy represents a rule used by an agent to decide on how to take action. A policy can be deterministic, it is usually denoted by μ

$$a_t = \mu(s_t), \tag{5.1.1}$$

it may also be stochastic, in which case it is usually denoted by π:

$$a_t \sim \pi(\cdot \mid s_t), \tag{5.1.2}$$

Since the policy is essentially the agent's brain, it is not unusual to use the word 'policy' instead of 'agent', e.g. to say 'the policy tries to maximize the reward'.

In DRL, we deal with parametric policies: policies whose output is a computable function that depends on a set of parameters (e.g. the weights and biases of a neural network) that can be tuned to change the behavior utilizing some optimization algorithm.

We usually use θ or ϕ to obtain the parameters of such strategies, which are then written as subscripts over the strategy symbols to highlight the connections:

$$\begin{aligned} a_t &= \mu_\theta(s_t) \\ a_t &\sim \pi_\theta(\cdot \mid s_t) \end{aligned} \tag{5.1.3}$$

5.1.1.4 Trajectories

A trajectory τ is compose of a sequence of states and actions in the world,

$$\tau = (s_0, a_0, s_1, a_1, \dots). \tag{5.1.4}$$

The first state of the world, s_0, is randomly sampled from the start-state distribution denoted by ρ_0:

$$s_0 \sim \rho_0, \tag{5.1.5}$$

The state transitions (what happens to the world between the state at time t, s_t, and the state at time $t+1$, s_{t+1}), are governed by the natural laws of the environment, and depend on only the latest action, a_t. They can be either deterministic,

$$s_{t+1} = f(s_t, a_t), \tag{5.1.6}$$

or stochastic,

$$s_{t+1} \sim P(\cdot \mid s_t, a_t). \tag{5.1.7}$$

5.1.1.5 Reward and Return

The reward function R in RL depends on the current state of the world, the action is just taken, and the next states of the world are denoted:

$$r_t = R(s_t, a_t, s_{t+1}). \tag{5.1.8}$$

However, frequently, this is simplified to just a dependence on the current state, $r_t = R(s_t)$, or state-action pair $r_t = R(s_t, a_t)$.

One type of return is the finite-horizon undiscounted return, which is computed by the sum of rewards obtained in a fixed window of steps:

$$R(\tau) = \sum_{t=0}^{T} r_t \tag{5.1.9}$$

Another type of return is the infinite-horizon discounted return, which is computed by the sum of all rewards ever obtained by the agent but discounted by how far off in the future they're obtained. This formulation of reward includes a discount factor $\gamma \in (0,1)$:

$$R(\tau) = \sum_{t=0}^{\infty} \gamma^t r_t. \tag{5.1.10}$$

Mathematically, an infinite-horizon sum of rewards could not converge to a finite value and is hard to deal with in equations. However, with a discount factor and under some reasonable conditions, the infinite sum converges.

5.1.1.6 The RL Problem

No matter choose which return measure (whether infinite-horizon discounted, or finite-horizon undiscounted), and which policy, the purpose in RL is to choose a policy that maximizes expected return when the agent acts according to it. To achieve expected return, we first have to introduce probability distributions over trajectories. Suppose that both the environment transitions and the policy are stochastic. In this case, the probability of a T -step trajectory can be defined as follows:

$$P(\tau \mid \pi) = \rho_0(s_0) \prod_{t=0}^{T-1} P(s_{t+1} \mid s_t, a_t)\, \pi(a_t \mid s_t). \tag{5.1.11}$$

The expected return $J(\pi)$ expressed as follows:

$$J(\pi) = \int_{\tau} P(\tau \mid \pi) R(\tau) = \underset{\tau \sim \pi}{\mathrm{E}}[R(\tau)] \tag{5.1.12}$$

Then, the central optimization problem in RL can be formulated as follows

$$\pi^* = \arg\max_{\pi} J(\pi) \tag{5.1.13}$$

5.1.1.7 Value Functions

Four main value function are introduced here.

1) The On-Policy Value Function $V^{\pi}(s)$:

$$V^{\pi}(s) = \mathop{\mathrm{E}}_{\tau\sim\pi}[R(\tau) \mid s_0 = s]. \tag{5.1.14}$$

Where it gives the expected return if the agent starts in state s and always acts according to policy π.

2) The On-Policy Action-Value Function $Q^{\pi}(s,a)$:

$$Q^{\pi}(s,a) = \mathop{\mathrm{E}}_{\tau\sim\pi}[R(\tau) \mid s_0 = s, a_0 = a]. \tag{5.1.15}$$

Where it gives the expected return if the agent starts in state s, takes an arbitrary action a which may not have come from the policy, and then forever after action according to policy π.

3) The Optimal Value Function $V^*(s)$:

$$V^*(s) = \max_{\pi} \mathop{\mathrm{E}}_{\tau\sim\pi}[R(\tau) \mid s_0 = s]. \tag{5.1.16}$$

Where it gives the expected return if the agent starts in state s and always acts according to the optimal policy in the environment.

4) The Optimal Action-Value Function $Q^*(s,a)$:

$$Q^*(s,a) = \max_{\pi} \mathop{\mathrm{E}}_{\tau\sim\pi}[R(\tau) \mid s_0 = s, a_0 = a]. \tag{5.1.17}$$

Where it gives the expected return if the agent starts in state s, takes an arbitrary action a, and then forever after action according to the optimal policy in the environment.

5.1.1.8 The Optimal Q-Function and the Optimal Action

There is an important relationship between the optimal action-value function $Q^*(s,a)$ and the action selected by the optimal policy. $Q^*(s,a)$ gives the expected return for starting in state s, taking (arbitrary) action a, and then acting according to the optimal policy forever after.

The optimal policy in s will select whichever action maximizes the expected return from starting in s. As a result, if we have Q^*, the optimal action, $a^*(s)$ can be directly computed through the solution of the following problem

$$a^*(s) = \arg\max_a Q^*(s,a). \tag{5.1.18}$$

It is worth to note that there may exist various actions that maximize $Q^*(s,a)$, and all of them are optimal, and the optimal policy may randomly choose any of them. However, there exist always an optimal policy that deterministically chooses an action.

5.1.1.9 Bellman Equations

All four of the value functions obey Bellman equations [109]. The Bellman equations for the on-policy value functions are expressed by:

$$V^{\pi}(s) = \mathop{\mathrm{E}}_{\substack{a\sim\pi \\ s'\sim P}} \left[r(s,a) + \gamma V^{\pi}\left(s'\right) \right], \tag{5.1.19}$$

$$Q^{\pi}(s,a) = \mathop{\mathrm{E}}_{s'\sim P} \left[r(s,a) + \gamma \mathop{\mathrm{E}}_{a'\sim\pi} \left[Q^{\pi}\left(s',a'\right) \right] \right], \tag{5.1.20}$$

where $s' \sim P$ is shorthand for $s' \sim P(\cdot \mid s,a)$, indicating that the next state s' is sampled from the environment's transition rules; $a \sim \pi$ is shorthand for $a \sim \pi(\cdot \mid s)$; and $a' \sim \pi$ is shorthand for $a' \sim \pi(\cdot \mid s')$.

The Bellman equations for the optimal value functions are expressed by:

$$V^*(s) = \max_a \mathop{\mathrm{E}}_{s' \sim P} \left[r(s,a) + \gamma V^* \left(s' \right) \right], \tag{5.1.21}$$

$$Q^*(s,a) = \mathop{\mathrm{E}}_{s' \sim P} \left[r(s,a) + \gamma \max_{a'} Q^* \left(s', a' \right) \right]. \tag{5.1.22}$$

The main difference between the Bellman equations for the on-policy value functions and the optimal value functions lies in the absence or presence of the max over actions. Its inclusion indicates the fact that whenever the agent gets to choose its action, to act optimally, it has to pick whichever action leading to the highest value.

5.1.1.10 Advantage Functions

The advantage function $A^\pi(s,a)$ corresponding to a policy π describes how much better it is to take a specific action a in state s, over randomly choosing an action according to $\pi(\cdot|s)$, supposing that the agent acts according to π forever after. Mathematically, the advantage function is denoted by

$$A^\pi(s,a) = Q^\pi(s,a) - V^\pi(s). \tag{5.1.23}$$

5.1.1.11 Markov Decision Processes

Markov Decision Processes (MDP) [109] refers to the fact that the system obeys the Markov property: transitions only depend on the most recent state and action, and no prior history. An MDP is a 5-tuple, $\langle S, A, R, P, \rho_0 \rangle$, where

1) S is the set of all valid states,
2) A is the set of all valid actions,
3) $R : S \times A \times S \to \mathbb{R}$ is the reward function, with $r_t = R(s_t, a_t, s_{t+1})$,
4) $P : S \times A \to \mathscr{P}(S)$ is the transition probability function, with $P(s'|s,a)$ being the probability of transitioning into state s' if the agent starts in state s and take action a, and ρ_0 is

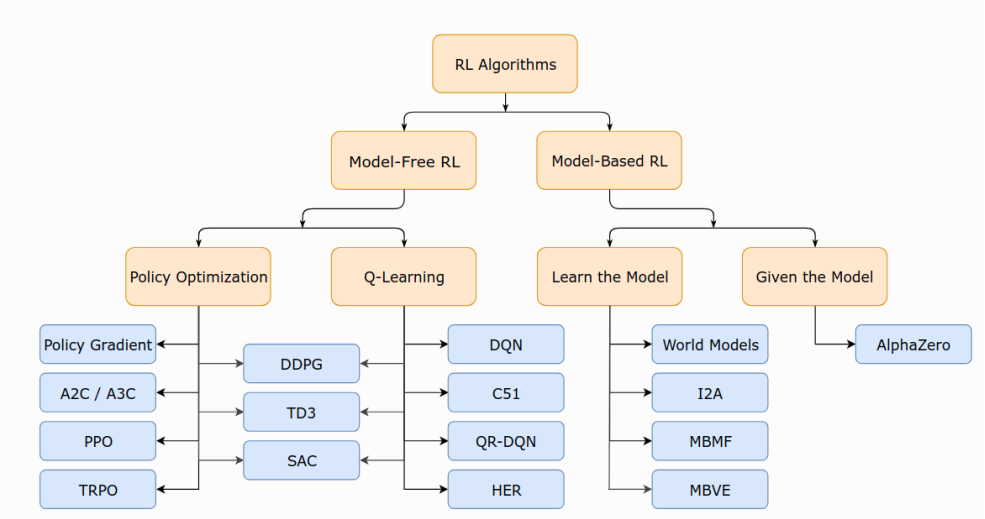


Figure 5.6: A non-exhaustive, but useful taxonomy of algorithms in modern RL in [108].

the starting state distribution.

5.1.1.12 Kinds of RL Algorithms

A non-exhaustive, but very useful taxonomy of algorithms of modern RL in [108] is illustrated in Figure 5.6. For further details, the reader can refer to [108].

5.1.2 Overview of the main Soft Actor-Critic (SAC)

Soft Actor Critic (SAC) [119] is an algorithm that optimizes a stochastic policy in an off-policy way, forming a bridge between stochastic policy optimization and DDPG-style approaches. A central feature of SAC is entropy regularization. The policy is trained to maximize a trade-off between the expected return and entropy, a measure of randomness in the policy. This has a close connection to the exploration-exploitation trade-off: increasing entropy results in more exploration, which can accelerate the learning later on. It can also prevent the policy from prematurely converging to a bad local optimum. There are various algorithm developments and applications about SAC in [101], [103], [117], [120], [121].

In this study, it is necessary to choose a DRL algorithm that can solve complex robotics tasks. We hope that the strategy can stably and efficiently explore an unknown environment to obtain the optimal strategy. The SAC algorithm can well meet this requirement by improving both the algorithm's robustness and its sample efficiency through training for a

Algorithm 1 Soft Actor-Critic

Input: initial policy parameters θ, Q-function parameters ϕ_1, ϕ_2, empty replay buffer $\mathcal{D}$
Set target parameters equal to main parameters $\phi_{\text{targ},1} \leftarrow \phi_1, \phi_{\text{targ},2} \leftarrow \phi_2$
repeat
    Observe state s and select action $a \sim \pi_\theta(\cdot \mid s)$
    Execute a in the environment
    Observe next state s', reward r, and done signal d to indicate whether s' is terminal
    Store (s, a, r, s', d) in replay buffer $\mathcal{D}$
    If s' is terminal, reset environment state.
    if it's time to update **then**
        for j in range(however many updates) **do**
            Randomly sample a batch of transitions, $B = \{(s, a, r, s', d)\}$ from $\mathcal{D}$
            Compute targets for the Q functions:

$$y(r, s', d) = r + \gamma(1-d)\left(\min_{i=1,2} Q_{\phi_{\text{tarr}\ i}}(s', \tilde{a}') - \alpha \log \pi_\theta(\tilde{a}' \mid s')\right), \quad \tilde{a}' \sim \pi_\theta(\cdot \mid s')$$

            Update Q-functions by one step of gradient descent using

$$\nabla_{\phi_i} \frac{1}{|B|} \sum_{(s,a,r,s',d) \in B} \left(Q_{\phi_i}(s, a) - y(r, s', d)\right)^2 \quad \text{for } i = 1, 2$$

            Update policy by one step of gradient ascent using

$$\nabla_\theta \frac{1}{|B|} \sum_{s \in B} \left(\min_{i=1,2} Q_{\phi_i}(s, \tilde{a}_\theta(s)) - \alpha \log \pi_\theta(\tilde{a}_\theta(s) \mid s)\right)$$

            where $\tilde{a}_\theta(s)$ is a sample from $\pi_\theta(\cdot \mid s)$ which is differentiable wrt θ via the reparametrization trick.
            Update target networks with

$$\phi_{\text{targ},i} \leftarrow \rho \phi_{\text{targ},i} + (1-\rho)\phi_i \quad \text{for } i = 1, 2$$

        end for
    end if
until convergence

maximum entropy [121].

To explain SAC algorithm [121], we first introduce the entropy-regularized reinforcement learning setting. In entropy-regularized RL, the equations of the value functions are slightly different.

5.1.2.1 Entropy-Regularized Reinforcement Learning

Let x be a random variable with probability mass or density function P. The entropy H of x is computed from its distribution P as follows

$$H(P) = \underset{x \sim P}{\mathrm{E}}[-\log P(x)] \tag{5.1.24}$$

In entropy-regularized reinforcement learning, the agent gets a bonus reward at each time step proportional to the entropy of the policy at that time step. This changes the RL problem to:

$$\pi^* = \arg\max_{\pi} \mathrm{E}_{\tau\sim\pi}\left[\sum_{t=0}^{\infty}\gamma^t\left(R\left(s_t,a_t,s_{t+1}\right)+\alpha H\left(\pi\left(\cdot \mid s_t\right)\right)\right)\right], \tag{5.1.25}$$

where $\alpha > 0$ is the trade-off coefficient. We can now define the slightly different value functions in this setting. V^π is changed to include the entropy bonuses from every time step:

$$V^\pi(s) = \mathop{\mathrm{E}}_{\tau\sim\pi}\left[\sum_{t=0}^{\infty}\gamma^t\left(R\left(s_t,a_t,s_{t+1}\right)+\alpha H\left(\pi\left(\cdot \mid s_t\right)\right)\right) \mid s_0 = s\right] \tag{5.1.26}$$

Q^π is changed to include the entropy bonuses from every time step except the first one:

$$Q^\pi(s,a) = \mathop{\mathrm{E}}_{\tau\sim\pi}\left[\sum_{t=0}^{\infty}\gamma^t R\left(s_t,a_t,s_{t+1}\right)+\alpha\sum_{t=1}^{\infty}\gamma^t H\left(\pi\left(\cdot \mid s_t\right)\right) \mid s_0 = s, a_0 = a\right] \tag{5.1.27}$$

With these definitions, V^π and Q^π are connected by:

$$V^\pi(s) = \mathop{\mathrm{E}}_{a\sim\pi}\left[Q^\pi(s,a)\right] + \alpha H(\pi(\cdot \mid s)) \tag{5.1.28}$$

and the Bellman equation for Q^π is expressed as follows,

$$\begin{aligned} Q^\pi(s,a) &= \mathop{\mathrm{E}}_{\substack{s'\sim P\\ a'\sim\pi}}\left[R\left(s,a,s'\right)+\gamma\left(Q^\pi\left(s',a'\right)+\alpha H\left(\pi\left(\cdot \mid s'\right)\right)\right)\right] \\ &= \mathop{\mathrm{E}}_{s'\sim P}\left[R\left(s,a,s'\right)+\gamma V^\pi\left(s'\right)\right] \end{aligned} \tag{5.1.29}$$

5.1.2.2 Soft Actor-Critic

The large continuous domain requires us to derive a practical approximation of soft policy iterations. To this end, we will use a function approximator for the soft Q function and strategy, instead of running the evaluation and improving to convergence, we will alternately optimize the two networks using stochastic gradient descent. We will consider a parameterized soft Q function and an operational strategy. The parameters of these networks are θ and ϕ. For example, the soft Q function can be modeled as an expressive neural network, and the policy is a Gaussian with mean and covariance given by the neural network. Next, we derive the update rules for these parameter vectors.

The soft Q-function parameters can be trained to minimize the soft Bellman residual

$$J_V(\psi) = \mathbb{E}_{s_t \sim \mathscr{D}} \left[\frac{1}{2} \left(V_\psi(\mathbf{s}_t) - \mathbb{E}_{\mathbf{a}_t \sim \pi_\phi} \left[Q_\theta(\mathbf{s}_t, \mathbf{a}_t) - \log \pi_\phi(\mathbf{a}_t \mid \mathbf{s}_t) \right] \right)^2 \right] \tag{5.1.30}$$

Besides, the value function is implicitly parameterized by the parameters of the soft Q-function and can be optimized using the Stochastic gradient:

$$\hat{\nabla}_\psi J_V(\psi) = \nabla_\psi V_\psi(\mathbf{s}_t) \left(V_\psi(\mathbf{s}_t) - Q_\theta(\mathbf{s}_t, \mathbf{a}_t) + \log \pi_\phi(\mathbf{a}_t \mid \mathbf{s}_t) \right) \tag{5.1.31}$$

here, instead of the replay buffer, the actions are sampled based on the current policy. The parameters of the soft Q-function can be learned to minimize the soft Bellman residuals

$$J_Q(\theta) = \mathbb{E}_{(\mathbf{s}_t, \mathbf{a}_t) \sim \mathscr{D}} \left[\frac{1}{2} \left(Q_\theta(\mathbf{s}_t, \mathbf{a}_t) - \hat{Q}(\mathbf{s}_t, \mathbf{a}_t) \right)^2 \right] \tag{5.1.32}$$

with

$$\hat{Q}(\mathbf{s}_t, \mathbf{a}_t) = r(\mathbf{s}_t, \mathbf{a}_t) + \gamma \mathbb{E}_{\mathbf{s}_{t+1} \sim p} \left[V_{\bar{\psi}}(\mathbf{s}_{t+1}) \right] \tag{5.1.33}$$

which again can be optimized with the stochastic gradients:

$$\hat{\nabla}_\theta J_Q(\theta) = \nabla_\theta Q_\theta\left(\mathbf{a}_t, \mathbf{s}_t\right)\left(Q_\theta\left(\mathbf{s}_t, \mathbf{a}_t\right) - r\left(\mathbf{s}_t, \mathbf{a}_t\right) - \gamma V_{\bar{\psi}}\left(\mathbf{s}_{t+1}\right)\right) \tag{5.1.34}$$

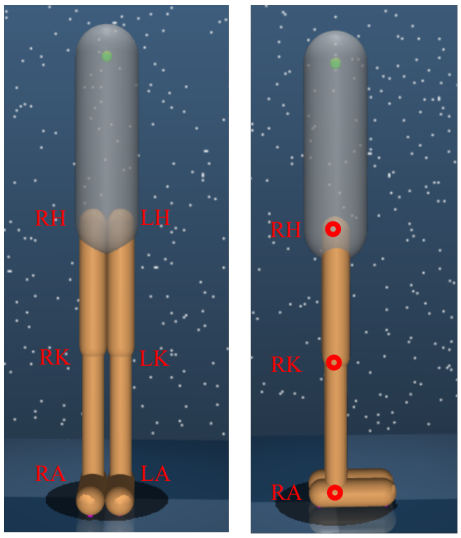
This update utilizes the target value network $V_{\bar{\psi}}$, where $\bar{\psi}$ can be an exponential moving average of the the value of the network weights that are stable for training [105]. Alternatively, we can periodically update the target weights to match the current value function weights. Finally, the strategy parameters can be learned by directly minimizing the expected Kullback-Leibler discrepancy given by

$$J_\pi(\phi) = \mathbb{E}_{\mathbf{s}_t \sim \mathscr{D}}\left[D_{KL}\left(\pi_\phi\left(\cdot \mid \mathbf{s}_t\right) \| \frac{\exp\left(Q_\theta\left(\mathbf{s}_t, \cdot\right)\right)}{Z_\theta\left(\mathbf{s}_t\right)}\right)\right] \tag{5.1.35}$$

There are several options for minimizing J_π. A typical solution for the policy gradient method is to use a likelihood ratio gradient estimation method that does not require back-propagation of the gradient through a network of policies and target densities. However, in the case of [121], the target density is a Q-function represented by a neural network and can be differentiated, then it is convenient to apply a reparameterization trick instead, resulting in low variance. Therefore, we use neural network transformations to reparameterize the policy:

$$\mathbf{a}_t = f_\phi\left(\varepsilon_t; \mathbf{s}_t\right) \tag{5.1.36}$$

The SAC policy is motivated by exploring explore more broadly while abandoning hopeless approaches; it does not assign a very high probability to any action within the scope of the action. This allows SAC to explore random behaviors more frequently during training, which may result in a faster learning speed with respect to the state-of-the-art methods that optimize traditional RL objective functions. The pseudo-code [121] of the SAC algorithm is described as Algorithm 1.



(a) In sagittal plane (b) In lateral plane

Figure 5.7: View of the walker model used in this study.' RA', 'LA', 'RK', 'LK', 'RH', and 'LH' stand for the abbreviations of right ankle, left ankle, right knee, left knee, right hip, and left hip, respectively.

5.2 Transition Analysis of Tracking Motion with Kinematic Results

In this study, we try to use model-free RL methods such as SAC to reproduce human-like 'hip-ankle' kinematic phase transition existing in the head back and forth tracking tasks depending on the speed [122]–[124]. Furthermore, it naturally emerges with two modes "in-phase or anti-phase" of hip and ankle movement as self-organization phenomena. The related human experiments have been reported in [122], [123] and its corresponding numerical model-based optimization study based on a double-inverted pendulum has been implemented in [124]. However, there is still no literature about RL-based postural coordination modes reproduction for head back and forth tracking tasks. Accordingly, our contribution in this study is to use SAC to reproduce postural coordination modes in the case of different tracking frequencies.

5.2.1 Model and Method

5.2.1.1 Model

In this section, the model used for our study is a walker agent built in [69] and illustrated in Figure 5.7. This model has six joint angles including right ankle, left ankle, right knee, left knee, right hip, and left hip joint angles, and seven links such as 2 feet, 2 lower leg, 2 upper leg, and torso.

5.2.1.2 Tracking Trajectory

The trajectory to be tracked is designed with sine function as follows:

$$T_{traget} = Asin(wt + \phi) + D, \tag{5.2.1}$$

where, T_{target} is the tracking trajectory signal, A is the amplitude of tracking trajectory (here $A = \ 0.08$); w is the oscillation frequency related parameter (here $w = \ 2\pi f$, where f is the sine function oscillation frequency); ϕ denotes the initial phase of the sine function (here $\phi = arcsin(1/2)$); D is the deviation part (here $D = 0.04$). In this study, we will focus on the tracking frequency influence on multi-segment coordination of the walker model sway motion during back and forth tracking tasks.

5.2.1.3 Reward function

Our target is to let the agent tracking a swing point and keep the foot at its initial point. Therefore, the reward function $\mathbb{R}$ is composed of four terms as follows:

$$\mathbb{R} = \alpha R_s + \beta R_t + \gamma R_b - \delta R_f - \varepsilon R_c \tag{5.2.2}$$

R_s is used to help the agent learn how to stand up straight. It is related to the height of the CoM of the torso and the length of the projection of the torso on the z-axis (see Figure 5.8).

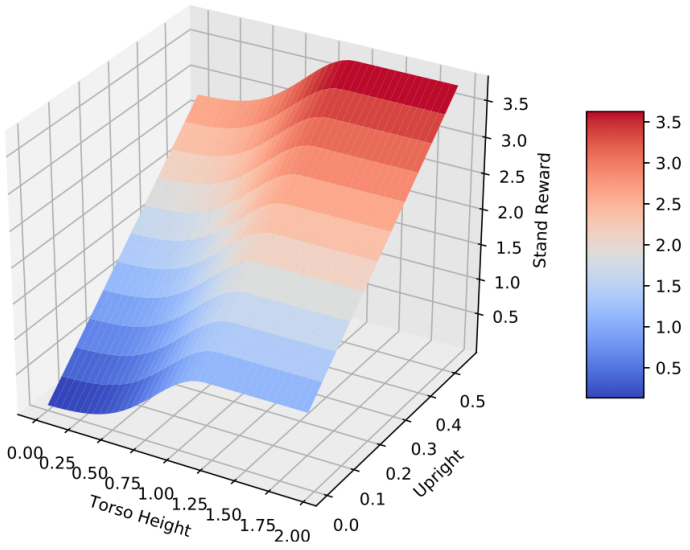


Figure 5.8: $3D$ view of the stand reward R_s, it is used to help the agent learn how to stand up straight. It is related to the height of the CoM of the torso and the length of the projection of the torso on the z-axis.

R_t is the reward obtained when the agent can track to the target point (see Figure 5.9).

R_b is used to limit the position of the agent's COM, making it easier to learn how to keep balance (see Figure 5.10).

R_f is used to let the agent keep both of its feet at their initial points (constrained feet positions). The initial position vectors of the left and right foot are denoted by $\vec{P_{L0}}$ and $\vec{P_{R0}}$. At every step, their position vectors become $\vec{P_L}$ and $\vec{P_R}$.

$$R_f = \|\vec{P_L} - \vec{P_{L0}}\|_2 + \|\vec{P_R} - \vec{P_{R0}}\|_2 \tag{5.2.3}$$

R_c is used to let the agent learn how to work efficiently. T_k is the torque of the motor in joint k.

$$R_c = \sum_{k=0}^{5} |T_k| \tag{5.2.4}$$

The process of the agent getting the reward is divided into two stages. When the training begins, the smaller but more readily available R_s and R_b can help the agent learn how to

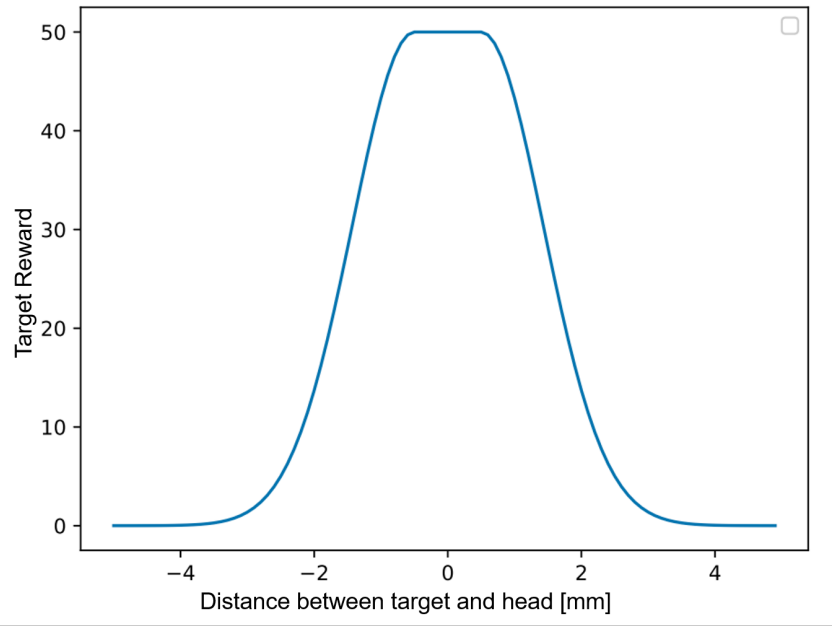


Figure 5.9: Evolution of the target reward R_t versus the distance between the target and head.

stand and keep balance first. Then, the agent has the chance to get the more challenging but more rewarding R_t.

5.2.2 Analysis of Postural Coordination Modes

To study the impact of tracking speed on the postural coordination transition, the target tracking frequency is set from 0.1 $[Hz]$ to 1.2 $[Hz]$. For each frequency, 6 seeds are trained for statistical reliability of training results. For each seed, 2 epochs are included. After training with the highest reward value, 20 epochs (trials) are repeated. Thus, we obtained 120-trial results for each frequency to analyze the transition of postural coordination modes.

Figure 5.18 shows the evolution of the target and head tracking trajectory for different tracking frequencies. As we can see here, the agent tracks the target with high accuracy. This indicates the agent has high motion adaptation after training. This is like the adaptation capacity of our humans for completing different level tasks. This in-phase motion between the target and the head illustrates the high and stable performance reproduced by the SAC learning algorithm.

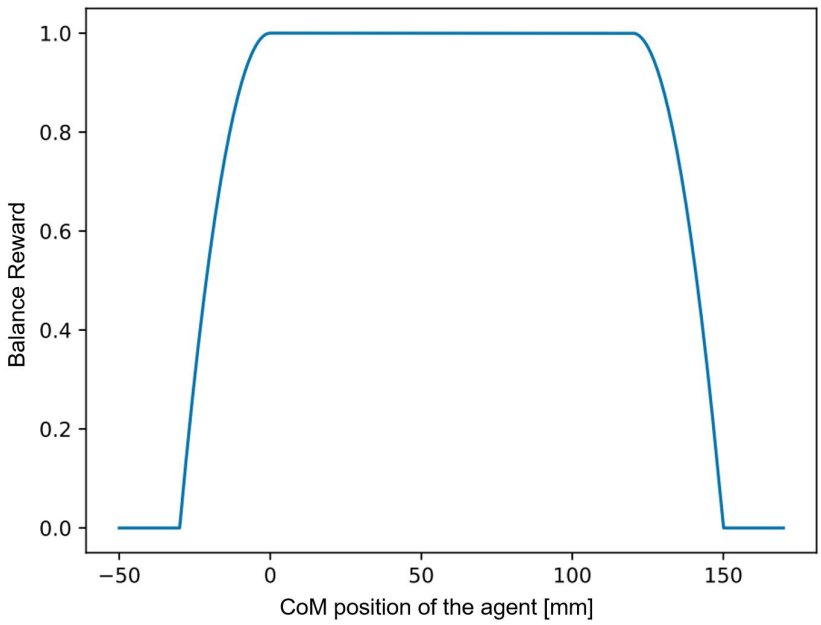


Figure 5.10: Evolution of the balance reward R_b versus the CoM postion of the agent.

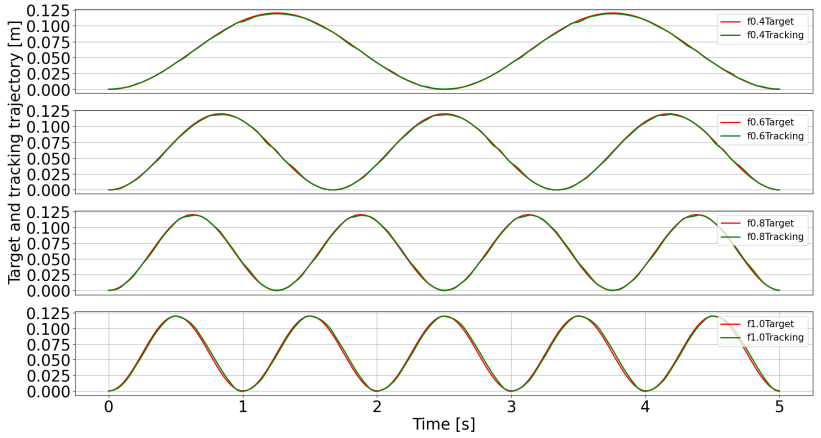


Figure 5.11: Position of tracking target and head for frequencies of 0.4, 0.6, 0.8, and 1.0 $[Hz]$.

Hip ankle joint correlation for frequency from 0.1 $[Hz]$ to 1.2 $[Hz]$ is shown in Figure 5.12. As we can observe here, for the tracking frequency ranges from 0.1[Hz] to 0.6[Hz]the positive ankle-hip correlation indicates the in-phase postural coordination. However, for the high tracking frequency which means high-speed body motions, ankle-hip joint coordination becomes negative. The postural transition between the ankle and hip switches into anti-phase. These phenomena have been observed from human experiments, And our contribution of this work is reproducing this type of human postural coordination modes through model-free DRL.

The ankle-hip phase portrait can also be seen as an evaluation index to analyze the

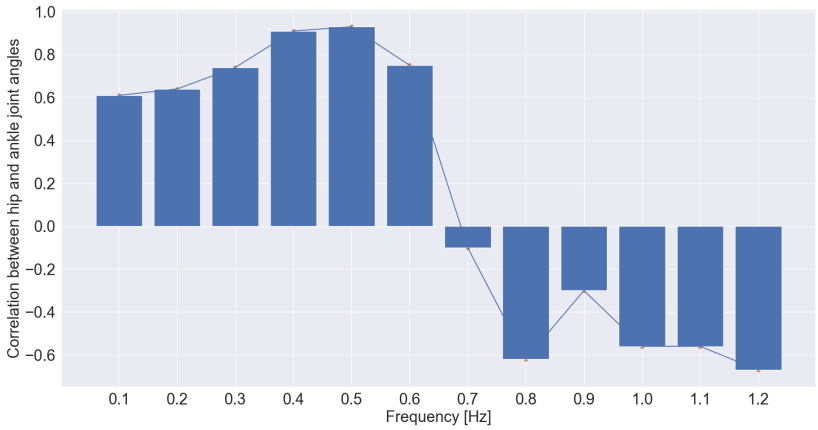


Figure 5.12: Transition of hip-ankle joint correlation for frequency from $0.1[Hz]$ to $1.2[Hz]$.

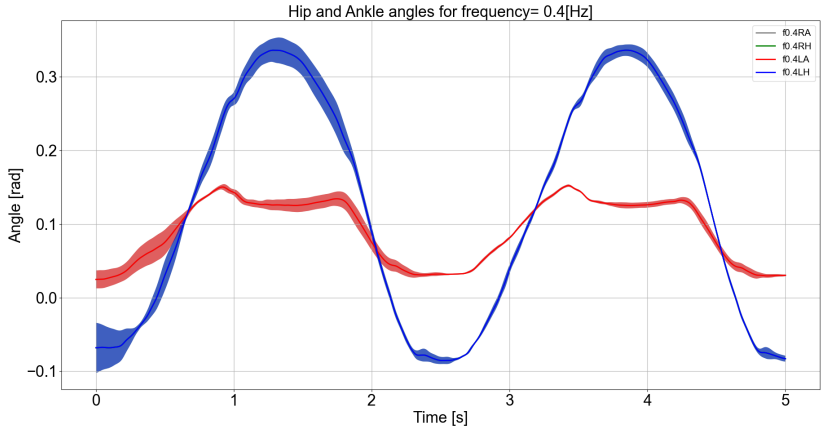


Figure 5.13: Evolution versus time of hip and ankle angles for a frequency of $0.4\ [Hz]$. 'RA', 'LA', 'RH', and 'LH' stand for abbreviations of right ankle, left ankle, right hip, and left hip, respectively.

coordination mode transitions. For instance, in Figure 5.17, the left figure shows the positive correlation, which means in-phase mode. And the right figure shows the negative correlation, which means anti-phase mode.

As mentioned previously, the in-phase and anti-phase postural transitions can be observed here. For frequencies 0.4 and 0.6 [Hz] (low-speed tracking), the agent performs in-phase coordination illustrated in Figures 5.13 and 5.14, while for frequencies 0.8 and 1.0 [Hz] (high-speed tracking), the agent tracks the target through anti-phase coordination mode illustrated in Figures 5.15 and 5.16. As the tracking target moves at a faster speed, the walker needs to well self-organize the body segments to follow the tracking trajectory. Thus, the anti-phase coordination between hip and ankle joint angles contributes more to tracking tasks for the high tracking frequency cases. This is reasonable because fast body movements may cause stable standing risks. To prevent this potential risks of falling down,

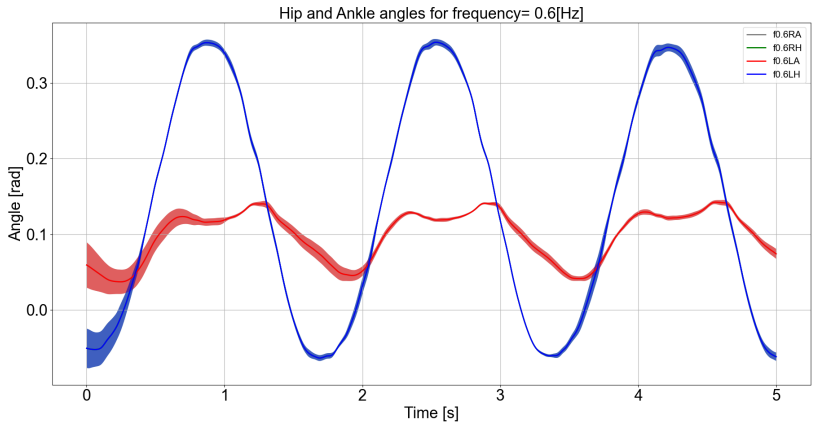


Figure 5.14: Evolution versus time of hip and ankle angles for a frequency of $0.6\ [Hz]$. 'RA', 'LA', 'RH', and 'LH' stand for abbreviations of right ankle, left ankle, right hip, and left hip, respectively.

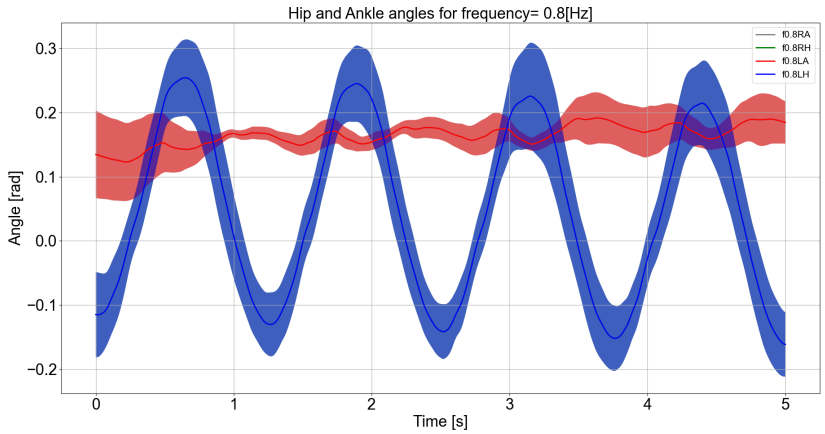


Figure 5.15: Evolution versus time of hip and ankle angles for a frequency of $0.8\ [Hz]$. 'RA', 'LA', 'RH', and 'LH' stand for abbreviations of right ankle, left ankle, right hip, and left hip, respectively.

the multi-segment coordination under anti-phase postural transition between ankle and hip is selected for the two tasks existing during the tracking process: (1) balance maintenance (Keep the body CoM above the base of support (BoS)); (2) tracking tasks with fast body movements.

According to the aforementioned multi-segment coordination analysis, the self-assembling, emergent phenomenon indicates the existence of multiple coordination modes and selective emergence. The in-phase and anti-phase transitions between hip and ankle joint angles change hierarchically based on the tracking frequency. This is not a sudden postural pattern switch, rather a more natural phenomenon as the frequency increases. It is worth noting that the frequency impact on postural coordination for human experiments and mode-based

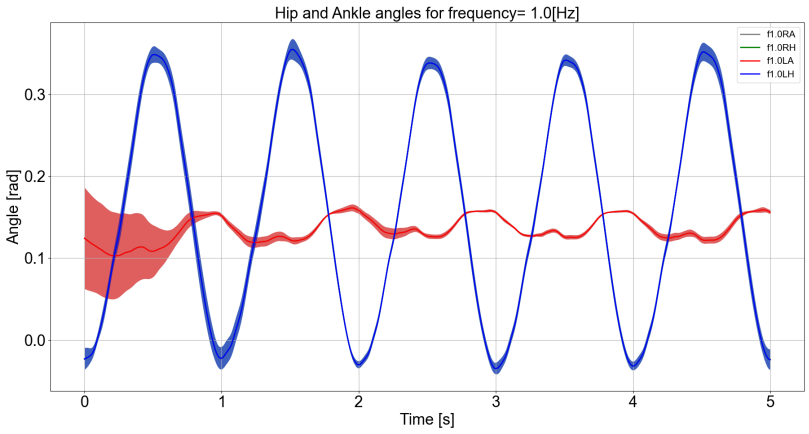


Figure 5.16: Evolution versus time of hip and ankle angles for a frequency of 1.0 $[Hz]$. 'RA', 'LA', 'RH', and 'LH' stand for abbreviations of right ankle, left ankle, right hip, and left hip, respectively.

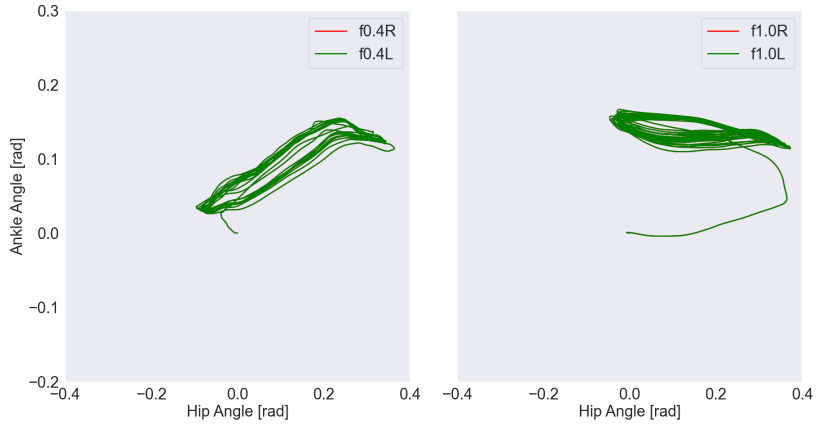


Figure 5.17: Hip ankle phase portraits for frequencies of 0.4 and 1.0 $[Hz]$.

simulation study in empirical work [122]–[124] is consistent with our study here.

The important remarked should be pointed out that multi-segment coordination can deal with balance stability maintenance as well as tracking tasks. The evolution of CoM for frequencies of 0.4, 0.6, 0.8, 1.0 $[Hz]$ is depicted in Figure 5.18. It is obvious to classify that in-phase postural coordination makes the large deviation of CoM for the frequencies 0.4 and 0.6 $[Hz]$ shown in Figure 5.18; while the anti-phase postural coordination decrease the deviation of CoM for the frequencies 0.8 and 1.0. Accordingly, anti-phase postural modes can reduce the CoM deviation. Especially, in the case of fast body movements, such as tracking the trajectory of sine oscillation at a fast speed (or a high frequency), anti-phase coordination exhibits the advantage both in completing high-speed tracking tasks and balance maintenance tasks. It is worth noting that the coordination mode transitions contribute to the decrease of CoM deviation. The emergent phenomena of

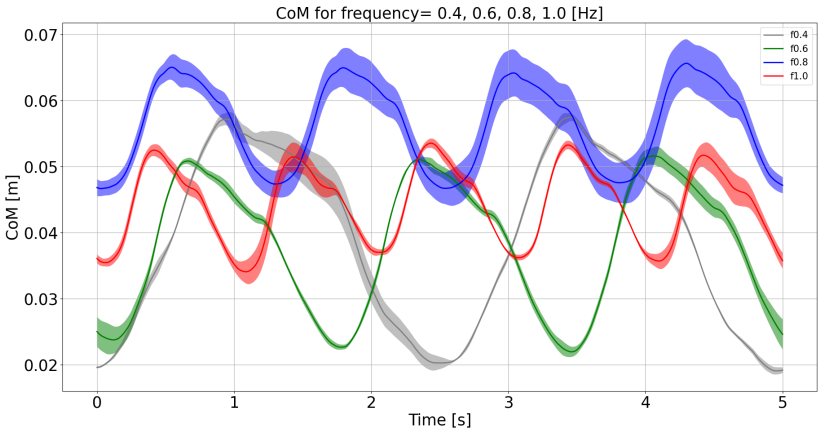


Figure 5.18: Evolution versus time of CoM for frequencies of 0.4, 0.6, 0.8 and 1.0 $[Hz]$.

postural coordination mode change from in-phase to anti-phase decrease the deviation of CoM. It is reasonable that for high-speed tracking tasks, the agent should reduce the CoM deviations for stable tracking performance.

5.3 Conclusions

In this chapter, we proposed a model-free learning method to study the multi-segment postural coordination existing in the head back and forth tracking tasks according to the different tracking frequencies. From the conducted analysis of kinematic results, the in-phase motion between the head and the target illustrates the robustness and good performance of the proposed SAC learning approach. Furthermore, the most important contribution of this study lies in the success of the walker model to track the target trajectory for the different frequencies through in-phase or anti-phase transitions between hip and ankle joint angles depending on the speed of the body movements. Finally, it is worth noting that with the same amplitude of tracking trajectory, the anti-phase postural coordination shows the advantage of balance maintenance through decreasing the deviation of CoM, compared with in-phase postural coordination.

Chapter 6

Conclusion and Future Work

6.1 Conclusions

In this thesis, we made effort to contribute to the balance recovery postural coordination reproduction through numerical simulations and the human supplementary experiments based on predictive and learning control strategies. The contributions of this thesis are summarized as follows.

In chapter 2, the double inverted pendulum without foot constraints has been proposed for the quiet standing balance recovery study considering the main quiet standing balance methods such as ankle strategy and hip-ankle strategy. And, the numerical linear model predictive control with the constraints of states and control has been designed to implement the ankle strategy and hip-ankle strategy for the different disturbing forces. Furthermore, many efficient dynamic indexes are used to evaluate the performance and mechanisms of different strategies.

Then, in chapter 3, the five-link, the three-joint simplified model including foot, low-body, upper-body, left-arm, right-arm segments and ankle, hip, arm joints has been built to study the arm usage in balance recovery control. The contributions of arm usage to balance recovery were verified with two methods: 1) arm swing type: active, passive, and fixed rotation; 2) different pushing forces. Furthermore, many kinematic and dynamic evaluation indexes are applied to compare which type of arm usage is efficient to balance

maintenance and which level of disturbing forces reaches the maximum limit of balance capacity. Besides, the influence of ankle capacity reduction on active arm usage has been discussed as well, which indicates the effectiveness of active arm usages in the cases of loss of ankle capacity. Finally, active arm usage has been concluded as the most efficient strategy for quiet standing balance recovery.

In chapter 4, postural spatial motor synergy is extracted from the kinematic results of human experiments, where balance recovery strategies such as ankle strategy, hip-ankle strategy, and arm strategy have been classified for the different levels of push forces. In the second part, we discussed the synergistic patterns extracted from the simulation results of balance recovery with the active arm strategy and analyzed the multi-joint coordination from postural and cybernetic aspects respectively through both spatial synergy and spatiotemporal synergy. And, we concluded that active arm usages could be an efficient solution to produce the synergistic patterns naturally. The proposed NMPC for reproducing active arm strategy exhibited strong robustness considering the mass estimation uncertainties with Gaussian distribution.

In chapter 5, we proposed a model-free learning method to study the multi-segment postural coordination existing in the head back and forth tracking tasks according to the different tracking frequencies. From the analysis of kinematic results, the high correlation between the torso CoM and the tracking position indicates the robustness and excellent performance of the SAC learning approach. Furthermore, the most important contribution of this study is that the model succeeds to track the target trajectory for the different frequencies through the in-phase or anti-phase transition between hip and ankle joint angles depending on the speed of the body movements. It is worth noting that with the same amplitude of tracking trajectory, the anti-phase postural coordination shows the advantage of balance maintenance through decreasing the deviation of CoM comparing with in-phase postural coordination.

6.2 Future Work

We successfully reproduced synergetic human balance recovery modalities through predictive and learning control strategies. However, only balance recovery motions in the sagittal

plane are considered. In the future, we will try to develop the controller to reproduce the human balance recovery motion to understand the mechanisms of the corresponding postural coordination of the multi-segment human body. Besides, we study the coordinated motion patterns for more subjects and with more different conditions to further understand the relationship between the extracted spatial synergies and balance postural control strategies. Finally, we further apply our control theory to design a new exoskeleton device to help the patient support their balance and develop the balance rehabilitation training strategies.

List of Publications

- Journal Papers

  - K. Shen, A. Chemori, and M. Hayashibe. "Reproducing Human Arm Strategy and its Contribution to Balance Recovery Through Model Predictive Control", *Frontiers in Neurorobotics*, April, 2021, doi: 10.3389/fnbot.2021.679570.
  - K. Shen, A. Chemori, and M. Hayashibe. "Human-like Balance Recovery Based on Numerical Model Predictive Control Strategy", *IEEE Access*, vol.8, pp.92050 – 92060, May, 2020.

- International Conference Papers

  - K. Shen, A. Chemori, and M. Hayashibe. "Effectiveness Evaluation of Arm Usage for Human Quiet Standing Balance Recovery through Nonlinear Model Predictive Control". *International Conference on Control and Robots (ICCR2020)*, Tokyo, Japan, Dec.26 - 28, 2020.

Bibliography

[1] U. Nations, "Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Volume II: Demographic Profiles," 2019.

[2] WHO, "https://www.who.int/initiatives/rehabilitation-2030," 2017.

[3] K. Shen, A. Chemori, and M. Hayashibe, "Human-Like Balance Recovery Based on Numerical Model Predictive Control Strategy," *IEEE Access*, vol. 8, pp. 92 050–92 060, 2020, ISSN: 21693536. DOI: 10.1109/ACCESS.2020.2995104.

[4] M. Vukobratovic, A. A. Frank, and D. Juricic, "on the Stability of Biped Locomotion," *IEEE Transactions on Biomedical Engineering*, vol. BME-17, no. 1, pp. 25–36, 1970, ISSN: 00189294.

[5] M. Vukobratovic and B. Borovac, "Zero-Moment Point — Thirty Five Years of Its Life," *International Journal of Humanoid Robotics*, vol. 01, no. 01, pp. 157–173, 2004, ISSN: 0219-8436. DOI: 10.1142/s0219843604000083.

[6] H. Hemami and P. C. Camana, "Nonlinear Feedback in Simple Locomotion Systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 6, pp. 855–860, 1976, ISSN: 15582523. DOI: 10.1109/TAC.1976.1101381.

[7] R. E. Goddard, H. Hemami, and F. C. Weimer, "Biped Side Step in the Frontal Plane," *IEEE Transactions on Automatic Control*, vol. 28, no. 2, pp. 179–187, 1983, ISSN: 15582523. DOI: 10.1109/TAC.1983.1103196.

[8] P. Gatev, S. Thomas, T. Kepple, and M. Hallett, "Feedforward ankle strategy of balance during quiet stance in adults," *Journal of Physiology*, vol. 514,

no. 3, pp. 915–928, 1999, ISSN: 00223751. DOI: 10.1111/j.1469-7793.1999.915ad.x.

[9] F. B. Horak and L. M. Nashner, "Central programming of postural movements: Adaptation to altered support-surface configurations," *Journal of Neurophysiology*, vol. 55, no. 6, pp. 1369–1381, 1986, ISSN: 00223077. DOI: 10.1152/jn.1986.55.6.1369.

[10] L. M. Nashner, C. L. Shupert, F. B. Horak, and F. O. Black, "Organization of posture controls: An analysis of sensory and mechanical constraints," *Progress in Brain Research*, vol. 80, no. C, pp. 411–418, 1989, ISSN: 18757855. DOI: 10.1016/S0079-6123(08)62237-2.

[11] A. D. Kuo and F. E. Zajac, "Human standing posture: Multi-joint movement strategies based on biomechanical constraints," *Progress in Brain Research*, vol. 97, no. C, pp. 349–358, 1993, ISSN: 18757855. DOI: 10.1016/S0079-6123(08)62294-3.

[12] A. D. Kuo, "An Optimal Control Model for Analyzing Human Postural Balance," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 1, pp. 87–101, 1995, ISSN: 15582531. DOI: 10.1109/10.362914.

[13] D. Winter, "Human balance and posture control during standing and walking," *Gait and Posture*, vol. 3, no. 4, pp. 193–214, 1995, ISSN: 0966-6362. DOI: https://doi.org/10.1016/0966-6362(96)82849-9. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0966636296828499.

[14] 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, ISSN: 00223077. DOI: 10.1152/jn.1996.75.6.2334.

[15] B. L. Day, M. J. Steiger, P. D. Thompson, and C. D. Marsden, "Effect of vision and stance width on human body motion when standing: implications for afferent control of lateral sway.," *The Journal of Physiology*, vol. 469, no. 1, pp. 479–499, 1993, ISSN: 14697793. DOI: 10.1113/jphysiol.1993.sp019824.

[16] S. I. A. Ishida, "Responses of the posture-control system to psedorandom acceleration disturbances," *Med. and Biol. Eng. and Compu*, vol. 18, pp. 433–438, 1980.

[17] R. Fitzpatrick, D. Burke, and S. C. Gandevia, "Loop gain of reflexes controlling human standing measured with the use of postural and vestibular disturbances," *Journal of Neurophysiology*, vol. 76, no. 6, pp. 3994–4008, 1996, ISSN: 00223077. DOI: 10.1152/jn.1996.76.6.3994.

[18] Y. Li, W. S. Levine, and G. E. Loeb, "A two-joint human posture control model with realistic neural delays," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 738–748, 2012, ISSN: 15344320. DOI: 10.1109/TNSRE.2012.2199333.

[19] Y. C. Pai and J. Patton, "Center of mass velocity-position predictions for balance control," *Journal of Biomechanics*, vol. 30, no. 4, pp. 347–354, 1997, ISSN: 00219290. DOI: 10.1016/S0021-9290(96)00165-0.

[20] S. Kajita, F. Kanehiro, K. Kaneko, and K. Fujiwara, "Biped Walking Pattern Generation by using Preview," *IEEE International Conference on Robotics and Automation*, pp. 1620–1626, 2003.

[21] C. Azevedo, P. Poignet, and B. Espiau, "Artificial locomotion control: From human to robots," *Robotics and Autonomous Systems*, vol. 47, no. 4, pp. 203–223, 2004, ISSN: 09218890. DOI: 10.1016/j.robot.2004.03.013.

[22] A. Hofmann, "Robust execution of bipedal walking tasks from biomechanical principles," *PhD Thesis*, 2006.

[23] C. G. Atkeson and B. Stephens, "Multiple balance strategies from one optimization criterion," *Proceedings of the 2007 7th IEEE-RAS International Conference on Humanoid Robots, HUMANOIDS 2007*, pp. 57–64, 2007. DOI: 10.1109/ICHR.2007.4813849.

[24] B. Stephens, "Humanoid Push Recovery," *2007 7th IEEE-RAS International Conference on Humanoid Robots*, pp. 589–595, 2007. DOI: 10.1109/ICHR.2007.4813931.

[25] B. J. Stephens and C. G. Atkeson, "Push recovery by stepping for humanoid robots with force controlled joints," *2010 10th IEEE-RAS International Conference on Humanoid Robots, Humanoids 2010*, pp. 52–59, 2010. DOI: 10.1109/ICHR.2010.5686288.

[26] B. J. Stephens, "Push Recovery Control for Force-Controlled Humanoid Robots," *PhD Theis*, no. June, 2011.

[27] C. Liu and C. G. Atkeson, "Standing balance control using a trajectory library," *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 3031–3036, 2009. DOI: 10.1109/IROS.2009.5354018.

[28] S. Kiemel, "Balance maintenance of a humanoid robot using the hip-ankle strategy," *Master Thesis*, 2012.

[29] D. N. Nenchev and A. Nishio, "Ankle and hip strategies for balance recovery of a biped subjected to an impact," *Robotica*, vol. 26, no. 5, pp. 643–653, 2008, ISSN: 02635747. DOI: 10.1017/S0263574708004268.

[30] Z. Aftab, T. Robert, and P. B. Wieber, "Simulating the effect of upper-body inertia on human balance recovery.," *Computer methods in biomechanics and biomedical engineering*, vol. 15 Suppl 1, no. September, pp. 148–150, 2012, ISSN: 14768259. DOI: 10.1080/10255842.2012.713653.

[31] A. Zohaib, T. Robert, and P. B. Wieber, "Ankle, hip and stepping strategies for humanoid balance recovery with a single Model Predictive Control scheme," *IEEE-RAS International Conference on Humanoid Robots*, pp. 159–164, 2012, ISSN: 21640572. DOI: 10.1109/HUMANOIDS.2012.6651514.

[32] T. R. Zohaib Aftab and P.-B. Wieber, "Balance recovery prediction with multiple strategies for standing humans," *PLoS ONE*, vol. 11, no. 3, pp. 1–16, 2016, ISSN: 19326203. DOI: 10.1371/journal.pone.0151166.

[33] D. S. Marigold and A. E. Patla, "Strategies for dynamic stability during locomotion on a slippery surface: Effects of prior experience and knowledge," *Journal of Neurophysiology*, vol. 88, no. 1, pp. 339–353, 2002, ISSN: 00223077. DOI: 10.1152/jn.00691.2001.

[34] H. Choi, S. Lee, T. Jin, and S. H. Lee, "Trajectory-free reactive stepping of humanoid robots using momentum control," *IEEE-RAS International Conference on Humanoid Robots*, vol. 2015-Decem, pp. 1173–1178, 2015, ISSN: 21640580. DOI: 10.1109/HUMANOIDS.2015.7363511.

[35] M. Shafiee-Ashtiani, A. Yousefi-Koma, M. Shariat-Panahi, and M. Khadiv, "Push recovery of a humanoid robot based on model predictive control and capture point," *4th RSI International Conference on Robotics and Mechatronics, ICRoM 2016*, pp. 433–438, 2017. DOI: 10.1109/ICRoM.2016.7886777. eprint: 1612.08034.

[36] L. Peneo, B. Clement, V. Moduano, E. Mingo Hoffman, G. Nava, D. Pucci, N. G. Tsagarakis, J. B. Mourert, and S. Ivaldi, "Robust Real-Time Whole-Body Motion Retargeting from Human to Humanoid," *IEEE-RAS International Conference on Humanoid Robots*, vol. 2018-Novem, pp. 425–432, 2019, ISSN: 21640580. DOI: 10.1109/HUMANOIDS.2018.8624943.

[37] K. Yamamoto, "Time-variant feedback controller based on capture point and maximal output admissible set of a humanoid," *Advanced Robotics*, vol. 33, no. 18, pp. 944–955, 2019, ISSN: 15685535. DOI: 10.1080/01691864.2019.1633403. [Online]. Available: https://doi.org/01691864.2019.1633403.

[38] L. M. Nashner, "The organisation of human postural movements A formal basis and experimental synthesis..pdf," *The Behavior and Brain Science*, no. 8, pp. 135–172, 1985.

[39] F. B. Horak, L. M. Nashner, and H. C. Diener, "Postural strategies associated with somatosensory and vestibular loss," *Experimental Brain Research*, vol. 82, no. 1, pp. 167–177, 1990, ISSN: 00144819. DOI: 10.1007/BF00230848.

[40] J. R. Chagdes, S. Rietdyk, M. H. Jeffrey, N. Z. Howard, and A. Raman, "Dynamic stability of a human standing on a balance board," *Journal of Biomechanics*, vol. 46, no. 15, pp. 2593–2602, 2013. DOI: 10.1016/j.jbiomech.2013.08.012.

[41] E. Chumacero, J. Yang, and J. R. Chagdes, "Effect of sensory-motor latencies and active muscular stiffness on stability for an ankle-hip model of balance on a balance board," *Journal of Biomechanics*, vol. 75, pp. 77–88, 2018. DOI: 10.1016/j.jbiomech.2018.04.045. [Online]. Available: https://doi.org/10.1016/j.jbiomech.2018.04.045.

[42] E. Chumacero, J. Yang, and J. Chagdes, "Numerical nonlinear analysis for dynamic stability of an ankle-hip model of balance on a balance board," *ASME Journal of Computational and Nonlinear Dynamics*, vol. 14, no. 10, pp. 1–11, 2019. DOI: 10.1115/1.4042693.

[43] P. E. Roos, M. P. McGuigan, D. G. Kerwin, and G. Trewartha, "The role of arm movement in early trip recovery in younger and older adults," *Gait and Posture*, vol. 27, no. 2, pp. 352–356, 2008, ISSN: 09666362. DOI: 10.1016/j.gaitpost.2007.05.001.

[44] M. Pijnappels, I. Kingma, D. Wezenberg, G. Reurink, and J. H. Van Dieën, "Armed against falls: The contribution of arm movements to balance recovery after tripping," *Experimental Brain Research*, vol. 201, no. 4, pp. 689–699, 2010, ISSN: 00144819. DOI: 10.1007/s00221-009-2088-7.

[45] P. J. Cordo and L. M. Nashner, "Properties of postural adjustments associated with rapid arm movements," *Journal of Neurophysiology*, vol. 47, no. 2, pp. 287–302, 1982, ISSN: 00223077. DOI: 10.1152/jn.1982.47.2.287.

[46] A. Ledebt, "Changes in arm posture during the early acquisition of walking," *Infant Behavior and Development*, vol. 23, no. 1, pp. 79–89, 2000, ISSN: 01636383. DOI: 10.1016/S0163-6383(00)00027-8.

[47] Y. Aoustin, M. F. Skii, Y. Aoustin, and A. M. Formal, "On optimal swinging of the biped arms To cite this version : HAL Id : hal-00462472 On optimal swinging of the biped arms," *HAL Id: hal-00462472 https://hal.archives-ouvertes.fr/hal-00462472*, 2008.

[48] M. Nakada, B. Allen, S. Morishima, and D. Terzopoulos, "Learning arm motion strategies for balance recovery of humanoid robots," *Proceedings - EST 2010 - 2010 International Conference on Emerging Security Tech-*

nologies, ROBOSEC 2010 - Robots and Security, LAB-RS 2010 - Learning and Adaptive Behavior in Robotic Systems, pp. 165–170, 2010. DOI: 10.1109/EST.2010.18.

[49] S. M. Bruijn, O. G. Meijer, P. J. Beek, and J. H. Van Dieën, "The effects of arm swing on human gait stability," *Journal of Experimental Biology*, vol. 213, no. 23, pp. 3945–3952, 2010, ISSN: 00220949. DOI: 10.1242/jeb.045112.

[50] M. Milosevic, K. M. McConville, and K. Masani, "Arm movement improves performance in clinical balance and mobility tests," *Gait and Posture*, vol. 33, no. 3, pp. 507–509, 2011, ISSN: 09666362. DOI: 10.1016/j.gaitpost.2010.12.005. [Online]. Available: http://dx.doi.org/10.1016/j.gaitpost.2010.12.005.

[51] K. J. Boström, T. Dirksen, K. Zentgraf, and H. Wagner, "The contribution of upper body movements to dynamic balance regulation during challenged locomotion," *Frontiers in Human Neuroscience*, vol. 12, no. January, pp. 1–10, 2018, ISSN: 16625161. DOI: 10.3389/fnhum.2018.00008.

[52] C. N. Objero, M. M. Wdowski, and M. W. Hill, "Can arm movements improve postural stability during challenging standing balance tasks?" *Gait and Posture*, vol. 74, pp. 71–75, 2019, ISSN: 18792219. DOI: 10.1016/j.gaitpost.2019.08.010.

[53] K. Shen, A. Chemori, and M. Hayashibe, "Effectiveness Evaluation of Arm Usage for Human Quiet Standing Balance Recovery through Nonlinear Model Predictive Control," *3rd International Conference on Control and Robots, ICCR 2020*, pp. 150–153. 2020. DOI: 10.1109/ICCR51572.2020.9344184.

[54] J. H. Allum, F. Honegger, and C. Pfaltz, "The role of stretch and vestibulospinal reflexes in the generation of human equilibrating reactions," *Prog Brain Res.*, vol. 80, pp. 399–409, 1989. DOI: 10.1016/s0079-6123(08)62236-0.

[55] N. Bernstein, "The Co-ordination and Regulation of Movement," *Oxford*, vol. 35, 1967.

[56] W. Hsu, L. Chou, and M. Woollacott, "Age-related changes in joint coordination during balance recovery," *Age*, vol. 35, pp. 1299–1309, 2013. DOI: 10.1007/s11357-012-9422-x.

[57] W. Lee, "Neuromotor synergies as a basis for coordinated intentional action," *Journal of Motor Behavior*, vol. 16, pp. 135–170, 1984. DOI: 10.1080/00222895.1984.10735316.

[58] E. Chiovetto, A. d'Avella, and M. Giese, "A Unifying Framework for the Identification of Motor Primitives," *arXiv:1603.06879*, 2016. DOI: 10.1080/00222895.1984.10735316.

[59] P. Haggard, K. Hutchinson, and J. Stein, "Patterns of coordinated multi-joint movement," *Journal of Motor Behavior*, vol. 107(2), pp. 254–266, 1995. DOI: 10.1007/BF00230046.

[60] J. Chai and M. Hayashibe, "Motor Synergy Development in High performing Deep Reinforcement Learning algorithms," *IEEE Robotics and Automation Letters*, vol. 5, pp. 1271–1278, 2020. DOI: 10.1109/LRA.2020.2968067..

[61] N. Yang, Q. An, H. Yamakawa, A. Tamura Y.and Yamashita, and H. Asama, "Muscle synergy structure using different strategies in human standing-up motion," *Advanced Robotics*, vol. 31, pp. 40–54, 2017. DOI: 10.1080/01691864.2016.1238781.

[62] H. Yamasaki and S. Shimoda, "Spatiotemporal modular organization of muscle torques for sit-to-stand movements," *Journal of Biomechanics*, vol. 49(14), pp. 3268–3274, 2016. DOI: 10.1016/j.jbiomech.2016.08.010.

[63] F. Ildar, E. Julia, V. O. Gijs, V. Mark, V. A. Edwin, and B. Etienne, "Assisting Human Balance in Standing With a Robotic Exoskeleton," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 414–421, 2018, ISSN: 23773766.

[64] S. Burmaoglu, V. Trajkovik, T. L. Tutukalo, H. Yalcin, and B. Caulfield, *Evolution map of wearable technology patents for healthcare field*. Elsevier Inc., 2018, pp. 275–290.

[65] A. S. Gorgey, "Robotic exoskeletons: The current pros and cons," *World Journal of Orthopaedics*, vol. 9, no. 9, pp. 112–119, 2018. DOI: 10.5312/wjo.v9.i9.112.

[66] K. Gordt, T. Gerhardy, B. Najafi, and M. Schwenk, "Effects of Wearable Sensor-Based Balance and Gait Training on Balance, Gait, and Functional Performance in Healthy and Patient Populations: A Systematic Review and Meta-Analysis of Randomized Controlled Trials," *Gerontology*, vol. 64, no. 1, pp. 74–89, 2017. DOI: 10.1159/000481454.

[67] G. Torres-Oviedo and L. H. Ting, "Muscle synergies characterizing human postural responses," *Journal of Neurophysiology*, vol. 98, no. 4, pp. 2144–2156, 2007, ISSN: 00223077.

[68] E. Todorov, T. Erez, and Y. Tassa, "MuJoCo: A physics engine for model-based control," *IEEE International Conference on Intelligent Robots and Systems*, pp. 5026–5033, 2012, ISSN: 21530858. DOI: 10.1109/IROS.2012.6386109.

[69] Y. Tassa, S. Tunyasuvunakool, A. Muldal, Y. Doron, S. Liu, S. Bohez, J. Merel, T. Erez, T. Lillicrap, and N. Heess, "Dm control: Software and tasks for continuous control," 2020. arXiv: 2006.12983.

[70] L. Grüne and J. Pannek, *Nonlinear Model Predictive Control Theory and Algorithms*. 2017, ISBN: 9783319460239. DOI: 10.1007/978-3-319-46024-6. [Online]. Available: http://www.springer.com/series/61.

[71] H. I. M. Kouchi M. Mochimaru and S. Mitani, "Anthropometric database for Japanese Population 1997-98," 2000.

[72] H. Chen and N. Sun, "Nonlinear Control of Underactuated Systems Subject to Both Actuated and Unactuated State Constraints with Experimental Verification," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 9, pp. 7702–7714, 2020, ISSN: 15579948. DOI: 10.1109/TIE.2019.2946541.

[73] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, "Robotics Modelling Planning and Control," *London, U.K.:Springer-Verlag*, 2009.

[74] M. W. Spong, S. Hutchinson, and M. Vidyasagar, "Robot Dynamics and Control," *Hoboken, NJ, USA: Wiley*, 2004.

[75] S. V. Raković and W. S. Levine, "Handbook Model Predictive Control," *Basel, Switzerland:Birkhäuser*, 2018.

[76] P. Lancaster and L. Rodman, "Algebraic Riccati Equation," *Oxford, U.K.:Oxford Univ. Press*, 1995.

[77] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000, ISSN: 00051098. DOI: 10.1016/S0005-1098(99)00214-9.

[78] H. Chen and F. Allgöwer, "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability," *ECC 1997 - European Control Conference*, no. July, pp. 1421–1426, 1997. DOI: 10.23919/ecc.1997.7082300.

[79] H. Chen and F. Allgöwer, "Nonlinear Model Predictive Control Schemes with Guaranteed Stability," *Nonlinear Model Based Process Control*, pp. 465–494, 1998. DOI: 10.1007/978-94-011-5094-1_16.

[80] V. Sairam and K. Vikram, "Stability analysis for linear/nonlinear model predictive control of constrained processes," *Proceedings of the American Control Conference*, vol. 3, no. June, pp. 1679–1683, 1998, ISSN: 07431619. DOI: 10.1109/ACC.1998.707292.

[81] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019, ISSN: 18672957. DOI: 10.1007/s12532-018-0139-4. [Online]. Available: https://doi.org/10.1007/s12532-018-0139-4.

[82] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, "Model Predictive Control: Theory Computation and Design," *Santa Barbara, CA, USA:Nob Hill Publishing*, 2019.

[83] C. F. Runge, C. L. Shupert, F. B. Horak, and F. E. Zajac, "Ankle and hip postural strategies defined by joint torques," *Gait and Posture*, vol. 10, no. 2, pp. 161–170, 1999, ISSN: 09666362. DOI: 10.1016/S0966-6362(99)00032-6.

[84] B. Colobert, A. Crétual, P. Allard, and P. Delamarche, "Force-plate based computation of ankle and hip strategies from double-inverted pendulum model," *Clinical Biomechanics*, vol. 21, no. 4, pp. 427–434, 2006, ISSN: 02680033. DOI: 10.1016/j.clinbiomech.2005.12.003.

[85] E. R. D. Gordon, C. Graham E., H. Joseph, K. Gary, and N. Saunders. Human Kinetics, 2014.

[86] A. Rajagopal, C. L. Dembia, M. S. DeMers, D. D. Delp, J. L. Hicks, and S. L. Delp, "Full-Body Musculoskeletal Model for Muscle-Driven Simulation of Human Gait," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 10, pp. 2068–2079, 2016, ISSN: 15582531. DOI: 10.1109/TBME.2016.2586891.

[87] R. P. Paul, *Robot Manipulators: Mathematics, Programming and Control*. 1981, ISBN: 9780262160827.

[88] E. Chumacero and J. Yang, "Basin of attraction and limit cycle oscillation amplitude of an ankle-hip model of balance on a balance board," *ASME Journal of Biomechanical Engineering*, vol. 141, no. 11, pp. 1–9, 2019. DOI: 10.1115/1.4043563.

[89] E. Chumacero and J. Yang, "Effect of disturbances and sensorimotor deficits on the postural robustness of an ankle–hip model of balance on a balance board," *Nonlinear Dynamics*, vol. 99, no. 3, pp. 1959–1973, 2020. DOI: 10.1007/s11071-019-05403-w.

[90] S. Kuindersma, R. Grupen, and A. Barto, "Learning dynamic arm motions for postural recovery," *IEEE-RAS International Conference on Humanoid Robots*, pp. 7–12, 2011, ISSN: 21640572. DOI: 10.1109/Humanoids.2011.6100881.

[91] M. L. Latash and V. M. Zatsiorsky, *Biomechanics and motor control defining central concepts*. 2016, ISBN: 9780128003848. DOI: 10.1016/C2013-0-18342-0.

[92] D. Fong, Y. Hong, L. Chan, P. Yung, and K. Chan, "A systematic review on ankle injury and ankle sprain in sports," *Sports Med*, vol. 37(1), pp. 73–94, 2007. DOI: 10.2165/00007256-200737010-00006.

[93] H. Negahban, A. Moradi-Bousari, S. Naghibi, J. Sarrafzadeh, M. Shaterzadeh-Yazdi, S. Goharpey, M. Etemadi, M. Mazaheri, and A. Feizi, "The eccentric torque production capacity of the ankle, knee, and hip muscle groups in patients with unilateral chronic ankle instability," *Asian journal of sports medicine*, vol. 4(2), pp. 144–152, 2013. DOI: 4(2),144152//doi.org/10.5812/asjsm.34515.

[94] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions A Royal Society*, 2016. DOI: 10.1098/rsta.2015.0202.

[95] A. F. Plates, DOI: website:https://www.amti.biz.

[96] 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," *PLOS Computational Biology*, vol. 14, no. 7, pp. 1–20, Jul. 2018. DOI: 10.1371/journal.pcbi.1006223.

[97] P. Stanley, E. F. Gaynor, and A. Thomas, "Anatomical Data for Analyzing Human Motion," *Research Quarterly for Exercise and Sport*, vol. 54, no. 2, pp. 169–178, 1983. DOI: 10.1080/02701367.1983.10605290.

[98] C. Jacob, C. Patricia, G. W. Stephen, and S. A. Leona, "Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences," *LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS*, 2003.

[99] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[100] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "OpenAI Gym," pp. 1–4, 2016. arXiv: 1606.01540. [Online]. Available: http://arxiv.org/abs/1606.01540.

[101] T. Haarnoja, V. Pong, A. Zhou, M. Dalal, P. Abbeel, and S. Levine, "Composable Deep Reinforcement Learning for Robotic Manipulation," *Proceedings - IEEE International Conference on Robotics and Automation*, no. 1, pp. 6244–6251, 2018, ISSN: 10504729. DOI: 10.1109/ICRA.2018.8460756. arXiv: 1803.06773.

[102] L. Zhongyu, C. Xuxin, P. Xuebin, A. Pieter, L. Sergey, B. Glen, and S. Koushil, "Reinforcement learning for robust parameterized locomotion control of bipedal robots," *2021 International Conference on Robotics and Automation (ICRA 2021)*, 2021. DOI: arXiv:2103.14295.

[103] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations," *arXiv*, 2017, ISSN: 23318422. DOI: 10.15607/rss.2018.xiv.049. arXiv: 1709.10087.

[104] OpenAI, M. Andrychowicz, B. Baker, M. Chociej, R. Józefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, and W. Zaremba, "Learning dexterous in-hand manipulation," *CoRR*, 2018. [Online]. Available: http://arxiv.org/abs/1808.00177.

[105] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015, ISSN: 14764687.

DOI: 10.1038/nature14236. [Online]. Available: http://dx.doi.org/10.1038/nature14236.

[106] T. V. Maia, "Reinforcement learning, conditioning, and the brain: Successes and challenges," *Cognitive, Affective and Behavioral Neuroscience*, vol. 9, no. 4, pp. 343–364, 2009, ISSN: 15307026. DOI: 10.3758/CABN.9.4.343.

[107] J. Hu, H. Niu, J. Carrasco, B. Lennox, and F. Arvin, "Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14 413–14 423, 2020. DOI: 10.1109/TVT.2020.3034800.

[108] "Https://spinningup.openai.com/en/latest/index.html,"

[109] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017, ISSN: 10535888. DOI: 10.1109/MSP.2017.2743240. arXiv: arXiv:1708.05866v2.

[110] V. François-lavet, P. Henderson, R. Islam, M. G. Bellemare, V. François-lavet, J. Pineau, and M. G. Bellemare, "An Introduction to Deep Reinforcement Learning. (arXiv:1811.12560v1 [cs.LG]) http://arxiv.org/abs/1811.12560," *Foundations and trends in machine learning*, vol. II, no. 3 - 4, pp. 1–140, 2018, ISSN: 19358245. DOI: 10.1561/2200000071.Vincent. arXiv: arXiv:1811.12560v1.

[111] D. Katić, "Reinforcement learning in humanoid robotics," *8th Seminar on Neural Network Applications in Electrical Engineering, Neurel-2006 Proceedings*, p. 85, 2006. DOI: 10.1109/NEUREL.2006.341182.

[112] "Reinforcement learning in robotics: A survey," *International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013, ISSN: 02783649. DOI: 10.1177/0278364913495721.

[113] M. P. Deisenroth, "A Survey on Policy Search for Robotics," *Foundations and Trends in Robotics*, vol. 2, no. 1-2, pp. 1–142, 2011, ISSN: 1935-8253. DOI: 10.1561/2300000021.

[114] X. B. Peng and M. van de Panne, "Learning locomotion skills using deep RL: Does the choice of action space maer?" *Proceedings - SCA 2017: ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, no. 1, 2017. DOI: 10.1145/3099564.3099567. arXiv: 1611.01055.

[115] X. B. Peng, Z. Ma, P. Abbeel, S. Levine, and A. Kanazawa, "AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control," vol. 40, no. 4, pp. 1–20, 2021. DOI: 10.1145/3450626.3459670. arXiv: 2104.02180.

[116] L. Liu, M. Van De Panne, and K. Yin, "Guided learning of control graphs for physics-based characters," *ACM Transactions on Graphics*, vol. 35, no. 3, 2016, ISSN: 15577368. DOI: 10.1145/2893476.

[117] R. Haarnoja, K. Hartikainen, P. Abbeel, and S. Levine, "Latent space policies for hierarchical reinforcement learning," *35th International Conference on Machine Learning, ICML 2018*, vol. 4, pp. 2965–2975, 2018. arXiv: 1804.02808.

[118] W. Konen and T. Bartz-Beielstein, "Reinforcement learning for games," p. 2641, 2009. DOI: 10.1145/1570256.1570375.

[119] T. Haarnoja, A. Zhou, K. Hartikainen, G. Tucker, S. Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel, and S. Levine, "Soft Actor-Critic Algorithms and Applications," *arXiv*, 2018. arXiv: 1812.05905.

[120] T. Haarnoja, H. Tang, P. Abbeel, and S. Levine, "Reinforcement learning with deep energy-based policies," *34th International Conference on Machine Learning, ICML 2017*, vol. 3, pp. 2171–2186, 2017. arXiv: 1702.08165.

[121] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," *35th International Conference on Machine Learning, ICML 2018*, vol. 5, pp. 2976–2989, 2018. arXiv: 1801.01290.

[122] B. G. Bardy, L. Marin, T. A. Stoffregen, and R. J. Bootsma, "Postural coordination modes considered as emergent phenomena," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 25, no. 5, pp. 1284–1301, 1999, ISSN: 00961523. DOI: 10.1037/0096-1523.25.5.1284.

[123] B. G. Bardy, O. Oullier, R. J. Bootsma, and T. A. Stoffregen, "Dynamics of human postural transitions," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 28, no. 3, pp. 499–514, 2002, ISSN: 00961523. DOI: 10.1037/0096-1523.28.3.499.

[124] V. Bonnet, S. Ramdani, P. Fraisse, N. Ramdani, J. Lagarde, and B. G. Bardy, "A structurally optimal control model for predicting and analyzing human postural coordination," *Journal of Biomechanics*, vol. 44, no. 11, pp. 2123–2128, 2011, ISSN: 00219290. DOI: 10.1016/j.jbiomech.2011.05.027.