BAYESIAN OPTIMIZATION STRATEGIES FOR HUMAN-IN-THE-LOOP SYSTEMS: THEORY AND APPLICATIONS IN PHYSICAL HUMAN-ROBOT-INTERACTION

by
HARUN TOLASA

Submitted to
the Faculty of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University
July 2024

ABSTRACT

BAYESIAN OPTIMIZATION STRATEGIES FOR
HUMAN-IN-THE-LOOP SYSTEMS:
THEORY AND APPLICATIONS IN
PHYSICAL HUMAN-ROBOT-INTERACTION

HARUN TOLASA

Mechatronics Engineering, M.Sc. Thesis, July 2024

Thesis Supervisor: Prof. Volkan Patoglu

Keywords: Human-in-the-loop experiments, Bayesian optimization, Gaussian process, haptic rendering, psycho-physics, multi-criteria optimization, assist-as-needed controllers, rehabilitation robotics, physical human-robot interaction.

Physical human-robot interaction (pHRI) involves direct physical interaction between humans and robots, sharing a workspace and performing tasks involving touch, force, and movement. Such interactions are essential for collaborative, surgical, wearable, and rehabilitation robots. Human-in-the-loop (HiL) studies are widely used in pHRI to develop safer, more collaborative, efficient, and personalized robots. HiL studies gather quantitative and qualitative evaluations based on user experiences to iteratively improve robot performance and operation. However, data collected from participants in HiL optimization studies are often limited due to fatigue and the limited attention span of users.

This thesis proposes sample-efficient Bayesian strategies for HiL optimization and explores their novel applications in various pHRI domains. In particular, this study (i) introduces Bayesian optimization (BO) based on qualitative user preferences to optimize the perceived realism of haptic rendering under conflicting haptic-visual cues. Then, it (ii) extends the Bayesian approach to compute Pareto solutions in HiL multi-criteria optimization problems with qualitative and quantitative performance metrics to customize the assistance level during

human motor learning tasks. Finally, (iii) it incorporates transfer learning to improve the sample efficiency of HiL optimizations and demonstrates the improved efficiency due to transfer learning during personalization of the assistance level provided to various upper-extremity movements using a rehabilitation robot.

First, we develop a qualitative feedback-based HiL optimization technique using sample-efficient Bayesian optimization that leverages qualitative pairwise comparisons and ordinal classifications. We apply this technique to investigate visual-haptic cue integration during multi-modal haptic rendering under conflicting cues, establishing a systematic approach to determine the optimal visual scaling for haptic manipulation to maximize the perceived realism of spring rendering for a given haptic interface. Our results demonstrate that parameters affecting visual-haptic congruency can be effectively optimized using our HiL optimization technique, ensuring consistently high perceived realism. Consequently, multi-modal perception can be successfully enhanced by solely adjusting visual feedback without altering haptic feedback, thereby extending the range of perceived stiffness levels for a haptic interface. We extend our findings to a group of individuals to capture a multi-dimensional psychometric field that characterizes the cumulative effect of feedback modalities utilized during sensory cue integration under conflicts.

Next, we extend HiL studies with multi-criteria BO, addressing the need for considering multiple and often conflicting optimization objectives in pHRI applications. We propose a HiL Pareto optimization approach to characterize the trade-off between the performance and the perceived challenge level of motor learning tasks to be used in assist-as-needed control. Our optimization employs a hybrid model that captures the user's performance by a quantitative metric, while the perceived challenge level is modeled as a qualitative metric gathered through preference-based qualitative feedback. Once the trade-off is characterized via HiL Pareto optimization, we demonstrate how this trade-off can guide the design of assist-as-needed training sessions of a motor learning task with optimal assistance levels. Furthermore, we provide evidence that the trade-off evolves as learning takes place, and the set of non-dominated solutions of multi-criteria optimization can be used to fairly evaluate the training progress, even when the user cannot perform the task without assistance. We show the feasibility and usefulness of the proposed approach through a case study involving a virtual manual skill training task administered to healthy individuals with haptic feedback.

Finally, we further enhance our HiL methods by incorporating transfer learning strategies to accelerate the convergence of the optimization, by leveraging data from previous experiments. These strategies significantly improve sample efficiency and reduce the need for extensive new data collection, addressing the impracticality of conducting HiL trials from scratch for each new task in pHRI applications. We demonstrate the applicability of transfer learning to multi-criteria BO through a HiL experiment conducted to personalize the assistance levels for robot-assisted upper-extremity rehabilitation with a diverse range of physical movements. In these optimizations, we utilize a quantitative metric to evaluate user effort and a qualitative metric to assess the user's perceived comfort. Once an initial Pareto curve is computed for a task through HiL optimization for the personalized assistance levels, we transfer these outcomes to subsequent tasks by correlating the similarities among physical tasks. This systematic method of transferring knowledge eliminates the need for empirical or sample-dependent correlation methodologies. Our results demonstrate that applying transfer learning can significantly accelerate HiL optimizations, making the HiL experiments more efficient and effective.

Overall, this thesis introduces novel BO strategies that can significantly broaden the scope of HiL solutions in pHRI applications, enhancing their applicability and feasibility. By systematically demonstrating the proposed methodologies across multiple HiL-based pHRI studies, this work not only showcases their effectiveness but also provides a comprehensive framework that can be adapted for HiL studies in various other pHRI fields. The insights and techniques presented in this thesis serve as a valuable guide for future research and development, paving the way for more efficient, personalized, and effective pHRI.

ÖZET

İNSANIN OPTİMİZASYON DÖNGÜSÜDE OLDUĞU SİSTEMLER İÇİN BAYES OPTİMİZASYON STRATEJİLERİ: TEORİ VE FİZİKSEL İNSAN-ROBOT ETKİLEŞİMİ UYGULAMALARI

HARUN TOLASA

Mekatronik Mühendisliği, Yüksek Lisans Tezi, Temmuz 2024

Tez Danışmanı: Prof. Dr. Volkan Patoğlu

Anahtar Kelimeler: İnsanın optimizasyon döngüsünde olduğu deneyler, Bayes optimizasyonu, Gauss süreci, dokunsal geri bildirim, psikofizik deneyler, çok kriterli optimizasyon, ihtiyaç kadar destek veren kontrolcüler, rehabilitasyon robotları, fiziksel insan-robot etkileşimi

Fiziksel insan-robot etkileşimi (fİRE), insanlar ve robotlar arasında bir çalışma alanını paylaşmayı ve insanla robot arasında güç iletimi gerektiren görevlerin yerine getirilmesini içerir. Bu tür etkileşimler işbirlikçi, cerrahi, giyilebilir ve rehabilitasyon robot uygulamaları için gereklidir. fİRE uygulamalarında, insanın optimizasyon döngüsünde olduğu çalışmalar, daha güvenli, işbirlikçi, verimli ve kişiselleştirilmiş robotlar geliştirmek için yaygın olarak kullanılmaktadır. İnsanın optimizasyon döngüsünde olduğu çalışmalarda, robot performansını sürekli olarak iyileştirmek için kullanıcı deneyimlerine dayalı niceliksel ve niteliksel değerlendirmeler toplanır. Ancak, bu tür çalışmalarda kullanıcıların yorulması ve dikkat sürelerinin sınırlı olması gibi nedenlerle kullanıcılardan toplanabilen veri miktarı kısıtlıdır.

Bu tez kapsamında, insanın optimizasyon döngüsünde olduğu çalışmalar için, numune verimli Bayes stratejileri önerilip bu yöntemlerin çeşitli fİRE uygulamalarındaki kullanımı araştırılmaktadır. İlk olarak, (i) çelişkili dokunsal-görsel ipuçları altında gerçekleştirilen dokunsal oluşturmanın algılanan gerçekçiliğini arttırmak için, niteliksel kullanıcı tercihlerine dayalı, Bayes optimizasyon (BO) yöntemi önerilmekte; sonrasında, (ii) insanın motor öğrenmesini hızlandırmak

amacıyla uygulanan yardım düzeyini kişiselleştirmek için, çok kriterli bir optimizasyon problemi ele alınmakta ve bu problemin çözümleri için, insanın optimizasyon döngüsünde olduğu ve hem niteliksel hem de niceliksel performans ölçümlerini ele alan, yenilikçi bir Pareto Bayes yaklaşımı önerilmekte; son olarak, (iii) insanın optimizasyon döngüsünde olduğu deneylerin numune verimliliğini artırmak için, aktarımlı öğrenim yöntemi uygulanmakta ve bu yöntemin etkinliği robot destekli üst ekstremite rehabilitasyonunda destek miktarının kişiselleştirilmesiyle gösterilmektedir.

İlk kısımda, ikili karşılaştırmalardan ve sıralı sınıflandırmalardan yararlanan, numune verimli Bayes optimizasyonu ele alınmış ve insanın optimizasyon döngüsünde olduğu niteliksel geri bildirime dayalı bir yöntem geliştirilmiştir. Bu yöntem, çelişen ipuçları altında çok modlu dokunsal oluşturma sırasında görsel-dokunsal işaret entegrasyonunu araştırmak için uygulanmış ve bir dokunsal arayüz ile yay oluşturmanın algılanan gerçekçiliğini en üst düzeye çıkarmak amacıyla, en uygun görsel ölçeklendirmeyi belirlemek üzere sistematik bir yaklaşım olarak ortaya konulmuştur. Sonuçlarımız, görsel-dokunsal uyumu etkileyen parametrelerin, önerilen yöntem kullanılarak, etkili bir şekilde artırılabileceğini ve tutarlı şekilde yüksek bir algılanan gerçekçiliğin sağlanabileceğini göstermektedir. Çok modlu algı, dokunsal geri bildirimi değiştirmeden yalnızca görsel geri bildirimi ayarlayarak, başarılı bir şekilde değiştirilebilmekte ve böylece dokunsal bir arayüz için algılanan sertlik seviye aralığı genişletilebilmektedir. Ayrıca bu çalışmadan elde edilen sonuçlar genellenerek, bir grup bireyin çelişen ipuçları altında duyusal işaret entegrasyonunda kullandığı yöntemlerin kümülatif etkisini karakterize eden çok boyutlu bir psikometrik alan belirlenmiştir.

İkinci kısımda, fİRE uygulamalarında genellikle birden fazla ve birbiriyle çelişen optimizasyon hedeflerine ihtiyaç duyulması nedeniyle, insanın optimizasyon döngüsünde olduğu çalışmalar için, karma performans ölçümlerine imkân veren, yenilikçi bir Pareto BO yaklaşımı önerilmiştir. İnsanın optimizasyon döngüsünde olduğu Pareto yaklaşımı, motor öğrenme görevlerinde performans ile algılanan zorluk seviyesi arasındaki ödünleşimi karakterize etmek ve kişiye özel şekilde ihtiyaç kadar destek sağlamak için kullanılmıştır. Önerilen yaklaşım, kullanıcının performansı için niceliksel ölçümlere, algılanan zorluk düzeyi içinse niteliksel kullanıcı geri bildirimine dayalı karma bir model üzerinde çalışmakta ve insanın optimizasyon döngüsünde olduğu Pareto yoluyla ödünleşimi karakterize etmektedir. Çelişen hedefler arasındaki ödünleşim, motor öğrenme görevinde ihtiyaç kadar desteğin belirlenmesine ve eğitim oturumlarında en uygun yardım seviyelerinde destek sağlanmasına imkân vermektedir. Üstelik, öğrenme gerçekleştikçe ödünleşim değişime uğramaktadır ve farklı zamanlara karakterize edilen Pareto çözümleri bu değişimi yakalayıp kullanıcının gelişiminin, yardım olmadan öğrenme görevini gerçekleştiremediği durumlarda bile, adil bir şekilde değerlendirilebilmesini sağlamaktadır. Önerilen yaklaşımın uygulanabilirliğini, sağlıklı bireylerle gerçekleştirilen sanal bir el becerisi eğitimi vaka çalışması aracılığıyla gösterilmiştir.

Son kısımda, aktarımlı öğrenme stratejileri, insanın optimizasyon döngüsünde olduğu yöntemlere eklenmiş ve önceki deneylerden elde edilen verilerden yararlanılarak deneylerin numune verimi arttırılmıştır. Aktarımlı öğrenme stratejileri sayesinde, insanın optimizasyon döngüsünde olduğu fİRE uygulamalarında, kapsamlı yeni veri toplama ihtiyacını azaltılmış ve farklı görevler içeren uygulamaların daha etkin olarak yürütülebilmesi sağlanmıştır. Robot destekli üst ekstremite rehabilitasyonu örnek bir uygulama olarak ele alınmış ve aktarımlı öğrenmenin, insanın optimizasyon döngüsünde olduğu Pareto BO ile, farklı fiziksel tedavi egzersizleri için, hastaya uygulanan destek seviyelerinin belirlenmesinde, etkin olarak uygulanabilirliği gösterilmiştir. Deneylerde, kullanıcı çabasını değerlendirmek için niceliksel bir metrik, kullanıcının algılanan konforunu değerlendirmek için ise niteliksel bir metrik kullanılmış ve fiziksel tedavi egzersizleri için insanın optimizasyon döngüsünde olduğu deneyler aracılığıyla başlangıç Pareto çözümleri hesaplanıp, bu çözümler hareketler arası benzerlikler doğrultusunda farklı egzersizlere aktarılmıştır. Hareketler arası benzerlikleri baz alan bu sistematik bilgi aktarım yöntemiyle ampirik veya numuneye bağlı korelasyon metodolojilerine olan ihtiyacı ortadan kaldırılmıştır. Sonuçlarımız, aktarımlı öğrenmenin insanın optimizasyon içinde olduğu deneyleri önemli seviyede hızlandırarak süreci daha verimli ve etkili hale getirebileceğini göstermektedir.

Sonuç olarak, bu tezde, insanın optimizasyon döngüsünde olduğu yöntemlerin kapsamı önemli ölçüde genişletilmiş ve bu yöntemlerin fİRE kapsamında uygulanabilirliğini artıran yeni BO stratejileri önerilmiştir. Sunulan yöntemler daha verimli, kişiselleştirilmiş ve etkili fiziksel insan-robot etkileşimlerini için gelecekte bu alanda yapılacak çalışmalar için faydalı bir rehber niteliğindedir.

In loving memory of my father and to my beloved family…

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Volkan Patoğlu, for his continuous support and guidance throughout my Master's studies. I am especially grateful for his academic insights, countless research discussions, and his encouragement. I consider myself incredibly fortunate to have had the privilege of working with him, as his mentorship has not only significantly improved my capabilities and knowledge in research but also provided invaluable lessons in both academia and engineering.

I would like to extend my sincere gratitude to my thesis committee members, Prof. Dr. Çağatay Başdoğan and Asst. Prof. Dr. Melih Türkseven, for their dedication, genuine interest in my research, and the valuable feedback they offered during my thesis defense.

I am also thankful to the HMI group members and alumni: Alperen Kenan, Bilal Çatkın, Görkem Gemalmaz, İlhami Osman Karakurt, Ömer Burak Aladağ, Özgür Taylan Kenanoğlu, Umut Kenanoğlu, and Uğur Mengilli. A special thanks to Bilal Çatkın and Uğur Mengilli for their collaboration and experience of working together. Also, a heartfelt thanks goes to all the participants who contributed their time and effort during my human experiments for my research.

I would also like to express my appreciation to my graduate friends Araz Haghnazari, Ali Enver Bilecen, Barış Almaç, Ceren Özsaltık, Cihan Arlı, Emre Yavaş, Feyza Teker, Işıl Demirkan, Mehmet Emin Mumcuoğlu, Mervenaz Şahin, Zeynep Özalp and Umut Barut for their friendship and support.

To my undergraduate colleagues, Muhammed Aybars Yalçın, İsmet Öztürk, Enise Kartal, and Arda Çalışkan, I am deeply grateful for the true friendship we have shared since our undergraduate years. I'm deeply grateful for all the moments we've shared even the exhausting times when we worked on projects or studied till morning. Despite everyone pursuing different paths in various places, I'm proud that we have managed to stay connected.

To my long-lasting friends Özgürcan Arslan and Aslı Çönkü, I am grateful for all the fun moments we've shared and for always supporting me whenever I needed it.

I would also like to thank my friends at Sabancı Outdoor Sports Club for the unforgettable camping and mountaineering experiences we've had.

Finally, and most importantly, I want to express my deepest gratitude to my family. Their unconditional support has been the foundation of my journey, and I especially thank my mother, father, and grandmother for their persistent love and encouragement throughout my studies. Without their support, none of this would have been possible.

This study was partially supported by TÜBİTAK Grant 22AG003.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Robotic systems are increasingly integrated with human activities to meet specific individual needs. Physical Human-Robot Interaction (pHRI) applications are expanding across healthcare, education, entertainment, and collaborative works.

Robots are utilized in rehabilitation and surgeries to automate and enhance healthcare. Assistive robots, such as upper or lower limb exoskeletons, are employed in rehabilitation [1], sports training [2], and aiding with daily tasks. These robots automate rehabilitation sessions, making them more feasible and reducing the need for constant physiotherapist involvement. Similarly, surgical robots are used in teleoperated surgeries to enable minimally invasive operations, allowing human experts to perform medical procedures with greater precision. Furthermore, haptic interfaces are integrated with surgical robotics for bilateral teleoperation, stimulating the sense of touch between the robot's end-effector and body tissue [3, 4].

Beyond their use in physical tasks, haptic interfaces are also used to stimulate virtual realities [5, 6]. These interfaces are used for haptic rendering which is a bilateral process making the dynamics of computationally rendered virtual environments (VEs) apparent to a human user. In most haptic rendering applications, haptic interfaces are paired with visual displays to simulate physical interactions to increase perceived realism. Notably, haptic studies have many promising applications in education such as robotic-guided medical training [7] and motor-task learning, improving the user experience and training outcomes.

As promising studies in pHRI continue to grow, the integration of robots with humans is becoming increasingly complicated, making personalization and human-centered design more critical than ever. Given the psychological and physical differences among individuals, humans can respond differently to the same settings of interactive devices. This variability highlights the importance of designing robotic systems to meet the unique needs and preferences of each user. Moreover, for robots to be widely accepted and effectively utilized, they must ensure that users feel both comfortable and safe during interactions. If a robot induces discomfort or a sense of insecurity, it is unlikely to gain user trust or widespread adoption.

To address the need for personalized and human-centered robotic systems, human-in-the-loop (HiL) approaches provide generalizable methodologies by employing online learning strategies [8, 9]. HiL approaches use human-sourced feedback to sequentially improve overall performance by conducting human trials. In these approaches, first, a human user experiences an interaction with a physical system. Then, based on the experience, the user provides quantitative or qualitative evaluations. These evaluations are processed by learning and optimization methodologies to make the system better for the user.

However, the HiL approach presents several challenges that cause many traditional reinforcement learning (RL) methodologies to be inapplicable. Firstly, individual-based objectives are often qualitative, including factors like safety, usability, and comfort. Secondly, individual responses can vary significantly under the same conditions, making simple aggregation of human evaluations impractical and leading to data insufficiency for traditional RL methods.

Humans often fail to evaluate pHRI systems numerically by specifying a reward as their evaluation strategy can be unreliable during HiL trials. Also, empirically quantifying qualitative metrics may cause pHRI systems to find a way to exploit rewards in an unintended manner. Although humans assign numerical rewards inconsistently while evaluating the performance of a system, they evaluate performance more consistently by providing qualitative feedback [10]. Mostly, qualitative feedback is acquired from humans with classifications and preferences by assigning discrimination tasks such as assigning a 2-alternative forced choice (2AFC) or ordinal ranking tasks. In particular, while deciding if robot behavior was "*good*" or "*bad*" during a HiL trial is an example of classification, deciding between trials "A" and "B" by comparing them is an example of preference.

In many applications of pHRI systems, optimizing a single criterion is often insufficient, as personalization requires balancing multiple design objectives, including feasibility, usability, comfort, safety, and performance. These design criteria frequently involve trade-offs as there are multiple metrics conflicting with each other. For example, reducing the cost of a physical device might also lower its performance. Similarly, enhancing user comfort could potentially compromise device efficiency or complexity. Such paradigms highlight the intricate challenges in personalizing pHRI systems, where the optimization process must consider and balance multiple, often conflicting, objectives to achieve the best overall outcome.

Besides the challenges with concurrent optimization objectives, another challenge in the HiL approach is insufficient available data. The convergence of learning and optimization includes sufficient sampling throughout the search space to reduce the uncertainty. Furthermore, an increment in the dimensionality of search space leads to more need for sampling known as the *curse of dimensionality.* Regardless of the objective of the HiL experiment, the available data are often limited due to human involvement. As the participants' energy, attention, and patience are limited, excessive participation usually causes negative effects on HiL approaches [11, 12]. For instance, if the trials of a wearable robot involve doing a physical task with the robot, excessive trials may cause fatigue or if the performance of the user is evaluated, then the reliability of the evaluations reduces as the user loses motivation to participate.

HiL studies are highly affected by insufficient sampling problems leading to the volume of the search space such as the curse of dimensionality. To reduce this uncertainty effectively, sample-efficient learning and optimization strategies are preferred in HiL studies [13]. Even more, HiL-related works have focused on maximizing information from each iteration. For instance, there are methods using information theory-based sampling strategies to predict the most beneficial parameters [14, 15]. Transfer learning is another strategy to increase sample efficiency as it is used to avoid the "*cold-start*" problem which refers to disregarding the usage of any built-in knowledge based on the prior experiment results.

We focus on Gaussian process (GP) models and Bayesian optimization techniques for optimization and learning strategies for HiL systems. BO is a commonly used method in HiL-related works as it is a non-parametric, derivative-free, global optimization method that can work with noisy inputs and provides analytical traceability [16]. BO is often utilized by a GP model in which the design of the model highly affects the efficiency of the optimization.

1.1 Contributions

In this thesis, we first introduce Bayesian optimization based on qualitative user preferences to enhance the perceived realism of haptic rendering under conflicting haptic-visual cues. We also extend this study to rigorously capture the multi-dimensional psychometric field, which characterizes the cumulative effect of feedback modalities during sensory cue integration under conflicts.

Next, we apply Bayesian optimization to compute Pareto solutions in HiL multi-criteria optimization problems, including qualitative and quantitative performance metrics, to tailor the assistance level during human motor learning.

Finally, we employ transfer learning to enhance the sample efficiency of multi-criteria HiL optimization and demonstrate its effectiveness in personalizing robot-assisted upper-extremity rehabilitation.

The main contributions of this thesis can be summarized as follows:

- We have developed a systematic method to determine the optimal visual scaling during multi-modal haptic rendering, maximizing the perceived realism of spring rendering through sample-efficient HiL optimization based on participants' qualitative feedback.
- We have extended non-parametric psychophysical methodologies to incorporate a wider range of classifications, enabling the collection of more information from users in each trial to build a more general and sample-efficient approach. Then, we refined these methods to suit multi-modal haptic rendering under visual-haptic incongruency, providing insights into multi-modal sensory integration by capturing the underlying psychometric field and perception thresholds.
- We have generalized our findings to a group of individuals to capture the underlying multi-dimensional psychometric field, characterizing the cumulative effects of feedback modalities utilized during sensory integration under conflicting cues. Results include reliable estimates of just noticeable

difference (JND) thresholds for stiffness under visual scaling and prominent features of sensory integration.

- We have demonstrated that HiL optimization allows for the determination of appropriate haptic-visual parameters, ensuring consistently high ratings of perceived realism in multi-modal rendering and providing a systematic and efficient means to study multi-modal sensory integration under conflicting cues.
- We have proposed a HiL Pareto optimization approach with hybrid (quantitative and qualitative) performance measures to characterize the trade-off between the performance and the perceived challenge level of a motor learning task.
- We have demonstrated how a sample design selection can be performed over the set of non-dominated solutions characterizing the trade-off between the performance and the perceived challenge level to effectively guide training with an AAN controller.
- We have proposed a novel and rigorous means of performance evaluation under assistance, through comparisons of the trade-off curves characterized at the different stages of training or among various users.
- We developed a novel methodology to transfer information from a HiL optimization to subsequent optimizations. Our approach customizes transfer learning for HiL studies by emphasizing the influence of online data over data obtained via earlier experiments. Furthermore, by characterizing the similarities among tasks, our approach eliminates the need for empirical and sample-dependent calculation of correlations.
- We have demonstrated the applicability of transfer learning boosted HiL multi-criteria optimization to personalize the assistance level during robot-assisted rehabilitation sessions.

1.2 Outline

The rest of the thesis is organized as follows:

Chapter 2 reviews related works with the Bayesian optimization and active learning strategies for human-in-the-loop experiments in physical human-robot interaction.

Chapter 3 presents the necessary statistical background for developing a simple Bayesian optimization model for both quantitative and qualitative metrics.

Chapter 4 presents an application of qualitative feedback-based Bayesian optimization for a perceptual study for improving the realism of haptic interfaces. The study presents a systematical method for increasing the realism of haptic stimuli by manipulating visuals. The study explores visual and haptic cue integration by modeling the perceived realism of participants involved in HiL experiments.

Chapter 5 introduces multi-criteria Bayesian optimization and its application in developing an assist-as-needed training session with a force-feedback controller for teaching a motor learning task in a virtual rendered environment. The study includes a performance comparison with an adaptive staircase-based training session.

Chapter 6 introduces a generalized transfer learning strategy to increase the efficiency of Bayesian optimization by reducing HiL trials. Furthermore, this chapter includes the application of transfer learning in developing personalized robot assisted rehabilitation session.

Chapter 7 summarizes the main findings and contributions of this thesis and discusses possible directions for future research, offering ideas on how our work can be further developed and improved to tackle challenges in pHRI.

Chapter 2

Literature Overview

This chapter reviews related works with Bayesian learning and optimization strategies and their applications in human-in-the-loop (HiL) experiments.

2.1 GP-Based Bayesian Optimization for a Quantitative Metric

Bayesian optimization (BO) is a sample-efficient global optimization approach often used with costly objective functions [16, 17, 18]. It is an iterative process that suggests parameters for trials and updates itself with the results of trials to suggest more promising parameters.

BO often relies on Gaussian Process (GP) models to predict outcomes by treating the latent function as a Gaussian random process. GP models use machine learning kernels to correlate parameters in the search space, predicting outcomes based on their relationships with previously measured parameters [19]. BO uses these GP predictions within acquisition functions to identify the most promising parameters for suggesting in each iteration.

The development of GP-based Bayesian optimization is not new [20], and it has many appliances in computer sciences such as hyper-parameter tuning of neural networks [21, 22]. However, it has recently gained popularity for real-time human-in-the-loop (HiL) optimization of assistive robotic devices, as it can significantly reduce the duration of these experiments. Bayesian optimization is used for reducing the metabolic cost during usage of assistive robotic devices [11, 12, 23], for which the evaluation of optimization metrics is costly or the number of trials is constrained by human involvement.

Zhang *et al.* used HiL Bayesian optimization for real-time adaptation of ankle exoskeleton [11]. Zahedi *et al.* [24] used Bayesian optimization to tune the variable damping of a controller based on human interaction.

2.2 GP-Based Bayesian Optimization for a Qualitative Metric

GP-based Bayesian optimization approaches have been extended to active learning with classifications instead of quantitative measurements [19, 25, 26]. These approaches often use a discriminative strategy [19] in which probabilistic functions are employed from decision theory.

GP models for binary classifications with "yes/no" type feedback have been introduced in [25] and it has been used for machine learning tasks such as image classification [27]. The binary classification method has been extended to ordinal classifications and later it has been extended to pairwise preferences in [26]. Pair-wise preferences are mostly selected in qualitative feedback-based GP applications as humans are more consistent with preference tasks than classification tasks. However, one can use pair-wise preferences and classifications for training the same GP model [28].

These GP-based Bayesian optimization approaches have been applied to HiL psychophysics studies [29, 30, 31, 32, 33] and HiL reward learning studies for personalization of robotic applications [14, 15, 28, 34]. Biyik [14] taught simple tasks to robots by using preference-based Bayesian active learning and an information theory-based sampling strategy. They demonstrated their methodology on reward learning-based HiL teaching. Tucker *et al.* [34] used a pair-wise preference-based Bayesian optimization to optimize the lower-body exoskeleton. Then, Tucker *et al.* [15] developed an efficient sampling methodology for high-dimensional gait personalization. Li *et al.* [28] used both ordinal classification and pair-wise preferences for HiL Bayesian optimization to tune the lower-body exoskeleton. They also provided a probabilistic safe sampling strategy and they use it to classify the safety of sampling space to avoid using unsafe parameters in HiL trials.

Fewer studies use preference-driven optimization for haptic rendering [35] and texture generation [36, 37]. Catkin *et al.* [35] proposed a preference-based HiL Bayesian optimization method for spring and friction rendering, demonstrating its effectiveness in capturing user preferences and personalizing parameters to enhance perceived realism. Their work showed that HiL optimization is an efficient approach to studying the impact of haptic parameters on perceived realism, even in multi-dimensional spaces.

2.3 GP-Based Multi-Criteria Bayesian Optimization

Bayesian optimization approaches have been extended to solve multi-criteria optimization problems [38, 39, 40, 41, 42, 43]. One means to address a multi-criteria optimization problem is to use a weighted sum of the cost functions to form a single aggregate cost function. Such scalarization approaches enable the original multi-objective problem to be formulated as a single criterion optimization problem. This approach has been applied to the Bayesian optimization setting by utilizing an aggregate cost function or a scalarized acquisition function for parameter sampling [39, 40, 44].

However, in all scalarization methods, since the relative weights of each cost function need to be pre-determined, the design preferences among the objective functions must be assigned a priori, before gaining sufficient knowledge of the trade-off involved.

On the other hand, Pareto optimization methods compute a set of non-dominated solutions that correspond to optimal designs for different design preferences among the optimizations metrics, and all such *non-dominated* solutions constitute the Pareto-front [45, 46, 47].

Unlike the single-shot scalarization-based optimization methods, Pareto methods fully characterize the trade-off among objectives. Once the Pareto-front solutions are computed, the designer can study these solutions to get an insight into the underlying trade-offs and make an informed decision to finalize the design by selecting *optimal* solutions from the Pareto set.

Pareto-optimization approaches have also been applied in the Bayesian optimization setting [38, 41, 42, 43]. Existing techniques can be loosely classified as hyper-volume improvement approaches [38, 42], information-theoretic methods [43], and wrapper methods via single-objective acquisition functions [41, 48].

In the hyper-volume improvement, the relative improvement amounts of all latent functions are estimated and multiplied with each other to find a hyper-volume. Then the sample point expected to produce the largest hyper-volume is selected. Hyper-volume improvement methodologies often use a multi-criteria extension of the expected improvement acquisition function to produce an effective sampling strategy [38, 42].

The information-theoretic methods derive a single acquisition function to maximize information gain for all objectives [43].

The wrapper methods utilize a multi-criteria optimizer to compute a surrogate

Pareto front characterizing the trade-off between the conflicting acquisition functions and sample parameters. Belakaria *et al.* used a sampling decision strategy among parameters in the Pareto front based on the highest volumetric variance [41, 48]. Recently, Astudillo *et al.* developed a novel multi-criteria framework for qualitative feedback-based Bayesian optimization [49].

2.4 Transfer Learning Approaches for GP-Based Methods

Applications in GP models in geophysics known as kriging [50] developed systematical extensions for sampling regions using correlated data. These extensions are generally known by cokriging models in which multiple kriging models are correlated with each other using variogram techniques.

Intrinsic coregionalization model (ICM) is a commonly used technique in cokriging [51]. ICM approach develops a latent function from the weighted sum of dependent latent functions in which each latent function requires the use of the same kernel function [52]. A generalized version of the ICM model, the linear coregionalization model (LCM), was developed to avoid restrictions for using the same kernel function.

Cokriging methodologies are adopted by GP regression-based machine learning and optimization techniques for modeling latent functions [53]. These multi-output GP models are commonly known as multi-tasking GP models [52, 54, 55]. Bonilla *et al.* used a kernel-based multi-tasking approach. Then Bonilla *et al.* developed a data-based correlating method between GP models using maximum likelihood estimation [52]. Golovin *et al.* suggested simplifying coregionalization models by only transferring pre-trained GP models to estimate a prior mean for the current GP model [56].

Due to the inclusion of large datasets, multi-tasking GP models are often used along with approximation methods to be a computationally feasible methods.

The Subset of regressor and projected process approximation is commonly used to divide a dataset into two parts and condition one part to another to transfer knowledge of the first dataset to the second one and using only the second one for computations [19]. These methodologies usually choose the division of datasets in an online manner [57]. Bayesian committee machine is another commonly used approximation methodology that divides datasets into many to train weaker GP models and combines predictions of each model with a Bayesian-based adaptive weighting method [58].

Alternative to multi-tasking GP approaches, more complex methodologies such as Semi-parametric approaches are developed for the Multi-tasking model [59, 60]. These approaches create an initial baseline based on either GP models or parametric functions.

Multi-tasking GP methodologies are used in Bayesian optimization to transfer offline data to an online GP model [61, 62]. This methodology is used in hyper-parameter tuning for complex machine learning models where measuring the performance of each hyper-parameter combination is excessively time-consuming [63]. Lübsen *et al.* developed a safety-focused sampling strategy for multi-tasking and transfer learning-based BO models [64].

Chapter 3

Preliminaries

This chapter includes the necessary information on the statistical background of GP-based Bayesian optimization methodologies. Based on the provided equations, one can build a simple GP model for both quantitative and qualitative metrics, tune the model automatically using the collected data and use the model for Bayesian optimization. More advanced knowledge on Bayesian optimization strategies are given in the latter chapters.

3.1 Probability Background

3.1.1 Bayes Rule

Let A and B be two random events, then conditional probability of A given that B is true denoted by $P(A|B)$

$$P(A|B) = \frac{P(A|B)P(A)}{P(B)} \tag{3.1}$$

3.1.2 Expected Value

Expected value is commonly used to find the mean and variance of a distribution. Let f be a continuous distribution with probability density function $P(f)$, then its mean is calculated by

$$E(f) = \int_{-\infty}^{\infty} fP(f)df \tag{3.2}$$

Variance of f denoted by $var(f)$ is calculated by $Var(f) = E(f^2) - E(f)^2$ where,

$$E(f^2) = \int_{-\infty}^{\infty} f^2P(f)df \tag{3.3}$$

3.1.3 Marginalization

Let $P(f, g)$ be the joint probability of two random variables, f and g, with continuous distributions. Then, the marginal probability of $P(g)$ is found by

$$P(g) = \int_{-\infty}^{\infty} P(g)P(f)df \tag{3.4}$$

The marginalization is also expressed as the expected value of $P(g)$ with respect to f.

$$E(P(g))_f = \int_{-\infty}^{\infty} P(g)P(f)df \tag{3.5}$$

3.1.4 Gaussian Distribution

Let f be a random variable with a Gaussian distribution. Then, f has a univariate Gaussian probability distribution function, $f \sim N(\mu, \sigma)$ with a probability distribution function as shown below.

$$P(f) = \frac{1}{\sigma\sqrt{2}} e^{-\frac{(f-\mu)^2}{2\sigma^2}} \tag{3.6}$$

3.1.5 Gaussian Process

Gaussian process (GP) is a type of random process used to model signals and functions via multi-variate Gaussian distribution.

Let $\boldsymbol{f}$ be a vector with n Gaussian random variable. Then, $\boldsymbol{f}$ has multivariate Gaussian distribution $\boldsymbol{f} \sim GP(\boldsymbol{\mu}, \Sigma)$

$$P(\boldsymbol{f}) = \frac{1}{(2\pi)^{\frac{n}{2}} \, |\Sigma|^{\frac{1}{2}}} \, e^{-\frac{1}{2}(\boldsymbol{f}-\boldsymbol{\mu})^T \, \Sigma^{-1} \, (\boldsymbol{f}-\mu)} \tag{3.7}$$

The covariance matrix Σ is shown as:

$$\Sigma = \begin{bmatrix} \sigma(f_1, f_1) & \cdots & \sigma(f_1, f_n) \\ \vdots & \ddots & \vdots \\ \sigma(f_n, f_1) & \cdots & \sigma(f_n, f_n) \end{bmatrix} \tag{3.8}$$

where the diagonal elements show the variances of each random variable and non-diagonal elements show the covariances.

3.2 Gaussian Processes Regression

Let $\boldsymbol{x} = \{x_1, x_2, .., x_n\}$ be a set consist of n input parameters and $D_y = \{y_1, y_2, .., y_n\}$ be a data set consist of n observed outputs. GP regression aims to create a non-parametric latent function $f(x)$ with a Gaussian process to find a relationship between the input parameter set $\boldsymbol{x}$ and output data set D_y. Furthermore, it aims to estimate an output $f(x_*)$ for any arbitrary input parameter x_*.

$f(x)$ has a prior distribution represented by a Gaussian process.

$$f(\boldsymbol{x}) \sim GP(\boldsymbol{\mu}, K) \tag{3.9}$$

In Eqn. (3.9), $\boldsymbol{\mu}$ is used to define an initial mean for the prior distribution, and K is used to define an initial covariance matrix. K is defined by a kernel matrix in which each element is created by a kernel function.

$$K = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \tag{3.10}$$

The radial basis function (RBF) kernel is one of the commonly used kernels in GP regression methods.

The RBF kernel is defined as:

$$k(x_i, x_j) = \sigma_{rbf}^2 e^{-\theta_{rbf}(x_i - x_j)^T (x_i - x_j)} \tag{3.11}$$

σ_{rbf} and θ_{rbf} are the hyper-parameters of RBF kernel in which σ_{rbf} is used to define magnitude of the covariance. θ_{rbf} is used to define the dependency between $f(x_i)$ and $f(x_j)$. One can increase the smoothness of the function by decreasing the value of θ_{rbf}. σ_{rbf} can be any real number except zero while θ_{rbf} must be a positive real number.

3.2.1 GP Regression for Noisy Observations

In many applications of GP regression, each output observation y_i in D_y is commonly assumed to be convoluted with Gaussian white noise $\epsilon_i \sim N(0, \sigma_w^2)$.

$$y_i = f(x_i) + \epsilon_i \tag{3.12}$$

The white noise assumption is used when the numerical observations are suspected to be noisy or reduce over-fitting and numerical calculation errors. While the white noise does not affect the mean of the initial distribution, it affects the covariance matrix. Let K_{noisy} represent the noise-added version of the kernel matrix. Then, K_{noisy} is found by: $\epsilon_i \sim N(0, \sigma_w^2)$.

$$K_{noisy} = K + \sigma_w^2 I \tag{3.13}$$

3.2.2 Bayesian Inference for GP Regression

The prior distributions of $f(x)$ and $f(x_*)$ are given as:

$$\begin{bmatrix} f(x) \\ f(x_*) \end{bmatrix} \sim Gp\left(\begin{bmatrix} \mu \\ \mu_* \end{bmatrix}, \begin{bmatrix} K & k_*^T \\ k_* & k_{**} \end{bmatrix}\right) \tag{3.14}$$

In Eqn. (3.14), $k_* = k(x_*, x_{1:n})$ is a vector of kernel function values between parameter x_* and parameters in $\boldsymbol{x}$ and k_{**} denotes $k(x_*, x_*)$. μ and μ_* denotes prior mean values.

By using the numerical observations D_y, posterior distribution of $f(x_*)$ is found by using a Bayesian approach. Accordingly, the posterior distribution of $f(x_*)$ has a Gaussian distribution and denoted as $f(x_*)_{|D_y}$.

$$\begin{aligned} f(x_*)_{|D_y} &\sim GP\left(\mu_{*|\boldsymbol{y}}, \sigma^2_{*|\boldsymbol{y}}\right) \\ \mu_{*|D_y} &= \mu_* + k_*(K + \sigma_w^2 I)^{-1}(D_y - \mu) \\ \sigma^2_{*|D_y} &= k_{**} - k_*(K + \sigma_w^2 I)^{-1} k_*^T \end{aligned} \tag{3.15}$$

In Eqn. (3.15), If the observations, D_y are not noisy, then $\sigma_w^2 I$ becomes zero.

3.3 Qualitative Feedback-based Gaussian Process

Let $\boldsymbol{x} = \{x_1, x_2, .., x_n\}$ be a set consisting of n input parameters and $D_q = \{q_1, q_2, .., q_n\}$ be the data set consist of users preferences and classifications based on a qualitative metric. The qualitative feedback-based GP model aims to find a quantified relationship between input parameters $\boldsymbol{x}$ and given qualitative data D_q.

We first model an initial latent function $f(\boldsymbol{x})$ a standardized Gaussian process model.

$$f(\boldsymbol{x}) \sim GP(\mu, K) \tag{3.16}$$

Then, we model a probability function to relate latent function $f(x)$ with the user's preferences and classifications. By using the user's qualitative feedback, we estimate a posterior distribution for the latent function. Let $P(D|f)$ be the probability of total qualitative feedback data for a given latent function.

$$P(f(\boldsymbol{x})|D_q) \propto P(D_q|f(\boldsymbol{x}))P(f(\boldsymbol{x})). \tag{3.17}$$

3.3.1 Modeling User's Qualitative Feedback

Let q_i be the i^{th} qualitative feedback of a user and $P(q_i|f)$ be the probability of occurrence of q_i based on the latent function. Assuming that each feedback is independent from each other, then $P(D_q|f)$ is found by

$$P(D_q|f(\boldsymbol{x})) = \prod_{i=1}^{n} P(q_i|f(x)) \tag{3.18}$$

For modeling probabilities for qualitative feedback, we preferred to use a normal distribution cumulative density function (CDF) Φ. One can prefer to use a sigmoid function instead of a normal CDF.

3.3.1.1 Probability Function for Binary Classifications

Let $B = \{b_1 = -1, b_2 = 1\}$ represent the binary classification corresponding to any type of "*yes/no*" decision. For example, if the acquired answer is "*yes*," the parameter x_i is classified as b_2; otherwise, it is classified as b_1. Let ϵ_b be Gaussian white noise with variance c_b^2, which affects the user's decision. The parameter x_i is then classified as b_j, as shown below.

$$P(q_{bi} = b_j | f(x_i)) = \Phi\left(\frac{b_j f(x_i))}{c_b}\right) \tag{3.19}$$

3.3.1.2 Probability Function for Ordinal Classifications

Let $O = \{o_1, o_2 ..., o_r\}$ be the finite set of r ordinal classifications and t_{o_j} be the threshold of each ordinal classification $-\infty = t_{o_0} < t_{o_1} < t_{o_2} < \cdots < t_{o_r} = \infty$. Let ϵ_o be a Gaussian white noise with variance c_o^2, affecting the user's decision while classifying. Then, the probability for parameter x_i belonging to the o_j^{th} ordinal class is defined as

$$P(q_{oi} = o_j | f(x_i)) = \Phi\left(\frac{t_{o_j} - f(x_i))}{c_o}\right) - \Phi\left(\frac{t_{oj-1} - f(x_i)}{c_o}\right) \tag{3.20}$$

where c_o is the noise level constant. To reduce this model to binary classification one can select a single threshold $t_1 = 0$

3.3.1.3 Probability Function for Pairwise Preferences

Let ϵ_p be a Gaussian white noise with variance c_p^2. Then, the probability of participants preferring x_{i1} over x_{i2} for a given latent function f is defined as

$$P(q_{pi} = (x_{i1}) \succ (x_{i2}) | f(x)) = \Phi\left(\frac{f(x_{i1}) - f(x_{i2})}{c_p}\right) \tag{3.21}$$

3.3.2 Laplace Approximation for Posterior Distribution

The aim of finding the posterior distribution of $P(f(x)|D_q)$ is to make estimations for the latent function distribution of any arbitrary parameter x_* in our search space. However, the exact distribution of Posterior, $P(f(x)|D_q)$ is not

Gaussian. Hence, the posterior distribution of a latent function of $f(x_*|D_q)$ is not analytically tractable.

To estimate $f(x_*|D_q)$ analytically there are multiple approximation method exists. Commonly used methodologies are expectation propagation EP and the Laplace approximation. We chose to use the latter one.

The Laplace method is used to approximate the posterior distribution of $P(f(\boldsymbol{x})|D_q)$ as a multi-variate Gaussian distribution. The Laplace approximation uses second-order Taylor expansion of posterior distribution around its mode [19]

$$f(\boldsymbol{x})_{|D_q} \sim GP(\hat{f}, (W + K^{-1})^{-1}) \tag{3.22}$$

$$\hat{f} = \operatorname{argmax}_{f(\boldsymbol{x})} \left(log(P(D_q|f(\boldsymbol{x}))P(f(\boldsymbol{x})))\right) \tag{3.23}$$

where W is the negative Hessian of $log(P(D_q|f(\boldsymbol{x})))$ defined as

$$W_{ij} = -\frac{\partial^2 \, log(P(D_q|f(\boldsymbol{x}))}{\partial f(x_i) \, \partial f(x_j)}.$$

3.3.3 Bayesian Inference for Qualitative Feedback based GP

From the posterior distribution of $f(\boldsymbol{x})|q = f_{|D_q}$, the output of the latent function $f(x_*) = f_*$ for any arbitrary parameter set $x_* \subset \boldsymbol{X}$ can be estimated. Let $f_{*|D_q}$ be the estimated output of the latent function for an arbitrary parameter set x_* based on information of q. Then, $f_{*|D_q}$ is another Gaussian distribution that can be found by marginalization.

$$P(f(x_*)|D_q) = \int_{-\infty}^{\infty} (P(f(x_*)|f(\boldsymbol{x}), q) P(f(x)|D_q)) \, df(\boldsymbol{x}) \tag{3.24}$$

The result of the integral in 3.24 is given as

$$f_{*|D_q} \sim GP(f; E(f_{*|D_q}), Var(f_{*|D_q})) \tag{3.25}$$

with

$$\begin{aligned} E(f_{*|D_q}) &= k_*\, K^{-1}\hat{f} \\ Var(f_{*|D_q}) &= k_{**} - k_*\,(K + W^{-1})^{-1} k_*^T. \end{aligned}$$

where $k_* = k(x_*, x_{1:n})$ and $k_{**} = k(x_*, x_*)$.

3.3.4 Estimating Qualitative Feedback

The Eqn. (3.25) can be further extended to develop a probability function to predict outcomes of user qualitative feedback q_* for an input parameter x_*.

$$P(q_*|D_q) = \int_{-\infty}^{\infty} P(q_*|f_*)P(f_*|D_q)df_* \tag{3.26}$$

3.3.4.1 Estimating Classifications

The posterior probability of classifying an arbitrary parameter x_{1*} as o_j^{th} class, denoted with the probability $P(q_{o*} = o_j|D_q)$

$$P(q_{o*} = o_j|D_q) = \Phi\left(\frac{t_{oj} - E(f_{*|D_q})}{\sqrt{Var(f_{*|D_q}) + c_o^2}}\right) - \Phi\left(\frac{t_{oj-1} - E(f_{*|D_q})}{\sqrt{Var(f_{*|D_q}) + c_o^2}}\right). \tag{3.27}$$

For the binary case, when there are two ranking $j = 1, 2$, thresholds of ordinal classification become $O = \{o_0 = -\infty, o_1 = 0, o_2 = \infty\}$. When the user provides classification as positive, $j = 2$, Eqn. (3.27) reduces to

$$P(q_{o*} = o_j|D_q) = 1 - \Phi\left(\frac{-E(f_{*|D_q})}{\sqrt{Var(f_{*|D_q}) + c_o^2}}\right) = \Phi\left(\frac{E(f_{*|D_q})}{\sqrt{Var(f_{*|D_q}) + c_o^2}}\right). \tag{3.28}$$

When the user provides classification as negative, $j = 1$, Eqn. (3.27) reduces to

$$P(q_{o*} = o_j | D_q) = \Phi\left(\frac{-E(f_{*|D_q})}{\sqrt{Var(f_{*|D_q}) + c_o^2}}\right). \tag{3.29}$$

3.3.4.2 Estimating Pairwise Preferences

The posterior probability of preferring an arbitrary parameter x_{1*} over the reference x_{2*}, denoted with probability $P(x_{1*} \succ x_{2*}|D_q)$.

Let $\Delta*$ be the difference between latent scores $f(x_{1*}) - f(x_{2*})$. For known a known difference Δ_*, the preference probability of the user is modeled as

$$P((x_{1*} \succ x_{2*}|\Delta_*) = \Phi(\Delta_*/c_p) \tag{3.30}$$

The posterior distribution of latent score difference, Δ_* has a Gaussian distribution with mean $E(\Delta_{*|D_q})$ and variance $Var(\Delta_{*|D_q})$.

$$E(\Delta_{*|D_q}) = E(f_{x_{1*}|D_q}) - E(f_{x_{2*}|D_q}) \tag{3.31}$$

$$Var(\Delta_{*|D_q}) = Var(f_{x_{1*}|D_q}) + Var(f_{x_{2*}|D_q}) - 2Cov(f_{x_{1*}|D_q}, f_{x_{2*}|D_q}) \tag{3.32}$$

The mean and variance of $\Delta f_{|D_q}$ is found using Eqn. (3.25). Then, $P(x_* \succ x_{ref}|D_q)$ is estimated with

$$P(x_{1*} \succ x_{2*}|D) = \int_{-\infty}^{\infty} \Phi(\Delta_*/c_p) P(\Delta_*|D_q) d\Delta_* \tag{3.33}$$

and the result of this integration is denoted as

$$P(x_{1*} \succ x_{2*}|D_q) = \Phi\left(\frac{E(\Delta_{*|D_q})}{\sqrt{Var(\Delta_{*|D_q}) + c_p^2}}\right). \tag{3.34}$$

3.4 Hyper-Parameter Tuning

Both GP regression and qualitative feedback-based models include hyper-parameters for modeling a latent function. These hyper-parameters can be defined by heuristic or empirical approaches or tuned systematically using acquired data D which consist of either numeric observations D_y or qualitative feedback D_q. In this section, we discuss the implementation of Maximum Likelihood Estimation (MLE) to statistically tune GP model hyper-parameters.

The goal of the MLE approach is to find a hyper-parameter set, $\boldsymbol{\theta_{hyp}}$ that maximizes log-likelihood $log\,(P(D|\boldsymbol{\theta_{hyp}}))$. As the evaluation of $log\,(P(D|\boldsymbol{\theta_{hyp}}))$ is generally costly, gradient-based search strategies are commonly used to find the maximum value of $log\,(P(D|\boldsymbol{\theta_{hyp}}))$.

3.4.1 Hyper-Parameter Tuning for GP Regression

Let the prior GP model be defined as $GP(\mu_y, K_y)$ where μ_y is the prior mean and K_y is the prior covariance matrix where K_y can be noisy. Let quantitative data D_y represented by a vector $\boldsymbol{y}$ with a size of n.

$$log\,(P(D_y|\boldsymbol{\theta_{hyp}})) = -\frac{1}{2}(\mu_0 - \boldsymbol{y})^T(K_y)^{-1}(\mu_0 - \boldsymbol{y}) - \frac{1}{2}log(|K_y|) - \frac{n}{2}log2\pi \quad (3.35)$$

Let θ_j be the j^{th} hyper-parameter that we want to find by MLE and let's assume that no hyper-parameters are used in the prior mean. Then, the gradient of $log\,(P(D_y|\boldsymbol{\theta_{hyp}}))$ is calculated as below.

$$\frac{\partial}{\partial\theta_j}log\,(P(D_y|\boldsymbol{\theta_{hyp}})) = \frac{1}{2}(\mu_0 - \boldsymbol{y})^T K_y^{-1}\frac{\partial K_y}{\partial\theta_j}(\mu_0 - \boldsymbol{y}) - \frac{1}{2}tr(K_y^{-}1\frac{\partial K_y}{\partial\theta_j}) \quad (3.36)$$

where tr denotes the trace of the matrix. This method is commonly used to tune kernel hyper-parameters such as σ_{rbf} and θ_{rbf} shown in Eqn. (3.11) as well as the standard deviation of white noise convolution σ_w.

3.5 Hyper-Parameter Tuning for Qualitative Feedback GP

Similar to the hyper-parameter tuning approach used in Section 3.4.1, MLE is used to tune hyper-parameters based on the qualitative feedback data.

Let prior GP model defined as $GP(0, K_q)$ and posterior GP model derived according to Eqn. (3.17) defined as $GP(\hat{f}, (K_q + W^{-1})^{-1})$ where K_q is the prior covariance matrix without noise addition. Let qualitative feedback data D_q and $\hat{f}$ have a size of n. Then, the log-likelihood is calculated as shown below.

$$log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right) = -\frac{1}{2}(\hat{f})^T(K_y)^{-1}(\hat{f}) + log\left(P(D_q|\hat{f})\right) - \frac{1}{2}log(|B|) - \frac{n}{2}log2\pi \tag{3.37}$$

where $B = I + WK$ and W is the negative Hessian, $W = -\nabla\nabla log(P(\boldsymbol{q}|\hat{f}))$.

Let θ_j be the j^{th} hyper-parameter that we want to find by MLE and let's assume that no hyper-parameters are used in the prior mean. Then, the gradient of $log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right)$ is calculated as below.

$$\frac{\partial log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right)}{\delta\theta_j} = \frac{\partial log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right)}{\delta\theta_j}\Big|_{explicit} + \sum_{i=1}^{n} \frac{\partial log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right)}{\delta\hat{f}_i}\frac{\partial\hat{f}_i}{\partial\theta_j} \tag{3.38}$$

The terms used in Eqn. (3.38) expanded as

$$\frac{\partial log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right)}{\delta\theta_j}|_{explicit} = \frac{1}{2}\hat{f}K^{-1}\frac{\partial K}{\partial\theta_j}K^{-1}\hat{f} - \frac{1}{2}tr\left((K^{-1}+W)^{-1})\frac{\partial K}{\partial\theta_j}\right)$$

$$\frac{\partial log\left(P(D_q|\boldsymbol{\theta_{hyp}})\right)}{\delta\hat{f}_i} = -\frac{1}{2}[(K^{-1}+W)^{-1}]_{ii}\frac{\partial^3}{\partial\hat{f}_i^3}log(P(D_q|\hat{f}))$$

$$\frac{\partial\hat{f}_i}{\partial\theta_j} = (I+KW)^{-1}\frac{\partial K}{\partial\theta_j}\nabla log(P(D_q|\hat{f}))$$

3.6 Bayesian Optimization

The Bayesian optimization uses the GP model of the latent function as defined in the Sections 3.2 and 3.3.

Algorithm 1 Pseudo-code for Simple Bayesian Optimization

```
initiate $S$: Parameter space, $f \sim GP(\mu_0, \sigma_0)$: GP prior, $M$: Space-filling iteration number , $N$: Total iteration number
1: for $i = 1, 2, \ldots, N$ do
2:     if $i \leq M$ then
3:         Select $x_i$ from $S$ according to space-filling strategy
4:     else
5:         Select $x_i$ where $x_i$ maximizes acquisition function
6:     Observe outcome data $d_i$ and append it to $D$
7:     Update posterior distribution of $f$ using all observed data, $D$
8: Return a solution, $x_{max}$ where $f(x_{max})$ has the largest posterior mean
```

Algorithm 2 presents the simple implementation of Bayesian optimization and its steps at each iteration: The first M iterations are conducted to explore the search space via space-filling methods (Lines 1–4); Then, for the next $N - M$ iterations, the algorithm suggests a parameter using the acquisition function (Line 5). Next, the suggested parameter is used in the HiL experiment trials and the corresponding outcome data is observed (Line 6). Then, the algorithm updates its GP posterior using all available data (Line 7). When the trials end, the algorithm suggests a globally optimized solution (Line 8).

3.6.1 Space-Filling Methods

Space-filling methods are applied in Bayesian optimization to have an initial guess about the search space. Commonly used approaches are random sampling, Sobol sampling and Latin hyper-cube.

3.6.2 Acquisition Function

The acquisition function is a heuristic approach used to sample a parameter more efficiently. Two important concepts of acquisition function are exploration and exploitation. Exploration allows the acquisition function to sample from different regions of search space while exploitation makes the acquisition function focus sampling in an optimal area. Hence, a trade-off between exploration and exploitation exists while designing an acquisition.

Commonly used acquisition functions are listed below.

3.6.2.1 Thompson Sampling (TS)

TS is a commonly used method for multi-armed bandit problems. TS randomly sample from the posterior random process distribution and selects a parameter corresponding to largest value in sampled process.

3.6.2.2 Upper Confidence Bound (UCB)

UCB chooses the largest parameter with an upper confidence interval threshold calculated using the latent function's posterior distribution.

$$\alpha_{UCB}(x_*) = \mu(x_*) + c\,\sigma(x_*) \tag{3.39}$$

where $\mu(x_*)$ denotes the mean of $f(x_*)$, $\sigma(x_*)$, is the standard deviation of the posterior probability distribution of $f(x_*)$, and c is a constant hyper-parameter used to define the exploration/exploitation ratio. As c increases (decreases) the weight of the standard deviation increases (decreases) thus, the sampling algorithm focuses more on exploration (exploitation).

3.6.2.3 Expected Improvement (EI)

EI acquisition function uses a probabilistic approach to find the expectancy of possible improvement from the best-sampled parameter.

$$\alpha_{EI}(x) = (E(f(x))_{max} - \mu(x_*))\Phi\left(\frac{E(f(x))_{max} - \mu(x*)}{\sigma(x)}\right) + \sigma(x)N\left(\frac{E(f(x))_{max} - \mu(x)}{\sigma(x)}\right) \tag{3.40}$$

where one can select $E(f(x))_{max}$ as the maximum expectation value from the posterior distribution of observed parameters or use y_{max} if the observations are qualitative. If all values of $\alpha_{EI}(x)$ are negative, the GP model is assumed to be converged.

Chapter 4

HiL Optimization of Perceived Realism of Multi-Modal Haptic Rendering under Conflicting Sensory Cues

In this chapter, we introduce a novel application of Bayesian optimization in visual-haptic studies for learning and increasing perceived realism.

4.1 Introduction

Humans receive real-world sensory cues through various feedback channels, combining complementary (non-redundant) signals and integrating redundant signals to form a coherent multi-sensory percept. This process results in a robust perceptual model of the world [65]. Bayesian models are commonly used to understand how sensory signals are (partially) integrated, offering insights into the neural mechanisms involved in perceptual decision-making [66]. Strong evidence in the literature suggests that neural mechanisms across a population of neurons implement *sensory integration* similar to maximum likelihood estimation. (MLE) [65]. During sensory integration, a percept is obtained through a weighted linear combination of redundant sensory signals. When these weights correspond to the (perceived) reliability of the signals, the most reliable percept is achieved. [65, 67]; otherwise, in terms of MLE, the result is considered sub-optimal. Furthermore, as the reliability of a signal decreases, its weight in the sensory integration process is also reduced, resulting in a shift in the dominant sensory modality, with the most reliable signal making the largest contribution to the percept. [68, 69, 70]. Additionally, when the discrepancy among sensory signals becomes significant, the reliability of the multi-sensory estimate may drop below that of a single-sensory (uni-modal) estimate. In such cases, the discrepant source may be excluded, or *vetoed*, rather than integrated into the perceptual model [65, 69, 71, 72].

Since independent sensory cues are required for optimal MLE-like integration behavior, its applicability for modeling conflicting cues is limited. In particular, while the MLE model is likely to stay valid for small and hard-to-detect discrepancies along the dimension of interest, this model of sensory integration fails to account for the breaking down of cross-modal interactions when information provided by each modality is highly conflicting [73, 74].

Several extensions of the MLE model exist in the literature that can account for the partial integration of cues across a wide range of inter-modal discrepancies and stimulus conditions [75, 76]. For instance, [75] utilizes prior knowledge about the correspondence between multi-modal cues when determining the degree of integration, while the causal inference model [76] considers possible causes of the underlying sensory events to enable partial integration.

Haptic perception mainly depends on the fusion of *force-related* cues and *movement-related* cues, which are complementary. Moreover, visual feedback commonly accompanies the haptic perception of movement, providing *additional* movement-related cues that are known to dominate other modalities under many circumstances [77, 78]. During the integration of visual and haptic sensory inputs, the

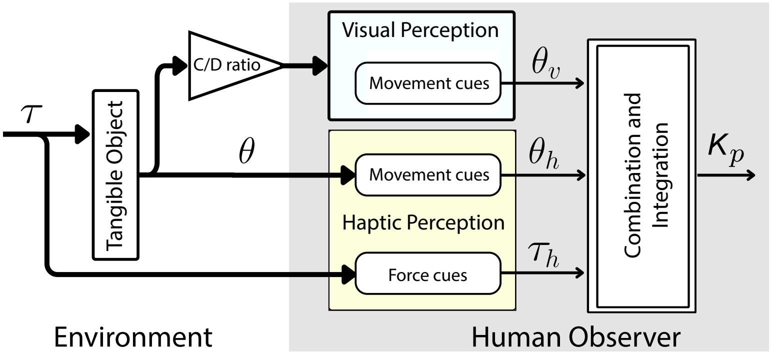

FIGURE 4.1: A schematic representation of visual-haptic sensory integration and combination to form a coherent compliance percept. In the figure, τ denotes the torque resulting from the rotation θ of the object. Symbol θ_h represents the haptic perception of movement cues, while τ_h denotes the haptic perception of force-related cues. The visual cues due to the scaled movement displayed on the monitor are denoted by θ_v, while the scaling factor is captured by the C/D ratio. Finally, K_p denotes the perceived stiffness formed by (partial) integration and combination of multi-modal cues.

percept has been shown to depend on each cue [65, 69], while the integration process may be suboptimal from a maximum reliability perspective [79, 80].

Haptic rendering is the bilateral process that makes the dynamics of computationally mediated virtual environments (VEs) apparent to a human user through a haptic interface. Haptic interfaces are commonly complemented with a visual display to improve immersion levels. Haptic rendering aims to maximize the perceived realism of the VEs by ensuring the perceived similarity of the rendering with respect to the reference model; however, high-fidelity rendering is restricted to a limited range of VEs due to underlying hardware limitations, such as the maximum force output or the minimum force resolution.

During a typical haptic rendering task, sensory information provided by the haptic interface and the visual display include both complementary and redundant cues. Movement-related percepts are generally formed through the integration of cues provided by both haptic and visual modalities [79], as depicted in Figure 4.1. Furthermore, it is possible to induce a controlled discrepancy between haptic and visual movement cues by introducing a scaling factor, called control-display (C/D) ratio, to the visual feedback provided during haptic rendering.

Haptic perception of a virtually mediated object under visual-haptic congruency has been of considerable interest, since it has been demonstrated that visual cues can override haptic feedback in displacement-related tasks, indicating the potential of visual manipulation to improve haptic rendering experiences [77]. Such approaches are commonly utilized for *pseudo-haptics* [81]. In the literature, it has been demonstrated that perceived compliance/stiffness can be increased by scaling the visual displacements [77, 82, 83, 84, 85, 86]. Similar approaches have also been applied to other haptic rendering tasks, such as rendering weight [78] and surface roughness [87, 88].

Overall, it is possible to capitalize on the multi-modal and redundant nature of the movement-related cues to enhance the *perceived* rendering range of any haptic interface by manipulating the visual cues. However, it remains an open challenge to determine the proper level of visual scaling to achieve the most realistic haptic rendering for a user. This goal necessitates an understanding of the psychometric model of multi-modal haptic rendering under conflicting sensory cues.

In this study, we explore the visual-haptic cue integration and propose a systematic means to determine the optimal visual scaling during haptic manipulation that maximizes the perceived realism of multi-modal spring rendering. Our approach is rooted in sample-efficient human-in-the-loop (HiL) optimization, where the rendering parameters are iteratively updated based on participants' qualitative feedback.

We extend our results to a group of individuals to capture the underlying multi-dimensional psychometric field that characterizes the cumulative effects of feedback modalities utilized during sensory integration under conflicting cues. Our results not only provide reliable estimates of just noticeable difference (JND) thresholds for stiffness under visual scaling but also capture several prominent features of sensory integration.

Overall, we demonstrate that the HiL optimization approach allows for the determination of appropriate haptic-visual parameters. This ensures a consistently high rating of perceived realism of multi-modal rendering. Additionally, it provides a systematic and efficient means to study multi-modal sensory integration under conflicting cues.

4.2 Related Work

The HiL setting where participants provide feedback in each trial is commonly employed by standard methods of classical psychophysics, such as the method of constant stimuli and the method of limits. However, these studies are not well-suited for evaluating stimuli with more than one dimension because the number of trials grows exponentially with the number of dimensions and the number of points per dimension [89]. Accordingly, several adaptive techniques, predominantly based on Bayesian methods, have been developed to achieve similar accuracy with classical psychophysics methods while using fewer trials [90].

Bayesian-based adaptive learning and optimization approaches have been used for psychophysics studies [29, 30, 91, 92, 93, 94, 95, 96, 97]. These methodologies can be loosely categorized as parametric and non-parametric approaches.

4.2.1 Parametric Approaches

The well-known adaptive methods, such as QUEST [91], and Psi [92], rely on the assumption that a parametric model for the psychometric function consistent with Weber's law exists. These parametric approaches assume that the stimulus varies on only one dimension and utilize a Bayesian update strategy to achieve sample efficiency. However, the extension of these methods to multi-dimensional stimuli is limited, as they evaluate any additional dimensions independently resulting in an inefficient search strategy over one-dimensional slices of the multi-dimensional psychometric field.

QUEST+ [94] and Psi-marginal [93] are more general parametric approaches that support multi-dimensional models but require the parametric form of the psychometric field to be specified *a priori*. However, the extension of one-dimensional psychometric curves to multi-dimensional psychometric fields is not straightforward, as there may exist nonlinear interactions between the additional variables and saturation behavior at low and high intensities [32].

Furthermore, since all parametric methods strongly depend on the model introduced before data collection, the conclusions that can be drawn from these approaches become limited if data violates the underlying assumptions of the model.

4.2.2 Non-parametric Approaches

Non-parametric approaches have been introduced to remove strong assumptions about the shape of the multi-dimensional psychometric fields, by modeling the psychometric function using a stochastic process. In most cases, a Gaussian process (GP) serves as a sample-efficient non-parametric model for complex functions [29, 30]. Furthermore, non-parametric approaches replace dense sampling with efficient active learning schemes, significantly improving the applicability of these methods for psychometric studies with multivariate stimulus settings [29, 30, 95, 96, 97].

The classical implementation of GP-based Bayesian optimization has been developed for quantitative metrics and is not directly applicable to qualitative evaluations. On the other hand, psychophysical studies are dominantly based on “yes/no” type classification tasks and most HiL optimization studies rely on qualitative comparative feedback from users to enhance the reliability of subjective metrics [10].

In this study, we utilize a non-parametric Bayesian optimization approach, as this selection allows us to model the multi-dimensional psychometric field between the rendering parameters and perceived realism without assuming a fixed parametric form. Our work is built upon the existing GP-based HiL optimization approach presented in [34, 35], since GP-based models are multi-dimensional by default and are flexible enough to model correct saturation behaviors.

While previous non-parametric psychophysical studies used either binary classifications [30, 32] or pairwise preferences [29, 33] as the qualitative feedback from the users, in this study, we extend this methodology to use a wider range of classifications to get more information from the user in each trial, such that we can build a more general and sample-efficient method. Our study significantly extends and refines these approaches to suit multi-modal haptic rendering under visual-haptic incongruency and provides insight into multi-modal sensory integration by capturing the underlying psychometric field and perception thresholds. The characterized psychometric field captures the cumulative effect of feedback modalities on perceived realism during sensory integration of conflicting cues.

4.3 Qualitative Feedback-Based HiL Optimization

The HiL optimization aims to learn a quantified relationship between the rendering parameters and user perception through a GP-based latent function. The GP-based latent function models users' perceptions without making strong parametric assumptions.

As users provide more qualitative feedback, such as classifications and comparisons, the GP-based latent function is updated according to Bayes' theorem, to capture their perception more accurately. During the HiL experiments, the GP-based latent function is used within a Bayesian optimization framework to efficiently learn the relationship by employing informed sampling techniques.

After posterior GP models of all users are trained, a generalizable latent perception model can be constructed by statistically averaging these posterior GP models, and a probabilistic relationship between the rendering parameters and corresponding qualitative feedback outcomes can be derived from the averaged posterior GP-based latent function to extract psychophysical thresholds of interest.

While a detailed mathematical model of the HiL optimization method based on qualitative feedback is presented in the Appendix, an overview can be provided as follows:

4.3.1 Latent Function Modeling

During the HiL optimization, subjects form a perception of the stiffness K_{p1} and K_{p2} from two different visual-haptic renderings and $K_{p_{ref}}$ from the reference rendering. Users are asked to evaluate the stiffness of renderings based on their perceived similarity with respect to the reference. Our experiments collect two ordinal classifications, q_{o_1} and q_{o_2}, and one pairwise preference, q_p, at each iteration. Users provide ordinal classification based on the similarity between a rendering and the reference, and they give a pairwise comparison between two renderings based on which one is more similar to the reference. Although pairwise comparisons provide no extra information when subjects classify parameters into separate categories, pairwise preferences are useful to capture small differences if two parameters are classified in the same category.

We train a GP-based latent function $f(x)$ to learn the relationship between rendering parameters $x = (K, \mathrm{C/D\ ratio})$, and users' perceived similarity by using the collected qualitative feedback data D. The latent function is modeled such that higher perceived similarity results in higher latent function scores. The prior probability of the latent function $P(f(x))$ is modeled using a normalized GP model.

We know that human decisions are not perfectly consistent; hence, this inconsistency is modeled by white noise interference to the decision process, as depicted

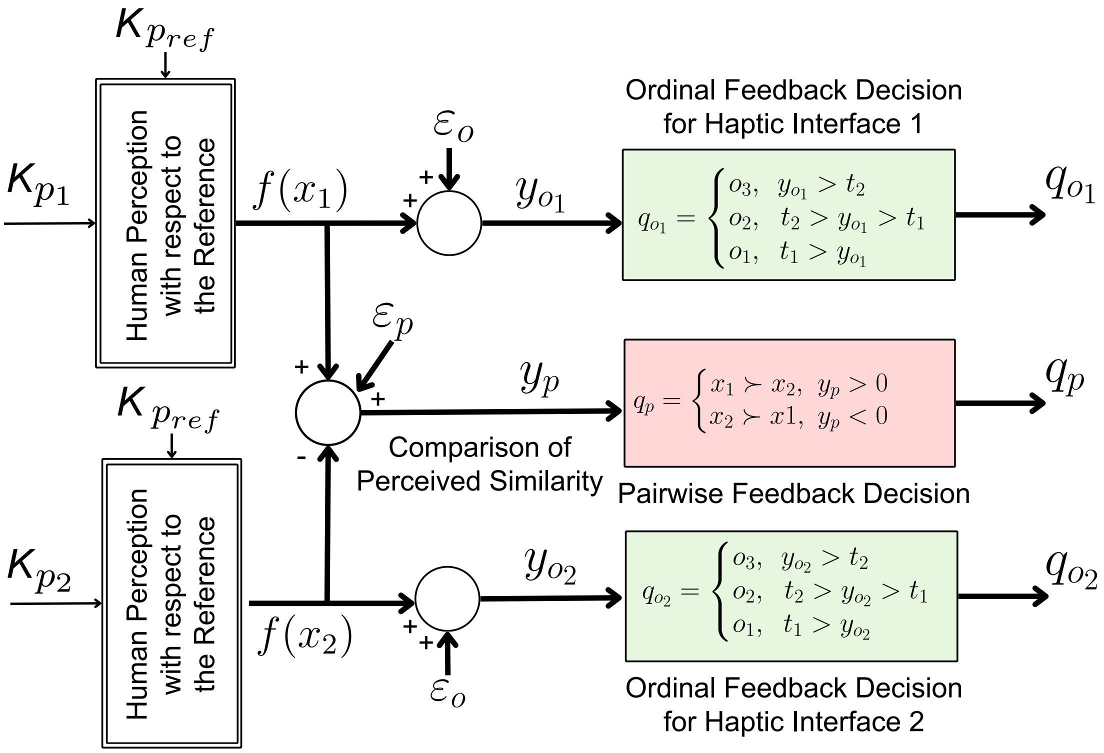

FIGURE 4.2: Human decision model for qualitative feedback

with noises ϵ_o and ϵ_p in Figure 4.2. The white noise aims to capture the limitations of human sensitivity and any random distractions that may take place during the experiment. Given that humans are typically more consistent while making pairwise preferences compared to ordinal classifications [28], distinct white noise parameters are used for pairwise preferences and ordinal classifications.

The users' ordinal classification decisions are modeled according to the threshold values between where the noise-interfered version of latent scores y_{o1} and y_{o2} fall. The binary classification used in [32] is a special case of ordinal classification with two categories separated by a threshold at zero. We preferred to apply three ordinal categories to enforce more reward to renderings perceived similar to reference and more penalty to renderings with high perceived distinctiveness. Similarly, the pairwise preference is modeled according to the noise-interfered version of the difference between latent function scores. When the noise-added difference is above zero, the user perceives the stiffness of the first rendering K_{p1} to be more similar to reference rendering $K_{p_{ref}}$.

4.3.2 Posterior Latent Function and Bayesian Inference

As the user provides more qualitative feedback, the latent function is updated according to Bayes' theorem

$$P(f(x)|D) \propto P(D|f(x))P(f(x)) \tag{4.1}$$

allowing it to capture the user's perception more precisely.

The aim of acquiring the posterior distribution of $f(x)$ is to enable predictions of the latent function scores $f(x_*)$ for any feasible arbitrary rendering parameter x_*. However, since the probability of qualitative feedback is not Gaussian, the posterior distribution of $f(x_*)$ is not analytically tractable using the posterior distribution of $f(x)$. Accordingly, the Laplace method is used to approximate the posterior distribution of $f(x)$ as a Gaussian distribution such that the posterior distribution $f(x_*)$ is also inferred as a Gaussian. Utilizing this commonly adopted method for posterior approximation, it is possible to make predictions for Bayesian optimization and extend the probabilistic derivations to capture psychophysical estimations based on human perception [14, 19, 28, 34].

4.3.3 Bayesian Optimization

The Bayesian optimization uses the GP model of the latent function as defined in the previous subsection. In each iteration, a query consisting of a two-parameter set is sampled by using the Upper Confidence Bound (UCB) acquisition function

$$\alpha_{UCB}(x_*) = \mu(x_*) + c\,\sigma(x_*) \tag{4.2}$$

Where $\mu(x_*)$ denotes the mean of $f(x_*)$, $\sigma(x_*)$, is the standard deviation of the posterior probability distribution of $f(x_*)$, and c is a constant hyper-parameter used to define the exploration/exploitation ratio. As c increases (decreases) the

Algorithm 2 HiL Optimization with Qualitative Feedback

```
initiate $S$: Parameter space, $f \sim GP(\mu_0, \sigma_0)$: GP prior, $M$: Space-filling iteration number , $N$: Total iteration number
1: for $i = 1, 2, \ldots, N$ do
2:     if $i \leq M$ then
3:         Randomly select two distinct points from $S$
4:     else
5:         Select two parameters
           $x_{i_{1,2}} := \mathrm{argmax}_{\mathrm{x} \in \mathrm{S}} \left(\alpha_{UCB}(x)_{i-1}\right)$
6:     Observe qualitative feedback ($q_{o_{i_1}}$, $q_{o_{i_2}}$ and $q_{p_i}$)
7:     Update posterior distribution of latent function model $f$
```

weight of the standard deviation increases (decreases) thus, the sampling algorithm focuses more on exploration (exploitation). To sample from different regions, we also impose a condition during the selection of the second parameter set, such that the prior covariance should be smaller than 0.2.

Algorithm 2 presents the basic steps at each iteration: The first M iterations are conducted to explore the search space via space-filling methods (Lines 1–4); i.e., random uniform sampling. Then, for the next $N - M$ iterations, the algorithm suggests two parameter sets using the most promising points according to the acquisition function (Line 5). Next, the suggested parameter sets are used in the visual-haptic rendering with the left and right knobs. After the subject tries the rendered parameters, qualitative feedback regarding the trial is transmitted to the algorithm (Line 6). Lastly, the algorithm updates its GP posterior according to the qualitative feedback data (Line 7). We use the GP update procedure for the parameter inference as in [14].

4.3.4 Aggregating Gaussian Process and Extracting Psychophysical Thresholds

After the data collection, the trained posterior models are used to create an averaged posterior GP model to form a general perception model. Then, this model is used to infer the classification and comparison decisions of users throughout the bimodal stimuli space.

4.3.4.1 Averaged Gaussian Process Posterior

The averaged posterior GP model is computed from the individual posterior GPs of the multi-modal experiment. We treat each participant's posterior GP results as independent and identically distributed measurements over the bi-modal space and use them to obtain the average perceived similarity posterior model as follows

$$\bar{f}_{*|D_{tot}} \sim GP\left(E(\bar{f}_{*|D_{tot}}), Var(\bar{f}_{*|D_{tot}})\right) \tag{4.3}$$

$$\begin{aligned} \text{with} \quad E(\bar{f}_{*|D_{tot}}) &= \frac{1}{n}\sum_{s=1}^{n} E(f_{s*|D_s}) \\ Var(\bar{f}_{*|D_{tot}}) &= \frac{1}{n^2}\sum_{s=1}^{n} Var(f_{s*|D_s}) \end{aligned}$$

where n denotes the number of subjects, while D_{tot} represents qualitative feedback data collected from all participants, D_s represents individual qualitative feedback data, and f_s represents individual perceived similarity scores with respect to the reference.

4.3.4.2 Post-Hoc Model Tuning

To create a more accurate averaged posterior model for perceived similarity, we apply post-hoc hyper-parameter tuning to select better parameters to use in averaged GP-model-based analysis. In our analysis we have used the MLE approach suggested in [19] to tune only θ hyper-parameter of RBF kernel 3.11 but one can extend this methodology to tune all hyper-parameters of the model.

We first define a log-likelihood for all posterior models and θ hyper-parameter. Let $P(D_{tot}|\theta_{rbf})$ define the probability of all qualitative feedback based on the implemented θ_{rbf} hyper-parameter. As the experiments are conducted independently, $P(D_{tot}|\theta_{rbf})$ can be written as a multiplication of probabilities of individuals' qualitative feedback based on provided θ.

$$P(D_{tot}|\theta) = \prod_{s=1}^{12} P(D_s|\theta_{rbf}) \tag{4.4}$$

Then, the log-likelihood can be written as summation per individual

$$log\left(P(D_{tot}|\theta_{rbf})\right) = \sum_{s=1}^{12} P(D_s|\theta_{rbf}) \tag{4.5}$$

The total sum of log-likelihood is found by using the MLE method provided in 3.5. Based on this approach we have selected a θ_{rbf} value where it maximizes the summation of log-likelihood.

4.3.4.3 Just Noticeable Difference Thresholds

Utilizing the averaged GP posterior model, one can estimate the probability of a participant preferring a rendered parameter set instead of the reference parameter set. The probability model quantifies the probability of each parameter set being as effective as the reference, in terms of perceived similarity to the reference. Let $\bar{f}_{*|D_{tot}}$ be the averaged posterior distribution of latent function f for arbitrary parameter set x_*, where $\bar{f}_{*|D_{tot}}$ is calculated according to Eqn. (4.3). Similarly, let $\bar{f}_{ref|D_{tot}}$ be the averaged posterior distribution of latent function for reference parameter x_{ref}. Then, the probability of a participant preferring the arbitrary parameter x_* instead of reference can be calculated as

$$P(x_* \succ x_{ref}|D_{tot}) = \Phi\left(\frac{E(\bar{f}_{*|q_{tot}} - \bar{f}_{ref|D_{tot}})}{\sqrt{Var(\bar{f}_{*|D_{tot}} - \bar{f}_{ref|D_{tot}}) + c_p^2}}\right) \tag{4.6}$$

where $E(\bar{f}_{*|D_{tot}} - \bar{f}_{ref|D_{tot}})$ is the difference between mean values of the posterior distributions, $Var(\bar{f}_{*|D_{tot}} - \bar{f}_{ref|D_{tot}})$ denotes the variance of the difference between $\bar{f}_{*|D_{tot}}$ and $\bar{f}_{ref|D_{tot}}$, and c_p is the coefficient used in defining the standard deviation of the white noise interference in Eqn. (6.36).

Preferring the rendered parameter set over the reference indicates that a participant either assigns a higher perceived similarity score to the rendered parameter set compared to the reference or is unable to distinguish between the rendered parameter set and the reference; hence, randomly selects one of the options. If the probabilities of selecting the rendered parameter set and the reference are equal when the participants cannot differentiate between the two options, and if we assume that the participants always prefer the reference when they notice a difference between the two choices, then utilizing the Bayesian inference, the probability of participants preferring the reference by detecting the difference can be computed as $1 - 2P(x_* \succ x_{ref} | q_{tot})$.

To determine a confidence interval where participants are more likely to prefer the reference, the JND threshold can be selected at 25%. This means that if a parameter set falls below this threshold, it's likely to be noticeably different from the reference with a probability greater than 50%.

4.4 Experimental Setup

4.4.1 Participants

Twelve participants (9 males and 3 females) with an average age of 25.91 ± 1.16 years took part in our study. Among them, only one person was left-handed, and none had any known sensory-motor disabilities. All participants signed an informed consent form approved by the Institutional Review Board of Sabanci University before the experiments began. None of the participants had significant prior experience with haptic interfaces or psychophysical studies. The only compensation provided to participants was coverage of their transportation costs.

4.4.2 Apparatus

The experiment setup, shown in Figure 4.3, consists of three identical haptic interfaces and a visual display. Equal-sized knob-shaped 3D-printed knobs were used as the end-effector of the haptic interfaces. The visual display was utilized for the visual-haptic congruency experiments, such that the visual feedback of all knob rotations was made available on the monitor, while the view of the haptic interfaces was kept hidden from the participants.

Each haptic interface includes a direct-drive Maxon RE40 brushed DC motor, which is equipped with a 1024 count/rev encoder. Motors operate in current mode and are controlled by a MaxPos digital positioning controller via an EtherCAT interface.

An open-loop impedance controller was implemented to render a torsional spring via a haptic interface. The controller is executed in real-time at a frequency of 1 kHz within the Matlab RealTime environment.

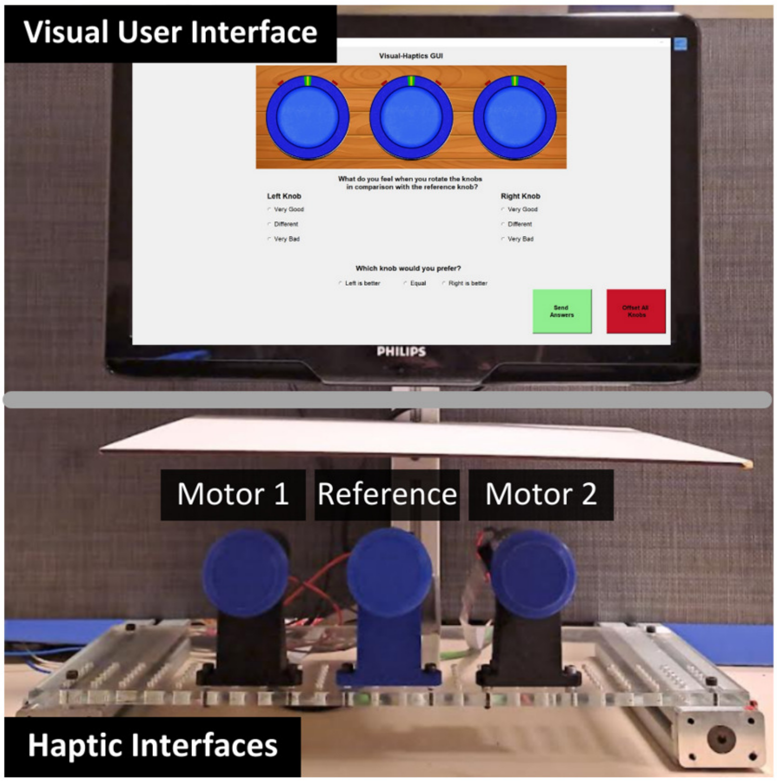

FIGURE 4.3: Three identical haptic interfaces and a visual display.

Haptic rendering of torsional spring with stiffness value K modeled according to Hooke's law:

$$\tau = K\,\theta \tag{4.7}$$

where θ is the rotation angle of the physical knobs and τ is the reference torque generated as a result of this rotation.

The stiffness range of $[0.1, 0.4]$ Nm/rad was utilized in the experiments and mapped to the normalized range of $[0, 1]$ during optimization. The reference stiffness was 0.2 N-m/rad, corresponding to the normalized value of 0.5.

4.4.3 Control-Display Ratio

The perceived stiffness of virtual torsional springs is studied by controlling the visual scaling factor, called control-display (C/D) ratio [78] of the knob rotations. The method involves applying a visual scaling to the actual knob rotation

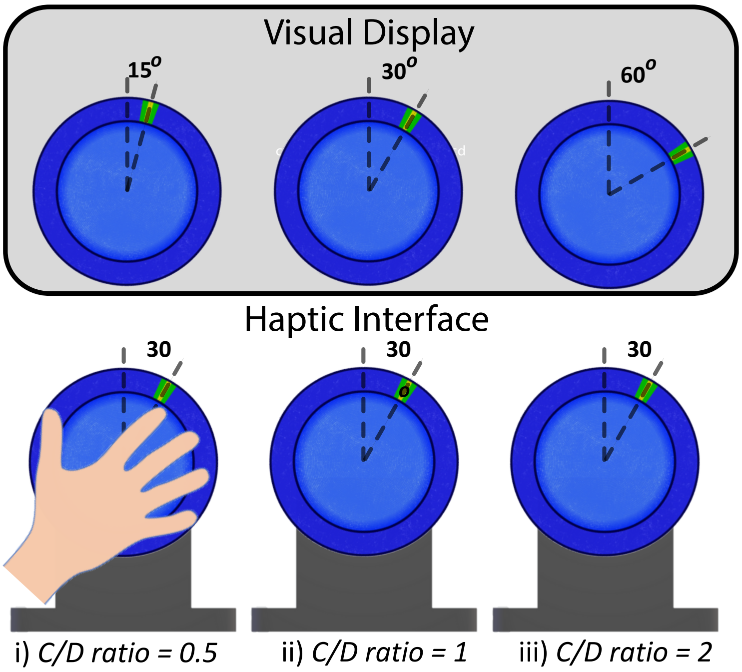

FIGURE 4.4: Knob rotations of the haptic interfaces and corresponding displayed visual rotations for the different C/D ratios.

of the haptic interface and allowing visual feedback only from the monitor, as depicted in Figure 4.4.

If the C/D ratio is less than one, the visualized movement is reduced, requiring the user to rotate the knob more to achieve the same screen rotation as with a higher C/D ratio. Contrarily, if the C/D ratio is greater than one, the movement is amplified, so the user needs to rotate the knob less to see the same screen rotation as with a lower C/D ratio.

The C/D ratio range of $[0.5, 2]$ is considered, such that two times visual amplification or reduction can be provided for a given knob rotation. During the optimization, C/D ratios are normalized to $[0, 1]$. The inverse of the following mapping is applied to normalize the C/D ratios: $C/D_{ratio} = 2^{(2\alpha-1)}$, where α denotes the normalized C/D ratio.

4.4.4 Hypotheses

We have designed an experiment to test the validity of the following hypotheses:

H1 The perception of users can be successfully manipulated by changing the visual modality without altering the rendered stiffness parameter, to make virtual environments feel stiffer or more compliant. It is possible to increase the perceived compliance via amplification and to increase the perceived stiffness via attenuation of the visual motion feedback.

H2 The visual-haptic incongruency limits the range of visual scaling for which the perception of users can be manipulated with high perceived realism.

To determine these limits, the JND thresholds for stiffness and C/D ratio can be estimated through the HiL optimization experiments.

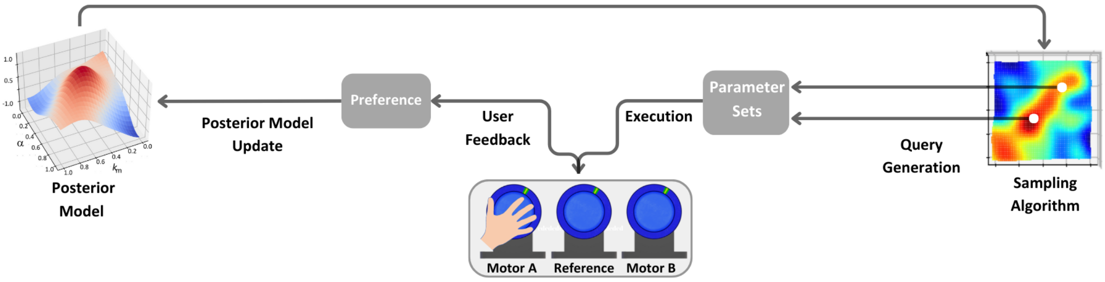

FIGURE 4.5: A schematic representation of the preference-based HiL optimization

H3 Movement-related cues from the haptics modality and visual cues are (partially) integrated through a weighted linear combination of redundant sensory cues. Accordingly, the perceived stiffness is formed as a linear combination of idealized visual stiffness and haptic stiffness values. Furthermore, as the incongruency level increases, the contribution of visual cues and their weights in the stiffness perception decreases.

4.4.5 Experimental Procedure

4.4.5.1 Setup and Overview

All volunteers participated in the experiment through a single seating and used their dominant hands. To minimize the effects of auditory cues, participants wore headsets playing pink noise. During the experiment, participants were instructed that the knobs control the visual display and asked to interact with the three knobs similarly. Participants were allowed to explore these systems as they preferred. To ensure that the visual cues were only provided from the display, all three haptic interfaces were covered throughout the experiment.

During all sessions, the middle knob was used as the reference rendering that maintained the same stiffness value. The visual feedback for the middle knob corresponded exactly to the physical rotation of the haptic interface, ensuring that the visual representation accurately reflected the actual displacement, without introducing any visual manipulation.

4.4.5.2 Procedure

Participants were asked to compare the stiffness of torsional springs rendered by three identical haptic interfaces, each coupled to the visual feedback displayed on the monitor. Participants provided feedback based on the perceived similarity to the reference. This task involved both an ordinal classification and a pairwise comparison to evaluate the similarity of the stiffness renderings.

All optimization trials followed the same experimental procedure for each trial, as depicted in Figure 4.5. In each trial, two different sets of parameters called the *query*, were sampled according to the UCB acquisition function (Query Generation in Figure 4.5). Then, the participants were presented with the parameters of the query through the left and right haptic interfaces, while the reference model was rendered by the haptic interface in the middle (Execution in Figure 4.5). Participants were asked to interact with the reference haptic interface at the beginning of each trial. After that, they were free to interact with the haptic interfaces on the left and right, and also go back to the reference, in any order they preferred. They were instructed to apply similar trajectories to all three knobs, ensuring they stay within the predetermined displacement range presented via the GUI as visual constraints. There was no time limitation for the trials.

Following the interactions with the haptic interfaces, participants were asked to compare the left and right knobs against the reference knob. Participants provided their responses to the two questions, based on their internal comparison (User Feedback in Figure 4.5).

For the first question, they evaluated the similarity of the feel of the right/left knobs, in comparison to the reference knob. For the ordinal classifications, they were asked to answer the question "*How do you feel when you interact with the knobs in comparison to the reference knob?*" by selecting one of the following responses:

- *Not Different*: the left (right) knob does not feel noticeably different compared to the reference knob.

- *Different*: the left (right) knob has a noticeably different feeling compared to the reference knob.
- *Very Different*: the left (right) knob has a significantly different feeling compared to the reference knob.

For the second question, they performed a pairwise comparison of the feelings of the left and right knobs with the reference knob and were asked to provide qualitative pairwise feedback by answering "*Which knob do you prefer?*" through one of the following responses:

- *Left*: the left knob feels more similar to the reference knob than the right knob.
- *Right*: the right knob feels more similar to the reference knob than the left knob.
- *Equal*: an equal level of similarity/dissimilarity exists for the left and right knobs compared to the reference knob.

Participants submitted their preferences through a GUI. Following this, the participants' preferences were utilized to update the GP for the perceived similarity with respect to the reference (Posterior Model Update in Figure 4.5).

4.4.6 Sessions

The experiment was organized into three distinct sessions: multi-modal optimization, constant stiffness optimization, and cross-comparison sessions. The

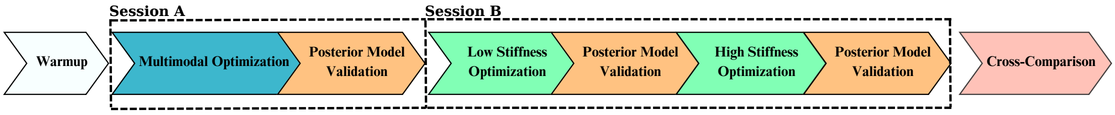

FIGURE 4.6: A schematic representation of the experimental procedure

constant stiffness optimization session was designed to validate the findings acquired from the multi-modal optimization.

Figure 4.6 depicts the experimental flow: Each experiment started with a warm-up and ended with a cross-comparison. Ten-minute breaks were scheduled between the sessions and the volunteers were also allowed to take a break whenever they wanted. Additionally, the order of the multi-modal optimization (Session A) and the constant stiffness optimization (Session B) sessions, as well as the order of the low- and high-stiffness experiments, were randomized for each participant to prevent any patterns or learning effects during the experiments.

4.4.6.1 Warm-Up

The warm-up took about five minutes and was used to familiarize the volunteers with the haptic rendering task and the operation of the visual-haptic interface. During warm-up, all subjects were provided with six different queries with 12 different renderings, and additional sets of queries were provided until volunteers felt ready.

4.4.6.2 Constant Stiffness Optimization

The constant stiffness optimization sessions were designed to optimize the C/D ratio while keeping the stiffness parameter fixed at a predetermined value. The constant stiffness optimization consisted of two sub-sessions, denoted as high- and low-stiffness optimizations. These sub-sessions were configured to provide stiffness levels approximately 1.3 times higher and 1.3 times lower than the reference stiffness value of 0.2 Nm/rad, respectively.

The constant stiffness optimization sessions consisted of 12 trials for the low-stiffness and 12 trials for the high-stiffness experiments. It took a total of 36 iterations, including six trials for the low-stiffness posterior model validation and six trials for the high-stiffness posterior model validation. The constant stiffness optimization sessions with posterior model validations took about 30 minutes.

4.4.6.3 Multi-modal Optimization

During the multi-modal optimization session, both the C/D ratio and the stiffness were concurrently optimized to explore the entire parameter space encompassing the C/D ratio and rendered stiffness variables.

The multi-modal optimization session consisted of 30 trials with 24 optimization trials and six posterior model validation trials. The multi-modal optimization session with the posterior model validation trials took about 30 minutes. The number of trials and the hyper-parameters used in the optimization were decided based on pilot studies.

4.4.6.4 Posterior Model Validation

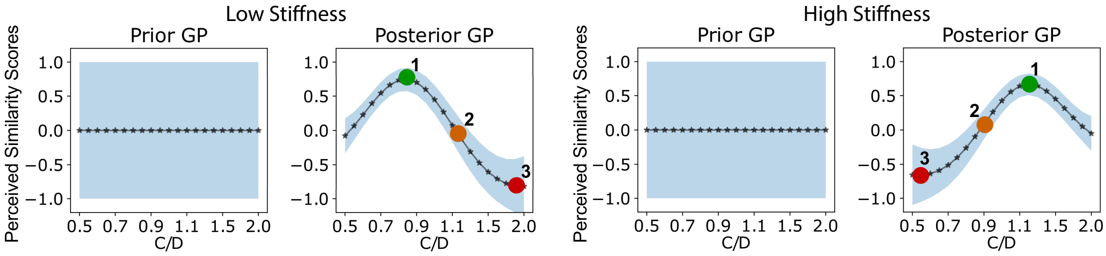

Figure 4.7: Sample prior and posterior models (captured by their mean and standard deviation) for perceived similarity depicted for the one-dimensional experiments, where the stiffness was kept constant at 0.151 N-m/rad for low and 0.262 N-m/rad for high stiffness cases, and the C/D parameter was varied. The green, yellow, and red points mark the maximum, mean, and minimum of the posterior, where the maximum and minimum points have the highest and lowest perceived similarity scores. The mean point has a perceived similarity score equal to the average of maximum and minimum scores.

Posterior model validations were conducted after each constant stiffness optimization session and the multi-modal optimization session. During the posterior model validations, a set of queries was generated with the parameters that give the maximum, mean, and minimum perceived similarity scores of the posterior model. Participants were asked to perform six pairwise comparisons of the maximum, mean, and minimum parameter sets for two blind orderings, for which the expected ordering was 1-maximum, 2-mean, and 3-minimum, as depicted in Figures 4.7 and 4.8. These validation sessions were aimed to confirm the reliability of the posterior model derived through preference-based learning.

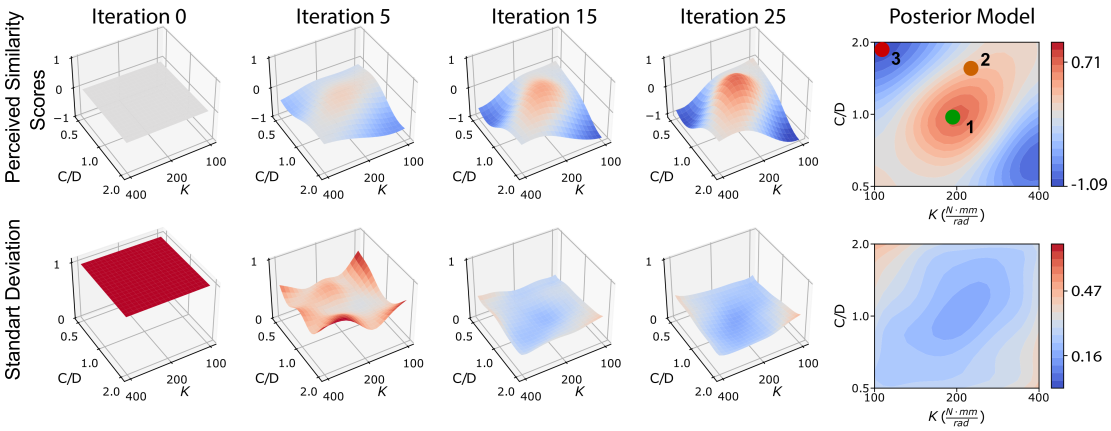

FIGURE 4.8: The progression of the posterior model of multi-modal perceived similarity depicted at various trials of the HiL optimization for a participant. The first row captures the mean, while the second row presents the standard deviation of the posterior model. The green, yellow, and red points mark the maximum, mean, and minimum values of perceived similarity scores. The mean point has a similarity score equal to the average of maximum and minimum scores.

4.4.6.5 Cross-Comparison

After the completion of all optimization sessions, cross-comparison trials were conducted to test if participants could distinguish their optimized parameters from the reference parameter. During the cross-comparison, participants were asked to compare two rendering models with each other presented by the left and right knobs. Participants followed the same procedure with the optimization sessions to provide pairwise and ordinal feedback as explained in Section 4.4.5.2. Queries were not generated by the acquisition function, instead, all optimized parameter sets for the low-stiffness, high-stiffness, and multi-modal settings, together with the reference parameters, were utilized. At each trial, participants provided pairwise and ordinal feedback based on perceived similarity to the reference knob in the middle. Each comparison was repeated two times for each parameter set, and all comparisons were completed in 12 trials, resulting in six pairwise comparisons and six ordinal classifications for a participant. The cross-comparison session took about 10 minutes.

4.4.7 Data Analysis

Statistical tests based on the generalized estimating equations (GEE) were applied to check for statistically significant differences in ordinal classifications of the low-, high-stiffness, and multi-modal optimal parameter sets.

The performance of all three learned perception models was studied by constructing confusion matrices. The parameters corresponding to the maximum, minimum, and mean values of the learned GP posterior model and the user ordering were utilized for multi-class classification.

4.5 Results

4.5.1 HiL Optimizations

The GP prior and posterior models, captured by their mean and standard deviation, for a sample participant during the low- and high-stiffness rendering experiments are given in Figure 4.7. The progressions of the GP over the trials during multi-modal experiments for a sample participant are presented in Figure 4.8. The last columns of Figures 4.7(a)-(b) and 4.8 also present the orderings of the participants, corresponding to the maximum (1), mean (2), and minimum (3) of the posterior.

The average optimized C/D ratio of the 12 volunteers for the low-stiffness setting is determined as 0.77 ± 0.12, while the average C/D value for the high-stiffness setting is identified as 1.35 ± 0.18. The average stiffness and C/D ratio values of the multi-modal setting are determined as 0.20 ± 0.03 (Nm/rad) and 1.06 ± 0.25, respectively.

4.5.2 Posterior Model Validation

To validate the optimized perception models, the participants ordered the minimum, mean, and maximum values of the posterior models. GP orderings were considered the ground truth and were used along with blind orderings provided by the participants for the multi-class classification. The accuracy of the orderings is presented as confusion matrices in Figure 4.9. Each cell in the matrices indicates the normalized number of correct classifications. Since the participants were allowed to indicate ties, the rows/columns of the confusion matrices may not sum up to one. The dominance of the diagonals shows how well the Gaussian posterior model can capture the underlying perception model.

(a) High Stiffness Posterior Model (b) Low Stiffness Posterior Model (c) Multimodal Posterior Model

FIGURE 4.9: Confusion matrices of the posterior model validation

4.5.3 Cross-Comparison

During the cross-comparison sessions, each optimal parameter set was classified 72 times. The optimal parameter sets were labeled as not different from the reference parameter set for 43.1%, 48.6%, and 68.1% of these trials for the low-stiffness, high-stiffness, and multi-modal cases, respectively.

GEE regression analysis indicates that while classifying whether an optimized parameter set is distinguishable from the reference parameter set, the multi-modal optimized parameter set has a statistically significant tendency to be labeled as “similar to the reference” with $p = 0.022$ and log-odd ratio $= 0.756$, while the analysis results are inconclusive for the low- and high-stiffness optimized parameter sets, with $p = 0.404$ and log-odd ratio $= -0.2796$, and $p = 0.83$ and log-odd ratio $= -0.0556$, respectively.

The analysis results indicate that the multi-modal optimized parameter set was often perceived as similar to the reference.

4.5.4 Averaged Gaussian Process Posterior

The averaged multi-modal perception model of 12 participants is computed according to Section 4.3.4.1 after post-hoc tuning of kernel hyper-parameter,θ_{rbf}.

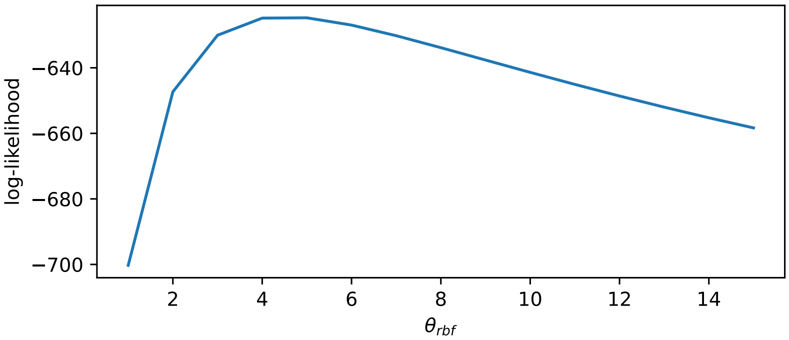

Figure 4.10: Post-hoc Kernel tuning for averaged posterior Model

According to the post-hoc MLE, θ_{rbf} is selected as 5 for the derivations of averaged posterior GP-based results.

The mean and standard deviation of averaged multi-modal posterior GP distribution are presented in figure 4.11 and figure 4.12 respectively.

In particular, Figure 4.11 presents the mean value plot annotated with the average values of the high, low, and multi-modal stiffness optimization results. The red regions in the figure 4.11, represent higher perceived similarity scores and the stars depict the mean parameter values for the low stiffness, high stiffness, and multi-modal optimizations.

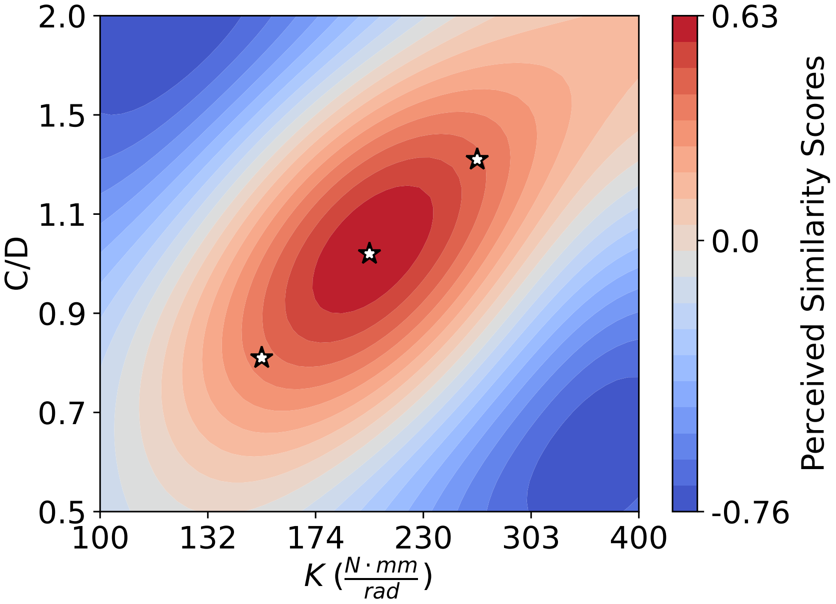

FIGURE 4.11: The mean values of the averaged GP posterior

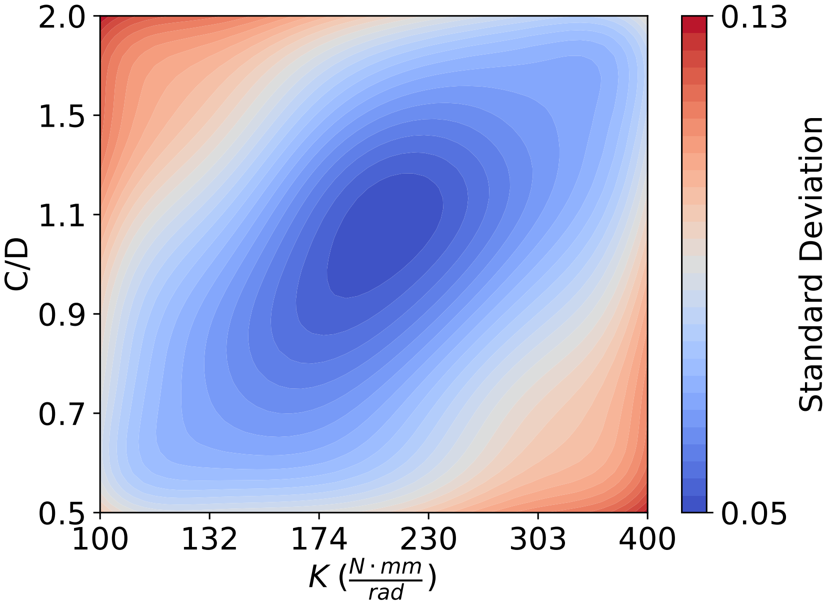

FIGURE 4.12: The standard deviation of the averaged GP posterior

Figure 4.12 presents the standard deviation of the averaged GP posterior. In figure 4.12, the red regions have higher standard deviation as they were sampled less frequently in comparison with blue regions which are estimated to have lower perceived similarity scores.

4.5.5 Posterior Probability Models for Classifications and Pair-wise Preferences

Figures 4.14 and 4.13 presents a re-interpretation of Figure 4.11 and 4.12 through the probability mapping discussed in Section 4.3.4.3.

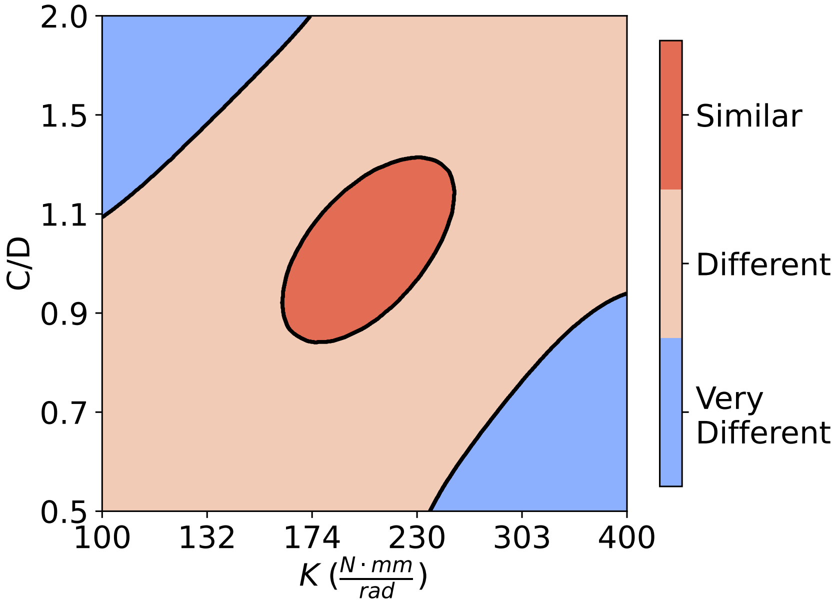

FIGURE 4.13: Ordinal classification map for discriminating a parameter from reference

Figure 4.13 shows a classification plot for estimated ordinal selections. The red regions show where participants can perceive a rendering as not different than the reference. Participants perceive the stiffness differently in beige regions and very differently in blue regions in comparison with the reference.

Figure 4.14 shows a probability plot for estimated pair-wise selections in which the region inside the JND curve captures the parameters that can be used instead of reference parameters. The lines on the figure are used to analyze the effect of the C/D ratio over the perceived stiffness.

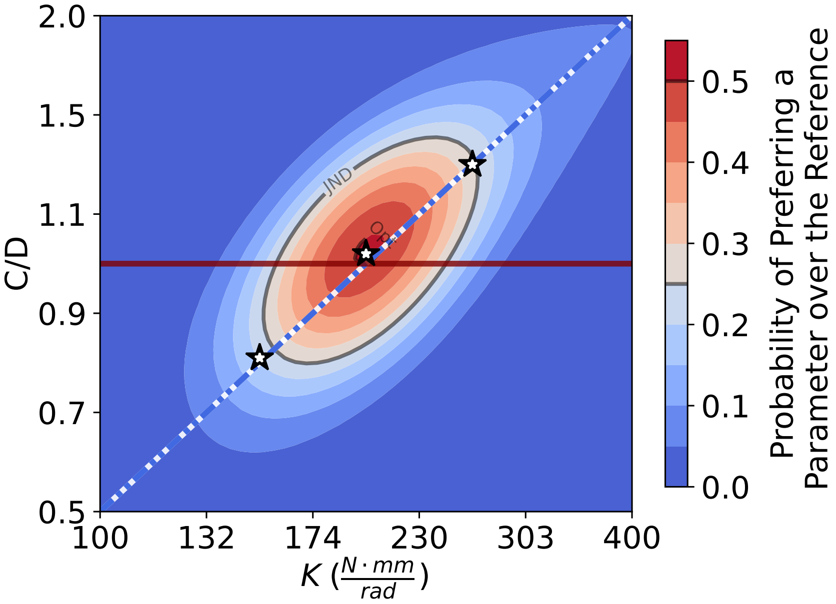

FIGURE 4.14: Probability map of selecting a parameter set in pairwise comparison to the reference

4.5.5.1 Just Noticeable Difference Thresholds

The estimated JND thresholds for perceived stiffness are depicted in Figure 4.14. The contours in this figure capture the probability of the rendered parameters being indistinguishable from the reference, as well as the boundary of JND[1]. The average values for the low, high, and multi-modal stiffness optimization results are also annotated with stars in the figure.

4.6 Discussion

In this section, we elaborate on the validity of each hypothesis presented in Section 4.4.4 and compare our results with the relevant literature.

---

[1] Probabilities exceeding 50% in Figure 4.14 are artifacts of the slight mismatch between the multi-modal optimized and the reference stiffness.

4.6.1 Hypothesis 1

The preference-based learning algorithm simultaneously constructs a Gaussian latent model of the perceived similarity of each individual, as depicted in Figures 4.7 and 4.8. Accordingly, it becomes possible to assign a perception score to any parameter set within the design space employing Gaussian regression, even if that particular parameter set has not been directly assessed. The posterior model validation results indicate that the GP model successfully captures the perceived similarity of the participants, as evidenced by the diagonal dominance in the confusion matrices, presented in Figure 4.9. These results demonstrate the efficiency of the GP in modeling the latent perception model, while concurrently optimizing the rendering model parameters.

Low standard deviations in the posterior plots indicate adequate convergence of the GP models. Compared with the standard deviation of the averaged model prior, the standard deviation of the averaged model posterior has reduced by more than 80% in the blue regions, indicating a high information gain. Accordingly, the blue dominant region Figure 4.12 indicates where the averaged multi-modal perception model can be utilized with high confidence.

The red dominant region of the averaged posterior GP model in Figure 4.11 indicates the stiffness and C/D ratio parameters for which the perceived similarity of the rendering compared to the reference is evaluated to be high. The average posterior model shows that the parameter sets that are close to the diagonal line defining the visual stiffness have the highest perceived similarity scores, compared to other parameter sets.

The high-stiffness optimization result further validates the effectiveness of visual scaling on perceived stiffness, since participants evaluated the perceived stiffness as not statistically significantly different from the reference model. This result demonstrates that it is possible to increase the perceived compliance via motion amplification with high perceived realism.

Overall, the results provide strong evidence that by proportionally increasing (decreasing) the C/D ratio, one can cause a haptic interface to feel more compliant (stiffer). Accordingly, the first hypothesis H1 is accepted.

4.6.2 Hypothesis 2

While the averaged posterior GP model in Figures 4.11 4.12 provides evidence of the feasibility of perceived stiffness modulation via visual scaling, further insights may be gained about the quality of such renderings by inspecting the perceived similarity of scores compared to the reference model.

The averaged posterior model in Figure 4.11 captures a significant decrease in the perceived similarity scores as the C/D ratio deviates further from the unity. For instance, the cross-comparison results of the multi-modal, high-stiffness, and low-stiffness optimization indicate that on average these parameter sets were declared to be similar to the reference model for 68.1%, 48.6% and 43.1% of the trials, respectively. The cross-comparison result for the multi-modal optimal parameter set is statistically significant, indicating high perceived similarity with respect to the reference. On the other hand, while the statistical tests for high-stiffness and low-stiffness optimal parameter sets are inconclusive, these results are in good agreement with Figures 4.11 4.12.

Figure 4.14 presents the JND threshold estimated for the perceived stiffness according to Section 4.5.5.1. The area inside the JND threshold curve captures the parameter sets that are indistinguishable from the reference rendering, while the parameter sets outside this closed curve have detectable differences. Please note that since users were not asked to simply compare stiffness but to rate the overall perceived similarity of the rendering, the existence of a detectable difference does not necessarily imply that the perceived stiffness of the rendered set is evaluated to be different; it only implies that users can detect some difference between the renderings, for instance, possibly due to detection of visual manipulation.

The JND threshold for the perceived stiffness depicted in Figure 4.14 provides an explanation for the cross-comparison results. In particular, while the average of the high-stiffness optimization results resides on the JND boundary, the average of the low-stiffness optimization results falls outside of the JND boundary; that is, on average the high-stiffness optimization parameter set is not reliably distinguishable from the reference model by the participants (declared to be similar for 48.6% of cross-comparison trials), while the low-stiffness optimization parameter set is distinguishable (declared to be similar for only 43.1% of cross-comparison trials). Furthermore, the multi-modal parameter set is not distinguishable from the reference set and is declared to be similar for 68.1% of cross-comparison trials, with statistical significance.

We further study the effect of visual-haptic incongruency on the JND thresholds by considering slices of Figure 4.14 along the straight lines depicted in the figure. The solid red line in Figure 4.14 captures the case when the C/D ratio is set to unity, while the dashed blue line considers the case when the stiffness and C/D ratio parameters are increased proportionally, such that the visual stiffness is kept constant. Figure 4.15 presents the cross-sections of these slices characterizing the effect of visual cue modulation on stiffness perception. The horizontal gray line in Figure 4.15 depicts the JND threshold.

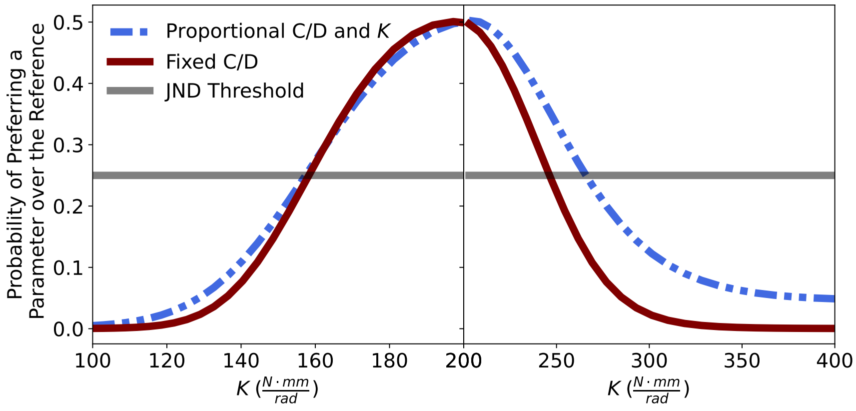

FIGURE 4.15: Probability of parameters to be preferred over the reference, computed by considering slices of Figure 4.14 along the straight lines depicted in that figure. The horizontal grey line depicts the JND threshold.

The virtual stiffness JND level in Figure 4.15 for the unity C/D ratio, corresponding to changing stiffness without any visual manipulation, is consistent with the results reported in the literature [98, 99]. The blue dashed curve in Figure 4.15 represents a proportional relationship between stiffness and C/D ratio parameters and provides supporting evidence that the perception of stiffness can be enhanced through visual scaling with high perceived realism. The blue dashed curve indicates that as the C/D ratio further deviates from unity, visual incongruency increases and finally interferes with the perceived realism of rendering. In particular, while increasing the perceived compliance, a proportional increase of the C/D ratio can extend the JND range from 20% to 35% of the reference stiffness in comparison to the case when the C/D ratio is set to unity. Hence, the JND threshold can be significantly increased through visual-haptic incongruency. Accordingly, by increasing the C/D ratio, one can cause a stiff haptic interface to feel more compliant in a manner that is indistinguishable from the reference model. However, the range of stiffness modulation with high perceived realism is not symmetrical; a similar effect through visual cue manipulation (decreasing the C/D ratio) is not observed (with a meaningful effect size) when aiming to achieve stiffer perceived renderings that are indistinguishable from the reference stiffness. In particular, under visual attenuation, while the users perceive the stiffness to be higher than the haptic stiffness, they also indicate that they can detect a difference compared to the reference model. This detectable difference is possibly due to the experimental setup necessitating excitations from the forearm rotations. A low C/D ratio requires a larger range of motion to achieve a similar visual excitation and this causes the forearm to approach its joint limits. Furthermore, such a scaling requires humans to perform more work for the same visual input. Overall, these effects may result in a detectable difference compared to the reference rendering when a C/D ratio is set lower than unity.

To summarize, our results suggest that while high C/D ratios can increase the perceived compliance in a manner that is indistinguishable from the reference model, low C/D ratios can increase the perceived stiffness, but users evaluate the overall perceived realism to be distinguishable from the reference model. Hence, the perceived realism due to visual-haptic incongruency is not symmetrical and favors increasing the perceived compliance.

Overall, the cross-comparison results and the JND threshold in Figures 4.14 and 4.15 demonstrate that the perceived realism under visual-haptic incongruency can be kept high for a limited range of visual scaling. Accordingly, the second hypothesis H2 is accepted.

4.6.3 Hypothesis 3

To study sensory integration using the averaged posterior model, in Figure 4.16, we consider vertical slices along the Figure 4.14 to determine the best C/D ratio for a given fixed stiffness and horizontal slices to determine the best stiffness for a given C/D ratio, respectively.

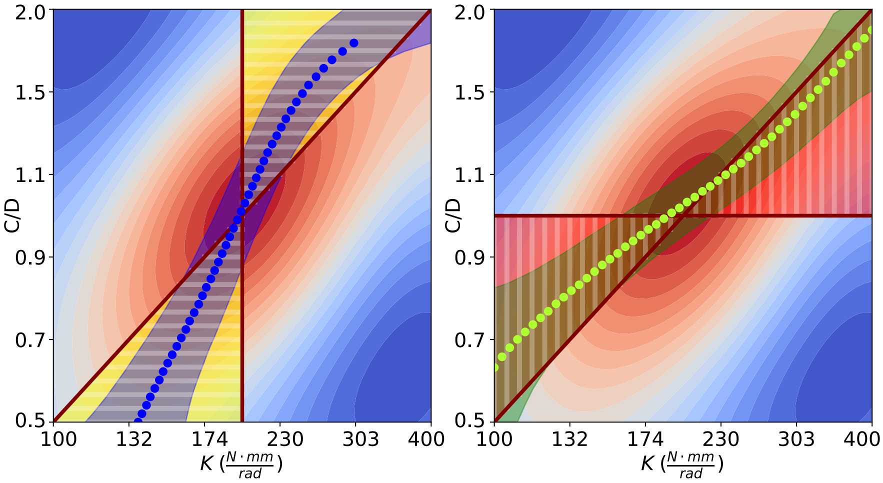

FIGURE 4.16: (a) The yellow area indicates the region of sensory integration for constant C/D values with variable stiffness. Blue points have the highest perceived similarity score along the horizontal axes, with the blue-shaded region around these points denoting one standard deviation confidence interval. (b) The red area indicates the region where sensory integration for constant stiffness with variable C/D ratios. Yellow points have the highest perceived similarity score along the vertical axes, with the green shaded region around these points denoting one standard deviation confidence interval.

Along the horizontal lines, as depicted in Figure 4.16(a) where a constant C/D ratio is considered, users determine the best stiffness that reduces the perceived difference between the rendered parameters and the reference rendering. If the users neglect the scaled visuals and only rely on the movement-related cues from the haptic modality, the perceived stiffness should be the reference stiffness. If users utilize only the visual movement cues presented on the screen and neglect the movement-related cues from the haptic modality, then the perceived stiffness should be the visual stiffness and lie on the diagonal line. If users utilize movement-related cues from both haptic and visual modalities and if sensory integration is employed, then the perceived stiffness should lie in the yellow triangular areas covering the range between the reference stiffness and visual stiffness.

Similarly, along the vertical lines, as depicted in Figure 4.16(b) where a constant stiffness is considered, users determine the best C/D ratio that reduces the perceived difference between the rendered parameters and the reference rendering. If users utilize movement-related cues from both modalities and if sensory integration is employed, then the perceived stiffness should lie in the red triangular areas covering the range between the haptic and the visual stiffness.

The points marked in Figures 4.16(a) and 4.16(b) depict the points with the highest probability along their respective (horizontal and vertical) search directions, while the shaded regions around these points denote one standard deviation confidence intervals. Since these points reside inside the colored triangular areas, it can be concluded that sensory integration exists for stiffness perception between visual and haptic cues, as captured by our probabilistic map.

Furthermore, the points in Figures 4.16(a) and 4.16(b) can also be used to estimate the relative contributions of haptic and visual cues utilized by stiffness perception under visual-haptic incongruency, if a weighted linear combination of the cues is assumed for (possibly partial) integration. For instance, the diagonal line in Figure 4.16(a) indicates the case when movement-related cues come only from the visual stimulus, resulting in the estimate from the visual stimulus S_v. The horizontal line in Figure 4.16(a) indicates the case when the movement-related cues come only from the haptic modality, resulting in the estimate from the haptic stimulus S_h. The points in Figure 4.16(a) satisfy the subjective equality, called PSE [69]. The visual weights can be computed as a result of linear interpolation, $w_v = (PSE - S_h)/(S_v - S_h)$, while haptic weights can be found as $w_h = 1 - w_v$.

Figure 4.17 presents the estimates of the contribution (or weight) w_v of visual cues utilized during sensory integration with respect to the C/D ratio, computed according to Figure 4.16(a). Only the parameters corresponding to the adequately explored regions of the parameter space are included in Figure 4.17 to preserve the reliability of the estimates[2].

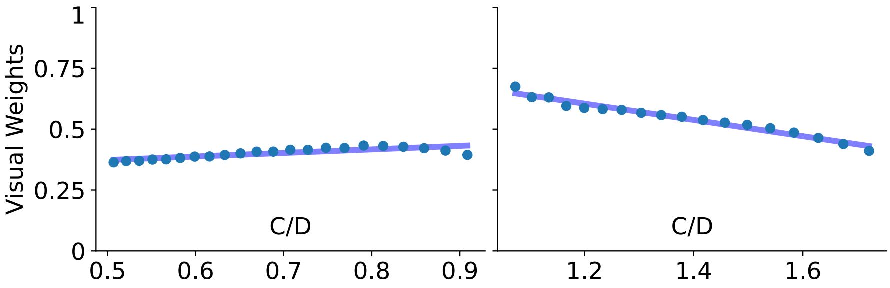

FIGURE 4.17: Estimation of the weight w_v of visual cues utilized during sensory integration with respect to C/D ratio.

The results indicate that as C/D ratio deviates further from unity increasing the level of incongruency, the contribution w_v of visual cues in sensory integration decreases. The crossing of the 0.5 thresholds in Figure 4.17 further indicates that the dominant sensory modality for movement-related cues switches from vision to haptics as the C/D ratio deviates from the unity. Moreover, the visual contributions are relatively higher for the range of C/D ratios that are aimed to increase perceived compliance, when compared to the range of C/D ratios aimed to increase perceived stiffness. These weights explain the non-symmetrical nature of the perceived realism presented in Figure 4.15, indicating that high C/D ratios result in a rapid decrease in the contribution of the visual cue. Finally, an extrapolation of the decreasing contribution trend in Figure 4.17 suggests that vetoing of visual modality is highly likely to occur for large incongruency

---

[2] C/D ratios near the boundaries of our search space are omitted from the estimates as these parts of the probability map are under-explored. Similarly, the C/D ratio near unity is not included as the estimation becomes very sensitive to experimental noise around this value.

levels. These results provide sufficient evidence for the last hypothesis H3 to be accepted.

4.6.4 Comparison of Results with the Related Work

Our results are in good agreement with [65, 69, 79, 80], indicating that the sensory integration process during multi-modal visual-haptic compliance estimation under visual scaling can be modeled through a weighted linear combination of redundant sensory signals and manipulated visual cues can expand the perceived stiffness range of haptic rendering.

While our study is similar to [79, 80], we explore a continuous range of visual scaling levels from a perceived similarity perspective and our results are in more line with those of [65, 69]. Our findings extend [80] and support [65, 69], indicating a decreased contribution of the manipulated visual cue in the multi-modal percept as its congruency is reduced and suggests possible vetoing of this unreliable cue when strong conflicts arise between the haptic modality and manipulated visual feedback.

Our study supports the conclusions of [77, 80, 83], indicating that the use of pseudo-haptic feedback can enhance haptic feedback during stiffness rendering. Our approach differs as we rely on qualitative feedback-based HiL optimization to maximize perceived similarity with respect to the reference and determine the optimal levels of the visual scaling ratio.

Our results support the observations in [84, 86] stating visual-haptic incongruency significantly influences stiffness discrimination performance when the discrepancy between visual feedback and hand movements is large. We extend these results by explicitly characterizing the relationship between incongruency levels and perceived realism.

Finally, our JND estimations are consistent with the results of earlier psychophysical experiments, which reported the JND threshold as 23% [99] or in the 8% – 22% range [98].

While our study determines the relative contributions of visual and haptic modalities with respect to the visual scaling under the assumption of a weighted linear combination of these cues, it makes no attempt to determine the underlying model of (possibly partial) sensory integration.

One interpretation of the results in Figure 4.17 is possible by utilizing the causal inference model [76], which generalizes the MLE model of multisensory cue integration by considering the causes of the underlying sensory events. In particular, the causal inference model infers the probability of a common cause versus two independent causes to derive optimal predictions. Accordingly, this model involves two parameters, capturing the probability of a common cause and the relative weight of each cue for the sensory integration case.

On the one hand, the causal inference model reduces to maximum likelihood estimation when there is a high probability of a common cause. In our experiments, participants were provided explicit instructions emphasizing the common cause between visual and haptic modalities. Furthermore, while the visual feedback was scaled, the synchronization between the haptic input and visual output was preserved throughout the experiments, resulting in a convincing causal reason for a common cause. Accordingly, if one considers that the probability of independent cause was low and the discrepancies were sufficiently small and hard to detect, then the weights in Figure 4.17 may be attributed to the weights of MLE-type sensory integration.

On the other hand, due to the experiment design, our results do not allow for the extraction of the two parameters of the causal inference model from the data; hence, the common cause assumption cannot be verified. Furthermore, discrepancies may become large as the C/D ratio further deviates from unity. As an alternative interpretation, if one assumes that visual feedback completely dominates movement-related cues, then the weights in Figure 4.17 can also be interpreted as the probability of a common cause.

Overall, it is likely that the underlying model of sensory integration lies somewhere between these two interpretations, performing partial sensory integration under conflicting cues.

Chapter 5

HiL Pareto Optimization for Assist-as-Needed Control during Motor Skill Training

In this chapter, we introduce Pareto-based HiL optimization with multi-criteria Bayesian optimization and its application in developing assist-as-needed virtual reality guided training sessions for motor-task learning.

5.1 Introduction

Physical human-robot interaction (pHRI) is commonly used for training tasks, such as human motor skill training or robot-assisted physical rehabilitation of patients. In such training applications, the efficacy of the training, characterized through performance evaluations with no assistance after the training, is of utmost importance. Haptic assistance is provided only during the training, when the user is coupled to a force-feedback robot, to ensure that the execution of the task can be completed with sufficient performance. For instance, consider a stroke patient going through physical rehabilitation, for whom the task execution may not be at all feasible without a proper level of assistance. On the other hand, too much assistance during training is known to be detrimental, as users may learn to rely on the existence of this support and slack [100]. Too much assistance may also cause the task to become not sufficiently challenging for the patient, resulting in a lack of engagement. As a consequence of excessive assistance, users may not learn how to successfully execute the task when no assistance is provided. Accordingly, there exists an inherent trade-off between the level of assistance provided and the level of perceived challenge for any given motor control task.

This trade-off has been widely acknowledged in the literature and it has been established that an optimal assistance level consists of the least amount of assistance that allows a user to able to execute a task with a sufficient level of success. Controllers that aim to provide such assistance are commonly referred to as assist-as-needed (AAN) controllers. AAN controllers aim to keep the assistance level at a proper level to maximally challenge but not overwhelm users with task difficulty or demotivate them with continual failures.

Most of the literature on AAN control focuses on the design of interaction controllers that can administer assistance forces safely and naturally, without overriding the user's intent. In particular, path-following controllers, such as velocity field controllers [101], and controllers with guaranteed coupled stability properties, such as passive velocity-field controllers [102], have been proposed and adopted in many related works [103, 104, 105, 106, 107]. While these studies on interaction control are indispensable for the safe and natural delivery of force feedback assistance, all these control approaches necessitate the proper level of assistance to be provided as input, typically by a domain expert.

In AAN controllers, the proper level of assistance is commonly decided empirically or heuristically, based on thresholds imposed on the measured performance/signals of the user. Most of the methods utilize sensor inputs or biosignals [108, 109, 110, 111], such as EMG and EEG, to adjust the level of assistance to promote voluntary participation based on thresholds. Adaptive controllers that utilize the dynamic model of the user and the device to minimize a cost function [112] or that use statistical estimates of psychophysical thresholds [113] have been proposed. Reinforcement learning has also been used to develop AAN controllers that do not depend on user or device-specific parameters [114, 115]. Moreover, in addition to quantitative measurements, more qualitative aspects, such as the psychological states of the user, have also been estimated by neural networks to adjust assistance levels [116, 117].

While these studies are promising, the determination of a proper level of assistance to be provided to a user is still an open challenge, commonly delegated to a domain expert. The problem is challenging, as each individual is unique; furthermore, user preferences and perceptions undergo continual changes as learning/recovery takes place during training. For instance, the perceived level of challenge of a task under a certain level of assistance is likely to decrease as a user gets better at performing the task. Moreover, other user-dependent metrics that are not easy to quantify, such as motivation and comfort level, may affect the perceived challenge level. Hence, the ideal assistance level necessitates personalization through a characterization of the trade-off between the perceived challenge level versus the task performance of each user.

Characterization of the trade-off between the perceived challenge level versus the task performance of users can guide training by establishing a proper level of assistance to be used by AAN controllers. This characterization can be performed at different stages of training, such that the changes in user preferences and perceptions can be captured as learning takes place and the level of assistance be adjusted accordingly.

Furthermore, such trade-off characterizations can also serve as a novel means of evaluation of user performance. While typical evaluations for manual skill training or physical rehabilitation are performed when no assistance is provided, as this captures the real-life case, such evaluations cannot capture improvements in performance during the early phases of training. For instance, if the task is sufficiently hard for a user, then it may not be possible for the user to execute it successfully without any assistance. In such cases, evaluations with no assistance will not capture any improvement, as the user's success will stay very low. Comparison of the characterized trade-offs between the perceived challenge level versus the task performance of a user at different stages of training provides a feasible alternative. A comparison of the time evaluation of the trade-off (i.e., Pareto optimization results forming a set of non-dominated solutions) provides a rigorous means to fairly evaluate the progress of user performance. Furthermore, Pareto results also enable comparisons among various users. It is important to emphasize that ensuring the fairness of comparisons of user performance under assistance necessitates the characterization of the trade-off via Pareto optimization since multiple variables, such as the level of assistance and the perceived challenge level, need to be considered simultaneously.

In this study, we propose a human-in-the-loop (HiL) Pareto optimization approach to characterize the trade-off between the user's performance and the perceived challenge level for a motor learning task implemented in a virtual environment with force-feedback to be used in AAN control. During the optimization, the user performance is captured by a quantitative metric, while the perceived challenge level is modeled as a qualitative metric captured through preference-based qualitative feedback. A multi-criteria Bayesian optimization technique is utilized for the HiL Pareto optimization of this hybrid model with both quantitative and qualitative metrics. The sample efficiency of Bayesian optimization is a crucial aspect, as it enables the trade-off to be characterized in a systematic and efficient manner, without exploring the whole search space and inducing fatigue to the user.

Once the trade-off is characterized via HiL Pareto optimization, we demonstrate how this trade-off can be used to guide training sessions with "optimal" assistance levels. In particular, we discuss how a set of optimal solutions can be selected from the set of non-dominated solutions to guide the training sessions. Finally, we show that the trade-off evolves in time as learning takes place and Pareto-front curves capturing the trade-off can be used to evaluate the progress of the user fairly and rigorously. We demonstrate the proposed approach through a pilot study of a virtual manual skill training task with haptic feedback administered to healthy individuals.

To the best of the authors' knowledge, Pareto optimization approaches have not been applied to HiL optimization. In this study, we apply a HiL Pareto optimization approach to characterize the trade-off between a user's performance and perceived challenge level for a motor learning task. During HiL Pareto optimization, we capture the user performance by a quantitative metric, while we model the perceived challenge level as a qualitative metric captured through preference-based qualitative feedback. We extend the Bayesian Pareto optimization approach in [41, 48] to a hybrid model with quantitative and qualitative metrics and apply it in the HiL context.

Our novel contributions can be listed as follows:

(i) We propose a HiL Pareto optimization approach with hybrid (quantitative and qualitative) performance measures to characterize the trade-off between the performance and the perceived challenge level of a motor learning task.

(ii) We demonstrate how a sample design selection can be performed over the set of non-dominated solutions characterizing the trade-off between the performance and the perceived challenge level to effectively guide training with an AAN controller.

(iii) We propose a novel and rigorous means of performance evaluation under assistance, through comparisons of the trade-off curves characterized at the different stages of training or among various users.

5.2 Human-in-the-Loop Pareto Optimization

The goal of HiL Pareto optimization is to determine optimal trade-offs for multiple conflicting objective functions while minimizing the total resource cost of experiments. Among various approaches, *multi-criteria* Bayesian optimization techniques hold promise for use in HiL applications due to their inherent sample efficiency [118], as the central idea of all Bayesian optimization approaches is to minimize the number of observations while rapidly converging to the optimal solution. Accordingly, Bayesian optimization provides a class of sample-efficient global optimization methods, where a probabilistic model conditioned on previous observations is used to determine future evaluations.

The statistical modeling in multi-criteria Bayesian optimization techniques is typically handled by one Gaussian process (GP) model for each objective to ensure the tractability of the problem, while the design of the acquisition function to capture the trade-off between multiple objectives results in various alternative techniques [38, 39, 40].

We utilize a wrapper method, called USeMO, which is based on single-objective acquisition functions [41, 48]. USeMO utilizes a multi-criteria sampling strategy that allows one to leverage acquisition functions from single-objective Bayesian optimization to solve the multi-objective Bayesian optimization problem as detailed below. We have preferred USeMO as its application to optimizations with hybrid (qualitative and quantitative) metrics is more accessible. Please note that USeMO is a sample optimization approach appropriate of HiL Pareto optimization, and alternative methods, such as [42, 43], may also be adapted for HiL Pareto optimization.

5.2.1 Virtual Motor Learning Task with Haptic Feedback

The motor learning task consists of a two-dimensional balancing game of an inverted pendulum on a cart, displayed on an LCD screen as depicted in Figure 5.1. The dynamics of the pendulum on a cart are rendered to the user through a single-axis force-feedback joystick. Two monsters, separated by constant distance, provide continual disturbances.

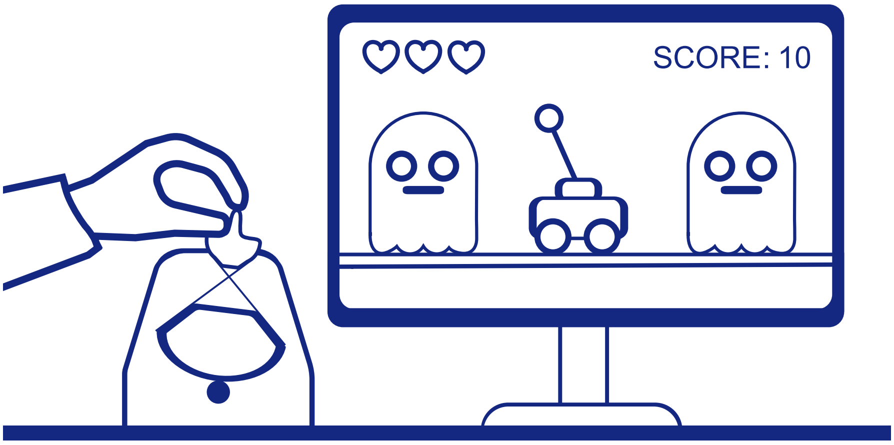

FIGURE 5.1: A sketch of a participant holding a force-feedback joystick and interacting with the pendulum game

A PD controller is implemented to determine proper cart positions to stabilize the pendulum and to move the cart to the middle of the monsters. Along with the cart and pendulum dynamics, a virtual spring is rendered between the stabilizing controller-determined position and the current position of the cart by the force-feedback joystick to provide assistance forces to the participant. The assistance level is adjusted through the stiffness constant of the virtual spring; the larger the spring constant, the more dominant the assistance provided.

Two conflicting optimization metrics are considered for the optimization: The performance metric is a quantitative measure based on the score of the participant, as detailed in Section 5.2.3. The perceived challenge level is a qualitative metric modeled utilizing ordinal classifications and pair-wise preferences provided by the participant, as detailed in Section 5.2.4. Sections 5.2.2 and 5.2.5 detail the characterization of the trade-off between the quantitative user performance and the qualitative perceived change level. It is important to emphasize that this trade-off is unique to every individual and needs to be characterized in a HiL manner.

5.2.2 Overview of HiL Multi-Criteria Bayesian Optimization

The multi-criteria Bayesian optimization approach, summarized in Algorithm 3, relies on two GP regression models: one based on the numeric game scores as detailed in Section 5.2.3, and another one based on qualitative feedback collected from the participants as detailed in Section 5.2.4. Both GP regression models possess their corresponding acquisition functions.

Algorithm 3 HiL Bayesian Pareto Optimization

```
initiate $A$: Feasible parameter space, $GP_{num}$: Quantitative GP regression model,
    $GP_{qual}$: Qualitative GP regression model, $N$: # of iterations, $N_0$: # of space-filling
    iterations
Assign acquisition function $\alpha_{num}$ to $GP_{num}$
Assign acquisition function $\alpha_{qual}$ to $GP_{qual}$
Create a sampling set $x$ with size M using Sobol sequence
for $n = 1, 2, \dots, N_0$ do
    Select parameter $x_n$ from $x$  and, use $x_n$ in the game trial
    Get the game score $s_n$, append it on s, and re-calculate y
    Get user's qualitative feedback $q_n$ and append on set $q$
Update $GP_{num}$ using $x$ and standardized score set $y$
Update $GP_{qual}$ using $x$ and qualitative feedback set $q$
for $n = N_0 + 1, N_0 + 2, \dots, N$ do
    Compute surrogate Pareto $x_p \leftarrow \mathrm{argmax}_{\mathrm{x \in A}} (\alpha_{num}, \alpha_{qual})$
    Select a parameter $x_n := \mathrm{argmax}_{\mathrm{x \in x_p}} (\sigma_{qual} \times \sigma_{num})$
    Use the parameter in the game iteration
    Get the game score $s_n$ append it on s, and re-calculate y
    Get user's qualitative feedback $q_n$ and append on set $q$
    Update $GP_{num}$ using $x$ and standardized score set $y$
    Update $GP_{qual}$ using $x$ and qualitative feedback set $q$
Compute Pareto front for the expected values of GP models
Plot the Pareto front and list the non-dominated solutions
```

During the initial iterations, parameters are selected via the Sobol sequence to ensure search-space exploration. For the rest of the iterations, new solutions are computed via the sampling strategy proposed in [41] and further detailed in Section 5.2.5.

In particular, the sampling strategy utilizes a multi-criteria optimizer to compute a surrogate Pareto front characterizing the trade-off between the conflicting acquisition functions. Then, from the set of non-dominated solutions on the surrogate Pareto front, a parameter with the highest volumetric variance is selected for the next sampling. At the end of the trade-off characterization session, Algorithm 3 utilizes a multi-criteria optimizer to compute the non-dominated solutions forming the Pareto front of the expected scores and expected perceived challenge level. Once the Pareto set is computed, promising non-dominated solutions can be selected as in Section 5.2.6 to guide the training sessions.

5.2.3 Quantitative Gaussian Process Regression Model

A GP regression model, GP_{num}, is designed to learn the relationship between the assistance level and the game score of the participant. The maximum applicable assistance during the game is represented by one, while zero represents the no assistance case. Let $A = \{x \subset \mathbb{R}^d : 0 \leq x_i \leq 1\}$ be the feasible parameter space and x_i be the assistance percentage. Let $\boldsymbol{x} = \{x_1, x_2, .., x_n\}$ be a set consisting of n assistance parameters used in prior game trials. Then, let $\boldsymbol{s} = \{s_1, s_2, .., s_n\}$ be a set consisting of n values of observed game scores corresponding to $\boldsymbol{x}$ and, let $\boldsymbol{y} = \{y_1, y_2, .., y_n\}$ be the statistically standardized version of $\boldsymbol{s}$. The dataset used to train the numerical GP regression model is represented with $\boldsymbol{D_N} = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$.

Finally, let $f_N(x)$ be the black-box function representing the relationship between assistance and standardized values of game score. To model the deviation of score measurements, we assume that standardized values of game scores are affected by Gaussian white noise with variance, σ_w^2 and we observe their noisy version y.

Then, the prior black-box function for game scores, f_N, can be modeled as

$$f_N(\boldsymbol{x}) \sim GP(0, K_N + \sigma_w^2 I) \tag{5.1}$$

with $K_N \in \mathbb{R}^{n \times n}, k_{N_{i,j}} = k_N(x_i, x_j)$, where radial basis function (RBF) kernel is used to capture the correlation of the parameters.

The game scores for unknown assistance parameters can be predicted via Bayesian inference. Let x_* be any arbitrary assistance parameters and let $f_{N*|\boldsymbol{D_N}}$ be the corresponding game score estimation based on previously acquired game scores. Then, Bayesian inference indicates

$$f_{N*|\boldsymbol{D_N}} \sim GP(\mu_{N*|\boldsymbol{D_N}}, \sigma^2_{N*|\boldsymbol{D_N}}) \tag{5.2}$$

$$\mu_{N*|\boldsymbol{D_N}} = k_{N_{*,1:n}}(K_N + \sigma_w^2 I)^{-1}\boldsymbol{y} \tag{5.3}$$

$$\sigma^2_{N*|\boldsymbol{D_N}} = k_{N_{*,*}} - k_{N_{*,1:n}}(K_N + \sigma_w^2 I)^{-1}k_{N_{1:n,*}} \tag{5.4}$$

5.2.4 Qualitative Gaussian Process Regression Model

A second GP regression model, GP_{qual}, is designed to learn the relationship between the assistance level and the perceived challenge level of the game by the participant. During the Pareto characterization session, after each game trial, the participant classifies the perceived challenge level of the game by selecting one of the following categories: *easy*, *moderate*, and *hard.* Then, except for the initialization, the participant compares the perceived challenge level of the game with that of the previous iteration. Based on the modeled probabilities of the answers, the qualitative GP regression model is updated.

The range of assistance is represented by $\boldsymbol{A} = \{x \subset \mathbb{R}^d : 0 \leq x_i \leq 1\}$, as in the quantitative GP regression model. Let $f_Q(x)$ be the latent function representing the participant's perceived challenge level of the game. The prior distribution of $f_Q(x)$ is modeled with normal GP distribution.

$$f_Q(\boldsymbol{x}) \sim GP(0, K_Q) \tag{5.5}$$

where $K_Q \in \mathbb{R}^{n \times n}, k_{Q_{i,j}} = k_Q(x_i, x_j)$ is the noiseless kernel matrix of the qualitative GP regression model generated by RBF kernel. It is worth noting that, the hyper-parameters of RBF functions used to utilize kernel matrices in both models need not be identical.

Let $\boldsymbol{q}$ be the set of qualitative feedback provided by the participant, where $\boldsymbol{q}$ consists of both ordinal classifications $\boldsymbol{q_o} = \{q_{o_1}, q_{o_1}, ..., q_{o_n}\}$ and pair-wise preferences $\boldsymbol{q_p} = \{q_{p_2}, q_{p_3}, ..., q_{p_n}\}$. Dataset including ordinal classifications defined as $\boldsymbol{D_O} = \{(x_1, q_{o1}), (x_2, q_{o2}), ..., (x_n, q_{on})\}$, and dataset including pair-wise comparisons defined as $\boldsymbol{D_P} = \{(x_1, x_2, q_{p_1}), (x_2, x_3, q_{p_2}), ..., (x_{n-1}, x_n, q_{p_n})\}$. Then, the dataset consisting of all qualitative feedback is defined as $\boldsymbol{D_Q} = \boldsymbol{D_O} \cup \boldsymbol{D_P}$.

The probability of latent function based on provided feedback is computed from the proportionality of $P(f_Q|\boldsymbol{D_Q}) \propto P(\boldsymbol{D_Q}|f_Q)P(f_Q)$, where $P(f_Q)$ is the prior unbiased probability of the regression model and $P(\boldsymbol{D_Q}|f_Q)$ is the probability of participant's all qualitative feedback being correct based on the given latent function.

Ordinal classifications and pair-wise comparisons are modeled, as in [14, 26, 28, 35], as follows: Let $\boldsymbol{O} = \{o_1, o_2, o_3\}$ be the finite set of three ordinal classifications representing the perceived challenge level from “easy" to “hard". Let t be the set of ordered thresholds used to distinguish the ordinal classifications $\boldsymbol{t} = \{t_0, t_1, t_2, t_3\}$ and $-\infty = t_{o_0} < t_{o_1} < t_{o_2} < t_{o_3} = \infty$. Then, the probability for a parameter x_i being classified with o_j^{th} ordinal class by the participant correctly is modeled with

$$P(q_{o_i} = o_j | f_Q) = \Phi\left(\frac{t_{o_j} - f_Q(x_i))}{c_o}\right) - \Phi\left(\frac{t_{oj-1} - f_Q(x_i)}{c_o}\right) \tag{5.6}$$

where Φ represents the cumulative distribution function of the Gaussian distribution and $c_o > 0$ is used to utilize the classification noise.

The probability of the participant correctly identifying the harder one between the two games is given by

$$P(q_{p_i} = (x_i \succ x_{i-1}) | f_Q) = \Phi\left(\frac{(f_Q(x_i)\ ) - f_Q(x_{i-1})\ )}{c_p}\right) \tag{5.7}$$

where $c_p > 0$ utilizes the noise in pair-wise preferences.

Under the assumption of independence of provided qualitative feedback, $P(\boldsymbol{D_Q}|f_Q) = P(\boldsymbol{D_O}|f_Q)P(\boldsymbol{D_P}|f_Q)$ is calculated by

$$P(\boldsymbol{D_Q}|f_Q) = \prod_{i=1}^{n} P(q_{o_i}|f_Q) \prod_{i=2}^{n} P(q_{p_i}|f_Q) \tag{5.8}$$

The posterior distribution of the regression model is computed by the Laplace approximation [19]. The perceived challenge level estimation of the participant $f_{Q*|\boldsymbol{D_Q}}$ for any arbitrary assistance parameter x_* can be computed using the posterior model as follows:

$$f_{Q*|\boldsymbol{D_Q}} \sim GP(\mu_{*|\boldsymbol{D_Q}}, \sigma^2_{*|\boldsymbol{D_Q}}) \tag{5.9}$$

$$\mu_{*|\boldsymbol{D_Q}} = k_{Q*,1:n}\, K_Q^{-1} \hat{f}_Q \tag{5.10}$$

$$\sigma^2_{*|\boldsymbol{D_Q}} = k_{Q**} - k_{Q*,1:n}\, (W^{-1} + K_Q)^{-1} k_{Q1:n,*} \tag{5.11}$$

where K is the noiseless covariance matrix calculated with radial basis function kernel, W is the negative Hessian matrix

$$W_{ij} = -\frac{\partial^2\, log(P(\boldsymbol{D_Q}|f(x))}{\partial f(x_i)\, \partial f(x_j)}$$

and $\hat{f}$ is the latent function that maximizes the log-likelihood

$$\hat{f} = argmax_{f(x)} \left(log(P(\boldsymbol{D_Q}|f_Q)P(f_Q)) \right)$$

5.2.5 Sampling Strategy for HiL Pareto Characterization

Let α_{num} be the acquisition function of GP_{num} and let α_{qual} be the acquisition function of GP_{qual}. In this study, for simplicity, we have used two upper confidence bound (UCB) acquisition functions with different hyperparameters. To sample an assistance level parameter x_n for $(n)^{th}$ game iteration, α_{num} and α_{qual} are utilized as

$$\alpha_{num}(x_*)_n = \mu_{*|\boldsymbol{D_{N_{n-1}}}} + \lambda_N \sigma_{*|\boldsymbol{D_{N_{n-1}}}} \tag{5.12}$$

$$\alpha_{qual}(x_*)_n = \mu_{*|\boldsymbol{D_{Q_{n-1}}}} + \lambda_Q \sigma_{*|\boldsymbol{D_{Q_{n-1}}}} \tag{5.13}$$

where λ_N and λ_Q are the hyper-parameters of GP_{num} and GP_{qual}, respectively, and x_* is an any arbitrary feasible assistance parameter. Next, based on the utilized acquisition functions, a computationally cheap Pareto problem is solved to find the parameter set x_p, lying on the surrogate Pareto front. Finally, a parameter is selected among x_p based on the maximum volumetric uncertainty, where the volumetric uncertainty is calculated by multiplying the standard deviations:

$$x_p = \mathrm{argmax}_{\mathrm{x \in A}} \left(\alpha_{num}, \alpha_{qual} \right) \tag{5.14}$$

$$x_n = \mathrm{argmax}_{\mathrm{x \in x_p}} \left(\sigma_{(x)|\boldsymbol{D_{N_{n-1}}}} \times \sigma_{(x)|\boldsymbol{D_{Q_{n-1}}}} \right) \tag{5.15}$$

5.2.6 Design Selection for Pareto-based AAN Training

All Pareto solutions characterizing the trade-off between the user performance and the perceived change level are optimal. Without additional preference information, all Pareto optimal solutions are considered equally good. The Pareto selection process imposes additional preferences to the set of Pareto optimal solutions to select among them. Pareto optimization methods allow the designer to impose additional constraints after the computation of the non-dominated solutions and inspection of the trade-off involved among the conflicting objectives. Being able to impose additional constraints to the set of optimal solutions, each corresponding to different design preferences, in a multi-shot manner, is among the most beneficial aspects of Pareto optimization approaches.

Once Pareto solutions characterizing the trade-off between the performance and the perceived challenge are computed, the designer can make a design selection to guide the training session. The Pareto selection process is not unique and alternative training methods may be devised from the same set of Pareto solutions. To the best of the authors' knowledge, this is the first study that makes use of Pareto characterizations of individuals to design AAN training protocols.

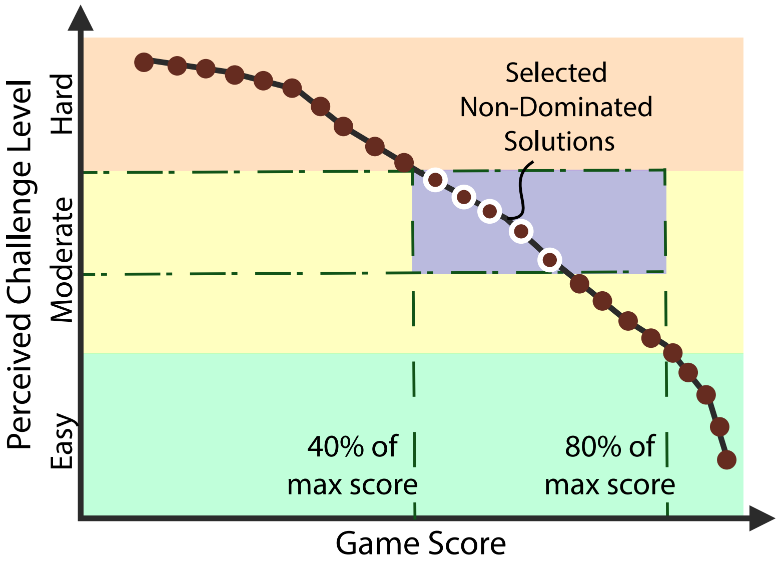

Figure 5.2: Selection of non-dominated solutions from the Pareto front by introducing design constraints after characterizing the trade-off

As a sample AAN training protocol, new constraints are imposed on the Pareto-front, as in Figure 6.4, to implement an AAN controller that keeps the task sufficiently challenging while allowing the user to achieve a sufficient level of performance. In particular, to facilitate a design selection among all non-dominated solutions, we introduce personalized thresholds to the set of non-dominated solutions to keep the perceived challenge level and quantitative performance sufficiently high. The performance is limited to 40–80% of the capacity of the individual, while the challenge level is limited to 40–80% of the perceived challenge level of the individual. All solutions lying in this range have been selected to be in the training session, to induce diversity in the training exercises. The proposed method emphasizes the importance of customization of the assistance based on individual performance characterizations. Pareto characterizations of individuals and sample results of the Pareto selection process are presented in Section 5.4.

5.3 Human Subject Experiment

5.3.1 Participants

Twenty participants, with an average age of 22.2 ± 1.3 years participated in the experiment. The participants played the game with their non-dominant hand. None of the participants had any known sensory-motor disability or significant prior experience with haptic devices. Before the experiments, participants signed an informed consent form approved by the Institutional Review Board of the university.

5.3.2 Apparatus and Game

The experimental setup consists of the virtual motor learning task with haptic feedback, as in Section 5.2.1, displayed on an LCD screen, as in Figure 5.3. A

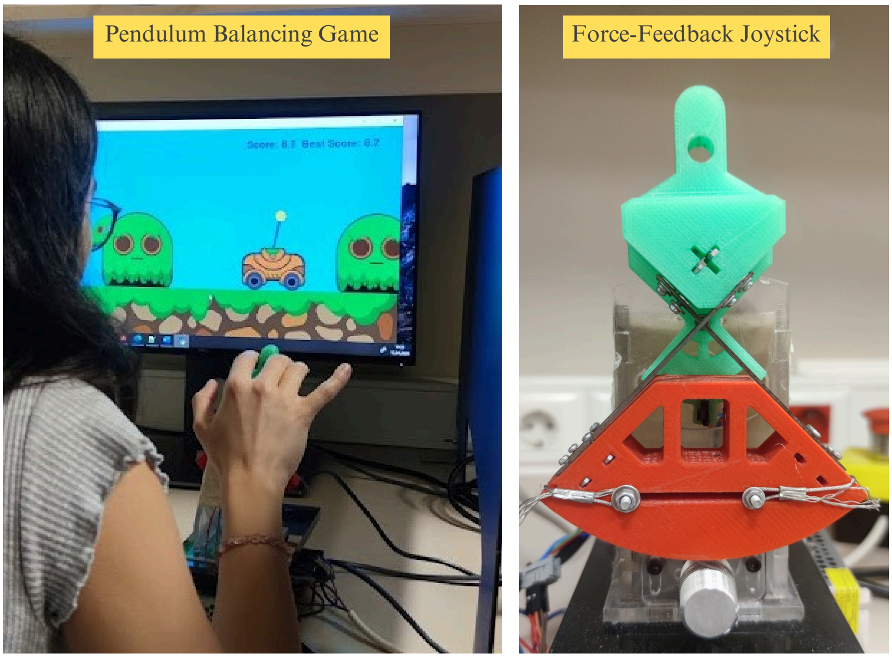

FIGURE 5.3: Pendulum balancing game and force-feedback Joystick

single-axis force-feedback joystick provides the interaction and assistance forces. The task fails if the pendulum rotates more than ±50° or the cart touches the monsters. The objective of the game is to survive for 25 s. The users have three chances for each game.

A force-feedback joystick with series elastic actuation is used as the interface for the pHRI task [5]. The force-feedback joystick is controlled under velocity-sourced impedance control [119], implemented in real-time at 1 kHz, utilizing a Texas Instruments C2000 microcontroller. The micro-controller communicates with the host computer displaying the game via serial communication.

5.3.3 Experimental Conditions

The human-subject experiment compares the training efficacy of the AAN training method based on Pareto optimization (test group) with a control group based on a commonly employed performance-based adaptive assistance controller. In particular, the following training protocols are compared:

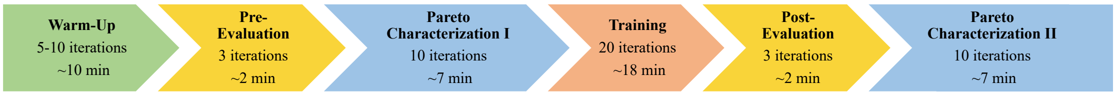

FIGURE 5.4: Experimental procedure

i) *Test Group – Pareto Approach:* As detailed in Section 5.2.6, the AAN training method based on Pareto optimization relies on HiL characterizations of the trade-off between the user performance and the perceived change level. After a trade-off is characterized, thresholds are introduced to select the subset of non-dominated solutions that span from 40% to 80% of the performance and the perceived challenge level of individual users. The set of Pareto solutions in this range are used in the training session, in a randomized order.

ii) *Control Group – Adaptive Assistance:* The control group is based on the commonly used approach of measuring progress based on the level of assistance provided. A performance-based adaptive assistance controller is implemented using an adaptive staircase approach, as in [109, 120, 121]. The assistance level of the control group is initialized at 50% and employs a two-up-one-down variation. In this adaptive approach, the assistance decreases by 10% after every two consecutive successful game trials or increases by 10% after each failed trial.

5.3.4 Experimental Procedure

The experiment consists of six sessions as in Figure 5.4: warm-up, pre-evaluation, pre-training HiL characterization, training, post-evaluation, and post-training HiL characterization.

The warm-up session includes a tutorial to help participants become familiar with the rules and mechanics of the game. Participants play the game with several levels of assistance. The warm-up session is concluded when a participant displays a clear understanding of the rules and can play the game without unpremeditated failure.

Pre- and post-evaluations measure the performance of the participants with no assistance. Participants with very high scores in the pre-evaluation session are excluded from the experiment, as their performance has already saturated. A comparison of pre- versus post-evaluation sessions provides a measure of participants' progress after the training session.

HiL Pareto characterization sessions are used to learn the trade-off between a participant's quantitative performance and the perceived challenge level of the game. For this purpose, two GP regression models are trained in each HiL session. During each iteration of the HiL learning, the participant plays the game with an assistance level determined by the optimization algorithm. Once the game ends, either by surviving for 25 s or failing to do so three times, the best score of the participant is used to train a quantitative GP regression model. After each game, two questions are asked to the participant to determine the perceived difficulty of the game. First, the participant is asked to rate the challenge level of the game by selecting one option from "hard", "moderate", or "easy", resulting in an ordinal classification. Next, (except for the first trial) the participant is asked to compare the challenge level of the last game with the previous one in a pair-wise manner. The answers to these queries are used to train a qualitative GP regression model. Finally, utilizing a multi-criteria sampling strategy, the HiL optimization algorithm updates the assistance level provided to the participant for the next iteration.

5.3.5 Hypotheses

We aim to test the validity of the following hypotheses:

H_1 The trade-off between the performance and the perceived challenge level of a task can be characterized by utilizing a HiL Pareto optimization approach.

H_2 The non-dominated solutions characterizing the trade-off between the performance and the perceived challenge can be used to guide training sessions with an AAN controller to match or improve the efficacy of training compared to performance-based adaptive assistance approaches.

H_3 Comparisons of the trade-off curves characterized at the different stages of training provide a rigorous means of evaluating training performance under assistance.

5.4 Results and Discussion

First, we check the overall learning of the participants after the training session. To this end, we compare the unassisted game scores of the participants before and after the training sessions. Figure 5.5 presents box plots of participants' game scores before and after the training sessions. Paired t-tests within the test and control groups indicate that there exists a statistically significant difference in scores with $t_{29} = 3.32$ and $p = 0.0012$ for the test group, and $t_{29} = 3.27$ and $p = 0.0014$ for the control group. The effect size is notable, as the post-training performances of test and control groups have improved by 37.5% and 36.7%, respectively, compared to their pre-training performances, with a confidence interval exceeding 99%. Moreover, unpaired t-tests between test and control groups indicate there exists no statistically significant difference between their pre- and post-training performances, respectively. Accordingly, the test and control groups display equal unassisted task performance before and after the training.

To further investigate the differences between the test and control groups, we utilize the advantages of the proposed Pareto approach to study the mean improvement in available performance for each group *under assistance*. Figure 5.6

presents the aggregate training efficacy for the test and control groups, by reporting mean percent changes in available performance under all assistance levels. This performance metric is defined as $100(Score_{post} - Score_{pre})/(1 - Score_{pre})$ to ensure that the saturation under high assistance levels is accounted.

Figure 5.6 indicates that the Pareto approach has achieved higher mean performance gains, especially within the 40%–90% assistance range, as most of the participants in the Pareto training group were trained with lower assistance.

Figure 5.7 presents the assistance level provided to the control group in comparison to the assistance level that would have been provided if these participants were in the test group. Figure 5.7 makes clear that most of the participants in the test group were provided with assistance levels higher than 55%, while the Pareto optimal assistance level for the test group was around 40%. Accordingly, the improvement of the available performance of the control group is lower than the test group for assistance levels exceeding 40%, as they have been trained with significantly higher assistance levels that were less challenging.

While there exists no statistically significant performance difference between the

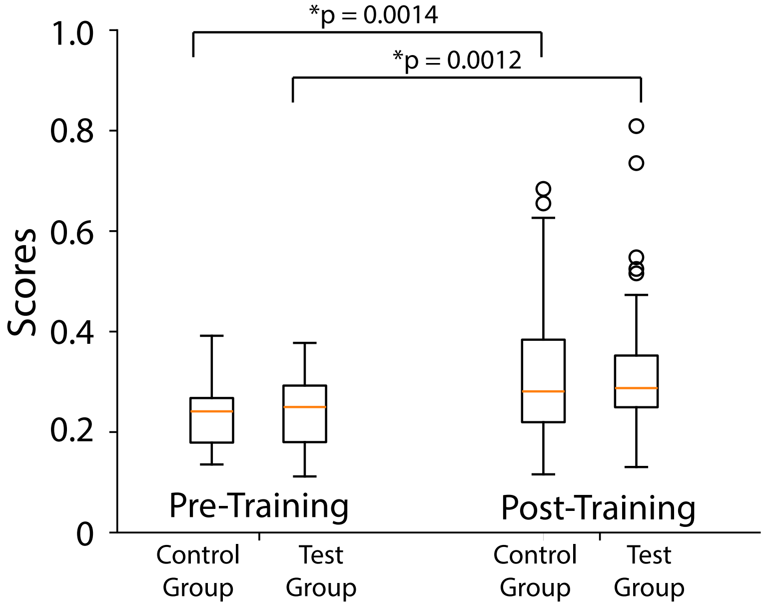

Figure 5.5: Comparison of unassisted pre- and post-training performance

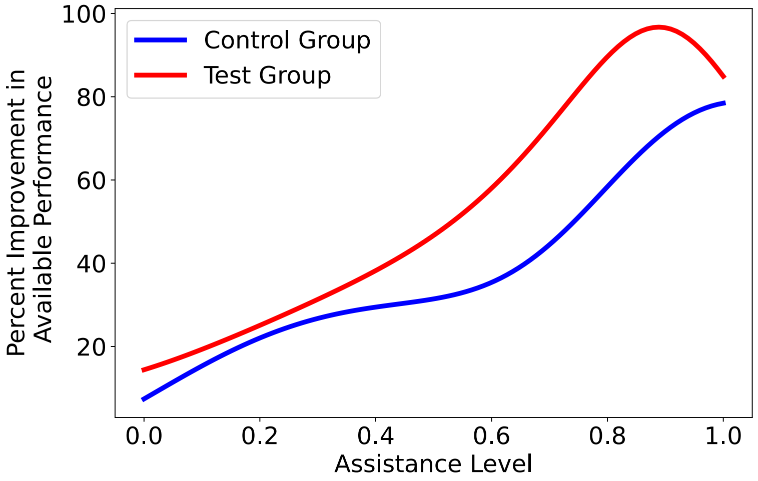

FIGURE 5.6: Percent improvement of available performance

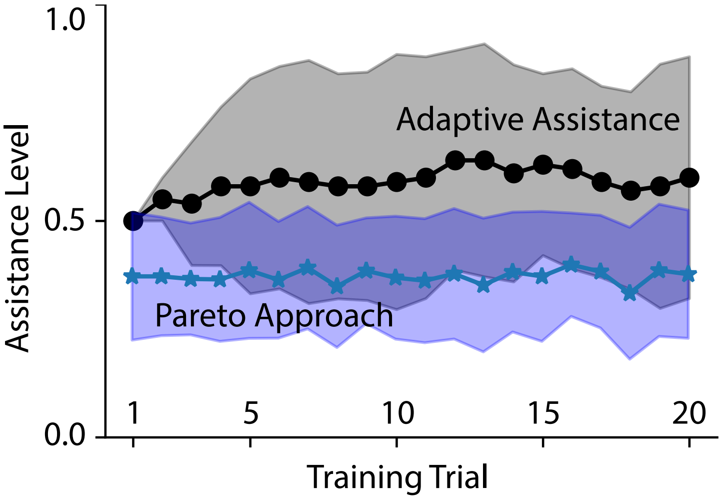

FIGURE 5.7: Assistance provided to participants in the control group versus assistance that would have been provided if they were in the test group.

groups under no assistance, possibly due to the limited number of training sessions, the results in Figure 5.6 indicate that the Pareto approach has resulted in better overall improvement at all assistance levels. Accordingly, there is experimental evidence that H_2 is valid, as utilizing the non-dominated solutions can provide training with high efficacy.

While aggregate results are useful for discussing the overall efficacy of the training protocols, the proposed Pareto approach emphasizes customization of training based on individual performances. Figure 5.8 presents the results of two sample participants with high and intermediate performance, while the results of all participants are provided in the Supplementary Document [122]. Figure 5.8

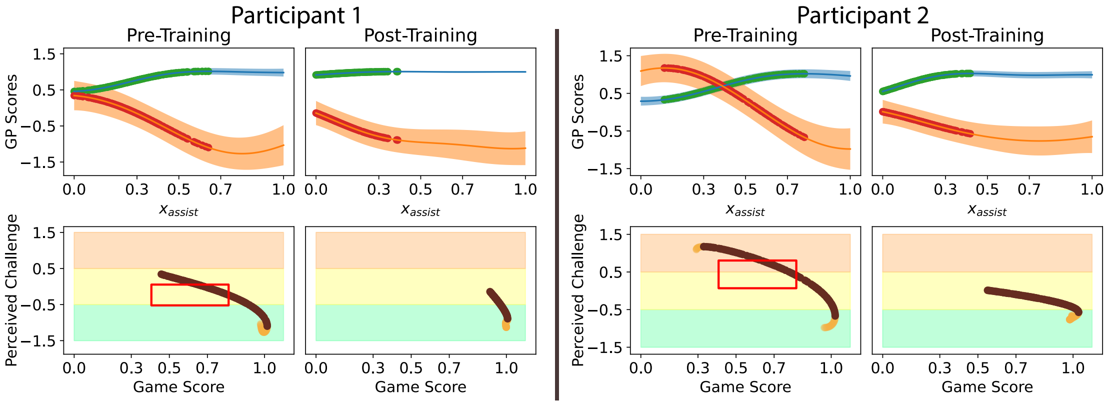

FIGURE 5.8: In the top rows, the orange and blue lines show the mean of the surrogate function for the perceived challenge and score, respectively, while the shading depicts their standard deviation. The bottom rows present the Pareto solutions are shown with dark red dots, while the dominated solutions in the feasible set are also shown with orange dots. The Pareto solutions within the red squares are used during the training of the participants in the test group.

presents the GP regression models and the Pareto fronts, before and after the training. In the top rows, the orange lines show the mean of the surrogate function for the perceived challenge level, while the orange shading depicts its standard deviation. Similarly, the blue lines represent the mean prediction for the game scores, while the blue shading depicts its standard deviation. The bottom rows present the Pareto plots characterized during pre- and post-training sessions, depicted in the left and right columns.

In the Pareto plots, the black points represent Pareto solutions, while the orange points depict all feasible solutions for the problem. Pareto solutions cover a non-trivial (and possibly disconnected as in Figure 5.10) subset of the set of feasible solutions, as they capture the non-dominated solutions of the trade-off between the expectations of the perceived challenge level and the score. For ease of visualization, the Pareto plots are divided into three regions; the red, yellow, and green shades represent the perceived challenge levels of hard, moderate, and easy, respectively. Finally, the red squares are used to indicate the Pareto solutions that have been used during the training of the participants in the test group, according to the Pareto selection criteria detailed in Section 5.2.6.

As hypothesized in H_1, Figure 5.8 provides evidence that the trade-off between the performance and the perceived challenge level of a task can be characterized via HiL Pareto optimization with hybrid performance measures. Thanks to the sample-efficiency inherited from the underlying Bayesian optimization approach [41], HiL Pareto characterization converged in 6.5 ± 1.4 min; hence, can easily be performed before the training sessions without inducing fatigue to the participants. The top rows in Figure 5.8 verify that the Bayesian multi-criteria optimization technique samples from the regions of surrogate functions that conflict with each other. The efficiency of the sampling strategy in locating the conflicting regions and the uniformity of samples in these regions are key aspects of the efficiency of the Pareto optimization technique which makes its use adequate for HiL optimization.

The bottom rows in Figure 5.8 properly characterize the trade-off between the performance and the perceived challenge level of a task. In particular, in the sample Pareto plots, both participants can achieve their best scores where their perceived challenge level is low and their performance decreases as their perceived challenge level increases. Furthermore, the trade-off curves also allow us to evaluate if the game is too challenging or easy for the participant, by checking their performance under high and low assistance levels, respectively. In such cases, it may be preferable to change the game to better cater to the abilities of the participant.

As hypothesized in H_3, comparisons of the trade-off curves characterized at the different stages of training provide a rigorous means of evaluating training performance. In the bottom rows of Figure 5.8, the changes between the pre-and post-training trade-off characterization plots capture the progression of the participants under assistance. For Participants 1 and 2, the shifts in the Pareto plots explicitly indicate that participants not only perceive the game as less challenging after the training, but also their scores improve significantly. Accordingly, their post-training Pareto curves have shifted towards right and downwards, compared to their pre-training Pareto curves.

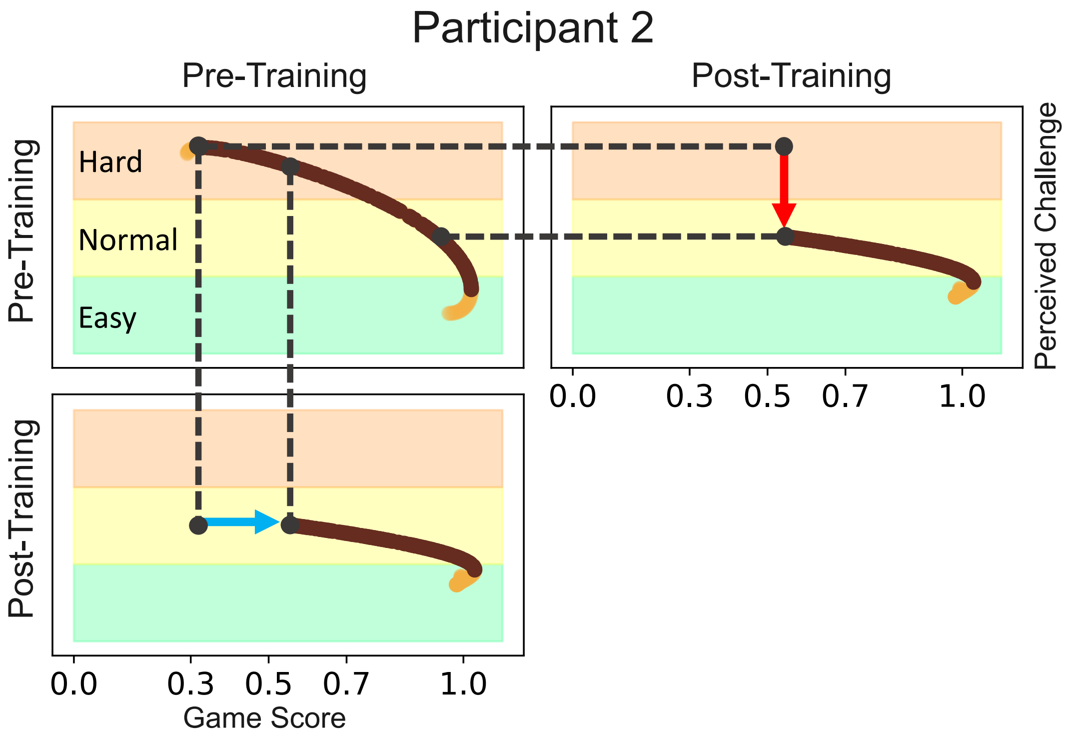

FIGURE 5.9: The shift indicated with the blue arrow shows that the participant's performance improved after training while the shift indicated with the red arrow shows that the participant perceives the game as less difficult.

For instance, Figure 5.9 shows that for Participant 2, the perceived level of challenge for the most difficult game has reduced from hard to moderate, indicating that the game became easier to play as learning took place. The positive progress in the performance can also be observed by the shrinkage of the Pareto front towards the right side, which captures the region for higher game scores. This shift indicates the game scores for the most difficult perceived challenge level have increased from 0.30 to 0.55 for Participant 2.

These improvements can also be observed in the GP regression models, before and after the training. For instance, one can observe by inspecting the surrogate function for the perceived challenge level and game scores of Participant 2 that the performance and challenge saturates around 70% assistance pre-training, while this saturation shifts to 45% assistance after the training. Hence, only the assistance levels from 0% to 40% belong to the post-training Pareto solutions.

Comparison of the pre-and post-training surrogate functions is especially useful to understand Pareto plots that consist of multiple disconnected sections.

Figure 5.10 presents the results for a sample of low-performing Participant 3, whose Pareto plots are harder to interpret. The surrogate functions for the perceived challenge level and game scores help to understand the results. While the worst performance of Participant 3 did not significantly improve from pre-training to post-training, the shifts in the surrogate functions of this participant indicate that Participant 3 requires less assistance to achieve a similar level of performance as the pre-training case. In particular, the increase in performance and the decrease in challenge level shifts from 70% assistance at the pre-training characterization to 50% assistance at the post-training characterization. Accordingly, the trade-off characterization captures improvements in the performance that cannot be captured by only observing the participants' performance without any assistance.

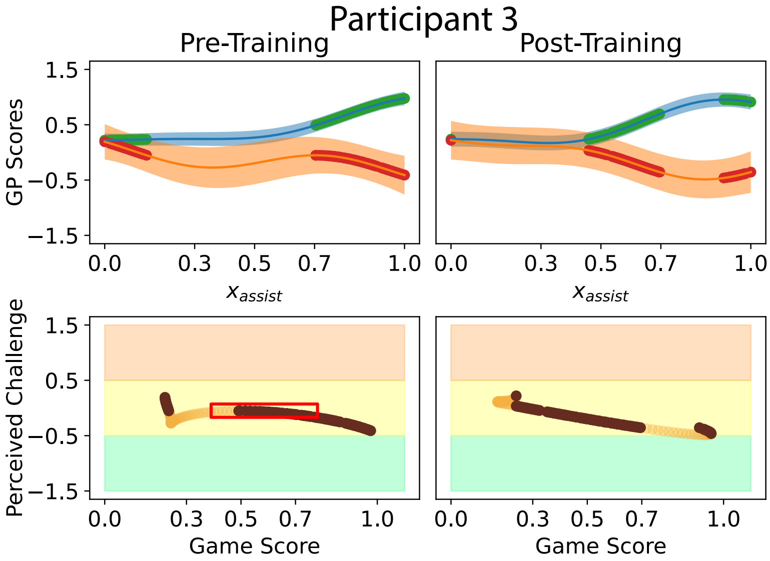

FIGURE 5.10: Results for a sample low-performing participant

Finally, the trade-off curves of different participants can also be compared with each other. For instance, the post-training Pareto front of Participant 2 is slightly better than the pre-training Pareto front of Participant 1, indicating that Participant 2 reaches a more advanced stage after the training, compared to the pre-training performance of Participant 1. Note that rigorous comparisons of different participants with each other, in general, is a challenging task, and a fair comparison of different participants is only possible by considering the non-dominated solutions, as they capture the best possible performances of each challenge level for all preferences.

Overall, this study demonstrates the feasibility of HiL Pareto optimization, its usefulness in improving the efficacy of training performance of AAN controllers, and its novel uses for performance evaluation and comparisons. Pilot experiments are provided with one qualitative and one quantitative cost function to demonstrate the use of two different forms of feedback mechanisms in HiL optimization.

Finally, it is important to emphasize that the proposed approach is generic and can be easily extended to Pareto solutions for any number and type of cost function. Similarly, while we have utilized USeMO as an efficient Pareto optimizer, similar results may be achieved with other Pareto optimization approaches. Besides, while this pilot study has been conducted for the single decision variable of *assistance level* to promote ease of presentation, the proposed Pareto optimization method trivially extends to a larger number of decision variables. Finally, the proposed Pareto training approach is only a sample design that utilizes Pareto solutions for a training protocol. More sophisticated approaches, for instance, an adaptive protocol that relies on instantaneous user performance to select the most appropriate Pareto solution on the fly, may perform better in terms of training efficacy. The determination and user evaluation of such alternative approaches are left as future work.

Chapter 6

A Generalized Transfer Learning Framework for Bayesian Optimization with Applications to Robot-Assisted Rehabilitation

Bayesian optimization (BO) is widely recognized for its effectiveness in personalizing physical human-robot interaction (pHRI). Numerous studies have demonstrated its efficiency in optimizing various aspects of these interactions, particularly in tasks requiring precise adjustments to meet individual needs. Typically, these studies have focused on optimizing single tasks, such as fine-tuning walking gait parameters for lower limb exoskeletons [11, 15, 28, 34].

However, the scope of interactive robotics often involves multiple distinct tasks, each requiring optimization. Conducting BO for each task independently, without leveraging prior knowledge, is not only inefficient but also impractical. To address this challenge, integrating prior knowledge or employing multi-task optimization strategies becomes essential to enhance the practicality and scalability of BO in complex, multi-task pHRI applications.

In this chapter, we first develop a generalizable transfer learning methodology for Gaussian Process (GP)-based models. This approach is designed to increase the effectiveness and practicality of human-in-the-loop (HiL) optimization strategies, particularly in applications where data efficiency and adaptability are crucial.

In our transfer learning methodology, we first focused on transferring from one GP model to another with fully Bayesian-based computation techniques to find the exact posterior. While doing it, we also divided equations such that different GP modeling strategies can be compatible without a need for developing a specific methodology to combine different GP modeling techniques such that one can transfer from a GP model with approximated posterior such as qualitative GP models to a GP model with exact posterior results.

Then, we extend this methodology to generalize to allow transfer from multiple GP models without adding heavy computational burdens. This extension significantly increased the practicality of our method for pHRI application where many offline HiL experiment results can be correlated with an online one.

We demonstrate the application of our transfer learning-based HiL optimization strategy in personalizing robot-assisted rehabilitation sessions. By transferring data from prior tasks, we accelerate the optimization process, allowing for more efficient personalization of assistance levels during rehabilitation.

Our novel contributions are listed below.

(i) We developed a generalizable transfer learning framework for Bayesian optimization to increase the practicality of HiL studies and their applications in pHRI that involve various tasks.

(ii) We propose a systematic and practical means of using Bayesian optimization in the personalization of robot-assisted rehabilitation sessions by implementing our transfer learning framework.

(iii) We developed a systematic methodology for tracking the patients' progress and comparing their progress with each other as well as among various physical tasks.

6.1 Generalized Transfer Learning Framework for GP Models

We develop a transfer learning strategy by defining a source GP model GP_s from which information is transferred and a target GP model GP_t to which the information is transferred. Each GP model has its own dataset and latent functions. The source dataset $\boldsymbol{D_s}$ includes parameters $\boldsymbol{x_s}$ and their corresponding outcomes, either numerical measurements $\boldsymbol{y_s}$ or qualitative evaluations $\boldsymbol{q_s}$. Similarly, the target dataset $\boldsymbol{D_t}$ includes parameters $\boldsymbol{x_t}$ and their outcomes $\boldsymbol{y_t}$ or $\boldsymbol{q_t}$. We define the latent function values for the source model as $\boldsymbol{f_s}$ and for the target model as $\boldsymbol{f_t}$.

To enable information transfer, we combine the source and target GP models into a multi-task GP framework, which includes latent function values $\boldsymbol{f_s}$, $\boldsymbol{f_t}$, and target latent function value $\boldsymbol{f_{t_*}}$ corresponding to any feasible parameter $\boldsymbol{x_{t_*}}$. While information transfer can occur between any GP models in a multi-task GP model, our focus is to find the posterior distribution of the target latent function value $\boldsymbol{f_{t_*|D_t,D_s}}$.

We apply Bayesian inference to conditioning source and target dataset with two steps to avoid the computational burden of calculating all information transfer between GP models in the multi-task GP model and produce a more generalizable approach. We first apply Bayesian inference techniques to transfer information from the source dataset and apply again to transfer information from the target dataset to make predictions for $\boldsymbol{f_{t_*}}$ by utilizing both datasets.

Finally, we extend our approach by merging it with a modified version of the Bayesian Committee Machine (BCM) [58], an approximation method for GP models with large data sets, to effectively compute when the source GP model includes various GP modeling techniques or consists of a large data set.

During the derivations, we model prior expectations of GP_t and GP_s as $\mathbf{0}$ for simplifying equations but one can extend our transfer methodology easily to GP models with non-zero prior.

6.1.1 Multi-Task Gaussian Process

The source GP model GP_s can consist of a single GP model or multiple GP models in the form of multi-tasking GP as shown in Eqn. (6.1) and Eqn. (6.2).

$$\boldsymbol{f_s} = [\boldsymbol{f_{s_1}}, \boldsymbol{f_{s_2}}, ..., \boldsymbol{f_{s_M}}] \sim GP(\mathbf{0}, Var(\boldsymbol{f_s})) \tag{6.1}$$

$$Var(\boldsymbol{f_s}) = \boldsymbol{K_{ss}} = \begin{bmatrix} \boldsymbol{K_{s_1,s_1}} & \cdots & \boldsymbol{K_{s_1,s_M}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{K_{s_M,s_1}} & \cdots & \boldsymbol{K_{s_M,s_M}} \end{bmatrix} \tag{6.2}$$

Let $\boldsymbol{f_s}$ represent the latent function values for parameters $\boldsymbol{x_s}$ used in training the $\boldsymbol{GP_s}$ and let $\boldsymbol{f_t}$ denote the latent function for the parameters $\boldsymbol{x_t}$ of GP_t. Finally, let $\boldsymbol{f_{t*}}$ be a latent function value for arbitrary parameter $\boldsymbol{x_{t*}}$ in GP_t. Then all latent functions can modeled in a single multivariate Gaussian distribution as shown below.

$$\begin{bmatrix} \boldsymbol{f_s} \\ \boldsymbol{f_t} \\ \boldsymbol{f_{t*}} \end{bmatrix} \sim GP\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{K_{ss}} & \boldsymbol{K_{st}} & \boldsymbol{K_{st*}} \\ \boldsymbol{K_{ts}} & \boldsymbol{K_{tt}} & \boldsymbol{K_{tt*}} \\ \boldsymbol{K_{t*s}} & \boldsymbol{K_{t*t}} & \boldsymbol{K_{t*t*}} \end{bmatrix} \right) \tag{6.3}$$

Variables in Eqn. (6.3) are defined below.

- $\boldsymbol{K_{ss}}$ is the kernel matrix between points of source GP's

- $\boldsymbol{K_{tt}}$ is the kernel matrix between points of transferred GPs
- $\boldsymbol{K_{st}}$ is the kernel matrix between points of source and transferred GP
- $\boldsymbol{K_{t_*t_*}}$ is the kernel matrix between arbitrary points of source and transferred GP
- $\boldsymbol{K_{tt_*}}$ is the kernel matrix between points and arbitrary points of source and transferred GP
- $\boldsymbol{K_{st_*}}$ is the kernel matrix between points of source GP and arbitrary points of source and transferred GP

We provide two methods for calculating the kernel matrix for the multi-tasking GP model. We have used the intrinsic coregionalization model for simplicity.

Intrinsic Coregionalization Model In the ICM approach, a single kernel is used for making kernel matrices between different GPs (tasks). For Gaussian processes i and j,

$$k(x, i), (x', j)) = w_{ij} * k(x, x') \tag{6.4}$$

w_{ij} is the correlation coefficient between Gaussian process i and j. It is important to select w_{ij} such that the overall multi-tasking matrix should be positive semi-definite.

As the ICM approach requires the utilizing a single kernel function for all GP models, it can limit the usage of the methodology. One can choose to use a linear coregionalization model as a generalized version of ICM to use different types of kernel functions to build a multi-tasking GP model.

Linear Coregionalization Model The LCM approach uses a linear weighted combination of multiple kernels to make kernel matrices between different GPs

(Tasks). For Gaussian processes i and j,

$$k((x,i),(x',j)) = \sum_{k=1}^{n_{GP}} w_{ij_k} * k(x,x') \tag{6.5}$$

w_{ij_k} is the correlation coefficient between Gaussian process i and j for the k^{th} kernel. This approach allows transfer between GP models employed by different kernel functions.

6.1.2 Data Transferred From a Source to a Target GP Model

In this section, we demonstrate how we apply Bayesian inference in two-step to conditioning datasets to find a posterior distribution of $\boldsymbol{f_{t*|D_s,D_t}}$.

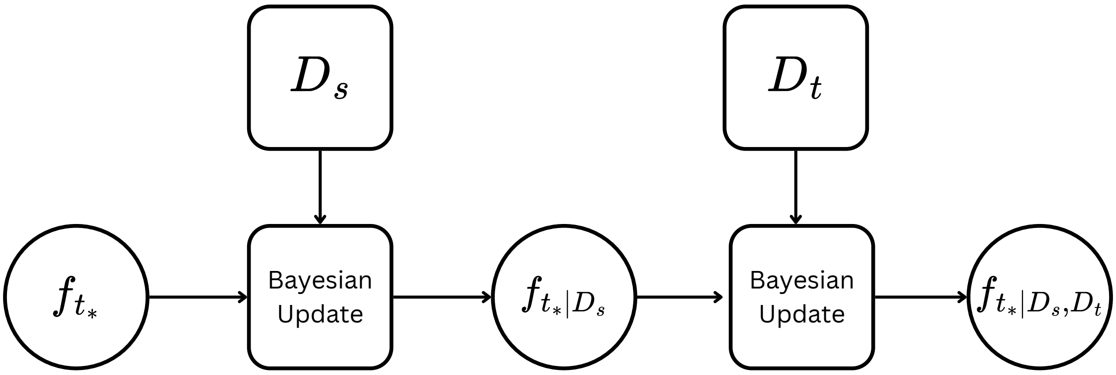

FIGURE 6.1: Data transfer from a source GP model to a target GP model

We first evaluate posterior distributions for source latent function values $\boldsymbol{f_s|D_s}$ using the source dataset $\boldsymbol{D_s}$. Then we use this posterior distribution to condition source data on target latent function values to get new probability distributions $\boldsymbol{f_t|D_s}$ and $\boldsymbol{f_{t*|D_s}}$. Hence, we transfer all information to the target GP model. Finally, we evaluate the posterior distribution of $\boldsymbol{f_{t|D_s,D_t}}$ and use it for conditioning the target dataset on $\boldsymbol{f_{t*|D_s}}$ to find a posterior distribution of $\boldsymbol{f_{t*|D_s,D_t}}$ which uses both information of source and target.

This approach is similar to the projected process approximation [19, 57], which is often used to speed up the computation of exact posterior predictions in GP models. The key difference is that our method calculates the posterior exactly, whereas the projected process approximation predicts a new distribution for the target latent function values only for the parameters in the target dataset, $\boldsymbol{f_{t|D_s}}$. As a result, in the projected process approximation, the posterior predictions are less affected by the source dataset when $\boldsymbol{x_{t_*}}$ is distant from the parameters in $\boldsymbol{x_t}$ [19].

Normally, the value $\boldsymbol{f_{t*|D_s,D_t}}$ is determined by applying Bayesian inference in a single step, conditioning on both the source and target datasets using a multi-task GP model. Both the single-step and two-step approaches calculate the exact posterior. However, while two-step conditioning doesn't increase computational efficiency for transferring between GP models with quantitative data, it significantly increases efficiency for GP models where posterior calculations are computationally expensive.

Additionally, the two-step method allows for calculating the posteriors of both the source and target models using independent techniques, making it possible to use different types of GP models within the same transfer process. Consequently, two-step conditioning allows for transferring between GP models that employ different strategies for approximating their posterior distributions. For instance, one model might use expectation propagation while another uses Laplace approximation, or the transfer could occur from a GP model that calculates its posterior directly to another that uses Laplace approximation.

6.1.2.1 Data Transferred Prior GP Model

We predict a prior distribution for a GP_t using the posterior model of GP_s and its data $\boldsymbol{D_s}$. Let inferred version of $\boldsymbol{f_t}$ and $\boldsymbol{f_{t_*}}$ according to the $\boldsymbol{D_s}$ represented as below.

$$\begin{bmatrix} \boldsymbol{f_{t|D_s}} \\ \boldsymbol{f_{t*|D_s}} \end{bmatrix} \sim GP\left(\begin{bmatrix} E(\boldsymbol{f_{t|D_s}}) \\ E(\boldsymbol{f_{t*|D_s}}) \end{bmatrix}, \begin{bmatrix} Var(\boldsymbol{f_{t|D_s}}) & Cov(\boldsymbol{f_{t|D_s}}, \boldsymbol{f_{t*|D_s}}) \\ Cov(\boldsymbol{f_{t*|D_s}}, \boldsymbol{f_{t|D_s}}) & Var(\boldsymbol{f_{t*|D_s}}) \end{bmatrix}\right) \tag{6.6}$$

Variables in Eqn. (6.6) are derived as below.

$$\begin{aligned}
&E(\boldsymbol{f_{t|D_s}}) = \boldsymbol{K_{ts}}\boldsymbol{K_{ss}^{-1}}E(\boldsymbol{f_{s|D_s}}) \\
&E(\boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{t*s}}\boldsymbol{K_{ss}^{-1}}E(\boldsymbol{f_{s|D_s}}) \\
&Var(\boldsymbol{f_{t|D_s}}) = \boldsymbol{K_{tt|D_s}} = \boldsymbol{K_{tt}} - \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}^{-1}} - \boldsymbol{K_{ss}^{-1}}Var(\boldsymbol{f_{s|D_s}})\boldsymbol{K_{ss}^{-1}})\boldsymbol{K_{st}} \\
&Var(\boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{t*t*|D_s}} = \boldsymbol{K_{t*t*}} - \boldsymbol{K_{t*s}}(\boldsymbol{K_{ss}^{-1}} - \boldsymbol{K_{ss}^{-1}}Var(\boldsymbol{f_{s|D_s}})\boldsymbol{K_{ss}^{-1}})\boldsymbol{K_{st*}} \\
&Cov(\boldsymbol{f_{t|D_s}}, \boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{tt*|D_s}} = \boldsymbol{K_{tt*}} - \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}^{-1}} - \boldsymbol{K_{ss}^{-1}}Var(\boldsymbol{f_{s|D_s}})\boldsymbol{K_{ss}^{-1}})\boldsymbol{K_{st*}}
\end{aligned}$$

This way, the transfer is simplified as creating a GP model with non-zero mean values. One can apply the calculations of $E(\boldsymbol{f_{t*|D_s}})$ and $Cov(\boldsymbol{f_{t|D_s}}, \boldsymbol{f_{t*|D_s}})$ for finding prior mean and kernel function values necessary for creating the GP model.

Transferring Data from Quantitative GP models For a quantitative source GP model, $E(\boldsymbol{f_{s|D_s}})$ and $Var(\boldsymbol{f_{s|D_s}})$ calculated as shown below.

$$\begin{aligned}
&E(\boldsymbol{f_{s|D_s}}) = \boldsymbol{K_{ss}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{y_s} \\
&Var(\boldsymbol{f_{s|D_s}}) = \boldsymbol{K_{ss}} - (\boldsymbol{K_{ss}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{K_{ss}}) = \sigma_{ws}^2\boldsymbol{K_{ss}}((\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1})
\end{aligned}$$

Then the transferred GP prior terms in Eqn. (6.6) reduces to

$$
\begin{aligned}
&E(\boldsymbol{f_{t|D_s}}) = \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{y_s}\\
&E(\boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{t*s}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{y_s}\\
&Var(\boldsymbol{f_{t|D_s}}) = \boldsymbol{K_{tt|D_s}} = \boldsymbol{K_{tt}} - \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{K_{st}}\\
&Var(\boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{t*t*|D_s}} = \boldsymbol{K_{t*t*}} - \boldsymbol{K_{t*s}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{K_{st*}}\\
&Cov(\boldsymbol{f_{t|D_s}}, \boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{tt*|D_s}} = \boldsymbol{K_{tt*}} - \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}} + \sigma_{ws}^2\boldsymbol{I})^{-1}\boldsymbol{K_{st*}}
\end{aligned}
$$

Transferring Data from Qualitative GP models If the source GP is a qualitative GP model and calculated by Laplace approximation, then $E(\boldsymbol{f_{s|D_s}})$ and $Var(\boldsymbol{f_{s|D_s}})$ calculated according to 3.3.

$$
\begin{aligned}
&E(\boldsymbol{f_{s|D_s}}) = \boldsymbol{\hat{f}_s} = \text{argmax}_{\boldsymbol{f_s}}\,(log(P(\boldsymbol{D_s}|\boldsymbol{f_s})P(\boldsymbol{f_s})))\\
&Var(\boldsymbol{f_{s|D_s}}) = (\boldsymbol{W_s} + \boldsymbol{K_{ss}^{-1}})^{-1}
\end{aligned}
$$

W_s denotes the negative hessian of $log(P(\boldsymbol{D_s}|\boldsymbol{f_s})P(\boldsymbol{f_s}))$. Accordingly, the transferred GP prior terms in Eqn. (6.6) reduces to

$$
\begin{aligned}
&E(\boldsymbol{f_{t|D_s}}) = \boldsymbol{K_{ts}}\boldsymbol{K_{ss}^{-1}}\boldsymbol{\hat{f}_s}\\
&E(\boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{t*s}}\boldsymbol{K_{ss}^{-1}}\boldsymbol{\hat{f}_s}\\
&Var(\boldsymbol{f_{t|D_s}}) = \boldsymbol{K_{tt|D_s}} = \boldsymbol{K_{tt}} - \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}} + \boldsymbol{W_s^{-1}})^{-1}\boldsymbol{K_{st}}\\
&Var(\boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{t*t*|D_s}} = \boldsymbol{K_{t*t*}} - \boldsymbol{K_{t*s}}(\boldsymbol{K_{ss}} + \boldsymbol{W_s^{-1}})^{-1}\boldsymbol{K_{st*}}\\
&Cov(\boldsymbol{f_{t|D_s}}, \boldsymbol{f_{t*|D_s}}) = \boldsymbol{K_{tt*|D_s}} = \boldsymbol{K_{tt*}} - \boldsymbol{K_{ts}}(\boldsymbol{K_{ss}} + \boldsymbol{W_s^{-1}})^{-1}\boldsymbol{K_{st*}}
\end{aligned}
$$

6.1.2.2 Bayesian Inference for GP Models After Data Transfer

Based on the transferred prior GP model, posterior inference can be conducted similarly as shown in 3.2. The goal of Bayesian Inference is to estimate the posterior distribution for the latent function $\boldsymbol{f_{t*|D_s,D_t}}$ for a given assistance level, $\boldsymbol{x_*}$ based on source GP data, $\boldsymbol{D_s}$ and transferred GP data $\boldsymbol{D_t}$.

The posterior distribution of $\boldsymbol{f_{t*|D_s,D_t}}$ shown as

$$\boldsymbol{f_{t*|D_s,D_t}} \sim GP(E(\boldsymbol{f_{t*|D_s,D_t}}), Var(\boldsymbol{f_{t*|D_s,D_t}})) \tag{6.7}$$

The generalized version of expectation and variance are given below.

$$E(\boldsymbol{f_{t*|D_s,D_t}}) = E(\boldsymbol{f_{t*|D_s}}) + \boldsymbol{K_{tt*|D_s}} \boldsymbol{K_{tt|D_s}^{-1}} \left(E(\boldsymbol{f_{t|D_s,D_t}}) - E(\boldsymbol{f_{t|D_s}})\right) \tag{6.8}$$

$$Var(\boldsymbol{f_{t*|D_s,D_t}}) = Var(\boldsymbol{f_{t*|f_t,D_s,D_t}}) + \boldsymbol{K_{tt*}} \boldsymbol{K_{tt}^{-1}} Var(\boldsymbol{f_{t|D_s,D_t}}) \boldsymbol{K_{tt}^{-1}} \boldsymbol{K_{tt*}} \tag{6.9}$$

$$Var(\boldsymbol{f_{t*|f_t,D_s,D_t}}) = \boldsymbol{K_{t*t*|D_s}} - \boldsymbol{K_{t*t|D_s}} \boldsymbol{K_{tt|D_s}^{-1}} \boldsymbol{K_{tt*|D_s}} \tag{6.10}$$

Equations 6.8 and 6.9 apply to all GP model types. Based on these equations one can estimate latent function values for new points based on both data of previous GP models and data of the current model.

6.1.2.3 Bayesian Inference for a Quantitative GP Models After Data Transfer

The posterior of the quantitative GP model after data transfer is computed according to equations in 3.2 by using transfer augmented version of kernel matrix components and prior mean values.

$$E(\boldsymbol{f_{t|D_s,D_t}}) = E(\boldsymbol{f_{t|D_s}}) + \boldsymbol{K_{tt|D_s}} (\boldsymbol{K_{tt|D_s}} + \sigma_{wt} \boldsymbol{I})^{-1} (\boldsymbol{y_t} - E(\boldsymbol{f_{t|D_s}})) \tag{6.11}$$

Where $\boldsymbol{y_t}$ is the observed numeric data and σ_{wt} is the standard deviation corresponding to noisy numeric data of the transfer GP model.

$$\begin{aligned} Var(\boldsymbol{f_{t|D_s,D_t}}) &= \boldsymbol{K_{tt|D_s}} - \boldsymbol{K_{tt|D_s}}(\boldsymbol{K_{tt|D_s}} + \sigma_{wt}\boldsymbol{I})^{-1}\boldsymbol{K_{tt|D_s}} \\ &= \sigma_{wt}^2 K_{tt|D_s}(\boldsymbol{K_{tt|D_s}} + \sigma_{wt}^2\boldsymbol{I})^{-1} \end{aligned} \tag{6.12}$$

Using the posterior mean and variance for the transferred quantitative GP model, the estimated mean shown in 6.8 becomes as shown below.

$$E(\boldsymbol{f_{t*|D_s,D_t}}) = E(\boldsymbol{f_{t*|D_s}}) + \boldsymbol{K_{tt*|D_s}}(\boldsymbol{K_{tt|D_s}} + \sigma_{wt}\boldsymbol{I})^{-1}(\boldsymbol{y_t} - E(\boldsymbol{f_{t|D_s}})) \tag{6.13}$$

The variance shown in Eqn. (6.9) becomes

$$Var(\boldsymbol{f_{t*|D_s,D_t}}) = \boldsymbol{K_{t*t*|D_s}} - \boldsymbol{K_{tt*|D_s}}(\boldsymbol{K_{tt|D_s}} + \sigma_{wt}\boldsymbol{I})^{-1}\boldsymbol{K_{t*t|D_s}} \tag{6.14}$$

6.1.2.4 Bayesian Inference for a Qualitative GP Models After Data Transfer

The posterior of the qualitative GP model after data transfer is computed similarly to the non-data transferred version. By using Laplace approximation results with a GP model as depicted in Eqn. (6.17).

$$\begin{aligned} log(P(\boldsymbol{f_t}|\boldsymbol{D_t}, \boldsymbol{D_s}) &\simeq log(P(\boldsymbol{\hat{f}_t}|\boldsymbol{D_t}, \boldsymbol{D_s})) \\ &+ (\boldsymbol{f_t} - \boldsymbol{\hat{f}_t})\frac{d^2 log(P(\boldsymbol{\hat{f}_t}|\boldsymbol{D_t}, \boldsymbol{D_s}))}{2!d\boldsymbol{f_t^2}}(\boldsymbol{f_t} - \boldsymbol{\hat{f}_t}) \end{aligned} \tag{6.15}$$

$$\boldsymbol{\hat{f}_t} = Argmax_{\boldsymbol{f_t}}(log(P(\boldsymbol{D_t}|\boldsymbol{f_t})) + log(P(\boldsymbol{f_t}|\boldsymbol{D_s})) \tag{6.16}$$

The posterior distribution of boosted reward GP is given as:

$$\boldsymbol{f_{t|q_t,D_s}} \sim GP(\boldsymbol{\hat{f}_t}, (\boldsymbol{K_{tt|D_s}^{-1}} + \boldsymbol{W_t})^{-1}) \tag{6.17}$$

$\boldsymbol{W_t}$ is the negative hessian of $log(P(q_t|\hat{f}_t))$

$$W_{t_{i,j}} = -\frac{\partial^2 log(P(\boldsymbol{q_t}|\boldsymbol{\hat{f}_t})}{\partial f_{t_i} \partial f_{t_j}} \tag{6.18}$$

Since the transferred posterior has a Gaussian distribution, $\boldsymbol{f_{t*|D_t,D_s}}$ also has a Gaussian distribution.

$$\boldsymbol{f_{t_*}|D_t, D_s} \sim GP(E(\boldsymbol{f_{t_*}|D_t, D_s}), Var(\boldsymbol{f_{t_*}|D_t, D_s})) \tag{6.19}$$

Mean and variance of $\boldsymbol{f_{t*|D_t,D_s}}$ is found by marginalizing $P(\boldsymbol{f_{t_*}, f_t|D_t, D_s})$ with respect to $\boldsymbol{f_t}$ then, finding $E(\boldsymbol{f_{t*|D_t,D_s}})$ and $E(\boldsymbol{f^2_{t*|D_t,D_s}})$

$$E(\boldsymbol{f_{t_*}|D_t, D_s}) = E(\boldsymbol{f_{t_*}|D_s}) + \boldsymbol{K_{t_*|D_s} K^{-1}_{tt|D_s}}(\boldsymbol{\hat{f}_t} - E(\boldsymbol{f_{t|D_s}})) \tag{6.20}$$

$$Var(\boldsymbol{f_{t*|D_t,D_s}}) = \boldsymbol{K_{t_*t_*|D_s}} - \boldsymbol{K_{t_*t|D_s}}(\boldsymbol{K_{tt|D_t}} + \boldsymbol{W_t^{-1}})^{-1}\boldsymbol{K_{tt_*|D_s}} \tag{6.21}$$

6.1.3 Modifying Bayesian Committee Machine for Transferring from Multiple Source

Our goal in utilizing the BCM is to approximate the latent function $\boldsymbol{f_{t*|D_t,D_s}}$ without needing a rigorous and computationally intensive data transfer from source to target. We divide the source GP model into multiple but easier-to-compute GP models and conduct the transfer process independently and in parallel for each source GP model. Then using a modified version of the BCM we estimate $\boldsymbol{f_{t*|D_t,D_s}}$.

For utilizing BCM for transfer learning we start with a probabilistic relationship between datasets, $\boldsymbol{D_s}$ and $\boldsymbol{D_t}$, and target latent function $\boldsymbol{f_{t_*}}$. This relationship is demonstrated below with a proportionality.

$$P(\boldsymbol{f_{t_*}}|\boldsymbol{D_t},\boldsymbol{D_s}) \propto P(\boldsymbol{D_s}|\boldsymbol{f_{t_*}},\boldsymbol{D_t})P(\boldsymbol{f_{t_*}}|\boldsymbol{D_t}) \tag{6.22}$$

Then we divide $\boldsymbol{D_s}$ into M subsets such that $\boldsymbol{D_s} = \{\boldsymbol{D_{s_1}}, \boldsymbol{D_{s_2}}, ..., \boldsymbol{D_{s_M}}\}$ and each $\boldsymbol{D_{s_i}}$ in $\boldsymbol{D_s}$ belong to different offline GP model. We approximate $P(\boldsymbol{D_s}|\boldsymbol{f_{t_*}},\boldsymbol{D_t})$ term in 6.22 as a product of $P(\boldsymbol{D_{s_i}}|\boldsymbol{f_{t_*}},\boldsymbol{D_t})$ assuming that each $\boldsymbol{D_{s_i}}$ are independent from each other.

$$P(\boldsymbol{D_{s_i}}|\boldsymbol{f_{t_*}},\boldsymbol{D_t}) \propto \frac{P(\boldsymbol{f_{t_*}}|\boldsymbol{D_{s_i}},\boldsymbol{D_t})}{P(\boldsymbol{f_{t_*}}|\boldsymbol{D_t})} \tag{6.23}$$

Merging of Eqn. (6.22) and Eqn. (6.23) results in

$$P(\boldsymbol{f_{t_*}}|\boldsymbol{D_t},\boldsymbol{D_s}) \propto \frac{\prod_{i=1}^{M} P(\boldsymbol{f_{t_*}}|\boldsymbol{D_{s_i}},\boldsymbol{D_t})}{P(\boldsymbol{f_{t_*}}|\boldsymbol{D_t})^{M-1}} \tag{6.24}$$

By following the previous sections 6.1.2.1 6.1.2.2, one can find exact distributions of each $\boldsymbol{f_{t_*}}|\boldsymbol{D_{s_i}},\boldsymbol{D_t}$ and their probability function, $P(\boldsymbol{f_{t_*}}|\boldsymbol{D_{s_i}},\boldsymbol{D_t})$.

The posterior distribution of $\boldsymbol{f_{t_*|D_t,D_s}}$ can be found as a Gaussian using the proportionality in 6.24 as the resultant probability on the right-hand side has a form of unnormalized Gaussian probability density function.

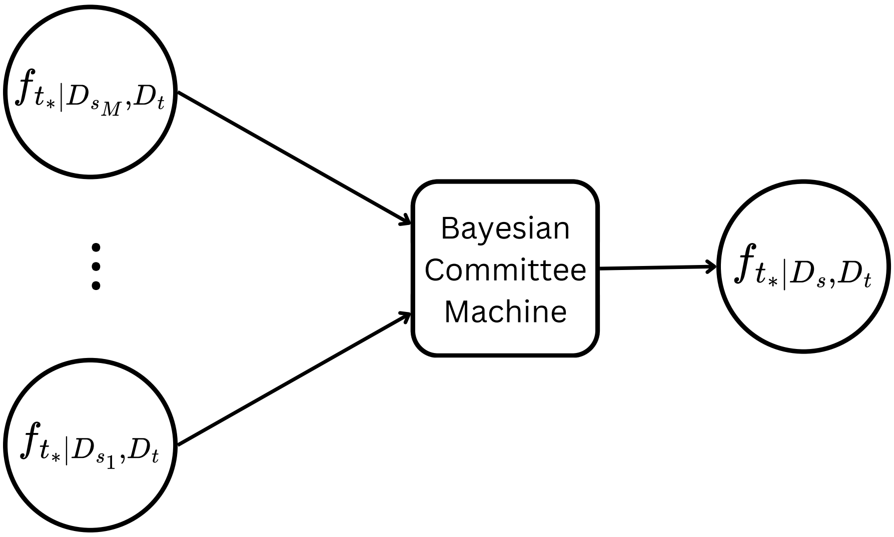

Figure 6.2: Bayesian committee machine approximation for combining M different prediction of a posterior target latent function

Distribution of $\boldsymbol{f_{t*}|D_t, D_s}$ is found as

$$\boldsymbol{f}_{\boldsymbol{t*|D_t,D_s}} \sim GP(E(\boldsymbol{f}_{\boldsymbol{t*|D_t,D_s}}), \boldsymbol{K}_{\boldsymbol{t*|D_t,D_s}}) \tag{6.25}$$

Where the posterior expectation and covariance are found from the equations shown below.

$$E(\boldsymbol{f}_{\boldsymbol{t*|D_t,D_s}}) = \boldsymbol{K}_{\boldsymbol{t*t*|D_t,D_s}}(-(M-1)\boldsymbol{K}^{-1}_{\boldsymbol{t*t*|D_t}}E(\boldsymbol{f}_{\boldsymbol{t*|D_t}})+\sum_{i=1}^{M}(\boldsymbol{K}^{-1}_{\boldsymbol{t*|D_t,D_{s_i}}}E(\boldsymbol{f}_{\boldsymbol{t*|D_t,D_{s_i}}}))) \tag{6.26}$$

$$\boldsymbol{K}_{\boldsymbol{t*t*|D_t,D_s}} = (-(M-1)\boldsymbol{K}^{-1}_{\boldsymbol{t*t*|D_t}} + \sum_{i=1}^{M}(\boldsymbol{K}^{-1}_{\boldsymbol{t*t*|D_t,D_{s_i}}}))^{-1} \tag{6.27}$$

6.2 HiL Pareto Optimization for Robot-Assisted Upper Extremity Rehabilitation with Transfer Bayesian Optimization

Assistive robotic systems in neurorehabilitation automate and enhance therapy sessions, reducing the need for constant physiotherapist intervention. These robots are included in passive, assisted, and resistive exercises. Passive exercises, where robots follow a pre-defined path, do not engage patients' active participation, limiting neuroplasticity recovery. In contrast, assisted and resistive exercises promote recovery by requiring patients to exert effort and actively participate [123, 124]. Numerous studies have confirmed the positive impact of patient effort and voluntary participation in rehabilitation [108, 125, 126], leading to the development of various strategies to encourage these behaviors. However, an ideal level of assistance that fits all patients does not exist due to individual physiological differences.

Determining the right amount of assistance is complex: too little assistance increases effort but insufficient effort prevents completing the exercise or demotivates a patient while too much assistance ensures completion but reduces patient engagement. The assist-as-needed (AAN) concept was introduced to address this challenge, aiming to provide minimal assistance to maximize the effectiveness of rehabilitation by maintaining an optimal balance between challenge and support.

Many studies are focused on developing an AAN strategy focusing on physical engagement. Velocity field-based controllers, which assist only when users deviate from their intended path, are commonly used in rehabilitation robotics to promote active participation [102, 103, 105, 106, 107]. Adaptive controllers dynamically adjust assistance based on user performance through systematic policies derived from reinforcement learning (RL) methods like RBF kernel networks and Q-learning [127, 128, 129]. Although These AAN-based control strategies are effective, they tend to produce solutions specific to individual patients or devices, making them less transferable across different settings.

More generalizable AAN strategies include sensor-dependent approaches, using EMG and EEG sensors to measure patient engagement through bio-signals [111, 130]. RL-based AAN strategies were developed to provide generalizable solutions, independent of specific devices or tasks [115]. For example, Pareek et al. trained an RL model using virtual simulation data, which was then applied in real-time to adjust assistance in a velocity field controller.

While most AAN approaches focus on physical engagement, some studies have begun to include psychological factors like motivation and interest, recognizing that mental engagement is also crucial for effective rehabilitation [116, 117, 131]. These studies suggest that for optimal rehabilitation, both physical and psychological aspects of patient engagement must be considered.

To create an effective AAN strategy, it is important to address both the physical and psychological needs of the user. Simply providing minimal assistance to ensure that a patient can complete an exercise might boost biomechanical participation for a single session. However, if the challenge is too great, it could lead to frustration or fatigue, ultimately reducing the patient's motivation to participate in future sessions.

In our work, we focus on personalizing robot-assisted rehabilitation by finding a trade-off between encouraging physical effort and maintaining the patient's motivation. Our goal is to increase biomechanical participation without causing the patient to feel overwhelmed or discouraged. By carefully adjusting the level of assistance based on both physical performance and psychological responses, we aim to keep patients engaged and motivated throughout their rehabilitation.

To achieve this, we employ a HiL Pareto-based personalization method for rehabilitation sessions, utilizing a multi-criteria BO technique. During HiL trials, we capture the user's biomechanical performance using quantitative metrics, while task comfort is assessed through qualitative evaluations. The evaluation of comfort is used during our strategy as perceiving a discomfort from an exercise indicates that patients are not willing to further participate.

We also integrated transfer learning into our HiL trials to accurately capture the trade-off between physical effort and perceived comfort without needing excessive HiL trials for each rehabilitation task. Integrating transfer learning in upper extremity rehabilitation is crucial for the practicality of our HiL Pareto-based personalization method. Since each rehabilitation session involves multiple exercises, personalizing each one individually during a single session would either require extensive participation in HiL experiments or alternatively, HiL trials can be prematurely concluded with sub-optimal optimization outcomes.

6.2.1 Assistance Controller

In this section, we detail the implementation of force-field [132] and velocity-field controllers [133], as adapted for the upper-limb exoskeleton Assist-On-Arm robot. These controllers, commonly used in assistance-as-needed contexts, promote active patient participation [134]. They operate based on time-independent path errors. The reference path is formed as a four-dimensional B-spline curve, fitted along four joint trajectories (shoulder adduction, flexion, internal rotation, and elbow flexion angles) recorded for each exercise set. The robot does not interfere with shoulder elevation and protraction during the exercises, thereby preserving the natural shoulder path [135].

For the assistance controller, we first created a potential force field controller to create a haptic rendering of a virtual tunnel such that users can only move in a restricted space. Then we have added an additional velocity field-based controller to assist the user in following a desired trajectory.

6.2.2 Potential Force Field Controller

Let the common position vector of the human user and robot be $\boldsymbol{\alpha}$ in joint space and the closest point to this position on the curve be $\boldsymbol{\alpha_c}$. The contour error between the robot position and the reference curve can be calculated as

$$\boldsymbol{e} = \boldsymbol{\alpha_c} - \boldsymbol{\alpha}, \tag{6.28}$$

and the unit normal vector at $\boldsymbol{\alpha_c}$ is given by

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{e}}{|\boldsymbol{e}|}. \tag{6.29}$$

Then, potential force field control as defined in [132] is given by the following rule:

$$\boldsymbol{\tau}_{\text{wall}} = \begin{cases} k\left(|\boldsymbol{e}| - \delta\right)\hat{\boldsymbol{n}} + b\left(\dot{\boldsymbol{\alpha}} \cdot \hat{\boldsymbol{n}}\right)\hat{\boldsymbol{n}} & \text{for } |\boldsymbol{e}| > \delta \\ 0 & \text{otherwise.} \end{cases} \tag{6.30}$$

Eqn. (6.30) defines a haptic viscoelastic wall in the form of a hyper-cylinder around the path where δ is the radius of the cylinder, k is the stiffness of the wall, b is the viscous damping of the wall, and $\dot{\boldsymbol{\alpha}}$ is the vector of joint velocities.

The user can move freely through the cylinder in all directions and the increased δ encourages repetitions without repetition while preventing an excess deviation from the path, such as a spastic movement.

6.2.3 Velocity Field Controller

Since the force field does not provide any support to complete the trial, an additional controller is implemented to assist users in following a desired path.

$$\boldsymbol{\tau}_{\text{flow}} = C_d \left(\dot{\boldsymbol{\alpha}}_{\boldsymbol{ref}} - \dot{\boldsymbol{\alpha}} \right) \tag{6.31}$$

$$\dot{\boldsymbol{\alpha}}_{\boldsymbol{ref}} = \begin{cases} \Gamma \left(|\boldsymbol{e}| \hat{n} - \frac{k_{sh}}{|\boldsymbol{e}|} \hat{t} \right) & |\boldsymbol{e}| > 0 \\ \Gamma \hat{t} & |\boldsymbol{e}| = 0 \end{cases} \tag{6.32}$$

In Eqn. (6.32), $\hat{t}$ represents the tangential direction to the reference path. Γ is the velocity flow constant used for adjusting the magnitude of the assistance while k_{sh} is a parameter used to balance the trade-off between tangential and normal forces and C_d is a constant used to amplify error deviation from reference velocity vector, $\dot{\boldsymbol{\alpha}}_{\boldsymbol{ref}}$.

6.3 HiL Optimization Strategy

To personalize assistance, we selected two separate assistance levels as optimization parameters. The first assistance level was applied during the initial half of the movement, while the second level was applied during the latter half of the movement path.

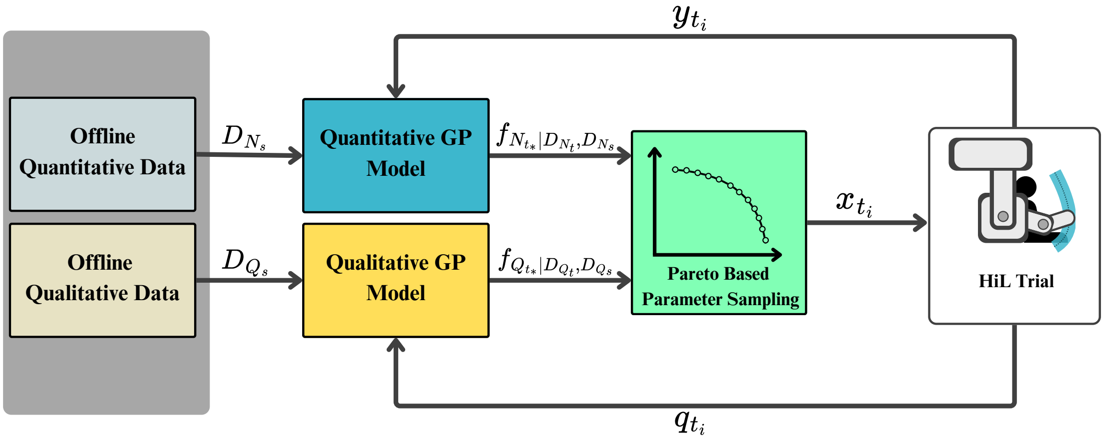

Figure 6.3: Multi-Criteria HiL Bayesian optimization with offline information transfer

Accordingly, we defined two quantitative and one qualitative optimization metric. The quantitative metrics are the average difference between the desired velocity and the average torque exerted by the assistive device, with the optimization goal of minimizing both. The velocity difference metric reflects the user's physical performance in task execution, while the torque metric ensures minimal assistance from the device. For the qualitative metric, we selected user comfort, as it is crucial for voluntary participation in the session. To simplify the multi-objective optimization problem, we combined the quantitative metrics into a single cost function through scalarization.

For the assistance suggestion strategy, we first eliminate assistance regions that are risky for the patient by applying empirical thresholds to avoid excessive or insufficient assistance levels. We then conduct multi-criteria Bayesian optimization using quantitative and qualitative GP models within the constrained assistance parameter regions.

We used transfer learning to start the optimization of both qualitative and quantitative metrics, building upon past data to make the optimization process more efficient. Furthermore, we developed a method for systematically correlating different rehabilitation tasks to adjust the rate of information transfer more precisely.

6.3.1 Quantitative GP Model

Let t_{l_i} and t_{u_i} represent the thresholds for the minimum and maximum assistance levels within a certain region, where maximum applicable assistance corresponds to 1 and the minimum assistance corresponds to 0. Let $x_i = (A_{i1}, A_{i2})$ be the set of assistance levels given for the i^{th} trial and $t_{l_i} \leq A_{i1}, A_{i2} \leq t_{u_i}$. From the assistance parameters, A_{i1} is provided to the user during the first half and A_{i2} is provided during the second half of the motion.

The cost function is derived from two criteria, tracking the desired velocity and less activation of the assistance controller. The controller exerts torque when a patient is distant from the desired track or the patient moves less than the desired velocity.

Let $\boldsymbol{X} = \{x_1, x_2, .., x_n\}$ consist of n assistance parameter set and let $\boldsymbol{D_y} = \{y_1, y_2, .., y_n\}$ be a set consisting of n measured cost evaluation corresponding to $\boldsymbol{X}$.

The each y_i in $\boldsymbol{D_y}$ are are calculated from cost function

$$y_i = \frac{1}{2}(\bar{v_i} - v_d)^2 + \frac{\lambda}{2}(\bar{\tau_i})^2 \tag{6.33}$$

where $\bar{v_i}$ is the average velocity of the end-effector, $\bar{\tau_i}$ is the average torque applied from the device to the human during the current HiL trial. We applied a constant hyper-parameter, λ, to balance the importance of moving at the desired velocity and minimizing the torque required to complete the task.

To simplify statistical modeling, we use statistical normalization before using quantitative observations. Let $D_{y_{norm}} = \{y1_{norm}, y2_{norm}, .., y_{n_{norm}}\}$ be the statistically normalized version of D_y.

$$y_{i_{norm}} = \frac{(y_{i_{meas}} - \frac{1}{n}\sum_{i=1}^{n}(y_{i_{meas}}))}{\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}(y_{i_{meas}} - \frac{1}{n}\sum_{i=1}^{n}(y_{i_{meas}}))^2}} \tag{6.34}$$

Finally, we model a latent function $f_{(x)_N}$ to represent the cost function. We observe a noisy version of the cost function values, $\boldsymbol{D_y}$. Thus white noise convolution is included in the prior GP model. We use GP equations given in 3.2 to develop and update the quantitative GP model.

6.3.2 Qualitative GP Model

Let $f(x)_Q$ be the latent function representing users perceived comfort for the trial. We sequentially observe the user's ordinal classifications corresponding to the trial's comfort level, q_o. The classifications are divided into three categories: the movement causes discomfort, the movement is neutral (neither uncomfortable nor comfortable), and the movement is comfortable. We asked for additional pairwise preferences if the provided ordinal classification is in the same category as the ordinal classification of the previous HiL trial.

Then, based on the provided classifications and pairwise preferences we build a GP model with Laplace approximation as shown in section 3.3.1.1. In our model, we used Gaussian CDF-based probability Φ for the probabilities of qualitative feedback.

The probability for parameter set x_i belonging to the o_j^{th} ordinal class is defined as

$$P(q_{oi} = o_j | f(x_i)) = \Phi\left(\frac{t_{o_j} - f(x_i))}{c_o}\right) - \Phi\left(\frac{t_{oj-1} - f(x_i)}{c_o}\right) \tag{6.35}$$

where c_o is the ordinal noise level constant. The probability of selecting a parameter set x_{i1} over another parameter set x_{i2} is defined by

$$P(q_{pi} = (x_{i1}) \succ (x_{i2}) | f(x)) = \Phi\left(\frac{f(x_{i1}) - f(x_{i2})}{c_p}\right) \tag{6.36}$$

where c_p is the preference noise constant.

6.3.3 Transfer Learning for Multi-Criteria Bayesian Optimization

We use transfer learning with both the qualitative GP model and quantitative GP model of multi-criteria Bayesian to increase the efficiency of the model. For the quantitative GP model GP_{Nt} we transfer quantitative data $\boldsymbol{D_{Ns}}$ of previous experiments and similarly, we transferred qualitative data $\boldsymbol{D_{Qs}}$ to the qualitative GP model GP_{Qt}.

To systematically adjust the correlation between physical exercises, we define a correlation metric using an RBF kernel based on the robot's motion while creating a desired trajectory. With this metric, we aim to develop a consistent and accurate methodology to adjust the transfer rate between tasks as the transfer rate directly affects the performance of optimization. For instance, excessive information transfer from an uncorrelated data source can introduce biases and misguide the optimization process, while insufficient transfer fails to leverage the full benefits of previous experiments.

Let $\boldsymbol{\bar{\alpha}_i}$ be the averaged reference joint value for i^{th} rehabilitation task and $\boldsymbol{\bar{\alpha}_j}$ be the averaged reference joint value for j^{th} rehabilitation task. Then we define the correlation coefficient for two rehabilitation tasks as

$$w_{ij} = e^{-\theta_{task}(\boldsymbol{\bar{\alpha}_i}-\boldsymbol{\bar{\alpha}_j})^T(\boldsymbol{\bar{\alpha}_i}-\boldsymbol{\bar{\alpha}_j})} \tag{6.37}$$

The θ_{task} hyper-parameter is determined empirically. For both qualitative and quantitative models, the correlation coefficient is used to create a multi-task GP model with the ICM method.

We divide all quantitative source data $\boldsymbol{D_{Ns}}$ into M sub-datasets $\boldsymbol{D_{Nsi}}$ in which each $\boldsymbol{D_{Nsi}}$ belongs to a different rehabilitation task. We re-train each quantitative source GP models GP_{Ns_i} with their data $\boldsymbol{D_{Ns_i}}$ and transfer $\boldsymbol{D_{Ns_i}}$ to GP_{Nt} by estimating a prior distribution of target latent function $\boldsymbol{f_{Nt|D_{si}}}$. Then we find $\boldsymbol{f_{Nt|D_{Ns_i},D_{Nt}}}$ by utilizing target data $\boldsymbol{D_{Nt}}$. We combine each $\boldsymbol{f_{Nt|D_{Ns_i},D_{Nt}}}$ with adjusted BCM method to develop a posterior model of $\boldsymbol{f_{Nt|D_{Ns},D_{Nt}}}$.

Algorithm 4 Multi-Criteria Bayesian Optimization with Transfer Learning

```
initiate $S$: Parameter space, $GP_{Nt}$: Quantitative target GP, $GP_{Qt}$: Qualitative target GP, $N_{tot}$: Total iterations, $N_0$: Exploration iterations, $\boldsymbol{D_s}$: Previous experiment data, $\alpha_{num}$: acquisition function for $GP_{Nt}$, $\alpha_{qual}$: acquisition function $GP_{Qt}$
Assign task coefficient between target and pre-existing source GP models
Transfer data $D_s$ to update target GP models $GP_{Nt}$ and $GP_{Qt}$
Create an initial sampling list $X_0$ with search space exploration methods
for $i = 1, 2, \ldots, N_0$ do
    Select a parameter $x_i$ from $X_0$
    Measure velocity and torque data and evaluate numeric cost, $y_i$
    Get qualitative feedback $q_i$ from participant
    Update Quantitative GP model, $GP_{Nt}$ using $x_i$ and $y_i$
    Update Qualitative GP model, $GP_{Qt}$ using $x_i$ and $q_i$
for $j = 1, 2, \ldots, N_{tot} - N_0$ do
    Compute surrogate Pareto $x_p \leftarrow \mathrm{argmax}_{\mathrm{x \in A}} (\alpha_{num}, \alpha_{qual})$
    Select a parameter $x_j := \mathrm{argmax}_{\mathrm{x \in x_p}} (\sigma_{qual} \times \sigma_{num})$
    Measure velocity and torque data and evaluate numeric cost, $y_j$
    Get qualitative feedback $q_j$ from participant
    Update Quantitative GP model, $GP_{Nt}$ using $x_j$ and $y_j$
    Update Qualitative GP model, $GP_{Qt}$ using $x_j$ and $q_j$
```

Similarly, we divide all quantitative source data $\boldsymbol{D_{Qs}}$ into M sub-datasets $\boldsymbol{D_{Qsi}}$. We re-train each qualitative source GP model GP_{Qsi} using their data $\boldsymbol{D_{Qsi}}$ to estimate new prior GP models GP_{Qti}. Then we train each GP_{Qti} using data of the target model $\boldsymbol{D_{Qt}}$ and combined each trained model with the adjusted BCM method to develop a posterior model of GP_{Qt}.

Optimization Algorithm 4 consists of the transfer learning integrated version of a multi-criteria Bayesian optimization algorithm. We begin by transferring data from the previous GP models to the quantitative and qualitative GP models. We use average movement angles of physical tasks in our kernelized distance function to define correlations between GP models.

First N_0 iterations are sampled with search-space sampling strategies such as Sobol sampling. Afterward, the rest of the iterations are sampled by multi-criteria acquisition function strategies. For the experiments, we adopted the same sampling strategy as shown in [41]. We capture the trade-off between the acquisition functions of the two target GP models by developing a Pareto front solution. Then calculate the volumetric variance of each parameter in the Pareto front and select the one with the largest variance for sampling. The process continues until the total number of iterations reaches N_{tot} or no promising parameter set is available, e.g., all expected improvement scores are beneath a certain threshold.

6.4 Experiment Design

In this section, details are given for experiment setup and experimentation design.

6.4.1 Experiment Setup

The experiment setup is created by the upper extremity rehabilitation device, Assist-on-Arm connected to a visual interface. Assist-on-arm uses Matlab Real-Time environment for controlling the device. The controller model includes predefined mods such as calibrating the device, path saving, and task repeating mode with a passive velocity field-based controller. The device includes both mechanical and virtual limitations for safety purposes.

The visual interface displays a bird-feeding task to gamify rehabilitation movements. It also visualizes a virtual tunnel with a trajectory spline at the center that guides the user to track the movement. Visualizations are created by Unity engine-based GUI.

FIGURE 6.4: Experimental setup including a visual display and assistive rehabilitation robot, AssistOn-Arm

6.4.2 Experimental Protocol

We conduct a plot experiment on one participant. The experiment consists of one brief warm-up session and three distinct HiL physical arm movement task practicing sessions. The participant was first informed about the assistive robot and practiced basic arm movements to become familiar with the robot. Then, the participant engaged in three distinct physical arm movement tasks with HiL settings. Each task is designed to engage different biomechanical patterns.

Task 1 (Vertical Arm Movement): The participant was asked to move their right arm vertically from a starting position near the right knee to a final position above the head. The movement was performed without bending the elbow throughout the trajectory.

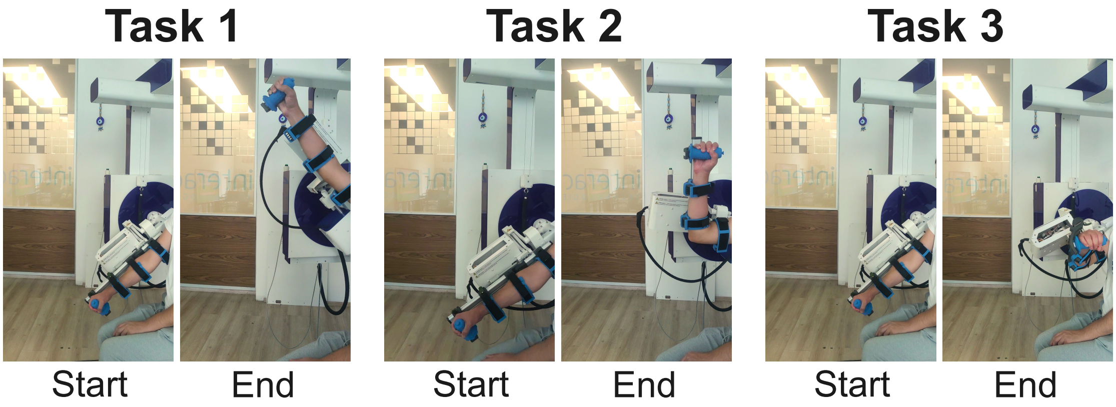

FIGURE 6.5: Start and end positions of the physical tasks

Task 2 (Vertical Arm Movement with Elbow Flexion): The second movement was similar to Task 1, with the participant moving their right arm vertically from the right knee to above the head. However, in this task, the participant was instructed to bend the elbow to a 90-degree angle when the arm reached a position parallel to the ground.

Task 3 (Cross-Body Movement): For the final movement, the participant was instructed to move their right arm diagonally across the body, from a starting position near the right knee to a final position at the left shoulder.

The first task was personalized using multi-criteria Bayesian optimization without using any data from previous optimizations. The second task was personalized with a transfer learning setting utilizing the data of the first task. Then, the third task is also personalized with a transfer learning setting by using data from the first two tasks. We used our kernelized correlation metric for the joint angles to determine the transfer rate between tasks.

For the third task, we applied transfer learning for multi-criteria Bayesian optimization by utilizing data from the first two tasks to increase the efficiency of the personalization process.

We conduct 20 HiL trials for the optimization of each task to ensure the convergence of all models. Optimization of each task took about 10 minutes and overall, the experiment took approximately 40 minutes.

6.5 Results and Discussion

First, We compared posterior GP models of task 2 and task 3 in which task 2's model used task 1 as the source, and task 3's model used both task 1 and 2 as sources. We make this comparison to examine how transfer learning affects the optimization's convergence. During the HiL experiments, we provided enough trials for each model to converge in its optimal region, both with and without transfer learning.

Next, to further analyze convergence, we defined a metric based on the probability of improvement in which we modeled the probability of the final converged optimal parameter surpassing the current optimal parameter. Accordingly, we first selected the parameter with the highest predicted latent value, x_{best}, from the GP model at the $20th$ iteration. For each iteration i, we selected the highest expected latent value $E(f_{max_i}))$, and evaluated the probability that the latent value of x_{best} at the i^{th} iteration, $f_{(x_{best})_i}$, being greater than the maximum expected latent value $E(f_{max_i})$. We represent this probability with $P(f_{(x_{best})_i} > E(f_{max_i}))$.

Since $E(f_{max_i})$ is the maximum latent value at the i^{th} iteration, the probabilistic metric can be at most 0.5, indicating when x_{best} becomes the highest evaluated parameter.

6.5.1 Correlations Between Tasks

According to our kernelized correlation metric, the correlation between task 1 and task 2 is found as 0.82 while the correlation between task 2 and task 3 is found as 0.76 indicating a high correlation between tasks. The correlation between task 1 and task 3 is found as 0.48 indicating these tasks are loosely correlated.

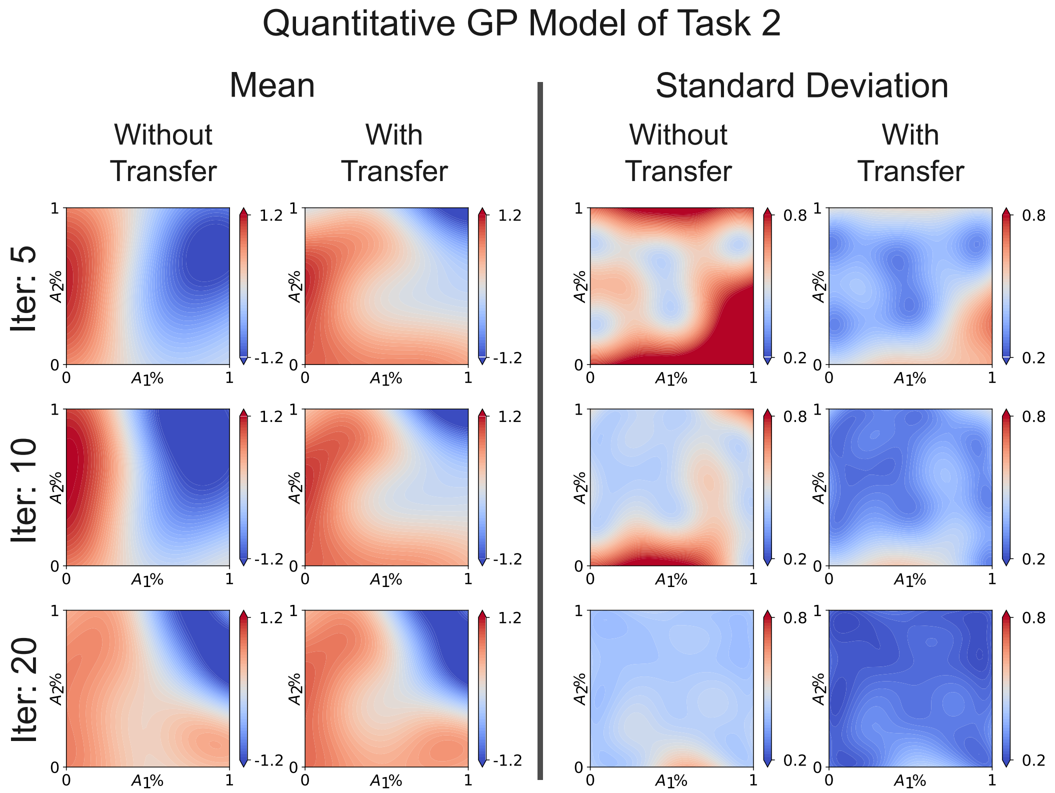

FIGURE 6.6: Quantitative GP model of task 2 with and without information transferred from the qualitative GP model of task 1

6.5.2 Comparison of Quantitative GP Models

For analyzing the plots of quantitative GP models, we used their statistically normalized versions to make a fair comparison between the models with and without transfer learning.

Figure 6.6 compares two versions of task 2's quantitative GP model: one with transfer learning and one without. It shows how these models change across different iterations. In the mean plots, the red areas indicate higher predicted latent scores, where users are expected to have more biomechanical performance. In the standard deviation maps, blue areas represent low uncertainty, while red areas represent regions of high uncertainty.

Throughout the iterations, the mean of the GP model with transfer learning does not deviate as much as the mean of the GP model without transfer learning. Mean of the GP model without transfer. However, when both of the models

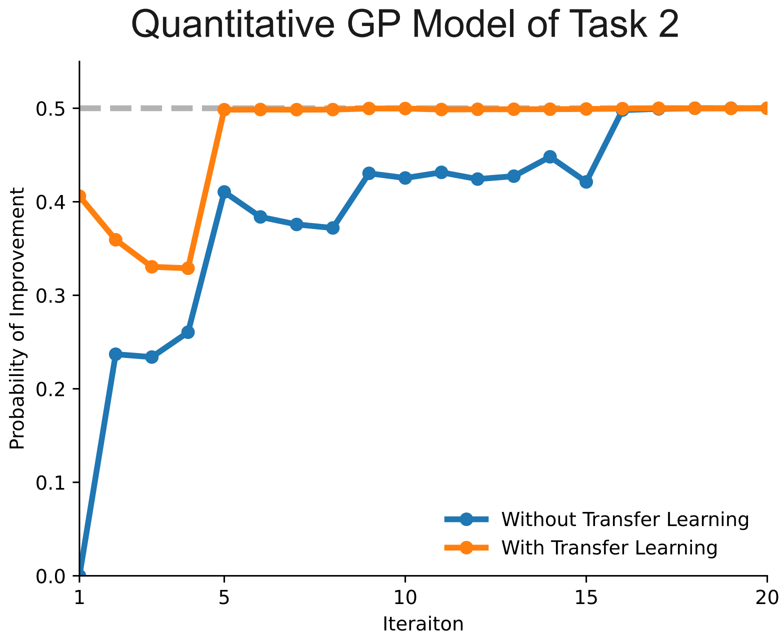

FIGURE 6.7: Convergence of quantitative GP models of task 2 with and without transfer learning

reach iteration 20, their mean plots converge to a similar region.

When comparing the standard deviations of the models, the GP model with transfer learning consistently showed lower standard deviations throughout all iterations. By iteration 5, the standard deviation of the transfer learning model had already reached the final standard deviation level of the model without transfer learning.

Figure 6.7, illustrates the convergence of GP models with and without transfer based on our probabilistic metric. We consider the models to have converged when the metric reaches the maximum score of 0.5 and stays consistent in the following iterations. The GP model with transfer converges after 5 iterations, while the model without transfer takes 17 iterations to converge.

The change in the two versions of quantitative GP models of task 3 is depicted in Figure 6.8. For task 3, the effect of transfer learning is similar to its effect on task 2. In the GP model without transfer, the initial best-predicted assistance

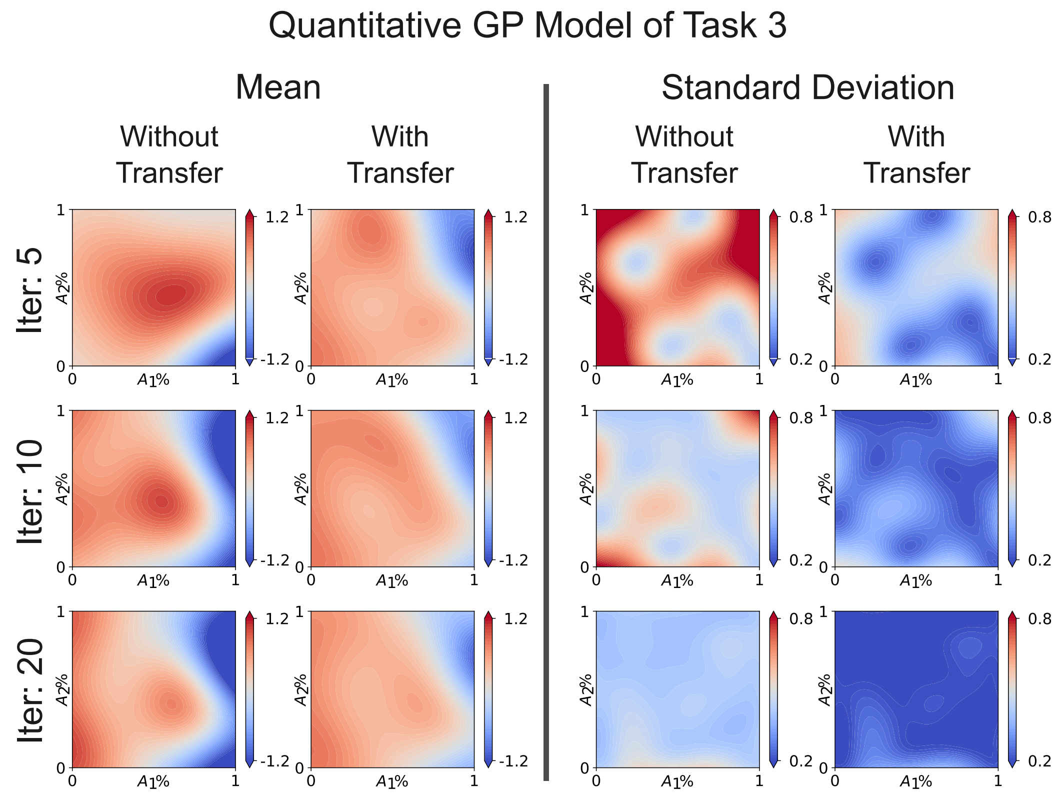

Figure 6.8: Quantitative GP model of task 3 with and without information transferred from the qualitative GP model of task 1 and task 2

parameter appears in the middle of the mean plot (shown by the intense red region). However, over time, the best-predicted parameter shifts toward zero assistance, showing that the participant biomechanically performs better with less help. In contrast, the mean plot of the GP model with transfer learning remains more stable throughout the iterations, while both posterior mean plots of GP models are similar.

The transfer learning utilizes information from two sources for the quantitative GP model of task 3 and consequently, the model is initialized with low standard deviation levels. The difference in standard deviations between the GP models with and without transfer of task 3 is more noticeable compared to the standard deviation difference between the quantitative GP models of task 2 with and without transfer learning.

Figure 6.9 shows that the transferred version of task 3 predicts the posterior optimal parameter x_{best} as near optimal from the beginning of the iterations and

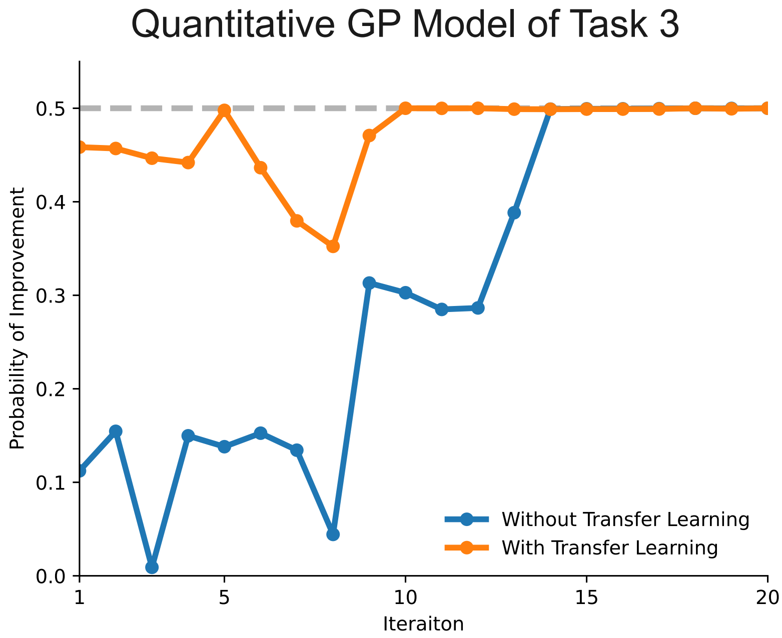

FIGURE 6.9: Convergence of quantitative GP models of task 3 with and without transfer learning

it converges around 10 iteration. Non-transferred version gradually increases the latent scores of its final optimal parameter and converges around 14. As a result, transfer learning increased the efficiency of the quantitative GP model by 28%.

6.5.3 Comparison of Qualitative GP Models

Figures 6.10 and 6.12 show the qualitative GP models for task 2 and task 3, both with and without transfer learning. For both tasks, the qualitative GP models without transfer learning start with a mean of 0 and a standard deviation of 1, while those with transfer learning inherit the posterior mean and standard deviation from their source GP models which makes them initialized with non-zero means and standard deviations below 1.

The changes in qualitative GP models of tasks 2 and 3 across the iterations followed a similar pattern to the quantitative GP models. In particular, the mean plots of the qualitative GP models with transfer learning showed less variation, while the GP models without transfer learning required some iterations for their means to be stabilized. Additionally, the standard deviation levels in the

GP models with transfer learning were lower than those in the models without transfer learning.

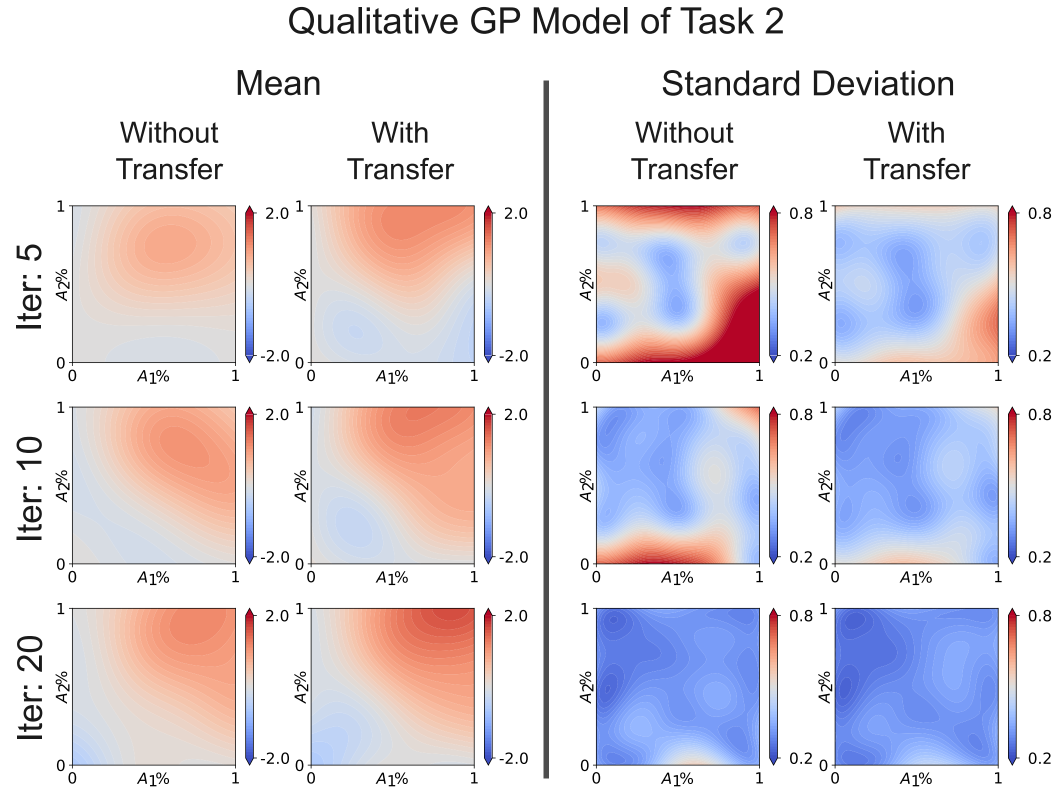

Figure 6.10: Qualitative GP model of second physical task with and without information transferred from task 1

Figure 6.11 shows the convergences of the qualitative GP model with and without transfer learning. According to the probabilistic metric, the model with transfer learning converges at 14 iterations while the model without transfer learning converges at 17 iterations. Similarly, Figure 6.13 shows the convergences of the two models in which GP models with and without transfer learning converge at trial 10. Although they converge at the same time, mean plots in Figure 6.12 show that the optimal regions of the GP model with transfer learning less changed during trials in comparison with the GP model without transfer learning.

Overall, with this work, we developed a transfer learning strategy to improve the efficiency of HiL Bayesian optimizations. Our approach allows for the transfer of information from GP models with both quantitative and qualitative metrics,

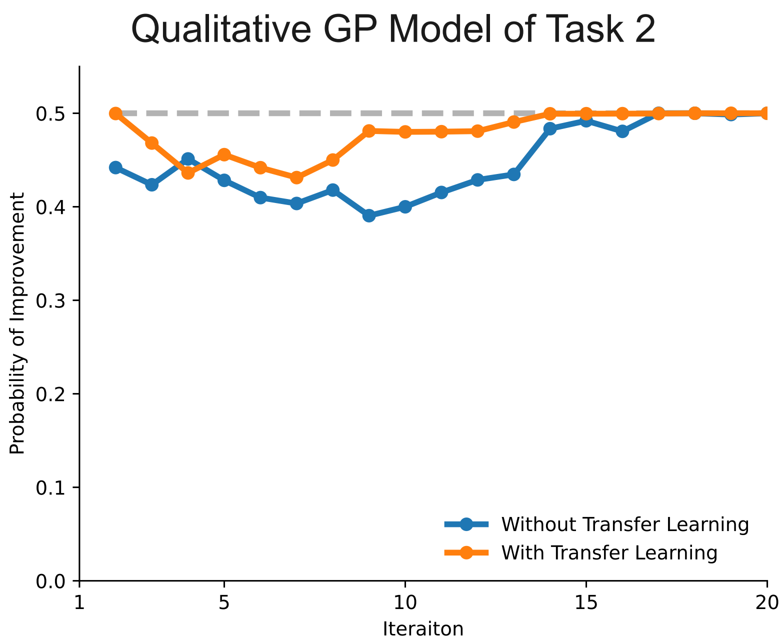

FIGURE 6.11: Convergence of quantitative GP models of task 3 with and without transfer learning

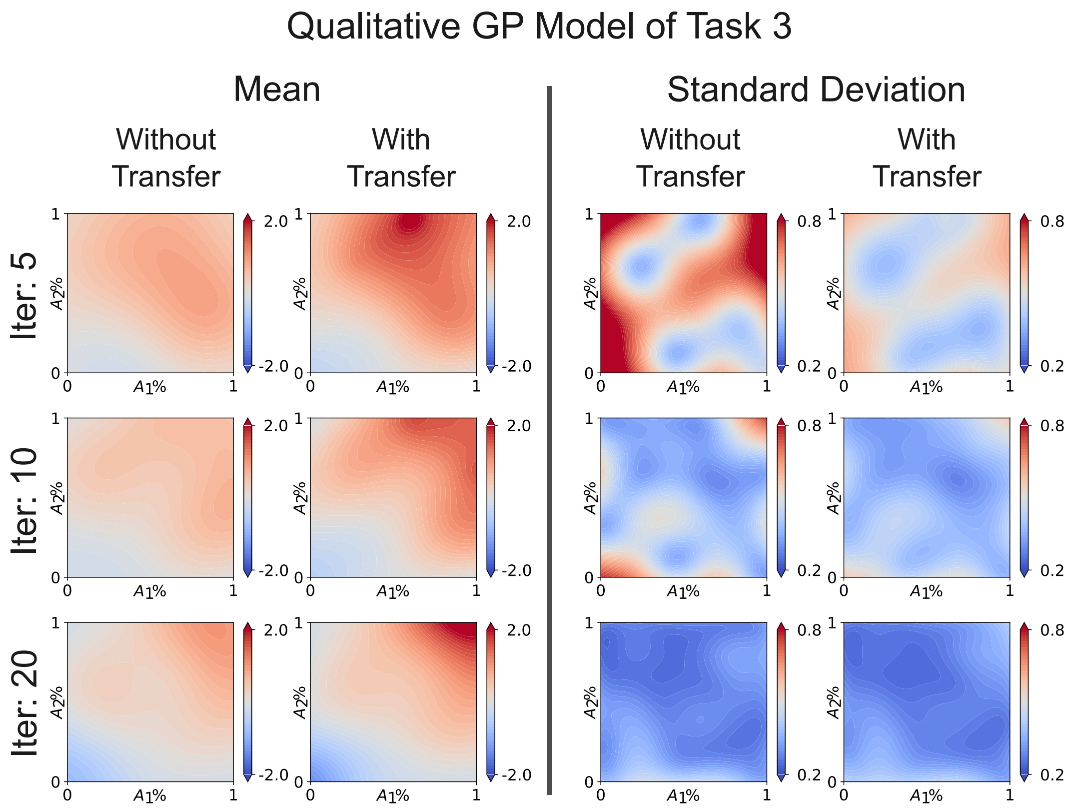

FIGURE 6.12: Qualitative GP model of task 3 with and without information transferred from task 1 and task 2

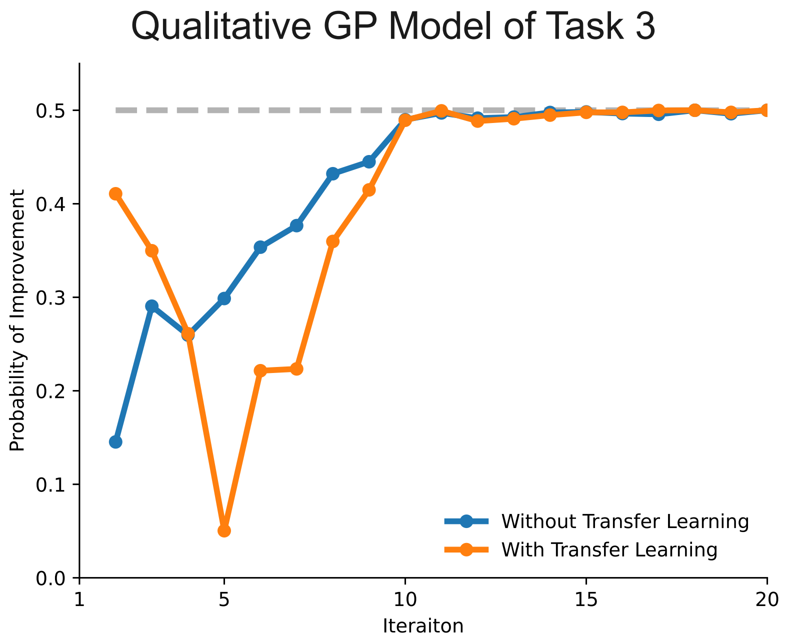

FIGURE 6.13: Convergence of quantitative GP models of task 3 with and without transfer learning

and it also facilitates transfer from multiple source models, which is challenging without an approximation method.

We first developed a method for transferring information from a single source to a target GP model. We then extended this to handle multiple sources using a modified version of BCM approximation to allow transferring information without calculating inter-dependencies between sources. The method also accurately calculates the effect of target data on posterior predictions exactly, without making any approximations.

To test our approach, we conducted an experiment where a participant performed three physical rehabilitation tasks with an assistive robot. The results show that transfer learning accelerates the convergence of HiL optimizations for both quantitative and qualitative metrics. Although both types of models required fewer iterations for convergence, the improvement was more for the quantitative GP models, in which, the number of trials needed for convergence was significantly reduced.

While our findings suggest that transfer learning improves the efficiency of HiL experiments, such as the experiments with robot-assisted rehabilitation, the experiment was a plot experiment for demonstrating our transfer learning approach on a real-life scenario and not enough to statistically conclude its effectiveness. Future research will focus more on human trials to validate the effectiveness of our transfer learning strategy.

Chapter 7

Conclusion

In this thesis, we introduced sample-efficient Bayesian strategies for Human-in-the-Loop (HiL) optimization and demonstrated their applications across various domains of physical Human-Robot Interaction (pHRI). We first presented a Bayesian optimization (BO) framework based on qualitative user feedback to improve the perceived realism of haptic rendering under conflicting haptic-visual cues. Then, we extended BO to solve multi-criteria optimization problems, including both qualitative and quantitative metrics to personalize assistance levels during human motor learning. Additionally, we developed a transfer learning strategy to boost the sample efficiency of HiL optimization and demonstrated its practical use in personalizing robot-assisted upper-extremity rehabilitation.

We first developed a BO framework utilizing qualitative feedback through pairwise comparisons and ordinal classifications. We applied our BO framework to a perceptual HiL study for exploring visual-haptic cue integration during multimodal haptic rendering under conflicting cues to establish a systematic approach to determine the optimal visual scaling for haptic manipulation. As a result of our study, we showed that HiL trials can effectively optimize visual-haptic congruency parameters to ensure high perceived realism. We showed that by solely adjusting visual feedback, a visual-haptic interface can change the perception of a haptic torsional spring rendering to be perceived as stiffer or less stiff. Consequently, by utilizing visual scaling we showed that the range of perceived haptic rendering levels can surpass the physical limits of the haptic device.

We expanded our research by aggregating data from HiL experiments to create a predictive probability model to estimate the perceptual decisions of a population. This model helps us study how multi-modal sensory cues affect people during sensory integration, especially when those cues conflict. The model can be used to explore important psychophysical thresholds such as the just noticeable difference (JND) and the point of subjective equality (PSE) which allows us to better understand how people process conflicting sensory information and how they prioritize one sensory cue over another when discrepancies occur.

We advanced our HiL research by including multi-criteria BO to tackle the challenges of optimizing multiple, often conflicting, objectives in pHRI applications. We proposed a HiL Pareto optimization approach to characterize the trade-off between performance and the perceived challenge level in motor learning tasks, specifically to provide minimal assistance. Our multi-criteria optimization model combined a quantitative metric to capture user performance with a qualitative metric to measure the perceived challenge level based on preference-based feedback. This trade-off, characterized by HiL Pareto optimization, informed the design of assist-as-needed training sessions with optimal assistance levels.

We tested our approach in a HiL experiment where participants learned a new motor skill using virtual reality training. This training was provided to healthy individuals, and we compared our method's performance with another leading adaptive teaching strategy to show its effectiveness. As the learning process continued, we found that the trade-off between objectives evolved, and our multi-criteria optimization method gave a fair way to measure progress, even when participants couldn't complete the task without assistance.

Finally, we enhanced our HiL methodologies with a transfer learning strategy, significantly improving sample efficiency by leveraging data from previous experiments. These strategies reduced the need for extensive new data collection, addressing the impracticality of starting HiL trials from scratch for each new task in pHRI applications. We demonstrated the effectiveness of transfer learning in multi-criteria BO through a HiL experiment to develop personalized robot-assisted rehabilitation sessions with various physical tasks.

Through HiL optimizations, we applied different levels of assistance to a rehabilitation robot while guiding a participant and, we used a quantitative metric to assess user effort and a qualitative metric to evaluate perceived comfort. Once an initial Pareto curve was established via HiL optimization, we transferred this knowledge to subsequent tasks by correlating similarities among physical tasks, eliminating the need for empirical or sample-dependent correlation methodologies. Our results indicated that transfer learning could significantly accelerate HiL optimizations, making the process more efficient and effective.

In summary, we developed advanced extensions for BO strategies that broaden the range of HiL solutions in various pHRI applications, making them more applicable and practical. We applied these enhanced BO methodologies across several HiL-based pHRI studies to showcase their feasibility and effectiveness. By doing so, we also provided a clear guide for how these methods can be used in other pHRI fields.

Bibliography

[1] M. Yalcin and V. Patoglu, “Kinematics and design of AssistOn-SE: A self-adjusting shoulder-elbow exoskeleton,” in *IEEE International Conference on Biomedical Robotics and Biomechatronics*, pp. 1579–1585, 2012.

[2] T. Morizono, K. Kurahashi, and S. Kawamura, “Realization of a virtual sports training system with parallel wire mechanism,” in *Proceedings of International Conference on Robotics and Automation*, vol. 4, pp. 3025–3030 vol.4, 1997.

[3] A. M. Okamura, “Haptic feedback in robot-assisted minimally invasive surgery,” *Current Opinion in Urology*, p. 102–107, 2009.

[4] M. Bergholz, M. Ferle, and B. M. Weber, “The benefits of haptic feedback in robot assisted surgery and their moderators: a meta-analysis,” *Scientific Reports*, vol. 13, no. 1, p. "19215", 2023.

[5] A. Otaran, O. Tokatli, and V. Patoglu, “Physical Human-Robot Interaction Using HandsOn-SEA: An Educational Robotic Platform With Series Elastic Actuation,” *IEEE Transactions on Haptics*, vol. 14, no. 4, pp. 922–929, 2021.

[6] F. Bernardoni, Özen, K. Buetler, and L. Marchal-Crespo, “Virtual reality environments and haptic strategies to enhance implicit learning and motivation in robot-assisted training,” in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pp. 760–765, 2019.

[7] C. Basdogan, C.-H. Ho, and M. Srinivasan, "Virtual environments for medical training: graphical and haptic simulation of laparoscopic common bile duct exploration," *IEEE/ASME Transactions on Mechatronics*, vol. 6, no. 3, pp. 269–285, 2001.

[8] W. Felt, J. C. Selinger, J. M. Donelan, and C. D. Remy, ""Body-in-the-loop": Optimizing device parameters using measures of instantaneous energetic cost," *PLOS One*, vol. 10, no. 8, pp. 1–21, 2015.

[9] J. R. Koller, D. H. Gates, D. P. Ferris, and D. C. Remy, "'Body-in-the-Loop' Optimization of Assistive Robotic Devices: A Validation Study," in *RSS*, 2015.

[10] E. Novoseller, *Online Learning from Human Feedback with Applications to Exoskeleton Gait Optimization.* PhD thesis, CalTech, 2021.

[11] J. Zhang, P. Fiers, K. A. Witte, R. W. Jackson, K. L. Poggensee, C. G. Atkeson, and S. H. Collins, "Human-in-the-loop optimization of exoskeleton assistance during walking," *Science Robotics*, vol. 356, no. 6344, pp. 1280–1284, 2017.

[12] Y. Ding, M. Kim, S. Kuindersma, and C. J. Walsh, "Human-in-the-loop optimization of hip assistance with a soft exosuit during walking," *Science Robotics*, vol. 3, no. 15, p. eaar5438, 2018.

[13] J. Hejna and D. Sadigh, "Few-shot preference learning for human-in-the-loop rl," 2022.

[14] E. Biyik, N. Huynh, M. Kochenderfer, and D. Sadigh, "Active Preference-Based Gaussian Process Regression for Reward Learning," in *RSS*, 2020.

[15] M. Tucker, E. Novoseller, C. Kann, Y. Sui, Y. Yue, J. W. Burdick, and A. D. Ames, "Preference-based learning for exoskeleton gait optimization," in *IEEE International Conference on Robotics and Automation*, pp. 2351–2357, 2020.

[16] P. I. Frazier, "A tutorial on bayesian optimization," 2018.

[17] E. Brochu, V. M. Cora, and N. de Freitas, "A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning," *CoRR*, 2010.

[18] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016.

[19] C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, vol. 7. MIT Press, 2005.

[20] H. J. Kushner, "A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise," *Journal of Basic Engineering*, vol. 86, no. 1, pp. 97–106, 1964.

[21] R. M. Neil, *Bayesian Learning for Neural Networks.* PhD thesis, University of Toronto, 1995.

[22] J. Wu, X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, and S.-H. Deng, "Hyperparameter optimization for machine learning models based on bayesian optimizationb," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, 2019.

[23] M. Kim, Y. Ding, P. Malcolm, J. Speeckaert, C. J. Siviy, C. J. Walsh, and S. Kuindersma, "Human-in-the-loop bayesian optimization of wearable device parameters," *PLOS One*, vol. 12, no. 9, 2017.

[24] F. Zahedi, D. Chang, and H. Lee, "User-adaptive variable damping control using bayesian optimization to enhance physical human-robot interaction," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2724–2731, 2022.

[25] C. Williams and D. Barber, "Bayesian classification with gaussian processes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1342–1351, 1998.

[26] W. Chu and Z. Ghahramani, "Preference learning with Gaussian processes," in *Int. Conf. on Machine Learning*, pp. 137–144, 2005.

[27] A. Kapoor, K. Grauman, R. Urtasun, and T. Darrell, “Active learning with gaussian processes for object categorization,” in *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, 2007.

[28] K. Li, M. Tucker, E. Biyik, E. Novoseller, J. Burdick, Y. Sui, D. Sadigh, Y. Yue, and A. Ames, “ROIAL: Region of Interest Active Learning for Characterizing Exoskeleton Gait Preference Landscapes,” in *IEEE Int. Conf. on Robotics and Auto.*, pp. 3212–3218, 2021.

[29] L. Owen, J. Browder, B. Letham, G. Stocek, C. Tymms, and M. Shvartsman, “Adaptive nonparametric psychophysics,” 2021.

[30] J. R. Gardner, X. Song, K. Q. Weinberger, Barbour, D. L., and J. P. Cunningham, “Psychophysical detection testing with bayesian active learning,” in *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*, 2015.

[31] J. Schlittenlacher, R. Turner, and B. Moore, “Audiogram estimation using bayesian active learning,” *The Journal of the Acoustical Society of America*, vol. 144, pp. 421–430, 07 2018.

[32] X. Song, K. Sukesan, and D. Barbour, “Bayesian active probabilistic classification for psychometric field estimation,” *Attention, Perception, and Psychophysics*, vol. 80, 2018.

[33] J. Browder, S. Bochereau, F. van Beek, and R. King, “Stiffness in virtual contact events: A non-parametric bayesian approach,” in *IEEE World Haptics Conference*, 2019.

[34] M. Tucker, M. Cheng, E. Novoseller, R. Cheng, Y. Yue, J. W. Burdick, and A. D. Ames, “Human preference-based learning for high-dimensional optimization of exoskeleton walking gaits,” in *IEEE International Conference on Robotics and Systems*, pp. 3423–3430, 2020.

[35] B. Catkin and V. Patoglu, "Preference-Based Human-in-the-Loop Optimization for Perceived Realism of Haptic Rendering," *IEEE Trans. on Haptics*, pp. 1–6, 2023.

[36] S. Lu, M. Zheng, M. C. Fontaine, S. Nikolaidis, and H. Culbertson, "Preference-driven texture modeling through interactive generation and search," *IEEE Transactions on Haptics*, vol. 15, no. 3, pp. 508–520, 2022.

[37] N. Heravi, H. Culbertson, A. M. Okamura, and J. Bohg, "Development and evaluation of a learning-based model for real-time haptic texture rendering," *IEEE Transactions on Haptics*, pp. 1–12, 2024.

[38] A. Mathern, O. Steinholtz, A. Sjöberg, M. Önnheim, K. Ek, R. Rempling, E. Gustavsson, and M. Jirstrand, "Multi-objective constrained Bayesian optimization for structural design," *Struct. and Multidiscip. Opt.*, vol. 63, 2021.

[39] S. Suzuki, S. Takeno, T. Tamura, K. Shitara, and M. Karasuyama, "Multi-objective Bayesian Optimization using Pareto-frontier Entropy," in *Int. Conf. on Machine Learning*, vol. 119, pp. 9279–9288, 2020.

[40] B. Paria, K. Kandasamy, and B. Póczos, "A Flexible Framework for Multi-Objective Bayesian Optimization using Random Scalarizations," in *Uncertainty in Artificial Intelligence Conf.*, vol. 115, pp. 766–776, 2020.

[41] S. B., A. D., N. K. J., and J. R. D., "Uncertainty-Aware Search Framework for Multi-Objective Bayesian Optimization," *CoRR*, 2022.

[42] M. Emmerich, J. W. Klinkenberg, and N. Bohrweg, "The computation of the expected improvement in dominated hypervolume of Pareto front approximations," in *Rapport technique, Leiden U.*, vol. 34, 2008.

[43] D. Hernandez-Lobato, J. Hernandez-Lobato, A. Shah, and R. Adams, "Predictive Entropy Search for Multi-objective Bayesian Optimization," in *Int. Conf. on Machine Learning*, vol. 48, pp. 1492–1501, 2016.

[44] J. Knowles, "Parego: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 10, pp. 50–66, 2006.

[45] P. Y. Papalambros and D. J. Wilde, *Principles of optimal design: modeling and computation.* Cambridge University Press, 2000.

[46] R. T. Marler and J. S. Arora, "The weighted sum method for multi-objective optimization: new insights," *Structural and multidisciplinary optimization*, vol. 41, no. 6, pp. 853–862, 2010.

[47] R. Unal, G. Kiziltas, and V. Patoglu, "A multi-criteria design optimization framework for haptic interfaces," in *Haptics Symposium*, pp. 231–238, 2008.

[48] S. B., A. D., N. K. J., and J. R. D., "Uncertainty-aware search framework for multi-objective bayesian optimization," in *AAAI Conf. on AI*, pp. 10044–52, 2020.

[49] R. Astudillo, K. Li, M. Tucker, C. X. Cheng, A. D. Ames, and Y. Yue, "Preferential Multi-Objective Bayesian Optimization," 2024.

[50] N. Cressie, "The origins of kriging," *Mathematical Geology*, vol. 22, no. 3, pp. 239–252, 1990.

[51] O. Babak and C. V. Deutsch, "An intrinsic model of coregionalization that solves variance inflation in collocated cokriging," *Computers Geosciences*, vol. 35, no. 3, pp. 603–614, 2009.

[52] E. V. Bonilla, K. Chai, and C. Williams, "Multi-task Gaussian Process Prediction," in *Advances in Neural Information Processing Systems* (J. Platt, D. Koller, Y. Singer, and S. Roweis, eds.), vol. 20, 2007.

[53] A. Forrester, A. Sobester, and A. Keane, "Multi-fidelity optimization via surrogate modelling," *Proc. R. Soc. A*, vol. 463, pp. 3251–3269, 10 2007.

[54] E. V. Bonilla, F. V. Agakov, and C. K. I. Williams, “Kernel multi-task learning using task-specific features,” in *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics* (M. Meila and X. Shen, eds.), vol. 2 of *Proceedings of Machine Learning Research*, pp. 43–50, PMLR, 2007.

[55] M. A. Alvarez, L. Rosasco, and N. D. Lawrence, “Kernels for vector-valued functions: a review,” 2012.

[56] D. Golovin, B. Solnik, S. Moitra, G. Kochanski, J. Karro, and D. Sculley, “Google vizier: A service for black-box optimization,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 1487–1495, Association for Computing Machinery, 2017.

[57] L. Csató, *Gaussian Processes - Iterative Sparse Approximations.* PhD thesis, Aston University, 03 2002.

[58] V. Tresp, “A Bayesian Committee Machine,” *Neural Computation*, vol. 12, pp. 2719–2741, 11 2000.

[59] Y. W. Teh, M. Seeger, and M. I. Jordan, “Semiparametric latent factor models,” in *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics* (R. G. Cowell and Z. Ghahramani, eds.), vol. R5 of *Proceedings of Machine Learning Research*, pp. 333–340, PMLR, 2005.

[60] K. Yu, V. Tresp, and A. Schwaighofer, “Learning Gaussian processes from multiple tasks,” in *Proceedings of the 22nd International Conference on Machine Learning*, ICML ’05, p. 1012–1019, Association for Computing Machinery, 2005.

[61] K. Swersky, J. Snoek, and R. P. Adams, “Multi-task bayesian optimization,” in *Advances in Neural Information Processing Systems* (C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, eds.), vol. 26, Curran Associates, Inc., 2013.

[62] B. Letham and E. Bakshy, "Bayesian Optimization for Policy Search via Online-Offline Experimentation," 2019.

[63] M. Feurer, J. Springenberg, and F. Hutter, "Initializing bayesian hyperparameter optimization via meta-learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 29, no. 1, 2015.

[64] J. Lübsen, C. Hespe, and A. Eichler, "Towards safe multi-task Bayesian optimization," in *Proceedings of the 6th Annual Learning for Dynamics amp; Control Conference* (A. Abate, M. Cannon, K. Margellos, and A. Papachristodoulou, eds.), vol. 242 of *Proceedings of Machine Learning Research*, pp. 839–851, PMLR, 2024.

[65] M. O. Ernst and H. H. Bülthoff, "Merging the senses into a robust percept," *Trends in Cognitive Sciences*, vol. 8, no. 4, pp. 162–169, 2004.

[66] R. van Ee, W. J. Adams, and P. Mamassian, "Bayesian modeling of cue interaction: bistability in stereoscopic slant perception," *Journal of the Optical Society of America A*, vol. 20, no. 7, pp. 1398–1406, 2003.

[67] J.-P. Bresciani, F. Dammeier, and M. O. Ernst, "Vision and touch are automatically integrated for the perception of sequences of events," *Journal of Vision*, vol. 6, pp. 2–2, 04 2006.

[68] M. J. Young, M. S. Landy, and L. T. Maloney, "A perturbation analysis of depth perception from combinations of texture and motion cues," *Vision Research*, vol. 33, no. 18, pp. 2685–2696, 1993.

[69] M. O. Ernst and M. S. Banks, "Humans integrate visual and haptic information in a statistically optimal fashion," *Nature*, vol. 415, pp. 429–433, 2002.

[70] S. Gepshtein and M. S. Banks, "Viewing geometry determines how vision and haptics combine in size perception," *Current Biology*, vol. 13, no. 6, pp. 483–488, 2003.

[71] H. H. Bülthoff and H. A. Mallot, "Integration of depth modules: stereo and shading," *J. Opt. Soc. Am. A*, vol. 5, no. 10, pp. 1749–1758, 1988.

[72] M. S. Banks and B. T. Backus, "Extra-retinal and perspective cues cause the small range of the induced effect," *Vision Research*, vol. 38, no. 2, pp. 187–194, 1998.

[73] R. G.H., "Auditory influences on visual temporal rate perception," *Journal of Neurophysiology*, vol. 89, pp. 1078–1093, 2003.

[74] B. J. P., E. M. O., D. K., B. G., M. V., and K. A., "Feeling what you hear: auditory signals can modulate tactile tap perception," *Experimental Brain Researcg*, vol. 162, pp. 172–180, 2005.

[75] N. W. Roach, J. Heron, and P. V. McGraw, "Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration," *Proceedings of the Royal Society B.*, vol. 273, pp. 2159—-2168, 2006.

[76] K. P. Kording, U. Beierholm, W. J. Ma, S. Quartz, J. B. Tenenbaum, and L. Shams, "Causal inference in multisensory perception," *Plos One*, vol. 2, no. 9, p. e943, 2007.

[77] M. A. Srinivasan, G. L. Beauregard, and D. L. Brock, "The impact of visual information on the haptic perception of stiffness in virtual environments," 1996.

[78] M. Samad, E. Gatti, A. Hermes, H. Benko, and C. Parise, "Pseudo-haptic weight: Changing the perceived weight of virtual objects by manipulating control-display ratio," in *Conference on Human Factors in Computing Systems*, pp. 1–13, 2019.

[79] M. Kuschel, M. Di Luca, M. Buss, and R. L. Klatzky, "Combination and integration in the perception of visual-haptic compliance information," *IEEE Transactions on Haptics*, vol. 3, no. 4, pp. 234–244, 2010.

[80] C. Basdogan, B. Ataseven, and M. A. Srinivasan, “Perception of soft objects in virtual environments under conflicting visual and haptic cues,” *IEEE Transactions on Haptics*, pp. 1–10, 2023.

[81] Y. Ujitoko and Y. Ban, “Survey of pseudo-haptics: Haptic feedback design and application proposals,” *IEEE Transactions on Haptics*, vol. 14, no. 4, pp. 699–711, 2021.

[82] A. Paljic, J. M. Burkhardt, and S. Coquillart, “Evaluation of pseudo-haptic feedback for simulating torque: A comparison between isometric and elastic input devices,” in *International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 216–223, 2004.

[83] M. Tatezono, K. Sato, K. Minamizawa, H. Nii, N. Kawakami, and S. Tachi, “Effect of haptic feedback on pseudo-haptic feedback for arm display,” in *IEEE ICROS-SICE International Joint Conference*, pp. 4332–4337, 2009.

[84] M. Hosseini, A. Sengul, Y. Pane, J. De Schutter, and H. Bruyninckx, “Haptic perception of virtual spring stiffness using ExoTen-glove,” in *International Conference on Human System Interaction*, pp. 526–531, 2018.

[85] E. Fakhoury, P. Culmer, and B. Henson, “The impact of visual cues on haptic compliance discrimination using a pseudo-haptic robotic system,” in *IEEE International Conference on Robotics and Biomimetics*, pp. 1719–1725, 2017.

[86] D. Matsumoto, K. Hasegawa, Y. Makino, and H. Shinoda, “Displaying variable stiffness by passive nonlinear spring using visuo-haptic interaction,” in *IEEE World Haptics Conference*, pp. 587–592, 2017.

[87] K. van Mensvoort, D. J. Hermes, and M. van Montfort, “Usability of optically simulated haptic feedback,” *International Journal of Human-Computer Studies*, vol. 66, no. 6, pp. 438–451, 2008.

[88] F. Argelaguet, D. A. Gómez Jáuregui, M. Marchal, and A. Lécuyer, "A novel approach for pseudo-haptic textures based on curvature information," in *EuroHaptics Conference*, pp. 1–12, 2012.

[89] S. A. Klein, "Measuring, estimating, and understanding the psychometric function: A commentary," *Perception and Psychopysics*, vol. 63, no. 8, pp. 1421–1455, 2001.

[90] M. R. Leek, "Adaptive procedures in psychophysics research," *Perception and Psychophysics*, vol. 63, no. 8, pp. 1279–1292, 2001.

[91] A. B. Watson and D. G. Pelli, "Quest: A bayesian adaptive psychometric method," *Perception & Psychophysics*, vol. 33, no. 2, pp. 113–120, 1983.

[92] L. L. Kontsevich and C. W. Tyler, "Bayesian adaptive estimation of psychometric slope and threshold," *Vision research*, vol. 39, no. 16, pp. 2729–2737, 1999.

[93] N. Prins, "The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less)," *Journal of vision*, vol. 13, no. 7, pp. 3–3, 2013.

[94] A. B. Watson, "Quest+: A general multidimensional bayesian adaptive psychometric method," *Journal of Vision*, vol. 17, no. 3, pp. 10–10, 2017.

[95] N. Houlsby, F. Huszár, Z. Ghahramani, and M. Lengyel, "Bayesian active learning for classification and preference learning," *arXiv preprint arXiv:1112.5745*, 2011.

[96] B. Settles, "Active learning literature survey," 2009.

[97] B. Letham, P. Guan, C. Tymms, E. Bakshy, and M. Shvartsman, "Lookahead acquisition functions for bernoulli level set estimation," in *International Conference on Artificial Intelligence and Statistics*, pp. 8493–8513, PMLR, 2022.

[98] H. Z. Tan, N. I. Durlach, G. L. Beauregard, and M. A. Srinivasan, "Manual discrimination of compliance using active pinch grasp: The roles of force and work cues," *Perception Psychophysics*, vol. 57, no. 4, pp. 495–510, 1995.

[99] L. Jones and I. Hunter, "A perceptual analysis of stiffness," *Experimental Brain Research*, vol. 79, pp. 150–156, 1990.

[100] A. Erdogan and V. Patoglu, "Slacking prevention during assistive contour following tasks with guaranteed coupled stability," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1587–1594, 2012.

[101] J. Moreno-Valenzuela, "Velocity field control of robot manipulators by using only position measurements," *Journal of the Franklin Institute*, vol. 344, no. 8, pp. 1021–1038, 2007.

[102] P. Y. Li and R. Horowitz, "Passive velocity field control of mechanical manipulators," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 4, pp. 751–763, 1999.

[103] R. Colombo, I. Sterpi, A. Mazzone, C. Delconte, and F. Pisano, "Development of a progressive task regulation algorithm for robot-aided rehabilitation," in *Conf. IEEE Eng. in Med. and Bio. Soc.*, pp. 3123–3126, 2011.

[104] A. Erdogan and V. Patoglu, "Online Generation of Velocity Fields for Passive Contour Following," in *IEEE World Haptics Conference*, pp. 245–250, 2011.

[105] U. Keller, G. Rauter, and R. Riener, "Assist-as-needed path control for the PASCAL rehabilitation robot," in *IEEE Int. Conf. on Rehab. Robotics*, pp. 1–7, 2013.

[106] H. J. Asl, M. Yamashita, T. Narikiyo, and M. Kawanishi, "Field-Based Assist-as-Needed Control Schemes for Rehabilitation Robots," *IEEE/ASME Trans. on Mechatronics*, vol. 25, no. 4, pp. 2100–2111, 2020.

[107] J. Lopes, C. Pinheiro, J. Figueiredo, L. Reis, and C. Santos, "Assist-as-needed Impedance Control Strategy for a Wearable Ankle Robotic Orthosis," in *IEEE Int. Conf. on Auto. Robot Systems and Comp.*, pp. 10–15, 2020.

[108] H. Krebs, J. Palazzolo, L. Dipietro, M. Ferraro, J. Krol, K. Rannekleiv, B. Volpe, and N. Hogan, "Rehabilitation Robotics: Performance-Based Progressive Robot-Assisted Therapy," *Autonomous Robots*, vol. 15, pp. 7–20, 2003.

[109] Y. Li, J. C. Huegel, V. Patoglu, and M. K. O'Malley, "Progressive shared control for training in virtual environments," in *World Haptics*, pp. 332–337, 2009.

[110] M. Sarac, E. Koyas, A. Erdogan, M. Cetin, and V. Patoglu, "Brain Computer Interface based robotic rehabilitation with online modification of task speed," in *IEEE Int. Conf. on Rehab. Robotics*, pp. 1–7, 2013.

[111] R. Yang, Z. Shen, Y. Lyu, Y. Zhuang, L. Li, and R. Song, "Voluntary Assist-as-Needed Controller for an Ankle Power-Assist Rehabilitation Robot," *IEEE Trans. on Biomedical Engineering*, vol. 70, no. 6, pp. 1795–1803, 2023.

[112] J. L. Emken, J. E. Bobrow, and D. Reinkensmeyer, "Robotic movement training as an optimization problem: designing a controller that assists only as needed," in *Int. Conf. on Rehabilitation Robotics*, pp. 307–312, 2005.

[113] V. Squeri, A. Basteris, and V. Sanguineti, "Adaptive regulation of assistance as needed in robot-assisted motor skill learning and neuro-rehabilitation," in *IEEE Int. Conf. on Rehab. Robotics*, pp. 1–6, 2011.

[114] C. Obayashi, T. Tamei, and T. Shibata, "Assist-as-needed robotic trainer based on reinforcement learning and its application to dart-throwing," *Neural Networks*, vol. 53, pp. 52–60, 2014.

[115] S. Pareek, H. J. Nisar, and T. Kesavadas, “AR3n: A Reinforcement Learning-Based Assist-as-Needed Controller for Robotic Rehabilitation,” *IEEE Robotics & Automation Magazine*, pp. 2–10, 2023.

[116] B. Zhong, W. Niu, E. Broadbent, A. McDaid, T. M. C. Lee, and M. Zhang, “Bringing Psychological Strategies to Robot-Assisted Physiotherapy for Enhanced Treatment Efficacy,” *Frontiers in Neuroscience*, vol. 13, 2019.

[117] A. Koenig, X. Omlin, L. Zimmerli, M. Sapa, C. Krewer, M. Bolliger, F. Mueller, and R. Riener, “Psychological state estimation from physiological recordings during robot-assisted gait rehabilitation,” *Journal of rehabilitation research and development*, vol. 48, pp. 367–385, 2011.

[118] R. Garnett, *Bayesian Optimization.* Cambridge University Press, 2023.

[119] F. E. Tosun and V. Patoglu, “Necessary and Sufficient Conditions for the Passivity of Impedance Rendering With Velocity-Sourced Series Elastic Actuation,” *IEEE Transations on Robotics*, vol. 36, no. 3, pp. 757–772, 2020.

[120] J. Anguera, J. Boccanfuso, J. Rintoul, O. Claflin, F. Faraji, J. Janowich, E. Kong, Y. Larraburo, C. Rolle, E. Johnston, and A. Gazzaley, “Video game training enhances cognitive control in older adults,” *Nature*, vol. 501, pp. 97–101, 09 2013.

[121] R. Gray, “Transfer of Training from Virtual to Real Baseball Batting,” *Frontiers in Psychology*, vol. 8, 2017.

[122] P.-I. 205, “Supplementary Document for Human-in-the-Loop Pareto Optimization for Assist-as-Needed Control during Training,” tech. rep., 2024.

[123] R. Teasell, N. A. Bayona, and J. Bitensky, “Plasticity and reorganization of the brain post stroke,” *Top Stroke Rehabil*, vol. 12, no. 3, pp. 11–26, 2005.

[124] C. Winstein, A. Wing, and J. Whitall, “Motor control and learning principles for rehabilitation of upper limb movements after brain injury,” *Handbook of neuropsychology*, vol. 9, pp. 79–138, 2003.

[125] P. Lum, C. Burgar, and P. Shor, “Evidence for improved muscle activation patterns after retraining of reaching movements with the mime robotic system in subjects with post-stroke hemiparesis,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 2, pp. 186–194, 2004.

[126] H. Krebs, N. Hogan, M. Aisen, and B. Volpe, “Robot-aided neurorehabilitation,” *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 1, pp. 75–87, 1998.

[127] E. T. Wolbrecht, V. Chan, D. J. Reinkensmeyer, and J. E. Bobrow, “Optimizing compliant, model-based robotic assistance to promote neurorehabilitation,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 3, pp. 286–297, 2008.

[128] D. Zanotto, P. Stegall, and S. K. Agrawal, “Adaptive assist-as-needed controller to improve gait symmetry in robot-assisted gait training,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 724–729, 2014.

[129] Y. Zhang, S. Li, K. J. Nolan, and D. Zanotto, “Adaptive assist-as-needed control based on actor-critic reinforcement learning,” in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4066–4071, 2019.

[130] M. Hamaya, T. Matsubara, T. Noda, T. Teramae, and J. Morimoto, “Learning assistive strategies from a few user-robot interactions: Model-based reinforcement learning approach,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3346–3351, 2016.

[131] M. Bragoni, M. Broccoli, M. Iosa, G. Morone, D. Angelis, V. Venturiero, P. Coiro, L. Pratesi, G. Mezzetti, A. Fusco, and S. Paolucci, “Influence of

psychologic features on rehabilitation outcomes in patients with subacute stroke trained with robotic-aided walking therapy," *American Journal of Physical Medicine Rehabilitation*, vol. 92, pp. e16–e25, 10 2013.

[132] A. Martinez, B. Lawson, and M. Goldfarb, "A controller for guiding leg movement during overground walking with a lower limb exoskeleton," *IEEE Transactions on Robotics*, vol. 34, pp. 183–193, 2 2018.

[133] A. Martinez, B. Lawson, C. Durrough, and M. Goldfarb, "A velocity-field-based controller for assisting leg movement during walking with a bilateral hip and knee lower limb exoskeleton," *IEEE Transactions on Robotics*, vol. 35, pp. 307–316, 4 2019.

[134] D. Novak and R. Riener, "Control strategies and artificial intelligence in rehabilitation robotics," *AI Magazine*, vol. 36, pp. 23–33, 12 2015.

[135] M. Ergin and V. Patoglu, "AssistOn-SE: A self-aligning shoulder-elbow exoskeleton," in *IEEE International Conference on Robotics and Automation*, pp. 2479–2485, 2012.