INVERSE OPTIMAL CONTROL FOR DIFFERENTIALLY FLAT SYSTEMS WITH APPLICATION TO LOWER-LIMB PROSTHETIC DEVICES

BY

NAVID AGHASADEGHI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Timothy Bretl, Chair
Professor Eric Perreault, Northwestern University
Professor Rayadurgam Srikant
Professor Douglas Jones

ABSTRACT

Powered prosthetic devices have shown to be capable of restoring natural gait to amputees. However, the commercialization of these devices is faced by some challenges, in particular in prosthetic controller design. A common control framework for these devices is called impedance control. The challenge in the application of this framework is that it requires the choice of many controller parameters, which are chosen by clinicians through trial and error for each patient. In this thesis we automate the process of choosing these parameters by learning from demonstration. To learn impedance controller parameters for flat-ground, we adopt the method of learning from exemplar trajectories. Since we do not at first have exemplar joint trajectories that are specific to each patient, we use invariances in locomotion to produce them from pre-recorded observations of unimpaired human walking and from measurements of the patient's height, weight, thigh length, and shank length. Experiments with two able-bodied human subjects wearing the Vanderbilt prosthetic leg with an able-bodied adaptor show that our method recovers the same level of performance that can be achieved by a clinician but reduces the amount of time required to choose controller parameters from four hours to four minutes.

To extend this framework to learning controllers for stair ascent, we need a model of locomotion that is capable of generating exemplar trajectories for any desired stair height. Motivated by this challenge, we focus on a class of learning from demonstration methods called inverse optimal control. Inverse optimal control is the problem of computing a cost function with respect to which observed trajectories of a given dynamic system are optimal. We first present a new formulation of this problem, based on minimizing the extent to which first-order necessary conditions of optimality are violated. This formulation leads to a computationally efficient solution as opposed to traditional approaches. Furthermore, we develop the theory of inverse optimal control for the case where the dynamic system is differentially flat. We demonstrate that the solution further simplifies in this case, in fact reducing to finite-dimensional linear least-squares minimization. We show how to make this solution robust to model perturbation, sampled data, and measurement noise, as well as provide a recursive implementation for online

learning. Finally, we apply our new formulation of inverse optimal control to model human locomotion during stair ascent. Given sparse observations of human walkers, our model predicts joint angle trajectories for novel stair heights that compare well to motion capture data. These exemplar trajectories are then used to learn prosthetic controllers for one subject. We show the performance of the learned controllers in a stair ascent experiment with the subject walking with the Vanderbilt prosthetic device.

In loving memory of my grandfather, Mohammad Reza Aghasadeghi.

ACKNOWLEDGMENTS

This dissertation would have never been possible without the help and support of the many people I was fortunate enough to have in my life. First and foremost, I would like to thank my advisor Tim Bretl for teaching me the very fundamentals of research and how to be a critical thinker, while at the same time giving me the flexibility to work on my interests, even if they seemed tangential at times. I am extremely grateful to Eric Perreault for believing in me and giving me a truly unique opportunity to apply my engineering knowledge to rehabilitation and prosthetics, and for graciously helping improve my career. I feel very fortunate to have worked in this exciting area of research and I owe much of it to both Eric and Tim. I also want to thank my committee members Dr. Srikant and Doug Jones for all their support and providing invaluable advice to improve my work and define my career path.

During my time in Champaign, I've had the pleasure of meeting a number of people who, whether I worked with them or not, all had a significant positive impact on my doctoral studies and my life. First, I need to acknowledge my research group that provided an exciting environment for doing research. Aadeel Akhtar, we started our journey very much the same way and you were always the biggest source of energy and encouragement throughout this process. Miles Johnson, Abdullah Akce, Dennis Matthews and Aaron Becker, the originals, you were all so refreshingly unique and a pleasure to work with. Joe Degol, David Hanely and Xinke Deng, I never thought quadrotors would be so much fun to work on. Jamie Norton, Jessica Mullins and Sean Yen thanks for bringing a different perspective to the lab and for your always positive attitudes. Finally, Andy Borum, Jacob Wagner, Mary Nguyen and Kevin Chen, it was a joy to work with you all. Throughout these years, I was lucky to have great officemates and CSL colleagues. It was a pleasure sharing these times with you Figen Oktem, Leila Fuladi, Vineet Abishek, Shiva Theja, Qiaomin Xie. Sara Bahramian, thanks for being the office mate and friend that I could talk to about anything. Also special thanks to my lab mate/roommates Javad Ghaderi and Fardad Raisali for making these years memorable ones. I cannot even imagine the last two years of my PhD without Ehsan

Salimi, Shayan Bordbar, Omid Alamdari and Jin Li. Ehsan, Shayan and Omid, the infinite sequence of soccer, FIFA and late night tea breaks we had kept me energetic and sane during tough times. Jin, your refreshing sense of humor and your ability to see the good in every situation made the last two years fly by with so much laughter and happiness.

I would like to acknowledge my funding sources, in particular the National Institute of Neurological Disorders and Stroke from the National Institute of Health for their generous fellowship award, which allowed me to pursue this interdisciplinary research.

Most importantly, I would like to thank my lovely family for their constant support and motivation. Mahshid, you have always been the best little sister I could ask for and such a big support in my life. Dad and Mom, nothing I ever say or do will even begin to repay what you have done for me. I owe all of this to your unconditional love and your belief in me. You have always been my biggest supporters and for that I am forever grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

There are over one million lower-limb amputees in the United States, and this number is expected to double by 2030 [1]. To restore locomotive functions to lower-limb amputees, different prosthetic devices such as the Rheo knee and the C-leg have been developed over the years [2]. Although these prostheses have made significant advances over previous mechanical systems, they are still limited by their passive nature. Passive prostheses cannot provide net power over a gait cycle, and therefore cause increased consumption of metabolic energy [3], and also have difficulty providing locomotive functions such as stair climbing, slope ascent and running. Therefore, powered prostheses have been suggested as a natural solution to overcome these limitations. Technological advances in the design of actuators, batteries and microprocessors along with improvements in control strategies have enabled the design of various powered prosthetic devices. These include prototype devices such as the Vanderbilt leg [4], but only one commercially available device called the BiOM prosthetic ankle. Although recent advancements are paving the way for the widespread use of powered prosthetic devices, we are still faced with important challenges that need to be addressed for such devices to become commercial.

In this thesis we focus on the problem of controller design for prosthetic devices, while many challenges related to the mechanical design of these devices such as power efficiency are still remaining and are equally important. Lower-limb prosthetic controllers often have a hierarchical structure that loosely resembles the structure and functionality of the human central nervous system (CNS) [5]. In this structure, a high-level controller is responsible for inferring the user's locomotive intent based on input signals from the device, the user and the environment. Once the high-level intent is inferred, a mid-level controller is designed to generate the desired output state for the given intent. Given these desired outputs states, a low-level device-specific controller is employed to execute the desired movement on the device. The active research in the area of prosthetic control is involved in solving challenges in one or a few of these levels. In this thesis, we will focus on the mid-level prosthetic controller.

One common strategy for mid-level control of lower-limb prostheses is impedance

control [4]. It proceeds by breaking a gait cycle into four phases, and by applying a proportional derivative (PD) feedback policy within each phase [4, 6, 7]. A challenge in the application of this framework is that it requires the choice of many controller parameters, in particular 12 parameters for the knee joint. Currently clinicians often choose these parameters by trial and error for each patient [4, 8, 9]. This tuning process takes four hours for each locomotion mode, for example level ground walking or stair/ramp ascent. Furthermore, in locomotion modes such as stair ascent, a separate controller has to be tuned for every desired stair height.

An approach to address the prosthetic controller tuning challenge is to learn controllers from demonstrations of locomotion. Techniques for *learning from demonstration* have been studied vastly in robotics. One line of work is learning from exemplar trajectories, or policy learning, which has been successfully applied in robotic applications [10–13]. These methods solve a regression problem to find controller parameters that would best reproduce an observed desired trajectory. An alternative approach is to learn a performance criterion given observations, and use the learned performance criterion (or cost function) to derive controllers. This approach is often called *inverse optimal control* or *inverse reinforcement learning* [13–32]. In these methods, the goal is to estimate a cost function, with respect to which observed trajectories are optimal, given a dynamic model of the system. This learned cost function is then expected to generalize to novel situations and produce appropriate controllers under circumstances not observed in training.

In the first part of this thesis, we address prosthetic tuning for flat-ground locomotion. Making the assumption that we can learn impedance controller parameters given an observation of average locomotion trajectories for flat-ground, we adopt the method of learning from exemplar trajectories. In our case, however, exemplar trajectories corresponding to a particular subject are not available. Thus we make use of invariances that exist in locomotion to produce subject-specific trajectories [33]. By tracking the invariant trajectories on a model which uses the amputee's physical characteristics, i.e. mass and length information of the individual, we produce exemplar joint trajectories corresponding to the size of the amputee. Once the amputee-specific trajectories are generated, we learn impedance controller parameters using a least-squares minimization. Experiments with two able-bodied human subjects wearing the Vanderbilt prosthetic leg with an able-bodied adaptor show that our method recovers the same level of performance that can be achieved by a clinician but reduces the amount of time required to choose controller parameters from four hours to four minutes.

To learn controllers for stair ascent, we cannot make the assumption of learning con-

trollers using an average observation of stair ascent, since stair ascent trajectories vary according to the stair height. Therefore, we propose to learn a cost function that is parameterized by the stair height and we adopt the method of inverse optimal control for stair ascent prosthetic learning. We propose to learn a cost function using sparse observations of stair ascent locomotion, and we hypothesize that the learned cost function can then generalize to produce stair ascent trajectories for novel step heights. Motivated by this application we extend the theory of inverse optimal control, developing a computationally efficient solution with theoretical guarantees on the quality of the estimated cost function.

The problem of IOC has been studied in many contexts. The works [28, 34] develop IOC approaches for systems modeled as Markov decision processes. IOC algorithms for non-linear stochastic continuous systems have been developed in the context of linearly-solvable Markov decision processes [30] and path integrals [31]. Additionally, IOC approaches for deterministic non-linear continuous-time systems have been proposed in [14]. A solution approach that is often used to solve these IOC problems is to iteratively solve an optimal control problem followed by minimization of an objective function that represents the quality of the estimated cost. The need to solve an optimal control problem introduces a major computational bottleneck for such IOC approaches. Moreover, the nature of this solution approach makes it difficult to provide theoretical guarantees on estimating the true cost parameters, i.e. ensuring identifiability of the IOC problem.

In chapters 3 and 4 of this thesis, we develop an approach to IOC for the class of deterministic non-linear continuous-time systems. The contribution of this work to the theoretical framework of IOC is twofold. First, we introduce a new solution approach to inverse optimal control based on linear least squares in chapter 3. We do so by minimizing the extent to which first-order necessary conditions of optimality are violated. Second, we develop a new formulation of inverse optimal control for differentially flat systems [35,36]. We show that for the class of differentially flat systems, the IOC problem is "easy" and can be solved efficiently using linear least squares minimization. Moreover, these contributions allow us to answer questions regarding the identifiability of the cost function, i.e. estimating the true cost function underlying optimal observations. We further develop a recursive solution to IOC, which allows for recursively estimating the cost function given new data.

We learn a model of human locomotion during stair ascent, using the proposed method of inverse optimal control for differentially flat systems. We incorporate a-priori knowledge about locomotion, by designing features that could model the intersegmental co-

ordination observed in human locomotion [33] and the principle of proximo-distal gradient [37]. We then learn the cost parameters using sparse observations of stair ascent trajectories, and show that our model can predict stair ascent trajectories corresponding to a *novel* stair height. This locomotion model is in turn used with the impedance parameter learning framework to learn a prosthetic controller for an able-bodied subject. We perform experiments with the subject walking with the Vanderbilt prosthetic leg, to demonstrate the performance of the learned controller. Moreover, using the learned model of locomotion, we design a different control architecture for controlling the knee, and show its performance in simulations of a dynamic bipedal model.

In summary, the contributions of this thesis are the following:

- **Prosthetic controller learning for flat ground:** In (Aghasadeghi et al. 2013 [38]) and (Aghasadeghi et al. 2015 [39]) we develop a method for learning impedance controller parameters for flat-ground prosthetic walking.

- **Approximately optimal inverse optimal control:** In (Aghasadeghi and Bretl 2012 [40]) and (Johnson, Aghasadeghi and Bretl 2013 [41]) we develop a computationally efficient solution to inverse optimal control based on the first-order necessary conditions of optimality.

- **Inverse optimal control for differentially flat systems:** In (Aghasadeghi and Bretl 2014 [42]) we develop the theory of inverse optimal control for differentially flat systems. We show that for this class of systems, the inverse optimal control problem reduces to a linear least squares solution. We provide theoretical guarantees on the identifiability of cost parameters, and develop filtering methods and recursive solutions for IOC.

- **Inverse optimal control for differentially flat systems with application to stair ascent prosthetic controller learning:** In (Aghasadeghi, Hargrove, Perreault and Bretl 2015 [43]) we develop the theory of inverse optimal control for differentially flat systems, and we use this theory to model locomotion during stair ascent. Using this model, we learn prosthetic controllers for stair ascent. We further develop a height-invariant prosthetic controller architecture, and show its feasibility in dynamic biped simulations.

The remainder of this thesis proceeds as follows. In chapter 2 we introduce the impedance parameter learning framework used to automate the process of prosthetic controller tuning. In chapter 3 we develop an alternative formulation of inverse optimal control based on minimizing the violations of the first-order necessary conditions

of optimality. We develop the theory of IOC further for differentially flat systems in chapter 4. We use the proposed method of IOC to model locomotion during stair ascent in chapter 5 and we conclude with the discussion in chapter 6.

CHAPTER 2

LEARNING IMPEDANCE CONTROLLER PARAMETERS FOR LOWER-LIMB PROSTHESES

2.1 Introduction

Impedance control [44] is an existing strategy for control of lower-limb prostheses [4, 45]. It typically proceeds by breaking a gait cycle into four phases, and by applying a proportional derivative (PD) feedback policy within each phase [4, 7], as shown below:

$$u(t) = k(\theta(t) - \theta^e) + b\dot{\theta}(t). \tag{2.1}$$

Here $u(t)$ denotes the joint torque and $\theta(t)$ and $\dot{\theta}(t)$ denote the joint angle and angular velocity. A challenge in applying this framework is that it requires choosing controller parameters k, b and θ^e for each phase. These parameters represent the stiffness, damping and equilibrium angle. Thus for a four phase control architecture for the knee and ankle, 24 controller parameters must be chosen.

Currently clinicians tune these parameters by trial and error for each patient [4, 8]. In this process, the amputee first walks on the prosthetic device with a pre-defined set of controller parameters [4]. The clinicians then modify the parameters to match the kinematics of unimpaired gait [6] and also to address the amputee's feedback. These steps are iterated until the device is fully tuned for the individual. This tuning process takes four hours on average for each individual for the impedance controller we consider in this work [46]. Moreover, multiple sessions are necessary to tune the device for different modes of locomotion such as stair ascent/descent and ramp ascent/descent.

In this chapter, we show how to learn impedance controller parameters k, b and θ_e from observed exemplar trajectories. Since we do not at first have exemplar joint trajectories that are specific to each individual, we use invariances in locomotion to produce them from pre-recorded observations of unimpaired human walking. To do so, we utilize locomotion outputs that have been shown to be invariant across individuals and walking speeds [33]. By tracking the invariant outputs on a model which uses the amputee's

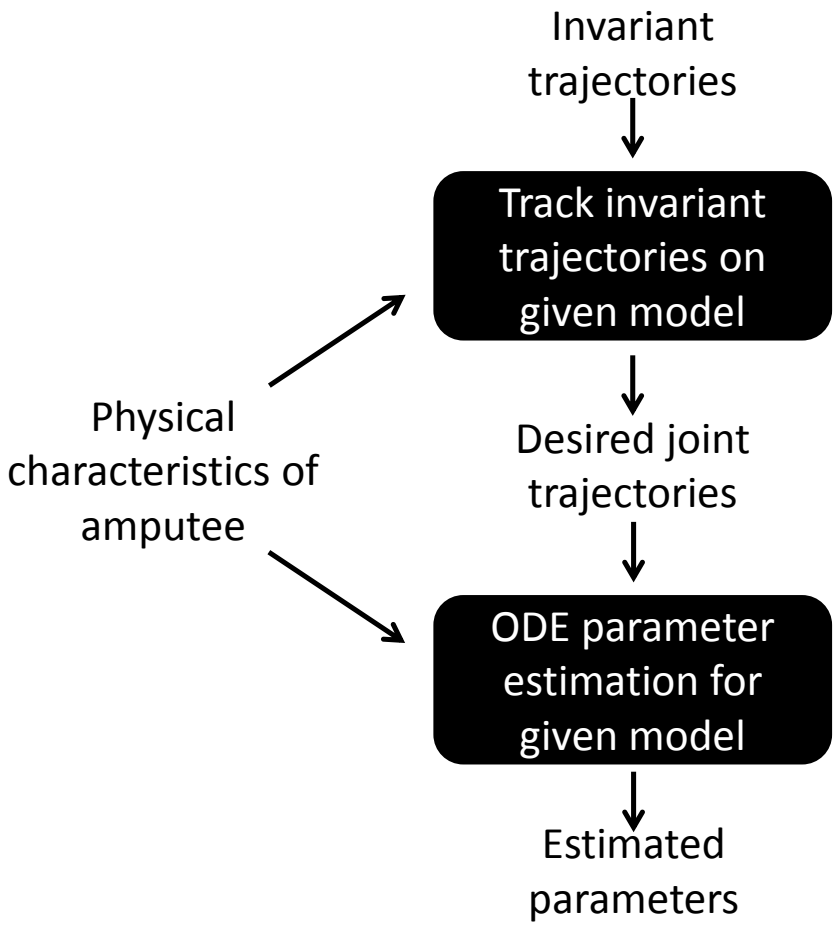


Figure 2.1: Overview of the proposed algorithm.

physical characteristics, i.e. mass and length information of the individual, we produce exemplar joint trajectories corresponding to the size of the amputee. Once the amputee-specific trajectories are generated, we estimate impedance controller parameters using a least-squares minimization. This leads to a computationally-efficient approach that produces controller parameters for an amputee walking at a desired cadence, while only using pre-recorded invariant locomotion trajectories and the amputee's physical characteristics. This approach is shown in the block diagram in Figure 2.1.

Biomechanical system identification is an alternative approach that in principle can be used to select impedance parameters for lower-limb impedance controllers [47,48]. These approaches consist of stationary and time-varying impedance identification techniques. In stationary impedance identification, the joint under study is perturbed multiple times, and the time-invariant impedance of the joint is estimated, potentially at many points in a locomotor study [49]. While these stationary identification methods are very well studied, they can be used to study joint impedances only in steady-state (stationary) con-

ditions [47, 50]. Time-varying impedance identification (TVID) techniques [51], on the other hand, continuously perturb the joints and can measure instantaneous impedance values in dynamic cases, i.e. when the joints are moving. Traditionally, these approaches were limited by the amount of experimental data that they required. However, recent work [52] has demonstrated the applicability of these methods to problems with short data segments as well. Moreover, using TVID for identifying impedance during human walking requires special equipment. A mechanical system was developed for ankle impedance estimation in [49, 53] and for knee impedance estimation in [54]. However, the development of these approaches is still a topic of active research and even if such systems were to be developed, the experimental cost could limit their use for individuals.

In this chapter, we apply the proposed impedance parameter learning algorithm to find controller parameters for two individuals. We perform experiments with the individuals walking on a lower-limb prosthetic device, and demonstrate the results.

The remainder of the chapter is organized as follows. In section 2.2 we discuss the methods, including the description of the biped model, the impedance controller and our proposed method of controller parameter learning. We then present both the simulated and experimental results in section 2.3, followed by a discussion in section 2.4.

2.2 Methods

In this section we describe our proposed framework for learning the parameters of a prosthetic impedance controller. We also discuss the dynamic biped model utilized to model the individual. These methods are applied to learn prosthetic controller parameters for two unimpaired individuals.

2.2.1 Powered Prosthetic Device

The powered prosthetic device used in this study, shown in Figure 2.2, was designed and fabricated at Vanderbilt University [55]. This device consists of an active ankle and knee joint controlled using brushless DC motors. The prosthetic device is equipped with a shank axial load sensor, encoders at the knee and ankle joints, and a six-axis inertial measurement unit. We used an able-bodied testing adaptor, seen in Figure 2.2, to test the prosthetic device with able-bodied individuals. In our experiments, we also used an electric goniometer to measure the knee angle of the non-prosthetic leg.

In this study, both the knee and ankle joints of the prosthetic device were controlled

using impedance control, which will be described later. We focus on learning impedance controller parameters for the knee, thus the controller parameters for the ankle were experimentally tuned for each individual.

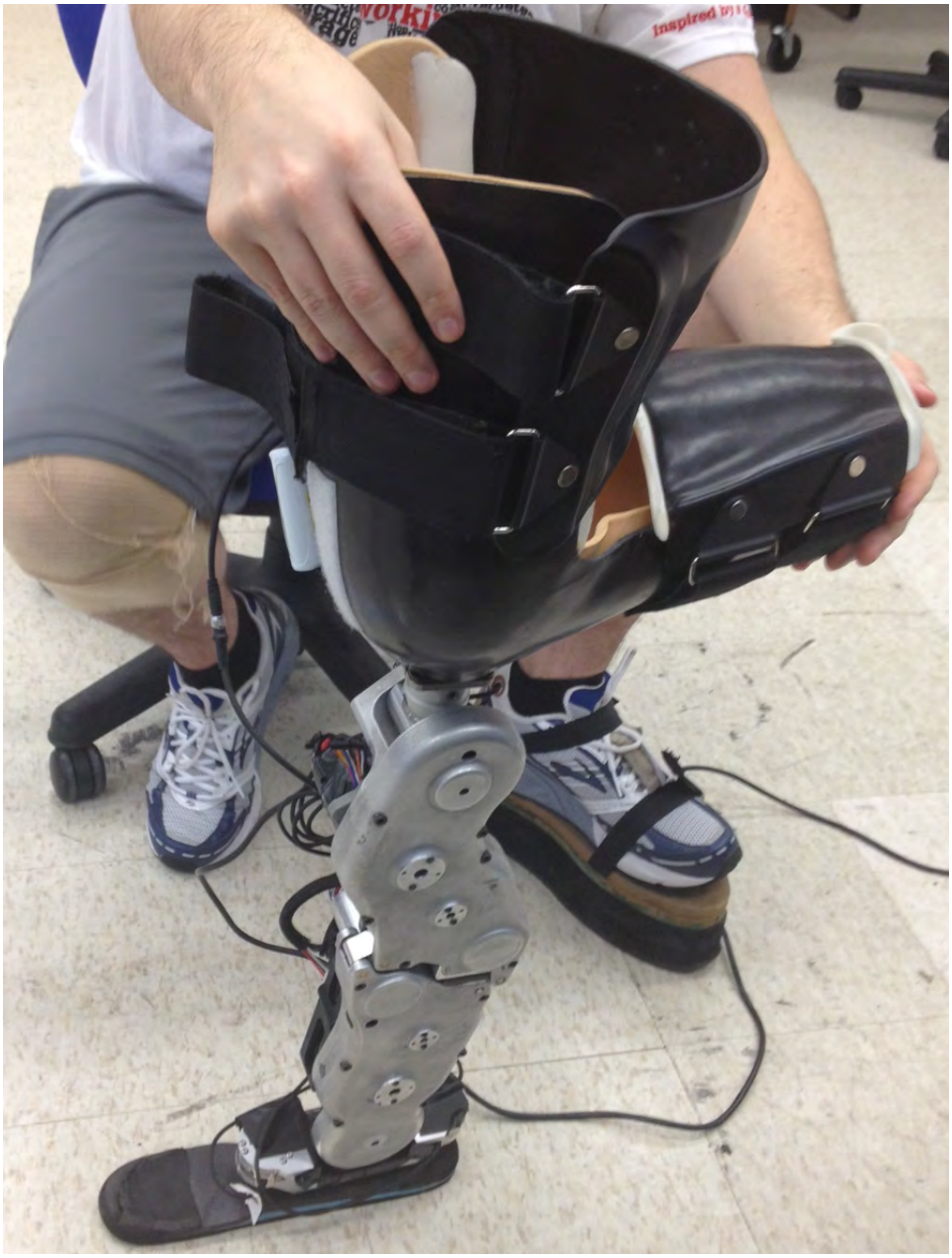


Figure 2.2: Vanderbilt prosthetic leg with the attached able-bodied adaptor.

Table 2.1: Table describing the mass (kg) and length (m) properties of the individuals (S1 and S2) and of the prosthetic (Pr)

*	L_{calf}	L_{thigh}	H	m_{foot}	m_{calf}	m_{thigh}	m_{torso}	W
S1	0.483	0.483	1.75	1.38	4.43	9.53	64.6	95.3
Pr	*	*	*	*	*	*	*	5.44

*	L_{calf}	L_{thigh}	H	m_{foot}	m_{calf}	m_{thigh}	m_{torso}	W
S2	0.483	0.483	1.96	1.51	4.85	10.4	70.7	104
Pr	*	*	*	*	*	*	*	5.44

2.2.2 Experimental Protocol

Two able-bodied male individuals participated in this experiment. Both individuals provided written informed consent to a protocol approved by the Northwestern University Institutional Review Board. The experiments were conducted in the laboratory of Center for Bionic Medicine at the Rehabilitation Institute of Chicago (RIC).

To learn controller parameters for each participant, we measured and computed the physical characteristics of the participants using the approach described in the next section. The resulting physical characteristics are detailed in Table 2.1. We then learned impedance controller parameters for both individuals using our proposed method, and tested the learned controllers in clinical experiments. To do so, the individuals were fitted to the prosthetic device on their right leg using the able-bodied testing adaptor. An electric goniometer was also placed on their left knee to measure the knee angle of the non-prosthetic leg. The knee angle of the prosthetic leg was measured using the encoder build in the prosthetic. The individuals were asked to walk between parallel bars and on a treadmill device for a duration of five minutes using the learned controllers. The individuals were instructed to walk with their desired cadence. After this experiment, clinicians experimentally modified the prosthetic controller parameters, starting with the learned controller parameters as an initial guess. Note that this method of tuning differs from previous experimental tuning, where clinicians started with a pre-defined set of controller parameters as an initial guess as opposed to starting from the learned controller for each individual. The individuals performed the same walking experiment with the experimentally modified controller parameters.

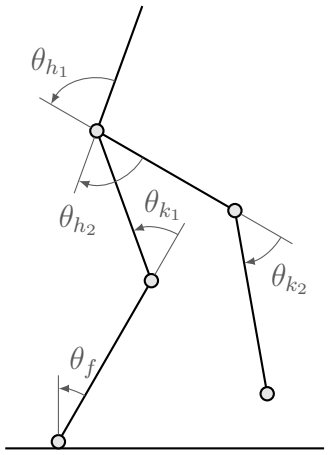


Figure 2.3: Five-link biped model considered in this chapter and the joints angles.

2.2.3 Biped Model Corresponding to the Physical Characteristics of the Amputee

We use a five-link biped to model the individual. We measure the following physical characteristics of the amputee:

- weight denoted by W
- height denoted by H
- thigh length denoted by L_{thigh}
- shank length denoted by L_{shank}.

The thigh and shank measurements will correspond to the sound leg of the individual. The size and weight of the prosthetic device are also measured and will be incorporated into the model.

We consider a planar point-feet biped model that has five serial chain links consisting of one torso, two thighs and two shanks as seen in Figure 2.3. The corresponding lengths and masses are denoted by $L_{torso}, m_{torso}, L_{thigh}, m_{thigh}, L_{shank}, m_{shank}$.

The configuration space of the biped is defined as $\mathcal{Q} = \mathbb{S}^5$ with coordinates $\theta = (\theta_f, \theta_{k1}, \theta_{h1}, \theta_{h2}, \theta_{k2})^T$. As illustrated in Figure 2.3, θ_f is the stance foot angle, θ_{k1} is the stance knee angle, θ_{h1} is the stance hip angle which is measured from the stance thigh to the torso, θ_{h2} is the swing hip angle which is measured from the torso to the swing thigh, and θ_{k2} is the swing knee angle. Throughout the chapter, a variable without a subscript denotes the vector consisting of the variable for all joints.

To compute the mass and length of all the links of the biped model using the measured physical characteristics of the amputee, we utilize the mass and length distributions discussed in [56, 57]. In particular we use fractions stating that the mass of the torso is 67.8%, mass of the thigh is 10% and mass of the combined shank and foot is 6.1% of the total weight. The mass of the shank of the impaired leg will be replaced with the total mass of the prosthetic device. Additionally all the length links besides the torso length are measured. The length of the torso is computed based on the length distribution law stating that the center of mass is located at a height equal to two-thirds of the height of the human. Also, as the length of the prosthetic device is adjusted to equal the sound leg, both legs will have the same length in our model.

The equations of motion of the biped modeling the continuous dynamics are given using the Euler-Lagrange formula

$$D(\theta)\ddot{\theta} + H(\theta, \dot{\theta}) = B(\theta)u, \tag{2.2}$$

where

$$H(\theta, \dot{\theta}) = C(\theta, \dot{\theta})\dot{\theta} + G(\theta).$$

Here $D(\theta)$ is the inertial matrix, $C(\theta, \dot{\theta})$ is the Coriolis matrix, $G(\theta)$ is the gravitational vector and $B(\theta)$ is the control torque distribution map which equals the identity matrix since the biped is fully controlled. These matrices are computed with respect to physical characteristics of the amputee.

We further introduce discrete dynamics to model the impact that happens instantaneously when the swing foot hits the ground. This causes an instantaneous change in the velocities of the biped and also an instantaneous switching of the stance and swing legs. We use the function Δ to denote the mapping between $(\theta, \dot{\theta})$ just before and just after the impact. Further details of the impact map can be found in [58, 59].

2.2.4 Impedance Control Framework

An existing approach to controlling lower-limb prosthetic [4] devices proceeds by breaking the gait into four phases, denoted $p = 1, ..., 4$ (see Figure 2.4). Each phase begins at time t_0^p and ends at t_f^p.

These phases consist of

1. P1 from heel strike to mid stance denoted by passing a threshold $\theta_f < T$,

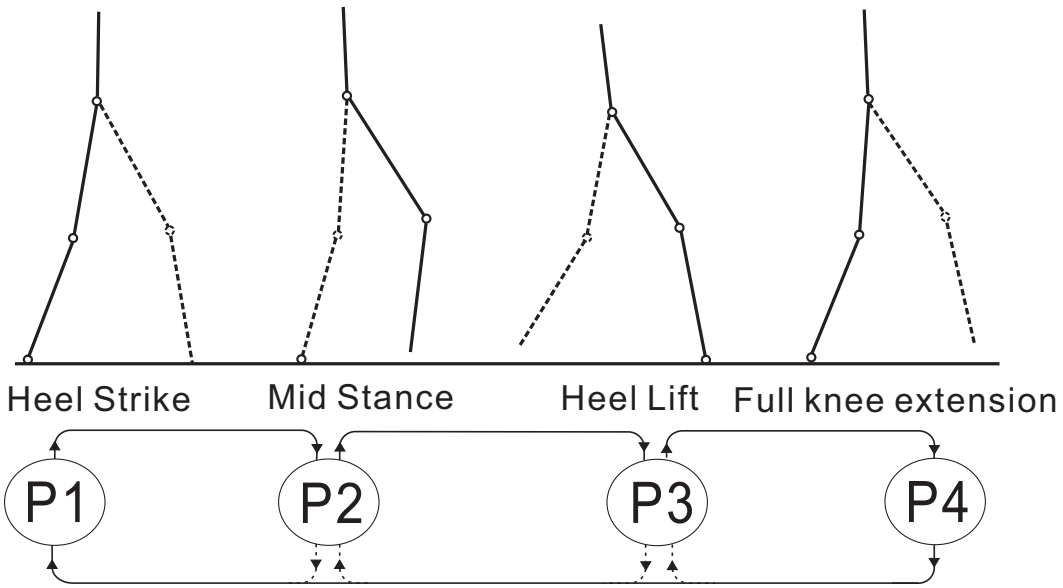


Figure 2.4: Separation of gait into four phases

2. P2 from mid stance to heel lift,

3. P3 from heel lift to full knee extension (i.e. $\dot{\theta}_{k2} < 0$),

4. P4 from full knee extension to heel strike.

The commanded torque of an impedance controller at the knee joint during a phase $p \in \{1, 2, 3, 4\}$ can be represented by

$$u_k^p(t) = k^p(\theta_k(t) - \theta_e^p) + b^p\dot{\theta}_k(t), \tag{2.3}$$

where $u_k(t)$ denotes the torque command and $\theta_k(t)$ and $\dot{\theta}_k(t)$ denote the angle and angular velocity of the knee joint, respectively, during either stance or swing phase. The impedance parameters for each phase p consist of

1. k^p denoting the stiffness,

2. b^p denoting the damping,

3. θ_e denoting the constant equilibrium angle.

These parameters are constant within each phase of the gait. We refer to the vector of all impedance parameters at the knee joint during phase p as

$$\beta^p = [k^p, \theta_e^p, b^p]^T. \tag{2.4}$$

In this chapter, we will consider having impedance control on one knee joint denoting the prosthetic knee with a total of 12 parameters to be chosen. The control input at other joints is described when necessary.

Note that the impacts in the hybrid system occur only during the transitions between the phases. Therefore, within each phase p, the biped system evolves from time t_0^p to t_f^p according to the following continuous dynamics, governed by the Euler-Lagrange equations

$$\begin{cases} D(\theta)\ddot{\theta} + H(\theta, \dot{\theta}) = u & \forall\ t \in [t_0^p, t_f^p] \\ \\ \big(\theta(t_0^p), \dot{\theta}(t_0^p)\big) = R\big(\theta(t_f^{p-1}), \dot{\theta}(t_f^{p-1})\big), \end{cases} \tag{2.5}$$

where we let $p - 1 = 4$ if $p = 1$. The initial condition to the ODE for each phase is given by the function R, which either models the impact or ensures continuity between phases and is defined as

$$R(\theta(t), \dot{\theta}(t)) = \begin{cases} \Delta(\theta(t), \dot{\theta}(t)) & \text{if at impact} \\ \\ (\theta(t), \dot{\theta}(t)) & \text{otherwise.} \end{cases}$$

For the prosthetic knee joint, we replace the impedance controller as the input to get

$$\begin{cases} \big[D(\theta)\ddot{\theta} + H(\theta, \dot{\theta})\big]_k = u_k & \forall t \in [t_0^p, t_f^p] \\ \\ u_k = k^p(\theta_k - \theta_e^p) + b^p\dot{\theta}_k \\ \\ \big(\theta(t_0^p), \dot{\theta}(t_0^p)\big) = R\big(\theta(t_f^{p-1}), \dot{\theta}(t_f^{p-1})\big), \end{cases} \tag{2.6}$$

where the notation $[v]_k$ denotes the row of the vector v corresponding to the knee joint.

2.2.5 Learning Impedance Controller Parameters

We learn impedance controller parameters from exemplar joint trajectories and using the discussed biped model corresponding to the individual. To produce exemplar joint trajectories specific to an individual, we make use of locomotion outputs that have shown to be invariant across individuals [33] and across different walking speeds. By

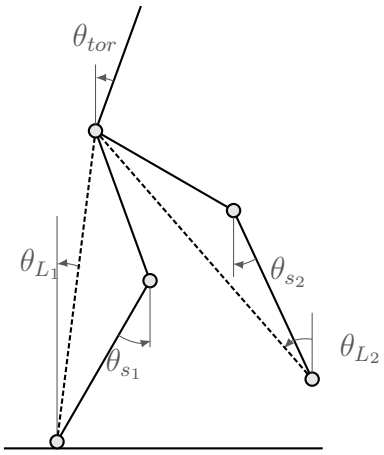


Figure 2.5: The chosen invariant locomotion outputs.

tracking these outputs on a biped model that corresponds to the individual, we produce joint trajectories for the individual, which we denote by θ^A. To track the invariant trajectories, we make use of an input/output (I/O) linearization controller [59, 60]. We make use of the following outputs that are invariant across individuals:

- y_1 denoting stance shank angle θ_{s1},
- y_2 denoting swing shank angle θ_{s2},
- y_3 denoting angle of the stance leg θ_{L1},
- y_4 denoting angle of the swing leg θ_{L2},
- y_5 denoting the angle of the torso from vertical θ_{tor}

The variables corresponding to the chosen invariant trajectories are shown in Figure 2.5. Due to the invariance of the kinematic trajectories with respect to speed, they can be scaled in time to provide trajectories for different speeds. We use invariant locomotion trajectories that were obtained from the unimpaired locomotion data using a high speed camera motion capture system. Details about these experiments can be found in [59]. Given noisy and sampled observations of the outputs y_i we first fit splines to the samples and subsequently obtain higher order derivatives of the outputs (as needed) using the spline estimates. We can also use an alternative function fitting approach that is described in [38].

As discussed in section 2.2.4 the dynamic equations for the prosthetic knee joint can be described by an implicit ODE (2.6) parameterized by impedance parameters

$\beta^p = [k^p, \theta_e^p, b^p]^T$ during each phase. After dropping the superscript p for notational simplicity, this ODE can be written as

$$\begin{cases} F_k(\beta, \theta^A, \dot{\theta}^A, \ddot{\theta}^A) = 0 & \forall \quad t \in [t_0, t_f] \\ \\ (\theta^A(t_0), \dot{\theta}^A(t_0)) = (\theta_0^A, \dot{\theta}_0^A), \end{cases} \tag{2.7}$$

where $F_k = \left[D(\theta^A)\ddot{\theta}^A + H(\theta^A, \dot{\theta}^A)\right]_k - u_k$ and $u_k = k(\theta_k^A - \theta_e) + b\ \dot{\theta}_k^A$.

We consider the problem of estimating unknown impedance parameters β_i given the parameterized ODE (2.7) and the produced joint trajectories θ^A. Two main estimation approaches have been developed to address this problem. The first class of methods are maximum likelihood (ML) estimators [61–64]. While these approaches have satisfactory theoretical properties, they are faced with some computational bottlenecks. These methods involve iteratively optimizing an objective function with respect to the parameter and then approximating the solution of an ODE with the current guess of the parameter. Approximating the solution of an ODE can be computationally expensive [65,66]. Moreover, due to the nonlinear nature of this optimization, the optimization often converges to a local solution [66].

Alternatively, another class of estimation approaches have been described in [66–68]. These approaches rely on a non-iterative least-squares approach to minimize the equation error (residual). We take this approach, with F_k denoting the residual which will be minimized. To minimize this function, we fit the joint trajectories with splines and subsequently find higher order derivatives of the joints trajectories. We then proceed to finding the unknown controller parameters by solving the following optimization:

$$\arg\min_{\beta} \int_{t_0}^{t_f} \left(F_k(\beta, \theta^A(t), \dot{\theta}^A(t), \ddot{\theta}^A\right)^2 dt. \tag{2.8}$$

Given the expression

$$F_k = \left[D(\theta^A)\ddot{\theta}^A + H(\theta^A, \dot{\theta}^A)\right]_k - \left[k(\theta_k^A - \theta_e) + b\dot{\theta}_k^A\right]$$

we can write the first component solely as a function of the known amputee joint trajectories and derivatives, i.e. $V(\theta^A, \dot{\theta}^A, \ddot{\theta}^A) = \left[D(\theta^A)\ddot{\theta}^A + H(\theta^A, \dot{\theta}^A)\right]_k$. Moreover, by defining $U(\theta^A, \dot{\theta}^A) = [\theta_k^A, 1, \dot{\theta}_k^A]$ and $\tilde{\beta} = [k, -k \times \theta_e, b]^T$ (as opposed to $\beta = [k, \theta_e, b]^T$),

we can then write the minimization (2.8) as

$$\arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} \left(V - U\tilde{\beta}\right)^2 dt. \tag{2.9}$$

Note that from (2.9) it is clear that the minimization is a linear least-squares optimization, which can be solved efficiently by the following closed-form expression:

$$\Big(\int_{t_0}^{t_f} U^T U dt\Big)\tilde{\beta} = \Big(\int_{t_0}^{t_f} U^T V dt\Big). \tag{2.10}$$

Given the expression (2.10) we can use tools from optimization theory to evaluate the properties of the solution. In particular, if the expression $\left(\int_{t_0}^{t_f} U^T U dt\right)$ is full rank, the solution will be unique. Otherwise, the solution is not unique and one can use insights from physiology of locomotion to choose a solution. In practice, we use standard Gaussian elimination or LU decomposition techniques to find numerically stable solutions to this equation. Finally note that while we are estimating the parameters $\tilde{\beta}$ instead of β, the estimated value of θ_e can be obtained by $-\tilde{\beta}(2)/\tilde{\beta}(1)$. Thus, by first obtaining estimates of desired joint trajectories $\hat{\theta}$ for an amputee, and then solving the optimization (2.9), we learn impedance controller parameters for the subject.

To estimate controller parameters for the Vanderbilt leg, we measured the intrinsic damping of the device during stance and swing to be 0.05 and 0.08 $N \cdot m \cdot s$ respectively. To account for the intrinsic damping of the prosthetic, we constrained our impedance estimation to have damping values larger than the intrinsic damping of the device. At the end we subtracted the intrinsic damping from the estimated damping parameters to arrive at the learned controller parameters.

2.2.6 Performance Measures

We evaluate the learned prosthetic controller for each subject using several performance measures. The first measure of performance is the *tuning time* [46]. For the automated controller parameter learning, the tuning time consists of the time required to run the learning software on a PC. For the experimental tuning, the tuning time is measured from the time the subject walks with the first set of controller parameters until the time when the clinicians decide that the prosthetic is tuned for that subject.

As another measure of performance, we evaluate a *measure of symmetry* between the prosthetic leg and the unimpaired leg [69]. To do so, we compute the coefficient of

determination (R^2) between the prosthetic knee and the unimpaired knee trajectories.

The third measure of performance we consider is *similarity to unimpaired walking*, which has been previously used in [4, 70]. We quantify this measure by finding the coefficient of determination (R^2) between the resulting prosthetic knee trajectories, and average unimpaired trajectories.

Finally, we evaluate the *number of falls and amount of necessary support*. The number of falls are counted during the entire experimental session for both learned and experimentally tuned parameters. The amount of support is also evaluated as the necessity of using the overhead harness in walking experiments.

2.3 Results

The two subjects performed the walking experiments outlined in the methods section at the Center for Bionic Medicine as seen in Figure 2.6. The resulting learned controller parameters and experimentally tuned controller parameters for these subjects are detailed in Tables 2.2 and 2.3.

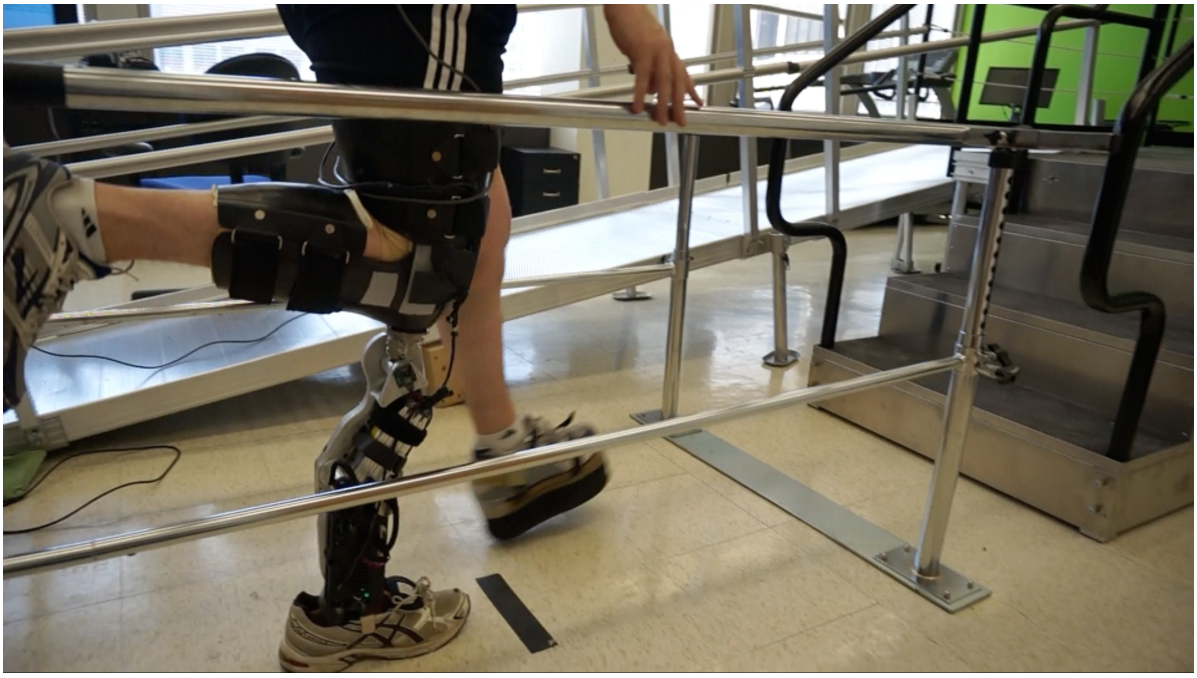


Figure 2.6: This figure shows the flat ground walking experiment performed at the Center for Bionic Medicine. In this picture, the subject is walking with the learned prosthetic controller parameters.

Table 2.2: Learned (top) and experimentally modified (bottom) controller parameters for subject 1

Phase	$k[N{\cdot}m]$	$b[N{\cdot}m{\cdot}s]$	θ_e *[Degree]*
P1	4	0.276	0
P2	0.94	0.05	0.4
P3	0.45	0.01	65
P4	0.33	0.02	0

Phase	$k[N{\cdot}m]$	$b[N{\cdot}m{\cdot}s]$	θ_e *[Degree]*
P1	4	0.276	0
P2	2	0.05	0.4
P3	0.45	0.01	65
P4	0.25	0.01	0

The learned controller parameters were first used in a dynamic biped simulation to evaluate the resulting behavior of the controller. The resulting knee angle trajectory for the learned controller parameters, for subject 1 as an example, is shown in Figure 2.7.

The controllers were next implemented on the prosthetic device, and the subjects performed the clinical experiments. The prosthetic knee angle trajectories corresponding to the learned controllers and experimentally tuned controllers are shown in Figures 2.8 and 2.9. It should be noted that the stance knee trajectories obtained from prosthetic walking in the experiments have a smaller flexion angle compared to the average unimpaired stance trajectory. This is commonly seen when walking with prosthetic devices [4, 69], and is also seen when walking with the experimentally tuned parameters. Therefore this is not an artifact of our learned impedance parameters.

The prosthetic controllers were evaluated using the discussed measures of performance. The *tuning time* for the learned controller parameters was 5 seconds. The tuning times for the experimentally modified parameters, starting with the learned ones as an initial guess, were 2.5 minutes and 4 minutes for subjects 1 and 2 respectively. This is as opposed to a tuning time of 4 hours when starting from a pre-defined set of controller parameters.

The *measure of symmetry* evaluated as the average coefficient of determination (R^2) between the prosthetic knee trajectories and the unimpaired knee trajectories was 0.9 and 0.86 for the learned controllers and 0.85 and 0.82 for the experimentally modified parameters for the two subjects respectively. The unimpaired knee trajectory for subject 1 is shown in Figure 2.10 as an example.

Table 2.3: Learned (top) and experimentally modified (bottom) controller parameters for subject 2

Phase	$k[N{\cdot}m]$	$b[N{\cdot}m{\cdot}s]$	θ_e *[Degree]*
P1	4	0.26	0
P2	1.4	0.05	4
P3	0.42	0.01	65
P4	0.29	0.02	0

Phase	$k[N{\cdot}m]$	$b[N{\cdot}m{\cdot}s]$	θ_e *[Degree]*
P1	4	0.26	0
P2	1.4	0.05	4
P3	0.3	0.01	65
P4	0.35	0.02	0

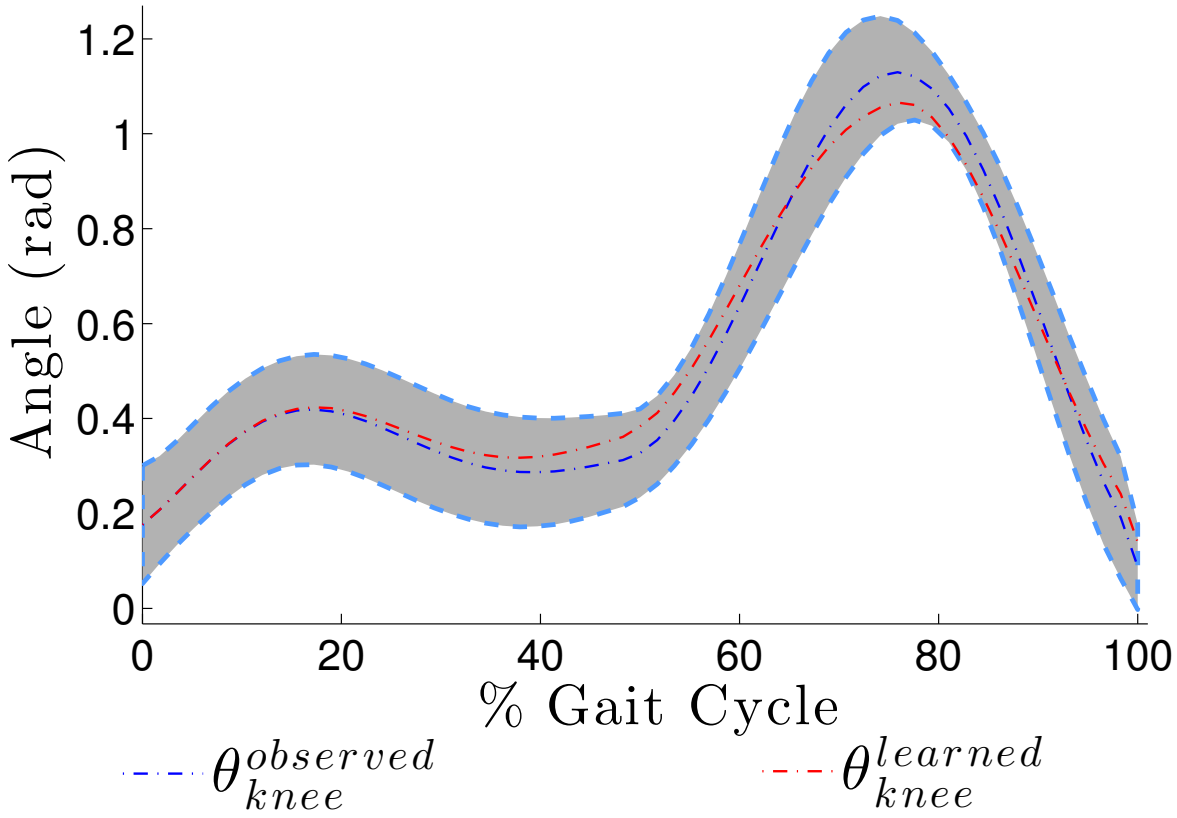


Figure 2.7: The observed unimpaired averaged knee trajectory is seen in blue, with one standard deviation shown in gray. The knee trajectory resulting from the learned impedance parameters for subject 1 is shown in red.

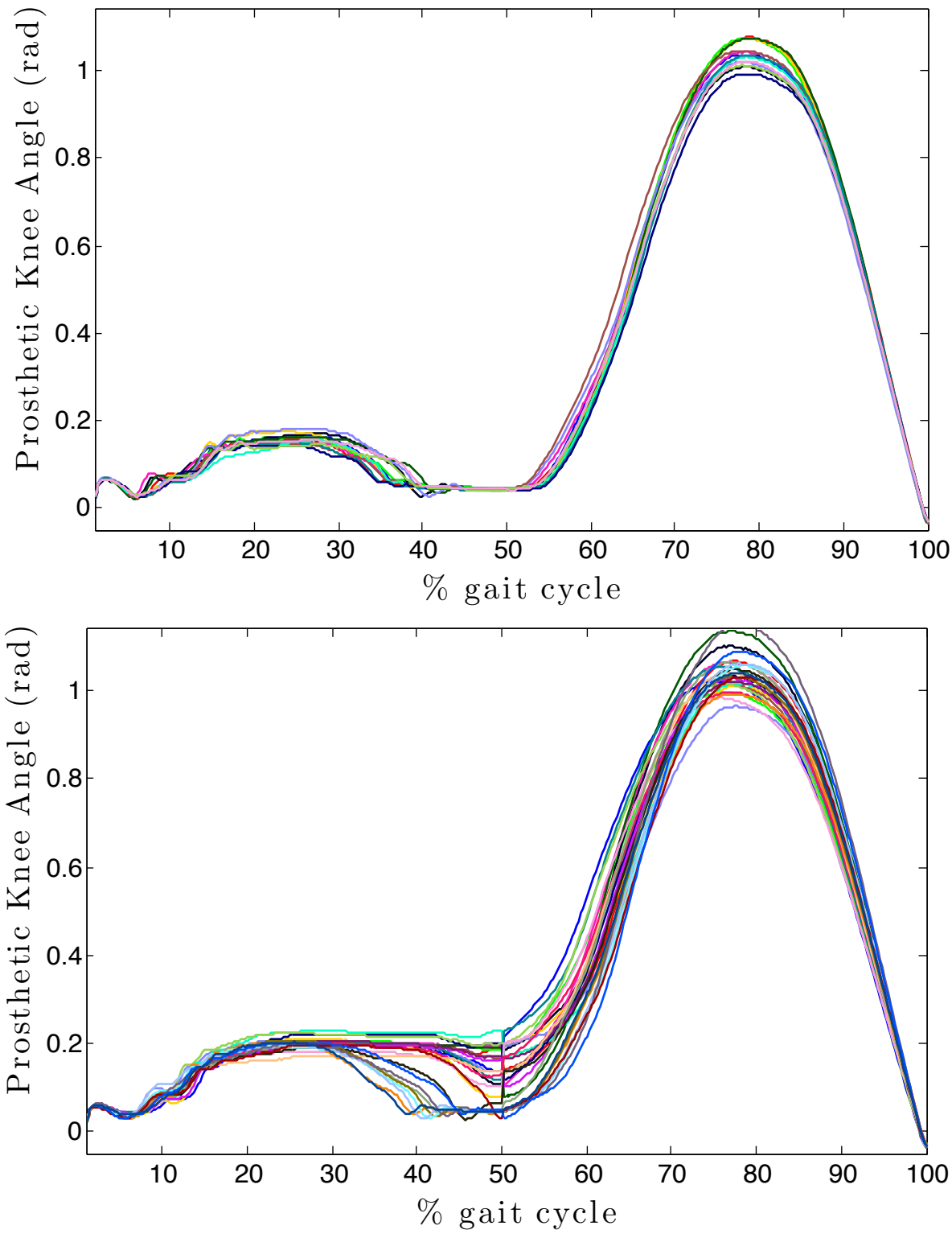


Figure 2.8: Resulting prosthetic knee trajectories with learned (top) and experimentally tuned (bottom) parameters for subject 1.

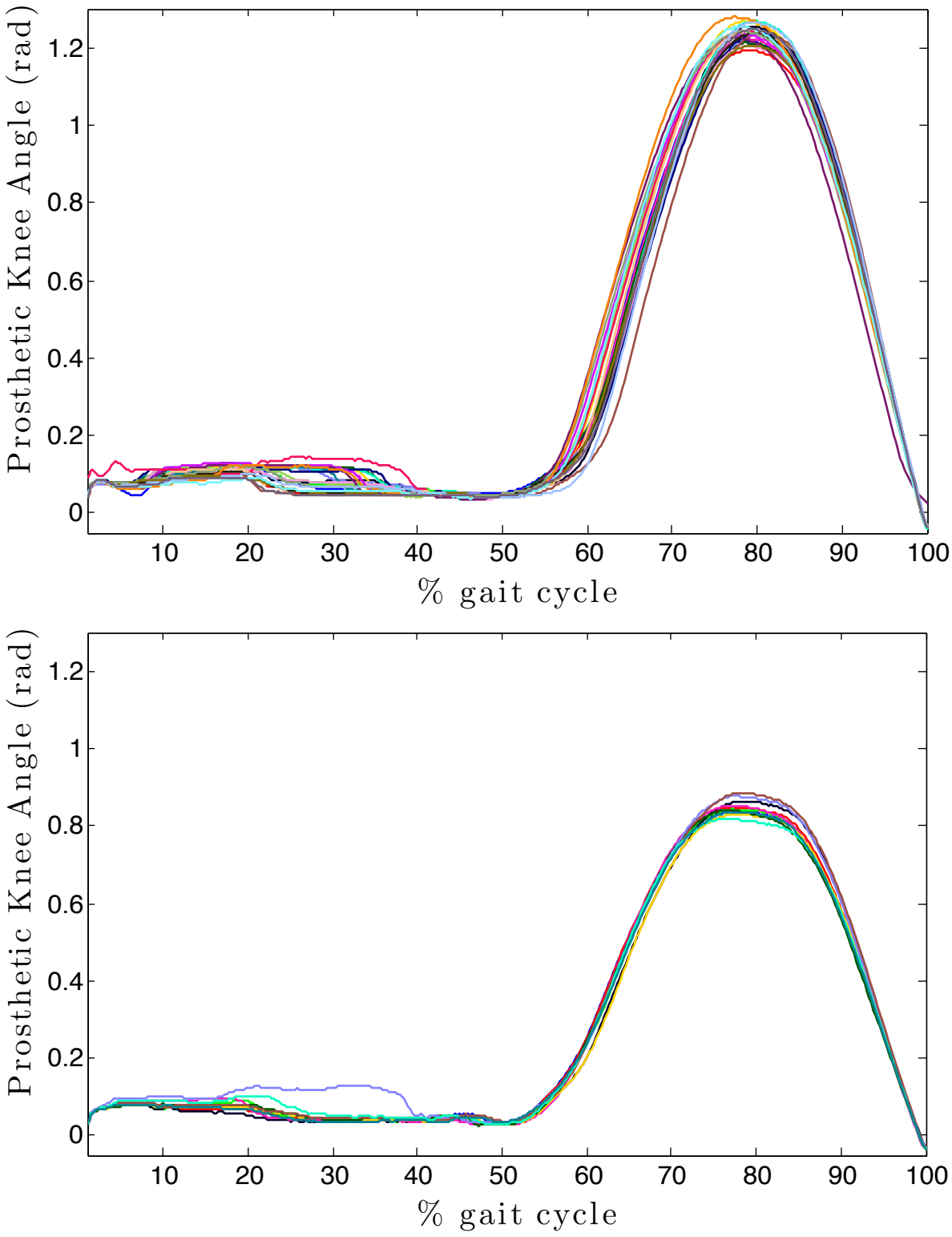


Figure 2.9: Resulting prosthetic knee trajectories with learned (top) and experimentally tuned (bottom) parameters for subject 2.

The *similarity to unimpaired walking*, also measured as the average R^2 between prosthetic knee trajectories and the average unimpaired knee trajectory, was 0.74 and 0.7 for the learned controllers and 0.85 and 0.73 for the experimentally modified parameters for both subjects.

Finally, the *number of falls* were counted to be zero for all experiments with both the learned and experimentally modified parameters. Also *the amount of necessary support* was determined by the subjects' ability to walk without any use of the overhead harness system in all experiments with both sets of controller parameters.

2.4 Discussion

In this chapter, we focused on the problem of choosing subject-specific controller parameters for lower-limb prosthetic devices. To address this problem, we proposed a method of learning impedance controller parameters that relied on pre-recorded observations of human locomotion and the individual's physical characteristics. Unlike biomechanical impedance identification approaches which compute the impedance of the joints using perturbation experiments, we used a model-based framework for learning controller parameters. In this approach, we did not estimate the true joint impedance; rather, we estimated impedance controller parameters that were necessary to replicate the observed kinematics of locomotion. Therefore, our approach is different from biomechanical impedance characterization. We demonstrated the feasibility of our approach in estimating controller parameters for two unimpaired individuals in less than four minutes as opposed to approximately four hours that is often needed for clinical tuning. We showed that these learned parameters performed similarly to clinically tuned parameters.

We formulated the problem of estimating impedance controller parameters as a linear least-squares minimization. This formulation leads to closed-form solution, which can be computed efficiently using standard convex optimization techniques. We evaluated identifiability properties of the impedance parameters, and provided conditions under which a unique estimate of the impedance controller parameter can be obtained. For cases where a family of solutions can produce similar kinematic trajectories, i.e. cases where the estimation does not have a unique solution, we can choose a solution that is bounded by an understanding of the physiology of locomotion [71].

An important advantage of our approach is that it can be readily scaled to learn parameters for different individuals and for varying locomotion speeds. This is due to our design choice in using locomotion outputs that are invariant across individuals and loco-

motion speeds. Moreover, this framework can be extended to learn controller parameters for other modes of locomotion such as stair ascent given invariant outputs corresponding to those locomotion modes. Preliminary results of the application of this approach to stair ascent parameter learning can be found in [42]. Therefore the scalability and generalizability of our approach make it suitable for use in clinical settings.

Our approach relies on three main assumptions. The first assumption is that the five-link biped model is sufficient to characterize the dynamics of the individual. The second assumption is that we can learn controller parameters using a single exemplar joint trajectory, as opposed to a cost criterion for locomotion. The third assumption is that using invariant locomotion outputs we can produce exemplar trajectories corresponding to a specific individual. Even with these assumptions, the evaluated performance measures suggest that the learned controller parameters perform similarly to experimentally tuned controllers. Moreover, the proposed framework can provide controller parameters in a much more timely manner, at worst in a few minutes, by using the learned parameters as an initial guess.

Nonetheless, our proposed approach might be improved by making less restrictive assumptions. In particular by using a different biped model, such as a seven-link biped that incorporates feet, we should be able to extend our learning framework to finding controller parameters for the ankle. Additionally, by using a cost function for locomotion as opposed to reproducing fixed exemplar trajectories, we should be able to measure transient conditions such as starting and stopping, rather than just the average steady state conditions considered in this work. This approach, however, would depend on the suitability of the selected cost function, which is still a topic of research. Lastly, one can identify alternative individual-invariant features in locomotion, and use the learned features to produce exemplar trajectories for each individual and potentially improve the approach. Learning such features can be possible given large amounts of locomotion data and the application of feature learning algorithms such as deep neural networks. However, this too is a topic of future research.

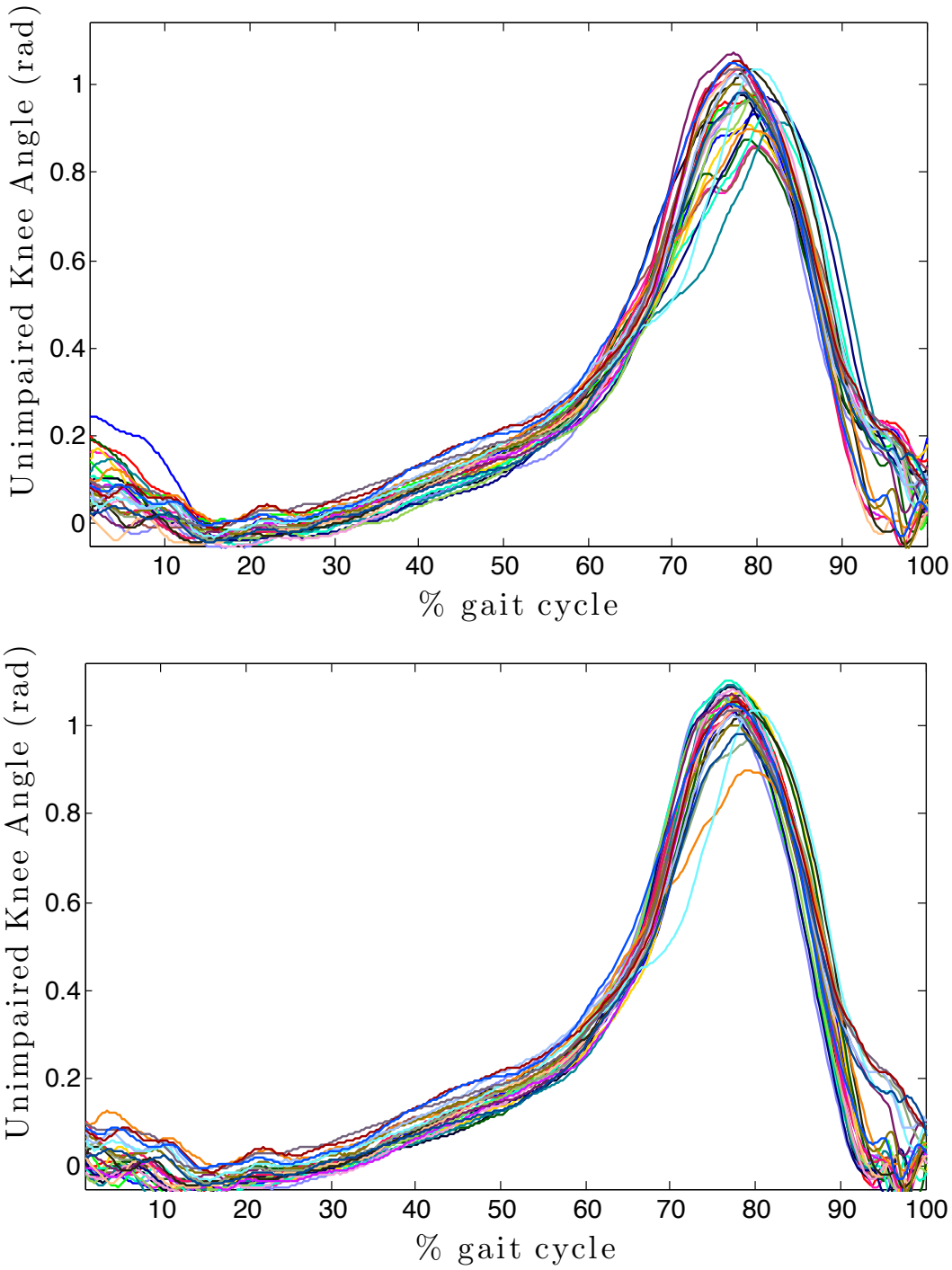


Figure 2.10: These figures plot the angular trajectory of the unimpaired knee trajectory of subject 1 when walking with the Vanderbilt prosthetic device. The top plot shows the unimpaired knee trajectory when the subject was walking with the device programmed with the learned impedance controller parameters. The bottom plot shows the unimpaired knee trajectory while the subject was walking with an experimentally tuned prosthetic device.

CHAPTER 3

APPROXIMATELY OPTIMAL FORMULATION OF INVERSE OPTIMAL CONTROL

3.1 Background on Inverse Optimal Control

Inverse optimal control is often used as a solution approach to the broad problem of learning from demonstration, which is often referred to as imitation learning or apprenticeship learning. The problem of learning from demonstration is to derive a control policy (a mapping from states to actions) from examples, or demonstrations, provided by a teacher. Demonstrations are typically considered to be sequences of state-action pairs recorded during the teacher's demonstration.

There are generally two methods of approach within learning from demonstration. One approach is to learn a map from states to actions using classification or regression. Argall et al. provide a survey of the work in this area [72]. The second general approach is to learn a cost function with respect to which observed input and state trajectories are (approximately) optimal, i.e. inverse optimal control [13–32]. These methods have primarily focused on finite-dimensional optimization and stochastic optimal control problems.

In the context of finite parameter optimization, Keshavarz et al. [17] develop an inverse optimization method which learns the value function of a discrete-time stochastic control system given observations. These ideas were extended to learn a cost function for a deterministic discrete-time system in Puydupin-Jamin et al. [18], and a hybrid dynamical system in [40]. Similarly, Terekhov et al. [19, 20] and Park et al. [21] develop an inverse optimization method for deterministic finite-dimensional optimization problems with additive cost functions and linear constraints. Other recent work formulates an optimization problem which simultaneously learns a cost function and optimal trajectories [73, 74].

A variety of methods have been developed in the context of stochastic optimal control problems and, in particular, Markov decision processes [13, 15, 23, 26–28, 30, 31]. Ng and Russell [23] developed a method for stationary Markov decision processes based on lin-

ear programming. Abbeel and Ng [15] extend that work by finding a cost function with respect to which the expert's cost is less than that of predicted trajectories by a margin. A further extension simultaneously learns the system dynamics along specific trajectories of interest [13]. Ramachandran et al. [27] take a Bayesian approach and assume that actions are distributed proportional to the future expected reward. The method developed in Ziebart et al. [28] works by computing a probability distribution over all possible paths which matches features along the observed trajectory. Dvjijotham and Todorov [30] develop a method of inverse optimal control for linearly-solvable stochastic optimal control problems. Their method takes advantage of the fact that, for the class of system model they consider, the Hamilton-Jacobi-Bellman equation gives an explicit formula for the cost function once the value function is known. Aghasadeghi and Bretl [31] develop a method of inverse optimal control that uses path integrals to create a probability distribution over all possible paths. The problem is then one of maximizing the likelihood of observations.

Learning from demonstration methods are applied in three different areas. First, learning from demonstration has been applied as a method of data-driven automation [12, 13, 15, 16, 23–26, 28, 30, 75, 76]. Tasks of interest include bipedal walking, navigation of aircraft, operation of agricultural and construction vehicles. Second, learning from demonstration methods have been applied to cognitive and neural modeling [14, 19–22, 29, 77–80]. Third, learning from demonstration methods have been applied to system identification of deformable objects [32], i.e. learning elastic stiffness parameters of objects such as surgical suture, rope, and hair.

3.2 Inverse Optimal Control: Problem Statement

In the rest of this chapter, we consider the following class of optimal control problems:

$$\begin{aligned} \underset{x,u}{\text{minimize}} \quad & \int_{t_0}^{t_f} \beta^T \phi[t, x(t), u(t)] dt \\ \text{subject to} \quad & \dot{x}(t) = f[t, x(t), u(t)] \\ & x(0) = x_{start} \\ & x(t_f) = x_{goal} \end{aligned} \tag{3.1}$$

where $x(t) \in \mathcal{X} \subset \mathbb{R}^n$ is the state, $u(t) \in \mathcal{U} \subset \mathbb{R}^m$ is the input, $\phi : \mathbb{R} \times \mathcal{X} \times \mathcal{U} \to \mathbb{R}^p$ are known basis functions, and $\beta \in \mathbb{R}^p$ is an unknown parameter vector to be learned. We

assume, without loss of generality, that $\|\beta\| \leq 1$. We assume that the system equations

$$\dot{x}(t) = f[t, x(t), u(t)] \tag{3.2}$$

are well posed, that is, for every initial condition x_{start} and every admissible control $u(t)$, the system $\dot{x}(t) = f[x(t), u(t)]$ has a unique solution x on $t \in [0, t_f]$. This is satisfied, for example, when f is continuous in t and u and differentiable ($\mathcal{C}^1$) in x, f_x is continuous in t and u, and u is piecewise continuous as a function of t [81, 82]. The objective basis function ϕ is assumed to be smooth in x and u. This problem also assumes there are no input and state constraints. These constraints are often important in practice, and will be the subject of future work.

The problem of *inverse optimal control* is to infer the unknown parameters with respect to which a given trajectory, the *observation*, is a local extremal solution to problem (3.1). This observed trajectory is denoted by

$$(x^*, u^*) = \{x^*(t), u^*(t) : t \in [0, t_f]\}. \tag{3.3}$$

For convenience, we will often drop the asterisk and refer to an optimal trajectory as (x, t). We also consider observing multiple trajectories, each local minimum of problem (3.1) for different boundary conditions. We will refer to a set D of M observations as

$$D = \left\{ \left(x^{*(i)}, u^{*(i)} \right) \right\} \quad \text{for } i = 1, \ldots, M \tag{3.4}$$

where each trajectory has boundary conditions

$$\left(x^{(i)}_{start}, x^{(i)}_{goal} \right) \quad \text{for } i = 1, \ldots, M. \tag{3.5}$$

An important quantity in the methods discussed in this chapter is the accumulated value of the unweighted basis functions along a trajectory, which we call the *feature vector* of a trajectory $\mu(x, u)$, and define by

$$\mu(x, u) = \int_{t_0}^{t_f} \phi[t, x(t), u(t)] dt. \tag{3.6}$$

In practice, one would generally have sampled observations of the behavior of the system, but for the analysis here, we assume we have perfect observations of the continuous-time system trajectories.

In practice, the solution of (3.1) is typically obtained using a numerical optimal con-

trol solver such as direct multiple shooting or collocation. In this work, we use the pseudospectral optimal control package GPOPS [83] to numerically solve the forward problem in the prior methods which require it.

3.3 Three Prior Methods

In this section we formally describe the three prior methods of inverse optimal control with which we compare the new method developed in Section 3.4. In their original form, the method of Abbeel and Ng, and the method of Ratliff et al. were developed in the context of Markov decision processes. The general structure and theoretical guarantees of the methods apply with slight modification to the deterministic continuous-time class of problems we consider in this chapter, specified in Equation (3.1).

3.3.1 Method of Mombaur et al.

The method of Mombaur et al. [14] works by searching for the cost function parameter β which minimizes the sum-squared error between predicted and observed trajectories. This method has two main components. In the upper-level, a derivative-free optimization technique is used to search for the cost function parameter β. In the lower-level, a numerical optimal control method is used to solve the forward optimal control problem (3.1) for a candidate value of β. We will now discuss the two levels in detail.

The objective of the upper-level derivative-free optimization is given by

$$\underset{\beta}{\text{minimize}} \int_{t_0}^{t_f} \|[x^\beta(t); u^\beta(t)] - [x^*(t); u^*(t)]\|^2 dt \tag{3.7}$$

where $[x^*(t); u^*(t)]$ is the vector concatenation of the state and input of the observed trajectory at time t, and $[x^\beta(t); u^\beta(t)]$ is the solution to the forward problem (3.1), given the parameter vector β. Mombaur et al. [14] discuss high performance derivative-free algorithms to minimize this upper-level objective. For our baseline analysis in this chapter, however, we use the Matlab `fminsearch` implementation of the Nelder-Mead simplex algorithm. Iterations of the Nelder-Mead algorithm constitute the upper-level of this method. Upon selecting a new candidate value of β, the lower-level proceeds by solving (3.1) for the candidate value, generating a predicted trajectory (x^β, u^β). Given this trajectory, the upper-level objective can be evaluated, and the search for a new candidate β continues. This method is easily extended for the case where multiple

trajectories are observed by considering the sum of predicted errors with respect to each observed trajectory in the upper-level objective.

3.3.2 Method of Abbeel and Ng

The method of Abbeel and Ng [15] was originally developed for infinite-horizon Markov decision processes with discounted reward. The goal of this method is to find a control policy which yields a feature vector close to that of the observed trajectory. The method is initialized by selecting a random cost function parameter vector $\beta^{(0)}$ and solving the forward problem (3.1) to obtain an initial predicted trajectory $(x^{(0)}, u^{(0)})$ and associated feature vector $\mu^{(0)}$. On the i-th iteration, solve the following quadratic program:

$$\begin{aligned} \underset{\beta^{(i)}, b^{(i)}}{\text{minimize}} \quad & \|\beta^{(i)}\|^2 \\ \text{subject to} \quad & (\beta^{(i)})^T \mu^* \leq (\beta^{(i)})^T \mu^{(j)} - b^{(i)} \\ & \qquad \text{for} \quad j = 0, \ldots, i-1 \\ & b^{(i)} > 0 \end{aligned} \tag{3.8}$$

where $b^{(i)}$ is the margin on the i-th iteration, and μ^* and $\mu^{(j)}$ are the feature vectors of the optimal trajectory and j-th trajectory, respectively (recall the definition given in (3.6)). If $b^{(i)} < \epsilon$, then terminate. Otherwise, given the result from the quadratic program, $\beta^{(i)}$, solve the forward optimal control problem, Equation (3.1), with $\beta = \beta^{(i)}$ to obtain the predicted trajectory $(x^{(i)}, u^{(i)})$ and associated feature vector $\mu^{(i)}$. Set $i = i + 1$ and repeat.

As shown in [15], this method terminates after a finite number of iterations (the theorems stated and proved in [15] carry over with minor modification to the deterministic continuous-time case, which we omit for brevity). Upon termination, this algorithm returns a set of policies Π, and there exists at least one policy in Π that yields a feature vector differing from the expert's by no more than ϵ.

3.3.3 Method of Ratliff et al.

The maximum margin planning method of Ratliff et al. [16] is an inverse optimal control method that learns a cost function for which the expert policy has lower expected cost than every alternative policy by a margin that scales with the *loss* of that policy. As in

the method of Abbeel and Ng, we begin with the following quadratic program:

$$\begin{aligned} \underset{\beta,b}{\text{minimize}} \quad & \|\beta\|^2 \\ \text{subject to} \quad & \beta^T\mu^* \leq \beta^T\mu(x,u) - b \\ & \qquad \text{for all} \quad (x,u) \in S, \end{aligned}$$

where S is some set of trajectories. However, instead of considering a finite collection of trajectories, consider all possible trajectories (x,u) s.t. $\dot{x}(t) = f(t,x(t),u(t))$. Then the constraints in the quadratic program are satisfied if

$$\beta^T\mu^* \leq \min_{\substack{(x,u) \\ \dot{x}=f(x,u)}} \left(\beta^T\mu(x,u) - b\right).$$

Also, instead of trying to find the maximum fixed margin b, consider a margin which depends on the trajectory, let $b = L(x,u)$, where $L(x,u)$ denotes a loss function. This loss function typically specifies the closeness of a trajectory (x,u) to the observed trajectory (x^*,u^*), i.e. the loss function is zero near the observed trajectory and increases gradually to 1 away from the observed trajectory. Finally, slack variables ζ are introduced to allow constraint violations. The problem is now given by

$$\begin{aligned} \underset{\beta,\zeta}{\text{minimize}} \quad & \zeta + \frac{\lambda}{2}\|\beta\|^2 \\ \text{subject to} \quad & \beta^T\mu^* \leq \min_{\substack{(x,u) \\ \dot{x}=f(x,u)}} \left(\beta^T\mu(x,u) - L(x,u)\right) + \zeta, \end{aligned} \tag{3.9}$$

where $\lambda \geq 0$ is a constant that trades off between the penalizing constraint violations and a desire for small weight vectors. Since the slack variables are tight, they can be pulled into the objective function to obtain:

$$J(\beta) = \lambda\|\beta\|^2 + \beta^T\mu^{*(i)} - \min_{\substack{(x,u) \\ \dot{x}=f(x,u)}} \left\{\beta^T\mu(x,u) - L(x,u)\right\}. \tag{3.10}$$

This convex program can be solved using subgradient descent. As shown in [16], a subgradient $g(\beta)$ of $J(\beta)$ is given by $g(\beta) = \mu^* - \widehat{\mu} + \lambda\beta$, where $\widehat{\mu}$ represents the solution to $\arg\min_\mu \left(\beta^T\mu + L(\mu)\right)$, i.e. the solution to the forward optimal control problem (3.1) with cost function augmented by the loss function. The unknown parameter β is then updated iteratively using subgradient descent $\beta^{(i+1)} = \beta^{(i)} - \alpha_i g(\beta^{(i)})$.

Ratliff et al. [16] show that for constant step size α, this method achieves linear

convergence to some neighborhood of the minimum cost. In addition, they show that for diminishing step size $\alpha_j = 1/j$, the method will converge to a local minimum at a sub-linear rate.

3.4 Proposed Method Based on Necessary Conditions for Optimality

The three methods described in the previous section shared common structure. In particular, each method solves a forward optimal control problem repeatedly in an inner loop. They do this in order to compare the observed trajectory (or feature vectors) with predicted trajectories given a candidate cost function. In this section, we derive another approach inspired by recent work in inverse convex optimization by Keshavarz et al. [17]. This approach was developed in collaboration with Miles Johnson in [41]. The key idea in our approach is that we assume that the observations are perfect measurements of the system evolution, and that the expert is only *approximately optimal*, where we define what it means to be approximately optimal below. Under this new set of assumptions, we can immediately say how optimal the agent is by looking at how well the demonstration trajectory satisfies the necessary conditions of optimality. To do this, we use the necessary conditions to define a set of residual functions. The inverse optimal control problem is then solved by minimizing these residual functions over the unknown parameters. In the remainder of this section, we will describe these different stages in detail.

3.4.1 Inverse Optimal Control Formulation

Consider a trajectory (x, u) of the system given in Equation (3.2). The minimum principle gives us necessary conditions for (x, u) to be a local minimum of Eq. (3.1) [84, 85]. The Hamiltonian function for the problem we consider is

$$H(x, u, p) = \beta^T \phi(t, x, u) + p^T f(t, x, u). \tag{3.11}$$

For a given β, if (x^*, u^*) is optimal, the necessary conditions for optimal control state that there exists a costate trajectory $p^* : \mathbb{R} \to \mathbb{R}^n$ such that

$$0 = \dot{p}^*(t)^T + \nabla_x H\left(x^*(t), u^*(t), p^*(t)\right) \tag{3.12}$$

$$0 = \nabla_u H\left(x^*(t), u^*(t), p^*(t)\right). \tag{3.13}$$

We apply these necessary conditions to our problem (3.1) to obtain

$$\begin{aligned} 0 &= \dot{p}(t)^T + \beta^T \nabla_x \phi\left[t, x(t), u(t)\right] + p(t)^T \nabla_x f\left[t, x(t), u(t)\right] \\ 0 &= \beta^T \nabla_u \phi\left[t, x(t), u(t)\right] + p(t)^T \nabla_u f\left[t, x(t), u(t)\right]. \end{aligned}$$

If $(x(t), u(t))$ is optimal, these conditions will be satisfied. If the trajectory is *approximately optimal*, these conditions are approximately satisfied. We formalize this by defining *residual functions* from these necessary conditions. Our method consists in minimizing the extent to which observed trajectories violate these necessary conditions, i.e. minimizing the extent to which the residual functions are not equal to zero (where we define what it means to be close to zero in Section 3.4.2). Let

$$z(t) = \begin{bmatrix} \beta \\ p(t) \end{bmatrix} \in \mathbb{R}^{k+n} \qquad\qquad v(t) = \dot{p}(t) \in \mathbb{R}^n.$$

The residual function $r[z(t), v(t)]$ is then defined as

$$\begin{aligned} r[z(t), v(t)] &= \begin{bmatrix} \nabla_x \phi\Big|_{(x,u)}^T & \nabla_x f\Big|_{(x,u)}^T \\ \nabla_u \phi\Big|_{(x,u)}^T & \nabla_u f\Big|_{(x,u)}^T \end{bmatrix} z(t) + \begin{bmatrix} I \\ 0 \end{bmatrix} v(t) \\ &= F(t) z(t) + G(t) v(t), \end{aligned} \tag{3.14}$$

where we have just rearranged the necessary conditions. The notation $(\cdot)\big|_{(x,u)}$ is shorthand for evaluating the particular function along the trajectory given in the observation, i.e. $\nabla_x \phi\Big|_{(x,u)} \equiv \nabla_x \phi\left[t, x(t), u(t)\right]$.

This formulation can also be extended to handle multiple observations. Consider M trajectories, which may have different boundary conditions, but have the same fixed final time t_f $\left\{\left(x^{(i)}, u^{(i)}\right)\right\} \quad i = 1, \ldots, M$ with each $(x^{(i)}, u^{(i)}) = \left\{x^{(i)}(t), u^{(i)}(t) : t \in [0, t_f]\right\}$. The development is the same and the unknown parameters $z(t)$ and $v(t)$ are simply extended to include the unknown costate trajectories of the M observations. The par-

ticular structure of the residual function is such that the amount of computation grows linearly with the number of observations, just as it does with the three prior methods of inverse optimal control. We now describe how our method minimizes these residual functions over the unknown parameters $\beta, p(t), \dot{p}(t)$ in order to best satisfy necessary conditions for optimality.

3.4.2 Residual Optimization

To solve for the unknown parameters $z(t)$ and $v(t)$ that cause the observed trajectories to best satisfy necessary conditions for optimal control, we solve the following problem:

$$\begin{aligned} \underset{z(t),v(t)}{\text{minimize}} \quad & \int_{t_0}^{t_f} \|r[z(t), v(t)]\|^2 dt \\ \text{subject to} \quad & \dot{z}(t) = \begin{bmatrix} 0 \\ I \end{bmatrix} v(t) \\ & z(0) = z_0 \qquad \text{(unknown)}, \end{aligned} \tag{3.15}$$

where

$$\|r[z(t), v(t)]\|^2 = z^T F^T F z + v^T G^T G v + z^T F^T G v, \tag{3.16}$$

where the argument t has been dropped for brevity. If $z(0)$ were known, this would be a standard LQR problem (with cross terms):

$$\begin{aligned} \min_{z(t),v(t)} \quad & \int_{t_0}^{t_f} \left\{ z^T Q z + v^T R v + z^T S v \right\} dt \\ \text{subject to} \quad & \dot{z}(t) = Az + Bv \\ & z(0) = z_0, \end{aligned} \tag{3.17}$$

where

$$\begin{aligned} A(t) &= 0 & B(t) &= \begin{bmatrix} 0 \\ I \end{bmatrix} \\ Q(t) &= F(t)^T F(t) & R(t) &= G(t)^T G(t) \\ S(t) &= F(t)^T G(t). \end{aligned}$$

Solving this LQR problem yields the linear control policy and quadratic value function

$$v(t) = K(t)z(t) \qquad\qquad V(z_0) = z_0^T P(0) z_0,$$

where

$$K(t) = -(G(t)^T G(t))^{-1} \left(G(t)^T F(t) + B(t)^T P(t) \right),$$

and where $P(t)$ represents the solution to the LQR Riccati equation. We complete our solution for $z(t)$ by solving the following problem

$$\underset{z_0}{\text{minimize}} \quad z_0^T P(0) z_0.$$

Without normalization, this quadratic program is satisfied by the trivial solution $z_0 = 0$. Normalization is performed by using prior knowledge about the problem domain. For example, when the forward optimal control problem has a quadratic cost function, one can often assume that one of the weights is equal to 1.

3.5 Simulation Experiments

3.5.1 Evaluating Methods under Noise-free Observations

To evaluate the performance of the three recent inverse optimal control methods described in Section 3.3 and the new method introduced in Section 3.4, we perform numerical simulations in which we observe optimal trajectories of three different systems and learn the objective function for each system. For each system, we collect the optimal trajectories, i.e. noise-free observations, by simulating the system acting under the optimal control policy for particular boundary conditions and fixed terminal time. We collect simulations for 50 random boundary conditions. These experiments form a baseline of comparison, whose results can be used to understand the fundamental behavior of each method.

3.5.2 Robustness under Cost Perturbation

To evaluate the robustness of the four methods, we perform the following perturbation to the inverse optimal control problem. Up to this point we have considered the true cost function to be perfectly modeled by the weighted combination of known basis functions,

as shown in Equation (3.1). In our perturbation simulations, we assume that the given cost basis functions are only an approximation to the true cost function. In particular, we set the true cost function to be

$$J(u) = \int_{t_0}^{t_f} \beta^T \phi[t, x(t), u(t)] + d^T \rho[x(t), u(t)] dt, \tag{3.18}$$

where $\rho : \mathcal{X} \times \mathcal{U} \to [0,1]^l$ are perturbation basis functions and $d \in \mathbb{R}^l$ are perturbation weights such that $\|d\| < \epsilon$ for some $\epsilon > 0$. In particular, we model a general perturbation with a linear combination of k-th order multivariate Fourier basis functions. The multivariate basis functions are defined as

$$\rho_i[z(t)] = \begin{cases} 1 & \text{if } i = 0 \\ 1 + \cos(2\pi a^i \cdot z) & \text{for odd } i \\ 1 + \sin(2\pi a^i \cdot z) & \text{for even } i \end{cases} \tag{3.19}$$

for $i = 1, \ldots l$, where $a^i = [a_1, \ldots, a_{n+m}]$, each $a_j \in [0, \ldots, l]$. Here z is the concatenation of the state and input vectors at time t, $z(t) = [x(t), u(t)]$. A particular set of basis functions is formed by systematically varying the elements in each a^i, and assuming only one nonzero element of a^i for each i.

3.5.3 Three Example Systems

The three systems we use are (a) linear quadratic regulation, (b) regulation of a kinematic unicycle, (c) calibration of an elastic rod. We now describe the forward optimal control problem of each of these systems.

Linear Quadratic Regulation

In our first system, we consider a linear system with quadratic cost

$$\begin{aligned} \underset{x,u}{\text{minimize}} \quad & \int_{t_0}^{t_f} x^T Q x + u^T R u \\ \text{subject to } & \dot{x}(t) = Ax(t) + Bu(t), \\ & x(0) = x_{start} \\ & x(t_f) = \text{Free}, \end{aligned} \tag{3.20}$$

Table 3.1: Results for perfect observations with known basis functions.

System	Error Type	Mombaur	Abbeel	Ratliff	New
LQR	computation (s)	280	68	117	4
	forward problems	129	28	48	0
	parameter error	7.03e-2	1.71e-1	6.99e-1	6.35e-8
	feature error	2.30e-3	3.07e-3	1.15e-1	2.81e-9
	trajectory error	1.36e-5	1.04e-4	2.64e-2	1.04e-16
Unicycle	computation (s)	448	63	280	2
	forward problems	133	20	100	0
	parameter error	3.27-2	5.12e-1	5.23e-1	2.54e-5
	feature error	3.53e-3	1.69e-2	1.42e-2	1.03e-5
	trajectory error	1.55e-5	1.12e-3	4.64e-3	8.09e-10
Elastic Rod	computation (s)	95	9	15	1
	forward problems	71	5	10	0
	parameter error	3.38e-2	8.92e-1	9.71e-1	3.96e-5
	feature error	6.77e-7	6.24e-3	4.48e-3	4.87e-7
	trajectory error	1.94e-5	7.95e-3	8.82e-3	6.14e-6

where states are denoted by $x(t) \in \mathbb{R}^n$ and control inputs are denoted by $u(t) \in \mathbb{R}^m$. The matrices A and B are assumed time-invariant, with elements drawn from a $N(0,1)$ Gaussian distribution for each trial. Matrix A is scaled such that $|\lambda_{max}(A)| < 1$, and controllability of the system is verified manually. The initial state of the system x_0 for each trial is drawn from a $N(0,5)$, and the final time $t_f = 10$ is fixed for all trials. Moreover, for each trial we select cost matrices Q and R, with diagonal elements generated according to the uniform distributions of $U[0,1]$ and $U[\epsilon,1]$ respectively, to obtain nonnegative-definite and positive-definite matrices Q and R.

Quadratic Regulation of the Kinematic Unicycle

As our second test system, we consider quadratic regulation of the kinematic unicycle

$$\begin{aligned} \underset{x,u}{\text{minimize}} \quad & \int_{t_0}^{t_f} x^T Q x + u^T R u \\ \text{subject to } \dot{x}(t) = & \begin{bmatrix} \cos x_3(t) \\ \sin x_3(t) \\ u(t) \end{bmatrix}, \\ & x(0) = x_{start} \\ & x(t_f) = \text{free}, \end{aligned} \tag{3.21}$$

where states are denoted by $x(t) \in \mathbb{R}^3$ (with $x_i(t)$ representing the i-th element of the vector $x(t)$), and control inputs are denoted by $u(t) \in \mathbb{R}$. We generate initial conditions and cost parameters in a similar manner to the linear quadratic regulation problem.

Calibrating an Elastic Rod

Our third test system considers a thin, flexible wire of fixed length that is held at each end by a robotic gripper, which we call an elastic rod [86]. Assuming that it is inextensible and of unit length, we describe the shape of this rod by a continuous map $q\colon [0,1] \to G$, where $G = SE(3)$. As defined in [86], let L_q denote the left translation map $L_q : G \to G$. Let e denote the identity element of G, and let $\mathfrak{g} = T_eG$ and $\mathfrak{g}^* = T_e^*G$. Abbreviating $T_eL_q(\zeta) = q\zeta$ as usual for matrix Lie groups, we require this map to satisfy

$$\dot{q} = q(u_1X_1 + u_2X_2 + u_3X_3 + X_4) \tag{3.22}$$

for some $u\colon [0,1] \to U$, where $U = \mathbb{R}^3$ and X_i are the usual basis for $\mathfrak{g}$. We refer to q and u together as $(q,u)\colon [0,1] \to G \times U$ or simply as (q,u). Each end of the rod is held by a robotic gripper. We ignore the structure of these grippers, and simply assume that they fix arbitrary $q(0)$ and $q(1)$. We further assume, without loss of generality, that $q(0) = I_{4\times4}$. Finally, we assume that the rod is elastic in the sense of Kirchhoff [87], so has total elastic energy

$$\frac{1}{2}\int_0^1 \left(\beta_1 u_1^2 + \beta_2 u_2^2 + \beta_3 u_3^2\right) dt$$

for given constants $\beta_1, \beta_2, \beta_3 > 0$. For fixed endpoints, the rod will be motionless only if its shape locally minimizes the total elastic energy. In particular, we say that (q,u) is

in static equilibrium if it is a local optimum of

$$\begin{aligned} \underset{q,u}{\text{minimize}} \quad & \frac{1}{2}\int_0^1 \left(\beta_1 u_1^2 + \beta_2 u_2^2 + \beta_3 u_3^2\right) dt \\ \text{subject to} \quad & \dot{q} = q(u_1 X_1 + u_2 X_2 + u_3 X_3 + X_4) \\ & q(0) = e, \qquad q(1) = b \end{aligned} \tag{3.23}$$

for some $b \in \mathcal{B}$.

As mentioned in Section 3.1, recent work [32] has tackled a similar problem which used a different model and a solution method analogous to the method of Mombaur et al.

3.6 Results and Discussion

3.6.1 Perfect Observations with Known Basis Functions

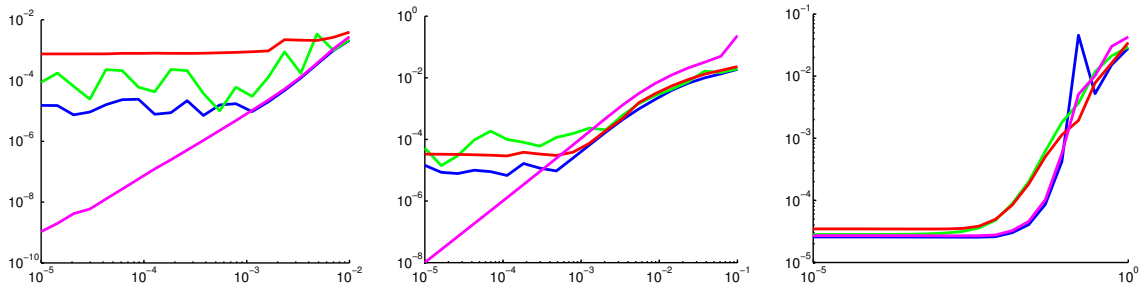
In this set of experiments, each algorithm was given one perfect observation of an optimal trajectory and learned the unknown cost function parameters β. After learning the cost function, predicted trajectories are computed. This allows us to compute other statistics such as the error in total cost, error in feature vectors, and sum squared error between observed and predicted trajectories.

Table 3.1 shows results averaged over 50 trials with randomly selected boundary conditions in each trial. In the method of Mombaur, the sum-squared error between predicted and observed trajectories converges near zero as the number of iterations increases. However, the inferred cost function parameters are not learned perfectly. Similarly, upon termination of the methods of Abbeel and Ratliff, the error between predicted and observed feature vectors is small, but the cost function parameters are not learned perfectly.

The new method developed in this work learns the unknown parameters perfectly (within the accuracy and precision tolerances of ODE and least squares solvers).

3.6.2 Perfect Observations with Perturbed Cost

In this set of experiments, the true cost function consists of a linear combination of known basis functions plus a bounded deterministic perturbation (see Section 3.5.2). For each system, one particular set of boundary conditions was selected, and observations



(a) LQR regulation, trajectory error.

(b) Unicycle regulation, trajectory error.

(c) Elastic rod calibration, trajectory error.

Figure 3.1: This figure shows how the sum-squared error between observed and predicted trajectories vary under perturbations of increasing magnitude. Blue, Green, Red, and Magenta curves correspond to the methods of Mombaur, Abbeel, Ratliff, and the new technique developed in this chapter, respectively.

of optimal trajectories are gathered for a range of perturbation magnitudes. Figure 3.1 shows the performance of each method over varying magnitude perturbations. These results generally show:

- All of the methods learn cost functions which are able to approximate the observation in terms of feature vector and trajectory errors.
- The performance of the iterative methods remains close to the results obtained with known basis functions for small perturbations, and then degrades at larger perturbations.
- The performance of our new method (KKT) continues to improve as the perturbation decreases, reflecting exact recovery of the cost function (to specified numerical method tolerances).

Note that in the case of the elastic rod, all of the methods, including our new approach, stop improving as the perturbation magnitude gets small. This trend occurs because the numerical method for solving the forward optimal control problem terminates before reaching the observed local minima under our standard convergence and tolerance parameters, which are fixed for all experiments.

3.7 Conclusion

In this chapter, we presented a new method of inverse optimal control, and compared the method to existing approaches using a set of canonical example systems. We com-

pared our new approach with the following methods: inverse reinforcement learning by Abbeel and Ng [15], maximum margin planning by Ratliff et al. [88], and inverse optimal control by Mombaur et al. [14]. These existing solution approaches search for values of the parameters that minimize the difference between predicted and observed trajectories (or state-action features). These approaches require solving a forward optimal control problem at each iteration. The approach presented in this work does not require the solution of any forward optimal control problem, and instead minimizes residual functions derived from first order necessary conditions for optimality. Our results show that the proposed method is better able to recover unknown parameters and is less computationally expensive than the existing methods. While our new method behaved well for the canonical systems used in this chapter, we acknowledge that there are a variety of situations in which it is not clear which method would be best.

There are a number of important questions we would like to pursue further. These include (a) investigating the existence and uniqueness of solutions of inverse optimal control, and (b) investigating how our method performs under imperfect observations, in particular sampled and noisy observations. In the next chapter, we show that by focusing on a class of systems called differentially flat systems, we can address these questions.

CHAPTER 4

INVERSE OPTIMAL CONTROL FOR DIFFERENTIALLY FLAT SYSTEMS

4.1 Introduction

In this chapter, we introduce an approach to inverse optimal control for the class of differentially flat systems [35, 36]. For such systems, the deterministic non-linear optimal control problem can be equivalently written as an *unconstrained* calculus of variations problem. We show that this equivalence has five important consequences for the corresponding inverse optimal control problem. First, we demonstrate that the solution of the inverse optimal control problem becomes "easy", in fact reducing to finite-dimensional linear least-squares minimization. Second, we precisely derive conditions for unique estimation of the cost parameters given a set of demonstrations and basis functions for the cost. Third, we show that the problem of IOC for differentially flat systems does not explicitly depend on the dynamic equations, thus leading to a robust estimation in the presence of model perturbations. In this sense our algorithm is similar to model-free IOC algorithms [31, 89]. Fourth, we demonstrate how standard filtering techniques can be applied to the problem of IOC, to obtain better solutions in the presence of sampled data and measurement noise. Fifth, a recursive version of the IOC solution can be formulated for differentially flat systems, making this approach suitable for online robot learning applications.

The remainder of this chapter proceeds as follows. In section 4.2 we reiterate the problem of inverse optimal control for deterministic nonlinear systems. We then consider the class of differentially flat systems and introduce the problem of inverse optimal control for these systems in section 4.3. For this class of systems, and using the IOC formulation developed in the previous chapter, we introduce a solution approach in section 4.4. We demonstrate how this method works on an example in section 4.5 and we conclude the chapter in section 4.6.

4.2 Inverse Optimal Control Problem Statement

Consider the following deterministic non-linear continuous-time optimal control problem:

$$\min_{x,u} \quad \int_{t_0}^{t_f} \beta^{*T} \phi(x,u) dt \tag{4.1}$$

$$\begin{aligned} \text{subject to} \quad & \dot{x} = f(x,u), \\ & x(t_0) = a, \\ & x(t_f) = b, \end{aligned} \tag{4.2}$$

where $x \in \mathbb{R}^n$ represents the state of the system, $u \in \mathbb{R}^m$ represents the inputs, $f(x,u) : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ represents the system dynamics and $t_0, t_f \in \mathbb{R}^+$ represent the initial and final time. The known basis functions $\phi(x,u) : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^p$ together with the cost parameters $\beta^* \in \mathbb{R}^p$ represent a cost function to be minimized. Note that for simplicity of presentation, additional state-input constraints on the optimization (4.1) have not been included.

4.3 Inverse Optimal Control for Differentially Flat Systems

We now introduce the problem of inverse optimal control for the class of *differentially flat* systems. We show how by focusing on the class of differentially flat systems, we can represent the optimal control problem (4.1) as an *equivalent* unconstrained calculus of variation problem. Subsequently for this class of systems, we introduce an alternative formulation of the IOC problem that is based on minimizing the extent to which first-order necessary conditions of optimality are violated.

A system is differentially flat [35, 36] if all the states x and inputs u can be obtained from the outputs without integration. More precisely, a system is differentially flat if for outputs $y = h(x, u, ..., u^{(j)}) \in \mathbb{R}^m$ there exists functions γ_x and γ_u such that

$$x = \gamma_x(y, \dot{y}, ..., y^{(k)}) \tag{4.3}$$

$$u = \gamma_u(y, \dot{y}, ..., y^{(k)}) \tag{4.4}$$

for some finite k-th order time derivative of the outputs denoted by $y^{(k)}$. Note that for a system to be differentially flat, the number of outputs q should equal the number of inputs m, i.e. the number of equations by which the ordinary differential equation (ODE) $\dot{x} = f(x,u)$ is under-determined. In this thesis, we will use the terms differentially

flat and flat interchangeably. Some examples of flat systems include fully actuated Lagrangian systems such as a fully actuated five-link biped, the car with N trailers, and a kinematic chain [90].

For such systems, the optimal control problem (4.1) is equivalent to the following calculus of variation problem on the flat outputs as discussed in [91, 92]

$$\begin{aligned} \min_{y,\dots,y^{(k)}} \quad & \int_{t_0}^{t_f} \beta^{*T} \phi(\gamma_x(y,\dots,y^{(k)}), \gamma_u(y,\dots,y^{(k)}))dt \\ \text{subject to} \quad & \gamma_x(y(t_0),\dots,y^{(k)}(t_0)) = a \\ & \gamma_x(y(t_f),\dots,y^{(k)}(t_f)) = b. \end{aligned} \tag{4.5}$$

Now the problem of inverse optimal control for flat systems is to find the unknown parameters β^* given the observed flat outputs y and known basis function $\phi(\gamma_x(Y), \gamma_u(Y))$. Here $Y = [1, y, \dot{y}, ..., y^{(k)}]^T$ denotes the outputs and their higher order derivatives (we have also added the constant 1 to specify that the functions γ_x and γ_u can have constant terms independent of the outputs). Note that since the optimization (4.5) is no longer constrained by the dynamics, the corresponding IOC problem does not explicitly depend on the dynamic equations and only requires observations of the flat outputs.

4.4 Proposed Solution for Inverse Optimal Control of Differentially Flat Systems

Next we describe the formulation of the IOC problem based on the least squares method introduced in the previous chapter and in our previous work [40, 41], for the case where the system is differentially flat. We proceed by writing the Euler-Lagrange equations corresponding to the first order necessary conditions of optimality of (4.5) for an unknown parameter β as follows:

$$\begin{aligned} & \beta^T \sum_{\ell=0}^{k} (-1)^\ell \frac{d^\ell}{dt^\ell} \frac{\partial}{\partial y_1^{(\ell)}} \phi\Big(\gamma_x(Y), \gamma_u(Y)\Big) = 0, \\ & \vdots \\ & \beta^T \sum_{\ell=0}^{k} (-1)^\ell \frac{d^\ell}{dt^\ell} \frac{\partial}{\partial y_m^{(\ell)}} \phi\Big(\gamma_x(Y), \gamma_u(Y)\Big) = 0, \end{aligned}$$

where $y^{(0)} = y$ and $\frac{d^0}{dt^0}y = y$. We can write the necessary conditions of optimality succinctly as

$$\beta^T \sum_{\ell=0}^{k} (-1)^\ell \frac{d^\ell}{dt^\ell} \Phi_{y^{(\ell)}} \Big(\gamma_x(Y), \gamma_u(Y) \Big) = \mathbf{0}, \tag{4.6}$$

where $\Phi_z(\gamma_x(Y), \gamma_u(Y))$ denotes the Jacobian of $\phi(\gamma_x(Y), \gamma_u(Y))$ with respect to the vector z, and $y^{(j)} = [y_1^{(j)}, y_2^{(j)}, ..., y_m^{(j)}]^T$ denotes the j-th derivative of the vector of outputs y. We further define the residual function $r(\beta, Y(t))$ as

$$r(\beta, Y(t)) = \beta^T \sum_{\ell=0}^{k} (-1)^\ell \frac{d^\ell}{dt^\ell} \Phi_{y^{(\ell)}} \Big(\gamma_x(Y), \gamma_u(Y) \Big).$$

To solve for the unknown parameters β, we formulate the IOC problem as minimizing the extent to which the first order necessary conditions of optimality are violated, i.e. by minimizing the L_2 norm of the residual as follows

$$\begin{aligned} \hat{\beta} = \arg\min_{\beta} \quad & \int_{t_0}^{t_f} \Big|\Big| r(\beta, Y(t)) \Big|\Big|^2 dt \\ \text{subject to} \quad & \beta_1 = 1. \end{aligned} \tag{4.7}$$

This IOC formulation can equivalently be seen as a parameter estimation problem for the ODE equation (4.6) given the observations Y. Different methods of ODE parameter estimation have been developed using maximum likelihood estimation [65] and least squares optimization [66]. The optimization (4.7) corresponds to a least squares method for ODE parameter estimation [66]. Also, note that since this formulation relies on the first-order necessary conditions of optimality, our approach can also be applied to observed *locally* optimal trajectories, similar to [93].

Note that the constraint on the parameter vector β (here on the first element of the vector β) is required to ensure that we find a non-trivial solution, since otherwise the all zero solution will minimize the objective. Also note that the optimization (4.5) has the same solution if we scale the cost by a positive number. Therefore, IOC in principle can only find β^* up to a scaling. The inclusion of the constraint thus ensures the possibility of having a unique solution. The choice of the constraint $\beta_1 = 1$ requires domain knowledge. However, the only domain knowledge necessary beforehand is that $\beta_1 \neq 0$, and the exact value of β_1 need not be known, again because cost functions scaled by a positive number lead to the same solution.

We should also note that additional constraints on the flat outputs can be added to optimization (4.5). Constraints that impose additional necessary conditions of optimality as a function of the unknown cost parameters can be incorporated in the residual minimization. Otherwise if they are not functions of the cost parameters, they will not change the IOC optimization.

4.4.1 IOC Solution Reducing to Linear Least Squares

In this section, we derive the least squares solution to optimization (4.7). First we define $A^T(Y(t)) = -\Big[\sum_{\ell=0}^{k}(-1)^\ell \frac{d^\ell}{dt^\ell}\Phi_{y^{(\ell)}}\Big(\gamma_x(Y),\gamma_u(Y)\Big)\Big]_{2:p}$, and $b^T(Y(t)) = \Big[\sum_{\ell=0}^{k}(-1)^\ell \frac{d^\ell}{dt^\ell}\Phi_{y^{(\ell)}}\Big(\gamma_x(Y),\gamma_u(Y)\Big)\Big]_1$. Here the notation $\Big[X\Big]_1$ denotes the first row of the matrix X and $\Big[X\Big]_{2:p}$ denotes the second through p-th rows of the matrix X. Furthermore, we write $\beta^T = [1, \tilde{\beta}^T]$. The optimization (4.7) can equivalently be written as follows when we replace the constraint $\beta_1 = 1$ in the objective

$$\hat{\beta} = \arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} ||b(Y(t)) - A(Y(t))\tilde{\beta}||^2 dt. \tag{4.8}$$

The minimization can be written as

$$\min_{\tilde{\beta}} \int_{t_0}^{t_f} (b^T b - 2\tilde{\beta}^T A^T b + \tilde{\beta}^T A^T A\tilde{\beta}) dt.$$

Differentiating the convex objective with respect to the constant $\tilde{\beta}$ and equating to zero yields the equation

$$\Big(\int_{t_0}^{t_f} A^T(Y(t))A(Y(t))\ dt\Big)\hat{\beta} = \Big(\int_{t_0}^{t_f} A^T(Y(t))b(Y(t))\ dt\Big). \tag{4.9}$$

Equation (4.9), which is the continuous-time analog to the discrete-time linear least squares solution, thus solves the IOC problem in closed form. Note that standard Gaussian elimination or LU decomposition techniques can be used to find numerically stable solutions to the equation (4.9). Also note that if $A^T A$ is not bounded, we can get a bounded signal by normalization [94]. Thus from now on without loss of generality we will always assume $A^T A$ is bounded.

4.4.2 Structural Identifiability of Cost Functions and the Null Lagrangian

The formulation of IOC for differentially flat systems enables us to analyze the identifiability of the cost parameters. To define identifiability for the IOC problem, let $Y(t|\beta, a, b)$ denote the solution to the calculus of variation problem (4.5) given cost parameters β and boundary conditions equal to a and b. We define the cost parameters to be identifiable if

$$Y(t|\beta', a, b) \equiv Y(t|\beta'', a, b) \Rightarrow \beta' = \beta'' \tag{4.10}$$

for some boundary conditions a and b. This is analogous to parameter identification definitions for dynamic systems and differential equations [95, 96]. The issue of cost parameter identifiability (4.10) can be decomposed into two parts. In the first part, we are concerned with whether the true cost parameters can be identified, given the structure of the cost function (structure of the model) and irrespective of what signal, i.e. optimal trajectory, we observe. This is often referred to as *structural identifiability* in the literature [95]. Once we establish the structural identifiability, we can evaluate the identifiability of the parameters given observations of optimal trajectories. We will discuss this part in the next section.

To illustrate the issue of structure identifiability for IOC cost functions, we will discuss a simple example. Consider the following optimization problem, with unknown cost parameters β_1, β_2 and β_3:

$$\min_{y,\dot{y},\ddot{y}} \int_{t_0}^{t_1} \phi(y, \dot{y}, \ddot{y}) = \min_{y,\dot{y},\ddot{y}} \int_{t_0}^{t_1} (\beta_1 \dot{y}^2 - \beta_2 y\ddot{y} + \beta_3 y^3) dt.$$

We write the Euler-Lagrange equations corresponding to this optimization as

$$\phi_y - \frac{d}{dt}\phi_{\dot{y}} + \frac{d^2}{dt^2}\phi_{\ddot{y}} = \beta_2\ddot{y} + 3\beta_3 y^2 - 2\beta_1\ddot{y} + \beta_2\ddot{y} = 2(\beta_2 - \beta_1)\ddot{y} + 3\beta_3 y^2 = 0. \tag{4.11}$$

It is clear from equation (4.11) that the cost parameters in this model are not identifiable. In particular, for any set of parameters satisfying $\beta_2 - \beta_1 = c$, where c is a constant, we get the same observations Y, thus violating the identifiability definition (4.10). Alternatively, the optimization

$$\min_{y,\dot{y},\ddot{y}} \int_{t_0}^{t_1} \psi(y, \dot{y}, \ddot{y}) = \min_{y,\dot{y},\ddot{y}} \int_{t_0}^{t_1} (\tilde{\beta}_1 \dot{y}^2 + \tilde{\beta}_2 y^3) dt$$

has the same Euler-Lagrange equation

$$\psi_y - \frac{d}{dt}\psi_{\dot{y}} + \frac{d^2}{dt^2}\psi_{\ddot{y}} = -2\tilde{\beta}_1\ddot{y} + 3\tilde{\beta}_3 y^2, \tag{4.12}$$

if we define $\tilde{\beta}_3 = \beta_3$ and $\tilde{\beta}_1 = \beta_1 - \beta_2$. This model on the other hand is identifiable. Thus, from the perspective of modeling behavior using a cost function, these two cost functions can model the same behavior, while the model using ψ as the cost function is identifiable.

To understand identifiability of cost parameters, we define *null Lagrangians* [97]. We say a cost function or a basis function is a *null Lagrangian* if its corresponding Euler-Lagrange equation is identically zero. We will denote the Euler-Lagrange operator for a basis function ϕ_i by $EL(\phi_i)$. Now consider the optimal control model (4.5), where the cost is defined as $\beta^T\phi$, i.e. a weighted linear combination of p basis functions $\phi_i, i = 1, ..., p$.

Then we claim that a necessary condition for identifiability of the cost parameters is that no linear combination of the basis functions is a null Lagrangian. To verify this claim, assume that without loss of generality, there exists a linear combination of the first $k < p$ basis functions such that it is null Lagrangian, i.e.

$$EL(\sum_{i=1}^{k} a_i\phi_i) \equiv 0 \Rightarrow EL(\phi_1) = \sum_{j=2}^{k} -\frac{a_j}{a_1}EL(\phi_j),$$

where not all a_i are equal to zero, and without loss of generality we assume that $a_1 \neq 0$. The above equality can be replaced in the equations for the necessary condition of optimality for the calculus of variation problem to arrive at

$$\sum_{i=1}^{p} \beta_i EL(\phi_i) = \sum_{j=2}^{k}(\beta_j - \beta_1\frac{a_j}{a_1})EL(\phi_j) + \sum_{l=k+1}^{p} \beta_l EL(\phi_l) = 0. \tag{4.13}$$

Thus for two sets of parameters $\beta' = [0, c_2, ..., c_k, c_{k+1}, ..., c_p]$ and
$\beta'' = [1, c_2+\frac{a_2}{a_1}, ..., c_k+\frac{a_k}{a_1}, c_{k+1}, ..., c_p]$, we get the same observations Y since the necessary conditions of optimality are identical. Thus the model would not identifiable.

In our first example, the linear combination of the first two basis functions shown as $\dot{y}^2 - y\ddot{y}$ was a null Lagrangian, since their corresponding Euler-Lagrange equation is identically zero.

Different works [97] have derived analytical conditions for verifying null Lagrangians. However, finding such conditions in the general case is still an open question. Nonetheless

it is often easy to find null Lagrangians for a particular example of a cost function. We take this approach in the next chapter, and eliminate the basis functions that lead to having a null Lagrangian in a locomotion modeling application.

We would like to emphasize that it is the formulation of IOC for differentially flat systems that allows us to clearly evaluate the structural identifiability of the cost function. In particular, if we consider the same problem without any consideration for the flatness of the system, we have

$$\min_{x_1,x_2,u} \int_{t_0}^{t_f} (\beta_1 x_2^2 - \beta_2 x_1 u + \beta_3 x_1^3) dt$$
$$\text{subject to } \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = f(x,u) = \begin{bmatrix} x_2 \\ u \end{bmatrix},$$

where letting $y = x_1$ leads to the same differentially flat example. Writing the necessary conditions of optimality for this problem we get

$$\begin{aligned} \dot{p}_1(t) - \beta_2 u + 3\beta_3^2 x_1^2 &= 0 \\ \dot{p}_2(t) - 2\beta_1 x_2 + p_1(t) &= 0 \\ -\beta_2 x_1 + p_2(t) &= 0. \end{aligned}$$

We see from the necessary conditions that it is not clear whether the problem is structurally identifiable. This confusion in particular is due to the introduction of the unknown infinite-dimensional co-states, $p_1(t)$ and $p_2(t)$, since identifiability of ordinary differential equations with infinite-dimensional unknown functions is still a topic of active research.

4.4.3 Identifiability Given Observed Optimal Trajectories

After addressing the structural identifiability problem, we evaluate the identifiability of the cost parameters given observations of optimal trajectories. To do so, we note that if the matrix $\int_{t_0}^{t_f} A^T A dt$ in (4.9) is full rank, a unique estimate $\hat{\beta}$ is obtained. Also note that the true cost parameter β^* satisfies the necessary conditions of optimality (4.6), and therefore guarantees the existence of a solution to the optimization (4.8). Therefore when the matrix $\int_{t_0}^{t_f} A^T A dt$ is full rank we can conclude that the unique estimate equals the true cost parameter, i.e. $\hat{\beta} = \beta^*$ and that the cost parameters are identifiable. In practice, one can check the condition number of the matrix.

4.4.4 Solution in the Presence of Noise and Sampled Data

Next, we will discuss two methods for computing an estimate of the cost parameters when the observed outputs are sampled and noisy. Both methods can lead to consistent estimation of the cost parameter, and the choice of the method depends on the application of interest. The first method will be based on spline-fitting the observed outputs, and the second method will make use of output filtering.

Solution via Spline-Fitting

Consider observing sampled noisy outputs. We model these observations as $z(k) \in \mathbb{R}^p$ where $k = 1, ..., N$ as follows:

$$z(k) = y(t_k) + \epsilon_k \quad \text{where } t_0 \leq t_1 \leq ... \leq t_N \leq t_f \quad \text{and } \epsilon_k \sim N(0, \Sigma_p), \Sigma_p > 0. \tag{4.14}$$

The problem of finding an estimate of the cost parameter β in (4.6) can be considered as a parameter estimation problem for ODEs. An approach to parameter estimation using sampled noisy observations [66] proceeds in two steps:

1. Find a consistent nonparametric estimate of $y(t)$, which we denote by $\tilde{y}(t)$ and its derivatives $\tilde{y}^{(n)}(t)$ using observations $z(k)$.

2. Use the continuous-time estimates of the outputs to obtain an estimate of the unknown parameters.

The problem of finding consistent nonparametric estimates has been largely studied and it is shown that such estimates can be constructed using splines and polynomial regression approaches [98], among others. Moreover, consistent estimates of the derivatives of $y(t)$ can also be obtained by differentiating $\tilde{y}(t)$ as described in [99].

Using the described two-step approach, it is shown [66] that we can find consistent estimates of the ODE parameters, which correspond to cost parameters in our work.

Solution via Output Filtering

In solving the IOC problem via optimization (4.8) we make use of higher order derivatives (HOD) of the observed outputs $y(t)$. Since we often only observe $y(t)$ and not

its derivatives, we have to rely on numerical differentiation methods to estimate these HODs. This approach was used to do IOC via spline fitting.

In this section we develop an alternative method of IOC based on filtering, which does not require computing HODs. Note that the matrix $A(Y(t))$ and vector $b(Y(t))$ each have entries that are functions of $[y, \dot{y}, ..., y^{(r)}]$ up to some finite derivative of order r. We proceed by filtering the entries of $A(Y(t))$ and $b(Y(t))$ using an r-th order stable filter

$$\frac{1}{\Lambda(s)} = \frac{1}{s^r + \lambda_{r-1}s^{r-1} + \cdots + \lambda_0}. \tag{4.15}$$

We apply this filter to both sides of the necessary conditions of optimality (4.6) as follows:

$$\frac{1}{\Lambda(s)}\mathcal{L}\{b(Y(t))\} = \frac{b_{\mathcal{L}}(s)}{\Lambda(s)}Y_{\mathcal{L}}(s)$$
$$\frac{1}{\Lambda(s)}\mathcal{L}\{A(Y(t))\}\tilde{\beta} = \frac{A_{\mathcal{L}}(s)}{\Lambda(s)}Y_{\mathcal{L}}(s)\tilde{\beta}$$

where $\mathcal{L}$ represents the Laplace operator, and $A_{\mathcal{L}}(s)$ and $b_{\mathcal{L}}(s)$ denote matrices of same size as $A(Y(t))$ and $b(Y(t))$ respectively. The entries of $A_{\mathcal{L}}(s)$ and $b_{\mathcal{L}}(s)$ are polynomial functions of s, and satisfy equations $A_{\mathcal{L}}(s)Y_{\mathcal{L}}(s) = \mathcal{L}\{A(Y(t))\}$ and $b_{\mathcal{L}}(s)Y_{\mathcal{L}}(s) = \mathcal{L}\{b(Y(t))\}$. Therefore, using this approach, we can treat every entry of $\frac{A_{\mathcal{L}}(s)}{\Lambda(s)}$ and $\frac{b_{\mathcal{L}}(s)}{\Lambda(s)}$ as a filter that is being applied to the observed outputs, and we eliminate the need for numerical computation of derivatives. Moreover, we can use the well-established theory of signal processing to design filters that can handle different kinds of measurement noise. By defining the filter outputs

$$\tilde{b}(t) = \mathcal{L}^{-1}\{\frac{b_{\mathcal{L}}(s)}{\Lambda(s)}Y_{\mathcal{L}}(s)\}, \quad \tilde{A}(t) = \mathcal{L}^{-1}\{\frac{A_{\mathcal{L}}(s)}{\Lambda(s)}Y_{\mathcal{L}}(s)\},$$

we can now write the following IOC optimization, which can be solved as in (4.8):

$$\hat{\beta} = \arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} ||\tilde{b}(t) - \tilde{A}(t)\tilde{\beta}||^2 dt. \tag{4.16}$$

To apply this method to sampled data $z(k)$ which are generated according to (4.14), with the assumption that $t_i - t_{i-1} = \Delta t$, we make use of the bilinear transformation [100] $s = \frac{2}{\Delta t}\frac{z-1}{z+1}$ to convert the continuous-time filter to a discrete-time filter.

4.4.5 Solution with Multiple Demonstrations

The proposed approach can naturally be extended to learning cost parameters given multiple demonstrations. Consider having N demonstrations of the outputs denoted by ${}_1Y, ..., {}_NY$, each satisfying the calculus of variation problem (4.5) with possibly different initial conditions, i.e. for $n = 1, ..., N$,

$$\begin{aligned} \min_{{}_nY} \quad & \int_{t_0}^{t_f} \beta^{*T} \phi({}_nY) dt \\ \text{subject to} \quad & \gamma_x({}_nY(t_0)) = a_n, \\ & \gamma_x({}_nY(t_f)) = b_n. \end{aligned} \tag{4.17}$$

Then these observations ${}_nY$ will satisfy first order necessary conditions analogous to equation (4.6). We can write all the equations succinctly as

$$\begin{aligned} & b({}_1Y) - A({}_1Y)\tilde{\beta} = 0, \\ & \vdots \\ & b({}_NY) - A({}_NY)\tilde{\beta} = 0 \end{aligned}$$

for $b(Y)$ and $A(Y)$ defined as before. We further define $\bar{b}_N(Y(t)) = [b({}_1Y)^T, ..., b({}_NY)^T]^T$ and $\bar{A}_N(Y(t)) = [A({}_1Y)^T, ..., A({}_NY)^T]^T$. Then, analogous to the case with $N = 1$, the IOC optimization becomes

$$\hat{\beta} = \arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} ||\bar{b}(Y(t)) - \bar{A}(Y(t))\tilde{\beta}||^2 dt, \tag{4.18}$$

and the unique estimate is obtained from

$$\hat{\beta} = \Big(\int_{t_0}^{t_f} \bar{A}^T \bar{A} \; dt \Big)^{-1} \Big(\int_{t_0}^{t_f} \bar{A}^T \bar{b} \; dt \Big),$$

given that the matrix $\int_{t_0}^{t_f} \bar{A}^T \bar{A} \; dt$ is full rank. Clearly, the discussed approach to output filtering can also be used in learning from multiple demonstrations.

4.4.6 Recursive Solution

Real-time learning is an important topic in robotics, since robots often encounter new situations and need to be equipped with the ability to self-improve [101]. In this section,

we develop a recursive method of IOC for flat systems. This method will allow us to update the estimate of the cost parameters upon receiving new observations without resolving the complete IOC problem.

As we described, to solve IOC given data $Y(s)$ from time t_0 to t (i.e. $t_0 \leq s \leq t$) we solve the following optimization

$$\hat{\beta}(t) = \arg\min_{\tilde{\beta}} \int_{t_0}^{t} ||b(Y(s)) - A(Y(s))\tilde{\beta}||^2 ds.$$

To develop a recursive version of the IOC, we propose minimizing the following objective with respect to $\tilde{\beta}$

$$\begin{aligned} J(\tilde{\beta}) = &\int_{t_0}^{t} ||b(Y(s)) - A(Y(s))\tilde{\beta}||^2 ds \\ &+ \frac{1}{2}(\tilde{\beta} - \beta_0)^T Q_0 (\tilde{\beta} - \beta_0), \end{aligned}$$

where we have added an extra cost to penalize deviation from some initial guess β_0 using a positive definite matrix $Q_0 > 0$. Note that given multiple demonstrations, we can solve the recursive IOC for the current demonstration, and use the estimated cost parameter as an initial guess when given the subsequent demonstration.

The objective $J(\tilde{\beta})$ is convex in $\tilde{\beta}$ and we set the gradient to zero to find the following equation that solves for $\hat{\beta}$ up to time t

$$\hat{\beta}(t) = \Big(\int_{t_0}^{t} A^T A ds + Q_0 \Big)^{-1} \Big(\int_{t_0}^{t} A^T b ds + Q_0 \beta_0 \Big).$$

We define the matrix $P(t)$ as

$$P(t) := \Big(\int_{t_0}^{t} A^T A ds + Q_0 \Big)^{-1}.$$

Note that $P(t)$ is well defined because $A^T A \geq 0$ and $Q_0 > 0$. To compute the cost parameter estimate $\hat{\beta}(t)$ recursively, we obtain an expression for $\dot{P}$ as

$$\begin{aligned} &PP^{-1} = I \Rightarrow 0 = \frac{d}{dt}(PP^{-1}) = \dot{P}P^{-1} + PA^T A \\ &\Rightarrow \dot{P} = -PA^T AP, \end{aligned}$$

and subsequently obtain a differential equation for $\hat{\beta}$ as

$$\begin{aligned}\dot{\hat{\beta}} &= \dot{P}\Big(\int_{t_0}^{t} A^T b ds + Q_0\beta_0\Big) + PA^T b \\ &= -PA^T AP\Big(\int_{t_0}^{t} A^T b ds + Q_0\beta_0\Big) + PA^T b \\ &= -PA^T A\hat{\beta} + PA^T b.\end{aligned}$$

Thus the recursive IOC solution can be obtained from

$$\begin{aligned}\dot{\hat{\beta}} &= -PA^T A\hat{\beta} + PA^T b \\ \dot{P} &= -PA^T AP.\end{aligned} \tag{4.19}$$

Using theorem 4.3.2 in [94] we can show that the recursive solution $\hat{\beta}(t)$ converges to a constant $\bar{\beta}$, and if the matrix $A(t)^T$ is persistently exciting, the solution converges to the true cost parameter β^*. We say $A(t)^T$ is persistently exciting if for some $\alpha_0, T_0 > 0$ we have

$$\int_{t}^{t+T_0} A^T(s)A(s)ds \geq \alpha_0 T_0 I, \forall t.$$

4.5 Unicycle Example

In this section, we show how to apply the proposed IOC by considering a unicycle example. We consider the following optimal control problem:

$$\begin{aligned}\arg\min_{x,u,v} \int_{t_0}^{t_f} &\frac{\beta_1}{2}(x_1(t) - x_1^*)^2 \\ &+\frac{\beta_2}{2}(x_2(t) - x_2^*)^2 + \frac{\beta_3}{2}v^2(t)dt, \\ \text{subject to} \quad \dot{x}(t) &= \begin{bmatrix} v(t)\cos(x_3(t)) \\ v(t)\sin(x_3(t)) \\ u(t) \end{bmatrix} \\ x_1(0) &= a_1, x_2(0) = a_2 \\ x_1(1) &= b_1, x_2(1) = b_2.\end{aligned} \tag{4.20}$$

Here $x(t) \in \mathbb{R}^3$ denotes the state of the unicycle at time t, and $u(t), v(t) \in \mathbb{R}$ denote the control inputs at time t. The unknown cost parameters consist of $\{\beta_1, \beta_2, \beta_3, x_1^*, x_2^*\}$. A unicycle is differentially flat [36] with flat outputs $y_1(t) = x_1(t)$ and $y_2(t) = x_2(t)$. Thus, the optimal control problem (4.20) can equivalently be written as

$$\begin{aligned}
&\arg\min_{y_1,y_2} \int_{t_0}^{t_f} \phi_U(y_1(t), \dot{y}_1(t), y_2(t), \dot{y}_2(t))dt = \\
&\arg\min_{y_1,y_2} \int_{t_0}^{t_f} \Big[\frac{\beta_1}{2}(y_1(t) - x_1^*)^2 + \frac{\beta_2}{2}(y_2(t) - x_2^*)^2 \\
&\qquad\qquad + \frac{\beta_3}{2}(\dot{y}_1^2(t) + \dot{y}_2^2(t))\Big]dt \\
&\text{subject to} \quad y_1(0) = a_1, y_2(0) = a_2 \\
&\qquad\qquad y_1(1) = b_1, y_2(1) = b_2.
\end{aligned} \tag{4.21}$$

To find the cost parameters, we write the first order necessary conditions of optimality for (4.21) as

$$\begin{aligned}
\beta_1 y_1(t) - \beta_1 x_1^* - \beta_3 \ddot{y}_1(t) = 0, \\
\beta_2 y_2(t) - \beta_2 x_2^* - \beta_3 \ddot{y}_2(t) = 0.
\end{aligned} \tag{4.22}$$

We redefine the unknown parameters as $\tilde{\beta}_1 = \beta_1, \tilde{\beta}_2 = \beta_2, \tilde{\beta}_4 = -\beta_1 x_1^*$ and $\tilde{\beta}_5 = -\beta_2 x_2^*$ and we fix $\beta_3 = 1$. Now the necessary conditions of optimality are linear in the unknowns $\tilde{\beta}_i$ (although the original objective was nonlinear in the parameters) and we can solve

$$\begin{aligned}
&\Big(\int_0^1 \begin{bmatrix} y_1(t) & 0 \\ 0 & y_2(t) \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1(t) & 0 & 1 & 0 \\ 0 & y_2(t) & 0 & 1 \end{bmatrix} dt\Big)\hat{\beta} = \\
&\Big(\int_0^1 \begin{bmatrix} y_1(t) & 0 \\ 0 & y_2(t) \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ddot{y}_1(t) \\ \ddot{y}_2(t) \end{bmatrix} dt\Big).
\end{aligned} \tag{4.23}$$

Now we consider the case in which we only observe sampled noisy observations $z(k)$ and demonstrate the use of output filtering. The spline-fitting approach can be applied in a similar and straight-forward way as well. We proceed by applying the stable third-

order filter

$$\frac{1}{\Lambda(s)} = \frac{1}{(s+1)(s+2)(s+3)}$$

to the Laplace transform of equations (4.22) to get

$$\begin{aligned}
\tilde{\beta}_1 \frac{1}{\Lambda(s)}(Y_1(s) - y_1(t_0)) + \tilde{\beta}_4 \frac{1}{s\Lambda(s)} = \\
\frac{s^2}{\Lambda(s)} Y_1(s) - \frac{s}{\Lambda(s)} y_1(t_0) - \frac{1}{\Lambda(s)} \dot{y}_1(t_0), \\
\tilde{\beta}_2 \frac{1}{\Lambda(s)}(Y_2(s) - y_2(t_0)) + \tilde{\beta}_5 \frac{1}{s\Lambda(s)} = \\
\frac{s^2}{\Lambda(s)} Y_2(s) - \frac{s}{\Lambda(s)} y_2(t_0) - \frac{1}{\Lambda(s)} \dot{y}_2(t_0).
\end{aligned}$$

The four filters $\frac{1}{s\Lambda(s)}, \frac{1}{\Lambda(s)}, \frac{s}{\Lambda(s)}$ and $\frac{s^2}{\Lambda(s)}$ are converted to discrete-time filters using the discussed bilinear transformation. The IOC problem is then solved using the filtered outputs as described in Section 4.4.4. A Monte Carlo simulation was performed using 300 different boundary conditions and cost parameters that were uniformly randomly generated. Each experiment was repeated with different noise levels ranging from a signal-to-noise ratio (SNR) of 0 to 30. Fig. 4.1 shows that as we expect, the estimated cost parameter approaches its true value.

4.6 Conclusion and Application to Locomotion Modeling

In this chapter we introduced a new formulation of the problem of inverse optimal control for the case where the system is differentially flat. We showed that IOC for this class of systems has a number of desirable properties, including an efficient finite-dimensional linear least squares solution, robustness with respect to model perturbations and noise in sampled observations.

An important motivation for developing the proposed inverse optimal control algorithm was to learn prosthetic controllers for other modes of locomotion, such as stair ascent. By applying the proposed IOC method to the outputs of a five-link differentially flat biped, we can produce exemplar locomotion trajectories for stairs and ramps of different inclinations. These exemplar trajectories can then immediately be used in the prosthetic learning framework introduced in chapter 2 to find prosthetic controller parameters for different modes of locomotion. The proposed method of IOC based on

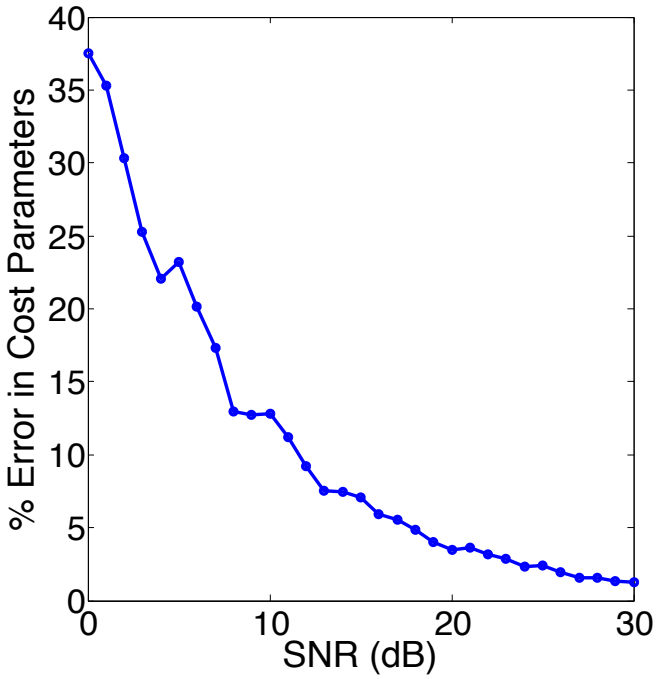


Figure 4.1: The results of estimating the cost parameters for an optimal control problem of a unicycle model. The observed trajectories were sampled, and corrupted by noise. These observations were then used in the IOC model to estimate cost parameters, and the percent error in estimating these parameters is shown vs. the SNR of the observed trajectory.

differential flatness is well-suited for this application. Due to the efficiency and scalability of the algorithm, we can learn cost functions composed of a large number of features. The learned cost function is also expected to work well for novel circumstances, due to the theoretical guarantees on estimation of unique cost parameters. Furthermore, the model-free nature of the algorithm is desirable since we often have data corresponding to multiple subjects, where the physical characteristics of the subjects, such as the mass of their body segments, are not exactly known. Using the proposed method, we do not need to model these unknown and complex physical dynamics of the subjects.

In the next chapter, we will describe how, by using sparse observations of stair ascent, we can learn an IOC model. Subsequently, we demonstrate the power of the model in predicting locomotion outputs for a novel stair height and verify the predictions by comparing to motion captured data.

CHAPTER 5

APPLICATION OF INVERSE OPTIMAL CONTROL TO PROSTHETIC CONTROLLER LEARNING DURING STAIR ASCENT

5.1 Introduction

To extend our method of prosthetic controller parameter learning to controllers for stair ascent, we need exemplar locomotion trajectories corresponding to the stair height that is observed at test time. To do so, we propose a model of human locomotion for stair ascent using the inverse optimal control method described in chapter 4. Using this model we produce exemplar joint trajectories corresponding to a desired stair height and subsequently learn controller parameters for that stair height. This proposed framework is demonstrated in Figure 5.1. This framework is used to learn prosthetic controller parameters for a subject, and experimental results with the Vanderbilt prosthetic leg are presented in this chapter.

We additionally show that the learned IOC model of locomotion can lead to a height-invariant controller for a prosthetic knee. Unlike the impedance control framework, which requires a unique choice of impedance controller parameters for each stair height, the proposed knee controller can generalize across different stair heights without changing any of the control parameters. We show the performance of this control strategy in a dynamic biped simulation and discuss its applicability to prosthetic control.

5.2 Inverse Optimal Control Model of Stair Ascent

To model locomotion using inverse optimal control, we have to make design choices regarding the dynamic model to be used and the basis functions used in to construct the cost function. We make use of a five-link fully actuated biped model shown in Fig. 5.2. Fully actuated mechanical systems are known to be differentially flat [102], thus we can use the IOC for differentially flat systems. While this is a simple model, we will demonstrate that through the use of the proposed IOC algorithm, we can learn a model

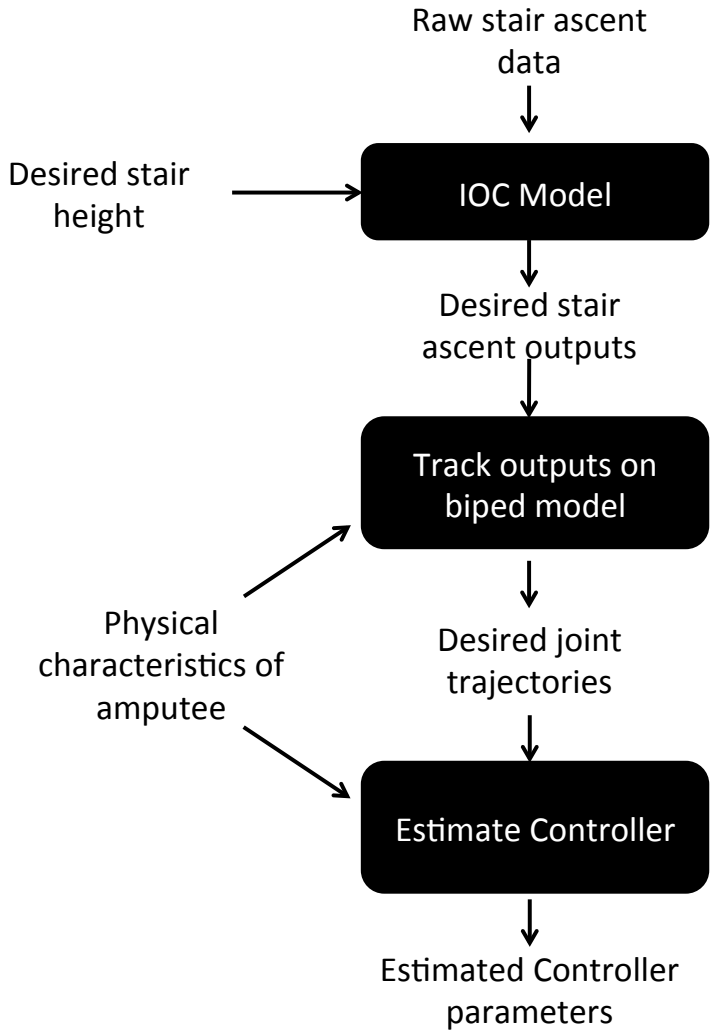


Figure 5.1: Prosthetic controller learning for stair ascent at a desired stair height.

for generating human locomotion trajectories.

The ability to model different behavior using inverse optimal control depends on having an appropriate set of basis functions for the cost. While learning basis functions would be desirable, learning appropriate basis functions still remains an open question. Furthermore, learning invariant basis functions requires a large number of observations, which is not available in our case since we are learning a model only using two locomotion examples.

Therefore in this work we draw inspiration from results in the locomotion literature to design basis functions. One important result is the existence of synergies between different body segments during locomotion. In particular, the work of [103] shows the existence of the law of kinetic covariance between the hip and knee torques. Moreover, the studies [33, 104–106] discuss an intersegmental coordination law that exists in different modes of locomotion. This law, which is termed the planar covariation of elevation angles, describes coordinated kinematic synergies between different segments of a biped, i.e. the thigh, shank and the foot, and is shown to be invariant across different subjects. These results suggest the possibility of designing a model to capture the coordination between the shank and the thigh elevation angles. Motivated by this result, we choose to work in the coordinate of elevation angles as the flat outputs, and the basis functions we design will be functions of the elevations angles as well. These angles are shown in Figure 5.2, where the subscripts T, S and tr denote thigh, shank and torso. More importantly, we design bases that are functions of both the thigh and the shank, thus modeling the relationship between these segments.

Another line of work [37] discusses the proximo-distal gradient. This principle suggests that proximal joints, i.e. joints that are closer to the point of attachment to the body such as the hip joint, are controlled in a feedforward manner, while distal joints, i.e joints that are more distant from the attachment point such as the knee and ankle, are more sensitive to local information such as perturbations from the environment. According to works such as [107], the modulation of a feedforward controller at the hip, for example modulation of central pattern generators, leads to producing locomotion for different modes and speeds. Inspired by these works, we make further design choices on the basis functions. In particular, we model the cost for stair ascent as having a component that is independent of the stair height, and a component that is modulated by the stair height. We assume that the cost component corresponding to the thigh is directly modulated by the stair height; however, the cost component corresponding to the shank is independent of the stair height, and the shank angle is implicitly modulated through the thigh.

We learn a cost function ϕ_L composed of the summation of the following basis functions

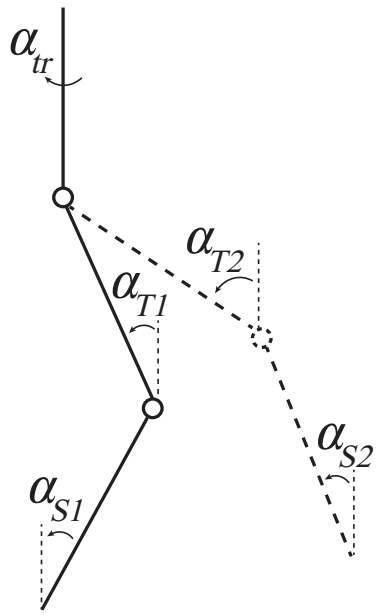


Figure 5.2: Elevation angles of the 5-link biped.

for both stance and swing phases separately. (We are omitting the stance and swing subscripts $i = 1, 2$ for notational simplicity.)

- $\phi_1(\alpha_S, \alpha_T)$: representing the component of the cost function capturing the relationship between the shank (α_S) and thigh (α_T) elevation angles.
- $\phi_2(\alpha_T, h)$: representing the component modulated by the stair height h.
- $\phi_3(\alpha_{tr})$: representing the cost for the torso.

Next we describe the different features that the functions ϕ_1, ϕ_2 and ϕ_3 are composed of weighted by cost parameters β_i.

Features Modeling Shank and Thigh Relationship

The cost component ϕ_1 equals the summation of the following features:

- $\beta_1(\alpha_S - \beta_2)^2$: denotes the cost of regulation of the shank angle to a set point.
- $\beta_3\dot{\alpha}_S^2$: denotes cost on angular velocity of the shank.
- $\beta_4\dot{\alpha}_T^2$: denotes cost on angular velocity of the thigh.
- $\beta_5\alpha_S\alpha_T + \beta_6\alpha_S\alpha_T^2 + \beta_7\alpha_S^2\alpha_T$: denotes monomial basis functions modeling the relationship between the shank and the thigh.

- $\beta_8 \sin(\alpha_S) + \beta_9 \cos(\alpha_S) + \beta_{10} \sin(\alpha_T) + \beta_{11} \cos(\alpha_T) + (\beta_{12} \sin(\alpha_S) + \beta_{13} \sin(\alpha_T))^2 + (\beta_{14} \cos(\alpha_S) + \beta_{15} \cos(\alpha_T))^2$: this feature corresponds to attractors for foot position, and can be derived by expanding cost on foot position regulation.

Features Modeling the Modulation of the Thigh

The thigh angle has an additional cost component $\phi_2(\alpha_T, h)$ which is a function of the stair height defined as

- $\beta_{16}(\alpha_T - \beta_{17}(h))^2$: denotes the cost of regulation of the thigh angle to a set point.

Features Modeling the Torso

We define the following features for the torso:

- $(\alpha_{tr} - \beta_{18})^2$: cost of regulation of torso to a set point.
- $\beta_{19} \dot{\alpha}_{tr}^2$: denotes cost on angular velocity of the torso.
- $\beta_{20} \sin(\alpha_{tr}) + \beta_{21} \cos(\alpha_{tr})$: denotes sinusoidal basis functions.

Now, given the discussed locomotion cost $\phi_L = \phi_1 + \phi_2 + \phi_3$, we consider the calculus of variation problem

$$\min_{\alpha_S, \alpha_T, \alpha_{tr}} \int_{t_0}^{t_f} (\phi_1(\alpha_S, \alpha_T) + \phi_2(\alpha_T, h) + \phi_3(\alpha_{tr}))dt. \tag{5.1}$$

The parameter $\beta_{15}(h)$ for swing is then varied to generate stair ascent trajectories for different inclinations. This parameter is independent of h during the stance phase. We apply the proposed method of IOC via cubic splines to solve for the unknown β cost parameters. In particular, we show how to solve two IOC problems for minimum and maximum stair heights, where the two problems share all cost parameters besides the height dependent cost parameter. We proceed by writing the first order necessary conditions of optimality for optimization (5.1).

$$\frac{\partial \phi_L}{\partial \alpha_S} - \frac{d}{dt}\frac{\partial \phi_L}{\dot{\alpha}_S} = 0 \Rightarrow \big[2\beta_1(\alpha_S - \beta_2) + \beta_5\alpha_T + \beta_6\alpha_T^2 + 2\beta_7\alpha_S\alpha_T + \beta_8\cos(\alpha_S)$$
$$- \beta_9\sin(\alpha_S) + 2\beta_{12}\cos(\alpha_S)(\beta_{12}\sin(\alpha_S) + \beta_{13}\sin(\alpha_T))$$
$$- 2\beta_{14}\sin(\alpha_S)(\beta_{14}\cos(\alpha_S) + \beta_{15}\cos(\alpha_T))$$
$$- 2\beta_3\ddot{\alpha}_S\big] = 0$$
$$\frac{\partial \phi_L}{\partial \alpha_T} - \frac{d}{dt}\frac{\partial \phi_L}{\dot{\alpha}_T} = 0 \Rightarrow \big[2\beta_{15}(\alpha_T - \beta_{16}(h)) + \beta_5\alpha_S + 2\beta_6\alpha_S\alpha_T + \beta_7\alpha_S^2 + \beta_{10}\cos(\alpha_T)$$
$$- \beta_{11}\sin(\alpha_T) + 2\beta_{13}\cos(\alpha_T)(\beta_{12}\sin(\alpha_S) + \beta_{13}\sin(\alpha_T))$$
$$- 2\beta_{15}\sin(\alpha_T)(\beta_{14}\cos(\alpha_S) + \beta_{15}\cos(\alpha_T))$$
$$- 2\beta_4\ddot{\alpha}_T\big] = 0.$$

The necessary conditions of optimality for the torso can be written in a similar fashion as well, independent of the shank and thigh angle. The above equations can be further expanded, collecting common terms together and redefining cost parameters to $\tilde{\beta}_i$ to arrive at the following equations:

$$\frac{\partial \phi_L}{\partial \alpha_S} - \frac{d}{dt}\frac{\partial \phi_L}{\dot{\alpha}_S} = 0 \Rightarrow \big[\tilde{\beta}_1 + \tilde{\beta}_2\alpha_S + \tilde{\beta}_4\alpha_T + \tilde{\beta}_5\alpha_T^2 + 2\tilde{\beta}_6\alpha_S\alpha_T + \tilde{\beta}_7\cos(\alpha_S)$$
$$+ \tilde{\beta}_8\sin(\alpha_S) + \tilde{\beta}_9\cos(\alpha_S)\sin(\alpha_T) + \tilde{\beta}_{10}\cos(\alpha_S)\sin(\alpha_S)$$
$$+ \tilde{\beta}_{11}\sin(\alpha_S)\cos(\alpha_T)\big] = \tilde{\beta}_{16}\ddot{\alpha}_S$$
$$\frac{\partial \phi_L}{\partial \alpha_T} - \frac{d}{dt}\frac{\partial \phi_L}{\dot{\alpha}_T} = 0 \Rightarrow \big[\tilde{\beta}_3\alpha_T + \tilde{\beta}_4\alpha_S + 2\tilde{\beta}_5\alpha_S\alpha_T + \tilde{\beta}_6\alpha_S^2 + \tilde{\beta}_9\cos(\alpha_T)\sin(\alpha_S) +$$
$$+ \tilde{\beta}_{11}\sin(\alpha_T)\cos(\alpha_S) + \tilde{\beta}_{12}\sin(\alpha_T)\cos(\alpha_T)$$
$$+ \tilde{\beta}_{13}\sin(\alpha_T) + \tilde{\beta}_{14}\cos(\alpha_T) + \tilde{\beta}_{15}(h)\big] = \tilde{\beta}_{17}\ddot{\alpha}_T.$$

Given these necessary conditions, we can now define the following vectors for maximum (denote by a superscript H for High) and minimum (denoted by a superscript L for Low) stair heights during swing:

$$A^T = \begin{bmatrix} 1 & 1 & 0 & 0 \\ \alpha_S^H & \alpha_S^L(t) & 0 & 0 \\ 0 & 0 & \alpha_T^H & \alpha_T^L \\ \alpha_T^H & \alpha_T^L & \alpha_S^H & \alpha_S^L \\ (\alpha_T^H)^2 & (\alpha_T^L)^2 & 2\alpha_S^H\alpha_T^H & 2\alpha_S^L\alpha_T^L \\ 2\alpha_S^H\alpha_T^H & 2\alpha_S^L\alpha_T^L & (\alpha_S^H)^2 & (\alpha_S^L)^2 \\ \cos(\alpha_S^H) & \cos(\alpha_S^L) & 0 & 0 \\ \sin(\alpha_S^H) & \sin(\alpha_S^L) & 0 & 0 \\ \cos(\alpha_S^H)\sin(\alpha_T^H) & \cos(\alpha_S^L)\sin(\alpha_T^L) & \cos(\alpha_T^H)\sin(\alpha_S^H) & \cos(\alpha_T^L)\sin(\alpha_S^L) \\ \cos(\alpha_S^H)\sin(\alpha_S^H) & \cos(\alpha_S^L)\sin(\alpha_S^L) & 0 & 0 \\ \sin(\alpha_S^H)\cos(\alpha_T^H) & \sin(\alpha_S^L)\cos(\alpha_T^L) & \sin(\alpha_T^H)\cos(\alpha_S^H) & \sin(\alpha_T^L)\cos(\alpha_S^L) \\ 0 & 0 & \sin(\alpha_T^H)\cos(\alpha_T^H) & \sin(\alpha_T^L)\cos(\alpha_T^L) \\ 0 & 0 & \sin(\alpha_T^H) & \sin(\alpha_T^L) \\ 0 & 0 & \cos(\alpha_T^H) & \cos(\alpha_T^L) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\tilde{\boldsymbol{\beta}}^T = \left[\tilde{\beta}_1,\ \tilde{\beta}_2,\ \tilde{\beta}_3,\ \tilde{\beta}_4,\ \tilde{\beta}_5,\ \tilde{\beta}_6,\ \tilde{\beta}_7,\ \tilde{\beta}_8,\ \tilde{\beta}_9,\ \tilde{\beta}_{10},\ \tilde{\beta}_{11},\ \tilde{\beta}_{12},\ \tilde{\beta}_{13},\ \tilde{\beta}_{14},\ \tilde{\beta}_{15}(h^H),\ \tilde{\beta}_{15}(h^L)\right],$$

and the following:

$$b = \begin{bmatrix} \ddot{\alpha}_S^H & \ddot{\alpha}_S^L & \ddot{\alpha}_T^H & \ddot{\alpha}_T^L \end{bmatrix}^T.$$

Here we have fixed $\tilde{\beta}_{16} = \tilde{\beta}_{17} = 1$. The unknown parameters can now be solved by the following equation:

$$\arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} ||b(t) - A(t)\tilde{\beta}||^2 dt. \tag{5.2}$$

Similar equations can be written for the stance phase, with the exception that we will only have one cost parameter $\tilde{\beta}_{15}$.

5.2.1 Results of Predicting Stair Ascent Trajectories Using Proposed and Existing IOC Approaches

We learn two IOC models for stair ascent locomotion using our proposed approach based on differential flatness and the method of Mombaur et al. [14]. The observed locomotion trajectories are motion-captured data collected in [108] and correspond to average joint trajectories of ten unimpaired subjects, climbing stairs at minimum, medium and maximum heights. We used output trajectories corresponding to the minimum and maximum stair heights as observations. The two IOC approaches are then applied to estimate the cost parameters given these two observations. To evaluate the performance of our estimated cost functions, we evaluate how well we can predict the observed trajectories in terms of R^2 and $RMSE$ as seen in Figures 5.3 and 5.4. Moreover, we used the learned models to predict a novel stair height (medium), and demonstrated the quality of fit to motion capture data ($R^2 = 0.97$ and $RMSE = 1.95$ degrees for the proposed approach and $R^2 = 0.88$ and $RMSE = 3.25$ degrees for the existing approach) as seen in Figures 5.5 and 5.6. To obtain locomotion trajectories corresponding to the new stair height we varied the cost parameter $\beta_{15}(h)$, which corresponds to the stair height, and evaluated the stair height that the generated trajectory accommodates. We can do this by looking at the height difference between the two feet at the end of the trajectory. We solved a numerical optimization in MATLAB to find the cost parameter $\beta_{15}(h)$ corresponding to the desired height. In future work, we can perform a regression with cost parameter $\beta_{15}(h)$ as the input and the resulting stair height as the output, and avoid having to solve this optimization repeatedly.

With this experiment, we would like to highlight some additional differences between the two IOC solutions. Note that the two IOC solutions were obtained using fundamentally different formulations of the estimation problem. The existing solution to IOC poses the problem as one of minimizing the trajectory error, i.e. minimizing the difference between the observed trajectories and the ones predicted by the model. In contrast, our proposed IOC formulation is based on minimizing the extent to which first order necessary conditions of optimality are violated, i.e. finding cost parameters that would make the norm of the residual as close as possible to zero. Due to this difference, the two approaches will often lead to different estimates of cost parameters when using real-world observations.

In our experiment with stair ascent locomotion data, we verified this difference between the approaches. First, we observed that the estimates of the cost parameters using the two approaches were significantly different. This can be due to the difference in for-

mulations, or due to the non-convex nature of the existing IOC approach. In particular, while the proposed IOC solution is always unique due to this closed-form expression, the existing IOC solution is based on a non-convex numerical optimization and can vary given different choices of initial cost parameters. To further illustrate the difference, we solved the non-convex optimization using the estimated cost parameters given by the proposed approach as the initial cost parameters. The final solution again was different than the closed-form solution given the proposed approach. This observation further verifies the fact that the two formulations lead to different models.

The question of which formulation is better suited for a task is a difficult one to answer. In modeling stair ascent locomotion, we found that the formulation based on minimizing violations of first-order necessary conditions of optimality performs better. This can be seen from the smaller error metrics obtained by our approach, in terms of both $RMSE$ and R^2. Moreover, we found that the cost parameters estimated using the existing trajectory error minimization approach led to ODEs that were unstable, as opposed to stable ODEs that were estimated using our proposed approach.

5.3 Trajectory Generation for Prosthetic Controller Learning

The learned model of stair ascent locomotion can be used in conjunction with the prosthetic controller learning framework discussed in chapter 2. We performed a prosthetic experiment with one male subject, with physical descriptions shown in Table 5.1. The experiments were held at the Rehabilitation Institute of Chicago using the Vanderbilt prosthetic device that was earlier discussed in chapter 2.

The staircase used in this experiment can be seen in Figure 5.7. The height of each stair was measured to be 19 centimeters (cm). We used this desired stair height in our learned locomotion model to produce desired stair ascent trajectories. The predicted joint trajectories, and in particular the knee trajectory that is of interest, can be seen in Figure 5.8.

Using these trajectories and the physical characteristics of the subject, we learned the impedance controller parameters using our proposed framework. Furthermore, similar to the walking experiments, we asked clinicians to experimentally tune the impedance controller parameters, given the learned parameters as an initial guess. The resulting learned and tuned impedance controller parameters are shown in Table 5.1.

Table 5.1: Learned (top) and experimentally modified (bottom) controller parameters for subject 1

Phase	$k[N{\cdot}m]$	$b[N{\cdot}m{\cdot}s]$	θ_e *[Degree]*
P1	4	0.276	0
P2	0.94	0.05	0.4
P3	0.45	0.01	65
P4	0.33	0.02	0

Phase	$k[N{\cdot}m]$	$b[N{\cdot}m{\cdot}s]$	θ_e *[Degree]*
P1	4	0.276	0
P2	2	0.05	0.4
P3	0.45	0.01	65
P4	0.25	0.01	0

5.4 Deriving a Height-Invariant Controller

We use the learned model of locomotion to learn a prosthetic knee controller that can generalize to different stair heights. The main intuition behind this approach is that since the locomotion model captures synergies between the thigh and the shank, we can find a controller for the shank (i.e. knee) that is a function of the shank angle and the thigh angle. Therefore, we can think of the thigh angle as passing information about the stair height to the shank (i.e. knee), and thus leading to a knee controller that is not a function of the stair height.

We demonstrate the feasibility of this controller in a dynamic biped simulation. Here the biped always starts at the same initial position. The biped controller is derived from the learned IOC model by solving the ordinary differential equations obtained from the necessary conditions of optimality. Once we have these ODEs, there are multiple ways of obtaining a controller. First, note that the theory of differential flatness provides a convenient method of finding a controller. In particular, by solving the differential equations we obtain the flat outputs, which can then be mapped to appropriate control inputs given that the system is differentially flat. Alternatively, we can use an input/output linearization controller to implement the desired trajectories. Proportional-derivative-integral (PID) controllers can also be used to track the desired trajectory generated by the differential equations. We use the first method and obtain controllers using the theory of differential flatness.

Once we obtain a controller using the discussed approach, we modulate the free

height parameter $\beta_{15}(h)$ during the swing phase to generate locomotion on different stair heights. We make use of the impact functions as described in [109] to model collisions with the ground. The resulting simulations over multiple stair heights can be seen in Figure 5.9.

Due to its properties, this controller is a suitable one for transfemoral prostheses, in particular because the amputee will modulate his thigh movement according to the stair height to be climbed. The prosthetic knee controller will then "react" to the modulation of the thigh, without any programmed knowledge of the stair height. Furthermore, this is a feasible strategy to be implemented on the Vanderbilt prosthetic leg. This device is equipped with an inertial measurement unit (IMU) on the shank, which is capable of providing an estimate for the shank angle. The shank angle estimate together with the knee angle measurement can be used to estimate the thigh angle. Note that this is possible because these angles form a triangle. Therefore, it is feasible to have an estimate of both the thigh angle and the knee (shank) angle, which is all the information the controller needs.

5.5 Conclusion

We further demonstrated that our approach can learn a model of stair ascent using sparse data, and generate locomotion trajectories for novel stair heights. Thus, this model enables the automated tuning of lower-limb prosthetic controllers for different locomotion modes. The scalability and robustness of this approach, along with its prediction power even with simple dynamic models, opens the door for many exciting robotic applications.

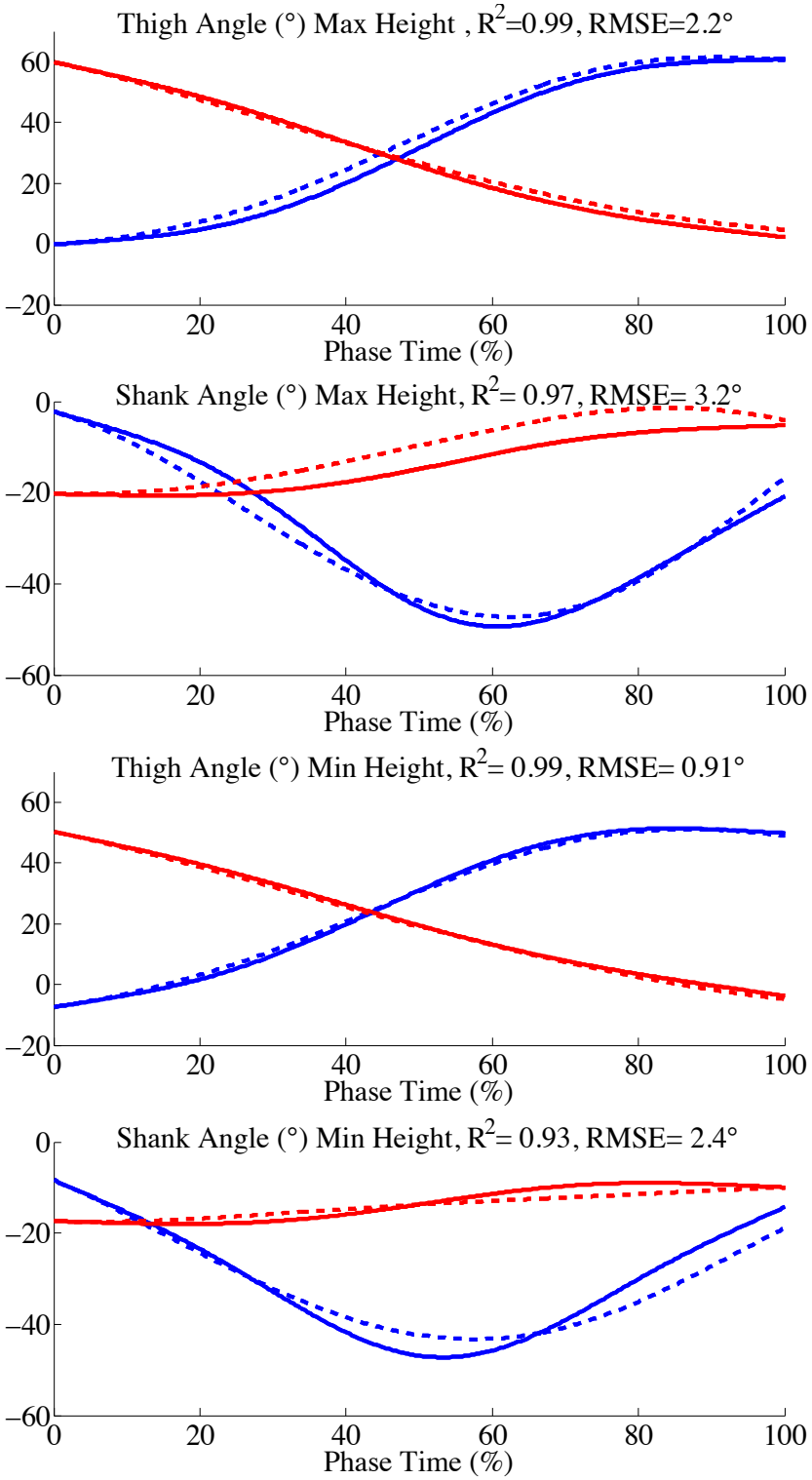


Figure 5.3: Proposed Solution - Prediction of stair ascent trajectories for a maximum and minimum step heights, which were used for IOC learning. Real (solid line) and predicted (dashed line) stair ascent trajectories. Swing is shown in blue and stance in red. Quality of predictions is noted by R^2 and $RMSE$ on each plot. Data for max and min stair heights were used in learning, and medium height was a novel prediction by IOC.

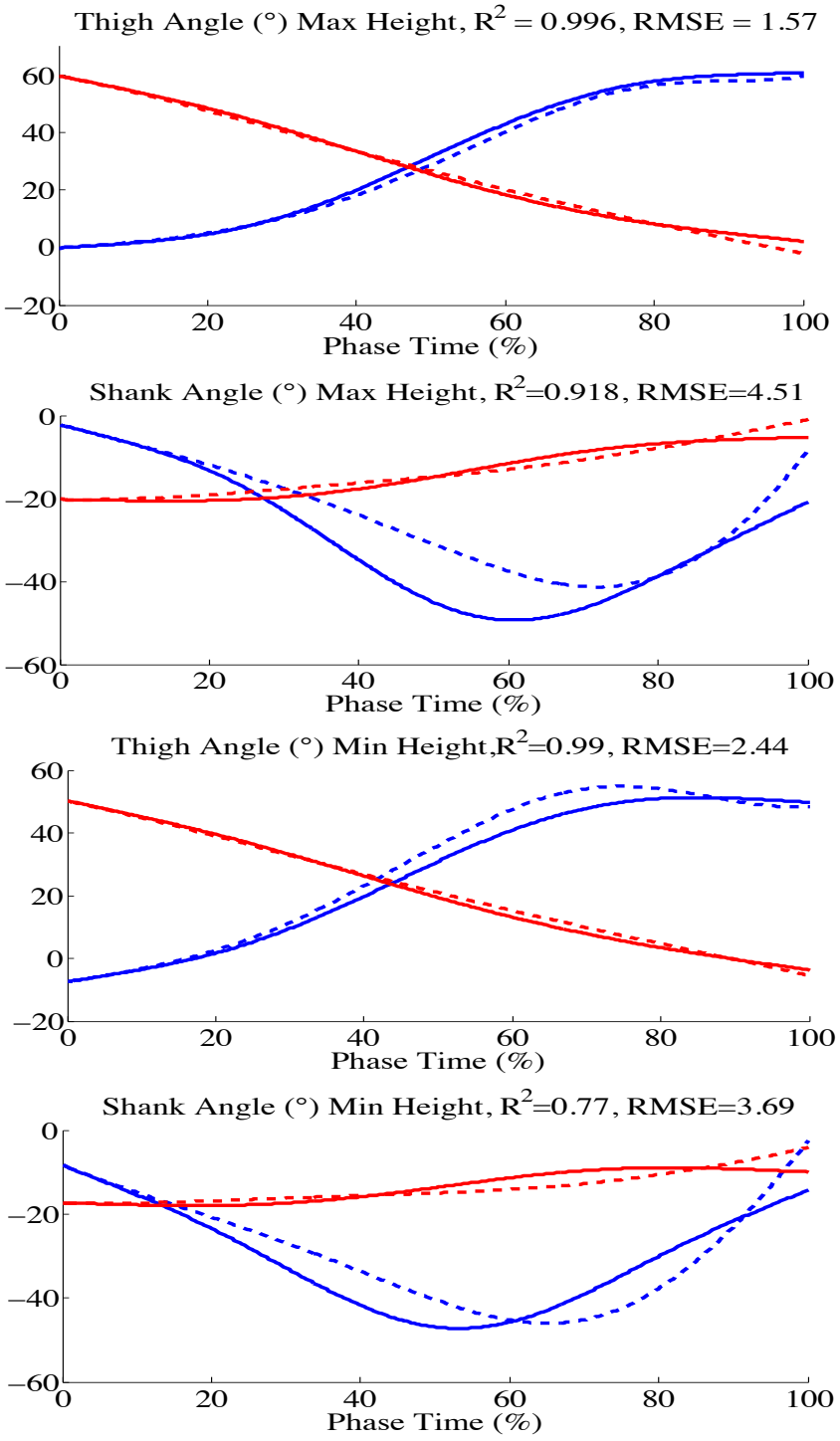


Figure 5.4: Existing Solution - Prediction of stair ascent trajectories for a maximum and minimum step heights, which were used for IOC learning. Real (solid line) and predicted (dashed line) stair ascent trajectories. Swing is shown in blue and stance in red. Quality of predictions is noted by R^2 and $RMSE$ on each plot. Data for max and min stair heights were used in learning, and medium height was a novel prediction by IOC.

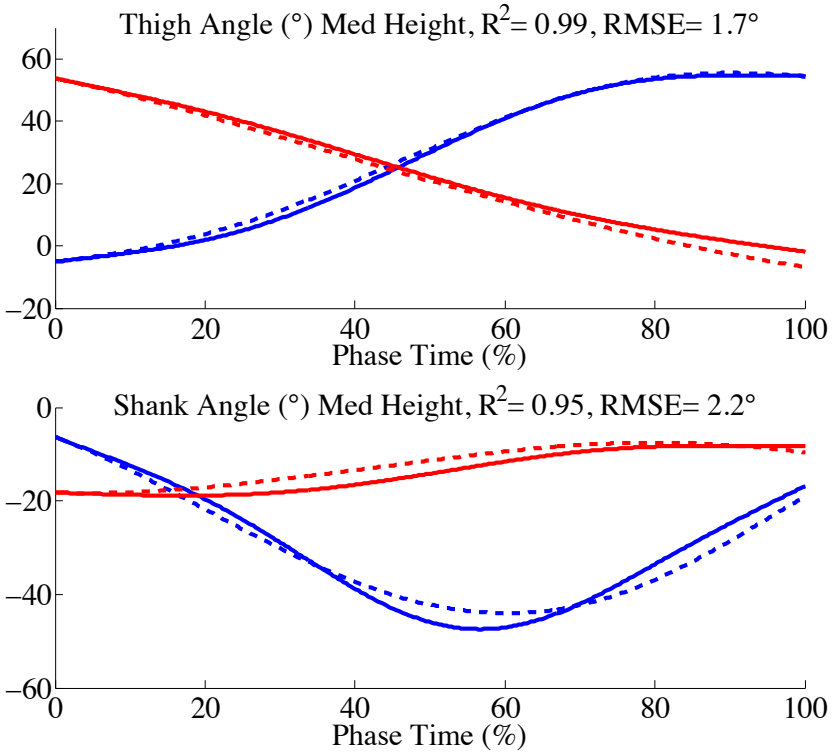


Figure 5.5: Proposed Solution - Prediction of novel stair ascent trajectories for a medium step height. Real (solid line) and predicted (dashed line) stair ascent trajectories. Swing is shown in blue and stance in red. Quality of predictions is noted by R^2 and $RMSE$ on each plot.

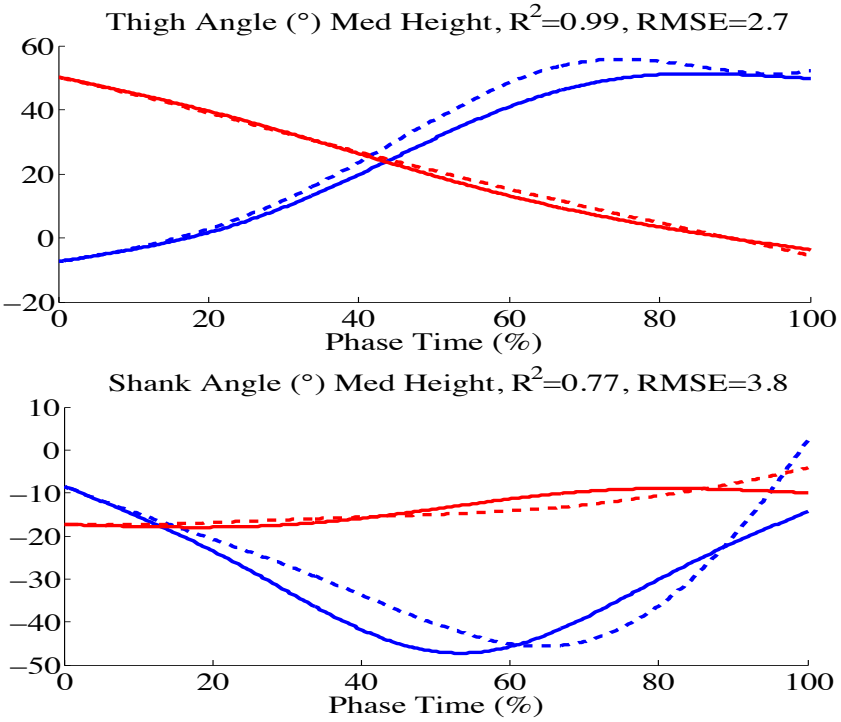


Figure 5.6: Existing Solution - Prediction of novel stair ascent trajectories for a medium step height. Real (solid line) and predicted (dashed line) stair ascent trajectories. Swing is shown in blue and stance in red. Quality of predictions is noted by R^2 and $RMSE$ on each plot.

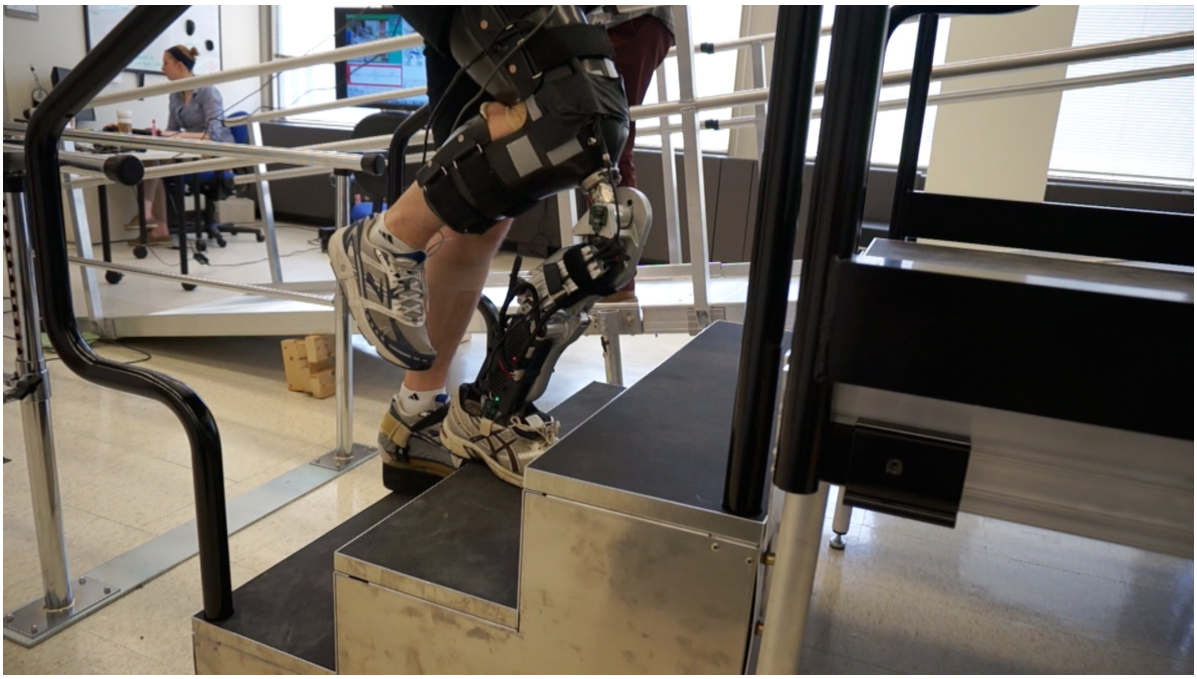


Figure 5.7: Photograph of one subject going up the stair case at the Center for Bionic Medicine. The length of the stairs were measured to be 19 cm and the proposed method of controller learning using inverse optimal control was used to control the prosthetic device seen in this picture.

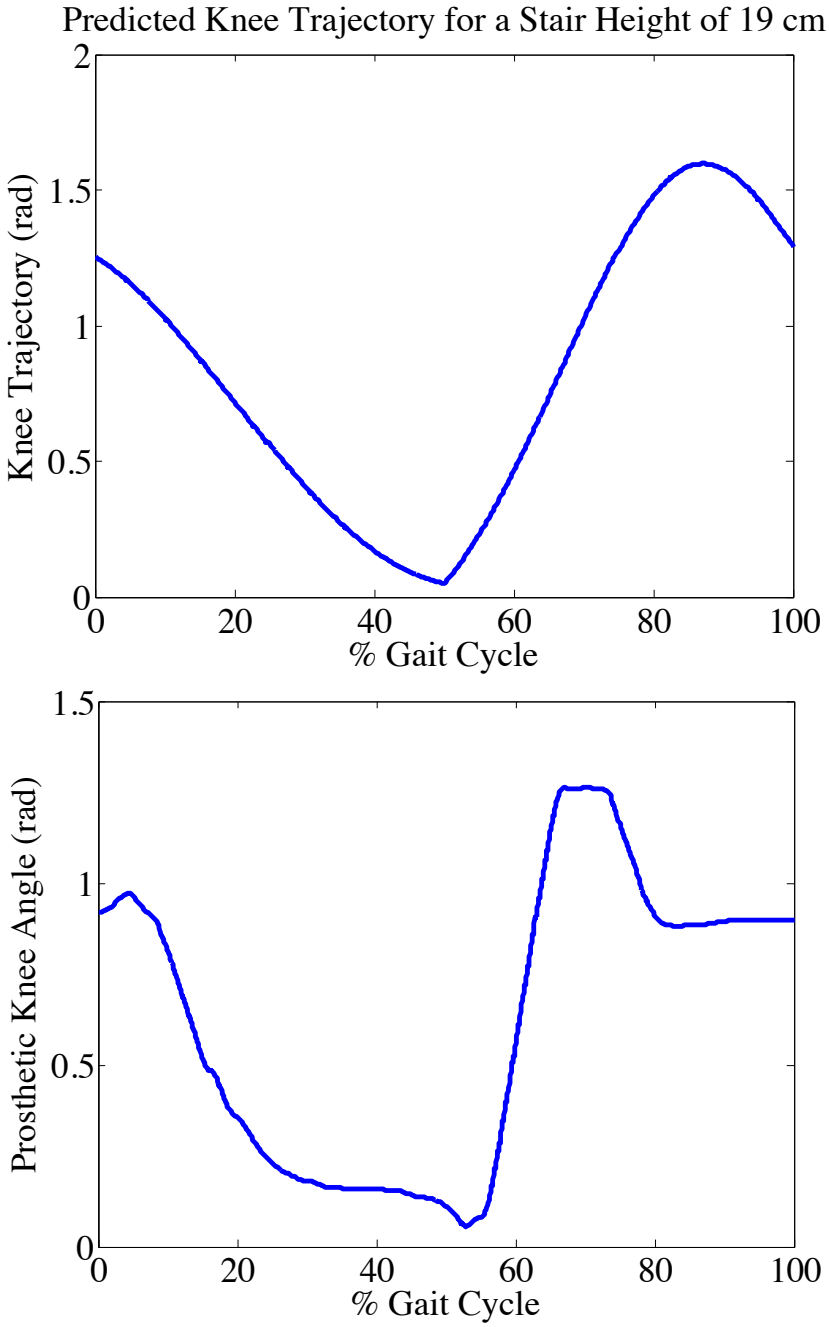


Figure 5.8: The knee trajectories corresponding to subject 1 climbing a stair of height 19 cm. The top plot shows the predicted trajectory using inverse optimal control and the bottom plot shows the resulting prosthetic knee trajectory using the learned controller parameters.

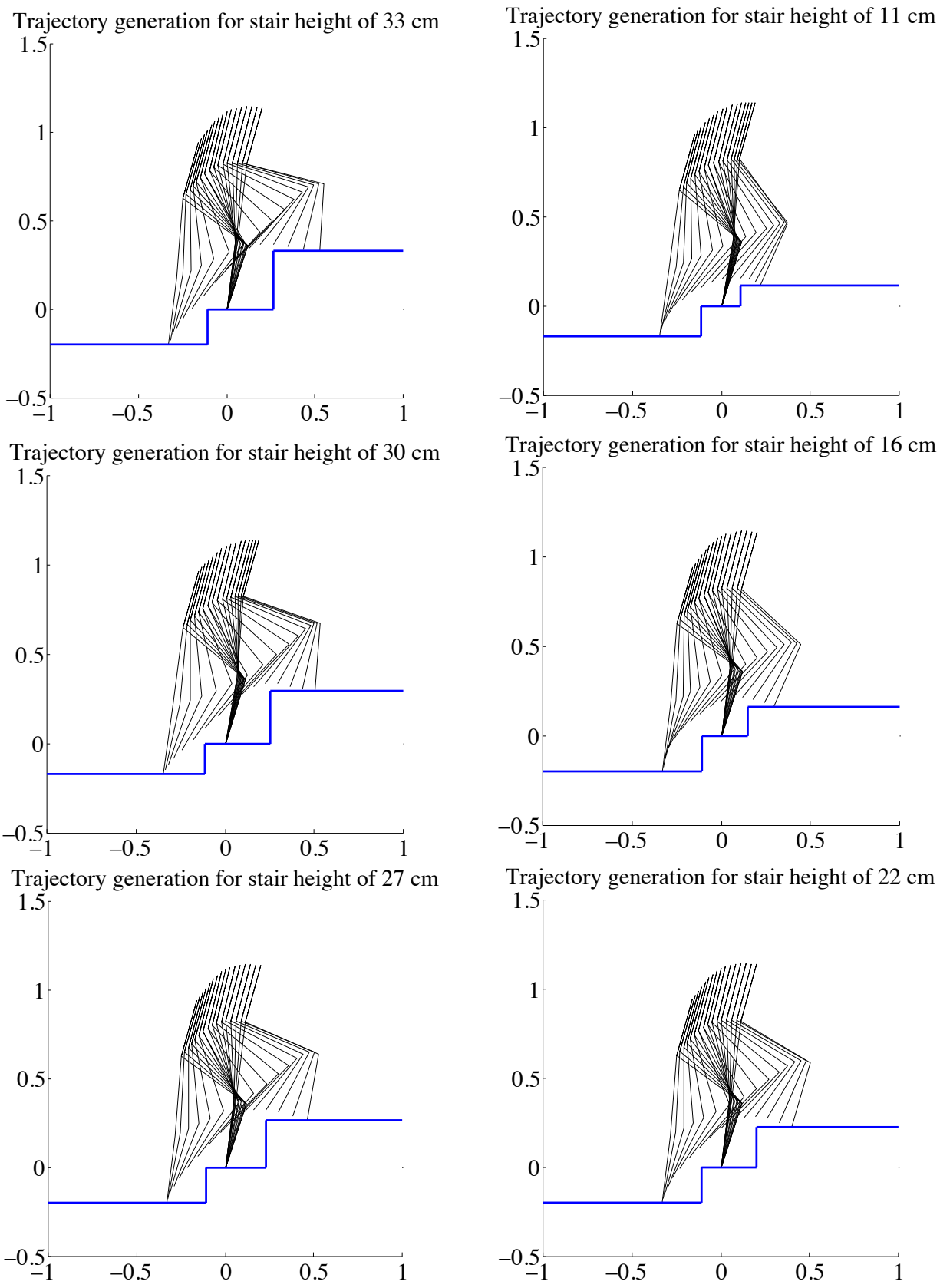


Figure 5.9: Dynamic simulation experiments with the learned height-invariant locomotion controller.

CHAPTER 6

DISCUSSION AND FUTURE WORK

6.1 Discussion of Prosthetic Controller Learning

In this thesis, we developed model-based controller learning approaches to automate the process of prosthetic controller tuning. We showed that these methods were capable of significantly reducing the tuning time, while learning prosthetic controllers with the same level of performance as the clinically-tuned controllers. To design these controller learning methods, we made different assumptions:

- We assumed that the five-link biped is a good model of the dynamic of the subject.
- We fixed the structure prosthetic controller to be an impedance controller.
- We assumed that we could find controller parameters by learning from exemplar trajectories corresponding to unimpaired individuals.

While these assumptions simplified the problem and produced clinically valid results in experiments, all of these assumptions can be modified to further improve prosthetic controller learning. In particular, we can consider more complex models, such as a seven-link biped or 3D bipedal models to better capture the dynamics of human walking.

Moreover, as suggested in the thesis, we can design novel control algorithms using model-based learning from human data. By learning invariant controllers across locomotion modes, e.g., different stair heights or different speeds, we can significantly reduce the number of parameters that need to be tuned for the prosthetic device. In designing new prosthetic controllers we can also utilize information such as the hip angle. While current impedance controllers do not use this information, the next generation of prosthetic controllers can be improved by modeling interactions with the hip joint.

Finally, instead of using exemplar trajectories from unimpaired individuals, observations of prosthetic locomotion can be used as exemplar trajectories. As we showed, there are certain features of the prosthetic locomotion, such as the stance maximum angle,

that are different than unimpaired locomotion even using a clinically tuned controller. By using prosthetic locomotion as observation, we can learn controller parameters that are better suited for prosthetic walking. Furthermore, by analyzing the invariant nature of prosthetic walking, we can obtain insights into why prosthetic walking does not exactly replicate unimpaired walking and potentially improve the control strategies.

6.2 Discussion of Inverse Optimal Control

In the next part of this thesis, we developed techniques for inverse optimal control. We demonstrated that by utilizing a new formulation of the problem and also properties of differentially flat systems, we can obtain computationally efficient solutions to the problem of IOC. Further, we can provide theoretical guarantees on the quality of the learned solution, thus making this approach very suitable for modeling locomotion and learning prosthetic controllers.

An important direction of future research is to automate the design of features for the cost functions. We demonstrated the importance of choosing the right features in modeling locomotion. In modeling locomotion, we used expert knowledge about locomotion from the literature to hand-design these features. Therefore, the ability to learn features is important for having a general purpose modeling tool.

Potentially promising approaches for learning features are artificial neural networks and deep learning approaches. These non-parametric methods can learn arbitrary high-level features, although they require large amounts of data. It would thus be extremely useful to study the incorporation of these non-parametric techniques into IOC methods and enable automatic feature learning. A connection to learning basis functions (features) using graphical models and neural networks can be made through the theory of optimal control for stochastic systems using path integrals. The works of [110, 111] discuss the connection between the optimal control problem and graphical model inference. Moreover, in our previous work [31] we have shown that for a class of stochastic dynamic systems, the problem of inverse optimal control using path integrals can be posed as:

$$\max_{\beta} \quad \frac{e^{\beta^T \phi(x^*,u^*)}}{\sum_{x,u} e^{\beta^T \phi(x,u)}}.$$

This form of a maximum entropy probability distribution is one that can be implemented using graphical models and neural networks. In particular, deep artificial neural networks can model this distribution and learn useful function representations of the

function $\phi(x, u)$, given enough data. This connection shows a promising direction for automatically learning features that can then be used for control of dynamic systems.

REFERENCES

[1] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, T. G. Travison, and R. Brookmeyer, "Estimating the prevalence of limb loss in the United States: 2005 to 2050," *Archives of physical medicine and rehabilitation*, vol. 89, no. 3, pp. 422–429, 2008.

[2] J. L. Johansson, D. M. Sherrill, P. O. Riley, P. Bonato, and H. Herr, "A clinical comparison of variable-damping and mechanically passive prosthetic knee devices," *American journal of physical medicine & rehabilitation*, vol. 84, no. 8, pp. 563–575, 2005.

[3] R. Waters, J. Perry, D. Antonelli, and H. Hislop, "Energy cost of walking of amputees: the influence of level of amputation," *J Bone Joint Surg Am*, vol. 58, no. 1, pp. 42–46, 1976.

[4] F. Sup, A. Bohara, and M. Goldfarb, "Design and control of a powered transfemoral prosthesis," *The international journal of robotics research*, vol. 27, no. 2, pp. 263–273, 2008.

[5] M. R. Tucker, J. Olivier, A. Pagel, H. Bleuler, M. Bouri, O. Lambercy, J. d. R. Millan, R. Riener, H. Vallery, and R. Gassert, "Control strategies for active lower extremity prosthetics and orthotics: a review," *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 1, 2015.

[6] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[7] D. Braun and M. Goldfarb, "A control approach for actuated dynamic walking in biped robots," *Robotics, IEEE Transactions on*, vol. 25, no. 6, pp. 1292–1303, 2009.

[8] F. Sup, H. A. Varol, and M. Goldfarb, "Upslope walking with a powered knee and ankle prosthesis: initial results with an amputee subject," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 19, no. 1, pp. 71–78, 2011.

[9] A. M. Simon, N. P. Fey, S. B. Finucane, R. D. Lipschutz, and L. J. Hargrove, "Strategies to reduce the configuration time for a powered knee and ankle prosthesis across multiple ambulation modes," in *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.

[10] C. G. Atkeson and S. Schaal, "Robot learning from demonstration," in *Machine Learning-International Workshop then Conference*. Morgan Kaufmann Publishers, Inc., 1997, pp. 12–20.

[11] A. Ijspeert, J. Nakanishi, and S. Schaal, "Learning attractor landscapes for learning motor primitives," *Advances in neural information processing systems*, pp. 1547–1554, 2003.

[12] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," *Robotics Research*, pp. 561–572, 2005.

[13] P. Abbeel, A. Coates, and A. Ng, "Autonomous helicopter aerobatics through apprenticeship learning," *International Journal of Robotics Research*, vol. 29, no. 13, pp. 1608–1639, 2010.

[14] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—an inverse optimal control approach," *Autonomous Robots*, vol. 28, no. 3, pp. 369–383, 04 2010.

[15] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *ICML '04: Proceedings of the twenty-first international conference on Machine learning*. New York, NY, USA: ACM, 2004, p. 1.

[16] N. Ratliff, J. Bagnell, and M. Zinkevich, "Maximum margin planning," in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 729–736.

[17] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *Intelligent Control (ISIC), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 613–619.

[18] A.-S. Puydupin-Jamin, M. Johnson, and T. Bretl, "A convex approach to inverse optimal control and its application to modeling human locomotion," in *IEEE International Conference on Robotics and Automation*, 2012.

[19] A. Terekhov, Y. Pesin, X. Niu, M. Latash, and V. Zatsiorsky, "An analytical approach to the problem of inverse optimization with additive objective functions: an application to human prehension," *Journal of Mathematical Biology*, vol. 61, pp. 423–453, 2010.

[20] A. Terekhov and V. Zatsiorsky, "Analytical and numerical analysis of inverse optimization problems: conditions of uniqueness and computational methods," *Biological Cybernetics*, vol. 104, pp. 75–93, 2011.

[21] J. Park, V. M. Zatsiorsky, and M. L. Latash, "Finger coordination under artificial changes in finger strength feedback: A study using analytical inverse optimization," *Journal of Motor Behavior*, vol. 43, no. 3, pp. 229–235, 2011.

[22] T. D. Nielsen and F. V. Jensen, "Learning a decision maker's utility function from (possibly) inconsistent behavior," *Artificial Intelligence*, vol. 160, pp. 53–78, 2004.

[23] A. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in *Proceedings of the Seventeenth International Conference on Machine Learning*, 2000, pp. 663–670.

[24] P. Abbeel, "Apprenticeship learning and reinforcement learning with application to robotic control," Ph.D. dissertation, Stanford University, 2008.

[25] J. Tang, A. Singh, N. Goehausen, and P. Abbeel, "Parameterized maneuver learning for autonomous helicopter flight," in *IEEE International Conference on Robotics and Automation*, May 2010, pp. 1142 –1148.

[26] U. Syed, M. Bowling, and R. E. Schapire, "Apprenticeship learning using linear programming," in *Proceedings of the 25th international conference on machine learning*, ser. ICML '08. New York, NY, USA: ACM, 2008, pp. 1032–1039.

[27] D. Ramachandran and E. Amir, "Bayesian inverse reinforcement learning," *IJCAI-07*, vol. 51, p. 61801.

[28] B. Ziebart, A. Maas, J. Bagnell, and A. Dey, "Maximum entropy inverse reinforcement learning," in *Proc. AAAI*, 2008, pp. 1433–1438.

[29] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey, "Human behavior modeling with maximum entropy inverse optimal control," *AAAI Spring Symposium on Human Behavior Modeling*, 2009.

[30] K. Dvijotham and E. Todorov, "Inverse Optimal Control with Linearly-Solvable MDPs," in *Proceedings of the Interntional Conference on Machine Learning*. Citeseer, 2010.

[31] N. Aghasadeghi and T. Bretl, "Maximum entropy inverse reinforcement learning in continuous state spaces with path integrals," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 1561–1566.

[32] S. Javdani, S. Tandon, J. Tang, J. O'Brien, and P. Abbeel, "Modeling and perception of deformable one-dimensional objects," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2011, pp. 1607 –1614.

[33] N. A. Borghese, L. Bianchi, and F. Lacquaniti, "Kinematic determinants of human locomotion." *The Journal of physiology*, vol. 494, no. Pt 3, pp. 863–879, 1996.

[34] P. Abbeel and A. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proceedings of the twenty-first international conference on machine learning*. ACM, 2004, p. 1.

[35] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, "Flatness and defect of non-linear systems: introductory theory and examples," *International journal of control*, vol. 61, no. 6, pp. 1327–1361, 1995.

[36] R. M. Murray, M. Rathinam, and W. Sluis, "Differential flatness of mechanical control systems: A catalog of prototype systems," in *ASME International Mechanical Engineering Congress and Exposition*. Citeseer, 1995.

[37] M. Daley, G. Felix, and A. Biewener, "Running stability is enhanced by a proximo-distal gradient in joint neuromechanical control," *Journal of Experimental Biology*, vol. 210, no. 3, pp. 383–394, 2007.

[38] N. Aghasadeghi, H. Zhao, L. J. Hargrove, A. D. Ames, E. J. Perreault, and T. Bretl, "Learning impedance controller parameters for lower-limb prostheses," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 4268–4274.

[39] N. Aghasadeghi, H. Zhao, L. J. Hargrove, A. D. Ames, E. J. Perreault, and T. Bretl, "Automating impedance controller tuning for lower-limb prostheses," submitted to *IEEE Transactions on Biomedical Engineering*, 2015.

[40] N. Aghasadeghi, A. Long, and T. Bretl, "Inverse optimal control for a hybrid dynamical system with impacts," in *Robotics and Automation (ICRA), 2012 IEEE/RSJ International Conference on*. IEEE, 2012.

[41] M. Johnson, N. Aghasadeghi, and T. Brelt, "Inverse optimal control for deterministic continuos-time non-linear systems," in *CDC'13*, 2013.

[42] N. Aghasadeghi and T. Bretl, "Inverse optimal control for differentially flat systems with application to locomotion modeling," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 6018–6025.

[43] N. Aghasadeghi, L. Hargove, E. Perreault, and T. Bretl, "Inverse optimal control for differentially flat systems with application to prosthetic controller learning for stair ascent," in preparation for *International Journal of Robotics Research*, 2015.

[44] N. Hogan, "Impedance control: An approach to manipulation," in *American Control Conference, 1984*. IEEE, 1984, pp. 304–313.

[45] K. H. Ha, H. A. Varol, and M. Goldfarb, "Volitional control of a prosthetic knee using surface electromyography," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 1, pp. 144–151, 2011.

[46] A. M. Simon, K. A. Ingraham, N. P. Fey, S. B. Finucane, R. D. Lipschutz, A. J. Young, and L. J. Hargrove, "Configuring a powered knee and ankle prosthesis for transfemoral amputees within five specific ambulation modes," *PloS one*, vol. 9, no. 6, p. e99387, 2014.

[47] R. Kearney, I. Hunter et al., "System identification of human joint dynamics." *Critical reviews in biomedical engineering*, vol. 18, no. 1, p. 55, 1990.

[48] E. Perreault, L. Hargrove, D. Ludvig, H. Lee, and J. Sensinger, "Considering limb impedance in the design and control of prosthetic devices," in *Neuro-Robotics*. Springer, 2014, pp. 59–83.

[49] E. Rouse, L. Hargrove, E. Perreault, and T. Kuiken, "Estimation of human ankle impedance during the stance phase of walking," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 22, no. 4, pp. 870–878, July 2014.

[50] P. Weiss, R. Kearney, and I. Hunter, "Position dependence of ankle joint dynamicsi. passive mechanics," *Journal of biomechanics*, vol. 19, no. 9, pp. 727–735, 1986.

[51] J. MacNeil, R. Kearney, and I. Hunter, "Identification of time-varying biological systems from ensemble data (joint dynamics application)," *Biomedical Engineering, IEEE Transactions on*, vol. 39, no. 12, pp. 1213–1225, 1992.

[52] D. Ludvig and E. J. Perreault, "System identification of physiological systems using short data segments," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 12, pp. 3541–3549, 2012.

[53] H. Lee and N. Hogan, "Time-varying ankle mechanical impedance during human locomotion," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 2014.

[54] M. R. Tucker, A. Moser, O. Lambercy, J. Sulzer, and R. Gassert, "Design of a wearable perturbator for human knee impedance estimation during gait," in *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.

[55] B. Lawson, H. A. Varol, A. Huff, E. Erdemir, and M. Goldfarb, "Control of stair ascent and descent with a powered transfemoral prosthesis," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 21, no. 3, pp. 466–473, 2013.

[56] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, 2nd ed. Wiley-Interscience, May 1990.

[57] S. H. Scott and D. A. Winter, "Biomechanical model of the human foot: Kinematics and kinetics during the stance phase of walking," *J. of Biomechanics*, vol. 26, no. 9, pp. 1091–1104, Sep. 1993.

[58] E. Westervelt, J. Grizzle, and D. E. Koditschek, "Hybrid zero dynamics of planar biped walkers," *Automatic Control, IEEE Transactions on*, vol. 48, no. 1, pp. 42–56, 2003.

[59] A. Ames, "First steps toward automatically generating bipedal robotic walking from human data," *Robot Motion and Control 2011*, pp. 89–116, 2012.

[60] S. S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*. Springer, June 1999.

[61] E. Baake, M. Baake, H. Bock, K. Briggs et al., "Fitting ordinary differential equations to chaotic data," *Physical Review A*, vol. 45, no. 8, pp. 5524–5529, 1992.

[62] K. Schittkowski, "Parameter estimation in systems of nonlinear equations," *Numerische Mathematik*, vol. 68, no. 1, pp. 129–142, 1994.

[63] M. Peifer and J. Timmer, "Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting," *Systems Biology, IET*, vol. 1, no. 2, pp. 78–88, 2007.

[64] Z. Li, M. Osborne, and T. Prvan, "Parameter estimation of ordinary differential equations," *IMA Journal of Numerical Analysis*, vol. 25, no. 2, pp. 264–285, 2005.

[65] J. Ramsay, G. Hooker, D. Campbell, and J. Cao, "Parameter estimation for differential equations: a generalized smoothing approach," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 5, pp. 741–796, 2007.

[66] N. Brunel, "Parameter estimation of odes via nonparametric estimators," *Electronic Journal of Statistics*, vol. 2, pp. 1242–1267, 2008.

[67] J. Varah, "A spline least squares method for numerical parameter estimation in differential equations," *SIAM Journal on Scientific and Statistical Computing*, vol. 3, p. 28, 1982.

[68] A. Poyton, M. Varziri, K. McAuley, P. McLellan, and J. Ramsay, "Parameter estimation in continuous-time dynamic models using principal differential analysis," *Computers & chemical engineering*, vol. 30, no. 4, pp. 698–708, 2006.

[69] S. M. Jaegers, J. H. Arendzen, and H. J. de Jongh, "Prosthetic gait of unilateral transfemoral amputees: a kinematic study," *Archives of physical medicine and rehabilitation*, vol. 76, no. 8, pp. 736–743, 1995.

[70] M. F. Eilenberg, H. Geyer, and H. Herr, "Control of a powered ankle–foot prosthesis based on a neuromuscular model," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 18, no. 2, pp. 164–173, 2010.

[71] S. Pfeifer, H. Vallery, M. Hardegger, R. Riener, and E. J. Perreault, "Model-based estimation of knee stiffness," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 9, pp. 2604–2612, 2012.

[72] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469 – 483, 2009.

[73] K. Hatz, J. P. Schloder, and H. G. Bock, "Estimating parameters in optimal control problems," *SIAM Journal on Scientific Computing*, vol. 34, no. 3, pp. A1707–A1728, 2012.

[74] M. Knauer and C. Büskens, "Bilevel optimization of container cranes," in *Progress in Industrial Mathematics at ECMI 2008*, ser. Mathematics in Industry, A. D. Fitt, J. Norbury, H. Ockendon, and E. Wilson, Eds. Springer Berlin Heidelberg, 2010, pp. 913–918.

[75] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *IEEE International Conference on Robotics and Automation*, May 2009, pp. 763 –768.

[76] G. Konidaris, S. Kuindersma, A. Barto, and R. Grupen, "Constructing skill trees for reinforcement learning agents from demonstration trajectories," in *Advances In Neural Information Processing Systems*, 2010.

[77] J. L. Yepes, I. Hwang, and M. Rotea, "New algorithms for aircraft intent inference and trajectory prediction," *Journal of Guidance Control and Dynamics*, vol. 30, pp. 370–382, 2007.

[78] D. Lee and Y. Nakamura, "Mimesis model from partial observations for a humanoid robot," *The International Journal of Robotics Research*, vol. 29, no. 1, pp. 60–80, 2010.

[79] D. H. Grollman and O. C. Jenkins, "Incremental learning of subtasks from unsegmented demonstration," in *International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, Oct. 2010.

[80] W. Li, E. Todorov, and D. Liu, "Inverse optimality design for biological movement systems," in *IFAC*, 2011.

[81] H. K. Khalil, *Nonlinear Systems*. Prentice-Hall, Inc., 2002.

[82] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2011.

[83] A. V. Rao, D. A. Benson, C. Darby, M. A. Patterson, C. Francolin, I. Sanders, and G. T. Huntington, "Algorithm 902: Gpops, a matlab software for solving multiple-phase optimal control problems using the gauss pseudospectral method," *ACM Trans. Math. Softw.*, vol. 37, no. 2, pp. 22:1–22:39, Apr. 2010.

[84] M. Athans and P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*. McGraw-Hill, 1966.

[85] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*. Hemisphere Publishing Co., 1975.

[86] T. Bretl and Z. McCarthy, "Equilibrium configurations of a kirchhoff elastic rod under quasi-static manipulation," in *WAFR*, 2012.

[87] J. Biggs, W. Holderbaum, and V. Jurdjevic, "Singularities of optimal control problems on some 6-d lie groups," *Automatic Control, IEEE Transactions on*, vol. 52, no. 6, pp. 1027–1038, 2007.

[88] N. Ratliff, J. A. D. Bagnell, and M. Zinkevich, "Maximum margin planning," in *International Conference on Machine Learning*, July 2006.

[89] A. Boularias, J. Kober, and J. Peters, "Relative entropy inverse reinforcement learning," in *International Conference on Artificial Intelligence and Statistics*, 2011, pp. 182–189.

[90] R. M. Murray, "Nonlinear control of mechanical systems: A Lagrangian perspective," *Annual Reviews in Control*, vol. 21, pp. 31–42, 1997.

[91] R. Mahadevan, S. K. Agrawal, and F. J. Doyle, "Differential flatness based nonlinear predictive control of fed-batch bioreactors," *Control Engineering Practice*, vol. 9, no. 8, pp. 889–899, 2001.

[92] J. Oldenburg and W. Marquardt, "Flatness and higher order differential model representations in dynamic optimization," *Computers & chemical engineering*, vol. 26, no. 3, pp. 385–400, 2002.

[93] S. Levine and V. Koltun, "Continuous inverse optimal control with locally optimal examples," arXiv preprint arXiv:1206.4617, 2012.

[94] P. A. Ioannou and J. Sun, *Robust adaptive control*. Courier Dover Publications, 2012.

[95] L. Ljung, *System identification*. Springer, 1998.

[96] L. Ljung and T. Glad, "On global identifiability for arbitrary model parametrizations," *Automatica*, vol. 30, no. 2, pp. 265–276, 1994.

[97] J. M. Ball, J. Currie, and P. J. Olver, "Null Lagrangians, weak continuity, and variational problems of arbitrary order," *Journal of Functional Analysis*, vol. 41, no. 2, pp. 135–174, 1981.

[98] A. Berlinet and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics*. Kluwer Academic Boston, 2004.

[99] S. Zhou and D. Wolfe, "On derivative estimation in spline regression," *Statistica Sinica*, vol. 10, no. 1, pp. 93–108, 2000.

[100] A. V. Oppenheim, R. W. Schafer, J. R. Buck et al., *Discrete-time signal processing*. Prentice Hall Upper Saddle River, 1999, vol. 5.

[101] S. Schaal, C. G. Atkeson, and S. Vijayakumar, "Scalable techniques from nonparametric statistics for real time robot learning," *Applied Intelligence*, vol. 17, no. 1, pp. 49–60, 2002.

[102] P. Martin, R. M. Murray, P. Rouchon et al., "Flat systems," lecture at Summer School on Mathematical Control Theory, Trieste, Italy, 2001.

[103] D. A. Winter, *Biomechanics and motor control of human gait: normal, elderly and pathological*. Transportation Research Laboratory, Wokingham, Berkshire, U.K., 1991.

[104] F. Lacquaniti, R. Grasso, and M. Zago, "Motor patterns in walking," *Physiology*, vol. 14, no. 4, pp. 168–174, 1999.

[105] L. Bianchi, D. Angelini, G. Orani, and F. Lacquaniti, "Kinematic coordination in human gait: relation to mechanical energy cost," *Journal of neurophysiology*, vol. 79, no. 4, pp. 2155–2170, 1998.

[106] Y. P. Ivanenko, A. d'Avella, R. E. Poppele, and F. Lacquaniti, "On the origin of planar covariation of elevation angles during human locomotion," *Journal of neurophysiology*, vol. 99, no. 4, pp. 1890–1898, 2008.

[107] M. MacKay-Lyons, "Central pattern generation of locomotion: a review of the evidence," *Physical Therapy*, vol. 82, no. 1, pp. 69–83, 2002.

[108] R. Riener, M. Rabuffetti, and C. Frigo, "Stair ascent and descent at different inclinations," *Gait & posture*, vol. 15, no. 1, pp. 32–44, 2002.

[109] Y. Hurmuzlu and D. B. Marghitu, "Rigid body collisions of planar kinematic chains with multiple contact points," *The international journal of robotics research*, vol. 13, no. 1, pp. 82–92, 1994.

[110] H. J. Kappen, V. Gómez, and M. Opper, "Optimal control as a graphical model inference problem," *Machine learning*, vol. 87, no. 2, pp. 159–182, 2012.

[111] B. van den Broek, W. Wiegerinck, and B. Kappen, "Graphical model inference in optimal control of stochastic multi-agent systems," *Journal of Artificial Intelligence Research*, vol. 32, no. 1, pp. 95–122, 2008.