User-Informed Exoskeleton Design Metrics and Multi-Modal Exoskeleton Control Strategies

by

Roberto Leo Medrano

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in the University of Michigan
2023

Doctoral Committee:

Associate Professor Elliott J. Rouse, Chair
Associate Professor Robert D. Gregg
Associate Professor Leia Stirling
Associate Professor Ram Vasudevan

Roberto Leo Medrano
leomed@umich.edu
ORCID iD: 0000-0002-8342-0494

DEDICATION

This dissertation is dedicated to my parents. In every conceivable way, I would not be who I am without you.

ACKNOWLEDGMENTS

I'm not one to mince words, so without further ado, the following are some of the (many) accomplished, talented, and simply awesome people without whom I would not be here today.

I'll start by thanking those in academia who have supported me on this journey. First, Robert Gregg, who I think of as my second advisor. You let me jump straight into some awesome technical work, and I'll forever be grateful for that opportunity afforded by the Locolab's close partnership with Neurobionics. Of course, I must also thank Gray Thomas (who is on literally every single one of my papers), for being an infinite source of technical wisdom, and for putting up with my near endless conversations and prompts (though I hope they were as equally exciting for you as they were for me). Turning to lab members past, I must thank Kim Ingraham, for providing both mentorship early in my graduate career when I needed it most and elder jaded-grad-student wisdom as a counterforce to my boundless 20-year-old optimism. I also thank Emma Reznick, for her invaluable Vicon experience and her dry wit, and Hannah Frame, for keeping the lab running smoothly during her tenure as research engineer, and for her endlessly amusing snarky comments. Thanks also to Ung Hee Lee for his insight into machine learning, and for being an all-around great team-mate. I also give thanks to Luke Mooney and the awesome team at Dephy, for providing the ankle exoskeletons that have enabled the bulk of my research (and for tolerating our incessant bug reports). I want to give a shout-out to Drew Margolin over at Cornell, and the serendipity of connecting flights to LAX. Finally, and almost certainly foremost: Elliott Rouse, the boss, my advisor. I chose to attend UM because I was inspired by your belief that engineering must actually help real people out there, and because of your leadership, I have always known that my choice was right. You've witnessed my growth these last five years and have provided such integral support and guidance in professional development, in encouraging me to be a better thinker and storyteller, and in granting me the latitude I needed to become a full-fledged independent PhD. I have had no greater mentor than you, and for that I will be eternally thankful.

To present friends, beginning with Yves Nazon, also a longstanding feature of the lab. I'll never forget the warm welcome you extended to me early on, and the great support you gave me during some darker moments is something I'll forever be grateful for. Same goes to Jose and Karinne, for your support throughout my journey, and for use of your car. To Nikko Van Crey, the self-proclaimed crazy uncle at Neurobionics, you were one of my first friends in Ann Arbor,

and your chaotic antics have never failed to brighten my days. Throughout these last five years, I have rested easy knowing I've had the support and love of some other bedrock friends, and to them, I give thanks. To Dan Reich and Shumrith Vinakollu: it's been a lot harder than our days in Perry 2 to find time for each other, and I'll never stop being grateful that we've managed to keep it going through our group calls and ivory tower political debates. To Neil Acharya: those Discord calls with you were a source of light throughout the pandemic, and I'll always remember that. To Richard Huckaby: your story and spirit has never ceased to inspire me to persevere, and know that you'll always have a listening ear and a life-long friend in me. Also, a big thank you to the people and communities at the gym, yoga, and pole, for keeping me sane throughout the last four years. Finally, the final friend shout-out goes to Youssef Tobah, perhaps my oldest friend: I've cherished each of our many years of friendship, and it never ceases to amaze me that, over a decade and at least three countries, we've kept it going strong.

A thank you to someone on the other side: Congressman John Lewis. You'll never know this, but that handshake with a young Georgia Tech freshman at his first protest at Hartsfield-Jackson Atlanta International airport will never be forgotten. I hope that no matter where my career takes me, I'll keep doing my part to help create that beloved community.

Finally, the greatest thank you to my family. To Franco, my bro(ther). Our bond is strong, and I want you to know your older sibling is proud of who you are and grateful for your love (and those great L4D2 sessions). To my extended family in Peru, both gone and still with us, thank you for your the many happy memories from my early life. And to my parents, Ursula and Roberto Medrano. I'm not sure anything I could write down right now could ever capture how grateful I am to you both. You've raised me to value service, hard work, kindness, and to understand that the right thing is worth doing simply because it is right. You've both sacrificed so much to help me get to where I am today and I'll never forget that. You came to this great country driven by that canonical immigrant hope that your children would see a brighter tomorrow, and, though I imagine you are both proud to see what I've accomplished, I am prouder still to say I am your son. Thank you.

PREFACE

Chapters 2-5 of this dissertation were each written as separate manuscripts. Given this, there may be some repetition between them, particularly between Chapters 4 and 5.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

LIST OF ACRONYMS

JND Just Noticeable Difference

TBE Timing Based Estimation

WF Weber Fraction

WAIC Watanabe Akaike Information Criterion

SD Standard Deviation

SEM Standard Error of the Mean

EKF Extended Kalman Filter

UKF Unscented Kalman Filter

EnKF Ensemble Kalman Filter

ABSTRACT

Powered exoskeletons have long captivated both the public consciousness and scientific interest with their potential to transform the mobility of the public. In particular, augmentative lower-limb exoskeletons have had notable success in making walking more efficient; however, this has not yet resulted in these technologies achieving widespread adoption. One primary obstacle to this goal is that the ideal metrics by which augmentative exoskeletons should be designed and evaluated to enable dissemination are not yet clear. Without a better understanding of how to develop these devices in such a way that they are voluntarily adopted, exoskeletons will remain confined to the laboratory. Another major challenge is that, while existing exoskeleton controllers are largely designed for a controlled laboratory setting, widespread adoption of these technologies will require controllers that can adapt their assistance in response to the transitory, ever-changing terrains of the real world.

This dissertation aims to provide solutions to these two major challenges. The works contained in this dissertation can be categorized into 1) investigations into the metrics by which augmentative exoskeletons are developed and judged, and 2) the development of a controller that continuously learns the wearer's gait in real-time and applies appropriate assistance. These two categories are comprised of four primary projects.

In the first project, I calculated the Just Noticeable Difference (JND) of changes in metabolic rate, which is the "gold standard" physiological metric by which augmentative exoskeletons are designed and evaluated [117]. This study revealed that the average JND was nearly twice as high as the average energetic benefit from modern devices, which calls into question the primacy of a metric that yields imperceptibly beneficial devices. In the second project, I characterized the economic value of modern bilateral ankle exoskeletons. The goal of this work was to investigate economic value as a potential alternative metric that more effectively reflects user willingness to use an exoskeleton. Designing exoskeletons according to this metric may then result in devices that are likelier to translate outside the laboratory by virtue of providing economic value to potential users. Beyond this result, this work also revealed that the average net benefit, in dollars, of the use of the bilateral exoskeletons tested in this study is near-zero. In the third project, I investigated the fundamental problem of estimating the user's gait state—the parametrization of a person's locomotion—continuously throughout a variety of walking tasks using the Bayesian framework

[118]. This project revealed that an Extended Kalman Filter (EKF) estimator was capable of tracking the gait state as it evolved throughout different speeds and inclines. Finally, in the fourth project, I built upon this work to develop a practical exoskeleton controller that learns the gait state in real-time and adapts biomimetic torque assistance accordingly [116]. This controller was validated both on controlled treadmill trials to establish a baseline, and on a series of "stress-tests" in which a participant walked in extreme outdoor terrain settings. Taken together, these projects provide a foundation for translating exoskeletons out of the laboratory and into the real world.

CHAPTER 1

Introduction

1.1 Research Motivation

Powered exoskeletons have long captured the interest of the public through their promise to let us exceed the limits of our human bodies. This fascination is reflected in the recent explosion in the number of scientific investigations into these technologies [150] as our world's top minds seek to create devices that will let us run faster, walk longer, and carry heavier. The augmentation from these exoskeletons has the potential to dramatically alter human society. For example, in the United States, nearly 28 million Americans work in fields that require heavy or very heavy labor [23, 22], such as supply chain workers, first responders, and builders. It is easy to imagine a future in which a worker, rather than using solely their own strength in their daily job, instead dons a powered exoskeleton that allows them to not only bear the physical demands of their career more easily, but also enables them to accomplish loading tasks that would otherwise be impossible without the aid of the device. What took five unaided workers, one worker wearing an exoskeleton could potentially do alone, drastically decreasing task completion times and allowing society to reap the benefits. The goal of superseding the limitations of our bodies motivates many research groups to overcome the various barriers that prevent exoskeletons from reaching widespread adoption.

In contrast to the classic image of exoskeletons as large, bulky devices, today's exoskeletons are designed to be lightweight, to avoid encumbering the wearer [121, 149, 133], powerful, to maximize the powered assistance delivered [21, 206], and untethered, to allow wearers to use these technologies outside the laboratory [133, 132, 92]. Commonly, these exoskeletons are designed to aid in locomotion, and do so by applying mechanical assistance to the joints of the leg during walking, thus reducing the demands on the wearer's neuromuscular system. Notable successes in augmentative exoskeleton development include an autonomous powered ankle exoskeleton [121], which uses an electric motor and belt transmission to apply assistive torques at the ankle, a soft exosuit that uses a Bowden cable transmission to apply torques at the knee and hip joints [133], and a modular multi-joint exoskeleton capable of assisting all three lower-limb joints depending

on the application [126]. Modern augmentative exoskeletons are commonly deemed successful if they reduce the wearer's metabolic cost (e.g. the amount of calories burned) during locomotion, as intuitively, upstream reduction in the energy required to perform a task should result in increased capabilities or performance from the wearer. Due to the focus on locomotion, the controllers for today's exoskeletons are developed to assist during steady-state walking, and are usually tested within the controlled conditions of a laboratory where this walking is easily realized on a treadmill. This regulated environment enables researchers to more easily design control strategies and paradigms that augment the wearer in sync with their gait.

Despite the advent of these high-performance devices, exoskeletons are still far from entering ubiquity in society. The field has made great strides in the design and control of exoskeletons within the laboratory, but for exoskeletons to truly impact society, they must leave the realm of academic research and enter people's daily lives. Thus, it is paramount that we understand and remove the obstacles preventing this widespread dissemination. There are two major *fundamental gaps* that have inhibited the spread of these technologies: 1) although the field has had great successes in augmenting wearers using the physiological metric of metabolic cost reductions, users may not perceive these benefits, which calls into question the primacy of this metric and underscores the need for a suitable standard according to which exoskeletons may be designed to have perceivable value to their wearers; 2) the controllers developed for augmentative exoskeletons are developed for tightly controlled laboratory conditions that may be advantageous for researchers, but are not designed for the non-steady gaits and terrains that feature in the real world, which reveals the need for a new generation of controllers that can function in practical environments. *To close these gaps*, in this dissertation I 1) investigate the metrics by which we design exoskeletons, with specific attention to how the user perceives and interprets the benefits from these technologies, and 2) develop exoskeleton controllers capable of augmenting the wearer during locomotion outside the laboratory. By addressing these gaps, we can make progress towards the eventual goal of getting exoskeletons out of the laboratory and into people's daily lives where they can fulfill their promise to transform human mobility.

1.2 State of the Art

1.2.1 Exoskeleton Metrics

The metrics by which we judge exoskeletons as successful influence their physical design, controllers, and potential widespread impact. In the case of rehabilitative technologies, the obvious physiological objective is to restore impaired gait function. These exoskeletons apply torque to compensate for weakened muscle-tendons [205, 204, 164, 27, 151], and are deemed successful

when they result in restored able-bodied kinematics or kinetics. Rehabilitative exoskeletons can thus be explicitly designed and controlled with this goal in mind. For augmentative exoskeletons, however, it is less clear what the metric of success should be. Users of augmentative technologies are typically able-bodied, and thus do not require powered assistance to return to a lost level of capacity. In the absence of as natural a goal as the restoration of normative gaits in the rehabilitative case, today's augmentative exoskeletons are developed and assessed based on their capacity to fulfill physiological objectives. Among these, the reduction of metabolic expenditure during locomotion (*i.e.* a reduction of the calories burned) has emerged as the "gold standard" for exoskeleton success [150]. Exoskeletons designed with this objective typically apply powered assistance at either the ankle joint [123, 122, 149, 110, 85, 202, 128, 32] or hip joint [132] to substitute for some of the mechanical power required from the user's muscles, and control their assistance in synchrony with the wearer to reduce the upstream metabolic expenditure. For example, Mooney *et al.* developed the first autonomous exoskeleton capable of reducing the wearer's metabolic rate below baseline unassisted levels [123]. Panizzolo *et al.* demonstrated metabolic rate reductions using a powered soft exosuit that applied assistance at the hip [132]. Collins *et al.* demonstrated that metabolic expenditure could be achieved using a passive exoskeleton that applied torque via a spring at the ankle[32]. This physiological objective is intuitively meaningful in that if an exoskeleton reduces the muscular power, and thus, the metabolic demands on the wearer, the user will have more energy to spare to perform strenuous locomotion tasks. An exoskeleton is thus deemed successful if it reduces the metabolic expenditure of the wearer below that of unassisted walking, which should be reflected in functional outcomes such as task completion time [169]. In addition, a user's metabolic expenditure is an objective measurement, and can be quantified in a laboratory setting [195, 154] using methods such as indirect calorimetry.

Beyond the metric of metabolic rate reduction, researchers in augmentative exoskeleton research have also considered other alternatives to evaluate exoskeletons. For example, Lewis *et al.* employed pneumatically driven hip exoskeleton to reduce net hip joint torques as an objective [103]. Muscle activation has also appeared as an alternative biomechanical metric. Lenzi *et al.* used a powered hip exoskeleton to reduce the activations of both ankle and hip muscles [101], while Zhu *et al.* demonstrated activation reductions using a powered knee exoskeleton [204]. Other researchers have also investigated operational metrics as a means to judge exoskeletons. Bequette, Norton, and Stirling employed a series of audiovisual tasks to evaluate a full lower-limb exoskeleton based on cognitive load [11]. King, Gonzalez, and Stirling used the metric of agility, quantified via IMU measurements, to evaluate the effect powered bilateral ankle exoskeletons on individuals' ability to quickly accelerate and perform turns. Acosta-Sojo and Stirling conducted an analysis of specific lower-limb muscle activations to inform operational outcomes when using powered bilateral ankle exoskeletons [4]. With operational outcomes in mind, Stirling *et al.* con-

ducted a systematic review to identify pertinent cognitive and human factors that may inform these outcomes, such as the mental burden to use a device or the trust placed in the robot [169]. In addition to the primary standard of energetic reductions, these alternative metrics provide an additional source of insight with which to evaluate exoskeletons.

With the metric of metabolic rate reduction as a guiding force, modern exoskeletons have augmented their wearers by reducing their metabolic rates by an average of approximately 14% relative to unassisted walking [123, 122, 132, 149, 110, 85, 202, 128]. The focus placed on this objective function has led to the rise of promising 'human-in-the-loop' optimization (HILO) techniques [95, 201, 42, 186, 104, 163, 61]. In this control paradigm, users experience powered assistance from the exoskeleton in the form of a torque profile applied to one of the user's joints. The exoskeleton controller uses online optimization algorithms to adapt the parameters of this torque profile based on a continuous measurement of the metabolic rate signal, eventually converging on an "optimal" profile that minimizes the wearer's metabolic rate. These techniques have quickly proliferated throughout the field of exoskeleton control. Koller *et al.* were among the earliest to implement this optimization by incorporating physiological feedback from the human via indirect calorimetry, and used gradient methods to optimize the onset timing of torque assistance from a pneumatic exoskeleton [95]. Zhang *et al.* applied HILO to simultaneously optimize four parameters of a torque profile applied by an exoskeleton emulator system, and reported greater metabolic cost reductions using this optimized profile versus a generic profile [201]. Ding *et al.* corroborated the benefits of individualized assistance in a study that optimized the peak torque and torque timing from a hip exoskeleton [42]. Recently, Gordon *et al.* used musculoskeletal modeling to estimate instantaneous metabolic rate rather than rely on the sparse and noisy measurements from indirect calorimetry, which led to reduced optimization times within HILO when it was implemented on a pelvis exoskeleton [61]. These successes have motivated the development of experimental infrastructure, such as tethered emulator systems [41, 21], whose purpose is to investigate exoskeleton design and control in terms of metabolic rate reductions. Furthermore, numerous studies have investigated the biomechanical underpinnings for metabolic rate reductions in the search for ever-greater energetic improvements [120, 193, 201, 85]. Finally, alternative metrics for success, such as net joint torque reduction [103] and muscle activation reduction [101, 204], are often simply proxies for metabolic rate which may be easier to measure. These developments sharply illustrate the primacy of the metric of metabolic rate, and emphasize the tight coupling between exoskeleton development and exoskeleton objectives.

Despite this primacy, it has not yet been established whether or not users can even perceive these reductions. While evidence exists that humans may be able to 'subconsciously' sense changes in their metabolic rate, it is not known whether these changes can be perceived *consciously*. Donelan *et al.* demonstrated that people choose a step width at which metabolic rate is minimized during

walking [43]. In a follow-up study, Selinger *et al.* demonstrated that exoskeleton wearers can re-optimize their gait patterns to minimize their metabolic rate when a resistive torque from a knee exoskeleton perturbed them from a normative gait pattern [155]. Finally, Wong *et al.* discovered that these optimizations in gait pattern did not spontaneously occur outside the laboratory without forced exploration of different gait patterns [195]. The presence of these gait optimization suggests that humans have, at some base level, the ability to sense metabolic rate; however, as these results demonstrated that humans do not spontaneously return to a minimal rate without forced exploration, exoskeleton wearers likely do not have a conscious perception of their metabolic rates.

Conscious perception is integral to decision-making, such as the decision to purchase, use, or otherwise voluntarily adopt augmentative exoskeletons; these conscious decisions are prerequisites for augmentative exoskeletons to reach widespread adoption within society. If the user is unable to perceive the metabolic benefits from an exoskeleton, the user is then unlikely to incorporate that value into decision-making. This would call into question the primacy of the metabolic cost reduction metric for the assessment and development of exoskeletons based on this objective, and could result in devices whose benefits are not perceptibly valuable despite being energetically valuable [193, 105, 42, 121, 99, 132]. Users must thus actually perceive the metabolic benefits from exoskeleton assistance before they can derive value from it.

However, this perception is not enough to ensure voluntary adoption on the part of the user; for this, exoskeletons must be designed to provide *value* to their wearers. Once the perception of the benefits of an augmentative exoskeleton is established, the value of these benefits must be great enough that users voluntarily choose to adopt these technologies. Exoskeletons must provide experiences that illustrate their value to potential wearers. Additionally, this value must be great enough such that it overcomes the potential "costs" of exoskeleton use, such as the known disadvantages of monetary cost or discomfort. A technology's perceived value has a significant impact on its adoption throughout society. Within the field of management science, Davis proposed the Technology Acceptance Model (TAM) [40], which demonstrated a significant correlation between the consciously perceived usefulness of software and users' intent to adopt the software [40]. Subsequently, King and He found that TAM was applicable to different technologies[93], including broadband internet [129], telemedicine [79], and smart watches [29]. These investigations suggest that when users are considering the potential adoption or purchase of an exoskeleton, they must perceive sufficient value to outweigh the cost of the exoskeleton's price, weight, or aesthetics. Thus, the metrics used to design augmentative exoskeletons must result in technologies that are perceived as valuable by users before voluntary adoption occurs.

1.2.2 Exoskeleton Control

For exoskeletons to see this widespread use, their controllers must be capable of functioning in the practical environments of the real world. Prior work has primarily focused on controllers suited to the steady-state, treadmill-based locomotion within a laboratory environment. This controlled setting makes it easier for researchers to develop and validate control strategies that provide useful augmentation in synchrony with the user's gait. In these settings, both the walking task and the *phase rate*, e.g. the rate at which the human continuously progresses through the gait cycle, remain constant. Given this steady-state locomotion, progression through the gait cycle can be reasonably estimated by normalizing time by the approximately constant stride period, which is itself modeled as a constant and quantified as the time between prior ipsilateral heel strike events. This normalized time is then used to predict the next heel strike event. Within the confines of the laboratory[180], this timing-based estimation (TBE) approach is effective for delivering torque assistance in the form of pre-defined torque profiles, which then deliver real-time assistance based on the normalized time elapsed. The aforementioned reductions in metabolic cost during treadmill walking [150] have been widely achieved using this torque profile paradigm. Furthermore, HILO has made great use of the TBE strategy to directly optimize torque profiles in real-time, as the unchanging treadmill conditions allow it to continuously update profiles solely on the metabolic signal without confounding from any other changing environmental variables [202, 42]. However, while the TBE paradigm for providing torque assistance is well-suited to the steady-state conditions of the laboratory, these conditions are not likely to feature in the practical environments of the real-world, in which locomotion is highly transitory and the walking task itself can change irregularly. For exoskeletons to fulfill their potential to transform the mobility of the public, their controllers must be capable of adapting to continuously changing gaits and tasks such that they can provide seamless augmentation in the real world.

Recent work has begun to address the challenge of building controllers capable of functioning during non-steady-state gaits by introducing the concept of *gait phase* to encode progression through the gait cycle. The gait phase variable monotonically increases from 0 at a heel strike to 1 at the next ipsilateral heel strike [184]. The time derivative of phase, phase rate, does not have to be constant; phase can thus parametrize locomotion at different speeds. Phase-based controllers for augmentative exoskeletons [157, 88, 89, 158, 200] and powered prostheses [75, 141, 140, 76] can thus adapt assistance when the user dynamically changes their speed by continuously adapting the rate of phase progression. Unlike TBE controllers, phase-based controllers must estimate the gait phase in real time using the sensors on the lower-limb robot, which can be done numerous ways. For example, Holgate *et al.*[75] used the relationship between shank angle and velocity in the phase plane to estimate gait phase using shank angle measurements. Quintero *et al.*[140] and Hong *et al.*[76] exploited a similar relationship between thigh angle and its integral in a phase

controller for a knee-ankle prosthesis, while Rezazadeh *et al.*[141] used thigh angle measurements and holonomic constraints to estimate phase. Machine learning techniques have also been used to extract estimates of phase from onboard measurements. Kang *et al.*[89] deployed a convolutional neural network to estimate phase using the sensors on a hip exoskeleton, which was then used to provide torque assistance. Shepherd *et al.*[158] implemented a similar network using the sensors on an ankle exoskeleton. This prior work highlights the potential for phase-based controllers to drive lower-limb robots throughout varying speeds. However, a non-constant phase alone cannot easily handle transitory walking tasks beyond level-ground walking.

Prior work has attempted to solve the challenge of estimating changing walking tasks using the framework of machine learning classifiers. These classifiers use patterns in the lower-limb robot's sensors to infer the wearer's intended task from among an enumerated list of candidate tasks, such as stair ascent or incline walking, and then apply the correct task-specific controller [180, 80, 197, 86, 196, 106, 78, 89, 158]. Although this strategy is capable of handling discrete changes in task, it is unable to easily detect a continuously varying task or tasks that are not explicitly contained within the candidate tasks. This continuous adaptation is critical to defining the behavior of the exoskeleton during the ever-evolving environments of the real world; in particular, gracefully handling mode transitions is critical to ensuring a seamless interaction between the robot and the wearer [169]. Recent work has extended the concept of continuous phase to develop gait models parameterized by *task variables*, such as ground slope or stride length, that parameterize a continuously changing walking task [47, 46]. These gait models take phase and task variables as input and predict different kinematic measurements. Because task variables can continuously change, they are potentially capable of encoding a wider range of walking tasks compared to categorizing by discrete tasks. Camargo *et al.*[24] used machine learning classification techniques to determine which of the ground incline, step height, or walking speed continuous task variables to estimate depending on ambulation mode, which highlights the potential for task variables to instantaneously parameterize locomotion; however, this approach was not validated on robotic hardware, and used multiple EMG electrodes, IMUs, and goniometers which may not all be available aboard a wearable robot. Embry *et al.*[46] used task variables and a continuous gait model of thigh angle to control a lower-limb prosthesis, but was limited to estimating task variables only once per stride. Notably, Holgate *et al.* [75] exploited the relation between tibia angle and angular velocity in the phase plane to continuously update estimates of the stride length task variable; however, this relation does not apply to non-steady-state walking, and was also not generalizable to other joints or task variables such as ground slope. For exoskeletons, and lower-limb wearable robots more generally, to seamlessly adapt their assistance to the varying terrains of the real world, new methods are needed for real-time, continuous task variable estimation.

The Bayesian filter framework is a promising strategy to continuously estimate both phase and

task variables using real-time sensor data. A Bayesian filter can use noisy measurements, such as those from a thigh IMU onboard an exoskeleton, and estimate hidden states that are not directly measurable, such as gait phase and continuous task variables. This framework is capable of identifying differences in measured gait biomechanics between and within strides to infer underlying changes in *gait state* variables (phase and task variables together). The classic implementation of the Bayesian filter is the well-studied Kalman filter, which is applicable to linear systems with Gaussian noise. However, for nonlinear systems, different implementations of the Bayesian filter with varying levels of simpliyfing assumptions are available [148]. The Extended Kalman Filter (EKF) locally linearizes the nonlinear system and thus features the strongest assumptions. The Unscented Kalman Filter (UKF) assumes less, and uses judiciously chosen points and weights to approximate the nonlinear measurement update via quadrature integration [185]. Finally, the Ensemble Kalman Filter (EnKF) has the weakest assumptions, solely relying on the assumption of a Gaussian state distribution to use Monte Carlo techniques to calculate the measurement update [48]. Bayesian filtering has seen success within the field of wearable robotics to estimate human gait. Clark *et al.* [30] used measurements of heel pressure and shank angle to estimate stance progress and control the kinematics of an ankle prosthesis, although it was not capable of estimating progression through the swing phase. Karulkar *et al.*[90] estimated stride length using a Kalman filter and hip and knee velocity measurements on an exoskeleton, although this estimate only updated non-continuously twice per step. Thatte *et al.*[173] recently developed an EKF-based controller for a lower-limb prosthesis that estimated phase and phase rate with more accuracy in non-steady-state locomotion compared to conventional TBE approaches. These results highlight the potential for a Bayesian filter implementation to estimate the human gait state to control augmentative exoskeletons in the continuously varying conditions outside the laboratory. Furthermore, they motivate further investigation into the fundamental problem of gait state estimation using the Bayesian framework to identify which assumptions and sensor measurements are necessary for successful estimation.

1.3 Dissertation Contributions

This dissertation contributes to the development of both metrics for success of augmentative exoskeletons and new control strategies for exoskeletons outside the laboratory. Both goals are important to solving the lack of exoskeleton adoption outside the laboratory, as to facilitate this adoption, exoskeletons must be designed and evaluated such that they are consciously valuable in their benefits to users, while also having controllers capable of providing consistent assistance throughout the transitory gaits and environments of the real world. The first goal of identifying suitable metrics for the assessment of augmentative exoskeletons is addressed by the second and

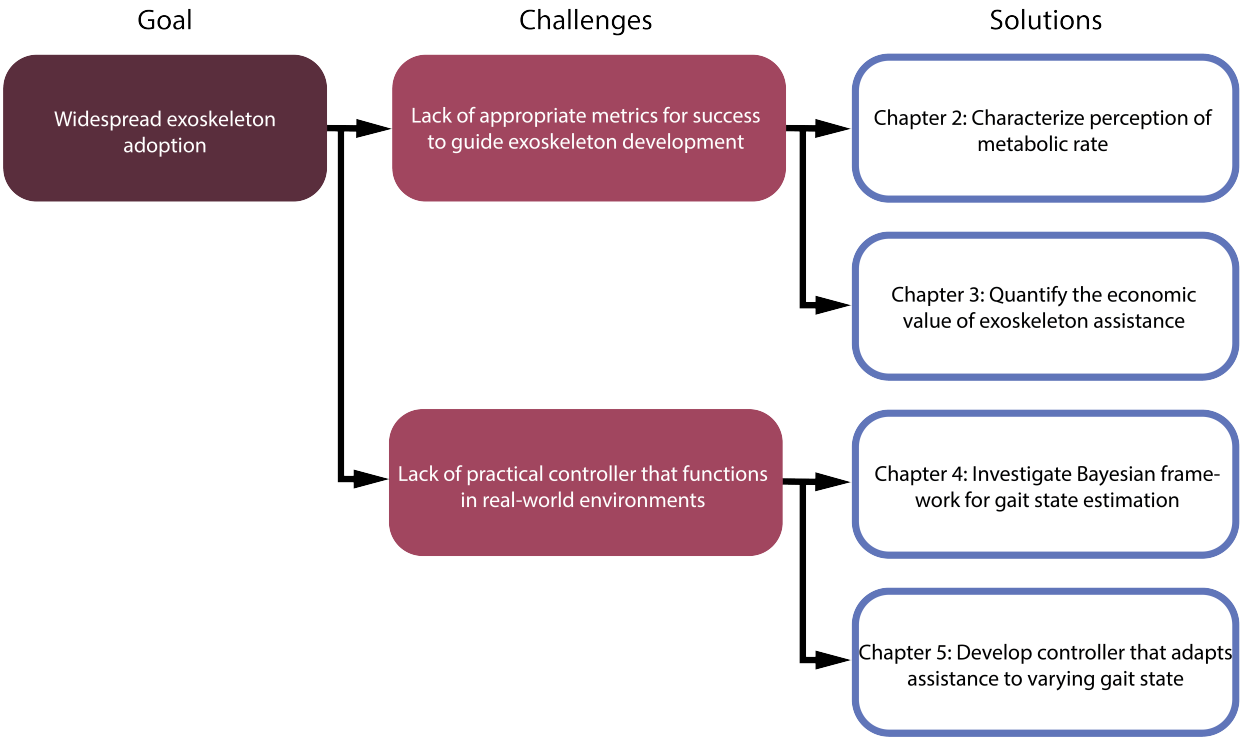

Figure 1.1: The central goal of this dissertation, the challenges to reaching that goal, and the four chapters presented to solve those challenges.

third chapters (Fig. 1.1). Chapter 2 quantifies the perceptual threshold for changes in the metabolic rate, which has been the field's predominant standard for evaluating exoskeleton success, with the motivation of elucidating whether this metric will lead to exoskeletons that are perceptually beneficial. Chapter 3 proposes an alternate metric for developing exoskeletons—that of economic value—that may be more conducive to voluntary adoption of exoskeletons, and evaluates state-of-the-art bilateral ankle exoskeletons using this metric. The second goal of developing exoskeleton controllers that function throughout practical applications is addressed by the fourth and fifth chapters (Fig. 1.1). Chapter 4 investigates the fundamental challenge of learning a vector representation of human gait capable of encoding real-world locomotion—the gait state—using Bayesian filtering techniques, with the goal of determining which Bayesian filter is best suited for learning realistic human gaits for real-time implementation on exoskeleton hardware. Chapter 5 then implements the best filter from Chapter 4 on bilateral ankle exoskeletons and demonstrates the ability to update torque assistance in response to accurate predictions of the gait state in real-time. By contributing to these goals, the investigations contained in this dissertation advance the overarching objective of translating these technologies out into the real world where they can transform the mobility of the public.

In Chapter 2 of this dissertation, I investigate the conscious perception of energetic changes

caused by powered exoskeletons by quantifying the Just Noticeable Difference (JND) of changes in metabolic rate. The JND acts as a threshold below which humans cannot consciously perceive the metabolic benefits from these technologies. This study refines our understanding of metabolic rate reductions as the "gold standard" for exoskeleton success by revealing if humans can actually sense the assistance from these technologies. This chapter has been published in the *Journal of NeuroEngineering and Rehabilitation* [117], with further analysis on the underlying psychophysical methodology that enabled this study having been presented at the *2020 IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics* [114] and *2018 Dynamic Walking* [113].

In Chapter 3 of this dissertation, I investigate the economic value of exoskeleton assistance. The goal of this work is to identify the actual value of exoskeletons using the metric of economic value as a potential alternative to metabolic cost reductions. This new metric is more viable compared conventional physiological metrics as it is both more accessible, in that it does not require specialized equipment, and more intuitive, as users likely have e better grasp of benefits expressed in terms of monetary currency rather than in terms of biomechanical quantities (*e.g.* calories burned). In this study, I leverage the Vickrey second-price auction within economics to measure participants' valuations of their walking times across different walking conditions, such as walking with and without exoskeleton assistance from bilateral ankle exoskeletons and walking with and without the exoskeleton itself. This chapter has been published in *Nature Communications Engineering*.

In Chapter 4 of this dissertation, I investigate the fundamental problem of gait state estimation with the Bayesian filtering framework. The goal of this work is to understand which implementation of the Bayesian filter is best suited to the problem, and to understand which gait state variables can be suitably estimated with different combinations of sensors that are available on various lower-limb powered exoskeletons. The results of this work result in a feasible implementation of the Kalman filter to continuously estimate the gait state throughout various real-world locomotion patterns and terrains, which will then adapt exoskeleton assistance in real-time within these practical environments. This chapter has been published in the *IEEE Robotics and Automation Letters*, and has been presented at the *2023 IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics* [118].

In Chapter 5 of this dissertation, I build upon the work presented in Chapter 4 to develop an Extended Kalman Filter (EKF)-based exoskeleton controller that continuously estimates both the phase variable and the task variables of ground incline and stride length, which are then used to update the torque profile of bilateral ankle exoskeletons throughout various controlled and outdoor environments. This work is significant as it demonstrates the EKF's ability to estimate both phase and task in real-time and presents a practical exoskeleton controller that is capable of adapting its augmentation throughout the non-steady-state walking conditions of the real world. This chapter

has been published in the journal *IEEE Transactions on Robotics* [116].

As a whole, this dissertation presents fundamental strategies for the translation of exoskeletons out of the laboratory and into the daily life of the public. This work contributes results about the conscious perception of exoskeleton benefits, a new user-focused economic design metric for wearable robots, and a theoretical analysis and practical evaluation of an adaptive multi-modal controller for ankle exoskeletons that all bring the field closer to the translation of exoskeletons.

CHAPTER 2

Can Humans Perceive the Metabolic Benefit Provided by Augmentative Exoskeletons?

2.1 Context

Chapter 2 of this dissertation focuses on investigating the conscious perception of energetic changes caused by powered exoskeletons, with the goal of elucidating if the dominant standard of metabolic rate changes will be perceived as useful and thus be conducive to exoskeleton success outside the laboratory. This is achieved through quantifying the Just Noticeable Difference (JND) of changes in metabolic rate, which serves as a threshold for human perception of the benefits of such technologies. The study refines the understanding of exoskeleton success by providing insight into whether humans can consciously sense the assistance from these devices. The chapter has been published in the *Journal of NeuroEngineering and Rehabilitation* [117], with additional analysis presented at the *2020 IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics* [114] and the *2018 Dynamic Walking* on the underlying psychophysical methodology utilized in this study.

2.2 Background

The purpose of augmentative exoskeletons is to help people exceed the limitations of their human bodies. These technologies apply mechanical assistance to the joints of the legs during locomotion, thereby reducing the physical demands on the wearer's neuromuscular system. The potential uses for these technologies are broad and impactful, including assisting people's abilities to walk, run, jump, and/or carry loads. Consequently, lower-limb exoskeletons may improve the mobility of people with disabilities, as well as those completing sustained, physically demanding activities (*e.g.* first responders, postal/supply chain workers, and military personnel, among others). Recently developed systems for human augmentation applications are untethered [133, 132, 92],

lightweight [121, 149, 133], and powerful [21, 206]. While recent work has been encouraging, an ongoing challenge has been quantifying the success of these systems; the quantification of an exoskeleton's ability to reduce the metabolic expenditure of walking (*i.e.* calories burned) has emerged as a focus of the field [150].

Exoskeleton researchers have focused on the reduction of metabolic rate because it is intuitive, measurable, and supported by previous research. State-of-the-art exoskeletons have consistently reduced the metabolic expenditure needed for walking by approximately 14% relative to not wearing an exoskeleton [123, 122, 132, 149, 110, 85, 202, 128]. These exoskeletons apply powered assistance at either the ankle joint [123, 122, 149, 110, 85, 202, 128] or hip joint [132] and implement control strategies that operate in tandem with the wearer to reduce their metabolic expenditure. Intuitively, if an exoskeleton is successful, the muscular effort required will be reduced, which should be reflected in an upstream reduction in the metabolic power required from the wearer. In addition, metabolic expenditure can be objectively measured in a laboratory setting, meaning it does not have the challenge of quantification that plagues other potentially subjective metrics of success (*e.g.* comfort, stability, or preference, among others).

There is mounting evidence that humans may be able to 'subconsciously' perceive their metabolic rate, but it is not yet known whether these changes can be perceived *consciously*. Donelan *et al*. showed that people choose step widths that minimize their metabolic rate during walking [43]. Subsequently, Selinger *et al.* demonstrated that exoskeleton wearers can re-optimize their gait patterns to minimize their metabolic rate when manipulated externally (with an exoskeleton) [155]. That is, the resistance of a knee exoskeleton was varied to incur a metabolic penalty during normative walking patterns, and participants needed to modify their gait patterns to reduce the superimposed metabolic burden. Participants converged to non-normative gait patterns that minimized metabolic rate, but this optimization did not occur spontaneously outside the laboratory [195]. Since exoskeleton wearers choose gait patterns that reduce metabolic rate, we believe this indicates people have some ability to sense this quantity (or something correlated). However, since people do not spontaneously stay or return to their lowest rate, it suggests that exoskeleton wearers do not have conscious knowledge of their metabolic rate or its gradient.

Conscious perception is a critical part of decision making. For an exoskeleton to appear valuable to its potential wearer, it must provide an experience that illustrates this value. Furthermore, this value must offset the potential "costs" of exoskeleton use. For example, without an intuitive and perceivable understanding of value, potential users may be unlikely to adopt exoskeletons with known disadvantages (*e.g.* monetary cost, discomfort, or being unfashionable). Previous research in the field of management science has investigated the implications of perceived value in technology adoption; one relevant framework proposed by Davis is the Technology Acceptance Model (TAM) [40]. In this model, Davis found a significant correlation between the consciously perceived

usefulness of software and users' intent to adopt the software [40]. More recently, King and He found that this relation was generalizable across many different technologies [93], such as broadband internet [129], telemedicine [79], and smart watches [29]. Thus, when potential exoskeleton users, manufacturers, and others are weighing the choice to adopt or purchase an exoskeleton, the consciously perceived value must outweigh the price, weight, aesthetics, and other costs of wearing a lower-limb exoskeleton.

The field of psychophysics focuses on quantifying human perception broadly [94]; for example, sensing may involve perception of images, temperatures, sounds [57], or metabolic rate. Forced comparison between two stimuli (such as asking which of two lines appears longer) across different trials is a powerful method for determining how humans can perceive changes in stimuli. This pair of stimuli is composed of a reference, which usually remains constant across trials, and a comparison, which changes from trial to trial. By analyzing large numbers of these comparisons, a perceptual model can be built that encodes and quantifies people's perceptual performance. The input to this model is the true difference between the reference and the comparison, and the output is a probability of the comparison being perceived as different from the reference. The model predicts that stimuli are accurately perceived when the difference between the stimuli is large, but that human perception becomes essentially random when the difference is small. These models are often visualized using a *psychophysical curve* [189], typically a sigmoid function. In general, a single psychophysical curve pertains to the specific reference stimuli about which the test is conducted.

The steepness of a psychophysical curve quantifies perceptual ability; namely, the smallest difference in stimuli that can be perceived reliably. Using a threshold for reliability of 75% [176], this delta is known as the *Just Noticeable Difference* (JND). The JND has been used to quantify meaningful differences in visual acuity [168], sound [166], taste [9], and weight [19]. Recently, wearable robotics researchers have begun to quantify the JND of various factors in the design and control of wearable robotic systems, including perception of prosthetic ankle stiffness by users [159] and clinicians [161], environment stiffness [6] and viscosity [5], electrical stimulation of the residual limb [2], and vibrations of an osseointegrated prosthesis [31].

In this study, we characterized exoskeleton users' conscious perception of their metabolic rate during assisted walking by quantifying the JND of metabolic rate changes. Understanding the human perceptual ability to sense this change is important because it has emerged as the gold standard by which exoskeletons are designed, controlled, and assessed. If exoskeletons are developed to impact a metric that is not perceivable by the user, it will likely hinder widespread success. To this end, we indirectly imposed different metabolic rates sequentially during walking by adjusting the assistance provided from bilateral ankle exoskeletons. Simultaneously, we recorded whether users perceived their metabolic rate to have increased or decreased as the control strategy changed.

We aggregated these data to estimate the JND for changes in metabolic rate. The contribution of this work includes new fundamental knowledge of how metabolic rate can be sensed during locomotion and a new benchmark for future exoskeleton developers who desire perceivable impact on metabolic expenditure. In addition, these results underscore the need for new metrics of exoskeleton success that are aligned with the value and experience of the user.

2.3 Methods

2.3.1 Participants

In this study, ten able-bodied participants (N = 10, 2 female, 8 male; age = 22.5 $\pm$ 3.17 years; mass = 70.9 $\pm$ 11.9 kg, Table. 2.1) walked using bilateral ankle exoskeletons on a treadmill. The required number of participants was chosen based on a power analysis to quantify a JND of 15% with 80% power and 5% type 1 error rate. We chose 15% as this was representative of the reductions achieved by the best performing lower-limb exoskeletons [99, 42, 105]. All participants provided written informed consent before participation. The study protocol was approved and overseen by the Institutional Review Board of the University of Michigan Medical School.

Table 2.1: Participant Data

Subject	Responses	Gender	Weight (kg)	Age (years)
1	53	M	52.2	20
2	100	M	74.0	23
3	34	F	72.0	21
4	100	M	86.0	24
5	100	M	78.5	21
6	100	M	74.0	23
7	100	M	82.5	24
8	100	M	59.0	19
9	100	F	53.5	20
10	100	M	77.0	30

2.3.2 Experimental Protocol

2.3.2.1 Walking Protocol

Participants experienced numerous metabolic rate changes in sequence that stemmed from the assistance provided by the ankle exoskeletons. Participants walked for 20 minute blocks, where each block consisted of 10 trials in series. Following each pair of trials, participants responded regarding which condition they perceived had a higher metabolic rate by agreeing or disagreeing to the binary question *"is the current level of exertion higher than the previous level of exertion?"*. Participants responded non-verbally with either a 'thumbs up' or a 'thumbs down.' Thus, each block consisted of nine comparisons across ten trials and participants completed approximately 11 blocks across three to four days of data collection. We chose the two minute walking duration for each trial to balance metabolic estimation quality with experiment duration; Zhang *et al.* demonstrated that the metabolic estimation error with two minutes of data is approximately 2% [33]. The two-minute trial duration also allowed the participants adequate time to experience and react to each walking condition.

Prior to the experiment, participants familiarized themselves with several aspects of the experimental protocol. Participants read a lay explanation of metabolic rate to familiarize themselves with the concept. Next, participants were primed to react to their feeling of general exertion by reading the instructions of the Borg Rating of Perceived Exertion [13, 15]. We chose to have the participants read information on the Borg Scale because it has previously been demonstrated to produce accurate estimates of exertion [14]. Finally, participants underwent a four-minute acclimatization period in which they were exposed to different representative exoskeleton behaviors that spanned what could be encountered during the experiment.

2.3.2.2 Exoskeleton Control

We used bilateral ankle exoskeletons (Dephy ExoBoot, Dephy Inc. Maynard MA) to manipulate the metabolic rate of the wearer. The exoskeleton (Fig. 2.6B) used electric motors ($\sim$300 W) and flat cable transmission ($\sim$15:1) to apply plantarflexion assistance during walking. The assistance was governed by parameterized current profiles that resembled a square pulse (see Fig. 2.1). The current profiles were governed by three parameters; we manipulated the onset timing, pulse magnitude, and pulse duration. We chose i) onset timings from a uniform distribution bounded between 25% and 50% of stride time, ii) pulse magnitudes from a uniform distribution bounded between 15 A and 25 A (corresponding to approximately 12 and 20 Nm with the ExoBoot's nonlinear transmission), and iii) pulse durations from a uniform distribution with variable bounds. The variable bounds for the pulse duration depended on the sample drawn from the onset timing distribution such that the square pulse had a minimum duration of 10% of stride, and a maximum

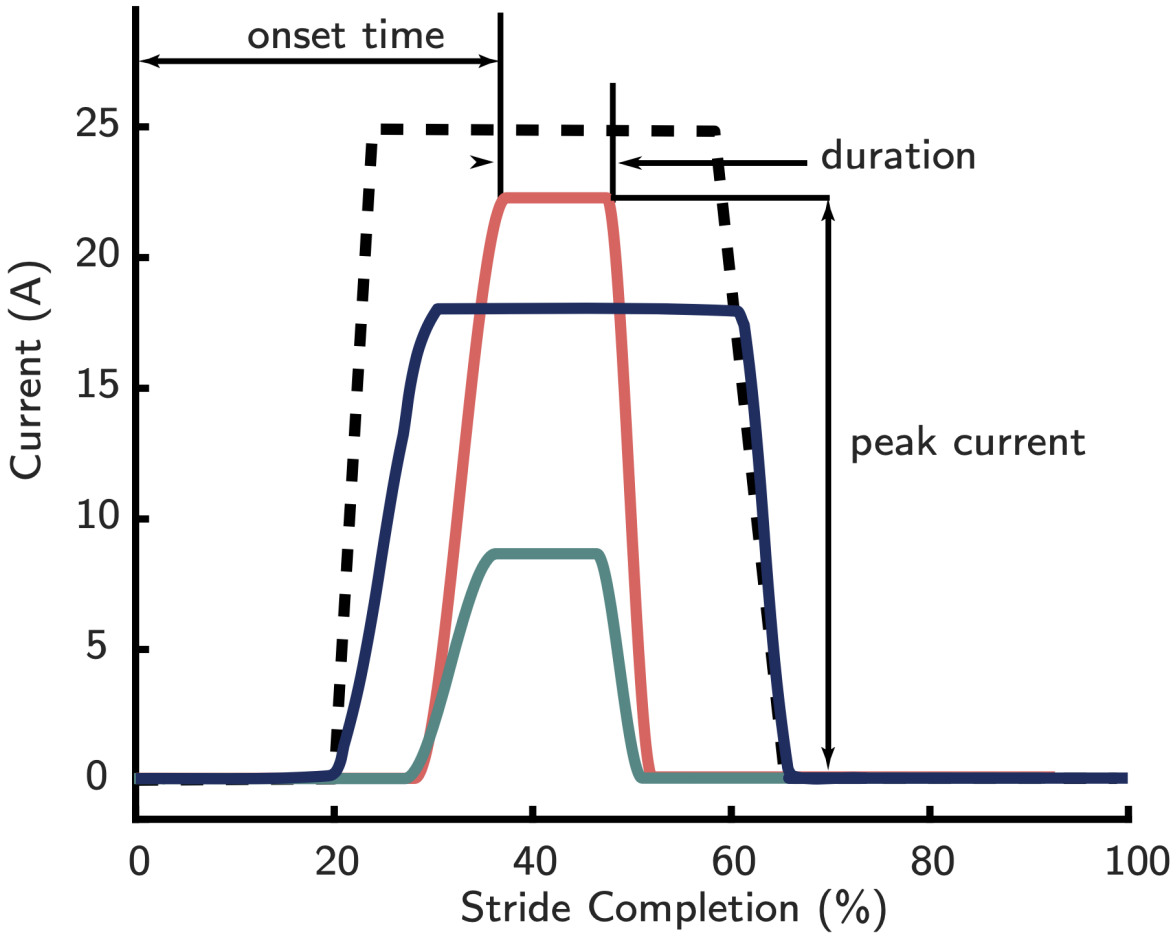

Figure 2.1: Sample exoskeleton current profiles used in this experiment (colored lines). The profiles resembled square pulses and were parametrized using the following parameters: the peak current of the profile, the onset time of the profile, and the duration of the profile. The total bounds of possible current profiles are delineated by the black dashed line. The profiles' currents were mapped to motor torques through the motor torque constant and exoskeleton's transmission.

duration of 60% of stride time. Onset timings that occurred earlier than 30% of stride were additionally constrained to have a minimum pulse duration of 20% of stride, which was imposed to prevent excessive device wear. We chose these bounds as they have been shown to significantly alter participant metabolic rate, and thus allow us to sample as wide an energetic range as possible [95, 147, 85] while balancing device integrity and user safety. The current profiles were described using the stride completion percentage to mitigate any variations in step length or cadence that occurred during the trial. Thus, we inferred the stride completion percentage using heel-strike events. We detected these events by thresholding the onboard accelerometers (MPU-9250, Invensense, San Jose, CA) [89, 24].

2.3.2.3 Metabolic Rate Sensing

Participants walked with a randomized torque profile for two minutes, which produced a first-order dynamic response in metabolic rate [154]. We measured participant metabolic rates through indirect calorimetry [66] (COSMED K5, Rome IT) (Fig. 2.6A). We estimated the user's steady state metabolic rate by fitting a first-order response [154, 95] to the breath-by-breath transient data gathered over the two minutes for each trial, with the steady state value representing the trial's metabolic rate. Prior to undergoing the walking protocol, participants stood still for four minutes to obtain their baseline metabolic rates. Each participant's standing metabolic rate was computed as the average rate over this four-minute interval. The standing metabolic rate was subtracted from each trial's metabolic rate measurement to isolate the metabolic effects of exoskeleton assisted locomotion (*i.e.* net metabolic rate).

2.3.3 Psychophysical Function Fitting

Within the field of wearable robots, established psychophysical protocols have previously been used to calculate the JNDs of different physical quantities [6, 159, 7, 31, 2]. A key aspect of the success of these methods is the ability to precisely control the magnitude of the stimulus under consideration. Therefore, to measure the JND of metabolic rate in the context of exoskeleton assistance, the change in metabolism must be precisely imposed. This poses a challenge for the assessment of the JND of metabolic rate, since the impact of an exoskeleton on metabolism varies with many controllable[95, 110, 85, 52] and uncontrollable [39, 71, 66] parameters. Thus, precise, repeatable control is not realizable in our application. Malcolm *et al.* [110] achieved a mean metabolic cost reduction of 6% by changing the timing characteristics of the exoskeleton, but with a reduction variability of $\pm 2\%$; this variability adds difficulty in regulating energetics as a stimulus quantity. Current psychophysical experimental protocols do not account for imprecisely controlled stimuli, such as the inability for an exoskeleton to predictably affect the metabolic rate during

locomotion. The effect of this imprecise control on the quantification of the JND is unknown. In addition, the optimal sampling strategy to estimate the JND with known statistical confidence when this type of variability is present is also unknown.

In work presented at the 2020 RAS/EMBS IEEE International Conference on Biomedical Robotics and Biomechatronics [114], we provide the foundation for assessing the effects of imprecise stimulus control when quantifying the JND in a broad class of psychophysical experiments; our specific focus is on quantifying the JND of metabolic rate during exoskeleton assistance. We investigate the two competing psychophysical sampling strategies using variable input stimuli, and demonstrate their ability to mitigate the negative effects of this variability. This protocol is then employed in the full study to quantify the JND of metabolic rate

Some common psychophysical concepts and definitions are useful for describing the following work. In the field of psychophysics, the fundamental relation between a stimulus and an observer's responses to that stimulus is given by the *psychometric function*. The psychometric function represents the probability that the observer will state that a comparison stimulus magnitude is higher than a fixed reference stimulus magnitude as a function of the comparison stimulus magnitude, and is well modelled by a sigmoidal function. Sigmoidal psychometric functions are naturally near zero for extreme negative comparisons and near unity for extreme positive comparisons, indicating that perception of the relative magnitude of stimulus and reference is nearly perfect when the difference between the two is large. A protocol in which a subject compares a stimulus to a reference and must provide a binary answer is termed a *forced-choice* protocol. In the *two-interval forced-choice* (2IFC) protocol used in this study, the subject experiences two stimuli sequentially and is asked to state which of the stimuli was higher in magnitude. From the psychometric function, we can calculate the *Just Noticeable Difference* (JND) for that function, which represents the minimum change in a stimulus quantity that must occur before an observer can reliably perceive that a change has happened. More precisely, the JND is typically defined as half the difference in stimulus between the points on the psychometric function with probability 25% and 75% [96].

To estimate the psychometric function, a typical experiment samples different stimulus levels, measures observer responses using a forced choice protocol, and fits maximum likelihood psychometric function parameters. A popular approach to choosing the stimulus levels is the *Method of Constant Stimuli* (MOCS), in which the comparison stimulus for each trial is randomly selected from a predefined vector of values. In this paper, we define the *conventional MOCS* sampling plan as a method in which the number of stimulus levels in this predefined vector is low ($\sim$5), but each is polled numerous times. By aggregating a subject's responses to whether these comparison stimulus levels are higher than a fixed reference, we obtain proportions that denote how likely a subject is to indicate a comparison value is higher than the reference. A variant of this sampling methodology is single presentation MOCS, which we refer to as *SP*. In SP, we instead present a much greater

number of stimulus levels ($\sim$100) only once, and fit the psychometric function to this binary data. A rich body of work has rigorously explored psychometric function estimation, including the optimal location of stimulus values [188], the number of samples taken [130, 18, 53, 94], and the type of function fitted [53, 178].

Each psychometric function and corresponding JND is specific to the fixed reference against which all comparison trials are compared. A common modelling choice that relates the JND to this reference is Weber's Law, which states that the ratio of JND to this reference is constant; this ratio is then defined as the *Weber Fraction* [127], and has been found to be constant (or near constant) for a wide range of reference stimulus levels far from the perceptual extremes in other sensing modalities [56]. A constant Weber Fraction relates the psychometric functions at different reference stimulus levels.[1]

This section presents a detailed description of 1) the perceptual model used to simulate psychophysical experiments with variable stimuli, 2) the binned MOCS sampling methodology and fitting protocol, 3) the SP sampling methodology and fitting protocol, and 4) the Monte Carlo experiment used to estimate and compare the performance of binned MOCS and SP under different magnitudes of stimulus variance. These comprehensive descriptions are intended to facilitate replication of these protocol in other physical experiments where the stimulus can be imprecisely controlled.

2.3.3.1 Perceptual Model

We begin by describing the perceptual model of the observer used in our simulation studies. This observer is modelled as responding to 2IFC discrimination tasks in which each of the responses is generated randomly, independent of the others. Using this observer model, we modelled the psychometric function Ψ describing these responses as a logistic function of stimulus x,

$$\Psi(x, \alpha, \beta, \gamma, \lambda) = \gamma + \frac{1 - \lambda - \gamma}{1 + e^{-\beta(x-\alpha)}}, \tag{2.1}$$

that was characterized by the following parameters: an experimental lapsing rate λ was fixed at the commonly used value of 0.02 [94]; a false positive rate γ, was fixed at the lapsing rate, since the experiment simulated discrimination tasks [94][53]; a slope parameter β that shapes the steepness of the logistic function; and the logistic function's threshold point α on the x-axis, which we fixed at the reference to simulate unbiased comparisons. This standard choice of function structure constrains the psychometric function to be odd symmetric about its inflection point, and is simple to manipulate analytically. The observer in this model responds to different values of comparison

[1] This is a very useful property for our purposes, but should be verified by fitting maximum likelihood models for different psychometric functions to test the assumption that it is constant with reference magnitude.

stimulus, referred to as stimulus levels; the value of Ψ at each of these levels denotes the probability that the observer will state that the comparison is higher than the reference at that level.

This psychometric function has a corresponding JND value, which is calculated by taking the difference between the values of x at $\Psi(x) = 0.75$ and $\Psi(x) = 0.25$ and dividing that difference by two [96][54]. The equation for JND then depends only on β,

$$\text{JND} = k/\beta, \tag{2.2}$$

with a scale constant k in terms of the fixed parameters.

$$k = \frac{1}{2} \ln \left[\frac{(0.75 - \gamma)(1 - \lambda - 0.25)}{(1 - \lambda - 0.75)(0.25 - \gamma)} \right]. \tag{2.3}$$

The Weber Fraction[2] (WF) is a function of both β and the reference α,

$$\text{WF} = \frac{\text{JND}}{\text{Reference Stimulus}} = \frac{k}{\alpha\beta}. \tag{2.4}$$

For both simulations, we defined a ground truth psychometric function with parameters α_0 = 10 and β_0 = 1, such that the probability of the simulated subject rating a stimulus s as larger than the reference α was modelled as

$$\Psi(s, \alpha, \frac{\beta_0 \alpha_0}{\alpha}, 0.02, 0.02). \tag{2.5}$$

This function's parameters also provided the benchmark $JND_0 = k/\beta_0$ (corresponding to the nominal reference α_0) for measuring estimation bias in the two sampling protocols. Subject responses to pairs of reference and stimuli were simulated randomly according to the probability in (2.5), as in the parametric bootstrap methods used in prior studies [188]. With prior knowledge of the benchmark psychometric function, we evenly spread the nominal stimulus levels S over the middle 98% support region of the psychometric function, which ensured an optimal sampling schema [188]. After the observer responses were generated at all stimulus levels, we used the Palamedes toolbox Maximum Likelihood Estimation (MLE) [94][53] to fit an overall psychometric function and JND to these simulated proportional response levels, which would then be compared against the ground truth function and JND.

2.3.3.2 The Impact of Variability

Our simulation study extends previous work by incorporating an analysis of stimulus variance during function fitting, which occurs when the experimenter does not have precise, direct control

[2]We assume the WF is locally constant for metabolic rates in the range of level-ground exoskeleton-assisted walking experiments [160, 127].

over the stimulus. We desired to identify the "optimal" psychophysical strategy to calculate the JND of a stimulus when confronted with this variability; that is, if the researcher can only afford N_t trials, what approach will yield a JND estimate with the least amount of bias or standard deviation? To aid in our analysis, we model this variability as normally distributed (Gaussian) noise. In our proposed physical experiment, σ represents the natural respiratory variability in metabolic rate during level ground walking, which results in an inconsistent imposition of the desired metabolic rate by the exoskeleton. Since this noise causes the reference to vary, the psychometric function (which is based on this reference) also varies. For each simulated trial's desired stimulus s and reference α_0 values, we add stimulus noise to the reference that perturbs the reference to $\hat{\alpha}$ and the stimulus to $\hat{s}$. We perform this corruption for all N_t trials, generating a vector of $\hat{\alpha}$ values that denotes the corrupted reference values for all trials. The effect of noise on the comparison manifests as noisy stimulus levels that the simulated observer experiences during each trial. For a given trial's corresponding comparison stimulus value s, we add stimulus noise to generate a new corrupt comparison stimulus $\hat{s}$. Similar to with the reference, for each stimulus level $S_{1,2,...N_s}$ where N_s is the number of stimulus levels, we perform this corruption for each separate stimulus value N_t/N_s times, to obtain a sub-vector containing noisy comparison values. We aggregate the N_s sub-vectors containing $\hat{s}_{1,2,...N_s}$ to obtain a vector of length N_t that represents the comparison stimulus values with added noise over all N_t trials.

2.3.3.3 Noise-Robust Methods

The impact of noise is to perturb the benchmark psychometric function and thus cause the observer to produce responses with probabilities far different than the baseline. As $\hat{\alpha}$ varies due to stimulus noise, it corrupts the benchmark Ψ by moving not only its reference from α_0 to $\hat{\alpha}$, but also its slope due to our constant Weber fraction model described in (2.4). To eliminate Ψ shifting in both slope and reference, we non-dimensionalize the threshold α_0 and slope β_0, which have stimulus units and inverse stimulus units respectively. For each given trial featuring noisy reference $\hat{\alpha}$ and noisy comparison stimulus $\hat{s}$, we non-dimensionalize using the following equations:

$$s^* = \frac{\hat{s} - \hat{\alpha}}{\hat{\alpha}}, \quad \alpha^* = 0, \quad \beta^* = \alpha_0 \beta_0. \tag{2.6}$$

This prevents Ψ from changing its shape due to noisy references, and re-characterizes all response probabilities as originating from a single logistic function.

After non-dimensionalizing the noisy reference and comparison stimuli, we generate observer response probabilities. For each trial $i = 1, 2, ...N_t$, we look up the corresponding $\hat{\alpha}_i$ and $\hat{s}_i$ values in their vectors and non-dimensionalize to obtain dimensionless quantities for both reference stimulus α_i^* and comparison stimulus s_i^*. By using (2.1) and evaluating $\Psi(s_i^*, \alpha_i^*, \beta^*, \gamma, \lambda)$, we

obtain the probability $\hat{p}_i$ that the simulated observer would state that $\hat{s}_i$ is greater than $\hat{\alpha}_i$. We perform this N_t times to obtain a third vector containing these probabilities for their respective noisy reference and comparison values. After all probability data are generated, we choose one of our two competing sampling methodologies to evaluate with the data.

We introduce the following two methods to mitigate the influence of noise on the stimulus quantities: an adaptation of conventional MOCS that bins noisy comparison stimulus values to nominal stimulus levels, and an SP approach that treats noisy comparisons as intentionally realized trials. To compare these two novel methods, we performed two simulations to characterize their performance in yielding accurate JND estimates with the presence of stimulus noise.

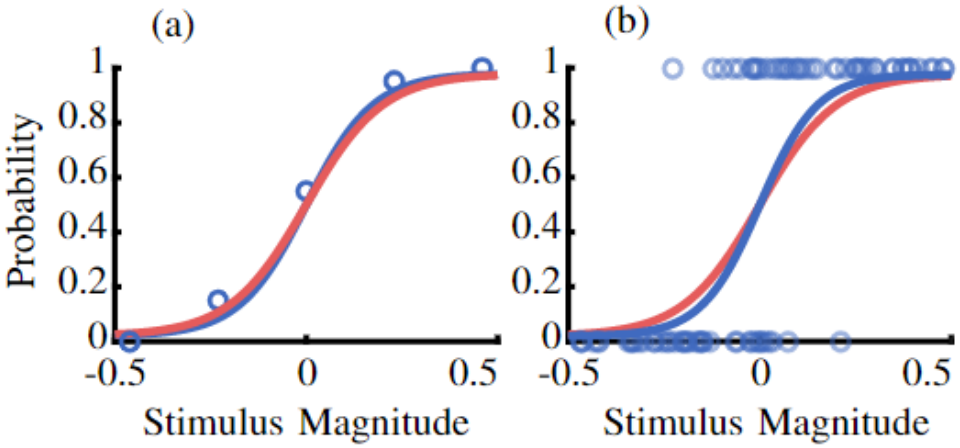

Figure 2.2: A sample psychometric curve (blue) fit to proportions of observer responses (blue circles) using the binning MOCS protocol (a). A sample psychometric curve fit to representative observer responses (blue circles) using the single-presentation protocol (blue) (b). The true underlying curve for both is shown in red.

2.3.3.4 Conventional MOCS with Binning

The first of these simulations (A) represents the adaptation of conventional MOCS to incorporate noise on both the reference and comparison stimulus values by binning noisy stimulus levels to their closest nominal values. Simulation A simulates MOCS with varying levels of noise and quantifies its effect on the number of trials needed for the psychometric function fitting procedure to converge to the true underlying JND value. To account for this noise in the conventional MOCS approach taken in Simulation A, we bin the actual non-dimensional noisy comparison stimulus values that the observer experienced to a bin centered on corresponding nominal non-dimensional

stimulus level values $S^*_{1,2,...N_s}$, expressed as follows:

$$S^*_{1,2,...N_s} = \frac{S_{1,2,...N_s} - \alpha_0}{\alpha_0}. \tag{2.7}$$

For each trial $i = 1, 2, ...N_t$, we generate a single binary response $\hat{v}_i$ from the observer using a Bernoulli distribution with probability $\hat{p}_i$ that the dimensionless comparison s^*_i is greater than the corresponding dimensionless reference α^*_i. We identify the appropriate bin for that response given by index k:

$$k = \arg\min_{j \in N_s} |S^*_j - s^*_i|, \tag{2.8}$$

where S^*_k denotes the dimensionless bin to which s^*_i is closest. In this simulation, all noisy s^* are placed in a bin; no s^*_i that are sufficiently far away from even the outermost bins are discarded. We count both the number of instances that the k^{th} bin was polled and the number of times the observer responded that the comparison was higher than the reference for the k^{th} bin. By dividing these two numbers, we obtain a response proportion for that bin; performing this binning procedure for all trials yields N_s proportions. We then fit a psychometric function using those proportions, S^*, fixed parameters α^*, γ, and λ, and free parameter β^*. Palamedes uses MLE to find the best fitting $\tilde{\beta}^*$ that corresponds to that function. We use (2.6) and (2.2) to solve for the resulting best-fit $J\tilde{N}D$. A representative function fit using this approach is shown in Fig. 2.2a.

2.3.3.5 Single Presentation MOCS

The second simulation (B) represents an alternative approach of incorporating noise into an experiment by using the single presentation MOCS (SP) protocol to treat noisy stimulus readings as "one-off" measurements dispersed evenly throughout a parameter space [178][124]. In the SP approach of Simulation B, the noisy comparison stimulus values experienced are instead treated as intentionally realized nominal values. In Simulation B, the number of stimulus levels N_s was equal to N_t, and thus each would be polled only once even without noise. The nominal stimulus levels in SP were also spread with even spacing over the middle 98% support region of the psychometric function. We follow a similar procedure as in Simulation A by generating responses using Bernoulli distributions using probabilities $\hat{p}$ for all noisy comparisons s^* and noisy references r^*. We poll each tuple of these quantities once and obtain binary responses at each s^* instead of proportions. In SP, we calculate $S^*_{1,2,...N_s}$ according to (2.7) using the respective $\hat{\alpha}_{1,2,...N_s}$ instead of α_0. We use this S^*, fixed parameters α^*, γ, and λ, and free parameter β^* to fit a logistic function using MLE. This yields the best fitting $\tilde{\beta}^*$ and $\tilde{\text{JND}}$. A representative function fit using SP is shown in Fig. 2.2b.

2.3.3.6 A Simulated Monte Carlo Comparison

To evaluate the competing simulation approaches, we varied both the noise level σ, which was scaled as a percentage of JND_0, over the following values: (0% 1% 10% 100% 1000%). We also varied the overall number of trials N_t, which ranged between 100 and 1000 in increments of 100, to measure their impact on the simulated JND estimates with these different sampling methodologies. We performed Simulation A with the number of stimulus levels N_s = 5 and N_s = 10, both reasonable numbers given prior psychophysical recommendations [94][159][53]. The number of times that each stimulus level was polled in Simulation A was then $N_t/5$ and $N_t/10$ respectively.

One thousand Monte Carlo replicates were completed for each sampling plan combination of simulation approach, number of trials, and magnitude of noise, which yields a distribution of JND estimates for that sampling methodology. The mean and standard deviation of these distributions can then be compared to the actual JND to calculate how accurate and precise that sampling plan. If a replicate failed to converge to a fit, that replicate was not included in further analysis. The complete process is illustrated in Fig. 2.3.

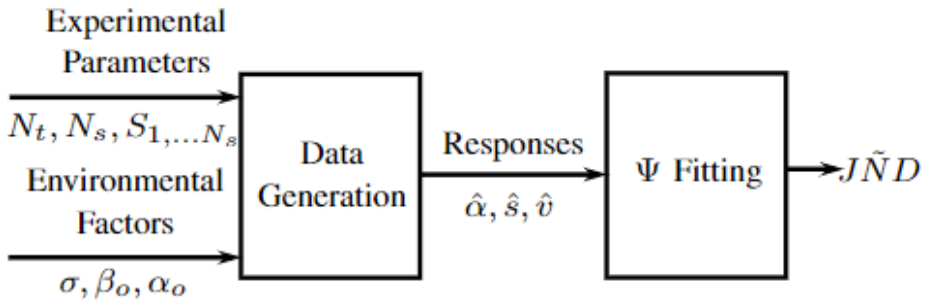

Figure 2.3: The simulated experimental process. The experimenter inputs the known or estimated noise value of σ as well as reasonable estimates of β_0 and α_0. These inform the experimenter such that they judiciously select the number of trials N_t distributed over the number of stimulus levels N_s with the option of binning the data in bins centered at $S^*_{1,2,\ldots N_s}$. Multiple simulated trials are run to generate tuples of noisy references ($\hat{\alpha}$), noisy comparisons ($\hat{s}$) and observer verdicts on the magnitudes compared ($\hat{v}$). Fitting a logistic function Ψ to these data yields a best-fit JND estimate $\tilde{\mathrm{JND}}$.

2.3.3.7 Results of Monte Carlo Simulation A

The mean bias of JND estimates for all sampling methodologies using binned MOCS in Simulation A are given in Fig. 2.4, which calculates bias of the estimates from the actual JND using the following equation:

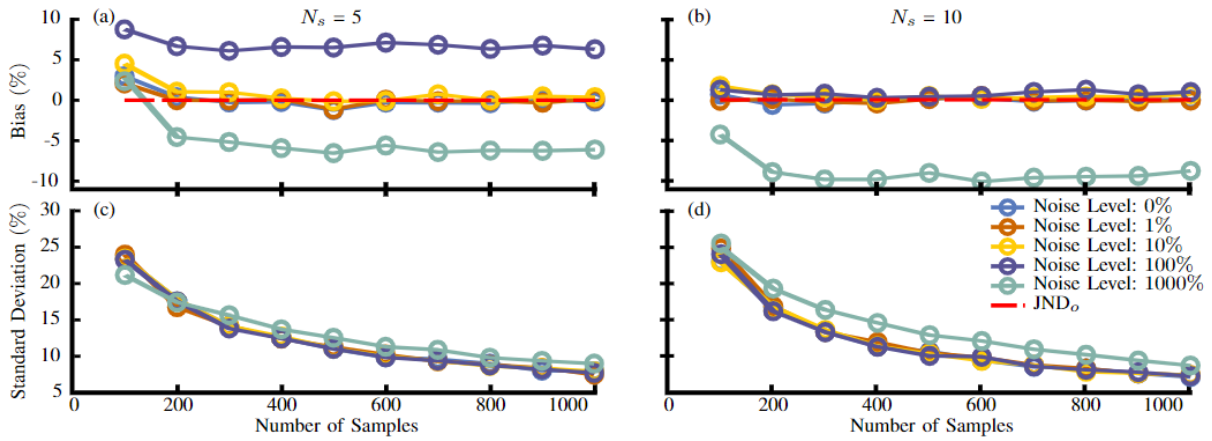

Figure 2.4: Results from Simulation A using the binned MOCS approach. (a) Mean bias estimates of JND with $N_s = 5$. (b) Mean bias estimates of JND with $N_s = 10$. (c) Standard deviation estimates of JND with $N_s = 5$. (d) Standard deviation estimates of JND with $N_s = 10$.

$$\text{MeanBias} = 100\% \times \frac{\tilde{\text{JND}} - \text{JND}_0}{\text{JND}_0} \tag{2.9}$$

The different curves represent different levels of simulated noise σ, expressed as a percentage of JND_0. The effect of noise manifests as a bias away from the true JND that increases with the value of σ. This bias is positive with increasing σ before reversing to a larger negative value once a sufficiently large σ has been reached. Increasing N_s from 5 (Fig. 2.4a) to 10 (Fig. 2.4b) mitigates the positive effect, and results in a larger negative bias at the 1000% noise level.

When the binned MOCS strategy is used and the stimulus noise is sufficiently low ($\sim$1% of JND), there is no apparent difference in fit behavior from an ideal noiseless case; increasing the number of trials N_t causes the mean JND estimate to converge to the true value, which is consistent with prior work [53].

The general trend in the standard deviations for JND estimates matches the expected inverse relation between deviation and increasing N_s [53] (Figs. 2.4c and 2.4d). Noise appears to have no impact on the standard deviation of JND estimates using binned MOCS, aside from an increase at the largest noise level.

2.3.3.8 Results of Monte Carlo Simulation B

A positive bias exists at high noise levels (100% and 1000%) and low N_s values for the mean bias of the JND estimates using SP (Fig. 2.5a). This low-N_s bias appears to depend on the handling of cases where the maximum likelihood solution does not converge. Unlike with the binned MOCS strategy, increasing noise level σ increases the standard deviation of the estimates (Fig.2.5b).

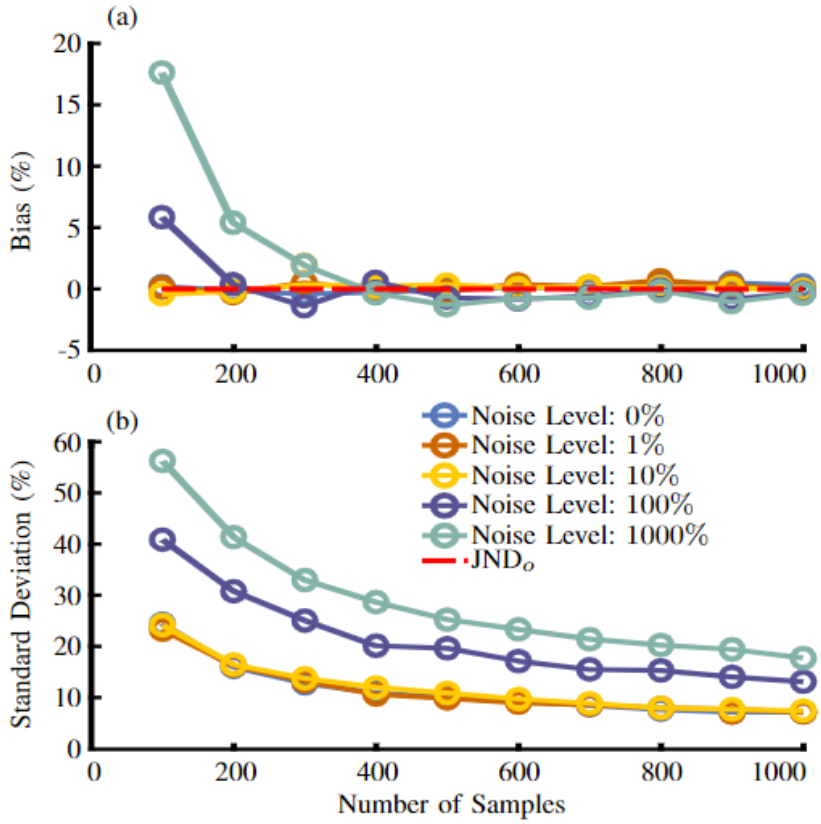

Figure 2.5: Results from Simulation B using the SP approach. Mean bias (a) and standard deviation (b) of Monte Carlo estimates of JND with $\alpha_0 = 10$.

2.3.3.9 Discussion of Experiment Designs

The effect of imprecisely controlled stimuli on psychophysical protocols is not understood, which poses a challenge in calculating the JND of variable signals such as that of metabolic rate. In this study, we characterized the impact of stimulus variability in the fitting process for psychometric functions. We modelled this variation as noise and incorporated it in both the reference stimulus and comparison stimulus values, and performed two simulations of alternative sampling methodologies for a 2IFC procedure. Our long term goal was to use the results of these simulations to inform a sampling strategy for that would mitigate the effects of stimulus variability, which we would then use in a physical experiment to calculate the JND of metabolic expenditure during exoskeleton-assisted locomotion.

Some common trends exist for the results of both simulations. The most obvious result is that as the number of samples N_t increases, the standard deviation of the sampling plan estimates decreases with an increasing number of samples, which is well established by the literature. An initial positive bias exists for both binned MOCS and SP at low samples before reaching steady-state behavior at 300 samples. This is likely due to our discarding of failed fits (MLE occasionally failed to converge) during our Monte Carlo simulations. The surviving fits are likelier to contain larger JND estimates, which skews the mean of the estimate distributions higher. This bias is eliminated by increasing the number of samples such that MLE can reliably find a fit.

The effects of this noise manifest differently in the two competing strategies explored in this study: binned MOCS and SP. In binned MOCS, extremely large stimulus noise causes bias in the mean JND estimates due to the effect of binning the noisy samples once they are perturbed far away from their nominal comparison values. Comparison values that are sufficiently outside the region of support of the psychometric function do not normally convey any useful information, as the sigmoid reaches asymptotic behavior and does not change value with changing stimulus in this region. However, by binning these values at the asymptotes to the outermost bins at the extremes of the stimulus levels, we disproportionately bias the edge bins towards 0% and 100% response probabilities when they would otherwise contain less extreme responses. These extreme responses, which would otherwise not contain useful information, artificially inflate the confidence Palamedes has in the outer bins when fitting functions, which return smaller JND values. The opposite effect happens at specific lower values of σ, which is caused by the inner bins being biased in an analogous way by the center bin's noisy comparison stimuli with response probabilities close to 50%. The center bin contains responses that are essentially guesses, so these responses corrupt the proportions of their adjacent bins by reducing the confidence of the observer's responses in the adjacent bins, and thus bias the JND to a larger value. The noise level at which these effects occurs depend on the relative magnitude of the JND to the noise, and the number of bins used.

In SP, the effect of noise is an increase in the standard deviation of the JND parameter and

thus a decreased confidence in our estimates, although this can be mitigated by increasing the number of samples taken. This effect is likely caused by the presence of perturbed comparison values s^* outside the psychometric function's optimal region of support. With increasing σ, more realized comparison stimuli fall in the asymptotic regions of the logistic function, and are therefore providing no useful information for the MLE function fitter. With fewer data points in the non-asymptotic region, confidence decreases, and so the standard deviation should increase with σ as observed.

With this knowledge, we can conclude that for experiments in which the noise vastly overpowers the signal, SP is the superior choice. Based on the results presented in Fig. 2.4, if the stimulus noise level is sufficiently low ($<\tilde{10}\%$), then the binned MOCS strategy can be used by assuming that the actual realized stimulus values are equal to the nominal values. In this scenario, binned MOCS is superior to SP, as the noise causes no mean bias or increase in variance, and the traditional trade-off between number of samples / fit confidence and experimental duration is present. When stimulus noise has overwhelming magnitude, such as in our experiment with metabolic rate, binned MOCS exhibits bias with large relative σ that does not disappear with increasing N_t, and depends on optimal bin placement. In a physical experiment, the real JND and Ψ are not known ahead of time, and thus the experimenter must make an educated guess as to the optimal location of the bins for the binned MOCS approach; if the guess is incorrect, the experimenter must expend more time to correct for bin placement. On the other hand, SP appears to avoid such biases, and more accurately reflects the reduction in estimate confidence due to the increase in non-informative tests.

To estimate the JND of metabolic rate, which denotes the magnitude of change necessary for consistent perception, our experimental protocol requires the normalization of metabolic rates. That is, to compare across sequential trials with differing references, normalization is needed to combine these data to obtain a single JND for each subject under the assumption of a constant Weber Fraction (see Limitations subsection) [114]. The Weber Fraction (WF) [127] is a metric that captures the differences in perceptual thresholds that are dependent on the magnitude of the reference stimulus used in the comparisons. By definition, the WF is the JND divided by the reference stimulus, thus it represents the percent change from the reference stimulus that is perceivable. For a wide range of stimulus magnitudes, the WF can be modeled by a constant [56].

For normalization, consider a sequential pair of metabolic rates A then B; we normalized B (the comparison) as a percent change in rate from A. Note that rate A is the reference that changes from trial to trial. We used the normalized metabolic rate differences and corresponding participant responses to fit a psychometric function. The psychometric function then provided JNDs with units of percent change of metabolic rate (rather than absolute units (W/kg)). The JND is then equivalent to the WF expressed as a percentage. This is advantageous as it allows the resulting

JNDs to characterize the perceptual thresholds of different tasks that may have different baseline rates (in W/kg), such as locomotion tasks at different walking speeds—the minimum perceivable change for those tasks would still be the percentage change relative to any baseline metabolic rate, as long as the comparison pertains to stimuli where the Weber Fraction is constant.

A logistic psychometric function was used to model participant responses. This model predicted the probability that the participant would choose "the comparison is *greater*" as a function of the normalized metabolic rate difference between the two trials. Using the convention from above, the psychophysical curve predicts the probability that rate B is greater than rate A, as a function of the relative difference between A and B. The logistic function of (reference-normalized) stimulus x had the following form,

$$\Psi(x, \alpha, \beta, \gamma, \lambda) = \gamma + \frac{1 - \lambda - \gamma}{1 + e^{-\beta(x-\alpha)}} \tag{2.10}$$

where $\Psi(x)$ was parametrized by the following variables: the experimental lapsing rate λ, which was fixed at the commonly used value of 0.02 [94]; the false positive rate γ, which was fixed at the lapsing rate since participants underwent a stimulus discrimination task [94, 53]; the logistic function's threshold point α on the x-axis, which anchors the center of the logistic curve and was set to 0; and the parameter β which governs the slope of the logistic function and is the only degree of freedom estimated during the fitting procedure.

Using the modeled psychometric curve, we quantified the JND which represents the minimum change in metabolic rate that must occur before an observer can reliably perceive with 75% accuracy [96]. The JND is calculated by taking the difference between the values of x at $\Psi(x) = 0.75$ and $\Psi(x) = 0.25$ and dividing the difference by two. By fixing the other parameters of $\Psi(x)$ at the values specified, the JND thus depended only on β

$$\text{JND} = k/\beta, \tag{2.11}$$

with a scale constant k in terms of the fixed parameters,

$$k = \frac{1}{2} \ln \left[\frac{(0.75 - \gamma)(1 - \lambda - 0.25)}{(1 - \lambda - 0.75)(0.25 - \gamma)} \right]. \tag{2.12}$$

Shallower slopes (indicating less sensitivity) caused higher JNDs, while steeper slopes (indicating higher sensitivity) caused lower JNDs.

2.3.4 Survey on Fitness Correlates

As a secondary outcome, we investigated potential links between the magnitude of the estimated JNDs and subject-specific fitness variables. At the close of the experiment, subjects completed

a survey in which they indicated how many hours per week they typically performed endurance activities (*e.g.* running), how many years they had spent performing these athletic activities, and their Body Mass Index (BMI). The intent of this survey was to provide initial investigation into potential underlying factors that may influence metabolic perception.

2.4 Statistics and Comparisons

A separate logistic model was fit for each participant using Bayesian analysis [97]. This approach yielded a posterior distribution of JND estimates for each participant. From this posterior distribution, we extracted the maximum likelihood estimate for each participant, which was considered the estimated JND [190]. Our approach of using Bayesian estimation enables quantification of both the JND value for each subject in addition to the uncertainty about our estimates. We chose Bayesian estimation because preliminary work indicated the conventional Maximum Likelihood Estimation approach could fail to converge [138]. We conducted our Bayesian analysis using the PyMC3 library in Python [146]. Each participant's prior distribution of JND estimates was chosen as a uniform distribution between 0% and 70%, representing a plausibly large range of perceptual abilities.

The posterior JND distributions were obtained by updating our prior distributions using the participant response data. We used the No-U-Turn Sampler (NUTS) [73] strategy—a Markov Chain Monte Carlo (MCMC) algorithm—to numerically approximate the posterior distribution of possible JND values; we used four sampling chains with 8000 tuning iterations and 4000 posterior predictive samples. We chose these values to balance computation time and accuracy. The JND with the highest likelihood in the posterior distribution was the nominal JND estimate. Each posterior also yielded a 95% credible interval for the JND estimates.

Our approach using Bayesian statistics enabled investigation of several assumptions made about the JND distributions. We compared three different JND models using the Watanabe-Akaike Information Criterion (WAIC) metric, which evaluates the predictive power of models and corrects for the number of model parameters to favor parsimony [182]. The three competing models were: i) Pooled—assuming all participant JNDs arose from a single posterior distribution, and which therefore does not allow for inter-participant differences (this model thus has a single constant WF); ii) Independent—assuming each participant has a single independent JND (and a constant WF), which allows for inter-participant differences in JND posterior distributions; and iii) Variable WF—assuming each participant can have two JNDs. The first JND was calculated using the metabolic data that corresponded to absolute reference costs in the lower half of that participant's reference cost magnitudes, and the second JND was calculated using the metabolic data in the upper half of reference magnitudes. Thus, this model featured a non-constant WF in which the JND

varies based on the absolute magnitude of the reference cost.

In each model, our parameter estimates were informed by the data, yielding posterior distributions over all possible parameter estimates. In the pooled model, the single JND estimate predicted the responses of all participants; in the independent model, each participant's responses were predicted by individual JND distributions which we estimated; and in the variable-WF model, each participant had two different JNDs for references above or below the average reference value. The pooled model featured 2000 tuning iterations and 2000 posterior predictive samples for reduced computational time, given the greater number of responses used as input. The remaining models used our default settings of 8000 tuning samples and 4000 posterior predictive samples due to the relative sparsity of the data and the complexity of the models.

2.5 Results

The average inter-participant JND was 22.7% (standard deviation (SD): 17.0%) (Fig. 2.7A). Many participants were highly attuned to the changes in their metabolic energetics, while others were less perceptive, as evidenced by the high standard deviation of the estimates (Fig. 2.7B). We used a one-sample Kolmogorov-Smirnov test in MATLAB to verify that the independent JND estimates from the participants were normally distributed. The lowest estimated JND was 6.4%, while the highest was 69.9%. The standard error of the mean (SE) was 5.35%. Our confidence in each participant's JND estimate was given by their respective JND posterior distributions, which represent the distribution over potential JNDs of each participant (Fig. 2.7C). Differences in the JNDs can then be observed by comparing the shapes of these distributions; for example, a narrow distribution with a defined peak at a low value represents a participant who is highly attuned to changes in energetics, while a flattened distribution with a peak at a high value denotes a participant with a greater JND and less sensitivity.

We used the WAIC metric [182] to identify the psychophysical model that best describes our data (Fig. 2.7D). This metric evaluates the predictive power of each psychometric model [55] and corrects for the number of parameters to favor parsimony. The best model was the independent-JND model with a constant WF (described in *ii* above in Statistics and Comparisons), which obtains the highest WAIC score and is outside the standard error regions of both competing models.

We found no significant relationships in our exploratory analysis of fitness correlates. We regressed lines of best fit to identify any relations between the magnitude of the independent JNDs and the fitness variables (Fig. 2.9). We conducted F-tests to detect significance of the slope ($\alpha = 0.05$). No significant relationship was found between the fitness variables and the magnitude of the estimated JNDs.

We examined the range of absolute metabolic rates experienced by participants in our protocol

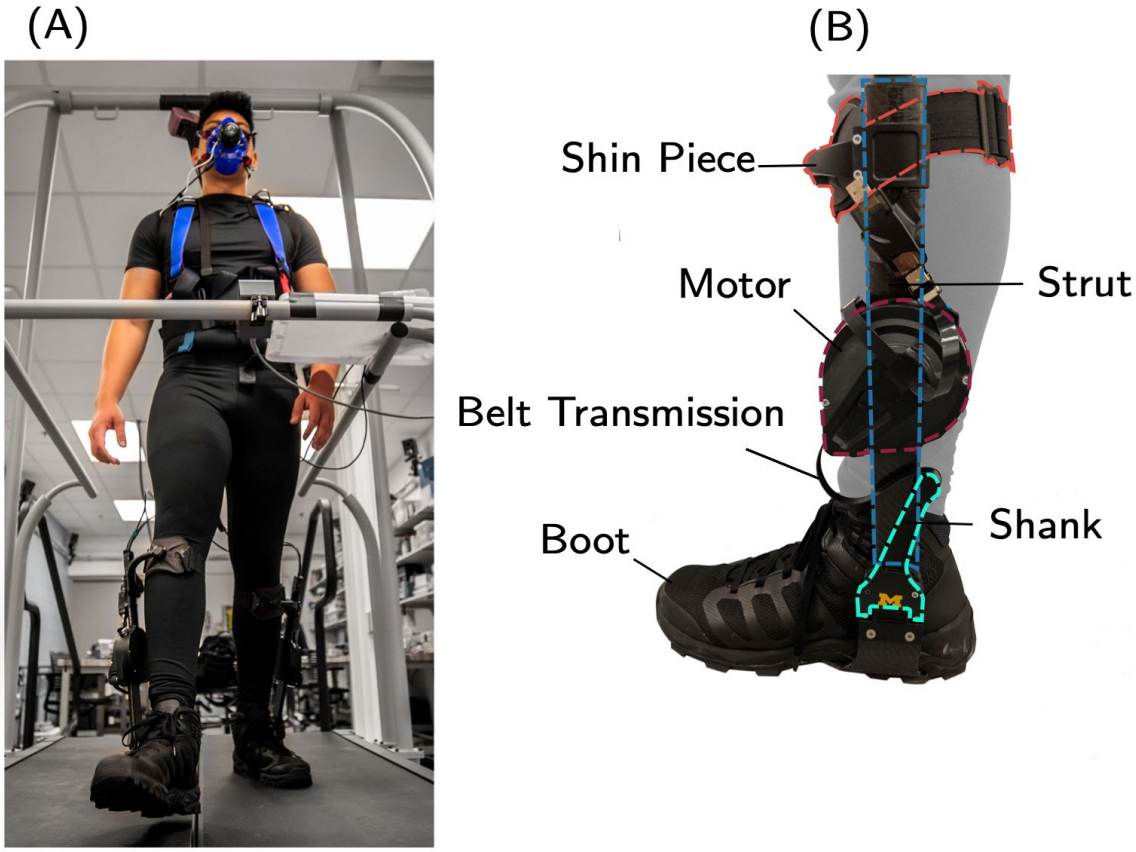

Figure 2.6: (A) The exoskeleton-human system (picture taken prior to the COVID-19 pandemic). Participants walked on a treadmill and experienced different changes to their metabolic rates, which were measured using indirect calorimetry. (B) The Dephy ExoBoot ankle exoskeleton used in the physical experiment. A brushless DC motor mounted on a rigid shank assists the user by generating torque through a belt drive transmission that applies force on a boot-mounted strut. The exoskeleton is securely attached to the user via a shank attachment that transmits the actuator's torque to the leg.

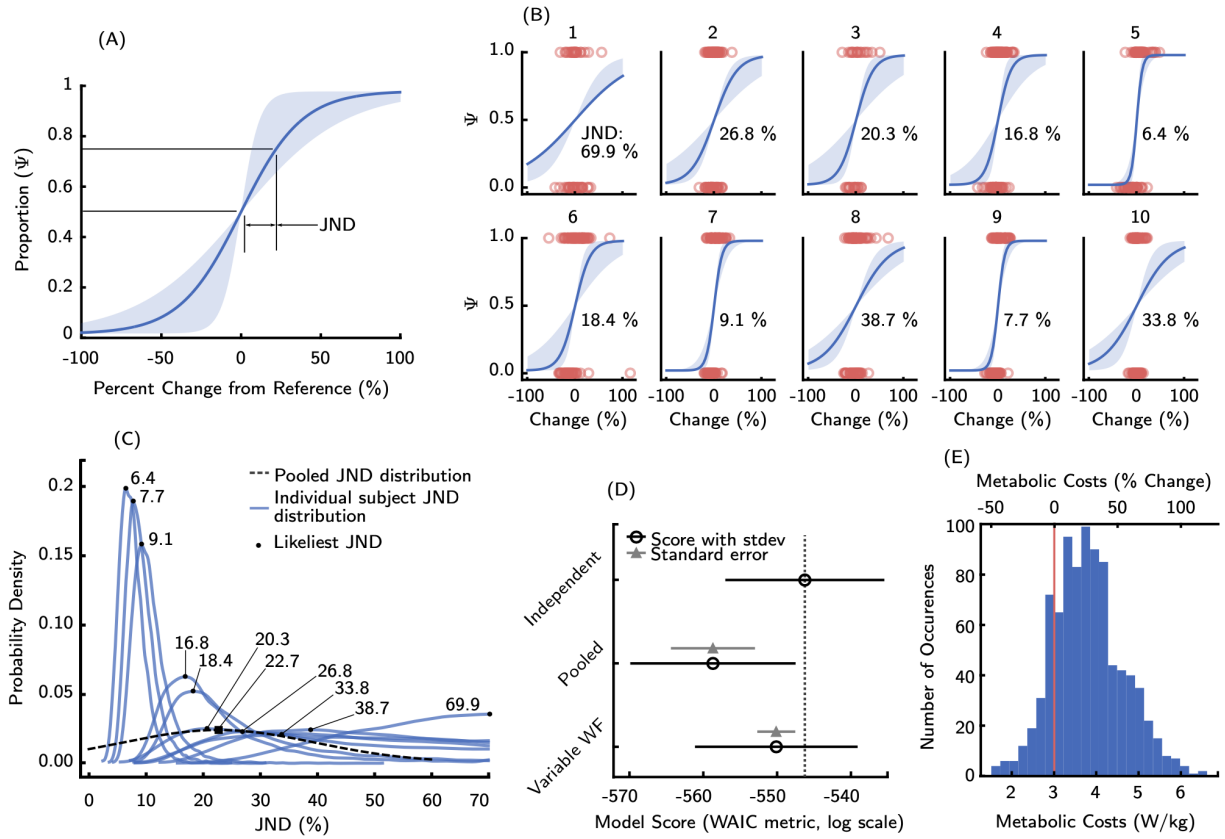

Figure 2.7: (A) The psychophysical curve corresponding to the inter-participant average (solid blue, 22.7%) with one inter-participant standard deviation (shaded, $\pm$ 17.0%). (B) Participant-specific data: the likeliest psychophysical curve for each participant (solid blue), participant responses (red circles), and the 95% credible interval of possible curves from the posterior distribution (shaded blue). (C) The posterior distribution for the inter-participant psychophysical curve (dashed black) vs. the posterior distributions for each participant with modeled inter-participant JND differences (blue). The inter-participant model posterior distributions show clear differences between participants and thus proved a better choice of model. (D) A comparison of different JND models using the Watanabe-Akaike Information Criterion (WAIC) metric. A higher WAIC score (black circle, standard deviations given by black lines) indicates a better model. The best model has a light gray dotted line through its empty circle to aid in comparison. Grey triangles indicate the difference in WAIC between that model and the top model (standard error given by grey bars). (E) The absolute range of reference costs aggregated across all participants. The vertical red line denotes the average net cost of walking at 1.25 m/s across different studies [38].

and verified that the metabolic rates humans experience while walking with an assistive exoskeleton were included in this range (∼10% reductions from unassisted walking, see Fig. 2.7E). As the WF is modeled as locally constant over this range, the JNDs estimated in this work pertain to the perception of energetics in both augmentative and resistive assistance regimes.

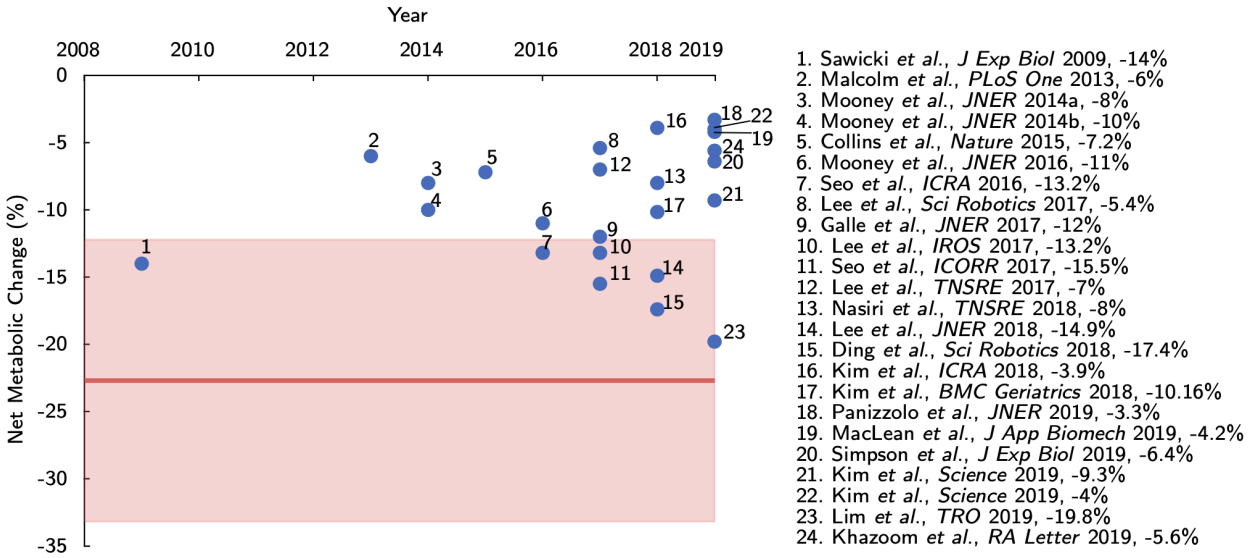

Figure 2.8: The average JND magnitude (solid red) plotted against the state-of-the-art in exoskeleton-driven metabolic rate reductions. The 95% confidence interval for the standard error of the mean of 5.35% is shown (shaded red). As of the date of writing, no published device (to the authors' knowledge) exists that would cause a perceivable short-term energetic benefit to the average wearer. Data reproduced with permission from [150].

2.6 Discussion

On average, modern augmentative exoskeletons do not yet provide a metabolic benefit sufficient to exceed the perceptual threshold of human energetics. We demonstrated that the average Just Noticeable Difference (JND) of metabolic rate was $22.7\% \pm 5.35\%$ (SE). This is substantially greater than the typical reductions obtained using state-of-the-art exoskeletons over the past decade (see Fig. 2.8) [150]. While some studies have shown metabolic reductions greater than 15% [99, 42, 105], most research has demonstrated more modest reductions. The mean reduction in metabolic rate over the past decade is ∼9.6% ± 4.5% (SD) (averaged from studies in Fig. 2.8). Based on the inter-participant psychophysical curve obtained in this work, there is a 61% likelihood an average user would perceive a 9.6% change in metabolic rate, when compared to walking without an exoskeleton (50% accuracy would be a random guess). Thus, based on the metabolic rate reductions provided to date [150], the typical benefit from these devices is unlikely to yet be

a critical factor in the short-term, conscious perception of exoskeleton use. The perceptual threshold presented in this work (*i.e.* the JND) can act as a useful benchmark for future exoskeletons designed to noticeably improve walking energetic efficiency.

For augmentative exoskeletons to demonstrate value to their wearers, the benefits provided should be perceivable in the short term. Given that the state-of-the-art exoskeletons do not yet provide a metabolic benefit that can be sensed reliably (*i.e.* metabolic benefit $> 22.7\%$), this reduction in metabolic rate is not likely to be a driving factor in short-term decision making. Given the short-term nature of the trials in this study, it is possible the reduction in metabolic rate is more perceivable over an extended period of use. While this could positively impact user experience, perception over a longer duration may also lead to challenges in experience and adoption. Prior work in economics has demonstrated that a benefit provided in the future is less valuable when compared to a more immediate benefit (*i.e.* temporal discounting) [199, 63, 64, 162, 35, 153, 17]. Thus, we believe exoskeletons will be most successful if the metrics used to develop these technologies are aligned with what is perceivable and valuable to the user in the short term. Understanding if and how longer-term energetic reductions are perceivable, in addition to the impact of temporal discounting, are important avenues of future study.

Metabolic rate reduction is currently the "gold standard" for augmentative exoskeletons, which is supported by its role in the reduction of joint mechanical power, previous biomechanical studies, and its objective measurability. However, our results demonstrate that the current reductions in metabolic rate are not yet broadly perceptible in the short term. The difficulty of perceiving changes to metabolic rate motivates the consideration of alternative metrics which may be more clearly perceivable by users, including reduction of muscle fatigue [112, 3, 36], peak joint forces in arthritic joints [3, 82, 115], and user preference [172, 84, 159, 179]. The development of perceivable and meaningful metrics to quantify success in future exoskeletons is an important challenge for the field.

Previous work investigating the perception of exertion has shown lower thresholds for exertion during exercise cycling [69]. Haile *et al.* applied the method of adjustment [94] to cycling intensity, arriving at a threshold of 0.15 L/min $\dot{VO}_2$, but did not provide a resting metabolic rate for their participants. Using the resting rate estimates provided in [39], this equates to a JND of $\sim$10%. There are several possible explanations for the differences from our results. For example, the method of adjustment can lead to lower JND estimates [191], and is known to be less reliable [167, 26, 135] than forced-choice experiments. Additionally, the exertion levels tested were substantially greater than what was tested in our experiment, and thus might have occurred in the perceptual regime where the WF was non-constant. Lastly, their method of moderating exertion used only a single variable (cycling resistance), which is susceptible to confounding factors. That is, cycling resistance will vary proportionally with muscle loading, which can be more easily sensed through the

Golgi tendon organs, mechanoreceptors, and other mechanisms. Thus, any perceptual thresholds calculated using these sensations could be confounded to underestimate the true JND of exertion because the participants could intuit a mapping between the easier-to-perceive cycling resistance and the harder-to-perceive metabolic effort.

Researchers have established that humans will seek energetically optimal gaits, even when metabolic rate changes are far below our estimate of the perceptual threshold for metabolic rate (*i.e.* $\sim$5% rather than 22.7%) [155, 1]. One potential explanation for these differences is that our experiment measures *conscious* perception of changes in metabolic rate, whereas this prior work has allowed for potential subconscious sensing contributions from sensorimotor system and autonomic nervous system [194, 156]. The literature suggests that humans rely on a combination of different afferent signals, such as heart rate or muscular strain, to generate a gestalt perception of exertion in ways that are not yet fully understood [70, 59, 152, 16, 14]. It is also not yet known whether the observed changes in locomotor mechanics that are correlated with metabolic rate are causally linked to those changes.

Participants varied greatly in their ability to perceive changes to their metabolic rate. In this study we investigated whether the JND was more appropriately modeled as a constant value or a person-specific value. Using the Watanabe-Akaike Information Criterion (WAIC) metric—a modern Bayesian tool for comparing the quality of models—we found that the data were better fit by the model where each participant had their own independent JND (see Fig. 2.7.D). While the inter-participant mean JND value was greater than the metabolic benefits provided by modern exoskeletons, our participant pool included three participants who had JNDs below 9.6% and thus would likely perceive benefits from these technologies [150] (see Fig. 2.7.C).

Although our study was not powered to identify significant relationships between fitness and perception, we investigated this potential link by regressing linear models between the estimated JND and relevant fitness variables (hours weekly spent performing endurance activities, years spend performing athletic activities, and BMI). We failed to detect any significant relationships between the fitness variables and the magnitude of the estimated JNDs. Future work is needed to study both the physiological mechanisms that underlie variability in perception of energetics, as well as discover methods to identify those users who may have better perception.

In our experiment, the WF—the ratio of the JND (in absolute units of W/kg) to its corresponding reference—is constant with respect to reference magnitude over the range of references tested. The Variable WF model we evaluated, in which each subject's JND was dependent on the absolute magnitude of the reference, was inferior to a model in which each subject had a single JND and constant WF. Given the inter-participant range of reference costs experienced (from $\sim$1.5 to 6.6 W/kg, see Fig. 3E), we believe our experiment occurred in the mid-section of the perceptual regime, rather than the extremes around which the WF will rise steeply in many cases [127].

This distribution of reference metabolic rates does skew towards relative increases away from the metabolic rates experienced during baseline unassisted walking, as it is far easier to add a metabolic penalty with exoskeleton use. However, in many psychophysical studies, the psychometric curve is assumed to be symmetric about the origin [53, 137, 189, 177, 102]—that is, subject perception does not depend on the sign of the comparison. Due to this symmetry and the success of the constant-WF model, we anticipate that the estimated JND pertains to both metabolic benefits and metabolic penalties. Consequently, our results may not only apply to the benefits from modern exoskeletons, but also extend to applications where metabolic penalties may be more relevant (*e.g.* athletic or endurance training).

2.6.1 Limitations

The posterior distributions for those participants with low and high JNDs were differently shaped, reflecting a limit on the maximum metabolic rate changes possible via exoskeleton assistance. The exoskeleton used in this experiment was capable of providing a peak torque of approximately 30 Nm ($\sim$ 10 J per stride), which limited the available metabolic rates that could be experienced. The ability to induce a wide array of metabolic rates is important for sampling the psychometric function. To obtain estimates of these functions that have low uncertainty, they must be sampled across both the constant and transitory regions of the psychometric curve [94]. The quality of the measurements is reflected in the posterior distributions for the JND estimates, with high quality measurements resulting in narrow posterior distributions. For participants with smaller JNDs, the limitation on available metabolic rates enabled the sampling of the majority of the relevant areas of the psychometric curve. This allowed us to exclude both excessively large and small estimates for those participants. In contrast, for participants with high JNDs, the imposed energetics spanned a comparatively narrower region of the psychometric function, which only excluded *lower* JNDs. The posterior distributions for participants with high JNDs was asymmetric, and thus contained greater uncertainty in the upper bound of the threshold. Consequently, any error would likely bias the true JND to be greater than what was measured in this study.

The indirect nature of manipulating energetics via an exoskeleton increased variability in each participant's JND distribution. In conventional psychophysical studies, researchers have more deterministic control over the applied stimulus under investigation. While exoskeletons are known to influence energetics indirectly through several controllable [95, 110, 85, 52] aspects of the torque profile, metabolic rate also depends on many uncontrollable factors that appear noise-like [39, 71, 66]. This added noise results in sub-optimal sampling of the psychophysical curve that reduces certainty in the corresponding JND estimates [114]. This uncertainty is reflected in the width of the posterior distributions of each participant.

The uncertainty of our results also stems from an experimental limitation in how many trials are feasible. Conventional best practice in the psychophysics literature would recommend $\sim$300 trials [53] when estimating the underlying psychophysical curve; however, in this study we were able to obtain $\sim$100 trials for each participant. The relatively low number of trials was due to the time necessary to obtain responses. In this protocol, participants experienced different metabolic rates in sequence, each of which requires two minutes to estimate the participant's metabolic rate. To obtain the necessary data for this experiment, participants walked during three sessions spread across three days, with each session lasting four hours. This is in contrast to many studies of human perception, which can obtain experimental data without the time delay of the human cardiopulminary system (τ_r = 42 s [201]). Consequently, the uncertainty of our estimates was increased by approximately 60% [114] due to the lower number of samples, which is reflected in the inter-participant distribution of the JNDs.

We found that despite the uncertainty in JND estimates, these estimates were relatively insensitive to assumptions in our approach. We used a uniform prior distribution in our analysis that encompassed available JNDs between 0% and 70%. We investigated the sensitivity of our results to the bounds of this prior distribution (*i.e.* 0% and 70%). We chose our lower bound to reflect perfect human perception, while the upper bound was informed by the reasonable assumption that a human could consistently detect changes in energetics just under those that result from switching from walking to running (a $\sim$ 100% change [171]). When the bounds of our uniform prior distribution were changed to [0%, 60%] and [0%, 80%], we found that the average inter-participant JND estimate shifted from 22.7% to 21.9% and 23.2%, respectively. These small shifts in the mean JND estimate indicate that our approach is robust to the exact shape of our prior distributions, and are thus well-informed by our sampled data.

Participants responded to questions about exertion, but we are unable to know what specifically drove their answers. Our study relies on participants honestly reporting perceived exertion and not confounding this report with other perceptions, which could include perceptions of assistive torque and assistance timing, as well as higher-level perceptions of the helpfulness of the actuation profile. Our study was designed to mitigate these confounding factors. Participants read a predefined script to help elucidate the concepts of metabolic rate and exertion. The prompt was designed using vocabulary consistent with the Borg Scale, used to assess exertion [13, 14, 152]. Additionally, the torque profile was designed to be intentionally complex (see Methods). That is, the participant's metabolic rate was induced by the complex interaction of three controller parameters, obscuring any foreseeable relationship with metabolic rate (*i.e.* "it feels more powerful when it is stronger, which must lower my exertion"). However, if participant's JNDs were affected by additional informative sources, this would also bias the true JND of metabolic perception to be greater than what was estimated.

2.7 Conclusion

Motivated by the need to develop augmentative exoskeletons that can realize their potential to impact society, we quantified the human ability to perceive the metabolic impact of these technologies. Participants were able to perceive an average change of 22.7% $\pm$ 5.35% (SE) in their metabolic rate with 75% accuracy in short-term testing. Thus, the average user cannot yet consciously perceive the metabolic benefits from the typical modern exoskeleton, which may hinder translation and adoption of these technologies. Our results provide a new benchmark for augmentative exoskeletons that will enable perceivable value to their users. The relatively insensitive perception of metabolic rate also suggests that alternative metrics for exoskeleton success, such as reduced muscle fatigue, loading, or user preference, may be more significant to user experience and exoskeleton success.

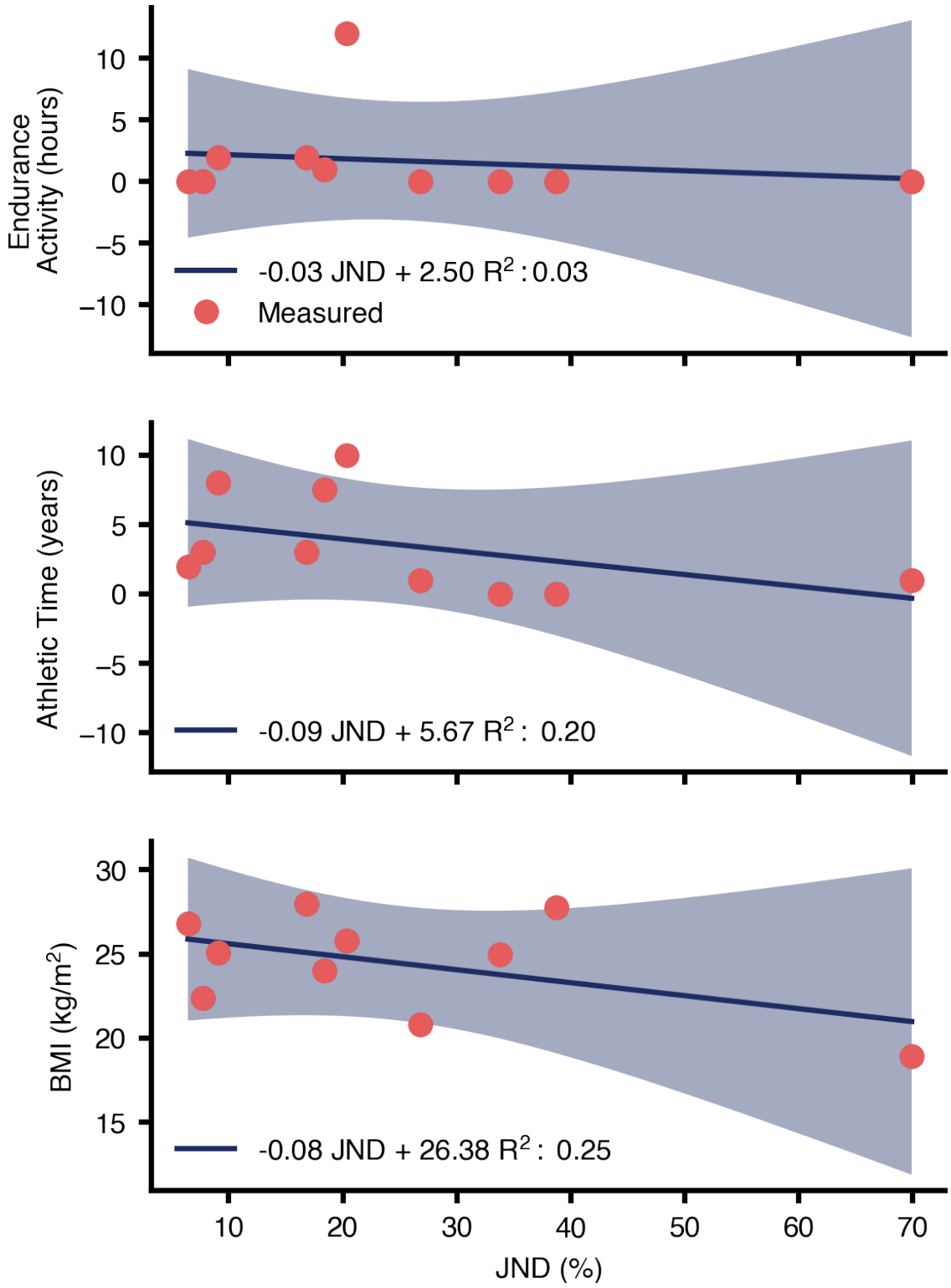

Figure 2.9: Fitness variables shown across estimated JNDs: Hours spend performing endurance activity *(e.g.* running exercise), years spent performing athletic activity, and Body Mass Index (BMI). The regressed best-fit lines are shown in blue, with individual responses in red. The 95% confidence intervals are shown by the shaded region. No relationships were statistically significant.

CHAPTER 3

The Economic Value of Augmentative Exoskeleton Assistance

3.1 Context

Chapter 3 of this dissertation delves into the economic value of exoskeleton assistance for the first time. The objective is to determine the actual value of exoskeletons using economic value as a viable alternative which may be more conducive to the adoption of exoskeletons outside the laboratory when compared to the current standard of metabolic cost reduction. The economic value metric is both more accessible, as it does not require specialized equipment, and more intuitive, as users can better comprehend benefits expressed in terms of monetary currency compared to biomechanical quantities, such as calories burned. This study employs the Vickrey second-price auction within economics to measure the participants' valuations of their walking times across various walking conditions, including walking with and without exoskeleton assistance from bilateral ankle exoskeletons and walking with and without the exoskeleton itself. This chapter is currently under submission to *Nature Communications Engineering*.

3.2 Introduction

Powered lower-limb exoskeletons have the potential to transform human mobility by extending the locomotor abilities of their wearers. These technologies augment lower-limb function by providing mechanical assistance to the joints of the legs in tandem with the human neuromotor system, and thus can make physically demanding tasks less challenging for both able-bodied and impaired individuals. Assistance provided by modern exoskeletons have been shown to reduce the caloric demands and muscular effort required for walking [123, 122, 132, 149, 110, 202, 21, 151, 49]. Consequently, exoskeletons may have beneficial implications for recreational users and workers in factory, military, or supply chain environments. In addition, exoskeletons can reduce muscle

activation, and thus may reduce joint loading [101, 103, 204, 115], potentially extending the physical capabilities of aging individuals. Rehabilitation-focused exoskeletons may also restore the mobility of people with neuromotor deficits who face weakness and impairments in balance, coordination, and joint mechanics following upper motor neuron disease (*e.g.* stroke). Exoskeletons can be used to assist the gait of these individuals, with several commercially-available lower-limb exoskeletons having been approved by the U.S. Food and Drug Administration.

The metrics by which we assess exoskeletons drive their design, control, and potential impact. For rehabilitative applications, the design of these technologies has a clear physiological objective: the restoration of impaired gait function. These exoskeletons are often considered successful if they restore able-bodied kinematics or kinetics, which provide clear physiological goals on which to base exoskeleton design and control decisions [205, 204, 164, 27, 151]. However, for augmentative applications in which the user is typically able-bodied, the metrics of success are less clear. Currently, augmentative exoskeletons are developed based on their ability to meet a physiological objective; the "gold standard" for exoskeleton success is the reduction of metabolic expenditure during locomotion (*i.e.* a reduction of the calories burned) [150]. This objective is both intuitively meaningful and objectively measurable [195, 154]. Modern exoskeletons have reduced the metabolic rate relative to unassisted walking by an average of approximately 14% [123, 122, 132, 149, 110, 85, 202]. This objective has led to the rise of promising 'human-in-the-loop' (HILO) optimization techniques, which directly modulate exoskeleton assistance based on the metabolic reductions experienced by the wearer [95, 201, 42, 186, 104, 163, 61]. Recent work has established experimental infrastructure that illustrates the tight coupling between metrics of success and exoskeleton development. An example of this infrastructure includes tethered emulator systems [41, 21], whose purpose is to inform exoskeleton design and control based on their ability to reduce the metabolic expenditure of their wearer. Numerous studies investigating the biomechanical underpinnings of metabolic cost reductions have also been conducted to find more optimal exoskeleton assistance settings [120, 193, 201, 85]. Finally, other metrics include net-joint torque reduction [103] and muscle activation reduction [101, 204], which are commonly used, easier-to-measure proxies for improvement in energetics.

Though the physiological benefits of exoskeletons have been demonstrated using metabolic metrics, these benefits have not widely translated to wearer perception of enhanced endurance and strength. These perceptions are important, as for augmentative exoskeletons to reach their potential in society, users will need to voluntarily accept these technologies into their lives. They thus must be developed to provide a perceivable benefit to their wearer, in addition to objective assessment of their impact. Our recent work has shown that during exoskeleton-assisted walking, the average user cannot yet perceive the benefit of most systems available today [117, 114, 113]. That is, the metabolic rate needed to be reduced by 23% (N = 10) before exoskeleton users could reliably per-

ceive this improvement (whereas most modern exoskeletons reduce the wearer's metabolic rate by 14% compared to unassisted walking [150]). These results agree with prior studies that showed humans were relatively insensitive to small changes in exertion in other exercise contexts [69]. Intuitively, if the user is unable to perceive the metabolic reduction provided by an exoskeleton, this value may be difficult to incorporate into decision-making during exoskeleton design, translation, and adoption. Consequently, assessing and developing exoskeletons based on reductions in metabolic rate could result in systems that are not perceived as valuable by users, despite significant energetic benefits [193, 105, 42, 121, 99, 132].

An alternative method for measuring success in exoskeleton development is to quantify the perceived *economic value* provided to the wearer during its use. Economic value, measured in monetary currency (*e.g.* US dollars), is assigned by the wearer and can reflect the multifaceted nature of exoskeleton user experience. Although exoskeletons can provide assistance that improves energetics, that assistance often comes at a cost to the wearer. Exoskeletons can add discomfort, weight, and audible noise, in addition to having aesthetic implications. While exoskeletons may potentially have universal positive value, the heterogeneity of the metabolic response to exoskeleton assistance, coupled with the known variety of responses to new innovations within the social sciences [144], could also imply a wide range of valuations for wearing an exoskeleton. If the user is able to assign economic value to the experience of exoskeleton use, they are able to inherently balance and quantify these trade offs. Thus, we posit that exoskeletons that maximize economic value may have a greater likelihood for adoption and use. Prior work in management science has established that the perceived value of different technologies has a significant impact on user intent to adopt those technologies into their daily lives [40, 93, 129, 79, 29]. When potential exoskeleton users, manufacturers, and others are weighing the choice to adopt or purchase an exoskeleton, the consciously perceived benefits must outweigh these costs. Thus, the perceived economic value of exoskeleton use is a potentially powerful metric for designing and controlling exoskeletons that quantifies meaningful, individualized benefits to wearers.

In this study, we introduce a tool for measuring the perceived economic value of exoskeleton use as a metric to evaluate their performance and user experience. We define and use this tool to quantify the economic value—termed *Marginal Value (MV)*—provided by bilateral ankle exoskeleton use during uphill walking. We leveraged the Vickrey second-price auction to measure participants' "price to walk," which was then aggregated and compared across conditions to obtain our economic value metric (*i.e.* MV). Our experiment revealed that while there was insignificant positive value of exoskeleton use across subjects, there was a large disparity between subjects. Some subjects reported substantial value provided by the exoskeletons, which represents an opportunity target these "responders" in future work. The near-zero net MV from the exoskeleton stems from two competing effects: the MV of the powered assistance alone was substantially positive,

but it was counteracted by the cost from wearing the device itself. The use of MV offers advantages over the more common metabolic rate metric, including its accessibility, in that it does not require specialized equipment, and its intuitiveness, as users and manufacturers are more likely to understand the value of monetary currency over biomechanical quantities (*e.g.* calories burned or muscle power). Our approach is also generalizable, and can be used to measure the value of not only different types of exoskeleton assistance, but also various technologies, activities, or experimental conditions.

Within the field of economics, the Vickrey second-price auction [183] provides a well-studied incentive structure for quantifying the economic value of abstract concepts. In a seller's Vickrey auction, participants compete via bidding to sell a good. The winner of the auction is the participant that bids the lowest amount to sell the good. However, the winner will earn the amount bid by the second-lowest bidder. The Vickrey auction is also "sealed-bid" in that each participant's bids are not revealed publicly. This second-price paradigm incentivizes subjects to bid their true perceived value for the good [183, 134, 107] rather than attempt to guess (and slightly underbid) all other bidders' lowest bids. Auction participants would be discouraged from bidding higher than what they believe the good is truly worth, for fear of losing the auction to a competitor with a lower bid, and the second-price nature breaks the link between the auction winner and their actual bid. Conversely, they will have no incentive to underbid below their true value since they would risk selling for a "loss" if they won the auction. Thus, the Vickrey auction provides a quantitative measurement for the value of arbitrary goods or services by measuring an individual's willingness to buy or sell [107, 181]. The specific incentive structure is thus less prone to biases when quantifying the value of a good when compared to alternative methods such as direct feedback [165, 44]. Prior researchers have also used Vickrey auction metrics to measure the value of abstract concepts or actions, such as food safety [50], GMO-free foods [145], the stigma resulting from HIV [74], personally identifiable information [165], and smartphone battery life [77]. In particular, Coursey *et al.* employed the Vickrey auction to quantify the willingness of participants to endure performing an unpleasant task, such as tasting a bitter liquid [34]. These examples highlight the promise of the Vickrey auction to quantify the perceived economic value provided by experiences, including exoskeleton use.

3.3 Methods

3.3.1 Participants

In this study, sixteen able-bodied participants (N = 16, 5 female, 9 male; age = 26.3 $\pm$ 4.61 years; mass = 77.1 $\pm$ 13.4 kg, Table. A.1) walked using bilateral ankle exoskeletons on a treadmill.

All participants provided written informed consent before participation. The study protocol was approved and overseen by the Institutional Review Board of the University of Michigan Medical School. Participants had no prior experience walking with the bilateral ankle exoskeletons featured in this study.

3.3.2 Experimental Protocol

3.3.2.1 Exoskeleton Apparatus

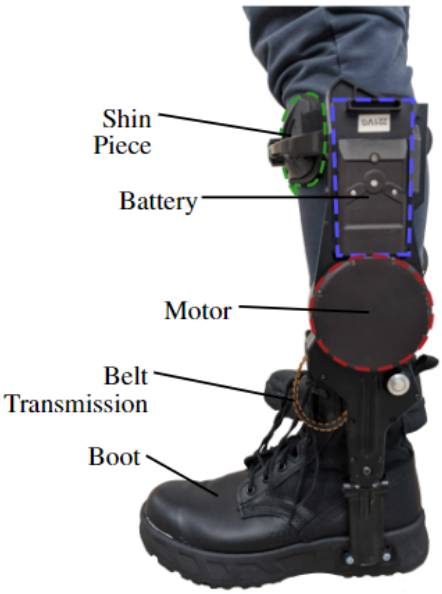

Figure 3.1: The Dephy ExoBoot used in this experiment. An electric motor applies plantarflexion torques at the ankle via a belt transmission.

Our approach quantifies the value of bilateral ankle exoskeletons that were designed to improve the energetic efficiency of human walking (ExoBoot, Dephy Inc. Maynard MA, Fig. 3.1). The commercially-available system utilizes an onboard brushless electric motor and flat cable transmission (for a mean transmission ratio $\sim$15:1, see Supplementary Fig. 3.3.) to apply ankle assistance during walking. The exoskeletons have a single powered degree of freedom (dorsi-plantar flexion) and a passive, unactuated degree of freedom (inversion-eversion). The transmission is unidirectional, which enables the system to apply plantar flexion assistance torque but it cannot provide dorsiflexion assistance. Each side of the exoskeleton applies a torque profile (Fig. 3.2a) that provides a burst of positive power (Fig. 3.2b) during the terminal stance phase of gait, augmenting the propulsive effort provided by the triceps surae. The average energy provided by the exoskeleton during the gait cycle is $13.4 \pm 2.9J$. During swing phase, the exoskeleton is able to add slack to

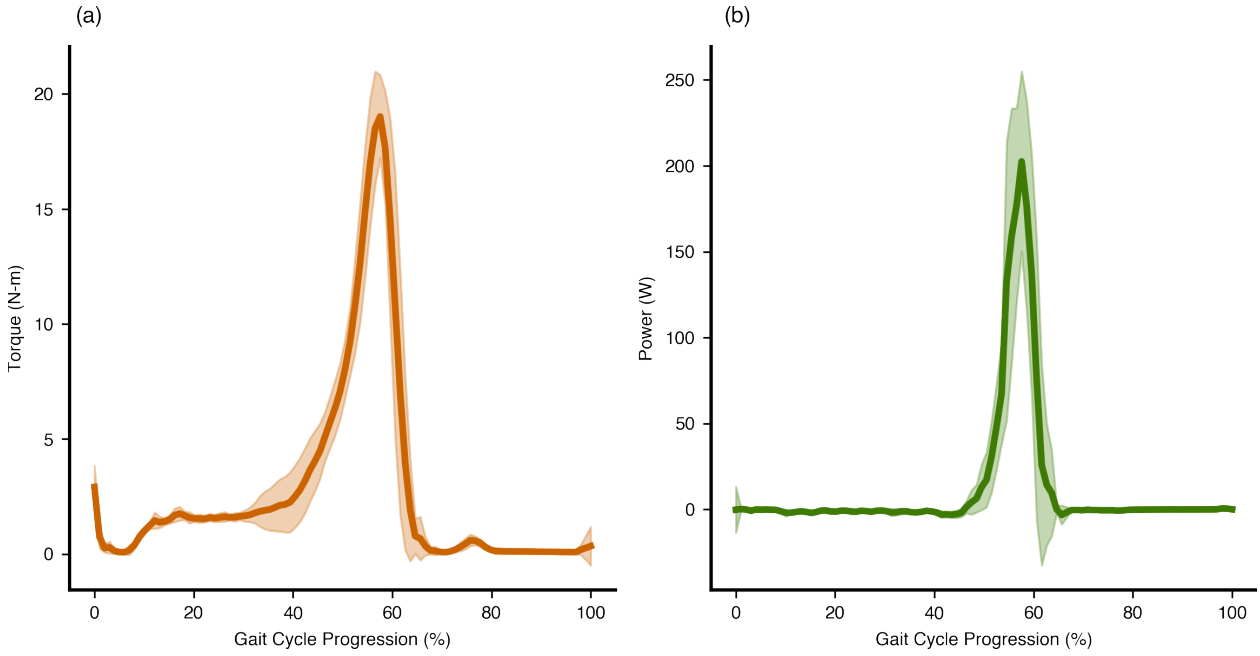

Figure 3.2: (a) The torque profile applied by the bilateral ankle exoskeletons during the uphill walking task. The torque applied stems from the proprietary control strategy implemented by Dephy Inc and represents the state of the art. The mean profile is shown in solid orange, with a single standard deviation shown by the shaded region. Torque was quantified by using the measured q-axis motor current and the experimentally-derived, q-axis torque constant to calculate motor torque [100], and then multiplying that value by the instantaneous transmission ratio. Gait cycle progression is defined as beginning and ending at sequential ipsilateral heel-strikes. (b) The power applied by the exoskeletons at the ankle during the walking trial. Power was calculated by multiplying the current-derived torque profiles in (a) by measured ankle angular velocities. The mean profile is shown in solid green, with a single standard deviation shown by the shaded region. The average energy provided by the exoskeleton, obtained by integrating the power curves over time is 13.4 ± 2.9 J.

the belt drive, thereby preventing any unwanted resistance to the foot; this capability stems from the unidirectional nature of the design. Gait progression, inferred from heel-strike timing events, was used to schedule how assistance was provided during each step. Similar exoskeletons have been shown to lower the user's metabolic expenditure during walking as well as reduce the biomechanical power requirements at the ankle and other joints of the legs [121, 123, 120]. The walking assistance controller was developed by Dephy Inc. as part of their commercial system. As one of the first commercially available exoskeletons, the Dephy Exoboot was chosen for its ease of use, robustness, and representation of the state-of-the-art.

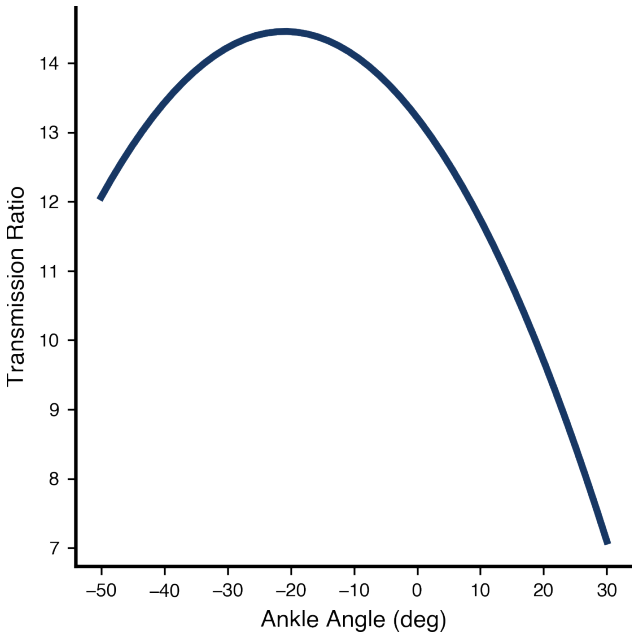

Figure 3.3: The transmission ratio curve of the exoskeleton device used in this experiment. Positive ankle angles denote dorsiflexion and negative angles denote plantarflexion. The curve was modeled as a second-order polynomial, and was obtained by taking the derivative of the best-fit third order polynomial that relates motor angle to ankle angle.

3.3.2.2 Vickrey Auction Protocol

The Vickrey auction [183] is a powerful economic tool for determining the true value placed on goods or actions. In this type of auction, participants compete to purchase (or sell) a good or item. For each participant, the auction's structure is designed such that the optimal strategy for obtaining the item is to truthfully represent their internal value with their bid (*e.g.* to bid an amount equal to the true worth of the item). This optimality stems from the *second-price* nature of the auction [183, 134, 107], in which the winner is the participant who bids the highest (or lowest, as in the selling implementation used in this study). However, rather than paying the highest bid, the winner of the auction instead pays the *second-highest* bid. Participants do not know the value of competing bids (sealed-bid). In theory, bidders are disincentivized to bid less than their true value, as they run the risk of not winning the auction (not acquiring the good). Similarly, participants should not bid more than their true value, which otherwise could cause them to pay more than the value of the item. The second-price nature breaks the link between the auction winner and their specific bid. The inverse is also true for the case of selling an item (second-lowest bid is paid, lowest bid wins); thus, the incentive structure that elicits truthful bidding still holds in the seller's auction. Due to the presence of this optimal strategy, the Vickrey auction provides a method for quantifying the value of abstract concepts [107, 181]. In our protocol, we use the Vickrey auction sequentially

to repeatedly sample individuals' valuations of their time during uphill walking. The participants bids across the sequential auctions denotes the wearer's "price-to-walk" curve.

3.3.2.3 Walking Protocol

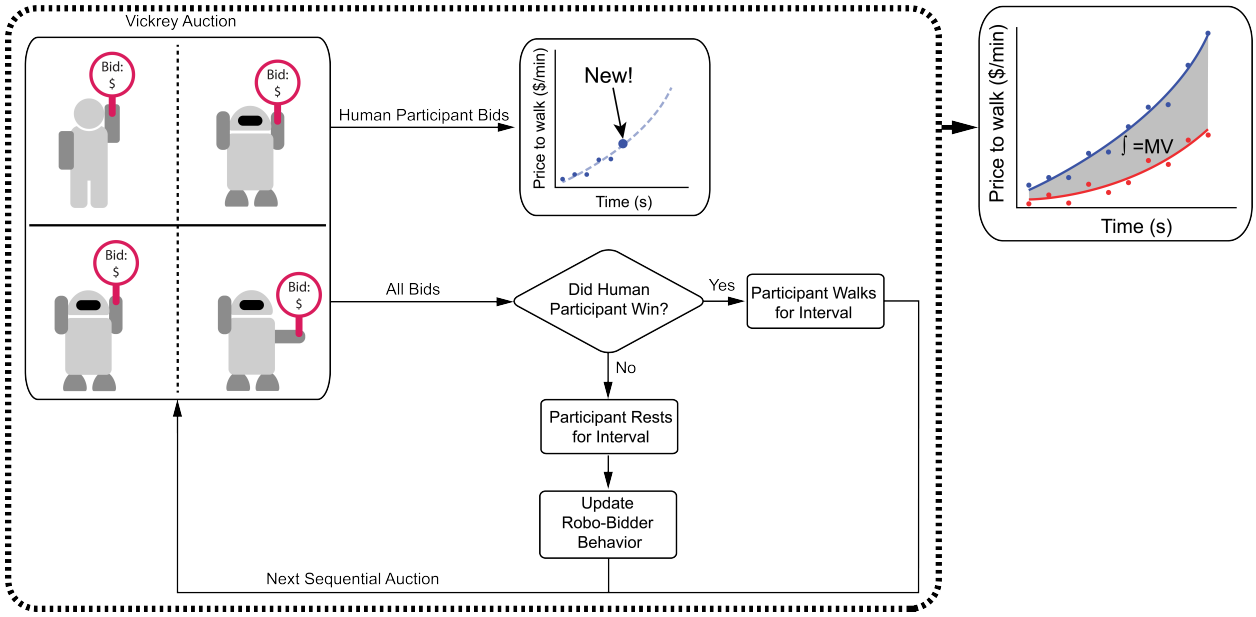

Figure 3.4: The experimental Vickrey auction protocol. In this protocol, human participants compete against computerized bidders (depicted in the 2×2 grid) in sequential Vickrey auctions for the prize of monetary compensation in exchange for undergoing a strenuous locomotion task. A human participant, shown in the upper-left quadrant, competes against computerized ('robo-') bidders, represented by the robots in the remaining three quadrants. Every two minutes, the human participants verbalize their bids to their experimenter, which are recorded and compared to the bids from the computerized opponents; the lowest bid is the winner of the auction, and the payout is the second-lowest bid. The second-price nature of the auction elicits truthful bids/prices from the participants. Should the human win, they accrue the payout and walk for the two-minute interval; should they lose, they instead miss out on the payout and rest until the next interval. Regardless of whether the human won or suffered a loss, their bid is aggregated to form their price to walk curve. This process repeats for roughly an hour. As the human walks more, they become more fatigued, causing their price to walk to increase; a trend well-modeled by a first-order exponential according to the observed data. Among other conditions, participants competed while wearing normal shoes ('no-exo' condition) and while receiving powered assistance from the exoskeleton ('exo-powered'). The integral of the difference between the price-to-walk curves for the no-exo and exo-powered conditions is the value added from the device's assistance.

Participants received guidance on the Vickrey auction protocol prior to the experiment. During this time, participants read a lay explanation of the Vickrey auction. Subsequently, participants underwent two mock Vickrey auctions, where the optimal strategy of truthful bidding was explained.

Subjects were also repeatedly informed of the optimality of the honest-bid strategy to improve the number of honest bids provided [111]. In the first mock auction, participants were presented with a miscellaneous office supply item and told to bid honestly on it as if they were competing to purchase it in a real Vickrey auction. Each participant wrote down their bid on an index card, while the researcher wrote down some artificial competing bids. The participant then revealed their bid, and the experimenter explained which bid won (the highest bid) and what the cost would be (the second-highest bid). Next, participants were told to imagine they were participating in the actual study, in which they would compete to sell their walking time in exchange for monetary compensation. Participants were prompted to consider bidding based on an hourly wage, although they were also told that any truthful bid would be accepted. Again, participants wrote down their bids while the researcher aggregated the competing bids. As in the previous mock auction, the participants revealed their bids, and the experimenter walked them through which bid won and the payout.

Following familiarization with the exoskeleton and auction mechanics, participants competed in sequential Vickrey auctions to sell their time while walking uphill. The protocol was organized into a set of Vickrey auctions experienced in series, each lasting two minutes. If the participant won the auction, they would walk for the two minute interval, and they would rest if they lost the auction. During walking, participants walked on a split-belt treadmill with a 10° incline (Fig. 3.4). Unlike in the mock auctions, the participants would accrue and receive monetary compensation according to the Vickrey auction protocol. At the end of each interval, the winning bid (second-lowest bid) was revealed. The sequential auctions lasted a randomly-specified duration between 50 and 70 minutes, with a mean duration of 60 minutes; participants walked on average a total of 31.3 minutes $\pm$ 11.5 minutes. The duration uncertainty was added to discourage subjects from waiting until their bidding competitors were "exhausted" to inflate payouts. The two-minute interval duration was chosen to minimize subject exploratory behavior. In total, participants experienced approximately 25 - 35 Vickrey auctions in series. By aggregating each subject's bids across the series auctions, the resulting price to walk curve captured the participants' valuation of their time to complete the experiment, which can be used to provide insight into the value provided by the exoskeleton in this application.

Our protocol was broken up across several days to reduce the effect of fatigue from the experiments. The order of the testing was randomized across participants. On one day, subjects completed the Vickrey auction protocol with normal shoes (*i.e.* the walking-no-exo condition); these data establish their baseline price to walk curve and cumulative price. On a different day, subjects instead walked with powered exoskeleton assistance (*i.e.* the exo-powered condition). Participants were not explicitly told that the exoskeleton would provide assistance; instead, they were told that the exoskeleton was applying a randomized torque profile that would either help or hinder their locomotion. This was done to reduce the chance that participants would experience a placebo

effect and perceive the exoskeleton as valuable due to information provided by the experimenter. By comparing the price to walk curves from the walking-no-exo and exo-powered conditions, we obtained a measurement for the value provided by the exoskeleton in this task. If undergoing the exo-powered condition on the second day, subjects were given time to re-familiarize and experience the exoskeleton's assistance.

In addition, each subject was randomly assigned to additional experimental sessions to investigate other attributes our approach. Participants were randomly assigned to groups that either quantified the inter-day variability of the price-to-walk measurements, or investigated the additional condition of walking with the exoskeleton without assistance (*i.e.* the exo-powered-off condition). Due to the number of experimental sessions required, particpants were assigned to only one of these groups. Of the sixteen total participants, ten participants walked in the exo-powered-off condition, while four participants repeated the walking-no-exo condition at least once (the remaining two subjects were not able to participate in the additional sessions). Within the experimental group that repeated the walking-no-exo condition, three participants repeated the condition twice more, while the last participant repeated the condition once. The intent of investigating the exo-powered-off condition was to support that the price-to-walk curves captured the cost of wearing the additional exoskeleton weight, when no assistance was provided. The goal of repeating the walking-no-exo condition was to provide insight into the test re-test variability of the price to walk curves and how they may be affected by inter-day confounding factors.

3.3.2.4 Robo-bidders

For a live Vickrey auction, multiple participants are needed; however, this constraint adds practical challenges in coordination and logistics. To this end, we utilized computerized bidding agents (robo-bidders) as a substitute for other human participants in the Vickrey auctions. The mechanics of the robo-bidders are modeled on the bidding behavior of three pilot subjects. The robo-bidders bid with prices following a first-order exponential function of the number of intervals spent walking. This model captured the effect of fatigue and increased their bids as the robo-bidder won the auction. If the robo-bidder did not win the auction (meaning the human participant won), their bid remained constant. Gaussian white noise (zero mean, standard deviation \$0.01) was added to corrupt the robo-bidder bids, to reduce the likelihood that the human participant would intuit the robo-bidder model. Robo-bidder behavior at time t_k during simulated walking was governed by the following equation:

$$y_k(t) = k \cdot e^{(b \cdot t_k)} + \sigma, \tag{3.1}$$

where y_k was the robo-bidder's price, k was the initial price, b was the rate at which the price increases, is time, and σ was drawn from normal distribution $\mathcal{N}(0, 0.01)$. Parameters k and b were set differently for each robo-bidder; full implementation details can be found in [TBD]. Having robo-bidders instead of human auction participants reduced the logistical difficulty of executing the Vickrey auction experiments over time, while still replicating modeled behavior of a human participant. Naturally, the robo-bidder bids competed in parallel with the human subject and approached an equilibrium behavior, which affected the total walking time of the human participant. In the application of auctions in series, an equilibrium would also likely be established between multiple human participants. The human participants were not initially made aware of the fact that they were competing against computerized opponents. They instead were told that they were competing in live Vickrey auctions against other humans located remotely, with the experimenters communicating live. This was done to prevent the participants from attempting to learn and exploit the price to walk curves of the robo-bidders in an attempt to maximize profits beyond the strategy of honest bidding. Participants were debriefed on the true nature of the robo-bidders at the conclusion of their participation in the experiment.

3.3.3 Analysis

3.3.3.1 Marginal Value

Our outcome metric is the marginal value (MV) that stems from difference in two price to walk curves obtained for different experimental conditions. Each price to walk curve was fit with a first-order exponential response of the form:

$$Y(t) = k \cdot e^{(b \cdot c \cdot t)}, \tag{3.2}$$

where Y is the participant's price, k is the initial price, b is the rate at which the price increases, c is a scaling factor equal to the participant's win-rate, and t is time. The scaling factor is necessary to control for the different durations of walking experienced by the participants (*i.e.* variations in the number of auctions won). After correcting for this win-rate, these curves are integrated and subtracted, yielding the area between the two curves. This area represents the marginal value added—in US dollars—from one exoskeleton condition compared to another. The expression for the MV is as follows:

$$MV = \overbrace{\int_{t_1}^{t_2} k_1 \cdot e^{(b_1 \cdot t)}\, dt}^{\text{exo condition 1}} - \overbrace{\int_{t_1}^{t_2} k_2 \cdot e^{(b_2 \cdot t)}\, dt}^{\text{exo condition 2}}, \tag{3.3}$$

where k_1, b_1 correspond to the first exoskeleton condition in the comparison, k_2, b_2 correspond to

the second condition, and t_1, t_2 are the bounds of the time domain for the integrals—0 and 30 minutes respectively, which roughly corresponds to the average time each participant walked in each trial. The MV is then equivalent to the value added or removed by the exoskeleton during a continuous thirty minute, uphill walking task. We commonly normalized the MV by the cumulative price from the walking-no-exo condition to obtain a percentage change. However, this metric can be converted to compare any two experimental conditions or other candidate control strategies that researchers wish to evaluate in terms of economic value.

3.4 Results

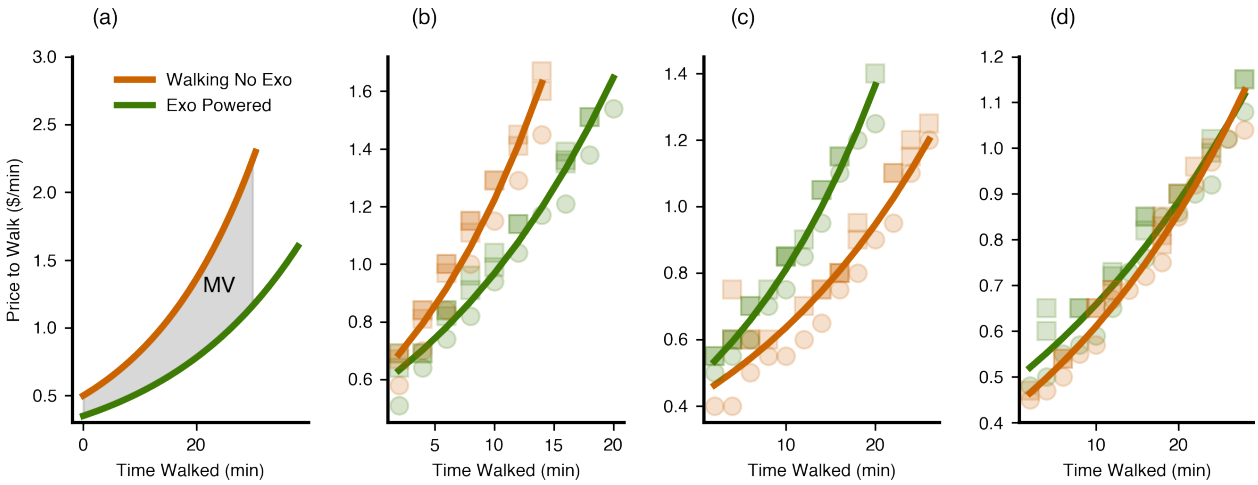

Figure 3.5: (a) Generic examples of price to walk curves for the walking-no-exo condition (green) and the exo-powered condition (orange) from different subjects. The marginal value (MV, the value of the exoskeleton assistance) is given by the area between the curves. In actual trials, the price to walk curves are estimated by fitting first order exponentials to subject bids. Circles denote winning bids, while squares denote losing bids. The participant in (b) shows a clear benefit from the device, the one in (c) shows a clear detriment, and the one in (d) is more ambiguous.

Within each trial, participant bids invariably trended upwards as participants became more fatigued and this trend was well represented by a first-order exponential. During each trial, participants either walked with normal walking shoes **('walking-no-exo')**, with the exoskeleton applying assistance **('exo-powered')**, or with the exoskeleton donned but applying no power **('exo-powered-off')**. Across all three walking conditions, the average bid for each two-minute interval was \$0.75, with a standard deviation of \$0.39. The maximum bid was \$5.00, while the minimum bid was \$0.10. As participants continued to walk during each trial, their bids increased at varying rates (Fig. 3.5). The first-order exponential model exhibited an R^2 of 0.87 ± 0.11, averaged over all conditions experienced by the sixteen participants. These first-order responses

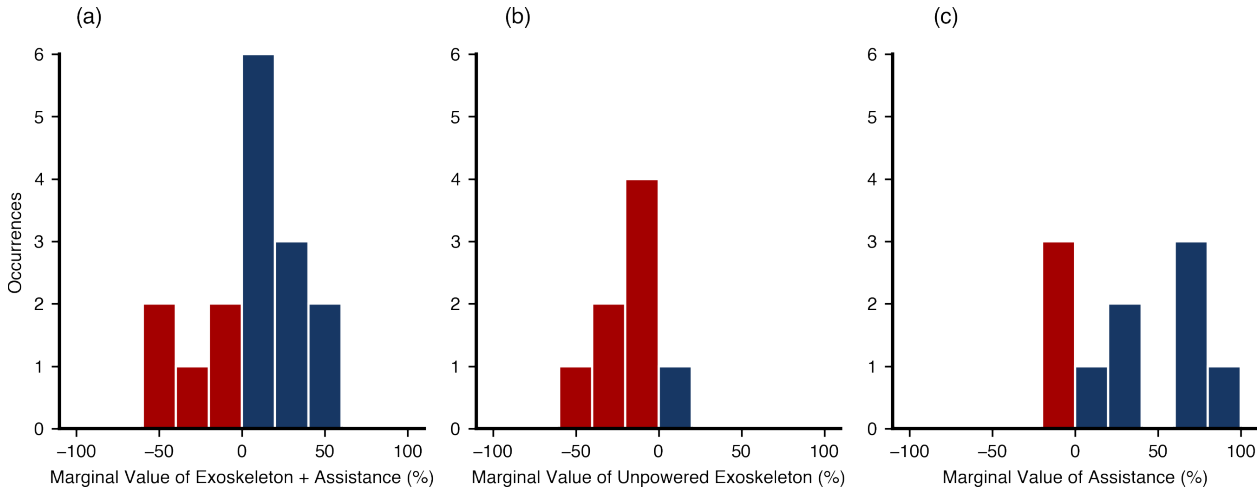

Figure 3.6: Histograms of the Marginal Values (MVs) for the different conditions. Positive MVs, indicating that value was added to the wearer are in blue, while negative MVs, indicating that costs were imposed on the wearer, are in red. (a) The MVs of exoskeleton + assistance for all sixteen subjects. The average MV was 5.81%, with an SD of 31.1%. (b) The MVs of the unpowered exoskeleton for the twelve subjects who participated in this condition. Aside from one subject, all participants experienced notable economic costs from this condition, reflected by the average MV being significantly negative (average: -31.8%, SD: 45.0%). (c) The MVs for the exoskeleton assistance alone (average: 33.8%, SD: 38.1%). While the average MV of the exoskeleton + assistance was not significantly positive, the assistance alone conferred a significant benefit.

denote the user's price to walk curves for each condition (the curves for the walking-no-exo and exo-powered conditions are shown in Fig. 3.5a). The area between the walking-no-exo and exo-powered curves denotes the Marginal Value (MV) of the exoskeleton and its assistance, which is the value obtained by the participant from the exoskeleton's use. Participant price-to-walk curves broadly demonstrated three potential outcomes; namely, a clear economic benefit from the exoskeleton's assistance (higher walking-no-exo curve than exo-powered curve, Fig. 3.5b, positive MV), a clear economic penalty (lower walking-no-exo curve than exo-powered curve, Fig. 3.5c, negative MV), and a negligible economic effect (similar walking-no-exo and exo-powered curves, Fig. 3.5d, near-zero MV).

Our approach was able to quantify the intuitive effects of added mass and assistance. Across subjects, the MV of the exoskeleton + assistance was positive but not significant, while the MV of the unpowered exoskeleton was significantly negative, and the MV of the assistance itself was significantly positive. The average inter-subject MV of exoskeleton use was 5.82%, with a standard deviation (SD) of 31.14% (N=16, Fig. 3.6a). The exoskeleton + assistance thus provided only a small value benefit to the average participant. Using a t-test, the average MV of exoskeleton + assistance (5.82%) was not significantly different from zero ($p = 0.24$). However, as denoted by the high standard deviation, some participants received large benefits from the device's assistance,

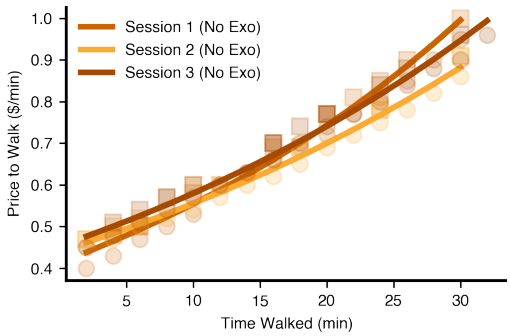

Figure 3.7: Representative price-to-walk curves for a single 'repeater' subject across different days. The agreement across the curves demonstrates the test re-test reliability of our approach. These sessions were completed over a seven day span.

while others experienced an economic penalty from exoskeleton use. The average MV for the unpowered exoskeleton was -31.8% with a standard deviation of 45.0% (N=13, 3.6b). Using the same t-test as in the exo-powered condition, we found this change to be significantly different from zero ($p = 0.03$). Additionally, the powered assistance alone from the exoskeleton provided a significant increase in value (mean: 33.8%, SD: 38.1%, $p = 0.01$, N=13, Fig. 3.6c).

The integral of each price to walk curve yields the cumulative price to walk for each condition (Fig. A.3). The average cumulative price for the walking-no-exo condition was $\$29.23 \pm \$28.11$ (Fig. A.3a), for the exo-powered condition $\$25.38 \pm \$13.87$ (Fig. A.3b), and for the exo-powered-off $\$49.68 \pm \$54.34$ (Fig. A.3c). Changes in participant value due to the different walking conditions can be measured by comparing the associated cumulative prices.

We also characterized the repeatability of our measurements with a subset of participants to verify that the changes in price to walk detected due to the changing walking conditions were attributable to the difference in perceived difficulty between conditions and not due to participant day-to-day variability (a representative repeater trial is shown in Fig. 3.7). For the four subjects that repeated the walking-no-exo condition at least once, the average intra-subject standard deviation of the cumulative price for the walking-no-exo condition was 3.38%, expressed as a percentage of each subject's average walking-no-exo cumulative price. This quantity represents the day-to-day fluctuation in the value demanded to walk, in addition to any user or experimentally derived noise.

3.5 Discussion

In this work, we use the Vickrey second-price auction as a method to capture the economic value or detriment provided by exoskeletons and their assistance. As part of our broader goal of emphasizing the user's role in the design, control, and evaluation of exoskeletons [117, 83], the intent of this work is to quantify the success of these technologies through a user-centered metric which

encompasses the different aspects of the user experience, including exertion, weight, comfort, and assistance. To this end, the wearer specified their "price to walk" during a series of Vickrey auctions from which we quantified the economic value of different conditions. Our approach is particularly relevant for assessing the success of augmentative exoskeletons, which do not have a clear, clinically-relevant, biomechanical or physiological objective (as in the case of prostheses or orthoses). Our work is motivated by the belief that obtaining accessible and relevant metrics of success is critical to the successful development and adoption of exoskeletons in the real world. This motivation is supported by a history of strong interconnection between the assessment of exoskeletons and their design and control architectures [95, 201, 42, 186, 104, 163, 61]. Our strategy quantifies success using monetary currency, as opposed to biomechanical or physiological quantities that may be more difficult for users, manufacturers, and those outside of the exoskeleton research field to interpret.

Our results showed that the economic value provided by the exoskeleton assistance was comparable to the cost incurred by wearing the unpowered system. We obtained these results by quantifying the marginal value (MV) between two conditions, and by varying which two conditions were compared, we isolated the value of the exoskeleton + assistance, the assistance alone, and the detriment of the unpowered exoskeleton. The average cumulative price of walking uphill for 30 minutes without the exoskeleton was \$29.23 (\$58.46/hr SD: \$57.49/hr). The MV of wearing the unpowered exoskeleton—that was not providing assistance—was -31.8%, which translates to a monetary cost of \$18.59/hr (SD: \$18.28/hr) for wearing the unpowered system. When assistance was applied by the exoskeleton, the MV increased to just above zero (\$3.40/hr, SD:\$3.35/hr) when compared to not wearing the exoskeleton. The marginal value added by the assistance alone was 33.8%, which translates to an added value of \$19.76/hr (SD: \$19.43/hr). Thus, the assistance applied by the exoskeleton offset the cost of wearing the system, and the net benefit was modest but positive. These results suggest that modern augmentative ankle exoskeletons may not provide a substantial benefit to their wearers during short-term uphill walking. This is particularly surprising because of the lightweight, refined design of the exoskeletons used, and the high physiological demands of the uphill walking task [131, 51], which we chose to increase the observed value of the exoskeleton (*i.e.* to improve the signal-to-noise ratio in our measurement of economic value).

Our approach and metrics can be used to guide the development of exoskeletons in new ways. Exoskeleton controllers can potentially be directly engineered to maximize the MV added (in dollars) from the assistance of the exoskeleton during walking. Similarly, exoskeletons can themselves be designed to minimize the economic cost they impose on the wearer, which includes not only weight, but also other factors such as inertia and discomfort. This information can be obtained using the MV of the unpowered exoskeleton when compared to the walking-no-exo condition. Additionally, this study suggests that exoskeleton development may be improved by identifying the

users who innately receive greater economic value from wearing the system. We found that nine participants obtained substantial positive economic value from wearing the system ($27.8\% \pm 16.3\%$, corresponding to $\$16.25/\text{hr} \pm \$9.53/\text{hr}$), even though the average across all participants was lower. Such "responders" may be more willing to adopt exoskeletons in the face of potential drawbacks, including cost and added mass, among others. Understanding why some users obtain this value would enable the targeting of these individuals to maximize the translation and impact of augmentative exoskeletons. For example, models such as the diffusion of innovation (DOI) theory within the social sciences categorize individuals based on their willingness to adopt a new innovation [144], such as exoskeletons. It is possible that the individuals who obtained positive value within our study fall into the "early adopter" category, and are more willing to find value in exoskeleton use. Future investigations could focus on the separation of these individuals *a priori* based on the social characteristics within DOI theory, or based on other human factors, such as fitness level or perceptual abilities [117]. Furthermore, investigations of whether users can be "trained" to receive a greater economic value from the exoskeleton are promising avenues of future study.

Participant results were consistent when repeated across days, supporting the quality of our measurements. Four participants repeated the walking-no-exo condition across several days, and the inter-day variance of the cumulative prices was low ($0.67, Table A.4). This consistency supports the ability for our approach to quantify perceived changes in value, with potentially minimal effect from inter-day confounding factors. Using this inter-day variance, we estimated the Minimum Detectable Change (MDC), defined as the minimum change in value not caused by chance. The MDC for the MV measurement was 9.4% (95% confidence interval, difference between two measurements, standard deviation 3.4%). To identify which subjects had a noticeable change in value, we compared each participant's MV against the inter-day variability's 95% confidence interval ($\pm 9.4\%$); nine of 16 subjects had MVs exceeding this threshold in the positive direction, while four subjects exceeded it in the negative direction.

Our auction strategy avoids several potential limitations of more direct methods to quantify value, such as a Likert scale test, direct feedback, or auctions without a real monetary payout [125]. The locomotion task implemented in our approach does not necessitate extreme fatigue and can remain relatively short, avoiding confounding factors (*e.g.* boredom or opportunity cost). Additionally, the incentive structure of the auction links participant bids to a specific consequence (*e.g.* walking uphill), and are thus less prone to biases, such as self-enhancement or social desirability, while still remaining intuitive [165, 44].

Individual bids and, subsequently, cumulative prices varied widely across subjects. Since participants were able to set their own bids, this added inherent variability across subjects. This variation may have resulted from each participant having different internal valuations of their time, which could have been driven by differing socioeconomic status, athleticism, opportunity cost, or other

factors. This large discrepancy motivated the creation of the marginal value (MV) metric, which normalizes by each participant's 'baseline' cumulative price from the walking-no-exo condition, expressing the change as a percent of the walking-no-exo condition.

We anticipate that the conclusions drawn in this study extend to other augmentative ankle exoskeletons, although further study is needed to quantify the value from other technologies. The ankle exoskeletons from this work were developed for commercial use, and represent a "best-in-class" technology; the system's refined design is lightweight, untethered, and provides substantial net-positive energy during each stride (13.4 ± 2.9 J). We expect that other, similar exoskeletons would have comparable marginal values, and thus may also not provide a substantial economic benefit to their wearers. Additionally, although the task of uphill walking does not represent all possible uses for augmentative exoskeletons, it enables an opportunity to quantify value provided by the exoskeleton during an intuitive application where it can provide substantial benefit. We expected that the increased energetic difficulty of uphill walking [131, 51] would enable the exoskeleton to more readily demonstrate its value to the wearer. Furthermore, we chose the uphill walking task to reduce experimental duration as the greater difficulty would cause participant bids to rise more quickly. We expect that for less strenuous tasks, exoskeletons will show reduced marginal value; this hypothesis is supported by the results of a separate study we conducted (see Supplementary Information) in which the MVs of the exoskeleton + assistance were lower for level-ground walking when compared to uphill walking.

The changes in exoskeleton conditions caused intuitive changes in value, supporting the validity of our economic value metric. That is, we believe the different exoskeleton conditions (walking-no-exo, exo-powered-off, and exo-powered) likely drove the changes in value measured, and these changes agree with the biomechanical demands of the conditions. The MV of the exo-powered-off condition was strongly negative, denoting a cost to wearing the system. This result is expected, since without assistance, wearing an exoskeleton is akin to wearing shank weights during locomotion, which would necessitate greater mechanical work from the triceps surae and cause an upstream increase in (by $\sim$ 16.8 W based on the results from [20, 121]). Additionally, when the assistance was added, the value increased to just above zero, indicating the assistance was useful in offsetting the challenge of wearing the unpowered system during uphill walking.

The MV for the exo-powered condition may shift with repeated sessions, as the participants adapt to the assistance. Prior work has found that training sessions across multiple days yield greater reductions in metabolic rate in naive users [132]. Accordingly, the MV for the exoskeleton may adjust as well while the user adapts to the experience of wearing an exoskeleton. For this initial investigation, we sought to understand the immediate economic value obtained from a first-time experience of exoskeleton use, analogous to a user assessing an exoskeleton when making the decision to adopt (*i.e.* a "test drive"). To calculate the MV, we thus conducted a single exo-powered

session with a brief adaptation period of a few minutes. Future work is needed to understand any adaptation of value that may occur over time, which would have implications in the longer-term value of these technologies.

The use of simulated bidding agents may have influenced the results but we expect the overarching conclusions of the study would not be affected. In our auction protocol, we used simulated bidding agents ("robo-bidders") to model human behavior, rather than implementing our approach with multiple human subjects. The intent for this choice was to reduce the logistical challenges of our approach, which would otherwise have required multiple treadmills and exoskeletons. To mitigate any effect of the simulated nature of the other participants, the human participant was informed that the robo-bidders were humans participating in the experiment in remote locations. The behavior of the robo-bidders—defined by a parameter that governs the rise and fall of their bids—was derived from pilot data obtained from human participants. To reduce the likelihood that the human participants could infer the robo-bidder behavior model, the robo-bidder bids were corrupted with noise. The use of robo-bidders also enabled us to standardize the interaction between the human participant and the other auction participants. Any series of auctions would naturally establish an equilibrium between the participants; thus, by using robo-bidders, we were able to control for this equilibrium, which strengthens our ability to compare across subjects. In addition, the use of robo-bidders enables our results to more readily be compared across other researchers, institutions, and exoskeletons that are assessed using comparable methods. The code to implement the robo-bidders is provided in [TBD].

During the experimental protocol, participants responded with bids as instructed, but we are unable to know for certain that their bids were truly honest (*i.e.* truthfully reflecting their internal sense of value). Our study relies on participants honestly reporting their bids, which is theoretically guaranteed as an optimal strategy by the nature of the Vickrey second-price auction for rational actors [183]. In addition to the natural structure of the Vickrey auction, which incentivizes truthful bidding, we minimized the potential influence of dishonest bidding by setting the auction interval to two minutes (as opposed to a shorter duration). We chose the two minute duration to increase the effort required, and thus minimize potentially 'dishonest' exploratory bidding, which would corrupt the price to walk curves. Additionally, subjects continuously received verbal instructions to always bid honestly as the best strategy, which has been demonstrated to increase the likelihood of honest bidding [111]. Finally, participants received the expected monetary compensation that resulted from their winning bids, which similarly incentivizes truthfulness.

3.6 Conclusion

We have developed a method to quantify the economic value of augmentative exoskeletons, and used these methods to assess the value provided by bilateral ankle exoskeletons during uphill walking. Our results underscore the challenge of developing exoskeletons that provide a clear, meaningful benefit when augmenting the healthy human neuromotor system. The value of the assistance provided by the exoskeletons was modest, and just offset the cost of wearing the unpowered system. Our results also suggest that more work is needed to identify why some participants received substantially greater value from the exoskeleton assistance, which could be used to identify or train individuals for maximizing the real-world impact of these technologies. Finally, the economic value metric we have developed can be readily used to compare different design and control strategies to develop exoskeletons that are maximally valuable to their wearers.

CHAPTER 4

Analysis of the Bayesian Gait-State Estimation Problem for Lower-Limb Wearable Robot Sensor Configurations

4.1 Context

Chapter 4 of this dissertation tackles the challenge of developing exoskeleton controllers capable of applying assistance in the real world by first investigating the fundamental problem of estimating the gait state (*i.e.* a vector encoding of human gait) using the the Bayesian filtering framework. The objective is to determine the most suitable implementation of the Bayesian filter for this problem and to identify which gait state variables can be estimated effectively using various sensor combinations available on lower-limb powered exoskeletons. The study results in a practical implementation of the Kalman filter for continuously estimating the gait state, which can then be used to adapt the exoskeleton assistance in real-time to practical environments. This chapter has been published in the *IEEE Robotics and Automation Letters* and presented at the *2023 IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics* [118].

4.2 Introduction

Lower-limb wearable robots, such as exoskeletons and prostheses, have the potential to transform the mobility of the public. Already, powered exoskeletons have been able to improve the physiological performance of users during locomotion, including reducing the metabolic cost [123, 202, 42, 132, 150] and muscular effort [62, 101, 206] below the levels of unassisted walking, while powered prostheses have allowed amputees to regain locomotion ability [5, 197, 198, 65, 80, 45]. However, these technologies have been largely limited to the controlled conditions of a laboratory. To truly impact society, these technologies must function with the unsteady, transitory gaits that arise in the real world, such as walking at variable speeds and inclines.

Measuring these task variables, along with the user's progression through a gait cycle, is difficult to do directly. Recent work has demonstrated success in estimating dynamically changing gaits using implementations of the Bayesian filtering framework [173, 90, 30, 116]. Thus, a systemic investigation into the fundamental problem of gait state estimation using the Bayesian framework can guide wearable robot controller design and may lead to better estimation of real-world walking conditions.

Gait variation can be represented in multiple ways. The concept of *gait phase* quantifies progression through the gait cycle. A phase variable ranges from 0 at heel strike to 1 at the next ipsilateral heel strike [184]. Its time derivative, phase rate, is not necessarily constant. Recent work in phase-based controllers for powered prostheses and orthoses relies on estimates of phase (and sometimes phase rate) [75, 140, 141, 88, 89, 157, 200, 76, 158] to handle gait speed variation. To encode other types of task variation, machine learning-based solutions often classify the locomotion task from signal patterns from onboard sensors [180, 80, 197, 86, 196, 106, 78, 89, 28, 158]. However, such controllers are typically only able to classify tasks within a pre-specified set of categories. Walking tasks have been parameterized by continuous *task variables* such as ground slope or walking speed [47, 46, 12], but these task variables have only been estimated once per stride in laboratory conditions. New methods for real-time, continuous task estimation are needed to enable seamless adaptation to real-world conditions.

Bayesian filtering allows using noisy measurements to estimate hidden states like gait phase and continuous task variables; this framework is well-suited to discriminate the differences in gait biomechanics between strides to notice changes in the underlying task. For linear systems with Gaussian noise, Bayesian filters are implementable as the classic Kalman filter. However, for nonlinear problems, simplifying assumptions are inevitable [148] with differing levels of assumption strength. The Ensemble Kalman Filter (EnKF) has the weak assumption of a Gaussian state distribution, and uses Monte Carlo methods to calculate the measurement update [48]. The Unscented Kalman Filter (UKF) assumes more, and uses quadrature integration and judiciously chosen points and weights to approximate the nonlinear measurement update [185]. The Extended Kalman Filter (EKF) assumes the most, and trusts a local linearization to perform the nonlinear measurement update.

Recently, Bayesian filtering has been applied to human gait estimation in wearable robotics. Thatte et al. [173] introduced a simple two-state EKF as a robust solution to estimating progression through the stance period using the hip, knee, and ankle angles and velocities from a knee-ankle prosthesis. Stance progress was also estimated using the heel pressure and shank angle of an ankle prosthesis [30]. However, the swing period was handled separately in these two approaches. Researchers were also able to estimate stride length using a twice per step filter update based on hip and knee velocities from an exoskeleton [90], but this estimate was not continuously updated.

Zhang et al. [203] proposed new methods using an EKF to estimate the gait phase and walking heading with greater accuracy, although this method relied on sensor thresholds set manually and did not update its task variables in real-time. Our recent work introduced a 4-state EKF to continuously estimate stride progress (both stance and swing), stride length, and ground inclination using foot and shank angles/velocities and filtered heel acceleration for real-time adaptive control of an ankle exoskeleton [116]. This simultaneous estimation approach begs the question of which measurements inform which states, which sensors are necessary for successful estimation, and to what extent the other Bayesian filters in the literature reflect the non-linearity of gait-state estimation.

In this paper, we A) systematically evaluate different implementations of the Bayesian filter to determine which simplifying assumptions (*i.e.*, those of the EnKF, UKF, or EKF) were best suited to the challenge of gait-state estimation; B) investigate which angular kinematic measurements of the leg segments provide the most information about the underlying gait state, which can inform a minimal realization of sensors on hardware that balances estimation quality with sensor complexity; and C) determine the relationship between sensor configurations and estimation of simplified models of the gait state without ground inclination, stride length, or both. We intend for these results to inform the design of future wearable systems that estimate human gait.

4.3 Methods

4.3.1 Process Model

Bayesian estimation employs a forward model that predicts the dynamic evolution of the measurable link angles according to a hidden gait-state that evolves from its initial conditions according to a dynamic model. We begin by defining the gait-state vector x to be estimated,

$$x(t) = \begin{pmatrix} p(t) & \dot{p}(t) & l(t) & r(t) \end{pmatrix}^T, \tag{4.1}$$

comprised of phase p, phase rate $\dot{p}$, stride length l, and the ground inclination r. The gait-state at time k evolves to time $k+1$ as

$$x_{k+1} = Fx_k + w_Q, \tag{4.2}$$

with w_Q distributed as zero-mean Gaussian process noise of covariance Σ_Q and state transition matrix F as

$$F = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.3}$$

This choice of F represents a simple numerical integration of phase rate using the time stride Δt. We model Σ_Q as a diagonal matrix $\mathrm{diag}[\sigma_{11}^2, \sigma_{22}^2, \sigma_{33}^2, \sigma_{44}^2] \times \Delta t$, with σ_{22}, σ_{33}, and σ_{44} being standard deviations for $\dot{p}$, l, and r, respectively. The phase variable p generally has no process noise σ_{11} to represent the noiseless integration of $\dot{p}$. The phase variable p is wrapped within $[0, 1)$ using the modulo operation.

4.3.2 Nonlinear Stride Length Transformation

Within walking, there exists an upper limit on a person's stride lengths, which is determined by the length of their legs. Additionally, we can choose to model the smallest possible stride length as 0, as negative stride lengths that could model backwards walking are instead handled by positive stride lengths and negative phase rates. To encode these choices, we model the stride length as the output of an arctangent transformation [184], to which the input is a 'pseudo-stride length' l_p. The arctangent transformation is defined as

$$l(l_p) = \frac{2}{\pi}\mathrm{atan}(\frac{\pi}{2}l_p) + 1. \tag{4.4}$$

This saturates the stride length output at 2 (meters) and floors it at 0. l_p is allowed to vary freely during EKF estimation. Our state vector x technically contains l_p instead of l, but for a more intuitive understanding of our estimation, we referred to it as containing stride length instead of its 'pseudo' counterpart. We also account for this change of variable when taking partial derivatives; for example, within the Jacobian H in the update step of the EKF, we multiply all partial derivatives with respect to l by $\frac{\partial l}{\partial l_p}$.

4.3.3 Measurement Model

In contrast to our simple linear process model, the gait-state uses a non-linear function $h(x)$ to predict both joint angle and angular velocity measurements. The joint angle measurements arise from a regressed gait model $h_{\mathrm{gait}}(x)$ trained on human biomechanical data, while the angular velocity measurements can be estimated using the differentiation chain rule. The gait model $h_{\mathrm{gait}}(x)$ is trained offline, but is evaluated in real-time. The continuous gait model $h_{\mathrm{gait}}(x)$ predicts global foot angle θ_f, global shank angle θ_s, global thigh angle θ_t, and global pelvis angle θ_p (all potentially

measurable by sensors on a lower-limb wearable robot), and this function is denoted by

$$\begin{pmatrix} \theta_f(t) \\ \theta_s(t) \\ \theta_t(t) \\ \theta_p(t) \end{pmatrix} = \begin{pmatrix} h_f(x(t)) \\ h_s(x(t)) \\ h_t(x(t)) \\ h_p(x(t)) \end{pmatrix} = h_{\text{gait}}(x(t)). \tag{4.5}$$

Additionally, given knowledge of the phase rate $\dot{p}$, we can also model the angular velocity of the foot $\dot{\theta}_f$, shank $\dot{\theta}_s$, thigh $\dot{\theta}_t$, and pelvis $\dot{\theta}_p$ using the differentiation chain rule:

$$\begin{bmatrix} \dot{\theta}_f \\ \dot{\theta}_s \\ \dot{\theta}_t \\ \dot{\theta}_p \end{bmatrix} = \begin{bmatrix} \partial\theta_f/\partial p \\ \partial\theta_s/\partial p \\ \partial\theta_t/\partial p \\ \partial\theta_p/\partial p \end{bmatrix} \dot{p} \tag{4.6}$$

where $\dot{p}$ is the estimate of the phase rate from the prediction stride and the partial derivatives of θ_f, θ_s, θ_t, and θ_p are available analytically from the regressed gait model. Phase substitutes in for stride time when parametrizing the gait cycle. The full observation function is then $h(x) = \left[\theta_f, \dot{\theta}_f, \theta_s, \dot{\theta}_s, \theta_t, \dot{\theta}_t, \theta_p, \dot{\theta}_p\right]^T$. We selected these signals to investigate as they are commonly available (via IMUs and joint encoders) in lower-limb wearable robots, where the exact combination of measurements depends on the hardware and use case. In this investigation, we analyzed the impact of using subsets of this 8x1 vector to drive gait state estimation (*e.g.*, using only shank angle and its velocity).

The gait model $h_{\text{gait}}(x)$ is regressed from labeled training data from a 10-subject able-bodied dataset [47], which contains walking data grouped by strides, over a range of speeds (0.8, 1, and 1.2 m/s) and ramps (-10 to 10 degree inclination in increments of 2.5 degrees). Ground-truth phase, phase rate, stride length, and ground inclination, along with the measurements of the global angles at each condition, are readily available from this dataset. The gait model $h_{\text{gait}}(x)$ takes as input the gait state vector x and outputs the best-fit estimates of the kinematic measurements.

For least-squares regression, $h_{\text{gait}}(x)$ is formulated as

$$h_{\text{gait}}(x) = \phi^T R^T(x), \tag{4.7}$$

where $\phi \in \mathbf{R}^{144\times 4}$ is a matrix of real-valued model parameters and $R : \mathbf{R}^4 \mapsto \mathbf{R}^{1\times 144}$ is a gait-state-dependent regressor row-vector. The parameters ϕ are chosen to minimize the sum squared

error for each equation of the form

$$\begin{pmatrix}\theta_f(t) & \theta_s(t) & \theta_t(t) & \theta_p(t)\end{pmatrix} = R(x(t))\phi, \tag{4.8}$$

for all times t in the training dataset (treating each instant of each step by each participant in each trial as a separate t).

The definition of $R(x)$ makes extensive use of the Kronecker product, $\otimes$ to construct large row-vectors from smaller row-vectors. To review, the Kronecker product of row-vectors $A \in \mathbf{R}^{1\times N}$ and $B \in \mathbf{R}^{1\times M}$, expressed as $A \otimes B \in \mathbf{R}^{1\times NM}$, is the block row-vector $(a_1 B \;\; a_2 B \;\; \cdots \;\; a_N B)$. In the case of matrices $A \in \mathbf{R}^{n\times N},\ B \in \mathbf{R}^{m\times M}$, this generalizes to

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1N}B \\ a_{21}B & a_{22}B & \cdots & a_{2N}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nN}B \end{pmatrix} \in \mathbf{R}^{nm\times NM}. \tag{4.9}$$

The regressor is defined as

$$R(x) = B_p(p) \otimes \Lambda_r(r) \otimes \Lambda_l(l) \otimes \Lambda_p(p), \tag{4.10}$$

that is, it combines the effects of the four simpler behaviors such that the final model depends on p, l, and r. The components are as follows:

- $B_p(p)$: The Boolean selector row $\mathbf{R} \mapsto \mathbf{R}^{1\times 4}$,

$$B_p(p) = \begin{cases} (1,0,0,0) & \text{if} \quad 0 \le p \le 0.1, \\ (0,1,0,0) & \text{if} \quad 0.1 < p \le 0.5, \\ (0,0,1,0) & \text{if} \quad 0.5 < p \le 0.65, \\ (0,0,0,1) & \text{if} \quad 0.65 < p \le 1, \end{cases} \tag{4.11}$$

  which divides the gait phase into four sections. In our gait model, we chose four Bernstein polynomials (see $\Lambda_p(p)$ below) to represent the kinematics; Bernstein bases have an equivalent span to polynomial bases and have previously been useful in gait modelling [47]. The sections were determined by inspection of the nominal biomechanical kinematics in our dataset.

- Λ_r : The ramp angle basis ($\mathbf{R} \mapsto \mathbf{R}^{1\times 3}$) is a second-order polynomial Bernstein basis in

ramp angle,

$$\Lambda_r(r) = \begin{pmatrix} (1-r)^2 & 2(1-r)r & r^2 \end{pmatrix}, \tag{4.12}$$

which allows for continuous adjustment to ground slope.

- Λ_l : The stride length basis ($\mathbf{R} \mapsto \mathbf{R}^{1\times 3}$) is a second-order Bernstein polynomial basis in stride length,

$$\Lambda_l(l) = \begin{pmatrix} (1-l)^2 & 2(1-l)l & l^2 \end{pmatrix}, \tag{4.13}$$

which allows for kinematic changes associated with stride length.

- Λ_p : The phase-polynomial basis $\mathbf{R} \mapsto \mathbf{R}^{1\times 4}$ is a Bernstein polynomial basis, defined as

$$\Lambda_p(p) = \begin{pmatrix} (1-p)^3 & 3(1-p)^2 p & 3(1-p)p^2 & p^3 \end{pmatrix}. \tag{4.14}$$

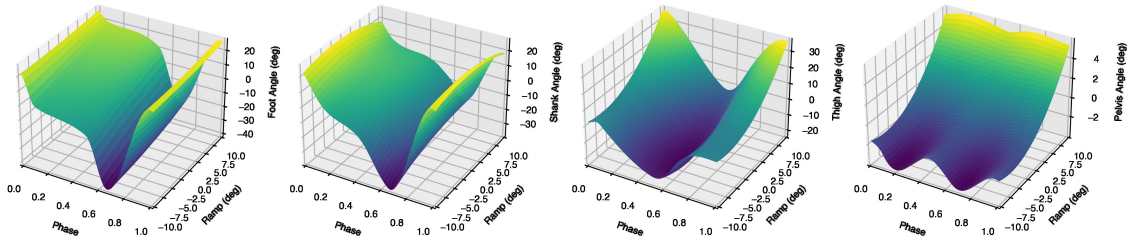

Figure 4.1: The regressed continuous gait models for θ_f, θ_s, θ_t, and θ_p, left to right respectively. As each model depends on three input variables (p, l, and r) and has one output variable (θ_f, θ_s, θ_t, or θ_p), the models reside fully in 4D-space and are thus difficult to plot. In this figure, the models are plotted over phase and ramp (with stride length constant at 1 meter).

4.3.4 Measurement Model Constraints

To ensure desirable gait model properties, such as continuity of the function $h_{\text{gait}}(x)$ (the global angle predictions) and realistic behavior with changing stride length, the elements of the parameter matrix ϕ are subject to constraints. To express the C^0 continuity of the model, we require $\lim_{p\to\rho^-} h_{\text{gait}}(x) = \lim_{p\to\rho^+} h_{\text{gait}}(x)$ for all $0 < \rho < 1$ as well as the special wrap-around case where $\lim_{p\to 1^-} h_{\text{gait}}(x) = \lim_{p\to 0^+} h_{\text{gait}}(x)$. This is trivially satisfied everywhere except at $\rho = 0.1,\ 0.5,\ 0.65$, and in the wrap-around case. In these four cases, equality constraints must be satisfied for all

possible stride lengths and ramp angles. We express this constraint on ϕ using a matrix equality:

$$\begin{pmatrix} (1\ -1\ 0\ 0) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes \Lambda_p(0.1) \\ (0\ 1\ -1\ 0) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes \Lambda_p(0.5) \\ (0\ 0\ 1\ -1) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes \Lambda_p(0.65) \\ \Big[(0\ 0\ 0\ 1) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes \Lambda_p(1) \\ \quad -(1\ 0\ 0\ 0) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes \Lambda_p(0)\Big] \end{pmatrix} \phi = 0_{36\times 4}, \tag{4.15}$$

where

$$\Lambda_{\forall r} = \begin{pmatrix} \Lambda_r(10) \\ \Lambda_r(0) \\ \Lambda_r(-10) \end{pmatrix} \text{ and } \Lambda_{\forall l} = \begin{pmatrix} \Lambda_l(0) \\ \Lambda_l(1) \\ \Lambda_l(2) \end{pmatrix} \tag{4.16}$$

serve to constrain all parts of the quadratic fits in r and l by specifying three (arbitrary but unique) points of each.

To express the C^1 continuity constraint, we exploit the linearity of the Kronecker product. For almost all p we can express the derivative

$$\frac{dh^T_{\text{gait}}(x)}{dp} = \left[B_p(p) \otimes \Lambda_r(r) \otimes \Lambda_l(l) \otimes \frac{d\Lambda_p(p)}{dp}\right] \phi, \tag{4.17}$$

since $\frac{dB_p(p)}{dp}$ is zero almost everywhere.

Note that $d\Lambda_p/dp : \mathbf{R} \mapsto \mathbf{R}^{1\times 4}$ is available analytically. The resulting continuity constraint is then

$$\begin{pmatrix} (1\ -1\ 0\ 0) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes d\Lambda_p/dp(0.1) \\ (0\ 1\ -1\ 0) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes d\Lambda_p/dp(0.5) \\ (0\ 0\ 1\ -1) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes d\Lambda_p/dp(.65) \\ \Big[(0\ 0\ 0\ 1) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes d\Lambda_p/dp(1) \\ \quad -(1\ 0\ 0\ 0) \otimes \Lambda_{\forall r} \otimes \Lambda_{\forall l} \otimes d\Lambda_p/dp(0)\Big] \end{pmatrix} \phi = 0_{36\times 4}, \tag{4.18}$$

simply requiring that the derivatives match on either side of each potential discontinuity in p (for all values of r and l).

To ensure constant-with-phase behavior when stride length is zero, *i.e.*, when the person is standing still, we require

$$\overbrace{I_4}^{\forall\ B_p(p),} \otimes \overbrace{\Lambda_{\forall r} \otimes \Lambda_l(0)}^{\forall r, \text{ if } l=0,} \otimes \overbrace{\begin{pmatrix} \Lambda_p(0) \\ \Lambda_p(\frac{1}{4}) \\ \Lambda_p(\frac{1}{2}) \\ \Lambda_p(\frac{3}{4}) \end{pmatrix}}^{\forall\ p} \phi = \overbrace{1_{48\times 1} \otimes \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}^T}^{\text{angles are constant}}. \tag{4.19}$$

The constraint takes a form similar to the Kronecker construction of the regressor, but where the regressor has $B_p(p)$, the constraint uses an identity matrix. The purpose of this is to expand the row dimension of the constraint so that it applies to any of the four possible cases for $B_p(p)$. Similarly, where the regressor has $\Lambda_p(p)$, the constraint has a block matrix that constrains the polynomial at four points (enough to ensure the third-order polynomial expression is equal to a constant everywhere). Taken together, these two components of the constraint equation force all four 3rd-order parts of the piecewise polynomial $\Lambda([p, \dot{p}, l = 0, r = 0]^T)\phi$ to be constant for all the measured kinematics. Any four unique phase points could replace the constants $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ and achieve the same effect of constraining the 3rd-order polynomials to be everywhere zero.

To enforce that the gait model has zero derivative with respect to stride length at zero stride length, we apply a constraint similar to (4.19), using $\frac{d\Lambda_l}{dl}$ instead of Λ_l:

$$\overbrace{I_4 \otimes \Lambda_{\forall r} \otimes \frac{d\Lambda_l}{dl}(0)}^{\forall B_p, r, \text{ if } l=0,} \otimes \overbrace{\begin{pmatrix} \Lambda_p(0) \\ \Lambda_p(\frac{1}{4}) \\ \Lambda_p(\frac{1}{2}) \\ \Lambda_p(\frac{3}{4}) \end{pmatrix}}^{\forall\, p} \phi \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}^{\text{ignoring pelvis}} = \overbrace{0_{48\times 3}}^{\text{d/dp=0}}. \tag{4.20}$$

We also constrain the pelvis to be a linear fit with respect to stride length by constraining it to have a constant derivative with respect to stride length (a constraint similar to (4.19) with $\frac{d\Lambda_l}{dl}$ instead of Λ_l).

We regressed the model using the constrained least-squares optimization function `lsqlin` in MATLAB (Fig. 4.1).

4.3.5 Heteroscedastic Noise Model

We developed a heteroscedastic measurement noise model that dynamically changes the measurement noise matrix Σ_R based on phase p. In this schema, Σ_R is defined as

$$\Sigma_R(p) = \Sigma_{R,\text{sensor}} + \Sigma_{R,\text{xsub}}(p), \tag{4.21}$$

where $\Sigma_{R,\text{sensor}}$ is the traditional measurement noise matrix that denotes how uncertain the measurements are, and $\Sigma_{R,\text{xsub}}$ denotes the uncertainty present due to inter-subject gait kinematic variability (subscript xsub for cross-subject). In this schema, the heteroscedastic model not only encapsulates the uncertainty present due to each person's unique gait, but also can continuously change its trust in the data to capture regions within the gait cycle where the measurements are more trustworthy due to smaller inter-subject gait variability (*e.g.*, flat-foot contact). The matrix $\Sigma_{R,\text{sensor}}$ was set as $\text{diag}[\sigma^2_{11,r}, \sigma^2_{22,r}, \sigma^2_{33,r}, \sigma^2_{44,r}, \sigma^2_{55,r}, \sigma^2_{66,r}, \sigma^2_{77,r}, \sigma^2_{88,r}]$, with each $\sigma_{xx,r}$ representing the standard deviation for $\theta_f, \dot{\theta}_f, \theta_s, \dot{\theta}_s, \theta_t, \dot{\theta}_t, \theta_p, \dot{\theta}_p$, respectively. In our implementation, the $\sigma_{xx,r}$ values that

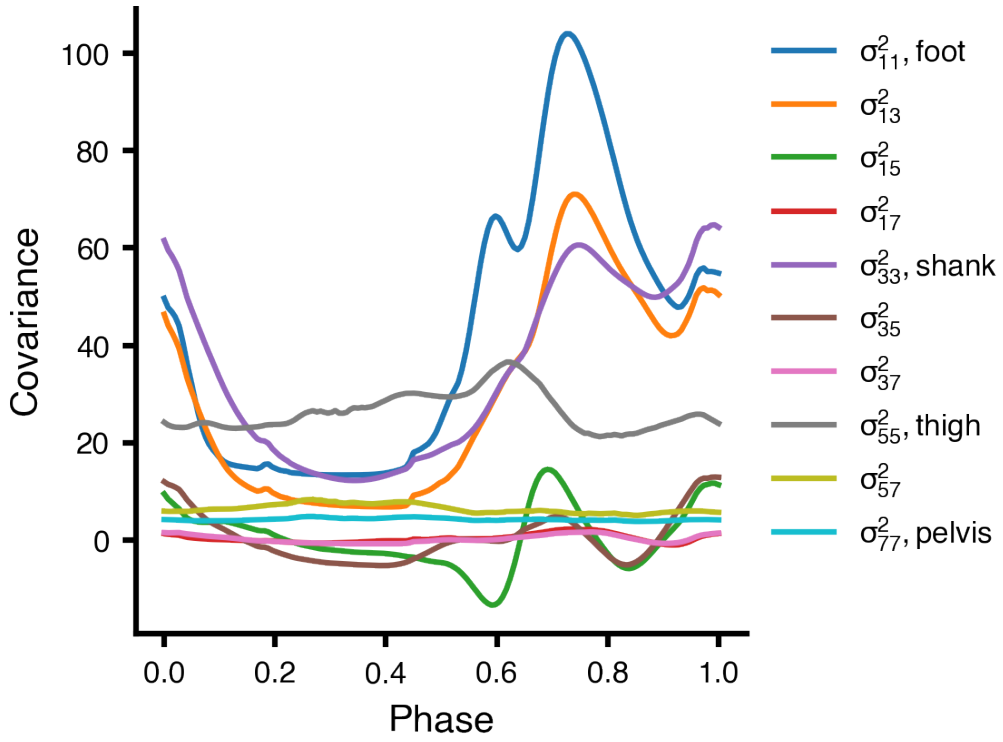

Figure 4.2: The heteroscedastic measurement noise model as a function of phase. Foot angle variance σ_{11}, shank angle variance σ_{33}, thigh angle variance σ_{55}, and pelvis angle variance σ_{77}, are shown, along with their respective covariances.

pertained to the angles and angular velocities were set equal to 1 and 10 respectively.

For $\Sigma_{R,\text{xsub}}(p)$ we used the prior dataset to calculate the covariance matrices of the measurement residuals y for the eight measured kinematic variables at each of 150 phase values (Fig. 4.2). The instantaneous $\Sigma_{R,\text{xsub}}(p)$ was then calculated in real-time using the estimate of phase.

4.3.6 Candidate Bayes Filter Implementations

The first filter we tested was the **Ensemble Kalman Filter (EnKF)**, which uses Monte Carlo methods to approximate the state mean and covariance at each time step. This allows for highly nonlinear measurement maps and makes the fewest assumptions of the filters we tested. In the EnKF, these expectations are tracked by N particles $X_k^{(i)}$, $i = 1, \ldots N$, which each evolve through the prediction and update steps of the Kalman Filter. There is a tradeoff between the number of particles used to track the states and the computational times involved in the filter. In our simulations, we use an EnKF with 1000 particles (EnKF1000) and one with 100 particles (EnKF100). In this formulation, each particle is fed through the dynamics and individually corrupted by samples from the process noise. Owing to the requirement that Σ_Q be strictly positive definite (due to directly needing to sample from the process noise matrix), σ_{11} was set to $1e-20$ to approximate the

noiseless integration while maintaining positive definiteness; σ_{22}, σ_{33}, and σ_{44} were set to $1e-2$, $1e-2$, and $1.5e-1$ respectively. State means and covariances were estimated by the empirical means and covariances of $X_k^{(i)}$ (see [48]).

The **Unscented Kalman Filter (UKF)** is capable of approximating the posterior mean and variance of nonlinear functions up to the 3rd order[185]. For the update step of the Kalman Filter, the UKF approximates the mean and covariances of the states using quadrature integration and carefully selected Sigma Points $\mathcal{X}^{[i]}$ and weights $w^{[i]}$. The Sigma Points $\mathcal{X}^{[i]}$ and $w^{[i]}$ were generated using $\alpha = 1e-3$, $\beta = 2$ for the Gaussian approximation, and $\kappa = 0$ according to Wan and van der Merwe [185].

The **Extended Kalman Filter (EKF)** locally linearizes the nonlinear parts of the system (the gait model) for a simple approximation of the normal Kalman Filter. It makes the strongest approximations about the problem, but offers the best computational performance. State means and covariances were updated using the linear Gaussian update equations using the Jacobian of h with respect to the gait-state vector x [87].

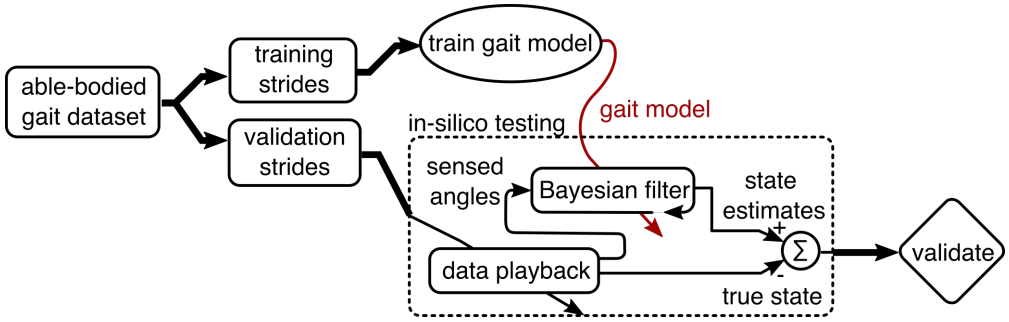

Figure 4.3: The Bayesian filter and gait model regression processes. In this cross-validation study, gait data from our dataset is segregated into training strides from non-excluded subjects that are used to regress the gait model, and validation strides from the subject being evaluated. The gait model is then used within each of the four candidate Bayesian filters. The validation strides contain sensor measurements that are input to the Bayesian filter, which then uses the gait model to yield gait state estimates that are compared to the true states from the validation strides.

4.3.7 Evaluation of Bayesian Filtering Problems

We evaluated each Kalman Filter with different measurement configurations: 1) full, in which the filter had access to foot, shank, thigh, and pelvis angle data (along with their respective velocities), 2) four configurations in which a different measurement was left out, 3) six configurations in which the different permutations of two measurements from the four kinematics were used (six configurations, *e.g.*, shank-foot, thigh-pelvis), and 4) four configurations in which each filter only had a single respective measurement. For the purposes of abbreviation, the sensors present in each

configuration are given by a string composed of the first letter of the angular measurement (foot, shank, thigh, pelvis) aside from the full configuration, which contains all four sensors (*e.g.*, the configuration of shank-foot is 'fs'). For the simulations where angle sensors were dropped from the filtering, the corresponding observation functions were ignored during the filter calculations.

For each combination of filter type and measurement model, we evaluated the combination's gait-state estimation on a simulated walking task that used the data from the same dataset used to regress the gait model. To simulate the filter's performance on unseen subjects, we performed a leave-one-out cross-validation on all ten of the dataset's subjects. For each subject, this cross-validation trained a new gait model and heteroscedastic noise model using the walking data from the remaining nine subjects. The subject's kinematic and walking data were then input to each filter, which estimated the underlying gait-state (Fig. 4.3). Errors for each of the states at each time point were calculated as the difference between the state estimate and its respective ground truth state measurement from the dataset. This process was repeated for all ten subjects, and the errors were aggregated to obtain overall distributions that described each filter combination's performance in estimating each element of the gait-state. For the EKF, in addition to the configurations above, we evaluated the estimation of limited subsets of the gait-state: 1) a subset where incline was excluded (Cancel Incline), 2) a subset where stride length was excluded (Cancel Stride Length), and 3) a subset where phase and phase rate only where estimated (Phase Only). This simulation was motivated by the potential for limited sensor configurations to still estimate parts of the gait-state vector. Ground truth measurements and state errors were computed as in the simulation with the full gait-state vector. To mitigate the effects of an increasing EKF bandwidth due to the removal of states from the filter, the elements of Σ_Q were scaled in the following way: for Cancel Incline, σ_{33} was scaled by 0.5; for Cancel Stride Length, σ_{44} was scaled by 0.5; for Phase Only, σ_{11} and σ_{22} were scaled by 0.5. To obtain an overall metric for estimation performance, we computed the Mahalanobis distance between the gait-state estimate and the ground truth state at each time point. This Mahalanobis distance was normalized using the average state covariance matrix $\bar{P}_k$ from the full-state, full-measurement EKF simulation; for the simulations where subsets of the gait-state were estimated, we used the subset of $\bar{P}_k$ that corresponded to those gait-states. This error metric captured the overall estimation performance of each combination of gait-state and measurement configuration. We then normalized these errors by the error from the full-state, full-measurement EKF simulation to aid in comparisons.

4.4 Results

Overall, the Extended Kalman Filter (Fig. 4.4A) was able to consistently estimate the gait-state despite its restrictive assumptions. With the full measurement configuration, the EKF featured an

average phase error of 0.01 ± 0.02 (SD), an average phase rate error of -0.01 ± 0.03 1/s, an average stride length error of 0.03 ± 0.10 m, and an average incline error of 0.08 ± 1.86 degrees. Gait-estimation performance was similar with the UKF (phase: 0.01 ± 0.01, phase rate: -0.01 ± 0.03, stride length: 0.03 ± 0.10, incline: 0.06 ± 1.96), while the EnKF implementations with 1000 particles (phase: 0.03 ± 0.06, phase rate: -0.02 ± 0.05, stride length: 0.03 ± 0.10, incline: 0.16 ± 2.66) and 100 particles (phase: 0.16 ± 0.14, phase rate: -0.06 ± 0.41, stride length: -0.45 ± 0.33, incline: 0.44 ± 7.31) were less reliable.

A representative trial of each filter's performance exemplifies these errors (Fig. 4.4B). For this representative trial, we computed the average root mean square error (RMSE) across all individual strides for each state variable. The EKF (phase RMSE: 0.01 ± 0.009, phase rate RMSE: 0.02 ± 0.02, stride length RMSE: 0.08 ± 0.06, incline RMSE: 0.72 ± 1.62 degrees) was again comparable to the UKF (phase RMSE: 0.01 ± 0.009, phase rate RMSE: 0.02 ± 0.02, stride length RMSE: 0.08 ± 0.06, incline RMSE: 0.73 ± 1.64); the EnKF with 1000 particles (phase RMSE: 0.03 ± 0.06, phase rate RMSE: 0.03 ± 0.04, stride length RMSE: 0.08 ± 0.07, incline RMSE: 1.2 ± 2.5) was also similar, while the EnKF with 100 particles (phase RMSE: 0.15 ± 0.15, phase rate RMSE: 0.27 ± 0.25, stride length RMSE: 0.38 ± 0.21, incline RMSE: 4.95 ± 4.42) was still the poorest estimator. As an additional comparison, we also computed the absolute peak errors for the different filters. The same relationships were borne out as with the RMSEs: the EKF (phase peak error: 0.04, phase rate peak error: 0.3, stride length peak error: 0.21, incline peak error: 9.8) was comparable to the UKF (phase peak error: 0.04, phase rate peak error: 0.29, stride length peak error: 0.4, incline peak error: 9.9), while the EnKF with 1000 particles (phase peak error: 0.17, phase rate peak error: 0.45, stride length peak error: 0.41, incline peak error: 19.9) and the EnKF with 100 particles (phase peak error: 0.42, phase rate peak error: 0.45, stride length peak error: 0.9, incline peak error: 9.9 ± 1.62) again performed poorly.

For the EKF implementation, removing the pelvis or shank measurements did not excessively deteriorate estimation of the full state vector in terms of the normalized Mahalanobis distance metric. In some sparse sensing configurations (s, t, f, sp, st, stp, ftp), gait-state estimation performance was improved by removing either stride length or incline from the state vector (Fig. 4.5). Canceling the incline state seemed to be more effective than cancelling stride length, with the exception of sp, f, fp, and the full configuration. The configurations of s, sp, and st were capable of estimating phase accurately without the task variables.

4.5 Discussion

Despite being the filter with the weakest assumptions and theoretically best ability to capture the gait-state probability distribution as it evolves through the nonlinear gait model, the EnKFs were

overall the worst performing filter based on gait-state errors. In particular, the EnKF with 100 particles (EnKF100) was the worst filter overall, as it featured the highest standard deviations for phase error and incline, and significant biases for stride length error. While increasing the number of particles from 100 to 1000 mitigates these errors, the EnKF1000 is at best roughly equal to the EKF or UKF. During the simulations, the added complexity in the EnKF1000 led to drastically increased run-times when compared to the EKF and UKF. Taken together, this indicates that the added complexity from an EnKF may be unnecessary for this gait-state estimation task, and instead a simpler EKF or UKF may suffice for the control of an exoskeleton.

The EKF and UKF have close RMSE state errors and are thus comparable in performance (Fig. 4.4A), particularly for the measurement configurations with only a single sensor removed. This has numerous implications for the potential applications of the Bayesian framework to the continuous control of lower-limb wearable robots. For example, the gait model h at the heart of the filters must be linearized in the case of the EKF, which can introduce significant error. However, the EKF performed comparably to the UKF, which can approximate models of up to the third order [185] with more points, and the EnKF, which can more closely approximate the output distributions during the update steps. This indicates that the simple linear approximation of the EKF is sufficient for gait estimation using the gait models developed in this work. The peak errors must also be considered for implementation, as even though they occur infrequently, unlikely events can still pose a risk to the control of an exoskeleton. The peak phase error, for instance, may be enough to surpass the perceivable threshold for timing discrepancy [136], which could momentarily cause the human to feel desynchronized from the device. Acceptability thresholds for the peak errors for the remaining task variables could be established using this perceptual metrics in future work, although such perception would be highly driven by the precise shape of the torque assistance applied. Indeed, it's possible that the torque assistance is itself designed to be robust to peak errors.

The magnitudes of the phase error from the EKF (0.01 ± 0.02) are comparable to recent methods based on Gaussian Processes [173] (phase RMSE: 0.04 ± 0.005) and machine learning [88, 89] (0.07 ± 0.03 from [88], 0.04 ± 0.006 from [89]), although direct comparison is impossible due to the different testing methods. The EKF incline errors (0.08 ± 1.86 degrees) are also similar to prior work that estimated ground inclination in real-time using machine learning [98] (incline RMSE: $2.15° \pm 0.29°$). As depicted in the representative trial (Fig. 4.4B) for the EKF, stride length was generally the state with the greatest estimation discrepancy. We also observed this discrepancy in our prior work [116], and it is likely due to the inter-subject variance in the gait model overpowering the deterministic effects of stride length on the measurable kinematics. However, in our prior work, this discrepancy did not seriously affect exoskeleton torque assistance, which depended on all the gait state elements, primarily phase.

In terms of the sensor configurations, predictably, the full sensor array produced the lowest errors. However, some measurement configurations, such as the foot-shank-thigh and foot-shank, have comparable errors to the full configuration, especially for the EKF. Taken together, these results point to the EKF as being the best choice for the gait-state estimation task, and motivated the simulation experiment using the EKF to identify which subsets of the state vector can be estimated using limited sensors. The simplicity of the EKF also lends itself to straightforward implementation on actual hardware (e.g., our recent work [116]), especially since the most informative kinematic measurements (e.g. foot, thigh) are usually readily available on numerous existing wearable robots. While prior work in gait phase estimation yielded good estimates of phase with thigh angle alone [184, 140, 108, 200, 76], these approaches employed normalization techniques that are not reproduced in our EKF's thigh-only estimation process.

When the measurement vector was limited in simulation, some states were still able to be effectively estimated using the EKF. In terms of the normalized estimation error outlined above, limited sensing configurations such as foot-shank-thigh and foot-thigh are comparable in error to the full-sensing configuration for the full-gait-state case. In particular, the pelvis measurement appears to be the least informative, as configurations with it have similar normalized errors as configurations without. Furthermore, the pelvis sensor alone is easily the worst single measurement across all four gait-state vectors. This suggests that this measurement can be dropped in a practical implementation of the EKF. Conversely, for the most limited case of phase-only state estimation using the EKF, the shank measurement is the optimal choice if only a single measurement is allowed. This indicates that the shank measurement is among the more informative measurements within the EKF framework, and should thus be prioritized in allocating sensors. The overall flexibility of the sensor configuration is a strength of the continuous gait model at the heart of the Kalman Filter. Unlike past methods for estimating the gait-state which required explicit rules to handle new measurements [141, 140], the EKF can incorporate new measurements by simply regressing a relation between the gait-state and the kinematic measurement offline, and then extending the measurement vector in the Kalman Filter. Similarly, in this study, we chose the task variables of stride length and incline as they can parametrize a broad array of gaits; however, the gait state vector could be easily extended to include other task variables depending on the application's demands, such as stride height for ascending stairs, or walking stability for slippery terrains.

4.6 Conclusion

We investigated the challenge of estimating gait behavior (phase and task variables) using Bayesian filtering with the sensors available to lower-limb wearable robots. We found this estimation problem to be tractable *in silico*, even with restrictive EKF assumptions and sparse sensing configura-

tions. Our results suggest that an EKF is a good choice for continuously estimating phase and task to drive and adapt an exoskeleton controller. Furthermore, reduced gait-states can be estimated with far fewer measurements using the proposed framework (for example, shank can be used alone if the user wishes to only estimate phase). Future work will involve testing these Kalman Filters on actual exoskeleton hardware, including modular joint configurations [126], to validate them in the real world.

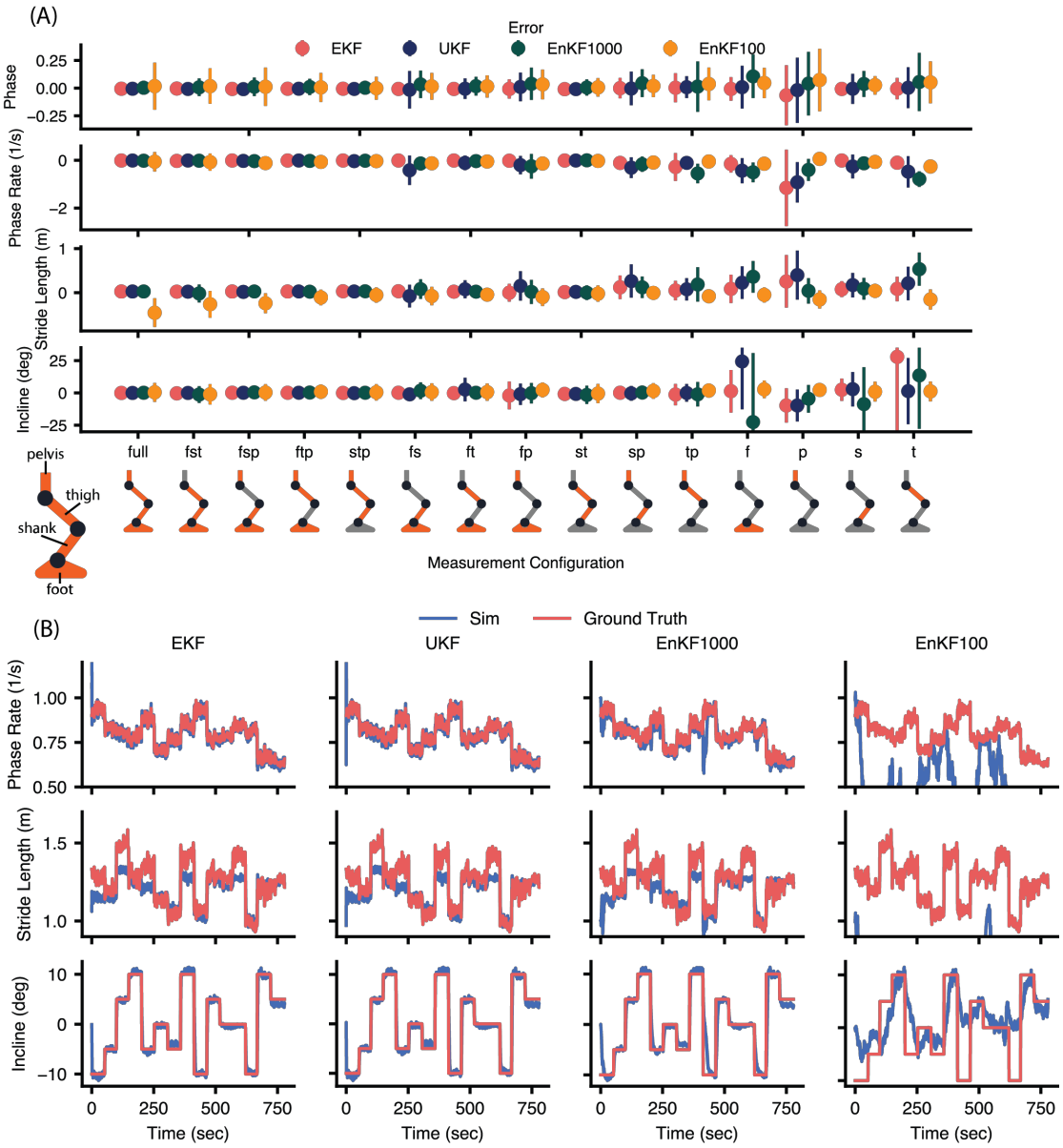

Figure 4.4: (A) The errors for all four elements of the gait-state for the four types of Kalman Filter, using each sensor configuration. The means are denoted by the solid circles, with standard deviations given by the vertical lines. Generally speaking, the EKF and UKF provide the best estimation for the four states. For the purposes of abbreviation, the sensors present in each configuration are given by a string composed of the first letter of the angular measurement (foot, shank, thigh, pelvis) aside from the full configuration, which contains all four sensors (*e.g.*, the configuration of shank-foot is 'fs'). (B) Representative results for the three Kalman Filters run in this simulation, with the full sensor configuration. These results were randomly selected from a single subject in the *in silico* cross-validation. The ground truth states are shown in solid red, with the filter estimates shown in blue. Phase is not shown since it is not perceptible at this time scale, but the estimation quality can be inferred from phase rate graphs.

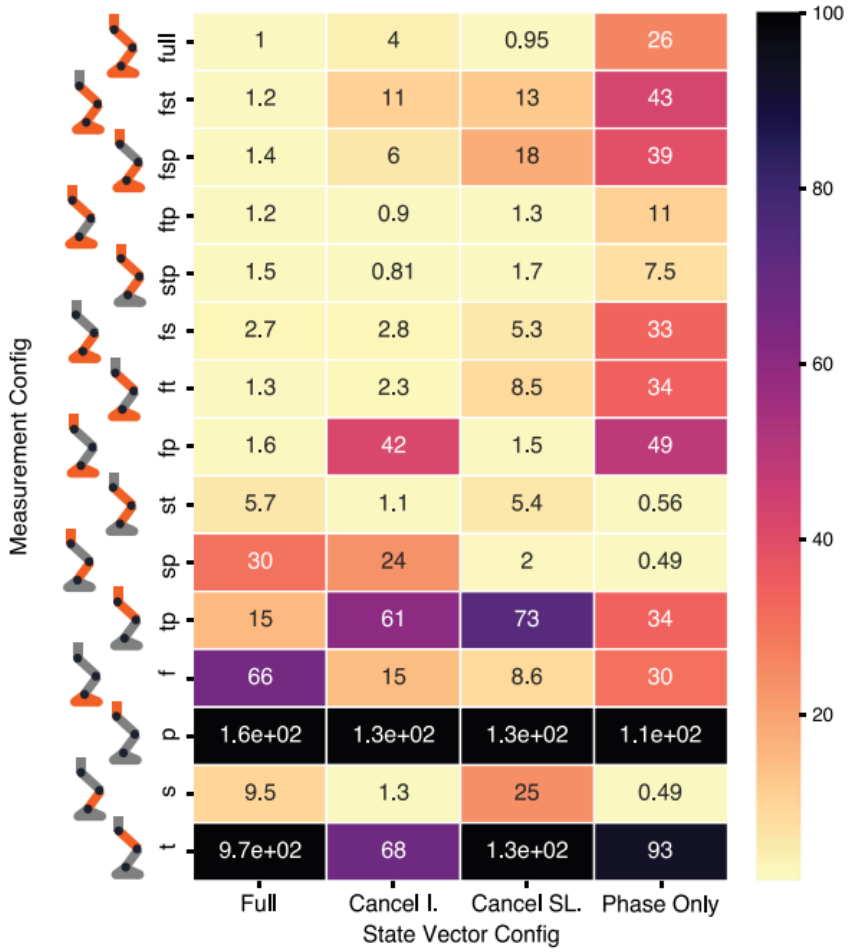

Figure 4.5: EKF estimation performance for the fifteen measurement configurations and four different state vector configurations. Combined state errors were calculated using the Mahalanobis distance weighted by the average state covariance matrix from the baseline case (full sensing, full state, top-left). Errors were then normalized by the error of the baseline state, with higher errors shown in darker colors.

CHAPTER 5

Real-Time Gait Phase and Task Estimation for Controlling a Powered Ankle Exoskeleton on Extremely Uneven Terrain

5.1 Context

Chapter 5 of this dissertation builds on the work presented in Chapter 4 to develop an exoskeleton controller based on the Extended Kalman Filter (EKF) as a solution to the lack of viable real-world controllers. The EKF-based controller continuously estimates the phase variable and the task variables of ground incline and stride length, which are then used to update the torque profile of bilateral ankle exoskeletons across different controlled and outdoor environments. This is a significant development, as it demonstrates the EKF's ability to estimate both phase and task in real-time and provides a practical exoskeleton controller that can adapt its augmentation to the non-steady-state walking conditions of the real world. The chapter has been published in the journal *IEEE Transactions on Robotics* [116].

5.2 Introduction

Robotic exoskeletons may someday allow us to overcome the limits of our natural bodies. Emerging lower-limb exoskeletons are capable of providing assistive joint torques to help their wearers walk and carry loads with promising outcomes, including reduced metabolic cost [123, 202, 42, 132, 150, 122] and muscular effort [62, 101, 206]. Most research to date has focused on steady-state locomotion in a controlled laboratory setting where the task and phase rate (rate of continuous progression through the gait cycle) are nearly constant. This regulated environment makes it easier to design control strategies that deliver appropriate torque assistance in synchrony with the user's gait. However, control strategies based on these assumptions perform poorly outside of the laboratory, where environments are uncertain and locomotion is highly non-steady and

transitory. In order for the field to study biomechanical outcomes outside of the laboratory, new control strategies are needed that explicitly account for continuously varying task and phase.

During steady-state locomotion in controlled laboratory settings, phase progression can be reasonably predicted using time normalized by the stride period. The stride period is usually estimated as the time between subsequent ipsilateral heel strike (HS) events, which can reasonably predict the next HS event during steady locomotion. This 'timing-based' approach is quite effective and widely used for controlling exoskeletons on treadmills [180]. Typically this phase estimate parameterizes a pre-defined torque profile to deliver real-time assistance through the exoskeleton's actuator(s). Recent work has demonstrated impressive reductions in the metabolic rate of level-ground treadmill walking by optimizing this torque profile in real-time [202, 42]. While this paradigm of timing-based estimation (TBE) for torque control works well in steady-state locomotion, it is not designed for more practical conditions outside the laboratory where both the periodicity of gait and the task can change sporadically.

Recent work has addressed the problem of *non-constant* phase progression in powered prostheses and orthoses by introducing phase-based[1] controllers [75, 140, 141, 88, 89, 157, 76, 200, 158], which continuously adjust the rate of phase progression to accommodate dynamic speed changes mid-stride. In contrast to the timing-based approach, these controllers estimate the gait phase online using the robot's sensors. This phase can be estimated from shank [75] or thigh motion [140]. Alternatively, Thatte *et al.* recently introduced an Extended Kalman Filter (EKF) to estimate phase and its derivative (phase rate) without relying on any particular sinusoidal pattern in the sensors [173]. This EKF was also shown to yield more accurate phase estimates in non-steady locomotion when compared to conventional timing- and EMG-based controllers. Building on this result, we pursue further improvements in estimation using the Kalman Filter framework by incorporating other variations in task beyond speed.

To handle transitions between tasks, several groups have proposed solutions rooted in machine learning classifiers. Such classifiers can detect the human's intended task from patterns in the exoskeleton's sensor signals to apply the correct task-specific controller (*e.g.*, stair ascent vs. level-ground walking) [180, 80, 197, 86, 196, 106, 78, 89, 158]. While this approach can identify discrete changes in task, it is less ideal for detecting continuous variations within a family of tasks or handling tasks outside the training data. Recently developed gait models have introduced task variables such as ground slope or stair height that continuously parameterize the instantaneous task [47, 46]. These variables capture more of the task's features, and can thus provide more tailored assistance to the user. Controllers based on these models have been limited to measuring the task

[1] We define 'phase-based' approaches as those where the rate of phase progression can change continuously within a stride, whereas 'timing-based' approaches have a fixed rate over the stride. Note that both would fall under the category of 'phase-based' approaches according to [180].

parameter only once per stride [46]—similar to the rate of phase progression before phase-based controllers were introduced. A notable exception to this paradigm is the work of Holgate *et al.* [75], which exploited the relation in the phase plane between tibia angle and angular velocity to continuously estimate gait phase and stride length. However, this relation does not hold for non-steady-state walking, nor does it extend to other joints or task variables (*e.g.*, ground inclination). Recent work has also combined ambulation mode classification with continuous task variable estimates for ramp incline, step height, and walking speed [24], but this approach uses multiple EMG electrodes, IMUs, and goniometers that may not all be available onboard an exoskeleton.

This paper introduces an EKF-based exoskeleton controller that continuously learns both the phase state (phase and phase rate) and task state (ramp and stride length) to modulate the torque profile of an assistive ankle exoskeleton in a biomimetic fashion. The EKF can indirectly estimate the gait and task parameters in real time using onboard sensors, letting the controller adapt its output quickly and in response to a continuously-varying environment. The contributions of our work include 1) introducing a new EKF phase estimator that also estimates task parameters in continuous time, 2) validating the quality of the state vector estimates using leave-one-out cross-validation based on previously-collected motion capture data of 10 able-bodied subjects walking on various inclines at various speeds, 3) validating the EKF estimates on an ankle exoskeleton used by another 10 able-bodied subjects on an instrumented treadmill which can vary speed and inclination, and 4) validating the EKF-based controller on an ankle exoskeleton during outdoor free-walking on continuously varying surfaces (the Michigan Mars Yard and the Michigan Wavefield). These contributions to exoskeleton control enable practical, real-world usage of exoskeletons that adapt their assistance during walking at non-steady-state conditions within a continuously evolving task.

5.3 Modelling and Estimating Gait

5.3.1 Gait Model

Our ankle exoskeleton controller is based on a biomechanical model of gait kinematics. This model predicts global shank angle θ_s, global foot angle θ_f (see Fig. 5.1), forward heel position p_f, and upward heel position p_u. As input, the model takes in a gait-state vector x, comprising a phase (or normalized time) signal p, its time derivative $\dot{p}$, a stride length signal l, and a ramp angle signal r. The phase variable ranges from 0 to 1 and increases monotonically throughout strides, resetting at ipsilateral heel-strikes. We denote this kinematic model,

$$\begin{pmatrix} \theta_s(t) & \theta_f(t) & p_f(t) & p_u(t) \end{pmatrix}^T = h_{\text{gait}}(x(t)). \tag{5.1}$$

This gait model is used to infer the gait-state vector from the measurable quantities in real-time using the framework of the Extended Kalman Filter. Using the gait-state estimate, the controller then applies the corresponding bio-mechanical torque using a second model, and this allows torque profiles to vary continuously with incline angle and stride length. We use global angles and foot positions, as opposed to joint angles, because of the convenient relationship between global foot angle and ramp inclination during stance, and because they can be either measured directly or estimated with IMUs.

5.3.1.1 Constrained Least-squares Regression

The model $h_{\text{gait}}(x)$ is based on labeled training data from a 10-subject able-bodied dataset [47]. This dataset contains individual stride walking data at 27 combinations of three speeds (0.8, 1, and 1.2 m/s) over nine slopes (-10 to 10 deg in increments of 2.5 deg), which were collected on a Bertec instrumented treadmill in a laboratory environment using Vicon motion capture. Each stride features 150 samples of kinematic and kinetic data, from which we calculated phase progression over the stride. Thus the dataset provides labeled tuples of $(\theta_s(t), \theta_f(t), p_f(t), p_u(t), x(t))$ for all (>25,000) individual strides.

We structure $h_{\text{gait}}(x)$ as

$$h_{\text{gait}}(x) = \phi^T R^T(x), \tag{5.2}$$

where $\phi \in \mathbf{R}^{160\times 4}$ is a matrix of real-valued model parameters and $R : \mathbf{R}^4 \mapsto \mathbf{R}^{1\times 160}$ is a gait-state-dependent regressor row-vector. The parameters ϕ are chosen to minimize the sum squared error for the equation

$$\begin{pmatrix}\theta_s(t) & \theta_f(t) & p_f(t) & p_u(t)\end{pmatrix} = R(x(t))\phi, \tag{5.3}$$

over all the sets of $(\theta_s(t), \theta_f(t), p_f(t), p_u(t), x(t))$ in the dataset.

The regressor $R(x)$ heavily uses the Kronecker product, $\otimes$, in its construction. The Kronecker product of row-vectors $A \in \mathbf{R}^{1\times N}$ and $B \in \mathbf{R}^{1\times M}$, denoted $A \otimes B \in \mathbf{R}^{1\times NM}$, is the block row-vector $(a_1 B \;\; a_2 B \;\; \cdots \;\; a_N B)$. For matrices $A \in \mathbf{R}^{n\times N},\ B \in \mathbf{R}^{m\times M}$, this generalizes to

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1N}B \\ a_{21}B & a_{22}B & \cdots & a_{2N}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nN}B \end{pmatrix} \in \mathbf{R}^{nm\times NM}. \tag{5.4}$$

To give a relevant example, letting $\Lambda_1(x_1) = [1,\ x_1]$ and $\Lambda_2(x_2) = [1,\ x_2]$, the combined model, $\Lambda_1(x_1) \otimes \Lambda(x_2) = [1, x_2, x_1, x_1 x_2]$, represents a basis for functions which depend linearly on both x_1 and x_2.

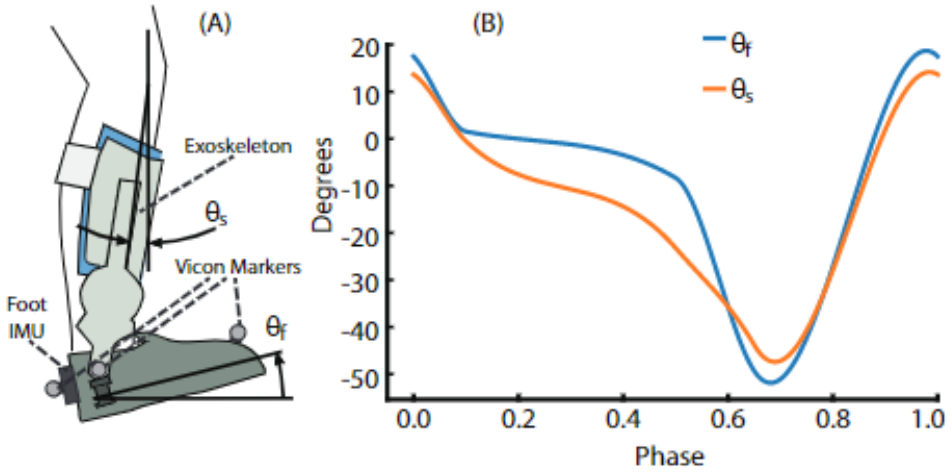

Figure 5.1: (A) A sagittal-plane view of the human leg. The foot angle θ_f is defined from the global horizontal, and the shank angle θ_s is defined from the global vertical. The Vicon markers used in the experiment were placed on the heel, the front of the foot (on the second metatarsal), and the ankle joint. The dedicated foot IMU was also placed on the back of the heel, while the exoskeleton system contained the IMU used for shank angle estimation. (B) Average θ_f and θ_s profiles over the gait cycle at 1 m/s, zero incline.

Using a series of Kronecker products, we define the regressor

$$R(x) = \Lambda_r(r) \otimes \Lambda_l(l) \otimes \Lambda_p(p), \tag{5.5}$$

which combines the effects of the four simpler behaviors such that the final model depends on p, l, and r. The components are:

- The ramp angle basis $\Lambda_r : \mathbf{R} \mapsto \mathbf{R}^{1\times 2}$ is a first-order polynomial Bernstein basis [192] in ramp angle,
$$\Lambda_r(r) = \begin{pmatrix} r & (1-r) \end{pmatrix}, \tag{5.6}$$
which allows for continuous adjustment to ground slope.

- The stride length basis $\Lambda_l : \mathbf{R} \mapsto \mathbf{R}^{1\times 2}$ is another first-order Bernstein polynomial basis in stride length,
$$\Lambda_l(l) = \begin{pmatrix} l & (1-l) \end{pmatrix}, \tag{5.7}$$
which similarly allows for kinematic changes associated with step length.

- Finally, the phase-polynomial basis $\Lambda_p : \mathbf{R} \mapsto \mathbf{R}^{1\times 2N}$ is a Fourier series basis of order N,

defined as

$$\Lambda_p(p) = \begin{pmatrix} 1, \cos(1 \cdot 2\pi p), \sin(1 \cdot 2\pi p), \dots \\ \cos(N \cdot 2\pi p), \sin(N \cdot 2\pi p) \end{pmatrix}, \tag{5.8}$$

where $N = 20$.

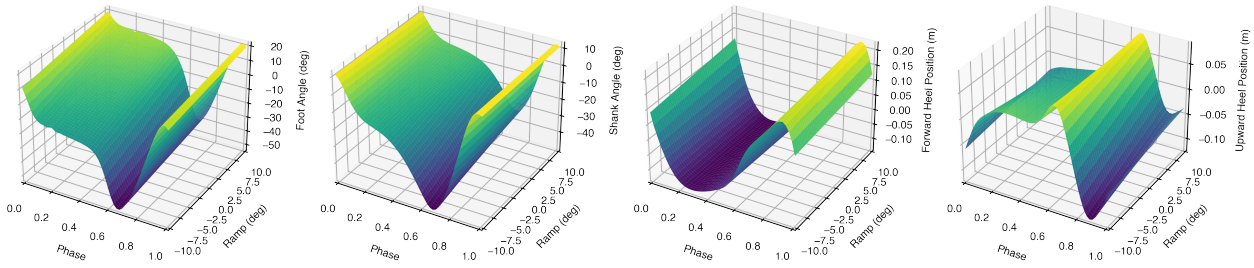

Figure 5.2: The regressed continuous gait models for θ_f, θ_s, p_f, and p_u (foot angle, shank angle, forward heel position, and upward heel position respectively). As the models themselves depend on three variables (p, l, and r) and each produce an output, they fully reside in 4D-space and are thus difficult to express in 3D-space. In this figure, the model relation between phase and ramp is shown (with stride length constant at 1 meter). Stride length merely changes the amplitudes of the gait model.

5.3.1.2 Least-squares Constraints

The elements of the parameter matrix ϕ are subject to constraints which ensure the resulting gait model $h_{\text{gait}}(x)$ is well-behaved in two ways. The constraints guarantee that the $h_{\text{gait}}(x)$ function 1) predicts constant kinematics if stride length is zero (*e.g.*, if the person is stationary, the measurements do not change with respect to phase), and 2) predicts the global foot angle is equal to the ramp angle when stride length is zero or when phase takes the value 0.2.

We first define a set of two constraints to ensure constant-with-phase behavior when stride length is zero, *i.e.*, when the person is standing still. The first constraint concerns the sinusoidal phase terms in the case case where stride length l is zero. In this case, we expect global foot angle θ_f, global shank angle θ_s, forward heel position p_f, and upward heel position p_u to have zero coefficients for the sinusoidal terms, which we express as a matrix equality,

$$\overbrace{I_2 \otimes \Lambda_l(0)}^{\forall r,\ \text{if } l=0,} \otimes \overbrace{\begin{pmatrix} 0 & I_{2N} \end{pmatrix}}^{\text{sinusoid terms}} \phi = \overbrace{0 \in \mathbf{R}^{4N \times 4}}^{\text{are zero}}. \tag{5.9}$$

To finish our pair of constraints for the case where the human is standing still (l=0), we consider the constant terms. In this case, the shank is vertical and the foot is aligned with the ramp, while the

heel positions are assumed to be at zero position. We choose—without loss of generality—$r = 0$ and $r = 10$ to express this constraint

$$\overbrace{\begin{pmatrix}\Lambda_r(0)\\ \Lambda_r(10)\end{pmatrix}}^{\text{if } r=0,10} \otimes \overbrace{\Lambda_l(0)}^{\text{and } l=0,} \otimes \overbrace{\begin{pmatrix}1 & 0_{1\times 2N}\end{pmatrix}}^{\text{constant terms}} \phi = \overbrace{\begin{pmatrix}0 & 0 & 0 & 0\\ 10 & 0 & 0 & 0\end{pmatrix}}^{\text{match priors}}. \tag{5.10}$$

Next, we constrain the model to predict that, regardless of stride length, the foot angle will be equal to r at $p = 0.2$ to represent flat-foot contact. This modeling choice enforces equivalency between the foot angle and the ground incline, which allows for better estimation of incline [118]. We express this constraint on the foot using the following equality:

$$\left[\overbrace{\Lambda_{\forall r}}^{\forall\, r} \otimes \overbrace{\Lambda_{\forall l}}^{\forall\, l} \otimes \overbrace{\Lambda_p(0.2)}^{\text{at } p=0.2}\right] \phi \overbrace{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}}^{\text{selecting } \theta_f} = \overbrace{\begin{pmatrix}0\\0\\10\\10\end{pmatrix}}^{\theta_f = r}, \tag{5.11}$$

where

$$\Lambda_{\forall r} = \begin{pmatrix}\Lambda_r(0)\\ \Lambda_r(10)\end{pmatrix}, \quad \text{and} \quad \Lambda_{\forall l} = \begin{pmatrix}\Lambda_l(0)\\ \Lambda_l(1)\end{pmatrix}, \tag{5.12}$$

are matrices in $\mathbf{R}^{2\times 2}$ that expand the constraint to affect all values of ramp and stride length. The values $r = 0$, $r = 10$, $l = 0$, and $l = 1$ are again chosen without loss of generality to constrain the entirety of these linear functions.

5.3.1.3 Complete Gait Model

We performed the regressions for the foot and shank angle models using the constrained least-squares optimization function `lsqlin` in MATLAB. The resulting models (Fig. 5.2) not only described how the measured kinematics varied with phase, but also with ramp and stride length. Least squares regression was used to fit a Fourier series modeling gait's cyclic relationship with phase, while also allowing simpler linear fits to the other task variables that capture the salient features of gait [118].

5.3.2 Biomimetic Exoskeleton Torque Profile

During the live trials, the exoskeleton provided torque assistance according to a profile parameterized by phase, stride length, and incline angle. We regressed this biomimetic torque model (Fig.

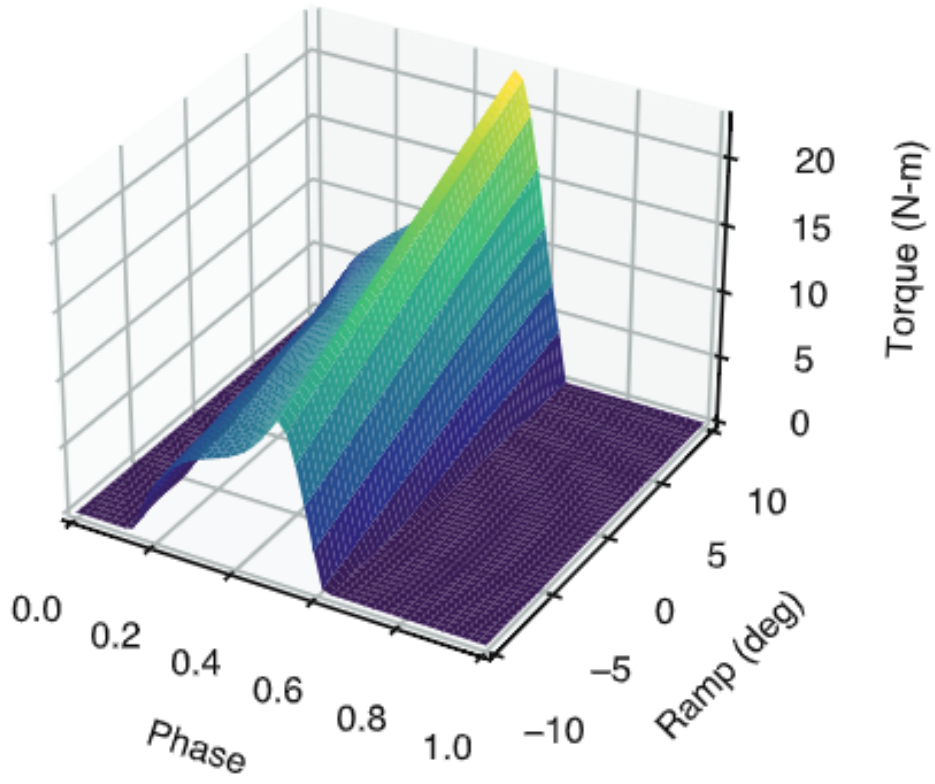

Figure 5.3: The regressed continuous biomimetic torque profile. Like the continuous gait model, the full torque profile depends on p, l, and r, and thus fully resides in 4D-space. The model relation between phase and ramp is shown (with stride length constant at 1 meter) in this figure for the purposes of visualization.

5.3) using scaled-down biological ankle torques (by a factor of 5) from the dataset [47]. The regression was performed using a near-identical regressor structure as in (4) without the constraints on stride length and incline and with the biological ankle torque in lieu of the measured kinematics. This yielded a torque profile encoded with the same structure as the gait model in (3). As the exoskeletons used in this experiment only applied plantarflexion torque, dorsiflexion torques were floored at zero for the biomimetic torque profile.

5.3.3 Dynamic Model and State Estimator

5.3.3.1 Primary Phase EKF Model

We use the standard EKF equations given in Appendix B. We define the state vector as the gait state in (2). The system process is encoded using the state transition matrix F:

$$F = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.13}$$

such that phase is updated by simple numerical integration of phase rate using the time stride Δt.

The state covariance matrix is initialized as $P = 1 \times 10^{-3} \cdot I_{4\times4}$. We define our process noise matrix Σ_Q as a diagonal matrix $\mathrm{diag}[0, \sigma_{22}^2, \sigma_{33}^2, \sigma_{44}^2] \cdot \Delta t$, where σ_{22}, σ_{33}, and σ_{44} are the standard deviations for $\dot{p}$, l, and r, respectively. Phase p has no process noise since it is defined using a noiseless integration of $\dot{p}$. The diagonal variances act as tunable parameters that modulate EKF performance; we empirically tuned the performance and found that $\sigma_{22} = 6\times10^{-4}$, $\sigma_{33} = 9\times10^{-4}$, and $\sigma_{44} = 6 \times 10^{-3}$ yielded good performance with respect to phase tracking and response time. These values were held constant for all subjects.

5.3.3.2 Measurement Model

Within the update stride of the EKF, our observation function $h(x)$ extends the directly measurable variables θ_f, θ_s, p_f and p_u. To encode time-dependent measurement information, we also model the velocity of the foot $\dot{\theta}_f$ and shank $\dot{\theta}_s$. These velocities are defined using the differentiation chain rule:

$$\begin{bmatrix} \dot{\theta}_f \\ \dot{\theta}_s \end{bmatrix} = \begin{bmatrix} \frac{\partial \theta_f}{\partial t} \\ \frac{\partial \theta_s}{\partial t} \end{bmatrix} = \begin{bmatrix} \frac{\partial \theta_f}{\partial p} \\ \frac{\partial \theta_s}{\partial p} \end{bmatrix} \dot{p} \tag{5.14}$$

where $\dot{p}$ is the estimate of the phase rate from the prediction stride and the partial derivatives of θ_f and θ_s are available analytically. The observation function is then $h(x) = \Big[\theta_f, \dot{\theta}_f, \theta_s, \dot{\theta}_s, p_f, p_u\Big]^T$, where θ_f, θ_s, p_f, and p_u, are available from the gait model $h_{\text{gait}}(x)$.

5.3.3.3 Nonlinear Stride Length Transformation

We choose to apply a nonlinear transformation to the stride length state. This transformation encodes the upper limit on a person's stride length. Furthermore, we model the smallest possible stride length as 0, which encodes backwards walking as having positive stride lengths and negative phase rates. In this transformation, the stride length is the output of an arctangent transformation

[184], in which the 'pseudo-stride length' l_p is input. Additionally, in our gait model regression, the stride lengths were normalized by participant leg length L. As part of the non-linear transformation, we denormalize by participant leg length to obtain stride length l in meters. The nonlinear transformation is defined as:

$$l(l_p) = L\left(\frac{4}{\pi}\text{atan}\left(\frac{\pi}{4}l_p\right) + 2\right). \tag{5.15}$$

This allows a maximum normalized stride length of 4 leg lengths and floors it at 0. l_p is then the state estimated by the EKF and contained in state vector x. Similarly, the gait model $h_{\text{gait}}(x)$ takes as input l/L instead of l. However, for ease of communication, we refer to x as containing stride length l rather than its 'pseudo', denormalized counterpart, and the gait model as taking the stride length input directly. To account for this change in the Jacobian H in the update step of the EKF, we pre-multiply all partial derivatives with respect to l by $\frac{\partial l}{\partial l_p}$.

5.3.3.4 Heteroscedastic Noise Model

EKFs generally encode measurement noise in a *constant* Σ_R matrix, which typically denotes how trustworthy the sensors used are. However, this constant model is unable to selectively change the trust in the measurements during regions of the state space where those measurements are known to be informative. For example, we expect that for phase values corresponding to flat-foot contact during locomotion, the measurements of foot angle will be highly informative for the ramp angle, given the position constraint of foot contact. To improve the performance of our phase EKF controller, we implemented a *heteroscedastic* measurement noise matrix, that can continuously change the measurement noise matrix Σ_R defined as follows:

$$\Sigma_R(p) = \Sigma_{R,\text{sensor}} + \Sigma_{R,\text{xsub}}(p). \tag{5.16}$$

In our measurement noise model, $\Sigma_{R,\text{sensor}}$ is the conventional noise matrix that denotes how uncertain the sensors are, and $\Sigma_{R,\text{xsub}}(p)$ represents the uncertainty present due to inter-subject gait kinematic variability (subscript xsub for cross-subject). $\Sigma_{R,\text{xsub}}(p)$ captures the regions within the gait cycle where measurements are more informative due to lower inter-subject variability (Fig. 5.4). To determine $\Sigma_{R,\text{xsub}}(p)$, we first calculated the covariance matrices of the residuals y between the measurements in the prior dataset and the regressed gait model at each of the dataset's 150 phase values. These covariances incorporated the full stride data from all ten subjects in the dataset across all the dataset's conditions. In real-time, these 150 matrices were used to map the real-time phase estimate to a corresponding covariance $\Sigma_{R,\text{xsub}}(p)$ via linear interpolation.

The matrix $\Sigma_{R,\text{sensor}}$ was defined as $\text{diag}[\sigma^2_{11,r}, \sigma^2_{22,r}, \sigma^2_{33,r}, \sigma^2_{44,r}, \sigma^2_{55,r}, \sigma^2_{66,r}]$, with each $\sigma_{xx,r}$

representing the standard deviation for $\theta_f, \dot{\theta}_f, \theta_s, \dot{\theta}_s, p_f, p_u$, respectively. In our implementation, these values were set to 1, 10, 7, 20, 0.01, and 0.08, respectively. These different values reflected the different measurement units (*e.g.*, deg vs. deg/s) and the different levels of sensor noise present in the measurements. For example, the shank angle was estimated with a secondary attitude EKF that updated at 100 Hz, whereas the foot angle was measured directly using a dedicated IMU that updated at 1 kHz, so we placed greater trust in the foot angle measurements.

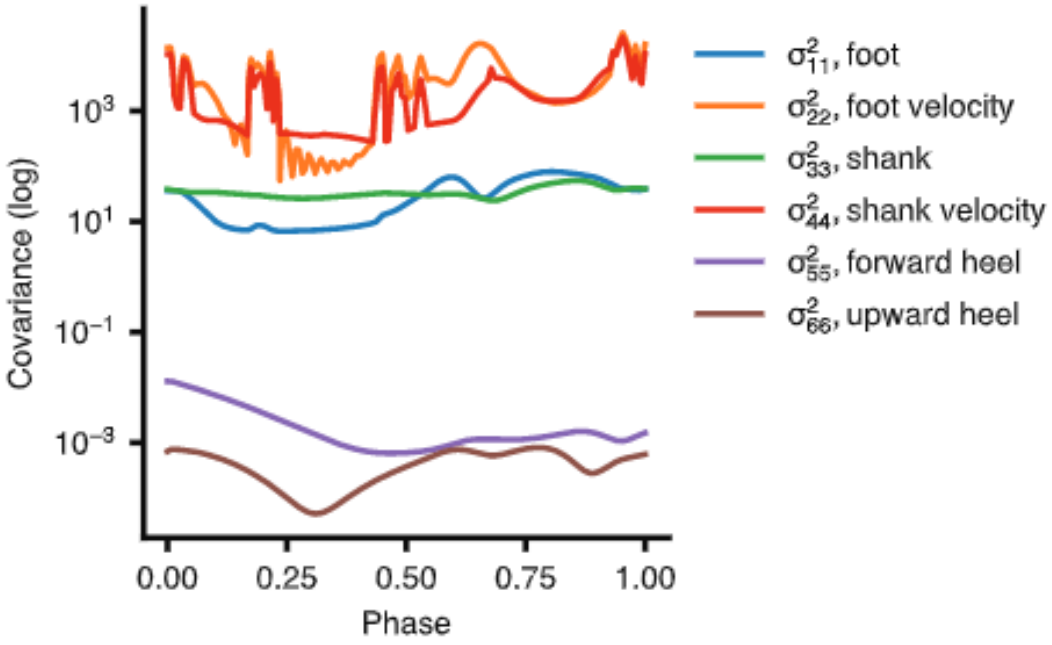

Figure 5.4: The heteroscedastic measurement noise model that modulates the measurement noise matrix as a function of phase. Foot angle variance σ_{11} (*i.e.*, the covariance with itself), foot angular velocity variance σ_{22}, shank angle variance σ_{33}, shank angular velocity variance σ_{44}, forward heel position variance σ_{55}, and upward heel position variance σ_{66}, are shown. For ease of viewing, the covariances between these measurements are not shown.

5.3.4 Heelstrike-based Estimation Backup

Depending on the process and measurement covariance matrices, it is possible that the EKF may 'get lost' in its state estimation and fail to track the actual states, for example, if the filter's effective bandwidth is too slow to respond to the measurements, or if it is too fast and thus sensitive to noise. To protect the EKF controller against getting lost in its estimation, we introduced a backup process that was permitted to reset the estimator state in the event of sub-par phase estimation. This backup system fuses a similar approach to the conventional "Timing Based Estimation" (TBE, [110]), which detects heel strikes and records the timestamps for each heel strike event, with the

EKF framework. Our backup plan is essentially a 'smaller' EKF that updates its phase rate estimates once per heel-strike, but estimates stride length and incline in real-time using the same measurements as the primary EKF.

If the backup describes a better fit to the subject's kinematics than the EKF, the backup resets the EKF state vector x with its own backup state vector. The backup system computes the sum of squares of the residuals (SSR) using the residual vector $\bar{y}$ each stride at heel-strike. It then compares its SSR to the SSR from the phase EKF (calculated using y from the EKF) over that stride. If the SSR from the backup is sufficiently smaller than the SSR from the EKF, *i.e.*, the EKF is performing poorly, then the backup overrides the EKF and places it back on course. We tuned the backup EKF to only activate once every 100 steps when the SSRs are of comparable magnitude.

5.4 Study Methods

We perform data-driven simulations and hardware experiments to address the following hypotheses: H1) Our EKF-based controller has a significantly lower phase RMS error compared to phase estimates from a HS-to-HS timing based controller (state-of-the-art) in a leave-one-out cross validation; H2) The inclusion of task variables significantly improves the phase RMS error in a leave-one-out cross validation; and H3) Our real-time EKF estimate of phase has an RMS error less than that of state-of-the-art timing based estimators in the presence of actuator torques. We also present the result of the EKF working in a practical outdoors setting with exoskeleton actuation. All human participants gave written, informed consent with approval from the University of Michigan Institutional Review Board (Study ID: HUM00158854).

5.4.1 EKF Simulation

We cross-validated our EKF *in silico* using the same walking dataset mentioned in Sec. 5.3.1.1 (N=10 subjects [47]), which includes trials at three speeds (0.8, 1, and 1.2 m/s) over nine slopes between -10 and 10 deg. The dataset contains walking data indexed by walking stride, which was used for training the gait model $h_{\text{gait}}(x)$. This stride data was concatenated to form a continuous walking sequence, which was then input into our EKF to analyze performance with realistic locomotion. The complete source code for our simulation is available in a ready-to-run computation capsule format through CodeOcean [119].

5.4.2 Cross-Validation of Simulated Phase Estimate Quality (H1)

To identify the ability of our EKF to adapt to new subjects while accurately estimating phase, we performed leave-one-out cross-validation on our EKF-based controller using the dataset, where ground truth phase was calculated using the normalized time between heel-strikes. For each subject, this cross-validation trained a new gait model using the individual stride data from the remaining nine subjects. Similarly, we also trained a new heteroscedastic model for each subject with the residual covariance data from the remaining nine subjects. We then used the concatenated stride walking data from the subject not included in the training set as the input to our EKF simulator, and computed the phase root-mean-squared-error (RMSE) of our EKF relative to the ground truth phase measurements for each stride in the dataset for that subject. To determine the improvement produced by our EKF controller, we compared the stride-wise phase errors from the EKF against the phase errors from a simulated TBE, which predicts the current phase in real-time using the normalized timings of previous heel strikes. This common method is causal whereas ground truth phase is non-causal. The TBE estimator was assumed to perfectly detect all heel-strikes from the ground truth dataset (an assumption that may not hold during practical implementations). Using a two-tailed, paired t-test, we compared differences in the stride-wise phase RMSE between the EKF and TBE approaches. We also computed the stride-wise stride length and incline RMSE relative to the ground truth values from the dataset. To assess how the resulting control torques were affected by the gait state estimates from the EKF, we computed the stride-wise RMSE between the torque profile resulting from the EKF state estimates and the torque profile resulting from the ground truth states.

5.4.3 Cross-Validation without Task Variable Estimation (H2)

In a separate leave-one-out cross-validation simulation, we assess the improvement to phase estimation quality that comes from the inclusion of the stride length and incline task variables in the state vector. This allows for a more direct comparison to the EKF featured by Thatte *et al.* [173], which only used phase variables in their filter. To effectively eliminate the stride length and ramp state variables, elements of the process noise matrix Σ_Q and covariance matrix P corresponding to these variables were initialized to extremely low values on the order of $1e{-}12$, which prevented the estimates from changing in real time. These values were not simply set to zero in order to avoid numerical issues in the real-time computation of the state estimate. We then computed the phase RMSEs for each stride using the phase estimates of this limited EKF and determined the significance of including the task variables by comparing these errors to those from the proposed EKF using a two-tailed, paired t-test.

5.4.4 Hardware Setup

We implemented our phase EKF on an ankle exoskeleton system (ExoBoot, Dephy, Inc., Maynard, MA) for our human participant experiments. We used separate sensors to measure each of the global angles due to the compliance of the exoskeleton's structure. We estimated the global shank angle using a custom attitude EKF [87] that ran concurrently with our phase EKF. The attitude EKF used readings from the onboard IMU (TDK Invensense MPU-9250, San Jose) mounted above the ankle joint of the exoskeleton. To measure global foot angle, we used an external IMU (3DM-GX5-25, LORD Microstrain, Williston) placed on the boot of the exoskeleton. This IMU estimated global orientation from which we obtained the global foot angle (*i.e.*, pitch angle). To account for inter-subject differences in exoskeleton position and resting shank angle, we calibrated both of these sensors prior to each trial by having the subjects stand still and recording the average angular measurements from these sensors, which were then subtracted from the live measurements. For the angular velocity measurements, we measured the shank angular velocity directly using the exoskeleton's gyroscope readings, and measured foot angular velocity using the foot IMU's gyroscope readings. To obtain the heel position signals, we applied a second-order linear filter to the foot IMU's accelerometer signals, which combined a double-integrator with a high-pass filter to convert forward and vertical acceleration signals into forward and vertical heel positions while also high-pass filtering away the effects of integration drift and gravity.

For the indoor hardware experiment, ten able-bodied, human subjects walked on an instrumented Bertec treadmill (Bertec, Columbus OH) with variable belt speed and inclination. The outdoor experiments were conducted with a single able-bodied human subject on both the University of Michigan Mars Yard (GPS: 42.29464, -83.70898), which demonstrates an unstructured outdoor environment, and the University of Michigan Wavefield (GPS: 42.29323, -83.71168), which provides a practical stress test on extremely unsteady terrain.

5.4.5 Treadmill Experiment (H3)

We characterized our controller's performance during a controlled treadmill experiment. In this experiment, ten able-bodied participants (5 male, 5 female, age: 24.2 ± 2.15 years, height: 1.75 ± 0.08 m, weight: 74.4 ± 14.1 kg) walked with the EKF-controlled exoskeleton for three segments. In the first segment, participants continuously walked on level ground for twenty seconds at 1.2 m/s, followed by twenty seconds at 0.8 m/s, followed by a return to 1.2 m/s for twenty seconds, then finally for 0.8 m/s for a final twenty seconds. In the second segment, the participants walked on the treadmill for 10 seconds at 1 m/s, then the treadmill inclined over a period of one minute and ten seconds to a maximum of 10 deg while the participant continued walking. The third segment was identical to the second, except that instead of the treadmill inclining to 10 deg, it declined

to -10 deg. The exoskeleton applied the adaptive biomimetic torque throughout the trials via PID current control to apply the desired torque at the ankle.

Participants were instrumented with Vicon markers to capture their kinematics, which were then used to establish ground truth estimates for phase, phase rate, stride length, and inclination, as well as for heel-strike events. Ground truth phase rate for each stride was estimated by taking the inverse of that stride's duration; ground truth phase was then calculated by integrating this rate over the duration of the stride beginning at heel-strike. Ground truth stride length was then calculated by dividing the known speed by the phase rate, while ground truth inclination was measured as the inclination of the treadmill. The Vicon experimental system featured 27 cameras and collected kinematic data at 100 Hz. We computed the errors and RMSEs for both EKF phase, phase rate, stride length and incline. We then compared the stride-wise phase RMSEs to those from a state-of-the-art TBE approach which thresholded the signals from the onboard device sensors to detect heel-strikes (similar to the approach from [122]); the thresholds were tuned to correctly detect heel-strikes across the walking tasks in this treadmill trial. We compared the two controllers using the same t-test as with the previous hypotheses.

5.4.6 Real-world Experiments

Finally, we characterized the ability of the EKF-based controller to adapt its gait state estimates and torque assistance in a highly irregular outdoor environment. One participant from H3 walked with the EKF-controlled exoskeleton on both the Michigan Mars Yard, which features a steep hill (approximately 20 deg) covered in small rocks, and the Michigan Wavefield, which features a sinusoidal hill pattern of inclines and declines (from approximately 30 to -30 deg) and has been previously used to test the stability of bipedal robots [37]. These locations were chosen to 'stress-test' the EKF with non-steady-state conditions outside its training dataset. During these tests, the process noise matrix Σ_Q was tuned to have $\sigma_{22} = 1e{-}3$, $\sigma_{33} = 2e{-}3$, and $\sigma_{44} = 5e{-}2$ for a faster filter response. At the Mars Yard, the participant walked in an alternating 'slow-fast-slow' speed pattern on level ground both prior to and after ascending/descending the hill. At the Wavefield, the participant performed the same alternating speed pattern before and after walking on the rapidly changing, sinusoidal hills. The separation of the speed and slope changes during these outdoor trials was done to isolate the EKF's ability to track each task variable independently. We recorded HD videos of these trial and compared the recorded events to the EKF's estimates of the gait state to assess the adaptations to the terrain features of the Mars Yard and Wavefield. To obtain ground truth phase values, we analyzed the video to determine the timings of heel-strike events, and interpolated phase values between them (assuming a constant phase rate). For this experiment, we compared our EKF against a TBE identical to the one in Sec. 5.4.2 (recall TBE is

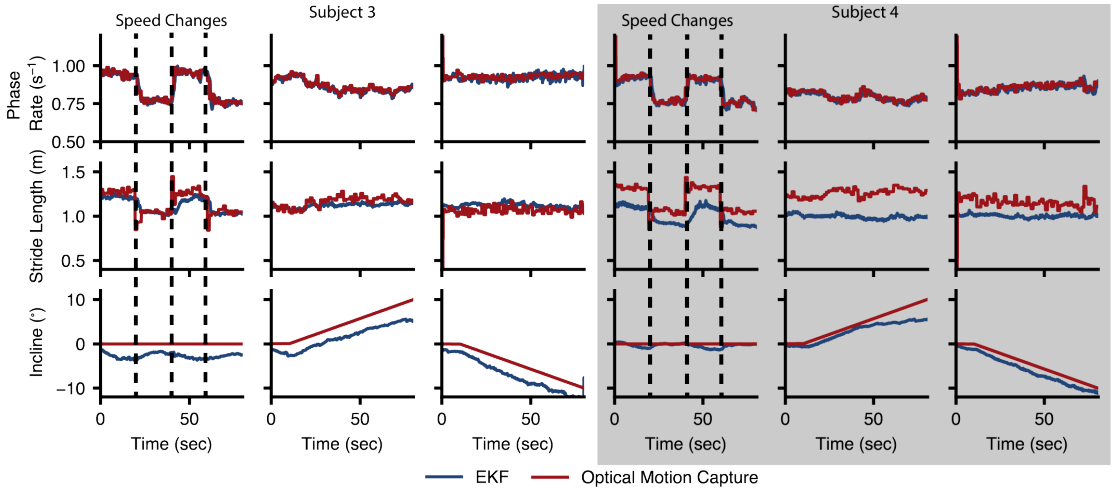

Figure 5.5: State estimates from the treadmill trials from two representative participants. Phase rate is shown in the first row, stride length in the second, and incline in the third. For clarity, the periods where the speed changed in the first trial are delineated by the black dashed lines. Overall, the EKF had excellent tracking of phase rate (and thus, phase), as evidenced in these representative trials. The EKF was responsive enough to correctly estimate Subject AB03's stride length but tended to underestimate this state for some participants (such as Subject AB04), in part due to individual differences from the trained gait model. Finally, the EKF adequately tracked ground incline, although it occasionally exhibited a negative bias due to changes in gait caused by the exoskeleton assistance (more obvious in Subject AB03).

causal whereas ground truth is not). We performed pairwise t-tests comparing the phase RMSEs calculated over each stride between the EKF and this TBE.

5.5 Results

5.5.1 Cross-Validations of EKF Performance (H1 and H2)

Our results support Hypothesis 1—that the EKF significantly outperforms TBE during an *in silico* leave-one-out cross validation. In terms of % of stride cycle progression, the inter-subject phase RMSE for the EKF ($1.6 \pm 0.31\%$) was significantly lower ($P < 0.001$) than the inter-subject RMSE for the TBE estimator ($2.1 \pm 0.1\%$). Additionally, our EKF predicted stride length with an inter-subject RMSE of 0.10 ± 0.02 m, ramp inclination with an inter-subject RMSE of 1.5 ± 0.23 deg, and desired torque from our biomimetic torque profile with an inter-subject RMSE of 1.2 ± 0.18 Nm. These RMSEs are likely due to inter-person variability in locomotion, which have been established to cause errors in kinematic predictions [46].

Table 5.1: Gait State RMSEs for All Subjects in Treadmill Experiment

Participant	EKF Phase RMSE (%)	EKF Phase Rate RMSE (%)	Stride Length RMSE (m)	Incline RMSE (°)	Desired Torque RMSE (N-m)	TBE Phase RMSE (%)
AB01	4.9	1.5	0.22	1.0	2.9	5.7
AB02	4.1	1.8	0.15	2.7	3.3	5.1
AB03	4.7	2.4	0.07	2.8	2.6	5.5
AB04	3.3	2.7	0.21	1.2	2.4	4.9
AB05	5.6	1.9	0.25	1.8	3.5	5.6
AB06	2.7	2.0	0.28	3.1	2.5	3.4
AB07	11.0	3.0	0.08	2.4	4.9	12.5
AB08	5.4	1.5	0.10	1.7	2.9	5.2
AB09	2.8	1.9	0.12	2.0	2.2	3.5
AB10	3.2	1.8	0.06	5.6	2.8	3.8
Mean (Standard Deviation)	4.8 (2.4)	2.1 (0.5)	0.15 (0.08)	2.4 (1.3)	3.0 (0.76)	5.5 (2.6)

The second cross-validation supported Hypothesis 2—that including task variable estimation in the EKF significantly improves phase estimation when compared to an EKF that does not feature task variable estimation (*e.g.*, [173]). The phase RMSE for the No-Task condition ($2.0 \pm 0.36\%$) was significantly higher than the phase RMSE for the proposed EKF condition ($1.6 \pm 0.31\%$, $P < 0.001$).

5.5.2 Treadmill Experiment (H3)

The treadmill test supported Hypothesis 3—that the EKF significantly outperforms TBE while applying torques. The EKF (phase RMSE: $4.8 \pm 2.4\%$, Table 5.1) resulted in significantly lower ($P < 0.002$) phase errors than the conventional TBE approach (phase RMSE: $5.5 \pm 2.6\%$). The EKF captured the pulse-like change in phase rate during the fast 1.2 m/s and slow 0.8 m/s sections of the trials, as shown in representative trials (Fig. 5.5). The EKF simultaneously provided live estimates of both stride length and inclination (cross-subject average stride length RMSE 0.15 ± 0.08 m, incline RMSE 2.4 ± 1.3 deg). The EKF clearly responded to the changes in ground inclination, although it often exhibited negative biases, as exemplified by the results from Subject AB04 in Fig. 5.5. The EKF underestimated changes in stride length, although with notable differences in tracking quality across subjects (Fig. 5.5). Additionally, the desired torque profile RMSE, which encodes the gait-state RMSEs, was only 3.0 ± 0.76 Nm compared to the $\sim$20 Nm peak magnitude of the torque profiles. From our torque profile, the average (across phase) change in torque from a stride length change of 0.15 m (the average stride length RMSE across subjects) was only 0.12 Nm. For additional context, we also computed the cross-subject stride-wise peak errors for each of the gait states throughout the treadmill trials. The peak phase error was $29.3 \pm 8.2\%$, the peak phase rate error was $41.2 \pm 25.9\%$, the peak stride length error was 0.70 ± 0.19 m, and the peak incline error was 5.0 ± 1.2 deg. These peak errors resulted in a peak torque error of 8.0 ± 1.65 Nm, These errors were predictably higher than the average error, although they occurred for short periods of time and were thus likely quickly recovered from using our approach.

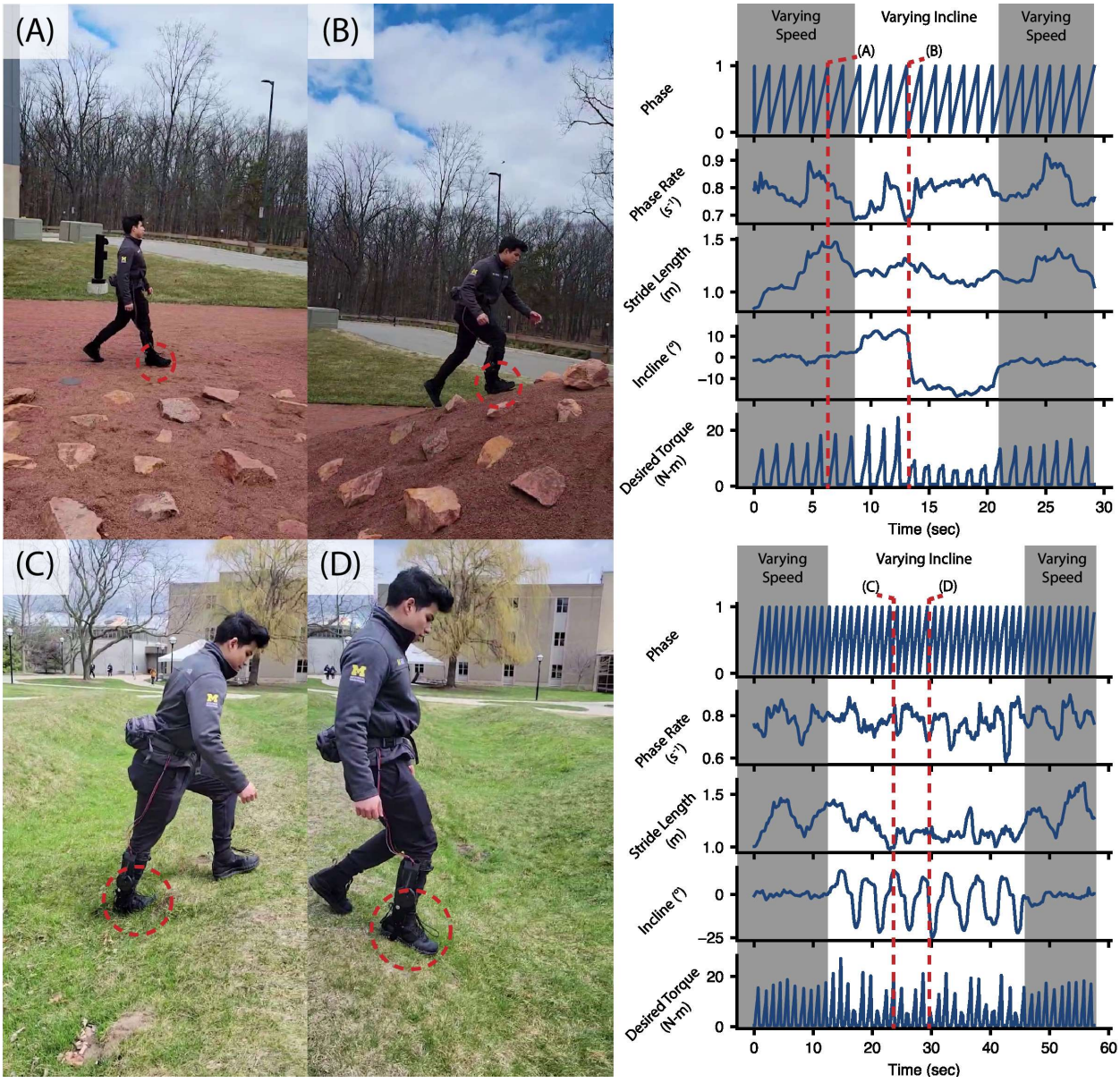

Figure 5.6: Stills from the analyzed videos taken during the Mars Yard and Wavefield trials, with a focus on the EKF's incline estimation performance. The incline estimation is primarily driven by the participant's foot angle on the outdoor slopes (red dashed circles). (A) The subject walking on level ground on the Mars Yard during a period of varying speeds, with the EKF correctly updating its estimate. (B) The subject walking up the Mars Yard, in which the EKF correctly increases its incline estimate. (C) The subject walking uphill on the Wavefield (D) The subject going down a decline on the Wavefield.

5.5.3 Real-world Experiments

In the Mars Yard trial, where the subject alternated speed and incline with exoskeleton assistance, the EKF was comparable to the conventional TBE ($3.8 \pm 2.0\%$ for the former, $4.4 \pm 3.1\%$ for the

latter, $P = 0.2$). However, unlike the TBE, the EKF state estimates tracked both the changing speeds prior to ascending the Hill and the changing ground slope from the Hill itself. Two stills that correspond to notable events from the trial are shown in Fig. 5.6A and B; in particular, the position of the subject's right foot is useful to demonstrate that the EKF is updating its incline estimate correctly. On the Wavefield, the EKF significantly outperformed the conventional TBE ($4.8 \pm 2.7\%$ for the former, $7.3 \pm 8.3\%$ for the latter, $P = 0.02$). The extreme conditions on the Wavefield (Fig. 5.6C and D) rendered the steady-state assumptions of the TBE a poor fit for the gait estimation task; in contrast, the EKF was able to track both phase and the rapidly changing ground incline of the Wavefield. The EKF's adaptation of its state estimates is also reflected in the assistance from the exoskeleton. The biomimetic torque profile broadly increased the magnitude of its torque assistance when ground inclination was positive, and decreased when it was negative. This trend is reflected in the torques shown in Fig. 5.6. A supplemental video of the real-world experiment is available for download.

5.6 Discussion

As expected, our EKF phase estimator outperformed the TBE estimator, with significantly lower phase RMSEs in the *in silico* steady-state treadmill trials. Fundamentally, the EKF observes the behavior between heel-strikes, allowing it to better predict the phase variability that accompanies natural human walking, even in this steady-state test with a constant belt speed, within each tested condition.

The presence of real-time task variable estimation within the EKF also significantly improved phase estimation. This is likely due to these task variables accounting for part of the variation from the kinematic sensor measurements, and therefore reducing the prediction errors. Without estimation of stride length or inclination, modeling error increased, resulting in greater variance in the phase estimation signal. Hence, the task variable states could improve phase estimation during variable-speed and variable-incline walking compared to the phase EKF presented in [173].

In the treadmill experiment, the EKF estimator outperformed the conventional TBE in addition to providing estimates of task variables in real-time. The EKF consistently detected heel-strikes with more accuracy than the conventional approach. Our EKF's phase RMSE (cross-subject average $4.8 \pm 2.4\%$) is comparable to recent online phase estimation performance achieved in hip exoskeletons [89], which used machine learning to estimate phase for subjects walking through circuits that featured different ramps (phase RMSE 5.04%[2]). Additionally, our phase RMSE is

[2]Note that comparing RMSE across distinct experiments is not a direct comparison of methods. We also feature a similar RMSE to that of another machine learning-based estimator that was implemented on the same ankle exoskeleton hardware [158] (3.9% stance phase RMSE).

near the perceptual threshold in timing error, suggesting that subjects may be unlikely to detect any errors in torque profile assistance due to the EKF's phase error [136]. However, it is worth mentioning that the peak error in phase ($29.3 \pm 8.2\%$) far exceeded this threshold, and thus could create the perception that the exoskeleton was out-of-synch with the user, which may diminish trust in the device. The EKF's mean incline RMSE (cross-subject average 2.4 ± 1.3 deg) was also comparable to recent results from an offline ML-based sensor fusion approach [24] (absolute incline error 2.3 ± 0.86 deg).

In the real-world tests, and in particular during the Wavefield stress test, the EKF estimator had a significant advantage over the conventional TBE as in the treadmill trial. The conventional TBE featured its poorest performance during the Wavefield trial (up to 7.3% from 2.1% RMSE in the steady-state cross-validation trials), due to the highly variable nature of the terrain and the non-steady-state steps taken by the participant. The EKF, in contrast, only increased to 4.8% phase RMSE from the steady-state cross validation phase RMSE of 1.6%. These results are comparable to a recent study [139] (phase RMSE: 4.12%) which estimated phase using a combination of IMUs and cameras to control a hip exoskeleton during stair and ramp ascent and descent.

Using an EKF confers several advantages for our adaptive torque controller. Because our simple gait model predicts angles and velocities in a fundamentally structured way, we are able to encode a large, continuous class of locomotion tasks. The EKF equations are intuitive to understand, simple to implement on hardware, and computationally lightweight, while still yielding comparable phase estimation performance to more complex machine learning approaches. Furthermore, the data-driven model component at the core of our controller may be easier to debug than so-called 'black-box' models that directly estimate gait parameters or control commands from sensor data. Similarly, it would be straightforward to substitute another gait model (for example, a Gaussian Process as in [173] or a neural network as in [24]) within our EKF depending on the preference of the experimenter, as the only requirement is that it take in the state vector as inputs and produce the kinematics as outputs.

Our EKF-based controller provides a solution to the challenge of estimating the state of an exoskeleton during practical trials. This challenge has been avoided in prior exoskeleton control work [123, 202, 42, 132] by testing in controlled steady-state conditions that allow TBE approaches to suffice. However, these conditions are unrepresentative of the real world. The performance of our controller throughout the extreme conditions of the Wavefield test demonstrates that it can provide phase estimates throughout non-steady-state conditions, which could potentially allow exoskeleton research to be translated out of the laboratory. Additionally, our explicit gait model is agnostic to the choice of exoskeleton hardware, so our estimator could serve as a 'plug-and-play' solution for researchers seeking to take their devices into the real world.

Our EKF can also scale well to additional sensors. In our current implementation, we use four

sensors (foot and shank angles, along with their derivatives, and forward and upward heel position), but it would be straight-forward to implement other sensors, such as instrumented insoles, to further improve phase estimation. We can use the current dataset to regress relationships between these sensors' measurements and our gait-state and simply extend the measurement vector in our EKF. We expect that this will further improve phase estimation due to the new information available from these sensors, motivating further study into the impact of different sensors (*e.g.*, [118]) as well as different model-fitting techniques to prevent overfitting (which we avoided by implementing a relatively simple linear model structure relative to the complexity of the dataset).

To further improve our estimates in future work, we would first target the inter-person kinematic variability that increases the variance of not only the estimated phase, but also the remaining state vector elements. Prior work has demonstrated that inter-person variability can cause large prediction errors in subject-independent kinematic models [46]. We suspect this is responsible for the large stride length RMSEs experienced by some subjects, such as Subject AB06 (0.26 m RMSE, Tab. 5.1), and for the smaller RMSE in stride length (0.10 ± 0.02 m) in the simulated walking cross-validation. The EKF's estimates of stride length are primarily informed by the amplitude of the forward and upward heel position measurements. When analyzing Subject AB06's forward and upward heel positions (Fig. 5.7A), there is a large discrepancy between the predicted position from the EKF's live gait model (blue), and the predicted position from the gait model when using the ground truth states from Vicon as inputs (green). For Subject AB06, the stride length that produced the best-matching predictions was smaller than the ground truth stride length. This is likely because the actual stride length is known with more certainty later in swing as the stride nears completion; during this region of phase in swing, the position measurements obtained by integrating the accelerometer readings from our IMU sensor (red) are smaller in amplitude than the actual positions measured by the Vicon signals (orange). The EKF sees these smaller amplitudes and estimates a smaller stride length, as it fits the measurements better. The discrepancy between the sensor measurements and ground truth measurements is far less pronounced earlier in stride, indicating the the integration was more accurate during this region. Our filter was designed to produce accurate integration of the forward and upward heel positions based on the generalized gait model. Subject AB06 simply walks differently later in the gait cycle than what our filter was designed to estimate and our model was designed to predict, in a way that the general gait model cannot capture with accurate state estimates. Notably, this only resulted in a desired torque RMSE of 2.6 Nm for Subject AB06, as the other gait state variables, in particular phase, were still estimated well. In contrast, Subject AB03 (Fig. 5.7B) walks nearly identically to the general gait model prediction, and thus their state estimates (in particular stride length) are more accurate. Simple gait personalization techniques have been shown to significantly reduce the error in gait models over a continuous range of tasks [142], so we suspect that the variance of estimated phase

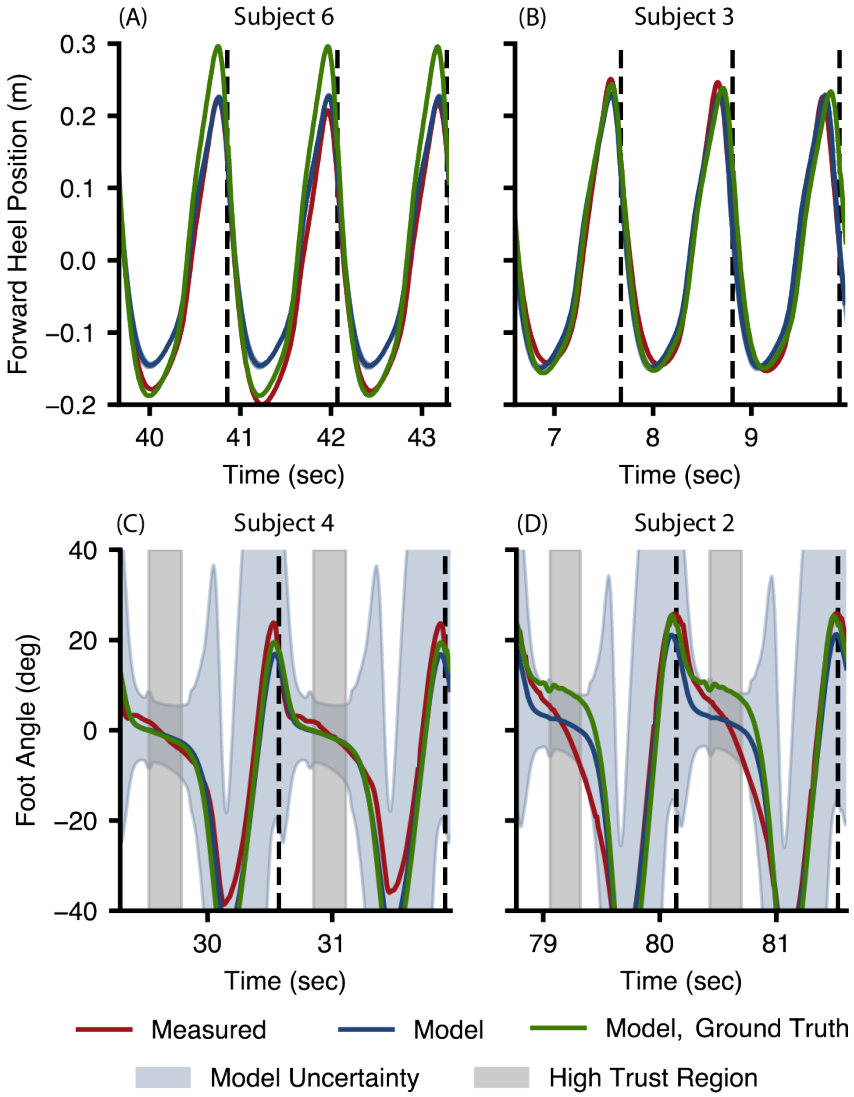

Figure 5.7: (A) Subject AB06's forward and upward heel position kinematics during the treadmill trials, as an example of how inter-person gait variability can lead to errors in state estimation. Based on the subject's kinematics as measured by the sensors (red), the EKF gait model predicts heel positions (solid blue) that imply a stride length state estimate below the ground truth measured by Vicon (orange). The gait model predictions using the ground truth state from Vicon (green) would more closely track the actual measurements from the sensors for Subject AB06. (B) Subject AB03's forward and upward heel kinematics during the treadmill trials better match the general gait model's predictions of heel position, leading to more accurate stride length estimates. (C) Subject AB04's foot and shank angle measurements exhibit a clear flat-foot contact period in early stance during regions of high model trust (gray shade) and thus result in accurate incline estimates; trust is indicated by the width of the confidence region (blue shade) about the model estimates. (D) Subject AB02's foot and shank angle measurements exhibit a less-defined flat-foot contact period where the noise model expects high confidence in the measurement, resulting in persistent negative biases in the incline estimate.

could be reduced with such a personalized gait model.

Unmodeled dynamics can also negatively impact our EKF performance. In particular, the exoskeleton's actuation imposes a disturbance on the system through the physical deflection of the device's IMU, which can lead to incorrect measurements of the shank angle. Our gait and heteroscedastic noise models assume that the measured shank angles will take on a specific, average profile, but actuation causes the measured shank profiles to be fundamentally different than what our models expect. In our implementation, we accounted for this by increasing the shank measurement covariance so that the EKF placed less trust in this signal. While this made the controller more robust to the disturbance caused by the exoskeleton torque, it also reduced the bandwidth of the controller, which otherwise could have been tuned to more quickly detect changes in gait.

Furthermore, the presence of exoskeleton assistance may have caused some participants to walk differently than how our gait model predicted, which would also lead to incorrect EKF behavior. For example, our gait model predicts that foot angle during early stance is relatively stationary and equal to the ground incline angle. Consequently, our noise model trusts the foot measurements during these regions of stance (Fig. 5.7, gray shade) as highly informative for estimating incline (Fig. 5.7, narrower blue bands around the predicted foot measurement during early stance). Some participants (Fig. 5.7C), walked similarly to this prediction while experiencing powered assistance, which led to more accurate incline estimates. However, other participants (Fig. 5.7D) notably had foot angle measurements that sloped downward more steeply during early stance and had a less-defined foot-contact region of stance, which was also reflected in a steeper shank angle. We suspect this is due to these participants allowing the exoskeleton torque to begin lifting their heel off earlier in stance. During the region of stance where the EKF expected 1) the foot measurement to be most informative about incline, and 2) the foot angle to be equal to the ground incline, the presence of exoskeleton assistance caused the foot measurement to have a negative value relative to the ground inclination. For Subject AB02 in Fig. 5.7D, this negative bias was as high as 7.3 deg. This was reflected in the ramp estimates of the EKF, which undershot the true ground incline by roughly this negative bias. Here again, individualized gait models may be a potential solution, as they can capture not only how every person walks differently, but also how each person uniquely adapts their gait to exoskeleton assistance.

Because the dataset used to regress our gait model only contained steady walking data, other tasks such as running or start/stopping are not explicitly modeled. While in theory the presented EKF can account for sudden stops by estimating $\dot{p}$ as zero, we believe the estimator will benefit from training with data including such gait transitions (*e.g.*, [143]). As datasets grow to include more behaviors, our intention is to extend our continuous gait model with new task variables representing these other dimensions of human locomotion.

5.7 Conclusion

We developed an exoskeleton torque controller based on an EKF which estimates phase, phase rate, stride length, and ramp in real time. This controller yields significantly reduced phase estimation errors compared to the state of the art. Furthermore, this controller improves upon the state of the art by allowing the assistive torque profile to adapt in real time in response to the state estimates. To the authors' knowledge, we are the first to estimate the gait phase variable along with stride length and ground inclination in real time and throughout an outdoor non-steady-state locomotion task on difficult terrains. This result represents a meaningful milestone for practical exoskeleton control and usage outside the laboratory.

CHAPTER 6

Conclusions and Future Directions

As a whole, this dissertation presents fundamental strategies for the translation of exoskeletons out of the laboratory and into the daily life of the public. This work contributes results about human's perception of exoskeleton benefits, a new user-focused economic design metric for wearable robots, and a theoretical analysis and practical evaluation of an adaptive multi-modal controller for ankle exoskeletons that all bring the field closer to the translation of exoskeletons. Greater voluntary adoption of these wearable technologies is likelier to occur with the refined understanding of how humans relate to state-of-the-art exoskeletons provided by this work. In particular, given the near-prerequisite for adoption that individuals perceive these technologies as valuable, this work has developed methods for ensuring that this perceptual value is quantifiable for exoskeletons. Furthermore, another prerequisite for mass adoption is that exoskeleton controllers be capable of adapting their powered assistance to the terrains and gaits of the real world, which requires the controller to first identify and learn these locomotion patterns. This dissertation has also not only presented methods to learn these patterns in real-time using the sensors onboard an exoskeleton, it has also demonstrated that it is possible to use the learned gait patterns to adapt the assistance from the exoskeleton, thus allowing this device to be used in real-world scenarios needed for large scale adoption.

This dissertation was motivated by the need to translate exoskeletons out of the controlled laboratory settings in which researchers are comfortable into the real world where they can actually benefit society. This work identified two key obstacles to this translation: 1) the lack of accessible, intuitive metrics by which to design and control exoskeletons that will actually matter to the general public, 2) the absence of exoskeleton controllers capable of augmenting human performance beyond the controlled conditions of the laboratory. This dissertation presented the results of four full experiments designed to address these gaps; the first two of these studies contributed scientific characterizations of the human relation to modern, metabolically advantageous exoskeletons, while the second two discussed the development of adaptive controllers capable of augmenting humans throughout a variety of locomotion modalities. In providing these solutions, this work advances the goal of exoskeleton translation to the public.

With the overarching mission of translation in mind, each chapter of this dissertation contributed to solving a different facet of this challenge. In Chapter 2, I characterized the conscious perception of metabolic rate changes, which was meant to answer the question of if humans can perceive the energetic benefits for modern exoskeletons. The results of this work then inspired the investigation in Chapter 3, which presented a multidisciplinary study that characterized the economic value (in dollars) of modern bilateral ankle exoskeletons designed to achieve metabolic reductions. Chapter 4 presented a study investigating the applicability of Bayesian analysis framework to learn a continuously evolving gait state vector that parametrizes different locomotion types throughout varying environments. Finally, Chapter 5 applies the conclusions of Chapter 4 to develop an Extended Kalman Filter (EKF)-based controller capable of applying torque over a variety of practical real-world environments. Taken together, this dissertation provides a solid foundation on which future research advances in exoskeleton translation can be developed.

6.1 Discussion of Contributions

The following subsections provide summaries of the major contributions from each of the previous chapters in this dissertation.

Chapter 2: Significance and Contributions

In Chapter 2, I described the quantification of the Just Noticeable Difference (JND) of changes in metabolic rate, which denotes the perceptual threshold below which changes in energetics are not perceived. This work was published in the *Journal of NeuroEngineering and Rehabilitation* (JNER) [117]. Metabolic rate reductions have emerged as the physiological "gold standard" by which modern exoskeletons have been designed and judged [150], and the capacity for an exoskeleton to reduce a user's energetic burden has become near-synonymous for the device's success. Researchers have adopted this metric as the standard largely because it is both intuitive, as it is a natural conclusion that the augmented human will walk for further or faster if they have more energy to spare, and objective, as metabolic rate is easily quantifiable within the laboratory. Despite these successes, it was unknown if users, when given a modern exoskeleton, would actually sense that they were being assisted, and indeed, this underscores a lack of consideration for the human system and thought process in the pursuit of strictly energetically efficient exoskeletons. This study revealed that, despite the well-established metabolic benefits from exoskeletons [123, 122, 132, 149, 110, 85, 202, 128], the average person would not actually sense that they were being assisted. Thus, this study filled a gap in the literature by providing crucial information on how users would consciously relate to modern exoskeletons. The JND itself acts as a natural

benchmark for exoskeletons designed to reduce metabolic cost, as the device can only be actively percieved as useful above the high magnitude of the JND. Beyond this primary result, this study also yielded a novel protocol that adapts classical psychophysical techniques for estimating JNDs to work with stimuli that are noisy, such as metabolic cost [114].

This study gave critical context to the long-standing metabolic rate reduction metric by which the majority of augmentative exoskeletons are designed and judged. The result that the average person is insensitive to changes in their metabolism casts doubt on the universality of the metabolic metric as the gold standard within the field, as if humans are unlikely to perceive these devices as immediately beneficial, it is also unlikely that they would voluntarily choose to adopt them into daily life [40]. Instead, the metrics by which we design exoskeletons should incorporate the user's perception in the evaluation in order to maximize the likelihood of mass adoption, and motivates further investigation into alternate biomechanical and physiological standards, such as muscle fatigue [3, 36], operational metrics (*e.g.* time until task completion) [169, 60, 11], or quantities such as user preference [84] that are more closely integrated with the conscious human system.

Chapter 3: Significance and Contributions

In Chapter 3, I described a multidisciplinary study which applied techniques from the field of economics to quantify the economic value of bilateral ankle exoskeletons and is currently in submission to *Nature Communications Engineering*. This work was heavily motivated by the results described in Chapter 2 which underscored the need for exoskeleton metrics of success that are conducive to mass adoption. By definition, users will only choose to adopt and use devices if they are perceived as actively valuable [40, 93, 129, 79, 29]. For such questions of value, it is only natural to turn to the field of economics for answers. By quantifying the *economic value* of modern exoskeletons, we can directly measure a variable that is tightly linked to the human's experience of wearing the exoskeleton; the user must holistically evaluate this experience and balance the perceived benefits and costs of the device, which is reflected in that user's value for the exoskeleton. Thus, quantifying the economic value of exoskeleton use presents a powerful, user-centric metric for the design and control of exoskeletons that directly reflects the likelihood of adoption.

To our knowledge, this is the first time researchers in our field have ventured into such a disparate discipline to answer a scientific question on exoskeletons. This foray into economics presents an exciting new frontier of research: investigations into how valuable users *actually* find exoskeletons. Perhaps no other quantity is as indicative of the potential for the mass adoption (which on some level must invariably involve a purchase) of these technologies than the simple value, in dollars, of the device itself. This study discovered that, to the average person, the value

of state-of-the-art bilateral ankle exoskeletons is negligible, which is particularly surprising given the refined nature of the robot and its ability to reduce the wearer's metabolic cost. It is noteworthy that, despite the *average* person receiving no benefit, there were numerous users in the study who received large economic value from the device (on the level of $16/hr). Additionally, this study acts as a proof-of-concept for the economic value metric, and motivates further investigations into the value of other devices or control strategies.

Chapters 4 and 5: Significance and Contributions

In Chapter 4, I described an analysis of the fundamental problem of gait state estimation with the Bayesian filtering framework, which has been published in the *IEEE Robotics and Automation Letters* [118]. In Chapter 5, I describe a novel Kalman-filter based controller that implements the results of Chapter 4 to continuously update the assistance from bilateral ankle exoskeletons on a variety of locomotion tasks; this chapter has been published in the *IEEE Transactions on Robotics* (T-RO) [116]. Turning to the challenge of controlling exoskeletons in the real world, we are first confronted with the wide variety of different locomotion modes possible beyond that of the steady-state walking that features in most scientific studies of exoskeletons. For these technologies to actually enter the real world, they must be capable of learning and adapting their assistance to the continuously evolving, transitory gaits and tasks that feature in practical settings, such as repeated starting and stopping on non-flat terrains. A person's *gait state*—the vector representation of their current locomotion mode—is typically difficult to measure directly using the sensors onboard wearable robots. However, prior research has been able to estimate subsets of the gait state using Bayesian filtering methods [173, 90, 30], although these past methods did not parametrize the gait cycle in its entirety and did not continuously learn all of the task parameters. Therefore, the combined works in Chapters 4 and 5 yielded a solution to the gait state estimation challenge in the form of an exoskeleton controller capable of learning the gait state throughout a wide range of locomotion modes. The study in Chapter 4 extended our understanding of the Bayesian filter's applicability to the gait state estimation challenge by characterizing the performance of different Bayesian Kalman filters (Extended, Unscented, and Ensemble) on continuously predicting the gait state during a variety of simulated walking tasks. We found that the Extended Kalman Filter (EKF), the simplest filter, provided the best estimates of the full gait state, which we then implemented on hardware in the study described in Chapter 5.

The key contribution of these works is an exoskeleton controller capable of functioning in realistic terrains outside of the laboratory, and therefore provides a solution to the challenge of learning a person's gait in the real world to adapt assistance. Our controller was tested in a variety of different scenarios, in particular the extreme conditions of the Michigan Wavefield, which features a set

of rapidly changing hills and thus represented a "stress-test" for our controller. In comparison to other standard approaches which only function on the controlled conditions of a laboratory treadmill [202, 42], our controller learned the user's gait during these trials and seamlessly updated its assistance. This contribution thus represents a meaningful milestone for practical exoskeleton controller design and implementation for the real world. Our EKF-based approach is also bolstered by the relative ease and intuition of the Kalman Filter equations at the controller's heart, which rendered the controller simple to implement and debug on hardware. Another key contribution of these works is the highlighting of personalization to improve estimation performance [46]. The controller leveraged a general gait model that mapped the gait state to the kinematics measurable by the exoskeletons, which was calculated using a pool of ten subject's data. We noted in Chapter 5 that some subjects' gaits deviated enough from the general gait model's expectations that performance deteriorated relative to those who hewed more closely to the average. Thus, implementing personalization techniques [142, 88] into controller to account for individuals' different strides is a key next step towards refining robust controllers for the real world.

6.2 Recommendations for Future Work

This dissertation, in pursuit of translating exoskeletons into ubiquity, also presents exciting future avenues of research. In keeping with the structure of solving the two primary obstacles to the goal (those of a lack of suitable metrics and practical controllers), I will discuss my thoughts and recommendations for future work in two sections. The first section will concern the challenge of obtaining useful metrics, and, more broadly, the relationship between humans and augmentative exoskeleton technology. The second section, devoted to addressing the challenge of controllers for the real world, will theorize about potential strategies for a controller capable of functioning in the near-infinite possibilities of the world outside the laboratory.

Suitable Metrics for Exoskeleton Design

The works in Chapters 2 and 3 were motivated by the open question of how to design exoskeletons people will want to use. In my mind, it is significant that, after a decade of nearly solely focusing on objective, physiological metrics like metabolism, we are only now as a field beginning to consider the question of how these benefits will actually relate to the human at the other side of the exoskeleton. Inspired by this revelation (and my own misgivings about academics being locked in ivory towers), I will rebrand this section as follows:

which I think better captures the key question at the heart of our recent efforts to create human-focused metrics to get exoskeletons out of the laboratory. I believe that if we are serious as a field about the potential of exoskeletons to transform human mobility, we must keep the average person front and center when it comes to the design of these devices, and this involves providing them with a sufficient and satisfactory reason to want to use our robots.

The work in Chapter 2 that describes the insensitivity of humans to modern, metabolically efficient exoskeletons in my opinion severely undercuts the case for metabolic rate reductions warranting as much attention as they have received by us scientists thus far. Given this insensitivity, exoskeletons designed with this metric that fail to meet the benchmark of the JND (which is essentially all untethered devices to date [150]) are unlikely to break into the public sphere; why would people want to use a device that renders no immediately observable benefit? Instead, the results of Chapter 2 encourage a reconsideration of which metrics will be most useful to the user. These metrics can be informed by similar investigations into the perceptual thresholds of other biomechanical or physiological quantities. For instance, the JND of mass manipulation has been shown to be about 8 grams (or 3% of the reference mass) [58]. Given this far keener sensitivity, exoskeletons designed with a task such as strength amplification [174, 72] or compliance shaping [174, 175], which aim to attenuate the physical properties (including inertia) of the environment felt by the wearer, may be likelier to be perceived as immediately beneficial by wearer's. Alternatively, we as a field could choose to shift our focus to the outcomes experienced by the user. This motivates a possible a shift to operational metrics [11, 169, 60] where the outcome of the task (*e.g.* how quickly the task is done, how many repetitions can be performed) is what matters. Perhaps the average user is insensitive to changes in metabolism, but he or she would notice if they completed a task more quickly.

There is another critical question implied by this shift in paradigm away from engineering metrics and towards the wearer's thought process: Where and when will humans find exoskeletons valuable? This consideration is informed by the economic value metric framework pioneered in Chapter 3, which directly centers the value of the device. In this respect, I've found that a comment from one of the reviewers for the Vickrey study on quantifying the value of exoskeletons during locomotion in Chapter 3, who as indicated by the editor, was not a scientist native to our field, encapsulates a longstanding thought of mine. The feedback is (verbatim) as follows:

> Vickrey second-price auctions are useful as valuation measuring rods only if it is interesting to learn the value of the item auctioned off. That might clearly be the case in my above example with a car, say if one is a company like Toyota and interested in figuring out demand for one's vehicles. It is furthermore compelling, as I said above,

> that someone may want to know the value that users of exoskeletons attach to such devices, in certain situations. However, I very much doubt that the situations studied in [the authors]'s experiment belongs to that class. **Subjects buy (or sell) the right (or obligation) to take a "walk" on a treadmill wearing an exoskeleton. That seems to be a rather uninteresting or irrelevant item to buy or sell.** For example, in the movie Edge of Tomorrow, Tom Cruise's character uses exoskeletons to fight off alien "Mimics," and it is clear that this is quite valuable to him. However, I doubt that he would have wanted to walk on treadmills instead. Accordingly, I feel that this study does not really produce any insights as regards the actual value of the use of exoskeletons **in meaningful situations.**

I think this feedback is notable for a few reasons. First, this represents the instinctive reaction of an outsider to our field to work with state-of-the-art exoskeletons, and the reviewer finds it a bit lacklustre that we are concerned with augmenting regular human walking. As scientists, I believe we have a responsibility (perhaps a pragmatic one) to consider what a regular member of the public (that is, someone who is not a wearable robotics researcher) thinks of our exoskeletons. If this reviewer's comment is anything to go by, state-of-the-art exoskeletons to reduce metabolic cost may be a harder sell than we would hope. Second, the reviewer identifies a use-case for exoskeletons in which they demonstrate obvious value (fighting aliens). In my opinion, future work in our field should clearly identify when and how exoskeletons are useful, instead of blindly assuming usefulness in all scenarios or assuming usefulness based on engineering principles. For instance, able-bodied individuals *can already walk normally* by definition; who is to say that the 10% reduction from a modern device will at all be useful when the baseline performance is already acceptable? If nothing else, considering use cases where the exoskeleton's benefit is incontrovertible dramatically raises the odds of use. The way I like to think about it is to consider the fundamental limitations of the human body when performing tasks. For example, no amount of caloric reduction will ever allow me to lift a car, as that task is impossible due to the physical makeup of the human physique. In my opinion, an exoskeleton that allows the wearer to supersede a limitation, rather than augmenting an already functional ability, is far likelier to be perceived as valuable, and this has direct implications for the design of exoskeletons. Perhaps, with the intermediate decade of scientific work as a foundation, exoskeleton designs akin to the original Hardiman [109] or the BLEEX system [91], which were designed to bear loads impossible for humans to sustain, can have a resurgence. This mode of thought lends itself to the operational metrics outlined previously, as human performance on tasks is now paramount. Increased performance based on operational metrics should have an obvious relationship to value, and future work should study the strength of this correlation, for example, by using the Vickrey auction protocol to evaluate whether individuals leverage the value from the device to consciously choose to walk further or faster. Similarly, the

Vickrey auction could be leveraged in future studies to identify whether exoskeleton assistance is valued differently during different tasks, which can also inform where we as a field should target our work for device adoption.

Finally, the works in Chapters 2 and 3 add to a growing body of work that reveals significant variation in individual responses to exoskeletons. Chapter 2 demonstrated that there were a select few individuals who in fact could perceive the metabolic benefits from state-of-the-art energetically efficient exoskeletons. Similarly, while the work in Chapter 3 demonstrated a negligible average benefit from exoskeleton assistance, this was due to roughly half the participants receiving a large net benefit being counteracted by the remaining half experiencing a significant penalty, rather than due to most participants receiving zero value. This corroborates past work that demonstrates varied biomechanical [4], physiological [121], and cognitive [117, 67] responses to exoskeleton augmentation. To quote Bequette *et al.*: "Because an exoskeleton should work for an individual, and will not be worn by an overall mean, it is important to consider the individual responses that comprise the population estimates" [11]. Future research should investigate why some individuals (commonly dubbed "responders") appear to be more amenable to the overall exoskeleton experience; perhaps initial efforts to disseminate exoskeleton technology into society will begin by targeting these responders due to this pre-existing aptitude [144]. As a suggestion for a potential next study, the Vickrey auction protocol outlined in Chapter 3 could be repeated, but with the additional measurement of the metabolic responses of each wearer, along with other quantities such as walking kinematics. Perhaps economic value correlates with increased metabolic benefit, and thus economic value can act as a faster proxy measurement for energetic reductions; similarly, perhaps responders discovered an adaptation to their kinematics that let them leverage the device for more energetic benefit or economic value, which would thus provide an explanation for their individual outsized benefits.

Taken as a whole, these recommendations present a notable shift in how we as a field approach the evaluation of exoskeletons towards a more human-focused paradigm in which physiological metrics are but one aspect of the human-exoskeleton experience holistically.

What Will a General Practical Controller Look Like?

Of course, when an exoskeleton is out in the real world, regardless of the metric by which it was designed, it must be capable of functioning robustly in all scenarios within its use case. Restricting ourselves to the case of augmenting locomotion (itself a strong assumption), the past work in this dissertation has advocated for a model-driven Kalman filter approach to solve the gait state estimation challenge. In my opinion, the Kalman filter gait state approach is promising as a foundation for future work, specific due to the presence of the continuous gait model at the heart of the filter.

Speaking from direct experience formulating, implementing, and (perhaps especially) debugging the EKF controller, the explicit, transparent nature of the gait model was a boon to the controller's development. Given the cyclic nature of locomotion, even on non-steady-state tasks, formulating the basis for the model as a series of Fourier functions was natural; similarly, we were able to exploit obvious relationships in the expected relationships between stride length (an amplitude scaling) and ground slope (a DC bias during stance) to express gait as a function of simple linear relations which in practice worked effectively.

I will provide further recommendations on improvements to this controller approach. An obvious extension to the gait model-based approach is to provide personalization to the wearer. Every person walks differently [46], and tailoring the gait model to reflect this fact will likely provide a straightforward next step to improve practical performance. For example, Principal Component Analysis techniques have already been applied to discriminate between individual walking gaits [81, 10, 207]; by including these principal components as gait state elements, an individualized gait model could be learned in real-time.

Where the really interesting challenges for practical controllers lie is in augmenting the user during tasks that are not as easily modellable as those previously discussed. For example, how are the non-steady transitions between repetitive steady-state tasks encoded? For challenges like these, where obvious relationships between the gait state and the measurable quantities from the exoskeleton as not as clear, a natural solution is to employ machine learning methods [88, 89, 158, 24]; no obvious relationship need exist, instead, simply label the data accordingly and turn the gait state estimation approach into a supervised regression task. Admittedly, having gained considerable practical experience with deep learning methods recently, this is an attractive solution for tasks such as stairs locomotion, where a continuous transition between this task and a task akin to level ground walking is not easy to intuit. Indeed, there is an increasing amount of open-source kinematics data [46, 143, 25] available with which to train such neural networks. In my opinion, this is perhaps the principal strength of the "black-box" approach that deep learning represents, as (theoretically) estimation can be improved by simply adding more data that covers a wider variety of subjects or tasks.

However, this black-box approach must reckon to what I believe is a fundamental weakness: the lack of guarantees that the data collected covers all possible use cases. Simply because we think we have seen enough data today guarantees nothing about the data we will see tomorrow [170]; the world is a complex, ever-changing place, and it is foolish to assume that datasets will contain every conceivable edge case for locomotion that exists now, to say nothing about future uncertainties. In contrast, a method like the Bayesian filter would behave in a predictable manner when confronted with such uncertainties due to its basis in explicit rules.

Having worked with both approaches, it is my belief that a practical controller for real-world

use must also explicitly consider confidence in its estimates and robustness to uncertainty when providing continuous augmentation. For example, a neural network to control the augmentation provided by an exoskeleton could also be trained to provide a level of confidence in its estimations, and shift its assistance accordingly; a high confidence level would result in full assistance, while a low confidence level could default to a "backup plan" driven by simpler rules, similar to the smaller EKF present in the work in Chapter 5. This would be advantageous, for instance, in gracefully handling transitions between tasks by defaulting to a safe mode of operation during these periods of higher uncertainty. This focus on robustness is especially critical when considering the human pilot's experience. The user should trust the device to provide seamless assistance when shifting from task-to-task [169], in order to maximize the likelihood that they perceive the device as beneficial. Similarly, the exoskeleton controller should not require any significant extra effort from the human to use, which promotes a controller that continuously adapts to the user. The Bayesian approach already embeds this confidence in its estimates in its assumptions, and thus a controller that merges both the black-box deep learning and Bayesian filter methodologies to leverage both approaches' advantages and maximize the odds of success (or at least, minimize the odds of failure) provides potential for a robust, practical controller. For example, a controller could leverage the comprehensive locomotion modes encoded in a neural network to provide a baseline gait state estimate, which could then act as an input to a Kalman filter. Such a controller's neural network and Kalman filter could even be trained end-to-end [68] to learn both the gait state and the appropriate level of confidence [8, 187].

6.3 Concluding Remarks

In lieu of a dry recapitulation of the past fifty-odd pages of dissertative work, I'd like to provide a brief anecdote that I think provides a better motivation of why this document came to be.

Within every mission, there come the occasional flashes of insight where, all at once, the true importance of your journey suddenly comes into sharp focus. For my own journey working with wearable robots, that moment came early in my academic career when I had the privilege of meeting disability rights advocate Shane Burcaw after a Robotics Colloquium here at the University of Michigan. Mr. Burcaw suffers from spinal muscular atrophy and has been wheelchair-bound since a young age. As I showed him some of my laboratory's powered ankle exoskeletons and prosthetic legs, it was hard to miss his sheer enthusiasm at the potential for such technologies to directly benefit people living with disabilities. While I had long been drawn to the abstract idea of wearable robotics to benefit the common good, here was concrete, human proof that I could, I would, meaningfully benefit the world with my work on wearable robots like exoskeletons.

We in robotics confront a vast array of technical challenges on a daily basis, so much so that

it's easy to lose sight of one basic truth: if we're not helping others with our work, why even do it? I choose to believe that all of us in this field, regardless of how we arrived here, whether that be from personally knowing someone with a disability, wanting to push the limits of the human body, or even having watched *Iron Man* as a nine-year-old kid and thinking it was the coolest thing ever, now choose to get up in the morning and put our heads to the grindstone willingly because we believe that our work can actually help real people in need like Shane.

It is this truth that has inspired the dissertation contained herein and my journey throughout the last five years and will continue to drive my work with powered exoskeletons.

APPENDIX A

Chapter 3 Supplementary Materials

A.1 Supplementary Text

A.1.1 Level-ground study

We conducted a separate, smaller study in which two participants (Subjects 2 and 7) underwent the main study's protocol, but on level-ground instead of at a 10° incline. This was done to investigate how their MVs changed when the walking was made easier and more representative of exoskeleton use in the community. Both participants' MVs diminished when the task was level-ground walking—Subject 2's MEV dropped from 36.35% to -28.10%, while Subject 7's MEV dropped from 26.68% to 12.62% (Fig. A.6). These results suggest that for everyday mobility tasks, which are likely to be more similar in difficulty to level-ground walking than uphill walking, energetic exoskeleton assistance may have less value to potential adopters. Furthermore, augmentative exoskeletons may thus provide maximal value during highly strenuous tasks. Future work should investigate if these results are consistent across more subjects and different walking conditions.

A.2 Supplementary Tables

Table A.1: Participant characteristics.

Participant	Gender	Weight (kg)	Age	Height (cm)
1	F	83.64	24	160.02
2	M	99.2	27	176
3	M	84	22	186
4	M	90.7	25	182.9
5	M	80.7	26	177.8
6	F	72.58	29	170.18
7	M	80	34	180.3
8	M	74.8	34	74.8
9	M	68.03	25	170.18
10	M	70.3	22	185.4
11	M	67.13	25	167.6
12	M	104.3	36	193
13	F	74.84	23	177.8
14	M	70.3	21	172.7
15	M	61.2	23	163
16	F	52.4	25	162.6

A.3 Supplementary Figures

Table A.2: Powered condition cumulative prices and marginal values.

Subject	MV (%)	Exo Powered ($)	Walking No Exo ($)		
			Day 1	Day 2	Day 3
1	1.07	13.34	13.48		
2	36.35	42.36	66.55		
3	26.14	15.46	20.93		
4	-13.85	17.66	15.09	15.94	
5	-8.00	25.03	23.16		
6	2.45	26.88	27.55		
7	49.27	36.53	72.02		
8	55.83	62.77	142.09		
9	11.70	11.40	12.91		
10	11.47	11.18	12.63		
11	14.86	17.36	20.64	19.81	20.71
12	-59.43	30.99	19.44		
13	-32.35	35.16	26.57		
14	18.04	11.51	14.04		
15	-46.56	33.42	21.97	23.30	23.14
16	26.08	14.86	19.24	20.23	20.85

Table A.3: Powered-off condition cumulative prices and marginal values.

Subject	MV (%)	Exo Powered Off ($)	Walking No Exo ($)
1	-0.53	13.56	13.48
2	-26.44	84.15	66.55
3	-3.53	21.31	20.59
4	-0.10	15.10	15.09
5	-7.40	24.88	23.16
6	17.44	22.75	27.55
7	-40.51	199.66	142.09
9	-26.64	16.34	12.91
12	-119.94	42.76	19.44
13	-111.95	56.31	26.57

Table A.4: Cumulative prices for participants who repeated the walking-no-exo condition

Subject	Average Cumulative Price ($)	St. Dev. Cumulative Price ($)	St. Dev. Cumulative Price (% of Average)	Prices ($)		
				Day 1	Day 2	Day 3
4	15.5	0.4	2.8	15.1	15.9	
11	20.4	0.4	2.0	20.6	19.8	20.7
15	22.8	0.6	2.6	22.0	23.3	23.1
16	20.1	0.7	3.3	19.2	20.2	20.9

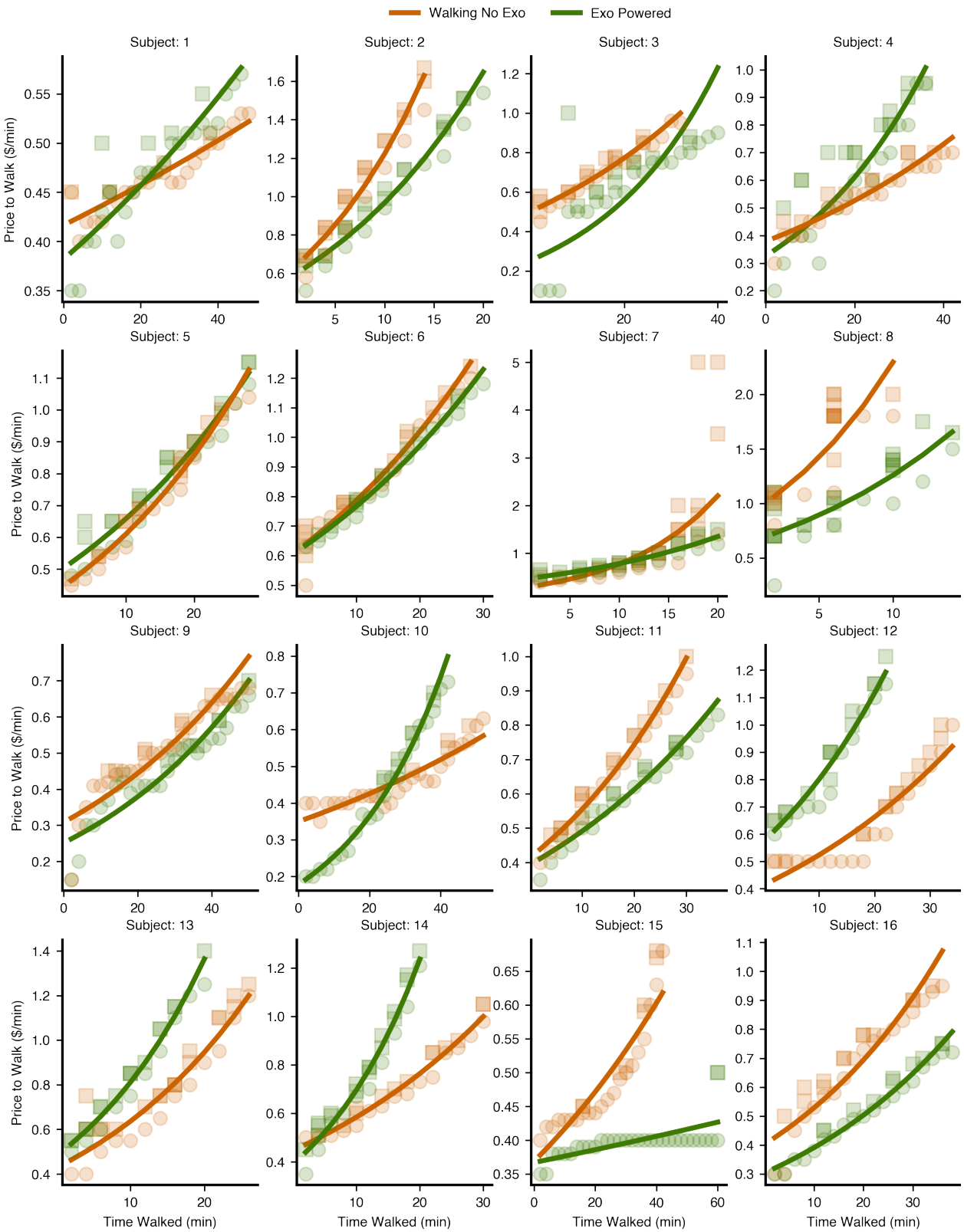

Figure A.1: The price to walk curves for the exo-powered (orange) and walking-no-exo (teal) from all sixteen subjects. Circles denote winning bids, while squares denote losing bids.

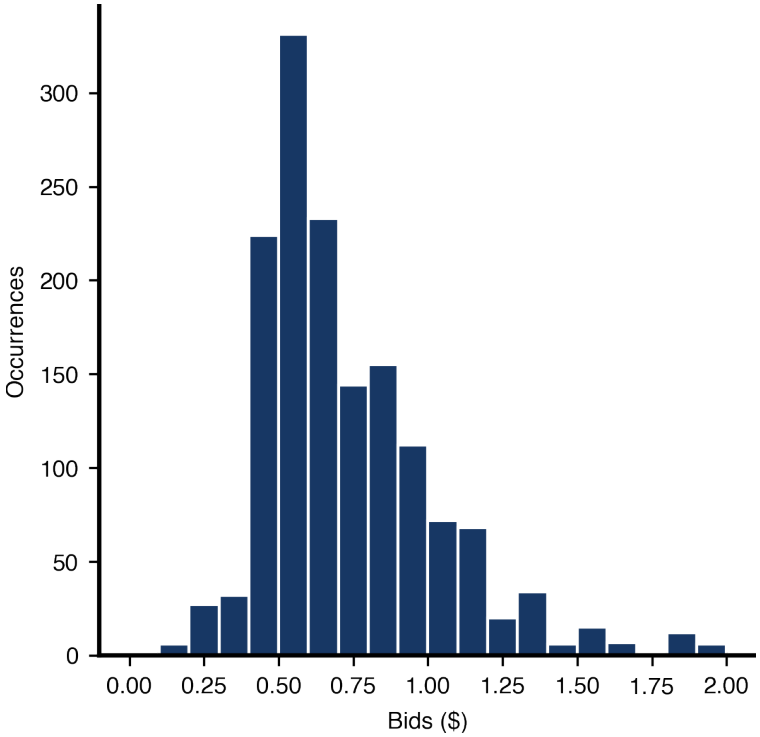

Figure A.2: A histogram of all the subject bids for the two-minute intervals across all tested conditions. The average bid was $0.75, with a standard deviation of $0.39.

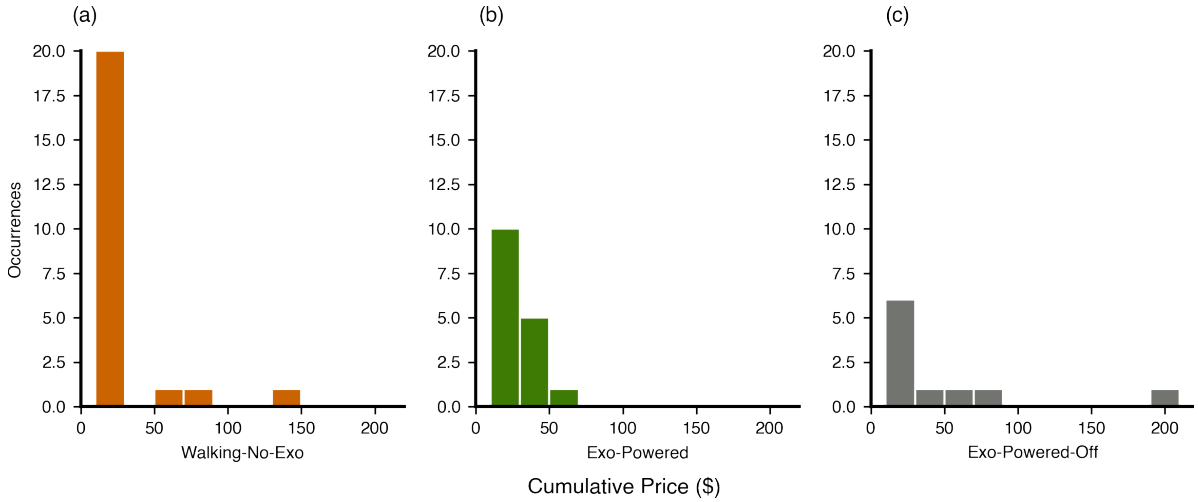

Figure A.3: Histograms of the cumulative costs for (a) the walking-no-exo condition, (b) the exo-powered condition, and (c) the exo-powered off condition.

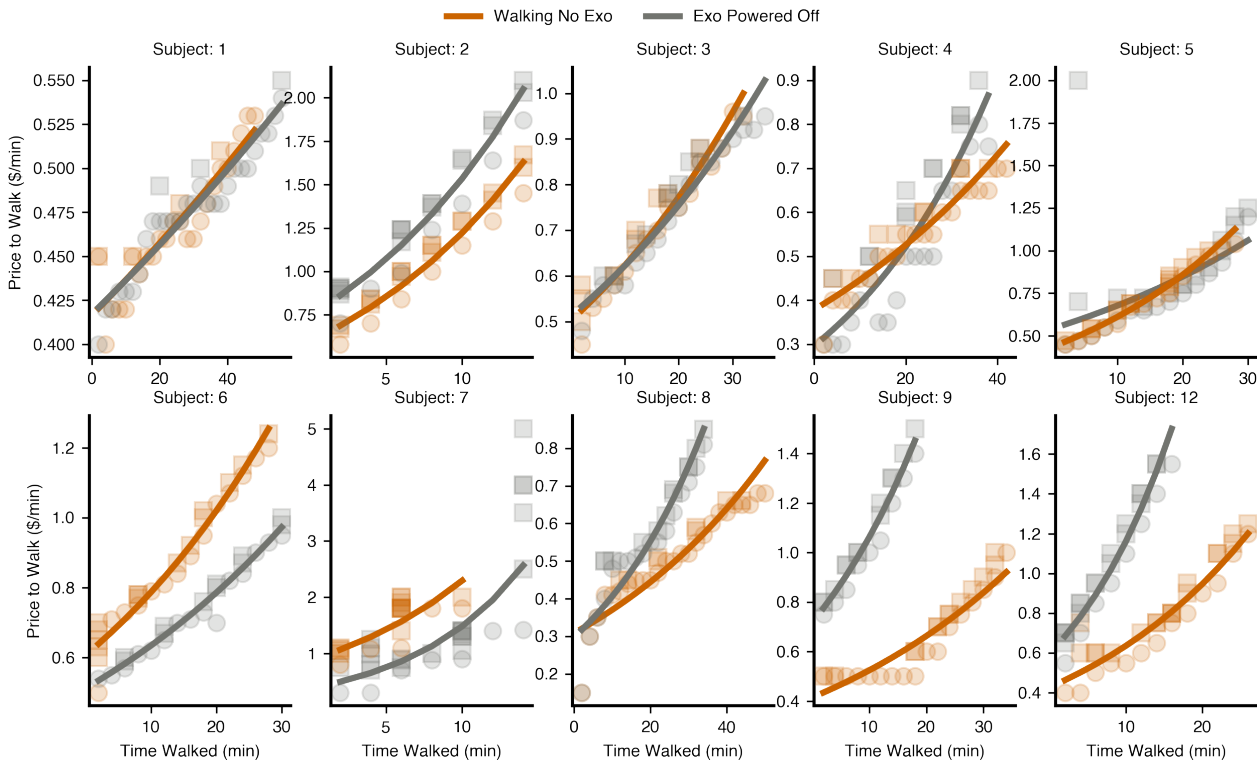

Figure A.4: The price to walk curves for the exo-powered-off (gray) and walking-no-exo (teal) for the ten subjects who completed the former condition. Circles denote winning bids, while squares denote losing bids.

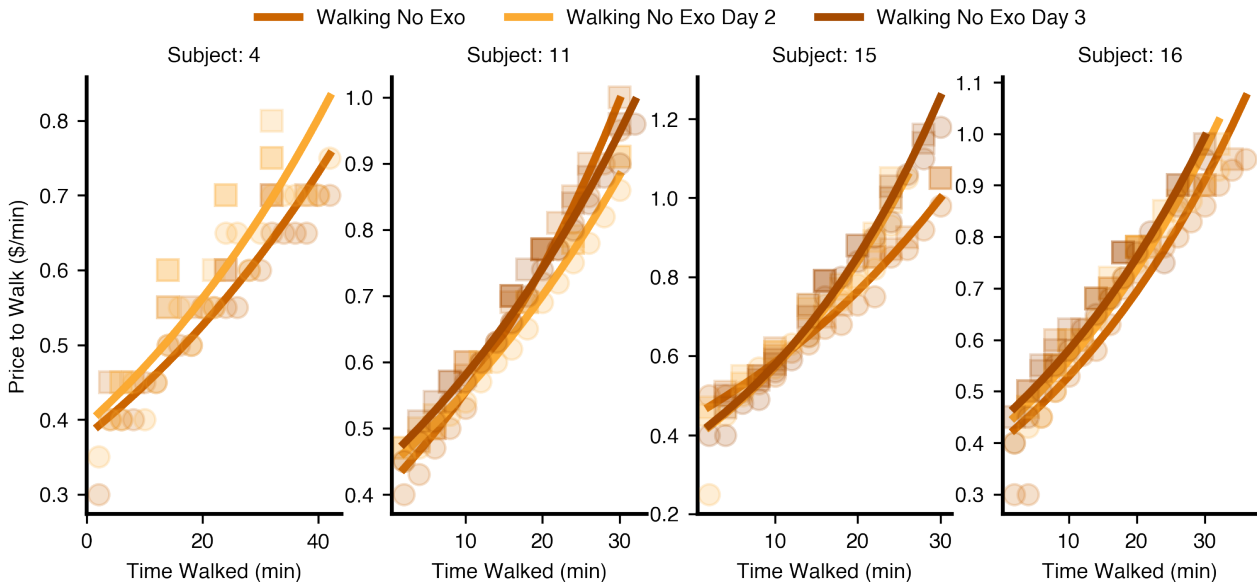

Figure A.5: The price to walk curves for the walking-no-exo conditions across different days for the four subjects who repeated the walking-no-exo condition. Circles denote winning bids, while squares denote losing bids. Subject 4 only repeated the walking-no-exo condition once, while the rest of the participants repeated it twice.

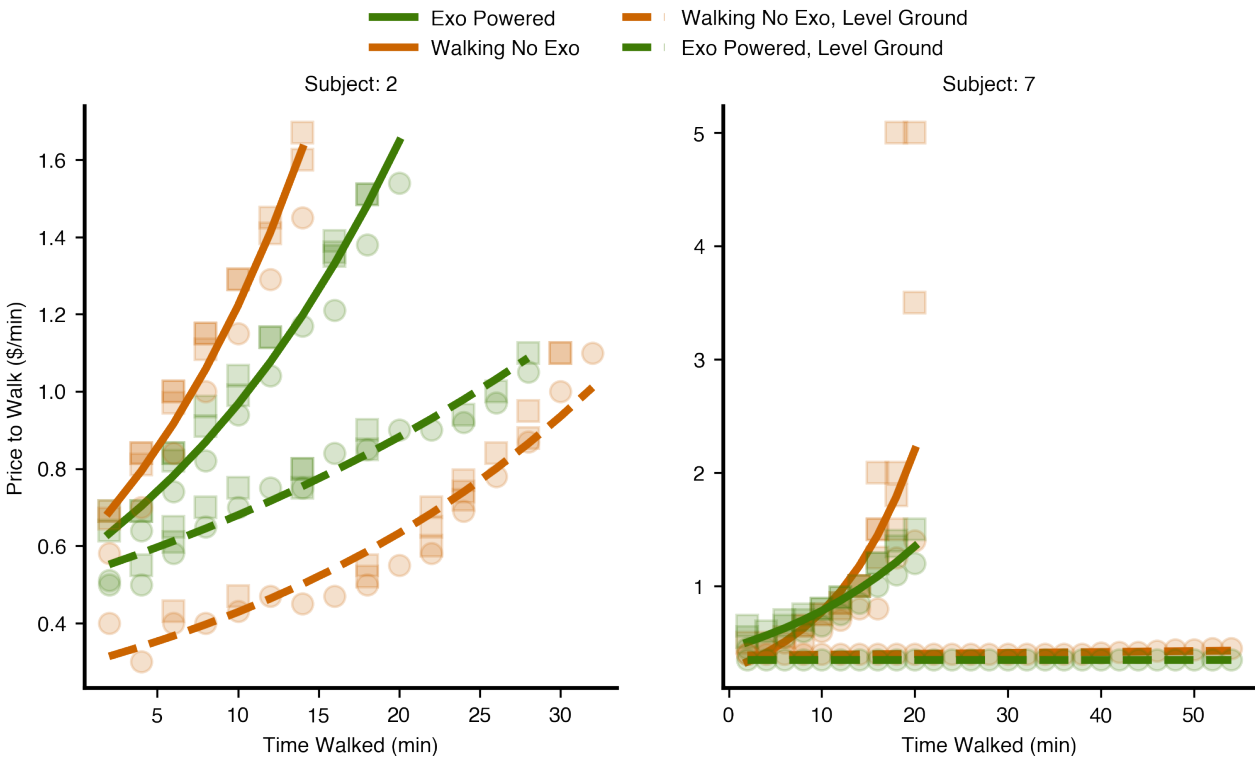

Figure A.6: The price to walk curves for two subjects (2 and 7) that repeated the walking-no-exo and exo-powered conditions on a level ground condition. The solid lines denote the standard 10° incline task featured in this work; the dashed lines are the level ground condition. Both participants' MEVs decreased during the level ground condition.

APPENDIX B

Extended Kalman Filter Implementation

Starting from the state estimate at the previous time $\hat{x}_{k-1|k-1}$ and the previous state covariance estimate $P_{k-1|k-1}$, our EKF implementation computes the current state estimate update using measurement z_k, dynamic model $f(\cdot)$, and measurement model $h'(\cdot)$. The process involves two steps. First, we propagate the (conveniently linear) dynamics from (4.3) across the time step,

$$\begin{aligned}\hat{x}_{k|k-1} &= F\hat{x}_{k-1|k-1},\\ P_{k|k-1} &= FP_{k-1|k-1}F^T + \Sigma_Q.\end{aligned}$$

We then correct the estimate based on the measurement z_k as

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(z_k - h(\hat{x}_{k|k-1})),\\ P_{k|k} &= P_{k|k-1} - K_kH_kP_{k|k-1},\\ \text{where } K_k &= P_{k|k-1}H_k^T\left[H_kP_{k|k-1}H_k^T + \Sigma_{R_k}(\hat{p}_{k|k-1})\right]^{-1},\\ \text{and } H_k &= \frac{\partial h'}{\partial x}|_{\hat{x}_{k|k-1}}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \frac{dl}{dl_p} & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.\end{aligned}$$

BIBLIOGRAPHY

[1] Sabrina J. Abram, Jessica C. Selinger, and J. Maxwell Donelan. Energy optimization is a major objective in the real-time control of step width in human walking. *J. Biomechanics*, 91:85–91, 2019.

[2] Rochelle Ackerley, Helena Backlund Wasling, Max Ortiz-Catalan, Rickard Brånemark, and Johan Wessberg. Case studies in neuroscience: Sensations elicited and discrimination ability from nerve cuff stimulation in an amputee over time. *J. Neurophysiology*, 120(1):291–295, 2018.

[3] Marko Ackermann and Antonie J. Van Den Bogert. Optimality principles for model-based prediction of human gait. *Physical Therapy*, 31(January 2008):1–8, 2011.

[4] Yadrianna Acosta-Sojo and Leia Stirling. Individuals differ in muscle activation patterns during early adaptation to a powered ankle exoskeleton. *Applied Ergonomics*, 98:103593, 2022.

[5] Alejandro F Azocar, Luke M Mooney, Levi J Hargrove, and Elliott J Rouse. Design and Characterization of an Open-source Robotic Leg Prosthesis. *Int. Conf. Biomedical Robotics and Biomechatronics*, 2018.

[6] Alejandro F Azocar and Elliott J Rouse. Stiffness Perception During Active Ankle and Knee Movement. *IEEE Trans. Biomedical Engineering*, 64(12):2949–2956, 2017.

[7] Alejandro F Azocar, Amanda L Shorter, and Elliott J Rouse. Damping Perception During Active Ankle and Knee Movement. *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 27(2):198–206, 2019.

[8] Philipp Becker, Harit Pandya, Gregor Gebhardt, Cheng Zhao, C James Taylor, and Gerhard Neumann. Recurrent kalman networks: Factorized inference in high-dimensional deep feature spaces. In *International Conference on Machine Learning*, pages 544–552. PMLR, 2019.

[9] J. G. BEEBE-CENTER and D. WADDELL. A general psychological scale of taste. *The Journal of psychology*, 26:517–524, 1948.

[10] Xianye Ben, Weixiao Meng, Kejun Wang, and Rui Yan. An adaptive neural networks formulation for the two-dimensional principal component analysis. *Neural Computing and Applications*, 27(5):1245–1261, 2016.

[11] Blake Bequette, Adam Norton, Eric Jones, and Leia Stirling. Physical and cognitive load effects due to a powered lower-body exoskeleton. *Human factors*, 62(3):411–423, 2020.

[12] T Kevin Best, Kyle R Embry, Elliott J Rouse, and Robert D Gregg. Phase-variable control of a powered knee-ankle prosthesis over continuously varying speeds and inclines. In *IEEE Int. Conf. Intell. Robots Syst.*, pages 6182–6189, 2021.

[13] E. Borg and L. Kaijser. A comparison between three rating scales for perceived exertion and two different work tests. *Scandinavian J. Medicine and Science in Sports*, 16(1):57–69, 2006.

[14] Elisabet Borg. *On perceived exertion and its measurement*. 2007.

[15] Elisabet Borg. So What's that On a Scale of 1 to 10. *Proc. 24th International Congress of Vexillology*, (August 2011):988–995, 2011.

[16] Gunnar A.V. Borg. Psychophysical bases of perceived exertion, 1982.

[17] Wouter Van Den Bos and Samuel M. McClure. Towards a general model of temporal discounting. *J. Experimental Analysis of Behavior*, 99(1):58–73, 2013.

[18] Thomas Brand and Birger Kollmeier. Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *The J. the Acoustical Society of America*, 111(6):2801–2810, 2002.

[19] Eric E. Brodie and Helen E. Ross. Sensorimotor mechanisms in weight discrimination. *Perception —& Psychophysics*, 36(5):477–481, 1984.

[20] Raymond C. Browning, Jesse R. Modica, Rodger Kram, and Ambarish Goswami. The effects of adding mass to the legs on the energetics and biomechanics of walking. *Medicine and Science in Sports and Exercise*, 39(3):515–525, 2007.

[21] Gwendolyn M. Bryan, Patrick W. Franks, Stefan C. Klein, Robert J. Peuchen, and Steven H. Collins. A hip–knee–ankle exoskeleton emulator for studying gait assistance. *International J. Robotics Research*, 2020.

[22] U.S. Department of Labor Bureau of Labor Statistics. Physically strenuous jobs in 2017. *The Economics Daily*.

[23] U.S. Department of Labor Bureau of Labor Statistics. The employment situation - may 2022. *News Release*.

[24] Jonathan Camargo, Will Flanagan, Noel Csomay-Shanklin, Bharat Kanwar, and Aaron Young. A machine learning strategy for locomotion classification and parameter estimation using fusion of wearable sensors. *IEEE Transactions on Biomedical Engineering*, 68(5):1569–1578, 2021.

[25] Jonathan Camargo, Aditya Ramanathan, Will Flanagan, and Aaron Young. A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions. *Journal of Biomechanics*, 119:110320, 2021.

[26] B. L. Cardozo. Adjusting the Method of Adjustment: SD vs DL. *The J. the Acoustical Society of America*, 37(5):786–792, 1965.

[27] Bing Chen, Bin Zi, Zhengyu Wang, Ling Qin, and Wei Hsin Liao. Knee exoskeletons for gait rehabilitation and human performance augmentation: A state-of-the-art. *Mechanism and Machine Theory*, 134:499–511, 2019.

[28] Wonseok Choi, Wonseok Yang, Jaeyoung Na, Juneil Park, Giuk Lee, and Woochul Nam. Unsupervised gait phase estimation with domain-adversarial neural network and adaptive window. *IEEE J. Biomedical and Health Informatics*, 2021.

[29] Stephanie Hui Wen Chuah, Philipp A. Rauschnabel, Nina Krey, Bang Nguyen, Thurasamy Ramayah, and Shwetak Lade. Wearable technologies: The role of usefulness and visibility in smartwatch adoption. *Computers in Human Behavior*, 65:276–284, 2016.

[30] Geoffrey Clark, Joseph Campbell, Seyed Mostafa Rezayat Sorkhabadi, Wenlong Zhang, and Heni Ben Amor. Predictive modeling of periodic behavior for human-robot symbiotic walking. In *IEEE Int. Conf. Robot. Autom.*, pages 7599–7605. IEEE, 2020.

[31] Francesco Clemente, Bo Hakansson, Christian Cipriani, Johan Wessberg, Katarzyna Kulbacka-Ortiz, Rickard Brånemark, Karl Johan Fredén Jansson, and Max Ortiz-Catalan. Touch and Hearing Mediate Osseoperception. *Scientific Reports*, 7:1–11, 2017.

[32] Steven H Collins, M Bruce Wiggin, and Gregory S Sawicki. Reducing the energy cost of human walking using an unpowered exoskeleton. 522(7555):212–215, 2015.

[33] Steven H. Collins, Juanjuan Zhang, Katherine L. Poggensee, Kirby A. Witte, Rachel W. Jackson, Pieter Fiers, and Christopher G. Atkeson. Supplementary Materials for Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, 356(6344):1280–1284, 2017.

[34] Don. Coursey, John L Hovis, and William D Schulze. The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value Author (s): Don L . Coursey , John L . Hovis and William D . Schulze Published by : Oxford University Press Stable URL : http://www.jstor.org/stable/1884223 Accessed :. 102(3):679–690, 2016.

[35] Thomas S. Critchfield and Scott H. Kollins. Temporal discounting: Basic research and the analysis of socially important behavior. *J. Applied Behavior Analysis*, 34(1):101–122, 2001.

[36] Harrison P Crowell, Gregory B Kanagaki, Meghan P O'donovan, Courtney A Haynes, Joon-Hyuk Park, Jennifer M Neugebauer, Edward R Hennessy, Angela C Boynton, Blake Mitchell, Andrew J Tweedell, and Henry J Girolamo. Methodologies for Evaluating the Effects of Physical Augmentation Technologies on Soldier Performance. *US Army Research Laboratory Aberdeen Proving Ground United States*, (May), 2018.

[37] Xingye Da, Ross Hartley, and Jessy W Grizzle. Supervised learning for stabilizing underactuated bipedal robot locomotion, with outdoor experiments on the wave field. In *IEEE Int. Conf. Robotics and Automation*, pages 3476–3483, 2017.

[38] Sauvik Das Gupta, Maarten F. Bobbert, and Dinant A. Kistemaker. The Metabolic Cost of Walking in healthy young and older adults – A Systematic Review and Meta Analysis. *Scientific Reports*, 9(1):1–10, 2019.

[39] Audra Davidson, Emily S. Gardinier, and Deanna H. Gates. Within and between-day reliability of energetic cost measures during treadmill walking. *Cogent Engineering*, 3(1):1–7, 2016.

[40] Fred D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, 13(3):319–339, 1989.

[41] Ye Ding, Ignacio Galiana, Alan Asbeck, Brendan Quinlivan, Stefano Marco Maria De Rossi, and Conor Walsh. Multi-joint actuation platform for lower extremity soft exosuits. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1327–1334. Ieee, 2014.

[42] Ye Ding, Myunghee Kim, Scott Kuindersma, and Conor J. Walsh. Human-in-the-loop optimization of hip assistance with a soft exosuit during walking. *Science Robotics*, 3(15):1–9, 2018.

[43] J. M. Donelan, R. Kram, and A. D. Kuo. Mechanical and metabolic determinants of the preferred step width in human walking. *Proc. Royal Society B: Biological Sciences*, 268(1480):1985–1992, 2001.

[44] David Easley and Jon Kleinberg. Networks, Crowds and Markets: Reasoning about a Highly Connected World. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(4):1073–1073, 2012.

[45] Toby Elery, Siavash Rezazadeh, Christopher Nesler, and Robert D Gregg. Design and validation of a powered knee–ankle prosthesis with high-torque, low-impedance actuators. *IEEE Trans. Robotics*, 36(6):1649–1668, 2020.

[46] K. Embry and R. D. Gregg. Analysis of continuously varying kinematics for powered prosthetic leg control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:262–272, 2020.

[47] Kyle R. Embry, Dario J. Villarreal, Rebecca L. Macaluso, and Robert D. Gregg. Modeling the Kinematics of Human Locomotion over Continuously Varying Speeds and Inclines. *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 26(12):2342–2350, 2018.

[48] Geir Evensen. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.

[49] Daniel P. Ferris, Keith E. Gordon, Gregory S. Sawicki, and Ammanath Peethambaran. An improved powered ankle-foot orthosis using proportional myoelectric control. *Gait and Posture*, 23(4):425–428, 2006.

[50] John A Fox, Dermot J Hayes, and Jason F Shogren. Consumer preferences for food irradiation: how favorable and unfavorable descriptions affect preferences for irradiated pork in experimental auctions. *Journal of risk and Uncertainty*, 24(1):75–95, 2002.

[51] S. Galle, P. Malcolm, W. Derave, and D. De Clercq. Uphill walking with a simple exoskeleton: Plantarflexion assistance leads to proximal adaptations. *Gait and Posture*, 41(1):246–251, 2015.

[52] Samuel Galle, Philippe Malcolm, Steven Hartley Collins, and Dirk De Clercq. Reducing the metabolic cost of walking with an ankle exoskeleton: interaction between actuation timing and power. *J. NeuroEngineering and Rehabilitation*, 14(1):1–16, 2017.

[53] Miguel A. García-Pérez and Rocío Alcalá-Quintana. Sampling plans for fitting the psychometric function. *Spanish J. Psychology*, 8(2):256–289, 2005.

[54] Rick M. Gardner, James A. Morrell, Deborah N. Watson, and Susan L. Sandoval. Subjective Equality and Just Noticeable Differences in Body-Size Judgments by Obese Persons. *Perceptual and Motor Skills*, 69(2):595–604, 2011.

[55] Andrew Gelman, Jessica Hwang, and Aki Vehtari. Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6):997–1016, 2014.

[56] G. a. Gescheider. Psychophysical Measurement of Thresholds, 1985.

[57] George A Gescheider. *Psychophysics: the fundamentals*. Psychology Press, 2013.

[58] Christos Giachritsis, Jorge Barrio, Manuel Ferre, Alan Wing, and Javier Ortego. Evaluation of weight perception during unimanual and bimanual manipulation of virtual objects. *3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics 2009*, pages 629–634, 2009.

[59] Alan St Clair Gibson, Denise A Baden, Mike I Lambert, E Vicki Lambert, X R Harley, Dave Hampson, Vivienne A Russell, and Tim D Noakes. The Conscious Perception of the Sensation of Fatigue. *Sports Medicine*, 33(3):1–10, 2003.

[60] Sarah Gonzalez, Paul Stegall, Stephen M Cain, Ho Chit Siu, and Leia Stirling. Assessment of a powered ankle exoskeleton on human stability and balance. *Applied Ergonomics*, 103:103768, 2022.

[61] Daniel FN Gordon, Christopher McGreavy, Andreas Christou, and Sethu Vijayakumar. Human-in-the-loop optimization of exoskeleton assistance via online simulation of metabolic cost. *IEEE Transactions on Robotics*, 2022.

[62] Keith E Gordon and Daniel P Ferris. Learning to walk with a robotic ankle exoskeleton. *J. biomechanics*, 40(12):2636–2644, 2007.

[63] Leonard Green, Nathanael Fristoe, and Joel Myerson. Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, 1(3):383–389, 1994.

[64] Leonard Green, Joel Myerson, and Edward McFadden. Rate of temporal discounting decreases with amount of reward. *Memory and Cognition*, 25(5):715–723, 1997.

[65] Robert D Gregg, Tommaso Lenzi, Levi J Hargrove, and Jonathon W Sensinger. Virtual constraint control of a powered prosthetic leg: From simulation to experiments with transfemoral amputees. *IEEE Trans. Robotics*, 30(6):1455–1471, 2014.

[66] Laura Guidetti, Francesco Bolletta, Maria Chiara Gallotta, Carlo Baldari, Marco Meucci, and Gian Pietro Emerenziani. Validity, reliability and minimum detectable change of COSMED K5 portable gas exchange system in breath-by-breath mode. *PLoS ONE*, 13(12):1–12, 2018.

[67] Aditi Gupta, Ryan McKindles, and Leia Stirling. Metrics for quantifying cognitive factors that may underlie individual variation in exoskeleton use. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 65, pages 216–220. SAGE Publications Sage CA: Los Angeles, CA, 2021.

[68] Tuomas Haarnoja, Anurag Ajay, Sergey Levine, and Pieter Abbeel. Backprop kf: Learning discriminative deterministic state estimators. *Advances in neural information processing systems*, 29, 2016.

[69] Luke Haile, Robert J. Robertson, Elizabeth F. Nagle, Maressa P. Krause, Michael Gallagher, Christina M. Ledezma, Kristofer S. Wisniewski, Alex B. Shafer, and Fredric L. Goss. Just noticeable difference in perception of physical exertion during cycle exercise in young adult men and women. *European J. Applied Physiology*, 113(4):877–885, 2013.

[70] David B. Hampson, Alan St Clair Gibson, Mike I. Lambert, and Timothy D. Noakes. The influence of sensory cues on the perception of exertion during exercise and central regulation of exercise performance. *Sports Medicine*, 31(13):935–952, 2001.

[71] J. O. Hangen, H., Melanson, E., Tran, Z., Kearney, J. T., Hill. Variability of resting metabolic rate. *American J. Clinical Nutrition*, 78:1141–1144, 2003.

[72] Binghan He, Gray C. Thomas, Nicholas Paine, and Luis Sentis. Modeling and Loop Shaping of Single-Joint Amplification Exoskeleton with Contact Sensing and Series Elastic Actuation. 2018.

[73] Matthew D. Hoffman and Andrew Gelman. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Machine Learning Research*, 15(2008):1593–1623, 2014.

[74] Vivian Hoffmann, Jacob R. Fooks, and Kent D. Messer. Measuring and mitigating HIV stigma: A framed field experiment. *Economic Development and Cultural Change*, 62(4):701–726, 2014.

[75] Matthew A. Holgate, Thomas G. Sugar, and Alexander W. Böhler. A novel control algorithm for wearable robotics using phase plane invariants. *IEEE Int. Conf. Robotics and Automation*, pages 3845–3850, 2009.

[76] Woolim Hong, Namita Anil Kumar, and Pilwon Hur. A phase-shifting based human gait phase estimation for powered transfemoral prostheses. *IEEE Robotics and Automation Letters*, 6(3):5113–5120, 2021.

[77] Simo Hosio, Denzil Ferreira, Jorge Goncalves, Niels Van Berkel, Chu Luo, Muzamil Ahmed, Huber Flores, and Vassilis Kostakos. Monetary assessment of battery life on smartphones. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1869–1880, 2016.

[78] Blair Hu, Ann M. Simon, and Levi Hargrove. Deep generative models with data augmentation to learn robust representations of movement intention for powered leg prostheses. *IEEE Trans. Medical Robotics and Bionics*, 1(4):267–278, 2019.

[79] P. H. Hu, P. Chau, O. R. Sheng, and K. Tam. Examining the technology acceptance model using physician acceptance of telemedicine technology. *J. Manag. Inf. Syst.*, 16:91–112, 1999.

[80] He Huang, Fan Zhang, Levi J Hargrove, Zhi Dou, Daniel R Rogers, and Kevin B Englehart. Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion. *IEEE Trans. Biomedical Engineering*, 58(10):2867–2875, 2011.

[81] P. S. Huang, C. J. Harris, and M. S. Nixon. Recognising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering*, 13(4):359–366, 1999.

[82] L. C. Hunter, E. C. Hendrix, and J. C. Dean. The cost of walking downhill: Is the preferred gait energetically optimal? *J. Biomechanics*, 43(10):1910–1915, 2010.

[83] K. A. Ingraham, C. D. Remy, and E. J. Rouse. The role of user preference in the customized control of robotic exoskeletons. *Science Robotics*, 7(64):eabj3487, 2022.

[84] Kimberly A. Ingraham, C. David Remy, and Elliott J. Rouse. User preference of applied torque characteristics for bilateral powered ankle exoskeletons. *Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatron*, pages 839–845, 2020.

[85] Rachel W. Jackson and Steven H. Collins. An experimental comparison of the relative benefits of work and torque assistance in ankle exoskeletons. *J. Applied Physiology*, 119(5):541–557, 2015.

[86] Chetas D. Joshi, Uttama Lahiri, and Nitish V. Thakor. Classification of gait phases from lower limb EMG: Application to exoskeleton orthosis. *IEEE Point-of-Care Healthcare Technologies*, pages 228–231, 2013.

[87] Simon J Julier and Jeffrey K Uhlmann. Unscented filtering and nonlinear estimation. *Proc. IEEE*, 92(3):401–422, 2004.

[88] Inseung Kang, Pratik Kunapuli, and Aaron J. Young. Real-Time Neural Network-Based Gait Phase Estimation Using a Robotic Hip Exoskeleton. *IEEE Trans. Medical Robotics and Bionics*, 2(1):28–37, 2019.

[89] Inseung Kang, Dean Molinaro, Srijan Duggal, Yanrong Chen, Pratik Kunapuli, and Aaron Young. Real-time gait phase estimation for robotic hip exoskeleton control during multimodal locomotion. *IEEE Robotics and Automation Letters*, 6(2):3491–3497, 2021.

[90] Roopak M Karulkar and Patrick M Wensing. Using footsteps to estimate changes in the desired gait speed of an exoskeleton user. *IEEE Robot. Autom. Lett.*, 6(4):6781–6788, 2021.

[91] H. Kazerooni. Exoskeletons for human power augmentation. *IEEE/RSJ Int. Conf. Intelligent Robots and Systems, IROS*, pages 3120–3125, 2005.

[92] Jinsoo Kim, Giuk Lee, Roman Heimgartner, Dheepak Arumukhom Revi, Nikos Karavas, Danielle Nathanson, Ignacio Galiana, Asa Eckert-Erdheim, Patrick Murphy, David Perry, Nicolas Menard, Dabin Kim Choe, Philippe Malcolm, and Conor J. Walsh. Reducing the metabolic rate of walking and running with a versatile, portable exosuit. *Science*, 365(6454):668–672, 2019.

[93] William R. King and Jun He. A meta-analysis of the technology acceptance model. *Information and Management*, 43(6):740–755, 2006.

[94] Frederick Kingdom and Nicolaas Prins. *Psychophysics: A Practical Introduction*, volume 53. 2013.

[95] Jeffrey R. Koller, Deanna H. Gates, Daniel P. Ferris, and C. David Remy. Confidence in the curve: Establishing instantaneous cost mapping techniques using bilateral ankle exoskeletons. *J. Applied Physiology*, 122(2):242–252, 2016.

[96] Tsuyoshi Kuroda and Emi Hasuo. The very first step to start psychophysical experiments. *Acoustical Science and Technology*, 35(1):1–9, 2013.

[97] Malte Kuss, Frank Jäkel, and Felix A. Wichmann. Bayesian inference for psychometric functions. *J. Vision*, 5(5):478–492, 2005.

[98] Dawit Lee, Inseung Kang, Dean D Molinaro, Alexander Yu, and Aaron J Young. Real-time user-independent slope prediction using deep learning for modulation of robotic knee exoskeleton assistance. *IEEE Robot. Autom. Lett.*, 2021.

[99] Sangjun Lee, Jinsoo Kim, Lauren Baker, Andrew Long, Nikos Karavas, Nicolas Menard, Ignacio Galiana, and Conor J Walsh. Autonomous multi-joint soft exosuit with augmentation-power-based control parameter tuning reduces energy cost of loaded walking. *Journal of neuroengineering and rehabilitation*, 15(1):1–9, 2018.

[100] Ung Hee Lee, Chen-Wen Pan, and Elliott J Rouse. Empirical characterization of a high-performance exterior-rotor type brushless dc motor and drive. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8018–8025. IEEE, 2019.

[101] Tommaso Lenzi, Maria Chiara Carrozza, and Sunil K Agrawal. Powered hip exoskeletons can reduce the user's hip and ankle muscle activations during walking. *IEEE Trans Neural Syst Rehabil Eng*, 21(6):938–948, 2013.

[102] H. Levitt. Transformed Up-Down Methods in Psychoacoustics. *The J. the Acoustical Society of America*, 49(2B):467–477, 1971.

[103] Cara L. Lewis and Daniel P. Ferris. Invariant hip moment pattern while walking with a robotic hip exoskeleton. *Journal of Biomechanics*, 44(5):789–793, 2011.

[104] Zhijun Li, Kuankuan Zhao, Longbin Zhang, Xinyu Wu, Tao Zhang, Qinjian Li, Xiang Li, and Chun-Yi Su. Human-in-the-loop control of a wearable lower limb exoskeleton for stable dynamic walking. *IEEE/ASME transactions on mechatronics*, 26(5):2700–2711, 2020.

[105] Bokman Lim, Jusuk Lee, Junwon Jang, Kyungrock Kim, Young Jin Park, Keehong Seo, and Youngbo Shim. Delayed output feedback control for gait assistance with a robotic hip exoskeleton. *IEEE Transactions on Robotics*, 35(4):1055–1062, 2019.

[106] Ming Liu, Fan Zhang, and He Helen Huang. An adaptive classification strategy for reliable locomotion mode recognition. *Sensors*, 17(9), 2017.

[107] David Lucking-Reiley. Vickrey auctions in practice: From nineteenth-century philately to twenty-first-century E-commerce. *J. Economic Perspectives*, 14(3):183–192, 2000.

[108] Rebecca Macaluso, Kyle Embry, Dario J Villarreal, and Robert D Gregg. Parameterizing human locomotion across quasi-random treadmill perturbations and inclines. *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 29:508–516, 2021.

[109] B. J. Makinson. Research and development prototype for machine augmentation of human strength and endurance. hardiman i project. Technical report, 1917.

[110] Philippe Malcolm, Wim Derave, Samuel Galle, and Dirk De Clercq. A Simple Exoskeleton That Assists Plantarflexion Can Reduce the Metabolic Cost of Human Walking. *PLoS ONE*, 8(2):1–7, 2013.

[111] Takehito Masuda, Ryo Mikami, Toyotaka Sakai, Shigehiro Serizawa, and Takuma Wakayama. The net effect of advice on strategy-proof mechanisms: An experiment for the vickrey auction. *Experimental Economics*, pages 1–40, 2022.

[112] Kirsty McDonald. Multi-objective prioritization in human walking. 2018.

[113] Roberto L Medrano, Gray C Thomas, and Elliott Rouse. Quantifying perception of metabolic changes from exoskeleton assistance. *Dynamic Walking 2018*.

[114] Roberto L Medrano, Gray C Thomas, and Elliott Rouse. Methods for measuring the just noticeable difference for variable stimuli: implications for perception of metabolic rate with exoskeleton assistance. In *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, pages 483–490. IEEE, 2020.

[115] Roberto Leo Medrano, Elliott J Rouse, and Gray Cortright Thomas. Biological Joint Loading and Exoskeleton Design. *IEEE Transactions on Medical Robotics and Bionics*, pages 1–4.

[116] Roberto Leo Medrano, Gray Cortright Thomas, Connor G. Keais, Elliott J. Rouse, and Robert D. Gregg. Real-time gait phase and task estimation for controlling a powered ankle exoskeleton on extremely uneven terrain. *IEEE Transactions on Robotics*, pages 1–13, 2023.

[117] Roberto Leo Medrano, Gray Cortright Thomas, and Elliott J Rouse. Can humans perceive the metabolic benefit provided by augmentative exoskeletons? *Journal of NeuroEngineering and Rehabilitation*, 19(1):1–13, 2022.

[118] Roberto Leo Medrano, Gray Cortright Thomas, Elliott J Rouse, and Robert D Gregg. Analysis of the bayesian gait-state estimation problem for lower-limb wearable robot sensor configurations. *IEEE Robotics and Automation Letters*, 7(3):7463–7470, 2022.

[119] Roberto Leo Medrano, Gray Cortright Thomas, Elliott J. Rouse, and Robert D. Gregg. Real-time phase and task estimation for controlling a powered ankle exoskeleton on extremely uneven terrain [Source Code], 2022.

[120] Luke M. Mooney and Hugh M. Herr. Biomechanical walking mechanisms underlying the metabolic reduction caused by an autonomous exoskeleton. *J. NeuroEngineering and Rehabilitation*, 13(1):1–12, 2016.

[121] Luke M Mooney, Cara H Lai, and Elliott J Rouse. Design and Characterization of a Biologically Inspired Quasi- Passive Prosthetic Ankle-Foot. *Conf Proc IEEE Eng Med Biol Soc*, 02139:1611–1617, 2014.

[122] Luke M Mooney, Elliott J Rouse, and Hugh M Herr. Autonomous exoskeleton reduces metabolic cost of human walking during load carriage. *J. NeuroEngineering and Rehabilitation*, 11(1):1–5, 2014.

[123] Luke M. Mooney, Elliott J. Rouse, and Hugh M. Herr. Autonomous exoskeleton reduces metabolic cost of walking. *Conf Proc IEEE Eng Med Biol Soc*, 11(1):3065–3068, 2014.

[124] Hans-georg Muller and Thomas Schmitt. Choice of Number of Doses for Maximum Likelihood Estimation of the ED50 for Quantal Dose-Response Data Author (s): Hans-Georg Müller and Thomas Schmitt Published by : International Biometric Society Stable URL : https://www.jstor.org/stable/2531635 REFE. 46(1):117–129, 1990.

[125] Helen R Neill, Ronald G Cummings, Philip T Ganderton, Glenn W Harrison, and Thomas McGuckin. Hypothetical surveys and real economic commitments. *Land economics*, pages 145–154, 1994.

[126] Christopher Nesler, Gray Thomas, Nikhil Divekar, Elliott Rouse, and Robert D Gregg. Enhancing voluntary motion with modular, backdrivable, powered hip and knee orthoses. *IEEE Robot. Autom. Lett.*, 2022.

[127] Kenneth H. Norwich. On the theory of Weber fractions. *Perception & Psychophysics*, 42(3):286–298, 1987.

[128] RW Nuckols, S Lee, K Swaminathan, D Orzel, RD Howe, and CJ Walsh. Individualization of exosuit assistance based on measured muscle dynamics during versatile walking. *Science Robotics*, 6(60):eabj1362, 2021.

[129] Sangjo Oh, Joongho Ahn, and Beomsoo Kim. Adoption of broadband internet in korea: the role of experience in building attitudes. *Journal of Information Technology*, 18(4):267–280, 2003.

[130] J. K. O'regan and R. Humbert. Estimating psychometric functions in forced-choice situations: Significant biases found in threshold and slope estimations when small samples are used. *Perception & Psychophysics*, 46(5):434–442, 1989.

[131] Amanda Louise Ryan Ortiz, Nicola Giovanelli, and Rodger Kram. The metabolic costs of walking and running up a 30-degree incline: implications for vertical kilometer foot races. *European Journal of Applied Physiology*, 117(9):1869–1876, 2017.

[132] Fausto A. Panizzolo, Gregory M. Freisinger, Nikos Karavas, Asa M. Eckert-Erdheim, Christopher Siviy, Andrew Long, Rebecca A. Zifchock, Michael E. LaFiandra, and Conor J. Walsh. Metabolic cost adaptations during training with a soft exosuit assisting the hip joint. *Sci. Rep.*, 9(1):1–10, 2019.

[133] Fausto A. Panizzolo, Ignacio Galiana, Alan T. Asbeck, Christopher Siviy, Kai Schmidt, Kenneth G. Holt, and Conor J. Walsh. A biologically-inspired multi-joint soft exosuit that can reduce the energy cost of loaded walking. *J. NeuroEngineering and Rehabilitation*, 13(1), 2016.

[134] Simon Parsons, Juan A Rodriguez-Aguilar, and Mark Klein. Auctions and bidding: A guide for computer scientists. *ACM Computing Surveys (CSUR)*, 43(2):1–59, 2011.

[135] D G Pelli and B Farell. Chapter 29 - Psychophysical Methods. *Handbook of Optics*, pages 1–13, 1996.

[136] Xiangyu Peng, Yadrianna Acosta-Sojo, Man I Wu, and Leia Stirling. Actuation timing perception of a powered ankle exoskeleton and its associated ankle angle changes during walking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:869–877, 2022.

[137] N Prins. Kingdom fa (2018). *Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. Frontiers in Psychology*, 9(1250):10–3389.

[138] Nicolaas Prins and Frederick A A Kingdom. Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, 9:1250, 2018.

[139] Yuepeng Qian, Yining Wang, Chuheng Chen, Jingfeng Xiong, Yuquan Leng, Haoyong Yu, and Chenglong Fu. Predictive locomotion mode recognition and accurate gait phase estimation for hip exoskeleton on various terrains. *IEEE Robotics and Automation Letters*, 7(3):6439–6446, 2022.

[140] David Quintero, Dario J. Villarreal, Daniel J. Lambert, Susan Kapp, and Robert D. Gregg. Continuous-Phase Control of a Powered Knee-Ankle Prosthesis: Amputee Experiments Across Speeds and Inclines. *IEEE Trans. Robotics*, 34(3):686–701, 2018.

[141] Siavash Rezazadeh, David Quintero, Nikhil Divekar, Emma Reznick, Leslie Gray, and Robert D. Gregg. A Phase Variable Approach for Improved Rhythmic and Non-Rhythmic Control of a Powered Knee-Ankle Prosthesis. *IEEE Access*, 7:109840–109855, 2019.

[142] Emma Reznick, Kyle Embry, and Robert D. Gregg. Predicting Individualized Joint Kinematics over a Continuous Range of Slopes and Speeds. *Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatron*, pages 666–672, 2020.

[143] Emma Reznick, Kyle Embry, Ross Neuman, Edgar Bolivar-Nieto, Nicholas P. Fey, and Robert D. Gregg. Lower-limb kinematics and kinetics during continuously varying human locomotion. https://figshare.com/s/3d4f7d8641d8b0f81b77, 2021.

[144] Everett M Rogers. *Diffusion of innovations*. Simon and Schuster, 2010.

[145] Matthew C. Rousu and Jayson L. Lusk. Valuing information on GM foods in a WTA market: What information is most valuable? *AgBioForum*, 12(2):226–231, 2009.

[146] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2:e55, 2016.

[147] Galle Samuel, Malcolm Philippe, and De Clercq Dirk. 2D Parameter sweep of bilateral exoskeleton actuation. Number July, pages 2–3. Dynamic Walking, 2014.

[148] Simo Särkkä. *Bayesian filtering and smoothing*. Number 3. Cambridge University Press, 2013.

[149] G. S. Sawicki and D. P. Ferris. Mechanics and energetics of level walking with powered ankle exoskeletons. *J. Experimental Biology*, 211(9):1402–1413, 2008.

[150] Gregory S. Sawicki, Owen N. Beck, Inseung Kang, and Aaron J. Young. The exoskeleton expansion: Improving walking and running economy. *J. NeuroEng. and Rehabilitation*, 17(1):1–9, 2020.

[151] Gregory S. Sawicki and Daniel P. Ferris. A pneumatically powered knee-ankle-foot orthosis (KAFO) with myoelectric activation and inhibition. *J. NeuroEngineering and Rehabilitation*, 6(1):1–16, 2009.

[152] Johannes Scherr, Bernd Wolfarth, Jeffrey W. Christle, Axel Pressler, Stefan Wagenpfeil, and Martin Halle. Associations between Borg's rating of perceived exertion and physiological measures of exercise intensity. *European J. Applied Physiology*, 113(1):147–155, 2013.

[153] Wolfram Schultz. Subjective neuronal coding of reward: Temporal value discounting and risk. *European J. Neuroscience*, 31(12):2124–2135, 2010.

[154] Jessica C. Selinger and J. Maxwell Donelan. Estimating instantaneous energetic cost during non-steady-state gait. *J. Applied Physiology*, 117(11):1406–1415, 2014.

[155] Jessica C. Selinger, Shawn M. O'Connor, Jeremy D. Wong, and J. Maxwell Donelan. Humans Can Continuously Optimize Energetic Cost during Walking. *Current Biology*, 25(18):2452–2456, 2015.

[156] Jessica C. Selinger, Jeremy D. Wong, Surabhi N. Simha, and J. Maxwell Donelan. How people initiate energy optimization and converge on their optimal gaits. *The J. Experimental Biology*, (September):jeb.198234, 2019.

[157] Keehong Seo, Young Jin Park, Jusuk Lee, Seungyong Hyung, Minhyung Lee, Jeonghun Kim, Hyundo Choi, and Youngbo Shim. RNN-based on-line continuous gait phase estimation from shank-mounted IMUs to control ankle exoskeletons. *IEEE Int. Conf. Rehabilitation Robotics*, pages 809–815, 2019.

[158] Max Shepherd, Dean Molinaro, Gregory Sawicki, and Aaron Young. Deep learning enables exoboot control to augment variable-speed walking. *IEEE Robot. Autom. Lett.*, 2022.

[159] Max K. Shepherd, Alejandro F. Azocar, Matthew J. Major, and Elliott J. Rouse. Amputee perception of prosthetic ankle stiffness during locomotion. *J. NeuroEngineering and Rehabilitation*, 15(1):1–10, 2018.

[160] Max K. Shepherd, Alejandro F. Azocar, Matthew J. Major, and Elliott J. Rouse. The Difference Threshold of Ankle-Foot Prosthesis Stiffness for Persons with Transtibial Amputation. *Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatron*, 2018-Augus:100–104, 2018.

[161] Max K. Shepherd and Elliott J. Rouse. Comparing preference of ankle–foot stiffness in below-knee amputees and prosthetists. *Scientific Reports*, 10(1):16067, dec 2020.

[162] Cathy A. Simpson and Rudy E. Vuchinich. Reliability of a measure of temporal discounting. *Psychological Record*, 50(1):3–16, 2000.

[163] Seungmoon Song and Steven H Collins. Optimizing exoskeleton assistance for faster self-selected walking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:786–795, 2021.

[164] Saivimal Sridar, Pham H. Nguyen, Mengjia Zhu, Quoc P. Lam, and Panagiotis Polygerinos. Development of a soft-inflatable exosuit for knee rehabilitation. *IEEE International Conference on Intelligent Robots and Systems*, 2017-Septe:3722–3727, 2017.

[165] Jacopo Staiano, Nuria Oliver, Bruno Lepri, Rodrigo De Oliveira, Michele Caraviello, and Nicu Sebe. Money walks: A human-centric study on the economics of personal mobile data. *UbiComp 2014 - Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 583–594, 2014.

[166] S. S. Stevens. On the psychophysical law. *Psychological Review*, 64(3):153–181, 1957.

[167] S S Stevens. Psychological Bulletin PROBLEMS AND METHODS OF PSYCHOPHYSICS 1. 55(4), 1958.

[168] Stanley S Stevens and Eugene H Galanter. Ratio scales and category scales for a dozen perceptual continua. *Journal of experimental psychology*, 54(6):377, 1957.

[169] Leia Stirling, Ho Chit Siu, Eric Jones, and Kevin Duda. Human factors considerations for enabling functional use of exosystems in operational environments. *IEEE Systems Journal*, 13(1):1072–1083, 2018.

[170] Nassim Nicholas Taleb. *The black swan: The impact of the highly improbable*, volume 2. Random house, 2007.

[171] Lennart P.J. Teunissen, Alena Grabowski, and Rodger Kram. Effects of independently altering body weight and body mass on the metabolic cost of running. *Journal of Experimental Biology*, 210(24):4418–4427, 2007.

[172] Nitish Thatte, Helei Duan, and Hartmut Geyer. A sample-efficient black-box optimizer to train policies for human-in-the-loop systems with user preferences. *IEEE Robotics and Automation Letters*, 2(2):993–1000, 2017.

[173] Nitish Thatte, Tanvi Shah, and Hartmut Geyer. Robust and adaptive lower limb prosthesis stance control via extended kalman filter-based gait phase estimation. *IEEE Robotics and Automation Letters*, 4(4):3129–3136, 2019.

[174] Gray Cortright Thomas, Orion Campbell IV, Nick Nichols, Nicholas Brissoneau, Binghan He, Joshua James, Nicholas Paine, and Luis Sentis. An Amplification Shaping Framework for Exoskeletal Human Strength Augmentation. 00(0):1–18, 2019.

[175] Gray Cortright Thomas, Jeremiah M Coholich, and Luis Sentis. Compliance shaping for control of strength amplification exoskeletons with elastic cuffs. In *Proc. 2019 IEEE/ASME Int. Conf. Advanced Intelligent Mechatronics*, pages 1199–1206. IEEE and ASME, July 2019.

[176] Warren S Torgerson. Theory and methods of scaling. 1958.

[177] Bernhard Treutwein. Adaptive psychophysical procedures. *Vision Research*, 35(17):2503–2522, 1995.

[178] Bernhard Treutwein and Hans Strasburger. Fitting the psychometric function. *Perception and Psychophysics*, 61(1):87–106, 1999.

[179] Maegan Tucker, Ellen Novoseller, Claudia Kann, Yanan Sui, Yisong Yue, Joel Burdick, and Aaron D. Ames. Preference-Based Learning for Exoskeleton Gait Optimization. 2019.

[180] Michael R Tucker, Jeremy Olivier, Anna Pagel, Hannes Bleuler, Mohamed Bouri, Olivier Lambercy, José del R Millán, Robert Riener, Heike Vallery, and Roger Gassert. Control strategies for active lower extremity prosthetics and orthotics: a review. *J. Neuroeng. Rehabilitation*, 12(1), 2015.

[181] Hal R. Varian and Christopher Harris. The VCG auction in theory and practice. *American Economic Review*, 104(5):442–445, 2014.

[182] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, 2017.

[183] William Vickrey. Counterspeculation, Auctions, and Competitive Sealed Tenders. *The J. Finance*, 16(1):8–37, 1961.

[184] Dario J Villarreal, Hasan A Poonawala, and Robert D Gregg. A robust parameterization of human gait patterns across phase-shifting perturbations. *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 25(3):265–278, 2017.

[185] Eric A Wan and Rudolph Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pages 153–158, 2000.

[186] Dong Wei, Zhijun Li, Qiang Wei, Hang Su, Bo Song, Wei He, and Jianqiang Li. Human-in-the-loop control strategy of unilateral exoskeleton robots for gait rehabilitation. *IEEE Transactions on Cognitive and Developmental Systems*, 13(1):57–66, 2019.

[187] Matthew L Weiss, Randy C Paffenroth, and Joshua R Uzarski. The autoencoder-kalman filter: Theory and practice. In *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pages 2176–2179. IEEE, 2019.

[188] Felix A Wichmann and N Jeremy Hill. The psychometric function: I. Fitting, sampling, and goodness of fit. 63(8):1293–1313, 2001.

[189] Felix A. Wichmann and N. Jeremy Hill. The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics*, 63(8):1314–1329, 2001.

[190] Felix A. Wichmann and Frank Jäkel. *Methods in Psychophysics*. 2018.

[191] Craig C. Wier, Walt Jesteadt, and David M. Green. A comparison of method-of-adjustment and forced-choice procedures in frequency discrimination. *Perception & Psychophysics*, 19(1):75–79, 1976.

[192] Wikipedia contributors. Bernstein polynomial — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Bernstein_polynomial&oldid=1033750467, 2021. [Online; accessed 31-August-2021].

[193] Kirby A. Witte, Pieter Fiers, Alison L. Sheets-Singer, and Steven H. Collins. Improving the energy economy of human running with powered and unpowered ankle exoskeleton assistance. *Science Robotics*, 5(40), 2020.

[194] Jeremy D Wong, Shawn M O'Connor, Jessica C. Selinger, and J. Maxwell Donelan. Contribution of blood oxygen and carbon dioxide sensing to the energetic optimization of human walking. *J. Neurophysiology*, page jn.00195.2017, 2017.

[195] Jeremy D. Wong, Jessica C. Selinger, and J. Maxwell Donelan. Is natural variability in gait sufficient to initiate spontaneous energy optimization in human walking? *J. neurophysiology*, 121(5):1848–1855, 2019.

[196] Aaron J. Young and Levi J. Hargrove. A classification method for user-independent intent recognition for transfemoral amputees using powered lower limb prostheses. *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 24(2):217–225, 2016.

[197] Aaron J. Young, Ann M. Simon, Nicholas P. Fey, and Levi J. Hargrove. Classifying the intent of novel users during human locomotion using powered lower limb prostheses. *Int IEEE EMBS Conf Neural Eng*, pages 311–314, 2013.

[198] Aaron J Young, Ann M Simon, and Levi J Hargrove. A training method for locomotion mode prediction using powered lower limb prostheses. *IEEE Trans Neural Syst Rehabil Eng*, 22(3):671–677, 2013.

[199] Elizabeth H. Zandstra, Krishna P. Miyapuram, and Philippe N. Tobler. *Understanding consumer decisions using behavioral economics*, volume 202. Elsevier B.V., 1 edition, 2013.

[200] Binquan Zhang, Min Zhou, Wanlu Xu, et al. An adaptive framework of real-time continuous gait phase variable estimation for lower-limb wearable robots. *Robotics and Autonomous Systems*, 143:103842, 2021.

[201] Juanjuan Zhang, Chien Chern Cheah, and Steven H. Collins. *Torque Control in Legged Locomotion*. Elsevier Inc., 1 edition, 2017.

[202] Juanjuan Zhang, Pieter Fiers, Kirby A Witte, Rachel W Jackson, Katherine L Poggensee, Christopher G Atkeson, and Steven H Collins. Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, 1284(June):1280–1284, 2017.

[203] Wenchao Zhang, Dongyan Wei, and Hong Yuan. The improved constraint methods for foot-mounted pdr system. *Ieee Access*, 8:31764–31779, 2020.

[204] H. Zhu, C. Nesler, N. Divekar, M. Ahmad, and R.D. Gregg. Design and Validation of a Partial-Assist Knee Orthosis with Compact, Backdrivable Actuation. *IEEE Int. Conf. Rehab. Robot.*, 2019.

[205] Hanqi Zhu, Jack Doan, Calvin Stence, Ge Lv, Toby Elery, and Robert Gregg. Design and validation of a torque dense, highly backdrivable powered knee-ankle orthosis. *IEEE Int. Conf. Robotics and Automation*, pages 504–510, 2017.

[206] Hanqi Zhu, Christopher Nesler, Nikhil Divekar, Vamsi Peddinti, and Robert Gregg. Design principles for compact, backdrivable actuation in partial-assist powered knee orthoses. *IEEE/ASME Trans. Mechatronics*, 2021.

[207] Qin Zou, Lihao Ni, Qian Wang, Qingquan Li, and Song Wang. Robust gait recognition by integrating inertial and RGBD sensors. *IEEE Trans. Cybernetics*, 48(4):1136–1150, 2017.