

Guiding Co-Adaptation in Physically interacting Human-Robot Teams

MSc Thesis

L.J. Flipse

Guiding Co-Adaptation in Physically interacting Human-Robot Teams

MSc Thesis

by

L.J. Flipse

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday December 16, 2021 at 9:00 AM.

Student number:	4364155
Project duration:	March 1, 2021 – December 16, 2021
Thesis committee:	Prof. dr. ir. D. A. Abbink, TU Delft, chair
	Dr. ing. S. Grammatico, TU Delft, supervisor
	Dr. ir. N. W. M. Beckers, TU Delft, supervisor
	Ir. M. Bianchi, TU Delft, supervisor
	Prof. dr. ir. M. A. Neerincx TU Delft, TNO, examiner

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Preface

Dear reader, this report contains my master's thesis, on which I have dedicated the largest part of my life for the past year. This work concludes both my masters of Systems & Control and Mechanical Engineering, combined into one thesis. This thesis additionally concludes more than 7 years of studying in Delft. This thesis has meandered into a research investigating how different types of adaptive behavior by a robot affect the co-adaptation between a physically interacting human and robot. I performed mathematical analysis, did computer simulations, performed physical validation tests and finally performed an experiment using 18 participants. To my own surprise, I did actually manage to finish this research in the exact week that I had planned it 9 months ago (see Appendix B). This report includes a scientific paper, in which I have condensed 9 months of work into 10 pages of text and figures. This paper is accompanied by 14 appendices, elaborating on technical details of my work, or giving background information.

Although this has been the toughest project I have ever performed in my life, with its ups and downs, it has been mostly ups for me, and looking back on it I have enjoyed working on this project a lot. Honestly, I could not have done this work without the help of a lot of people. First of all, I'd like to thank my daily supervisors Mattia and Niek for all that you have done for me over the past months. You have supported me both in a critical and motivational way, helping me push the research to a higher level. Mattia, I highly appreciate you taking the time to be involved in my thesis, although residing on the other side of the Atlantic ocean. I have found your critical feedback to help me forward very helpful and I'd like to wish you all the best in the rest of your Ph.D. Niek, I really enjoyed our low-key contact, allowing me to easily and quickly ask all the questions I had. Your feedback always helped me to get moving again, even though I felt like I was getting stuck. I'd like to wish you all the best on your adventure to Norway. Secondly, David and Sergio, I'd like to thank you for being my supervisors. Sergio, thank you a lot for taking me on your team. I have really valued your feedback and supervision. David, I (together with the other students) remain astonished after each meeting at how good you are at motivating people and how everybody manages to leave the room after a meeting feeling good.

I would like to also thank some people on a personal note. Wouter, I think I cannot thank you enough for driving me to the university and back for the last half-year, I'm going to miss the daily lunches! Also to everyone with which we shared "taxi Wouter", I had a lot of fun getting ready for a day of hard work together with you. Secondly, to my roommate, Carel, thank you for never stopping with asking questions, and please never stop. Thanks for helping me with the mathematical notation. To Anne-Martijne and Wouter, with which I spent the last half-year in the Haptics Lab, it helped a lot to be able to ventilate frustrations with people who understand what you are going through. I wish you all the best in finishing your studies. Lotte, thank you for your unconditional support and for the cover page. Finally, I'd like to thank everyone who voluntarily participated in my experiment, especially my uncle Jos and grandfather Lou who drove all the way from Zeeland to participate.

At last, to you my dear reader, I'd like to thank you for taking the time to read my work and I hope that you enjoy it as much as I do.

L.J. Flipse
Delft, December 2021

Contents

I	Scientific Paper	1
II	Appendices	15
A	Background	17
A.1	Haptic Shared Control	17
A.2	Co-adaptation	18
A.3	Interaction Strategies	18
B	Planning	21
C	Optimal Control Versus Differential Game	23
D	Mathematical Proofs	25
D.1	Stability proof	25
D.2	Estimating the human cost function	26
E	Simulation Results	29
F	Comparison with state-of-the-art algorithm	33
F1	Stability proof	33
F2	Comparison	36
G	Sensitivity Analysis	39
G.1	Gain matrix K	39
G.2	Normalization parameter κ	40
G.3	Velocity gain bias $\tilde{L}_{h,2}$	41
H	Non-linear Model Identification	43
H.1	Linear model filtering	43
H.2	Grey-box Parameter Estimation	43
H.3	Optimization definition	45
H.4	Results	45
I	Multi-sine signal design	49
I.1	Nonlinear Model Identification Input Signal	49
I.2	Reference Trajectory Design	50
J	Algorithm	51
K	Controller Validation	53
K.1	Virtual Human	53
K.2	Real Human	56
K.3	Conclusion	56
L	Questionnaire	57
M	Informed Consent Form	59
N	Qualitative Results Experiment	63
N.1	Qualitative Results	63
N.2	Analysis	66

O	Figures	67
O.1	Trial data	67
O.2	Participant data	70
O.3	Experiment overview	74
	Bibliography	81

I

Scientific Paper

Guiding Co-Adaptation in Physically Interacting Human-Robot Teams

L.J. Flipse

Delft University of Technology

Delft, The Netherlands

l.j.flipse@student.tudelft.nl

Supervision: Dr. Ir. N.W.M. Beckers, Ir. M. Bianchi, Dr. Ing. S. Grammatico, and Prof. Dr. Ir. D.A. Abbink

Abstract—To make the cooperation within a physical human-robot team as efficient as possible, the team members must be able to co-adapt. We developed and evaluated a robot that adapts to a human, using an adaptation strategy, in such a way as to guide the co-adaptation to have a positive effect on human task contribution and team performance. A novel adaptive control algorithm for the robot was designed, estimating and adapting to the human control strategy, and using a Nash equilibrium to compute the robot’s control inputs. The stability of the controller was theoretically proven and physically validated. Two robot adaptation strategies, positive and negative reinforcement, were compared in an experiment in which 18 participants participated. The negative reinforcement adaptation strategy assists humans on an assist-as-needed basis, whereas the positive reinforcement strategy is designed to intrinsically motivate humans to contribute to the control task. Results show a significant increase in performance in the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy, whereas both conditions show a significant increase in performance compared to manual control. Results additionally show a significant decrease in both estimated (by the robot) and perceived (by the human) control share in the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy. In conclusion, to guide the co-adaptation to increase both performance and engage humans to actively contribute to a control task, a robot should be designed to adapt using a positive adaptation strategy.

Index Terms—Physical human-robot teams, co-adaptation, game theory, adaptive control

I. INTRODUCTION

One of the reasons for the human dominance on Earth is our species’ ability to learn how to efficiently cooperate [1, 2]. Humans can quickly adapt to one another in a dynamic environment, successfully coordinating actions [3]. However, in the current technological era, this ability is no longer our prerogative as robots are becoming more and more intelligible and capable of adapting when interacting with humans; they are turning from tools into functional team members [4]. In a physical human-robot team, shown in fig. 1, a human and a robot can use visual feedback and force feedback (often called haptic feedback) to determine their actions, cooperating towards a common goal. To make this cooperation between the human and the robot as efficient as possible, the robot must be able to adapt to the adapting human [5]. *Co-adaptation*, is defined following van Zoelen et al. [6] as “a process in which at least two parties change their

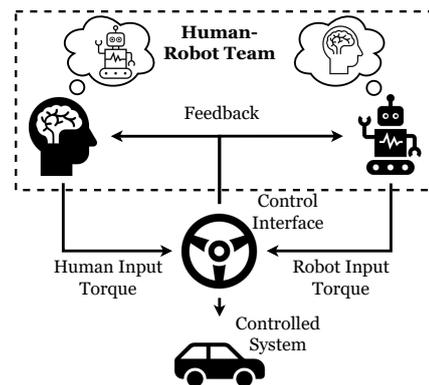


Figure 1: Schematic representation of a physical human-robot team, both tasked with controlling a controlled element. To effectively adapt to a team member while physically interacting in a cooperative task, a mental model of the other player must be established. Here, we focus on physical interaction, hence a mental model needs to be formed based on the information in the force feedback.

behavior and/or mental models concurrently as a consequence of changes in the task or team situation while collaborating”. One of the main challenges in co-adaptation is designing a computational mental model of the human that allows the robot to understand and meaningfully adapt to the adapting human. To this end, the mechanics of co-adaptation between a human and a robot should be better understood. The main problem in understanding the mechanics of co-adaptation in a human-robot team is that human adaptation is an inherently unobservable process [7]. Even a whole field of study on human cognition has emerged from the desire to understand human adaptation towards technology [8].

The inability to understand and predict human adaptation in human-robot teams results in foreseen and unforeseen effects of co-adaptation. Foreseen effects of co-adaptation are often positive. For example in Davies et al. [9], AI and mathematicians successfully cooperated to solve open problems in distinct areas of mathematics, by leveraging the strengths of mathematicians (e.g., their intuition) and AI (e.g., its pattern recognition capabilities). On the other hand,

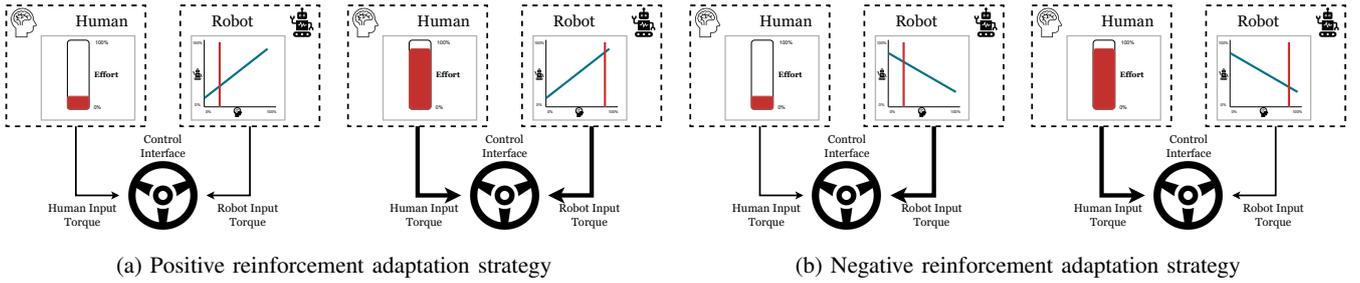


Figure 2: The figures illustrate the fundamental idea of the two adaptation strategies. Figure 2a shows that for the positive reinforcement adaptation strategy, as the human increases its control effort, the robot mimics the human behavior by increasing its control effort as well. Figure 2b shows that for the negative reinforcement adaptation strategy, as the human increases its control effort, the robot exchanges control authority by decreasing its control effort.

unforeseen effects usually lead to undesired situations. First, the human may adapt to a state of reduced vigilance, trust, or situation awareness. For instance, in 2016 a Tesla Model S collided with a tractor-trailer, causing a fatal accident to the driver of the Tesla. The accident was a result of the failure of the human driver to continually monitor the highly reliable autopilot system [10]. Secondly, letting the robot provide too much assistance may cause the human to detach from the joint task. Marchal-Crespo and Reinkensmeyer [11] describes this as the *slacking hypothesis*, where a robot could potentially decrease human motor learning if it encourages a decrease in motor output, effort, or attention. Finally, humans may adapt to misuse the shared control system. In a real driving experiment, Melman et al. [12] shows that using haptic steering guidance, compared to no steering guidance, caused human drivers to increase their speed by 7km/h on average, reducing potential safety benefits of the haptic system to the user. To prevent such undesired and often unforeseen effects, the robot must be designed to guide the co-adaptation to have a desirable effect on human behavior and team performance.

Different types of robot adaptation behavior may have different effects on human behavior. E.g. in motor learning, different types of motivation lead to different effects; where reward leads to memory retention, punishment accelerates learning [13]. This paper distinguishes between two fundamental robot *adaptation strategies* (see fig. 2), prescribing how the robot adapts its behavior to the adapting human, based on principles from motor learning and human factors. The *negative reinforcement* adaptation strategy aims to complement the human's actions on an assist-as-needed basis [14]. This adaptation strategy allows the robot to assist the human by taking a complementary role, allowing the human to adopt a role as leader, follower, collaborator, or competitor. The drawback is that, when the robot is designed to assist the human too much, the human might detach as predicted by the slacking hypothesis [11]. The *positive reinforcement* adaptation strategy revolves around the idea of intrinsic motivation of the human by copying or *mirroring* [15] the human. This way, the robot takes on a pure collaborator role, encouraging the human to actively cooperate. Differences

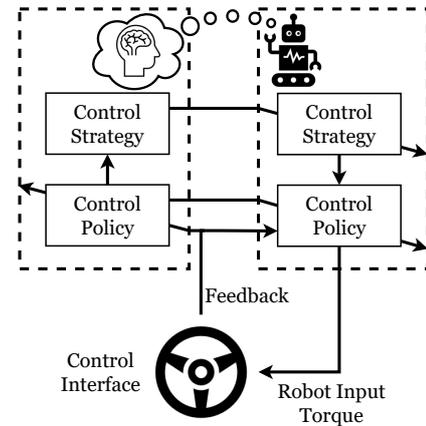


Figure 3: In the shared control task, the robot computes its input torque based on a control policy (i.e., its controller gains), which is determined from its control strategy (i.e., its objectives). Both the control strategy and control policy depend on the model that the robot has of the human team member. The human model consists of a control policy and a control strategy, and it is updated online using measurements.

in team performance for different humans may however be enlarged.

To determine how the human adapts, and to subsequently choose appropriate actions, the robot must not only observe but also understand why the human applies a certain control action. In other words, the robot needs to infer the human's objectives. Abbink et al. [16] proposes a framework consisting of multiple task levels, describing how at each task level, humans and robots can observe and learn to adapt their behavior to one another. As shown in fig. 3, this paper only considers the strategic and operational task levels. Human motor coordination in isolation is often captured by modeling the human as an optimizer of a cost function [17], which describes its *control strategy*. This control strategy determines which *control policy* is selected, relating feedback from the control task to control actions. Multiple studies have shown that the decision-making of humans, physically interacting

with humans [18] or with robots [19], can be approximated by a Nash equilibrium solution [20].

This paper studies co-adaptation by focusing on a haptic shared control [15] problem, motivated by automotive applications. In particular, the goal for the human-robot team is to control a steering wheel. To perform control, the robot must stabilize the jointly controlled system while co-adapting with the human. In Li et al. [21] a differential game theory [22] approach is used in combination with adaptive control [23], using observations to form a computational mental model of the human for the robot, whilst guaranteeing stability. These observations can be through visual feedback, haptic feedback, and other types of feedback. Instead, this work uses state observations. The robot updates its mental model of the human using observations of the system state, without the need for force-sensing measurements.

This paper extends the line of work by Li et al. [21] and proposes a novel algorithm, for which stability is proven, to determine a human’s control strategy. The control task is additionally extended from arm reaching movements, which consider a static reference signal, to a compensatory tracking task, which considers a dynamic reference signal, for which an adaptive control algorithm is designed. Moreover, to the writer’s best knowledge, this paper performs a human-in-the-loop experiment for the first time, examining how human behavior, team performance, and co-adaptation are affected by robot adaptation strategies. Using the human-in-the-loop experiment, this paper aims to answer the following research question.

How should a robot be designed to adapt to a human in a physically interacting human-robot team to guide the co-adaptation to maximize team performance and human task contribution?

A. Hypotheses

It is expected that when a human and a robot physically interact,

- H_1) there exists a significant increase in control performance for both the negative reinforcement adaptation strategy and the positive reinforcement adaptation strategy, compared to manual control;
- H_2) there exists a significant increase in control performance using the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy;
- H_3) there exists a significant decrease in estimated (by the robot) and perceived (by the human) control share using the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy.

B. Contributions

The main contributions of this paper are as follows:

1. Design and validation of a stabilizing adaptive control algorithm that estimates the control strategy and input of a human in a compensatory tracking task in real-time and adjusts

its control strategy accordingly, allowing for co-adaptation. The robot control inputs are determined by combining a state-feedback differential game controller (to determine the robot control actions) with an averaged normalized gradient algorithm and an observer (to estimate the human control actions).

2. Experimental testing of robot adaptation strategies and their effect on human behavior, team behavior, and co-adaptation. Using the aforementioned control method, the robot control strategy can be adapted based on an estimate of the human control strategy, through an adaptation strategy. Two different types of adaptation strategies are designed and their effects on the human and system behavior are experimentally tested in a human-in-the-loop experiment.

C. Outline

This paper is outlined as follows. First, section II presents the adaptive controller design. In section II-A, section II-B and section II-C the theoretical and physical design is discussed. Section II-D validates the control algorithm using physical tests. Section III describes the human-in-the-loop experiment that was performed to test the control algorithm on a group of participants, performing a compensatory tracking task. In Section III-A, the different adaptation strategies that were presented to participants are designed. Section III-B describes the experimental design. Section IV presents the results from the experiment, and in section V these results are discussed. Finally, conclusions, relating to the hypotheses, are drawn in section VI. The planning for the research is found in appendix B.

II. CO-ADAPTIVE CONTROLLER DESIGN

This section presents the design and validation of the adaptive controller. In this paper, it is assumed that on a control interface, which is a steering wheel with haptic force feedback such as is shown in fig. 5a, a human and a robot simultaneously input steering torques $u_h(t)$ and $u_r(t)$, controlling the steering angle $x(t)$, steering rate $\dot{x}(t)$ and steering acceleration $\ddot{x}(t)$. The control interface is modeled as a linear mass-spring-damper system:

$$u_r(t) + u_h(t) = I\ddot{x}(t) + D\dot{x}(t) + Kx(t), \quad (1)$$

where I, D, K are scalar values belonging to the inertia, damping, and stiffness of the steering wheel. In a compensatory tracking task, the control objective is to minimize the distance between the current steering angle $x(t)$ and a reference steering angle $r(t)$. The control equations for compensatory tracking of a reference trajectory $\mathbf{r}(t) \in \mathbb{R}^2$ are linearized around $\boldsymbol{\xi}(t) = 0, u_r(t) = 0, u_h(t) = 0$, assuming that for a high controller frequency and a slow

moving reference the reference signal remains approximately constant ($\mathbf{r}(t) \approx \mathbf{c}$):

$$\begin{aligned} \dot{\boldsymbol{\xi}}(t) &= A\boldsymbol{\xi}(t) + B(u_r(t) + u_h(t)), \quad \boldsymbol{\xi}(t) = \begin{bmatrix} x(t) - r(t) \\ \dot{x}(t) - \dot{r}(t) \end{bmatrix}, \\ A &= \begin{bmatrix} 0 & 1 \\ -\frac{K}{T} & -\frac{D}{T} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{T} \end{bmatrix}. \end{aligned} \quad (2)$$

A. Differential game control

The goal of this section is to design the robot's control inputs $u_r(t)$ such that it stabilizes the system whilst being able to adapt to a human. These control inputs will be computed according to the principles of differential game theory [22]. Game theory allows to directly incorporate the decision-making of a human into the decision-making of the robot. Multiple studies have shown that the decision-making of humans, physically interacting with humans [18] or with robots [19], can be approximated by a Nash equilibrium solution [20]. In appendix C a comparison is made between differential game control and linear quadratic control.

Differential game theory assumes the human inputs $u_h(t)$ to be known a priori, and since human internal processes are inherently unobservable [24], this poses a problem. This problem is often solved by approximating the human's control inputs using a model of the human's actions or control strategy, e.g. driving based on driver risk field [25] or motion primitives [26]. However, using a human model based on behavioral principles often takes the assumption that the human can be captured accurately using a model and assumes the human (parameters) to remain unchanged. An alternative approach, which is pursued in this paper, is estimating the human's actions or strategy in real-time, using state-observations to update the estimate of the human [27, 21].

Game theory models each *player* in the differential game as a cost function, describing its control strategy, which is minimized selfishly. If none of the players can adjust their inputs to their advantage, the Nash equilibrium is reached [22]. Todorov and Jordan [17] have shown that human motor coordination in isolation can be captured using feedback control with a linear quadratic cost function J_h , composed of a penalty on the error state $\boldsymbol{\xi}(t)$ and the input $u_h(t)$. Since the real human cost function and control input is unknown, these are estimated through observations. The robot's cost function J_r , prescribing the robot control strategy, is determined analogously to that of the estimated human cost function \hat{J}_h as

$$\begin{aligned} J_r &= \int_{t=0}^{\infty} \boldsymbol{\xi}^T(t) Q_r \boldsymbol{\xi}(t) + u_r^T(t) u_r(t), \\ \hat{J}_h &= \int_{t=0}^{\infty} \boldsymbol{\xi}^T(t) \hat{Q}_h \boldsymbol{\xi}(t) + \hat{u}_h^T(t) \hat{u}_h(t), \end{aligned} \quad (3)$$

where the robot and estimated human cost function weights $Q_r, \hat{Q}_h \in \mathbb{R}^{2 \times 2}$ are diagonal matrices indicating the penalty on the two components (steering angle and steering rate) of the error state. The Nash equilibrium of the game is reached when the conditions $J_r(u_r^*, \hat{u}_h^*) \leq J_r(u_r, \hat{u}_h^*)$ and

$\hat{J}_h(u_r^*, \hat{u}_h^*) \leq \hat{J}_h(u_r^*, \hat{u}_h)$ are satisfied. The control inputs $u_r(t), \hat{u}_h(t)$ for the robot and estimated human are computed using an error state feedback architecture as:

$$u_r(t) = -L_r \boldsymbol{\xi}(t), \quad \hat{u}_h(t) = -\hat{L}_h \boldsymbol{\xi}(t), \quad (4)$$

where $L_r, \hat{L}_h \in \mathbb{R}^{1 \times 2}$ are the robot and estimated human controller gain respectively.

Let the coupled algebraic Riccati equation for the linear system in eq. (2) and eq. (4) be defined as:

$$\begin{aligned} \mathbf{0}_n &= (A - BB^T P_r)^T \hat{P}_h + \hat{P}_h (A - BB^T P_r) - \hat{P}_h BB^T \hat{P}_h + \hat{Q}_h, \\ \mathbf{0}_n &= (A - BB^T \hat{P}_h)^T P_r + P_r (A - BB^T \hat{P}_h) - P_r BB^T P_r + Q_r, \end{aligned} \quad (5)$$

where $P_r, \hat{P}_h \in \mathbb{R}^{2 \times 2}$ are unknown symmetric matrices and $\mathbf{0}_n \in \mathbb{R}^2$ a column vector of zeros. The controller gains that satisfy the Nash equilibrium solution are then given by theorem 1.

Theorem 1: Let P_r, \hat{P}_h be a stabilizing solution of the coupled Riccati equation in eq. (5). Then define the optimal robot and estimated human controller gains $L_r, \hat{L}_h \in \mathbb{R}^{1 \times 2}$ as:

$$L_r = B^T P_r, \quad \hat{L}_h = B^T \hat{P}_h. \quad (6)$$

Then (L_r, \hat{L}_h) is a feedback Nash equilibrium.

Proof. See Engwerda [28].

B. Observing human controller gains

To observe the human controller gains, a state-observer is used. It is assumed that the human control input is not directly measured, but it can be detected through the dynamics of the system. Additionally, the system states and reference states and their derivatives are assumed measurable. The state-observer equations are given, for some $\Gamma \in \mathbb{R}^{2 \times 2}$ such that $A - \Gamma$ is negative definite, as:

$$\begin{aligned} \dot{\boldsymbol{\xi}}(t) &= A\boldsymbol{\xi}(t) + B(u_h(t) + u_r(t)), \\ \dot{\hat{\boldsymbol{\xi}}}(t) &= A\hat{\boldsymbol{\xi}}(t) + B(\hat{u}_h(t) + u_r(t)) - \Gamma \tilde{\boldsymbol{\xi}}(t), \\ \dot{\tilde{\boldsymbol{\xi}}}(t) &= (A - \Gamma)\tilde{\boldsymbol{\xi}}(t) + B\tilde{u}_h(t), \end{aligned} \quad (7)$$

where $\tilde{\boldsymbol{\xi}}(t) = \hat{\boldsymbol{\xi}}(t) - \boldsymbol{\xi}(t)$ is the estimation error, where $\tilde{u}_h(t) = \hat{u}_h(t) - u_h(t)$ is the input estimation error, and where $\hat{\boldsymbol{\xi}}(t)$ is the estimated system error state. To determine the estimated human controller gain \hat{L}_h , a *Normalized Gradient Algorithm* as described in Toa [29] is used to establish a dynamic update-law $\dot{\hat{L}}_h$. This approach is similar to Shen and Cruz Jose [23], which considers a similar problem in discrete time. This paper however considers a continuous-time approach, such as in Li et al. [21]. In theorem 2, the update law is presented. The estimated human cost function weights \hat{Q}_h can additionally be computed from the coupled Riccati equation in eq. (5), following the procedure in appendix D. Simulation results are presented in appendix E. A comparison with the proposed algorithm and that described in Li et al. [21] is made in appendix F.

Theorem 2: Let the symmetric gain matrix $K = K^T > \mathbf{0} \in \mathbb{R}^{2 \times 2}$, design parameter $\kappa > 0$ and normalization parameter $m(t) = \sqrt{1 + \kappa \boldsymbol{\xi}^T(t) \boldsymbol{\xi}(t)}$ be given. Then, the estimated

human controller input \hat{u}_h , estimated as in eq. (4) and updated using the following update-law

$$\dot{\hat{L}}_h(t) = K \frac{\xi^T(t) \tilde{u}_h(t)}{m^2(t)}, \quad (8)$$

converges to the real human controller input u_h and stabilizes the system described in eq. (7).

Proof. See appendix D.

C. Physical design

To verify the control algorithm in the previous section, a SensoDrive SENSO-Wheel [30] is used. This force feedback steering wheel is shown in fig. 5a. The steering wheel has torque and angle resolution of $0.03Nm$ and 0.009° respectively and a cycle time of $300 - 1000\mu s$.

1) *Filtering sensory noise:* The algorithm in eq. (8) requires not only the steering angle to be known, but also the steering rate and steering acceleration to be available, these signals are computed by numerical differentiation of the steering angle. This however amplifies measurement noise and must be accounted for using filtering. To this end, a digital Biquad filter is used [31]. Appendix H demonstrates that the steering acceleration signal becomes fairly distorted through filtering, placing a physical limit on the performance of the control algorithm.

2) *Nonlinear dynamics compensation:* The dynamical model in eq. (1) assumes a linear model. The SENSO-Wheel violates this assumption. Two nonlinear dynamical components are identified. The nonlinear friction f_{fric} component, depending on the steering rate $\dot{x}(t)$, is modeled by combining a Coulomb friction model and the Stribeck friction model, according to Specker et al. [32]. The nonlinear gravity component f_{grav} , depending on the steering angle $x(t)$, models the steering wheel's asymmetries. The nonlinear components are modeled as follows:

$$\begin{aligned} f_{nl}(\mathbf{x}(t)) &= f_{fric}(\dot{x}(t)) + f_{grav}(x(t)) \\ &= \tau_c \tanh\left(\frac{\dot{x}(t)}{\dot{x}_t}\right) + \tau_s \frac{\dot{x}(t)}{\dot{x}_{sp}} e^{-\left(\frac{\dot{x}(t)}{\sqrt{2}\dot{x}_{sp}}\right)^2 + \frac{1}{2}} + \quad (9) \\ &mg(\delta w \cos x(t) + \delta h \sin x(t)). \end{aligned}$$

The procedure for the identification of the parameters of the nonlinear dynamical components is described in appendix H, resulting in a variance-accounted-for [33] of 71.8% and 93.3% of the model for the steering angle and steering rate respectively. Through feedforward compensation, the steering wheel model is linearized, resulting in the total control input $u_{r,t}(t)$ using

$$u_{r,t}(t) = u_r(t) - f_{nl}(\mathbf{x}(t)). \quad (10)$$

3) *Control algorithm:* The pseudo-code in algorithm 1 summarizes the control algorithm. A more extensive version of the control algorithm is presented in appendix J.

Algorithm 1 Adaptive Control Algorithm

Input: Measured angle $x(t)$, reference state $\mathbf{r}(t)$

Output: Robot control input $u_{r,t}$

- 1: Initialize steering rate $\dot{x}(t)$ and acceleration $\ddot{x}(t)$ filter parameters,
 - 2: Initialize $\hat{L}_h, \hat{Q}_h, L_r, Q_r$ (eq. (5), eq. (6)), $\hat{\xi}(t), \Gamma$ (eq. (7)), K, κ (eq. (8)) and t_{end} .
 - 3: **while** $t \leq t_{end}$ **do**
 - 4: Measure steering angle $x(t)$, compute and filter steering rate $\dot{x}(t)$ and acceleration $\ddot{x}(t)$,
 - 5: Form error states $\xi(t), \dot{\xi}(t)$ from $\mathbf{r}(t)$ (eq. (2)) and compute $\hat{\xi}(t), \dot{\hat{\xi}}(t)$ (eq. (7)),
 - 6: Compute $\hat{u}_h(t)$ (eq. (4)) and $\tilde{u}_h(t)$ (eq. (7)),
 - 7: Update $\hat{L}_h(t)$ from $\hat{L}_h(t)$ (eq. (8)) and $\hat{\xi}(t)$ from $\dot{\hat{\xi}}(t)$ (eq. (7)),
 - 8: Compute $L_r(t)$ (eq. (5), eq. (6)), compute $u_r(t)$ (eq. (4)) and solve \hat{Q}_h (eq. (5)),
 - 9: Calculate $f_{nl}(\mathbf{x}(t))$ (eq. (9)) and determine $u_{r,t}$ (eq. (10)).
 - 10: **return** $u_{r,t}$
 - 11: **end while**
-

D. Validation

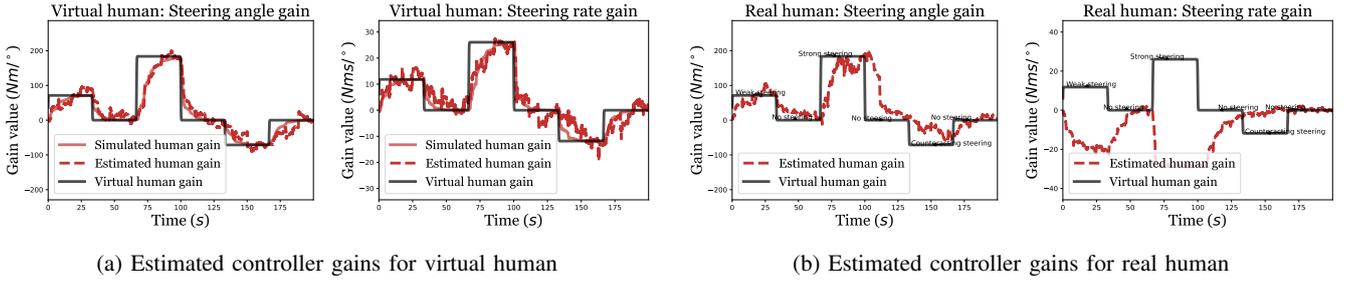
To guarantee the stability of the controller, convergence of the estimated human controller gain \hat{L}_h to the real human controller gain L_h is essential. This is validated using physical tests. Validation of the control algorithm is split up into two parts. First, the control algorithm is tested using a *virtual human*. By adding control inputs to the steering wheel, of which the control algorithm is unaware, human inputs are augmented to the steering wheel. Secondly, the control algorithm is tested on a *real human*. In this scenario, the human inputs cannot be determined analytically but different behaviors towards the controller are tested. In appendix K the validation is extensively described.

A sum-of-sines signal [34] was used to generate the reference signal $\mathbf{r}(t)$ as follows

$$\mathbf{r}(t) = \sum_i^N A_i \sin(\omega_i t + \theta_i), \quad \mathbf{r}(t) = [r(t) \quad \dot{r}(t)]. \quad (11)$$

The forcing function spectra were created using a second-order Butterworth filter [35]. To evenly distribute the power over the frequencies in the reference signal was maximized using crest factor minimization [36]. The design of the reference signal is shown in appendix I.

1) *Virtual human:* In fig. 4a validation of the control algorithm is visually represented. The virtual human was modeled as a set of fixed gains $L_{h,vir}$ in an experiment of 200s. These gains represented light steering, strong steering, and actively counter-steering alternated with no steering action. In the figures, the virtual human is illustrated the black line. The algorithm (interrupted red line) can detect the virtual human gains correctly in considerable time (typically within 15s), albeit with low precision. When compared with



(a) Estimated controller gains for virtual human

(b) Estimated controller gains for real human

Figure 4: Correct identification of human controller gains is essential to guarantee stability for the control algorithm. Figure 4a shows that the robot correctly identifies the controller gain for the steering angle and rate when using a virtual human. Figure 4b shows that, in this control task, a real human neglects the steering rate information and therefore can only be modeled as a controller gain on the steering angle.

a simulation of the algorithm (transparent red line), the effect of noise on the filtered acceleration signal becomes evident.

2) *Real human*: To test how a real human interacts with the control algorithm, an interface is designed, giving visual feedback of the current steering wheel angle and the current reference steering wheel angle. In section III-B3 the interface is discussed. In the validation test, a real human was asked to perform light steering, strong steering and actively countersteering behavior alternated with no steering action, similar to the virtual human. Results are shown in fig. 4b. It is observed that the trend of the estimated human controller gain for the steering wheel angle matches the virtual human controller gain. However, the steering rate gain (which was capped at $-22.9 \text{ Nms}/^\circ$ for safety reasons) does not match the trend of the virtual human, indicating that the human neglects the derivative information of the error signal.

3) *Validation conclusion*: It can be concluded from the validation tests that the control algorithm can estimate the human gain in considerable time, stabilizing the system and thus validating the control algorithm. On the other hand, looking at the test on a real human, given this specific control task, humans cannot be modeled as both a controller gain on the steering wheel angle error and steering rate error but only as a controller gain on the steering wheel angle. This poses no surprise as only visual feedback of the steering wheel angle and steering error is presented to the human. Finally, the control algorithm is physically limited by the signal noise on the computed steering wheel acceleration.

III. EXPERIMENTAL METHODS

This section discusses the human-in-the-loop experiment performed using the control algorithm presented in section II. First, the design of two adaptation strategies, describing how the robot adapts its control behavior to that of the human, is discussed. Then, the design of the experiment is discussed.

A. Adaptation strategies

How the robot performs its feedback is encoded through the robot control strategy, which is captured by the robot cost function. The values of the weights in the cost function weight matrix Q_r determine how much the error signal ($\xi(t)$)

is penalized traded off with how much control effort ($u_r(t)$) is penalized. To encode how the robot adapts its behavior to that of the human, the robot can adapt its control strategy to the robot control strategy. In mathematical terms:

$$Q_r(t) = f(\hat{Q}_h(t)). \quad (12)$$

The design of the negative and positive adaptation strategies are briefly discussed below. Figure 5b and fig. 5c summarize how the robot control strategy (cost function weights) and control policy (controller gains) are determined from the estimated human control strategy and policy (see fig. 3).

1) *Negative reinforcement*: The main goal of the negative reinforcement adaptation strategy is maintaining a certain level of performance, which is encoded through a sharing variable C_0 as

$$Q_r(t) + \hat{Q}_h(t) = C_0. \quad (13)$$

This sharing rule produces a spectrum of control behaviors, ranging from solo driving ($\hat{Q}_h = \mathbf{0}$ or $\hat{Q}_h = C_0$), to cooperation ($C_0 \leq \hat{Q}_h \leq \mathbf{0}$), to competition ($\hat{Q}_h \leq \mathbf{0}$ or $\hat{Q}_h \geq C_0$).

2) *Positive reinforcement*: The main goal of the positive reinforcement adaptation strategy is to intrinsically motivate the human to participate in the control task. This is done by increasing the robot cost function weights proportional to those of the estimated human as

$$Q_r(t) = C_1 + \hat{Q}_h(t), \quad (14)$$

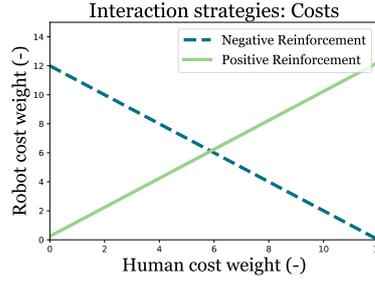
where C_1 is a minimal baseline performance weight matrix, which may be equal to zero. This sharing rule however does not show versatile behavior as in the negative reinforcement adaptation strategy, but only promotes cooperative behavior.

B. Experimental design

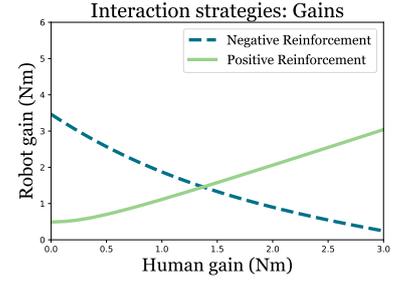
1) *Independent variables*: The experiment considered one independent variable, the experimental conditions, being the *type of control*. These conditions are manual control, positive reinforcement, and negative reinforcement. To randomize the experiment, a Latin-square [37] is used, therefore a multiple of six participants, due to the six possible combinations of conditions, are recruited for the experiment



(a) The experiment consisted of a steering task, in which both the participant and the control algorithm were tasked with compensatory tracking of the yellow circle, controlling the grey plus.



(b) Two experimental conditions are compared in the experiment. For both conditions, the distribution of cost function weights are demonstrated.



(c) The resulting controller gain distribution is presented for the two experimental conditions.

Figure 5: Experimental set-up and experimental conditions.

2) *Dependent variables*: Three metrics are measured in the experiment and statistically compared after the experiment, being root-mean-square steering angle error, estimated control share, and perceived control share. The root-mean-square value of the steering angle error ($RMSE$) is determined as

$$RMSE = \sqrt{\frac{\sum_{t=0}^N \xi^2(t)}{N}}, \quad (15)$$

where $\xi(t)$ is the steering angle error at measurement t . The estimated control share $\alpha_r(t)$ and perceived control share α_h indicate the ratio of the amount of control effort contributed by the human and the robot, computed as

$$\alpha_r(t) = \frac{\hat{L}_h(t) - L_r(t)}{\hat{L}_h(t) + L_r(t)}, \quad \alpha_h = \frac{s_h - s_r}{s_h + s_r}, \quad (16)$$

where the estimated human gain and robot gain ($\hat{L}_h(t), L_r(t)$) are measured during the experiment and the perceived human and perceived robot control contribution (s_h, s_r) are obtained from the questionnaire after each condition. As a metric, the median value of the estimated control share $\alpha_r(t)$ is used. A control share value of $\alpha = 1$ corresponds to the situation where the human performs all the control effort, whereas $\alpha = -1$ corresponds to the situation where the robot performs all the control effort. All values $-1 \leq \alpha \leq 1$ indicate cooperation, where $\alpha = 0$ corresponds to pure cooperation. Finally, a value $|\alpha| > 1$ corresponds to competition.

3) *Experimental set-up*: The experimental set-up consists of the SENSO-Wheel and a visual interface, which is shown in fig. 5a. The visual interface shows the controlled element, represented as a grey cross, and a reference position, represented as a yellow circle, on a screen. The controlled element is related to the steering angle $x(t)$ through a pure gain K , relating the steering angle $x(t)$ and the reference steering angle $r(t)$ to a position $p(t)$ and reference position $p_r(t)$ on the screen as

$$p(t) = Kx(t), \quad p_r(t) = Kr(t). \quad (17)$$

Questionnaire question

On a scale of -2 to 2, during the previous 4 trials...
It felt like I was causing the movement of the grey cross.
It felt like the steering wheel was causing the movement of the grey cross.
It felt like I was contributing to the movement of the grey cross.
It felt like the steering wheel was controlling my movement.

Table I: Questions asked after each condition in the questionnaire.

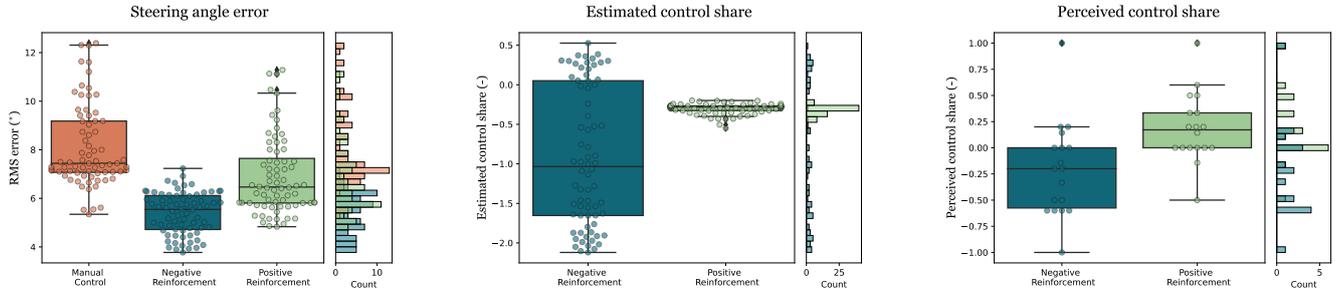
4) *Experiment procedure*: In the experiment, participants were asked to track a reference trajectory as well as possible. The three conditions were presented to the participants in random order and each condition was repeated four times. Each trial was performed for a duration of 77.5s, the same length as the multi-sine signal discussed in section II-D. This multi-sine signal was then re-used in the repetitions, but either mirrored in time, direction, or both. Between each condition, the participants were asked to fill in a questionnaire. The questions in the questionnaire are presented in table I. In appendix L and appendix M, the questionnaire and informed consent form are found.

IV. RESULTS

For the experiment, 18 participants were voluntarily recruited between the age of 24 and 89 years old ($Mdn = 25, std = 16.4$), of which 13 identified as male and 5 identified as female. Raw data of the distribution of controller gains throughout the experiment is presented for one participant in ???. All statistical analysis is performed according to the method by Field [38]. The distribution of metrics for each condition is tested for normality first using a Kolmogorov-Smirnov test. If the distribution of a metric for a condition is significantly non-normally distributed, a non-parametric test is used to test for differences between distributions. All tests were performed in IBM SPSS [39]. The qualitative results and analysis are presented in appendix N. Some additional figures are presented in appendix O.

Metric	No.	Comparison	Test	Significance	Test statistic
RMS steering angle error	1	Between Conditions	Kruskall-Wallis	$p \leq 0.001$	$H(2) = 99.6$
	2	Positive Reinforcement & Negative Reinforcement Manual Control & Positive Reinforcement	Mann-Whitney Mann-Whitney	$p \leq 0.001$ $p \leq 0.001$	$z = -6.093, r = -0.718$ $z = -4.53, r = -.534$
Estimated control share	3	Between Conditions	Kruskall-Wallis	$p \leq 0.001$	$H(2) = 155.3$
Perceived control share	4	Positive Reinforcement & Negative Reinforcement	Mann-Whitney	$p \leq 0.001$	$z = -3.728, r = -.439$
	4	Between Conditions	Kruskall-Wallis Independent sample t-test	$p \leq 0.001$ $p = .005$	$H(2) = 31.4$ $t(34) = -2.99, r = .0696$

Table II: Overview of the results of statistical tests, testing significant differences between distributions of metrics.



(a) Box-plots visualizing the difference in distributions of RMS steering angle error between the conditions.

(b) Box-plots visualizing the difference in distributions of estimated control share (from the perspective of the robot) between the conditions.

(c) Box-plots visualizing the difference in distributions of perceived control share (from the perspective of the human) between the conditions.

Figure 6: The boxplots visually show the distributions of the metrics for different conditions.

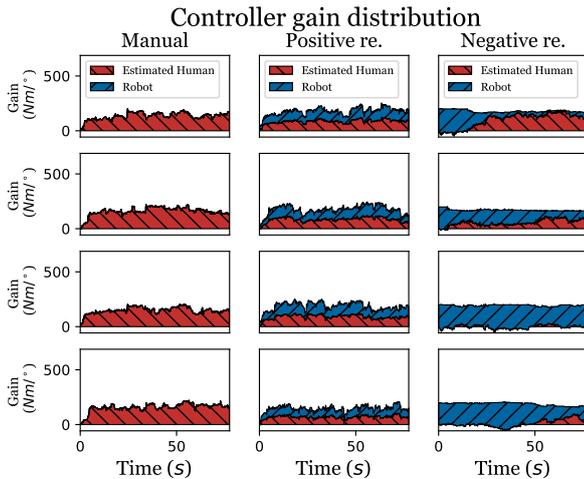


Figure 7: Overview of the distribution of the estimated human gain $\hat{L}_h(t)$ and the robot gain $L_r(t)$ for participant 15.

A. Root-mean-square steering angle error

The Kolmogorov-Smirnov test for the root-mean-square (RMS) steering angle error for the Manual Control condition $D(72) = 0.206, p \leq .001$ and for the Positive Reinforcement condition $D(72) = 0.145, p \leq .001$ were both significantly non-normal. The Kolmogorov-Smirnov test for the RMS steering angle error for the Negative Reinforcement condition $D(72) = 0.090, p \geq .2$ was not significantly non-normal. Thus, a non-parametric Kruskal-Wallis test is used to test differences between the distributions of the conditions. All

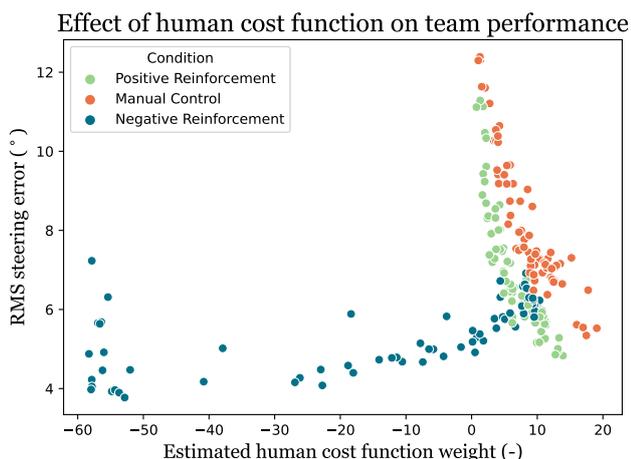
statistical results are found in table II. There exists a significant effect of the conditions on the RMS steering angle error. The RMS steering angle error is significantly lowest using the negative reinforcement adaptation strategy ($Mdn = 5.55^\circ$), followed by the positive reinforcement strategy ($Mdn = 6.47^\circ$). The significantly highest results are obtained in the manual control conditions ($Mdn = 7.45^\circ$).

B. Estimated control share

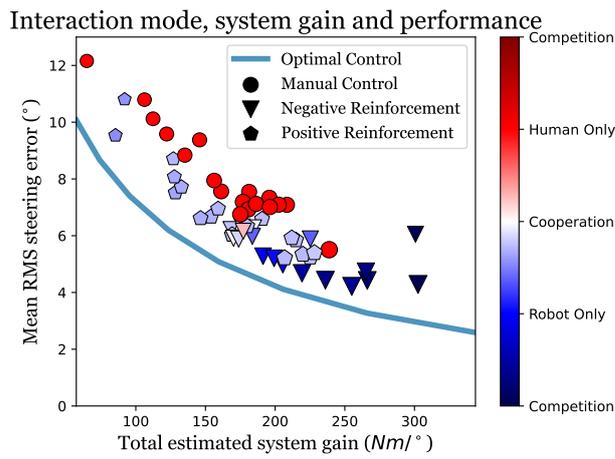
The Kolmogorov-Smirnov test for the median estimated control share for the Negative Reinforcement condition $D(72) = 0.147, p = .001$ and for the Positive Reinforcement condition $D(72) = 0.167, p \leq .001$ were both significantly non-normal. Thus, a non-parametric Mann-Whitney test is used to test differences between the distributions of the conditions. All statistical results are found in table II. There exists a significantly higher estimated control share using the positive reinforcement adaptation strategy ($Mdn = -0.281$), than using the negative reinforcement adaptation strategy ($Mdn = -1.03$).

C. Perceived control share

The Kolmogorov-Smirnov test for the perceived control share for the Negative Reinforcement condition $D(18) = 0.137, p \geq .2$ and for the Positive Reinforcement condition $D(18) = 0.175, p = .15$ were both not significantly non-normal. Thus an independent sample t-test was used to test the difference between the distributions of the conditions. All statistical results are found in table II. The perceived control share was significantly higher in the positive reinforcement



(a) Reducing control effort, leading to a lower estimated cost function weight, is rewarded with higher control performance in the negative reinforcement adaptation strategy. The positive reinforcement and manual control conditions exhibit the opposite phenomenon.



(b) Competition leads to the highest total system controller gain in the negative reinforcement adaptation strategy, which results in the highest control performance. The control performance and system controller gain in the positive reinforcement adaptation strategy is slightly increased compared to the manual control condition.

Figure 8: Reducing control share and/or changing control strategy is rewarded in the negative reinforcement strategy, leading to higher performance, whereas in manual control and positive reinforcement this is penalized leading to high human task engagement.

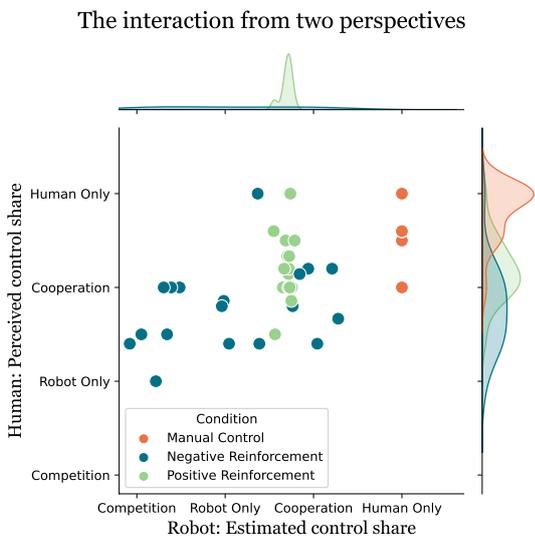


Figure 9: The perspectives of both the team members in the human-robot team are shown. The perspective of the robot is the *estimated control share* and the perspective of the human is the *perceived control share*. Interestingly, although these perspectives usually coincide, sometimes they do not.

condition ($Mdn = 0.187, SE = 0.078$) than in the negative reinforcement condition ($Mdn = -0.207, SE = 0.107$).

V. DISCUSSION

The objective of this paper was to design a robot that can adapt to a human to guide the co-adaptation of a human and a robot such that both human and team behavior are positively influenced. Two fundamental adaptation strategies were designed for the robot and in an experiment the effects of these adaptation strategies on human and team behavior were examined. Results show a significant increase in control performance for both the positive and negative reinforcement adaptation strategy compared to manual control, with the highest control performance in the negative reinforcement adaptation strategy. Results additionally show a significant decrease in both estimated (by the robot) and perceived (by the human) control share in the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy.

A. Spectrum of human and robot behavior

Each human showed to be very different from one another, each with their skill level, interaction strategy, and preferred role. Due to this variety of behaviors, the control algorithm also shows a large population of behaviors. As a consequence a large variety in the data is observed in fig. 6, fig. 9 and fig. 8. Despite these differences, some results can be generalized.

1) *Foreseen and unforeseen effects:* Co-adaptation leads to foreseen and unforeseen effects. The foreseen effects of the co-adaptation were a reduction of human control effort, captured as the human control strategy, and increased performance, shown in fig. 8a. The control task was also qualitatively rated

as being simpler when sharing control compared to manual control. Unforeseen effects are mainly observed in the negative reinforcement adaptation strategy. Figure 8a demonstrates that in the negative reinforcement adaptation strategy a reduction of the estimated human cost function weights results in higher team performance, therefore rewarding the human to reduce its control effort or adapt its control strategy. Figure 8b demonstrates that this reduction of cost function weights leads to a higher total system controller gain, leading to high performance. A negative cost function weight however leads the robot to think that the human is actively competing with the robot, increasing its controller gain such that the total system gain highly increases. This in turn causes the error signal, which is used to update the estimated human controller gain (see eq. (8)) to become very small, thus leading to an inability of the robot to again increase the estimated human controller gain. This can be seen as a design flaw of the control algorithm.

2) *The slacking hypothesis*: The slacking hypothesis for motor learning [11] states that when a robot delivers too much assistance it could potentially decrease motor learning if it encourages a decrease in motor output, effort, or attention. Figure 8a and fig. 8b show that in the negative reinforcement adaptation strategy, as the human decreases its control effort, this leads to higher robot assistance and higher performance, thus encouraging the human to decrease control effort. This confirms the slacking hypothesis for co-adaptation. Some humans, such as in fig. 7, noticed, usually unconsciously, that lowering control effort increased team performance, and a decrease in human control effort is observed over trials.

3) *Two perspectives*: Since there are two team members in the same game, cooperation is perceived from two perspectives. The perspective of the robot is measured by estimating the human control gains and calculating the robot control gains, and the perspective of the human is quantitatively established through the questionnaire. The metric that was used for these perspectives was the estimated (for the robot) and subjective (for the human) control share, indicating the mode of interaction between the human and the robot. Figure 9 shows that the perspectives usually correspond to one another. However, there seems to exist a tendency of the human to rate itself as doing more than the robot believes. Note that humans were not asked in the questionnaire whether they believed to be competing with the robot.

4) *Qualitative analysis*: After the experiments, participants were asked which adaptation strategies they liked best and which adaptation strategies for the robot resulted in the best cooperation. Interestingly, opinions varied a lot from participant to participant. Often, humans reported that the more subtle controllers were more pleasant. If a participant with a high skill level interacted with the negative reinforcement adaptation strategy controller, the amount of feedback by the controller would be fairly low, resulting in a liking towards this adaptation strategy. On the other hand, the participants with a slightly lower skill level clearly showed a liking toward the positive reinforcement adaptation strategy. For these

participants, the controllers with a negative reinforcement adaptation strategy were perceived as too dominant. An additional clear difference exists between participants in how much difference they felt between different conditions. Again, this difference is attributed to the skill level of the participants. Participants that naturally performed much lower in the control task usually noticed the negative reinforcement adaptation strategy to be very dominant. On the other hand, participants that naturally performed fairly well noticed a much lower difference between the conditions.

B. Implications

The control algorithm that was presented in this paper shows stabilizing behavior, whilst being able to adapt to a human, interacting with the control algorithm. Implicitly, the control algorithm allows to keep the human in the loop, whilst guaranteeing safety using optimal control. The negative reinforcement adaptation strategy has proven to yield the best overall team performance. Thus, in the situation where a human-robot team must perform a task, such that a certain level of performance is guaranteed, this strategy would be preferred. However, in terms of the cooperation of the robot with the human, this adaptation strategy may not deliver the desired results. It may in that case be even better to fully automate the task, as the slacking hypothesis [11] predicts that the human disappears from the control task. The positive reinforcement adaptation strategy on the other hand showed to increase performance and reduce control effort compared to the manual control condition while engaging the human to contribute to the control task. Moreover, this system was often rated as most favorable by the participants. However, from a safety perspective, this adaptation strategy might not always perform desirable, if the skill of the human operator is not adequate.

C. Recommendations

This paper builds on a few assumptions and has a limited research scope, which should be examined in future works. For further study, the following points are recommended:

1. Add visual preview to make the task more realistic towards car driving. This has implications for the control algorithm, however. Driver preview models such as the van der El [40] model can be used, modeling the human as a two degree of freedom controller, where the reference signal is low-pass filtered.

2. Trigger adaptation by either changing the environment, the robot, or the human to investigate co-adaptation behavior. In this paper, only co-adaptation toward some equilibrium is observed. By triggering some changes in either the environment or in one of the team members, leading to co-adaptation, changes in the equilibria may be observed.

3. Design of different adaptation strategies. In this paper only two fundamental adaptation strategies are designed, being a positive and negative reinforcement strategy. By combining these ideas with the findings in this research, adaptation

strategies can be designed leading to more desirable co-adaptation. Some ideas are lowering the robot cost when competition is detected, giving way to humans. Other ideas include designing the adaptation strategy as a hysteresis model, forcing the human to input effort into the system, but maintaining some level of control effort when the human reduces its effort to maintain system performance.

4. Include visual information about the interaction strategy of the robot, to make the interaction between the human and the robot more transparent. It may be questionable whether pure haptic feedback is enough information for the human to form a good mental model of the robot. Adding visual information may help improve this process for humans.

5. Consider other types of solutions for differential games which are built on other assumptions of human behavior. Na and Cole [41] examines these assumptions in simulation experiments, comparing between a Nash, Stackelberg, and Pareto equilibrium. There may be other types of solutions that may apply and are worth examining.

6. Research the time scale of strategic adaptation. The algorithm in this paper operates at a fairly high timescale, where the human controller gain is usually estimated to reach an equilibrium around 15s. Interaction on the strategic level may however be a much slower process. This is motivated by the trial-by-trial meta-adaptation reported in the negative reinforcement adaptation strategy visualized in fig. 7.

VI. CONCLUSION

This paper has successfully extended the line of work by Li et al. [21] by designing an adaptive robotic controller that jointly performs a compensatory tracking task while interacting and adapting to a human. Additionally, to the writer's best knowledge, a human-in-the-loop experiment was performed for the first time, examining how human behavior, team performance, and co-adaptation are affected by robot adaptation strategies. Using the human-in-the-loop experiment, this paper has investigated the following research question: How should a robot be designed to adapt to a human in a physically interacting human-robot team to guide the co-adaptation to maximize team performance and human task contribution?

To answer this question, two interaction strategies, positive and negative reinforcement, were designed and compared in a human-in-the-loop experiment. From the experiment data, summarized in table II, the following conclusions are drawn:

- 1) There exists a significant increase in control performance for both the negative reinforcement adaptation strategy and the positive reinforcement adaptation strategy, compared to manual control;
- 2) There exists a significant increase in control performance using the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy;
- 3) There exists a significant decrease in estimated (by the robot) and perceived (by the human) control share using

the negative reinforcement adaptation strategy compared to the positive reinforcement adaptation strategy.

In conclusion, to guide the co-adaptation to both increase performance and engage humans to actively contribute to a control task, a robot should be designed to adapt using a positive adaptation strategy.

REFERENCES

- [1] R. Bregman, *De meeste mensen deugen: een nieuwe geschiedenis van de mens.* de Correspondent, 2019.
- [2] Y. N. Harari, *Sapiens: A brief history of humankind.* Random House, 2014.
- [3] C. S. Burke, K. C. Stagl, E. Salas, L. Pierce, and D. Kendall, "Understanding team adaptation: A conceptual analysis and model," *Journal of Applied Psychology*, vol. 91, no. 6, pp. 1189–1207, 2006.
- [4] S. Ososky, D. Schuster, E. Phillips, and F. Jentsch, "Building appropriate trust in human-robot teams," *AAAI Spring Symposium - Technical Report*, vol. SS-13-07, pp. 60–65, 2013.
- [5] E. M. van Zoelen, K. van den Bosch, and M. Neerinx, "Becoming Team Members: Identifying Interaction Patterns of Mutual Adaptation for Human-Robot Co-Learning," *Frontiers in Robotics and AI*, vol. 8, no. July, pp. 1–17, 2021.
- [6] E. M. van Zoelen, K. van den Bosch, M. Rauterberg, E. Barakova, and M. Neerinx, "Identifying Interaction Patterns of Tangible Co-Adaptations in Human-Robot Team Behaviors," *Frontiers in Psychology*, vol. 12, no. July, pp. 1–16, 2021.
- [7] M. Marcano, S. Diaz, J. Perez, and E. Irigoyen, "A Review of Shared Control for Automated Vehicles: Theory and Applications," *IEEE Transactions on Human-Machine Systems*, pp. 1–17, 2020.
- [8] G. Boy, "Theories of Human Cognition: To Better Understand the Co-Adaptation of People and Technology," *Knowledge Management, Organizational Intelligence and Learning, and Complexity. In: Kiel, LD (ed.) Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford, UK*, vol. III, 2002.
- [9] A. Davies, P. Veličković, L. Buesing, S. Blackwell, D. Zheng, N. Tomašev, R. Tanburn, P. Battaglia, C. Blundell, A. Juhász, M. Lackenby, G. Williamson, D. Hassabis, and P. Kohli, "Advancing mathematics by guiding human intuition with AI." *Nature*, vol. 600, no. 7887, pp. 70–74, 2021.
- [10] V. A. Banks, K. L. Plant, and N. A. Stanton, "Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016," *Safety Science*, vol. 108, no. August 2017, pp. 278–285, 2018.
- [11] L. Marchal-Crespo and D. J. Reinkensmeyer, "Review of control strategies for robotic movement training after neurologic injury," *Journal of NeuroEngineering and Rehabilitation*, vol. 6, no. 1, 2009.

- [12] T. Melman, J. C. de Winter, and D. A. Abbink, "Does haptic steering guidance instigate speeding? A driving simulator study into causes and remedies," *Accident Analysis and Prevention*, vol. 98, pp. 372–387, 2017.
- [13] J. M. Galea, E. Mallia, J. Rothwell, and J. Diedrichsen, "The dissociable effects of punishment and reward on motor learning," *Nature Neuroscience*, vol. 18, no. 4, pp. 597–602, 2015.
- [14] L. Crespo and D. Reinkensmeyer, "Haptic guidance can enhance motor learning of a steering task," *Journal of Motor Behavior*, vol. 40, no. 6, pp. 545–557, 2008.
- [15] D. A. Abbink, M. Mulder, and E. R. Boer, "Haptic shared control: Smoothly shifting control authority?" *Cognition, Technology and Work*, vol. 14, no. 1, pp. 19–28, mar 2012.
- [16] D. A. Abbink, S. Member, T. Carlson, M. Mulder, J. C. F. D. Winter, F. Aminravan, T. L. Gibo, and E. R. Boer, "A Topology of Shared Control Systems — Finding Common Ground in Diversity," vol. 48, no. 5, pp. 509–525, 2018.
- [17] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature Neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002.
- [18] V. T. Chackochan and V. Sanguinetti, "Incomplete information about the partner affects the development of collaborative strategies in joint action," *PLoS Computational Biology*, vol. 15, no. 12, pp. 1–23, 2019.
- [19] X. Ji, K. Yang, X. Na, C. Lv, and Y. Liu, "Shared Steering Torque Control for Lane Change Assistance: A Stochastic Game-Theoretic Approach," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3093–3105, 2019.
- [20] N. Jarrassé, T. Charalambous, and E. Burdet, "A Framework to Describe, Analyze and Generate Interactive Motor Behaviors," *PLoS ONE*, vol. 7, no. 11, p. e49945, nov 2012.
- [21] Y. Li, G. Carboni, F. Gonzalez, D. Campolo, and E. Burdet, "Differential game theory for versatile physical human–robot interaction," *Nature Machine Intelligence*, vol. 1, no. 1, pp. 36–43, jan 2019.
- [22] J. F. Nash, "The Bargaining Problem," *Econometrica*, vol. 18, no. 2, p. 155, apr 1950.
- [23] D. Shen and B. Cruz Jose, "Adaptive state feedback nash strategies for linear quadratic discrete-time games," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 17, no. 1 PART 1, 2008.
- [24] K. Brown, K. Driggs-campbell, and M. J. Kochenderfer, "Modeling and Prediction of Human Driver Behavior : A Survey," 2020.
- [25] S. Kolekar, J. de Winter, and D. Abbink, "Human-like driving behaviour emerges from a risk-based driver model," *Nature Communications*, vol. 11, no. 1, pp. 1–19, dec 2020.
- [26] M. Flad, L. Frohlich, and S. Hohmann, "Cooperative shared control driver assistance systems based on motion primitives and differential games," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 5, pp. 711–722, 2017.
- [27] J. E. Vos, "Reducing conflicting forces with co-adaptive haptic shared control using online time-varying operator identification," 2018.
- [28] J. Engwerda, "Algorithms for computing Nash equilibria in deterministic LQ games," pp. 113–140, 2007.
- [29] G. Toa, *Adaptive control design and analysis*, 2003, vol. 37, no. 3.
- [30] SensoDrive GmbH, "Software Manual SENSO-Wheel."
- [31] V. Lesnikov, T. Naumovich, and A. Chastikov, "Topography of z-plane which is discretized due to quantization of coefficients of digital biquad filters," *2016 International Siberian Conference on Control and Communications, SIBCON 2016 - Proceedings*, no. 6, pp. 2014–2017, 2016.
- [32] T. Specker, M. Buchholz, and K. Dietmayer, "A new approach of dynamic friction modelling for simulation and observation," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 19, pp. 4523–4528, 2014.
- [33] J. J. McDowell, "Wilkinson'S Method of Estimating the Parameters of Herrnstein'S Hyperbola," *Journal of the Experimental Analysis of Behavior*, vol. 35, no. 3, pp. 413–414, 1981.
- [34] H. J. Damveld, G. C. Beerens, M. M. Van Paassen, and M. Mulder, "Design of forcing functions for the identification of human control behavior," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 4, pp. 1064–1081, 2010.
- [35] S. Mahata, S. K. Saha, R. Kar, and D. Mandal, "Optimal design of fractional order low pass Butterworth filter with accurate magnitude response," *Digital Signal Processing: A Review Journal*, vol. 72, pp. 96–114, 2018.
- [36] P. Guillaume, J. Schoukens, R. Pintelon, and I. Kollái, "Crest-Factor Minimization Using Nonlinear Chebyshev Approximation Methods," *IEEE Transactions on Instrumentation and Measurement*, vol. 40, no. 6, pp. 982–989, 1991.
- [37] J. De Winter and D. Dodou, *Applied Sciences and Technology, Human Subject Research For Engineers, A Practical Guide*, 2017, no. May.
- [38] A. Field, *Discovering Statistics Using SPSS ISM (London, England) Introducing Statistical Methods Series*, 2009.
- [39] IBM Corp., "IBM SPSS Statistics for Windows," 2017. [Online]. Available: <https://hadoop.apache.org>
- [40] K. van der El, "How humans use preview information in manual control," Ph.D. dissertation, 2018.
- [41] X. Na and D. J. Cole, "Linear quadratic game and non-cooperative predictive methods for potential application to modelling driver-AFS interactive steering control," *Vehicle System Dynamics*, vol. 51, no. 2, pp. 165–198, 2013.

II

Appendices



Background

The [World Health Organization](#) estimates that annually over 1.2 million people die in road traffic accidents [47]. To put this in perspective, that is more than twice the number of deaths from war, crime, and terrorism, combined [19]. An additional 50 million are injured each year due to traffic accidents. Unless this trend is stopped, by 2030 road traffic accidents will be the fifth leading cause of death in the world. It is estimated that human driver-related factors (fatigue, inattention, etc.) account for 67% to as much as 90% of the road traffic accidents [11, 42, 44].

To prevent these human errors, safety measures have been implemented in vehicles over the last decades in the form of *Advanced Driver Assistance Systems* (ADAS). Advanced Driver Assistance Systems are defined as “*electronic systems that are designed to support the driver in his/her driving task*” [27]. Examples of ADAS are adaptive cruise control and lane assist systems. The deployment of vehicles with increased automation correlates with a decrease in fatal road traffic accidents [25].

However, recent accidents with vehicles with a high degree of automation motivate that full autonomy of intelligent vehicles is not expected to happen in the near future [32]. For example, an accident in the 2007 DARPA Urban Challenge shows that tasks, such as detecting traffic lights, which humans perform naturally, suddenly become very hard technical problems [6, 15]. These technically complex challenges are to be overcome before certain aspects of the driving tasks are to be automated. Moreover, adding more technology does not remove the possibility of human error, but it changes or relocates it [10]. In 2016 a Tesla Model S collided with a tractor-trailer, causing a fatal accident to the driver of the Tesla. The accident was a result of the failure of the human driver to continually monitor the highly reliable Autopilot system [3].

A.1. Haptic Shared Control

An attractive approach to bridge the transition from assisted control to automatic control, often dubbed the uncanny and unsafe valley of assistance and automation [14], is haptic shared control. In haptic shared control, “*both the human and robot exert forces on the control interface, of which its output is the direct input to the controlled system*” [2]. Multiple studies have shown that the use of haptic shared control leads to reduced workload and increased performance compared to manual control [32, 37]. Shared control additionally introduces the benefit of redundancy gain [9], where if one controller fails, the other can still operate the system.

However, despite these advantages, there exist some risk factors that must be addressed in the design of the shared control system. Itoh et al. [20] state that conflicting torques can result in annoyance, discomfort, and possibly dangerous situations. De Winter and Dodou [9] state that either the human or the shared control system may potentially disengage due to impaired vigilance or due to failure or design flaws of components of the shared control system. These potential risks are to be addressed in the design of the haptic shared control system.

In haptic shared control, a human and a robot work together as a team. It is the designer’s task to design the robot such that the team optimally performs. To establish an effective team, the human and the robot must:

1. share a common goal;
2. be able to communicate;

3. be able to adapt to one another.

By design, haptic shared control covers the first two points, e.g. in a haptic car lane following system the human and the robot are tasked to follow the car lane and can communicate using forces through the control interface. The last point is usually not explicitly incorporated in the design of haptic shared control systems, although human behavioral adaptation [33] is usually observed in experimental studies.

A.2. Co-adaptation

In Boink et al. [5] a very simple control architecture for the automated vehicle was chosen in a haptic shared control curve negotiation task. Results show a lot of force conflicts, even up to a magnitude of five times as large as in manual control. Individualization of the control algorithm additionally failed and force conflicts could not be reduced. Participants were reported to cut corners and often kept to their preferred trajectory as in manual control, even to the expense of higher interaction forces. This demonstrates that if the decision-making of the human and the automated vehicle are not properly aligned, this leads to the disuse of the haptic shared control system.

Based on these findings, Scholtens et al. [39], van Paassen et al. [45] proposed and validated a more sophisticated control architecture, which uses a *human-compatible reference trajectory*. This control algorithm uses a pre-computed reference trajectory based on a human driver model. Results show that compared to a simpler control algorithm without a pre-computed reference trajectory [1], force conflicts are highly reduced in both frequency and magnitude. Although the results seem promising, force conflicts were mainly reduced when the human accepted the pre-computed reference trajectory, which is a person-dependent preference. Especially right curves were rejected, leading to a large adaptation of trajectories. Another study on haptic shared control using a driver model to generate feedback torques focussed on model parameter uncertainty [38]. It was demonstrated that when comparing the identified driver model to the human driver behavior, human driver adaptation to the steering assistance system was observed, which may influence the driver model parameters.

To model human behavioral adaptation, a game-theoretical approach is often used. Game theory [35] assumes that multiple decision-makers in an environment influence each other's actions, each optimizing their selfish objectives. Although this recent approach to haptic shared control has mainly been studied in simulation experiments [16, 21, 34], some experimental studies have been performed [13, 23]. Experimental results show a high reduction of force conflicts, especially in corner-cutting, between the human and the automated vehicle. According to Ji et al. [23], the human driver's decision-making process in the shared control task can be roughly described by the game-theoretical strategy. However, a limitation to these studies is their assumption that the human control strategy, describing the human's preferences, skill, and attitude towards the automated vehicle, is unchanging and known beforehand. Using these strict assumptions, adaptation actually cannot be captured by the shared control system.

To this end, some works have presented shared control systems that can estimate the human control input in real-time, and adapt the robot control input to this estimation [26, 46]. In Vos [46], an Extended Kalman Filter is used to estimate feedback and neuromuscular parameters. Results show a reduction of 19% in conflicting forces. However, in this work, the effect of the forces from the controller was not taken into account in the control architecture, leading to a poor model fit. Li et al. [26] presents an algorithm that can estimate a human's control strategy in real-time and adapt a robot's control strategy accordingly, through a control effort sharing rule. This way, co-adaptation is facilitated on a strategic level.

A.3. Interaction Strategies

Usually in human-robot interaction, robots are designed to take on a specific pre-defined role [22]. Traditionally, robots are designed to passively follow the human operator [17]. The latest years have seen an increase in robots that are meant to be partners instead of pure machines. Not much is known about how robots should be designed to cooperate well with humans, or what it means to have good co-operation.

This question is addressed in Chackochan and Sanguinetti [7]. In a human-human interaction experiment, where some uncertainty of the partner's action was presented to participants. Different groups of two humans were provided with different amounts of information about the partner: haptics only, visio-haptic (where the interaction force was also presented), and partner visible (where additionally, the position of the partner was visible). Results showed, in combination with simulations based on a mathematical model based on game theory, that for different amounts of information about the partner, different strategies existed. Incomplete

information about the partner caused an asymmetry in the roles the players adapted, meaning that a leader-follower relationship emerged. However, when the information was more reliable, this effect faded and the interaction is best described by a Nash equilibrium. Moreover, Chackochan and Sanguineti [7] describes the existence of an asymmetry in the roles of the partners as a non-optimal collaboration.

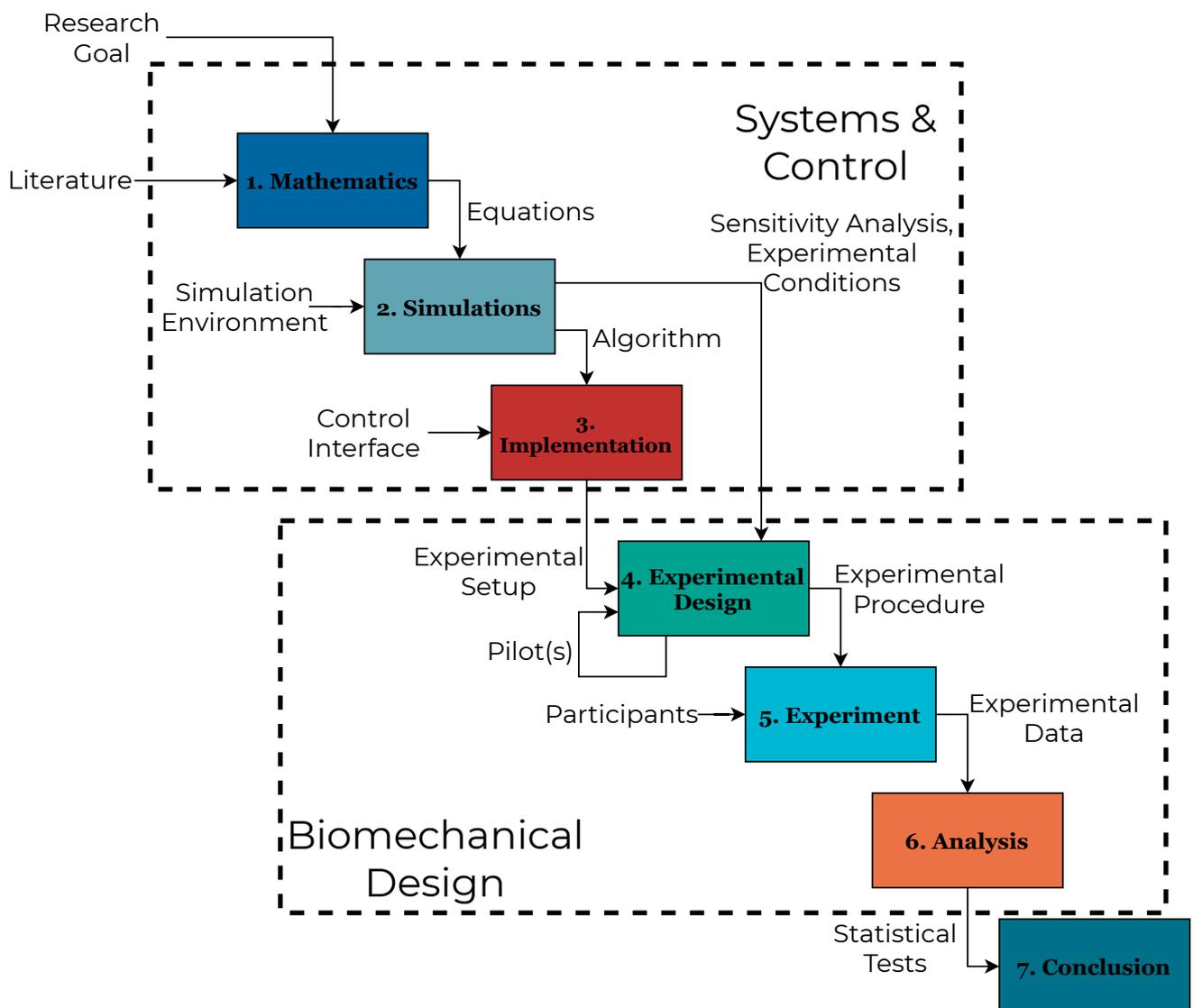
However, other works demonstrate that when there exists a difference in skill, the existence of a leader-follower relation can lead to better performance. In Noy et al. [36] an experiment with pairs of two humans is performed, based on the mirror game, where players imitate each other. Two experimental conditions were tested, one where one player was assigned leader and the other follower, and another with joint-improvisation where no leader was selected. Both participants that had expert experience in improvisation and those that had no prior experience in improvisational arts were selected. Results showed that experts were significantly more precise and scored higher performance when no leader was selected. When specific roles were selected, a significant loss of precision and performance was observed. However, novices were observed to obtain lower performance than the experts, especially when no roles were designated.

Groten et al. [17] similarly demonstrated person-dependent differences in the dominance behavior that participants showed. In a dynamic tracking task, controlling the position of a virtual mass, two participants jointly controlled the position of the virtual mass. Results showed that participants preferred to work with a dominance difference, indicating the spontaneous emergence of roles. The dominance behavior was not only found to be person-dependent but also showed consistency, meaning that humans tend to maintain their roles.

In the experiment of Mörtl et al. [31], the interaction between a human and a physical robot was considered. The human-robot team was tasked with jointly moving a table to four parking configurations. The interaction between the human and the robot was shaped using an 'effort sharing policy', which indicated how the roles were distributed. Three different role allocation strategies were considered, being constant role allocation, continuous role allocation, and discrete role allocation. Results showed a statistically significant increase in completion time for the continuous role allocation compared to the constant and discrete role allocation. However, subjective scores showed that the perceived effort was higher and the predictability was lower for the continuous role allocation.

B

Planning



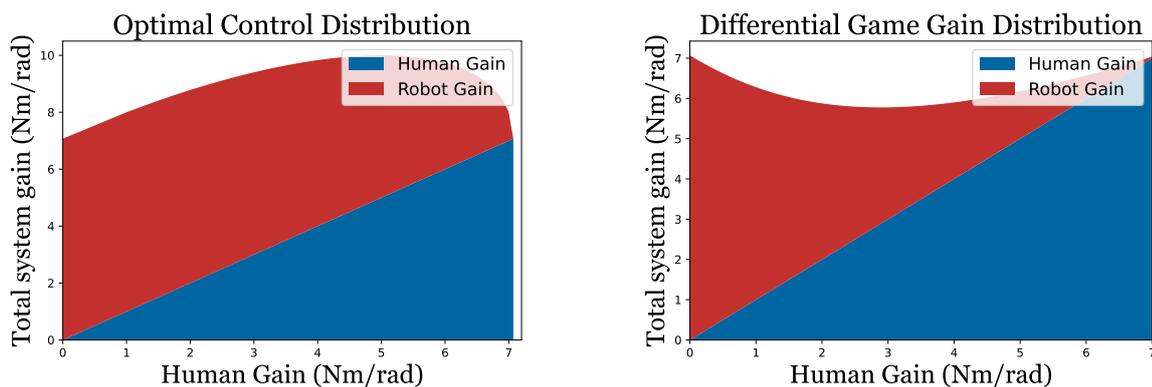
C

Optimal Control Versus Differential Game

This appendix compares optimal control and differential game control solutions, using examples to illustrate these differences. The fundamental difference between optimal control and differential game control is how the controller gains are related to the control strategy. In optimal control, the controller gains are computed to stabilize the system. Differential game control however assumes multiple *players* to control the same system and therefore takes these other players into account when delivering feedback. Therefore, optimal control considers the Riccati equation to compute the feedback controller gains, whereas in differential game solutions the double Riccati equation is solved, which incorporates the controller gains of the other players. In Figure C.1 this difference is clearly visualized. In the figures, the total system cost function weights are kept constant as:

$$Q_h + Q_r = C, \quad (C.1)$$

where Q_h, Q_r are the human and robot cost function weight matrices. The logical similarity between the distribution of controller gains is that when $Q_r = \mathbf{0}$ or $Q_h = \mathbf{0}$, the controller gains are computed the same for optimal and differential game control. The main difference is that when the cost weights are distributed between players, the total system gain of the optimal control controllers increases, whereas the differential game control solution yields a total system gain that remains approximately constant or even decreases. This is easily explained since in optimal control, now two players are trying to stabilize the total system. Whereas, in the differential game solution, the controller gains are determined given that the other player is now also controlling the same system, thus distributing the controller gains. In Figure C.1 the difference is easily seen.

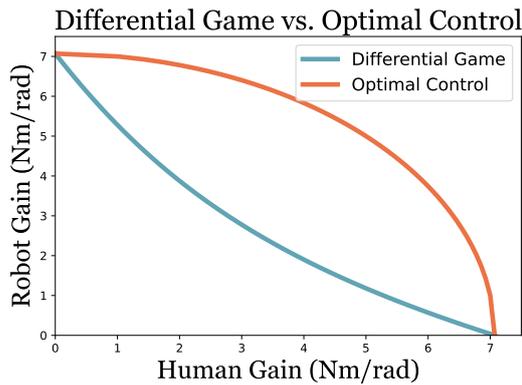


(a) Distribution of controller gains in optimal control.

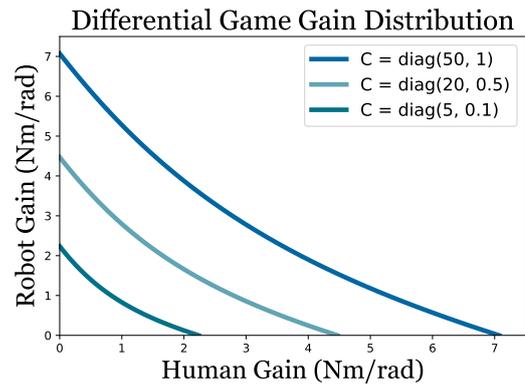
(b) Distribution of controller gains in differential game control.

Figure C.1: In the optimal control solution, the total controller gains increase due to the stabilizing character of the solution. In differential games, the load of stabilizing is distributed, thus lowering the total controller gains.

Figure C.2a visualizes the differences in one figure. Figure C.2b shows isolines, visualizing how or different sharing rules, the controller gains are distributed.



(a) Distribution of controller gains in optimal control versus in differential game control.



(b) Distribution of controller gains in differential game control for different sharing values.

Figure C.2: Comparison of control input responses for different *players* in different control architectures

Some conclusions can be drawn:

- Differential game acknowledges that two controllers are jointly stabilizing the system instead of two optimal controllers both stabilizing the system. In the latter case, in an extreme situation, the system could be destabilized due to overshooting as a consequence of the combined actions.
- The differential game solution is less robust to changes by *another player* since the gains are lower. Imagine one player disconnecting, whereas the other expects the other player to jointly control the system, this may lead to destabilizing the system.
- Both control architectures are based on different assumptions, where for human-robot interactions the assumptions of the differential game may describe the system somewhat more accurately. However, other types of solutions must not be overlooked and may describe the joint system even more accurately.

D

Mathematical Proofs

D.1. Stability proof

Let the following observer equations be defined as in Equation 7 in the paper:

$$\begin{aligned}\dot{\xi}(t) &= A\xi(t) + B(u_h(t) + u_r(t)), \\ \dot{\hat{\xi}}(t) &= A\hat{\xi}(t) + B(\hat{u}_h(t) + u_r(t)) - \Gamma\tilde{\xi}(t), \\ \dot{\tilde{\xi}}(t) &= (A - \Gamma)\tilde{\xi}(t) + B\tilde{u}_h(t).\end{aligned}\tag{D.1}$$

Normalized Gradient Descent Observer (Theorem)

Theorem D.1. Let the symmetric gain matrix $K = K^T > \mathbf{0} \in \mathbb{R}^{2 \times 2}$, design parameter $\kappa > 0$ and normalization parameter $m(t) = \sqrt{1 + \kappa \xi^T(t)\xi(t)}$ be given. Then, the estimated human controller input \hat{u}_h , estimated using the following update-law

$$\dot{\hat{L}}_h(t) = K \frac{\xi^T(t)\tilde{u}_h(t)}{m^2(t)},\tag{D.2}$$

converges to the real human controller input u_h and stabilizes the system described in Equation (D.1).

Proof. As in Toa [43], the derivative of $\hat{L}_h(t)$ is chosen, in the steepest descent direction, to minimize a normalized quadratic cost function. This normalized quadratic cost function is chosen as

$$J(\hat{L}_h) = \frac{\tilde{u}_h^2(t)}{2m^2(t)} = \frac{(\hat{L}_h\xi(t) - u_h(t))^2}{2m^2(t)},\tag{D.3}$$

where $m(t) = \sqrt{1 + \kappa \xi^T(t)\xi(t)}$. The steepest descent direction is given by the derivative with respect to \hat{L}_h :

$$\frac{\partial J(\hat{L}_h)}{\partial \hat{L}_h} = \frac{\partial}{\partial \hat{L}_h} \frac{(\hat{L}_h\xi(t) - u_h(t))^2}{2m^2(t)} = \frac{\xi^T(t)\tilde{u}_h(t)}{m^2(t)}.\tag{D.4}$$

Then, using a symmetric positive definite gain matrix $K = K^T > \mathbf{0} \in \mathbb{R}^{2 \times 2}$, the following update rule for the estimated human gains is established:

$$\dot{\hat{L}}_h(t) = K \frac{\xi^T(t)\tilde{u}_h(t)}{m^2(t)},\tag{D.5}$$

where $\tilde{u}_h(t)$ is deconstructed from the observation error in Equation (D.1), using the pseudo-inverse of matrix B in the following way:

$$\tilde{u}_h(t) = (B^T B)^{-1} B^T (\dot{\hat{\xi}}(t) - (A - \Gamma)\tilde{\xi}(t)).\tag{D.6}$$

Now, define the following Lyapunov candidate $V(\tilde{L}_h)$ be defined as:

$$V(\tilde{L}_h) = \tilde{L}_h K^{-1} \tilde{L}_h^T,\tag{D.7}$$

where $\tilde{L}_h = \hat{L}_h - L_h$ is defined as the gain estimation error. Assuming that a static human gain ($\dot{L}_h = \mathbf{0}$) is estimated, meaning that $\dot{\tilde{L}}_h = \dot{\hat{L}}_h$, the derivative with respect to time of this Lyapunov candidate is computed as:

$$\frac{d}{dt}V(\tilde{L}_h) = \frac{\partial V(\tilde{L}_h)}{\partial \tilde{L}_h} \frac{\partial \tilde{L}_h}{\partial t} = \frac{2\tilde{L}_h \xi(t) \tilde{u}_h(t)}{m^2(t)} = -\frac{2\tilde{u}_h^2(t)}{m^2(t)}. \quad (\text{D.8})$$

This shows that, since the Lyapunov candidate is monotonically decreasing, using the update law, $\tilde{u}_h(t)$ asymptotically converges to zero ($t \rightarrow \infty$, $\tilde{u}_h(t) \rightarrow 0$). To show state stability, again a Lyapunov candidate is given by $V(\xi(t)) = \frac{1}{2}\xi^T(t)\xi(t)$. The time derivative of this Lyapunov candidate is calculated as:

$$\begin{aligned} \frac{d}{dt}V(\xi(t)) &= \frac{\partial V(\xi(t))}{\partial \xi(t)} \frac{\partial \xi(t)}{\partial t} = \xi^T(t)\dot{\xi}(t) = \xi^T(t)(A - BL_r - BL_h)\xi(t), \\ &= \xi^T(t)(A - BL_r - B\hat{L}_h + B\tilde{L}_h)\xi(t), \\ &= \xi^T(t)(A - BL_r - B\hat{L}_h)\xi(t) + \xi^T(t)B\tilde{L}_h\xi(t), \\ &= \xi^T(t)(A - BL_r - B\hat{L}_h)\xi(t) - \xi^T(t)B\tilde{u}_h(t). \end{aligned} \quad (\text{D.9})$$

As a property of the Nash Equilibrium equations, the robot's controller gain L_r is computed such that the closed-loop system $A - BL_r - B\hat{L}_h$ is stabilized, meaning that $A - BL_r - B\hat{L}_h \leq \mathbf{0}$. Then, let λ_{max} be defined as the maximal eigenvalue of the closed-loop system $A - BL_r - B\hat{L}_h$. Then, the following inequality holds:

$$\begin{aligned} \frac{d}{dt}V(\xi(t)) &\leq \lambda_{max}|\xi(t)|^2 - \xi^T(t)B\tilde{u}_h(t), \\ &\leq \lambda_{max}|\xi(t)|^2 + |\xi^T(t)||B\tilde{u}_h(t)|. \end{aligned} \quad (\text{D.10})$$

Now, following Young's inequality [4] with ε , for $\varepsilon > 0$ the following inequality holds [30]:

$$\begin{aligned} \frac{d}{dt}V(\xi(t)) &\leq \lambda_{max}|\xi(t)|^2 + \frac{\varepsilon}{2}|\xi^T(t)|^2 + \frac{1}{2\varepsilon}|B\tilde{u}_h(t)|^2, \\ &= (\lambda_{max} + \frac{\varepsilon}{2})|\xi(t)|^2 + \frac{1}{2\varepsilon}|B\tilde{u}_h(t)|^2. \end{aligned} \quad (\text{D.11})$$

Then, choosing ε such that $(\lambda_{max} + \frac{\varepsilon}{2}) < 0$, it is deduced that the system is input-to-state stable (p. 107, Theorem 4.6 in Khalil [24]). From Equation (D.8) it followed that as $t \rightarrow \infty$, $\tilde{u}_h(t) \rightarrow 0$. From definition of input-to-state stability, for every bounded $\tilde{u}_h(t)$, the error states $\xi(t)$ are bounded. Additionally, as $\tilde{u}_h(t)$ asymptotically decreases to 0, this means that by definition $t \rightarrow \infty$, $\xi(t) \rightarrow 0$, stabilizing the system. This concludes the proof. \square

D.2. Estimating the human cost function

Let the coupled Riccati equations [12] be given by:

$$\begin{aligned} \mathbf{0}_n &= (A - BB^T P_r)^T \hat{P}_h + \hat{P}_h (A - BB^T P_r) - \hat{P}_h BB^T \hat{P}_h + \hat{Q}_h \\ \mathbf{0}_n &= (A - BB^T \hat{P}_h)^T P_r + P_r (A - BB^T \hat{P}_h) - P_r BB^T P_r + Q_r \end{aligned} \quad (\text{D.12})$$

Cost Function Observer (Theorem)

Theorem D.2. Define system matrices $A \in \mathbb{R}^{2 \times 2}$, $B \in \mathbb{R}^{2 \times 1}$, estimated human and robot cost function weights $\hat{Q}_h, Q_r \in \mathbb{R}^{2 \times 2}$, estimated human and robot controller gains $\hat{L}_h, L_r \in \mathbb{R}^{1 \times 2}$ according to the observer equation in Equation (D.1) as

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{D}{I} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad Q_h = \begin{bmatrix} q_{h,1} & 0 \\ 0 & q_{h,2} \end{bmatrix}, \quad \hat{L}_h = [\hat{L}_{h,1} \quad \hat{L}_{h,2}], \quad L_r = [L_{r,1} \quad L_{r,2}], \quad (\text{D.13})$$

where the robot's controller gains are computed as $L_r = B^T P_r$ by solving the following single Riccati equation (from Equation (D.12)) for P_r :

$$\mathbf{0}_n = (A - B\hat{L}_h)^T P_r + P_r (A - B\hat{L}_h) - P_r BB^T P_r + Q_r. \quad (\text{D.14})$$

Then, the human cost function is determined similarly by solving the equations in Equation (D.12) for \hat{Q}_h using the previous result:

$$\hat{Q}_h(t) = \begin{bmatrix} \hat{L}_{h,1}(\hat{L}_{h,1} + 2(K + L_{r,1})) & 0 \\ 0 & \hat{L}_{h,2}(\hat{L}_{h,2} - 2(D + L_{r,2})) - 2I\hat{L}_{h,1} \end{bmatrix}. \quad (\text{D.15})$$

Proof. The solution for the Nash equilibrium in a feedback game is given by the solution of the coupled Riccati equations in Equation (D.12). The solution for optimal gains is given by $L_h = B^T P_h$, $L_r = B^T P_r$. The estimated human gain \hat{L}_h is similarly computed as $\hat{L}_h = B^T \hat{P}_h$. \hat{P}_h is a symmetric matrix and thus the gains are computed as:

$$\hat{L}_h = [0 \quad \frac{1}{I}] \begin{bmatrix} \hat{a}_h & \hat{b}_h \\ \hat{b}_h & \hat{c}_h \end{bmatrix} = \frac{1}{I} [\hat{b}_h \quad \hat{c}_h]. \quad (\text{D.16})$$

One can exploit the structure of Q_h , which must by definition be a symmetric matrix. It follows from Equation (D.12) that the following equation must hold:

$$\begin{aligned} \hat{Q}_h &= -(A - BB^T P_r)^T \hat{P}_h - \hat{P}_h (A - BB^T P_r) + \hat{P}_h BB^T \hat{P}_h \\ &= -(A - BL_r)^T \hat{P}_h - \hat{P}_h (A - BL_r) + \hat{L}_h^T \hat{L}_h \end{aligned} \quad (\text{D.17})$$

Filling in using Equation (D.13), this yields for \hat{Q}_h :

$$\begin{bmatrix} \hat{q}_{h,1} & 0 \\ 0 & \hat{q}_{h,2} \end{bmatrix} = - \begin{bmatrix} 0 & \alpha_1 \\ 1 & \alpha_2 \end{bmatrix} \begin{bmatrix} \hat{a}_h & \hat{b}_h \\ \hat{b}_h & \hat{c}_h \end{bmatrix} - \begin{bmatrix} \hat{a}_h & \hat{b}_h \\ \hat{b}_h & \hat{c}_h \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \alpha_1 & \alpha_2 \end{bmatrix} + \begin{bmatrix} \hat{L}_{h,1} \\ \hat{L}_{h,2} \end{bmatrix} \begin{bmatrix} \hat{L}_{h,1} & \hat{L}_{h,2} \end{bmatrix}, \quad (\text{D.18})$$

where $\alpha_1 = \frac{-K - L_{r,1}}{I}$ and $\alpha_2 = \frac{-D - L_{r,2}}{I}$. Writing this out gives:

$$\begin{bmatrix} \hat{q}_{h,1} & 0 \\ 0 & \hat{q}_{h,2} \end{bmatrix} = \begin{bmatrix} \hat{L}_{h,1}^2 - 2\hat{b}_h \alpha_1 & \hat{L}_{h,1} \hat{L}_{h,2} - \hat{a}_h + \hat{b}_h \alpha_2 - \hat{c}_h \alpha_1 \\ \hat{L}_{h,1} \hat{L}_{h,2} - \hat{a}_h + \hat{b}_h \alpha_2 - \hat{c}_h \alpha_1 & \hat{L}_{h,2}^2 + 2\hat{c}_h \alpha_2 - 2\hat{b}_h \end{bmatrix} \quad (\text{D.19})$$

Filling in (from Equation (D.16)) that $\hat{b}_h = I\hat{L}_{h,1}$ and $\hat{c}_h = I\hat{L}_{h,2}$, combined with filling in α_1 and α_2 , yields:

$$\begin{bmatrix} \hat{q}_{h,1} \\ \hat{q}_{h,2} \end{bmatrix} = \begin{bmatrix} \hat{L}_{h,1}(\hat{L}_{h,1} - 2(-K - L_{r,1})) \\ \hat{L}_{h,2}(\hat{L}_{h,2} + 2(-D - L_{r,2})) - 2I\hat{L}_{h,1} \end{bmatrix} \quad (\text{D.20})$$

And consequently:

$$\hat{a}_h = \hat{L}_{h,1} \hat{L}_{h,2} + \hat{L}_{h,1}(-D - L_{r,2}) - \hat{L}_{h,2}(-K - L_{r,1}) \quad (\text{D.21})$$

This concludes the proof. \square

Simulation Results

In this appendix, two simulation examples are presented. The first example presents the situation where the robot control strategy (captured in Q_r) is constant ($Q_r = c$). The second example considers the situation where the robot control strategy changes based on the estimated human cost function, using the following adaptation strategy:

$$Q_r + \hat{Q}_h = C, \quad (\text{E.1})$$

where C is a sharing variable, prescribing the amount of effort that is shared by both team members. Results show that both examples converge to the same solution, validating the controller theoretically.

Simulation (Example)

Example E.1. Simulation example with fixed robot cost function weights

Given a simple 1D linear tracking task, tracking a reference state $\mathbf{r}(t) \in \mathbb{R}^2$ of a mass with damping only, with acceleration control $\mathbf{u}(t) \in \mathbb{R}$, the states $\mathbf{x}(t) \in \mathbb{R}^2$ and error states $\boldsymbol{\xi}(t) \in \mathbb{R}^2$ system matrices $A \in \mathbb{R}^{2 \times 2}$ and $B \in \mathbb{R}^{2 \times 1}$ and true cost matrix $Q_h \in \mathbb{R}^{2 \times 2}$ and sharing rule $C \in \mathbb{R}^{2 \times 2}$ are given as:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -10 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 20 \end{bmatrix}, \quad Q_h = \begin{bmatrix} 10 & 0 \\ 0 & 0.4 \end{bmatrix}, \quad Q_r = \begin{bmatrix} 15 & 0 \\ 0 & 0.6 \end{bmatrix} \quad (\text{E.2})$$

For a reference position $x_r(t) = 0.5 \cos(\frac{2\pi}{5} t)$, the reference vector is given by $\mathbf{r}(t) = [x_r(t) \quad \dot{x}_r(t)]^T$. The following matrices are chosen the following way:

$$\hat{Q}_h = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{\mathbf{x}} = \tilde{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}, \quad K = \begin{bmatrix} 300 & 0 \\ 0 & 50 \end{bmatrix} \quad (\text{E.3})$$

Such that $A - \Gamma$ is chosen negative definite.

The true states $\mathbf{x}(t)$ and their estimates $\hat{\mathbf{x}}(t)$ are then obtained as follows:

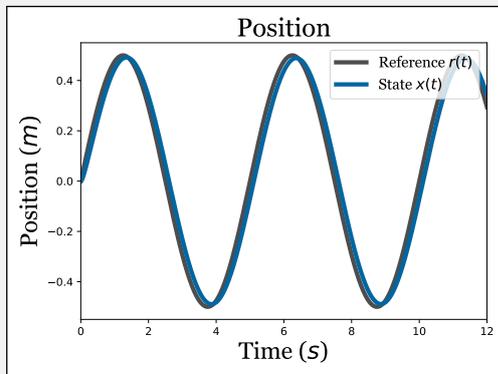


Figure E.1: Position tracking performance.

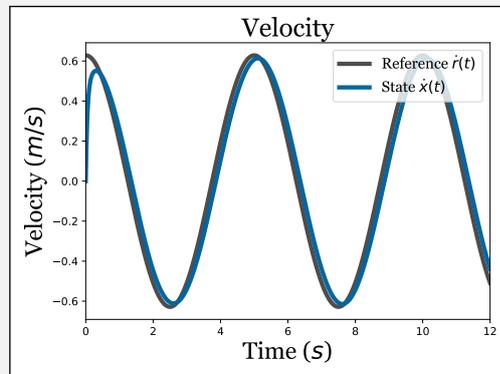


Figure E.2: Velocity tracking performance.

Cost and gain for position:

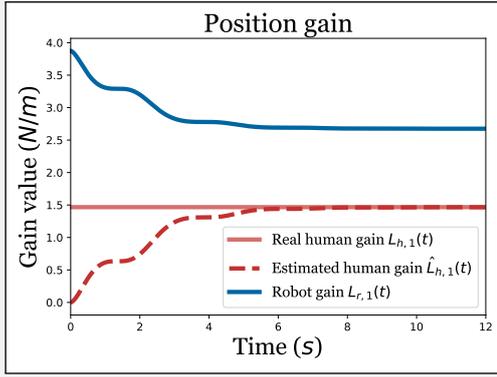


Figure E.3: Position controller gain estimation.

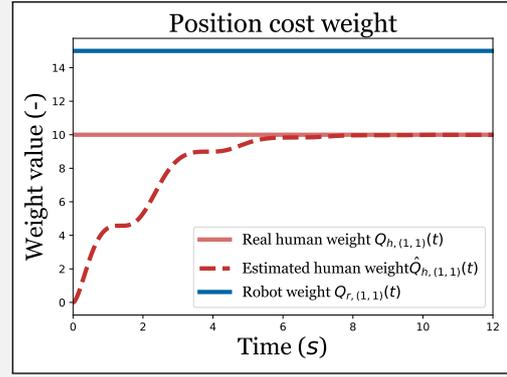


Figure E.4: Position cost function weight estimation.

Cost and gain for velocity:

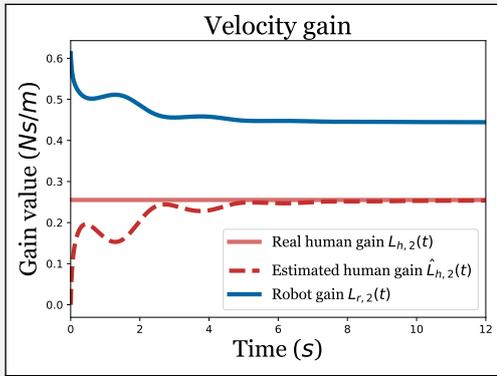


Figure E.5: Velocity controller gain estimation.

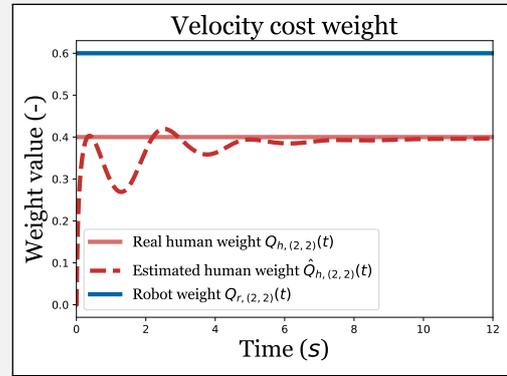


Figure E.6: Velocity cost function weight estimation.

Simulation (Example)

Example E.2. Simulation example with varying robot cost function weights

Given a simple 1D linear tracking task, tracking a reference state $\mathbf{r}(t) \in \mathbb{R}^2$ of a mass with damping only, with acceleration control $\mathbf{u}(t) \in \mathbb{R}$, the states $\mathbf{x}(t) \in \mathbb{R}^2$ and error states $\boldsymbol{\xi}(t) \in \mathbb{R}^2$ system matrices $A \in \mathbb{R}^{2 \times 2}$ and $B \in \mathbb{R}^{2 \times 1}$ and true cost matrix $Q_h \in \mathbb{R}^{2 \times 2}$ and sharing rule $C \in \mathbb{R}^{2 \times 2}$ are given as:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -10 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 20 \end{bmatrix}, Q_h = \begin{bmatrix} 10 & 0 \\ 0 & 0.4 \end{bmatrix}, C = \begin{bmatrix} 25 & 0 \\ 0 & 1 \end{bmatrix}, Q_r = C - Q_h \quad (\text{E.4})$$

For a reference position $x_r(t) = 0.5 \cos(\frac{2\pi}{5} t)$, the reference vector is given by $\mathbf{r}(t) = [x_r(t) \quad \dot{x}_r(t)]^T$. The following matrices are chosen the following way:

$$\hat{Q}_h = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \hat{\mathbf{x}} = \tilde{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Gamma = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}, K = \begin{bmatrix} 300 & 0 \\ 0 & 50 \end{bmatrix} \quad (\text{E.5})$$

Such that $A - \Gamma$ is chosen negative definite.

The true states $\mathbf{x}(t)$ and their estimates $\hat{\mathbf{x}}(t)$ are then obtained as follows:

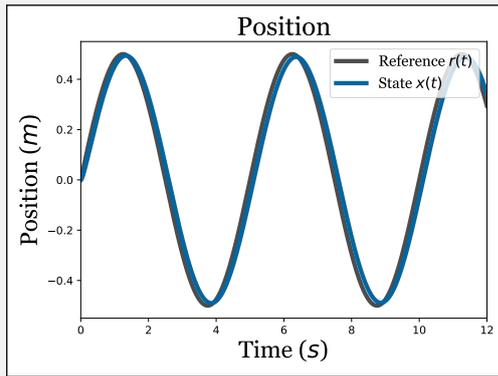


Figure E.7: Position tracking performance.

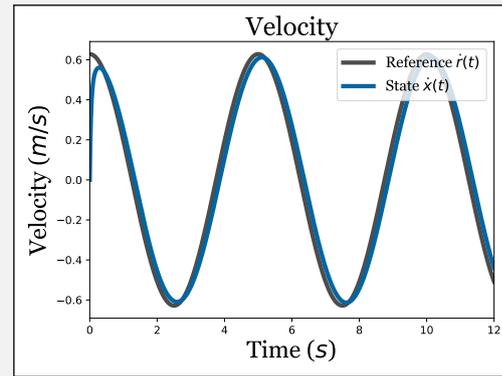


Figure E.8: Velocity tracking performance.

Cost and gain for position:

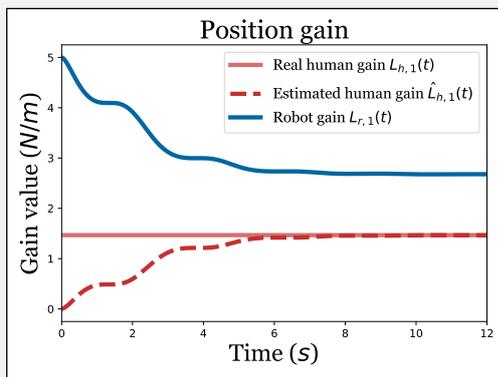


Figure E.9: Position controller gain estimation.

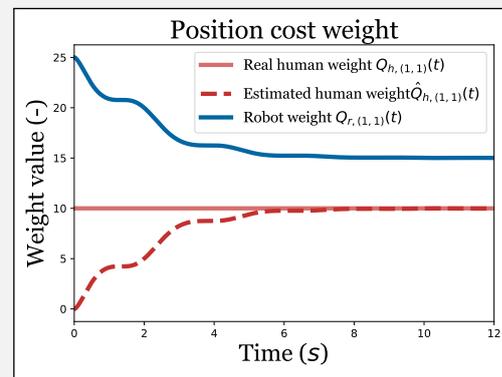


Figure E.10: Position cost function weight estimation.

Cost and gain for velocity:

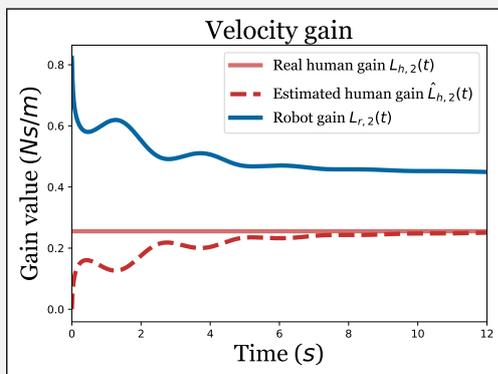


Figure E.11: Velocity controller gain estimation.

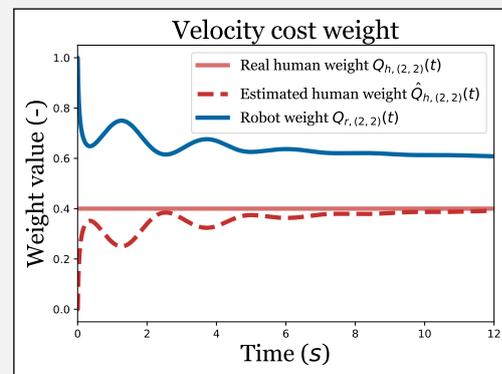


Figure E.12: Velocity cost function weight estimation.

F

Comparison with state-of-the-art algorithm

This appendix compares the proposed control algorithm of this research with the state-of-the-art in the literature. The pioneering work by Li et al. [26] considers the same control architecture as this thesis, however, the update rule for the estimated human controller gain \hat{L}_h is established constructively using a Lyapunov candidate. The stability proof is summarized in the following section.

F.1. Stability proof

Let the following observer equations be given:

$$\begin{aligned}\dot{\hat{\xi}}(t) &= A\hat{\xi}(t) + B(u_h(t) + u_r(t)) \\ \dot{\hat{\xi}}(t) &= A\hat{\xi}(t) + B(\hat{u}_h(t) + u_r(t)) - \Gamma\tilde{\xi}(t) \\ \dot{\tilde{\xi}}(t) &= (A - \Gamma)\tilde{\xi}(t) + B\tilde{u}_h(t)\end{aligned}\tag{E.1}$$

where $\tilde{u}_h = \hat{u}_h - u_h$, and where the estimated human control input and robot control input are computed using

$$u_r = -L_r\hat{\xi}, \quad \hat{u}_h = -\hat{L}_h\hat{\xi}.\tag{E.2}$$

Li et al. [26] algorithm (Theorem)

Theorem F.1. *Let the symmetric gain matrix $\alpha = \alpha^T > \mathbf{0} \in \mathbb{R}^{2 \times 2}$ be given, then the controller gain update law is given as:*

$$\dot{\hat{L}}_h = \alpha B^T (\tilde{\xi} - \hat{\xi}) \hat{\xi}^T,\tag{E.3}$$

stabilizing the system described in Equation (E.1).

Proof. Using the following (positive semi-definite) Lyapunov function candidate

$$V = \frac{1}{2}\hat{\xi}^T\hat{\xi} + \frac{1}{2}\tilde{\xi}^T\tilde{\xi} + \frac{1}{2\alpha}\tilde{L}_h\tilde{L}_h^T,\tag{E.4}$$

where $\tilde{L}_h = \hat{L}_h - L_h$. Then for convergence of the error states and estimated human controller gain it must hold that $\dot{V} \leq 0$:

$$\begin{aligned}\dot{V} &= \hat{\xi}^T\dot{\hat{\xi}} + \tilde{\xi}^T\dot{\tilde{\xi}} + \frac{1}{\alpha}\dot{\tilde{L}}_h\tilde{L}_h^T \leq 0, \\ &= \hat{\xi}^T(A\hat{\xi} + B(u_r + u_h)) + \tilde{\xi}^T((A - \Gamma)\tilde{\xi} + B\tilde{u}) + \frac{1}{\alpha}\dot{\tilde{L}}_h\tilde{L}_h^T, \\ &= \hat{\xi}^T(A\hat{\xi} - B(L_r + L_h)\hat{\xi}) + \tilde{\xi}^T((A - \Gamma)\tilde{\xi} - B\tilde{L}_h\tilde{\xi}) + \frac{1}{\alpha}\dot{\tilde{L}}_h\tilde{L}_h^T, \\ &= \hat{\xi}^T(A - B(L_r + L_h))\hat{\xi} + \tilde{\xi}^T(A - \Gamma)\tilde{\xi} - \tilde{\xi}^TB\tilde{L}_h\tilde{\xi} + \frac{1}{\alpha}\dot{\tilde{L}}_h\tilde{L}_h^T.\end{aligned}\tag{E.5}$$

Now, we can only assume that $A_{cl} = A - B(L_r + \hat{L}_h)$ is stable, and thus negative definite. We can use the fact that $L_h = \hat{L}_h - \tilde{L}_h$. Secondly, we use for simplicity $\Gamma_0 = A - \Gamma$, in which Γ has to be designed such that Γ_0 is stable and thus negative definite. Also, $\dot{\hat{L}}_h = B^T \dot{\hat{P}}_h$. Then:

$$\begin{aligned}\dot{V} &= \xi^T (A - B(L_r + L_h))\xi + \tilde{\xi}^T (A - \Gamma)\tilde{\xi} - \tilde{\xi}^T B\tilde{L}_h\xi + \frac{1}{\alpha}\dot{\hat{L}}_h\tilde{L}_h^T, \\ &= \xi^T (A - B(L_r + \hat{L}_h - \tilde{L}_h))\xi + \tilde{\xi}^T \Gamma_0\tilde{\xi} - \tilde{\xi}^T B\tilde{L}_h\xi + \frac{1}{\alpha}\dot{\hat{L}}_h\tilde{L}_h^T, \\ &= \xi^T A_{cl}\xi + \tilde{\xi}^T \Gamma_0\tilde{\xi} + \xi^T B\tilde{L}_h\xi - \tilde{\xi}^T B\tilde{L}_h\xi + \frac{1}{\alpha}B^T \dot{\hat{P}}_h\tilde{L}_h^T.\end{aligned}\quad (\text{E6})$$

Now, we know that the first terms are per definition negative semi-definite due to the stability properties of A_{cl} and Γ_0 . Thus, if these terms are omitted from the equation, then if it holds that $\dot{V} \leq 0$, then $\dot{V} \leq 0$ will also occur with these terms. Now assuming the human gain fixed, then $\dot{\hat{P}}_h = \mathbf{0}$. This means that $\dot{\hat{P}}_h = \dot{\hat{P}}_h - \dot{\hat{P}}_h = \dot{\hat{P}}_h$. Now these assumptions give:

$$\begin{aligned}\dot{V} &\geq \xi^T B\tilde{L}_h\xi - \tilde{\xi}^T B\tilde{L}_h\xi + \frac{1}{\alpha}B^T \dot{\hat{P}}_h\tilde{L}_h^T, \\ &= (\xi - \tilde{\xi})^T B\tilde{L}_h\xi + \frac{1}{\alpha}B^T \dot{\hat{P}}_h\tilde{L}_h^T \leq 0.\end{aligned}\quad (\text{E7})$$

In order to come to a conclusion for the update rule $\dot{\hat{P}}_h$ the following must hold:

$$\frac{1}{\alpha}B^T \dot{\hat{P}}_h\tilde{L}_h^T \leq (\tilde{\xi} - \xi)^T B\tilde{L}_h\xi, \quad (\text{E8})$$

Now, note that $(\tilde{\xi} - \xi)^T B = B^T (\tilde{\xi} - \xi)$ and that $\tilde{L}_h\xi = \xi^T \tilde{L}_h^T$. Then we obtain:

$$\begin{aligned}\frac{1}{\alpha}B^T \dot{\hat{P}}_h\tilde{L}_h^T &\leq B^T (\tilde{\xi} - \xi)\xi^T \tilde{L}_h^T, \\ \dot{\hat{P}}_h &\leq \alpha(\tilde{\xi} - \xi)\xi^T.\end{aligned}\quad (\text{E9})$$

□

E1.1. Remarks

There is however a side note to the Li et al. [26] algorithm, as the algorithm that is presented in the paper, shown in Equation (E3), does not obtain desired results. This is demonstrated in Example E1.

Simulation (Example)

Example E1. *Simulation example showing that the Li et al. [26] algorithm does not converge*

Given a simple 1D linear tracking task, tracking a reference state $\mathbf{r}(t) \in \mathbb{R}^2$ of a mass with damping only, with acceleration control $\mathbf{u}(t) \in \mathbb{R}$, the states $\mathbf{x}(t) \in \mathbb{R}^2$ and error states $\boldsymbol{\xi}(t) \in \mathbb{R}^2$ system matrices $A \in \mathbb{R}^{2 \times 2}$ and $B \in \mathbb{R}^{2 \times 1}$ and true cost matrix $Q_h \in \mathbb{R}^{2 \times 2}$ and sharing rule $C \in \mathbb{R}^{2 \times 2}$ are given as:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -10 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 20 \end{bmatrix}, \quad Q_h = \begin{bmatrix} 10 & 0 \\ 0 & 0.4 \end{bmatrix}, \quad C = \begin{bmatrix} 25 & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{E10})$$

For a reference position $x_r(t) = 0.5 \cos(\frac{2\pi}{5} t)$, the reference vector is given by $\mathbf{r}(t) = [x_r(t) \quad \dot{x}_r(t)]^T$. The following matrices are chosen the following way:

$$\hat{Q}_h = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{\mathbf{x}} = \hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 20 & 0 \\ 0 & 8 \end{bmatrix} \quad (\text{E11})$$

Such that $A - \Gamma$ is chosen negative definite.

The true states $\mathbf{x}(t)$ and their estimates $\hat{\mathbf{x}}(t)$ are then obtained as follows:

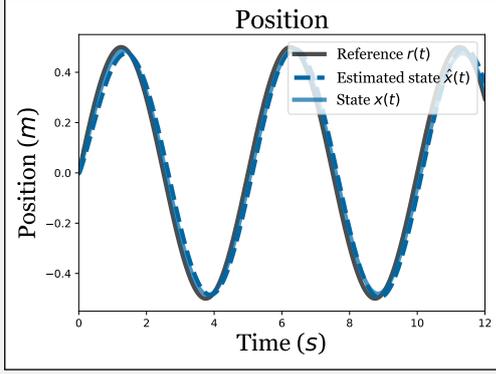


Figure E1: Position tracking performance.

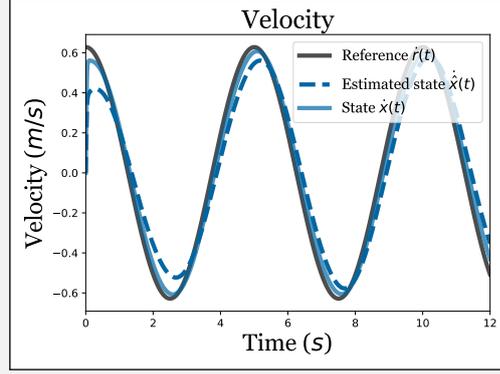


Figure E2: Velocity tracking performance.

Cost and gain for position:

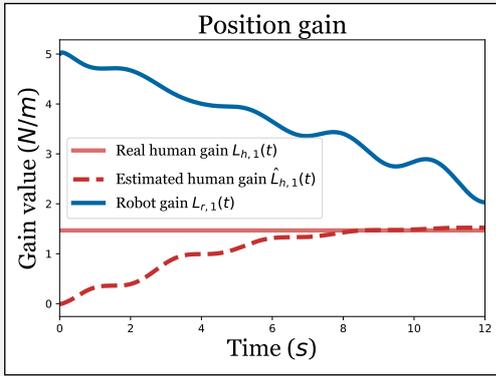


Figure E3: Position gain estimation.

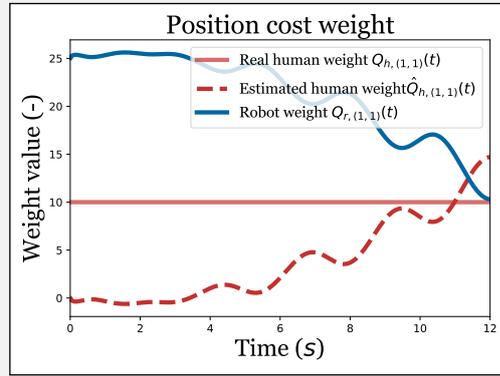


Figure E4: Position cost function weight estimation.

Cost and gain for velocity:

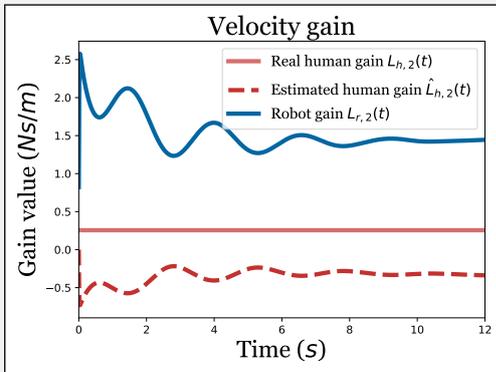


Figure E5: Velocity gain estimation.

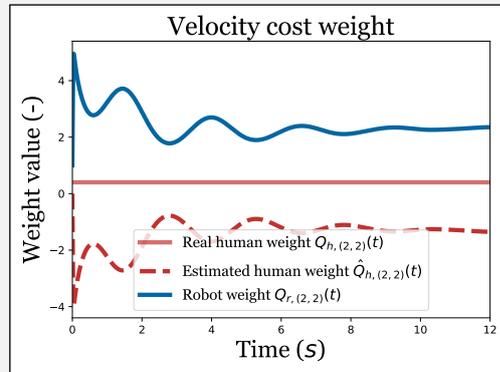


Figure E6: Velocity cost function weight estimation.

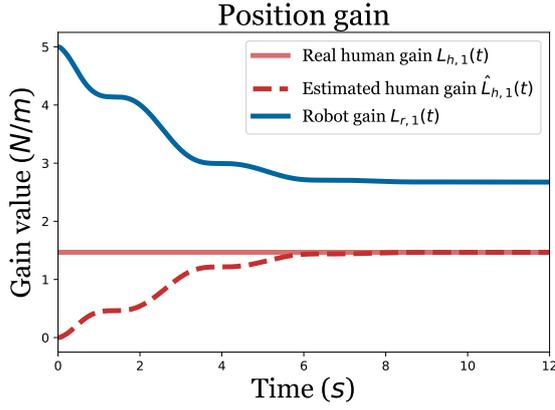
However, by using the following update function:

$$\dot{\hat{L}}_h = \alpha B^T \tilde{\xi} \tilde{\xi}^T, \quad (E12)$$

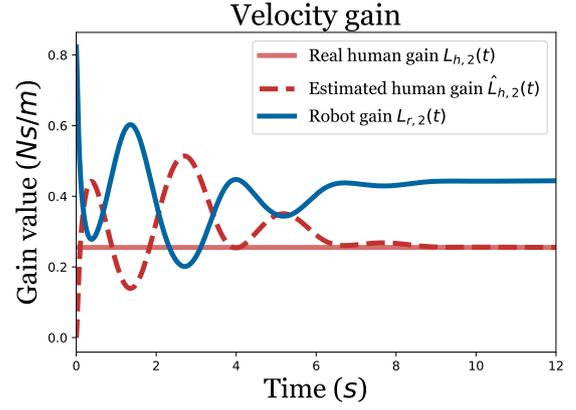
combined with making the \hat{P}_h symmetric and making \hat{Q}_h a diagonal matrix, using:

$$\hat{P}_{h,1,2} = \hat{P}_{h,2,1}, \quad \hat{Q}_{h,1,2} = \hat{Q}_{h,2,1} = 0, \quad (E13)$$

the algorithm converges. This is demonstrated in Figure E7.



(a) The position error controller gain converges.

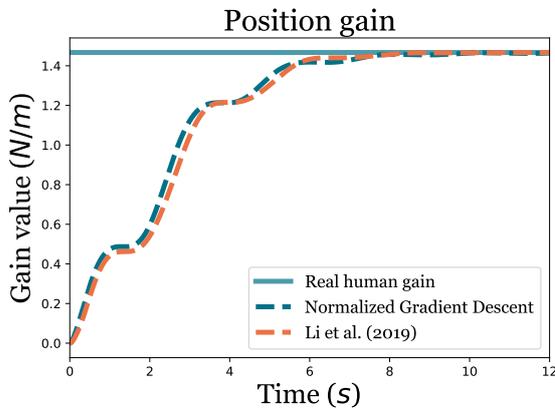


(b) The velocity error controller gain converges.

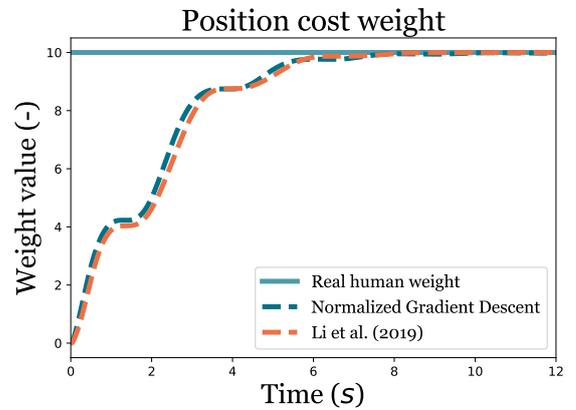
Figure E7: Under strict assumptions the Li et al. [26] algorithm converges.

E.2. Comparison

Figure E8 and Figure E9 show how the human gains and cost weights are estimated for the different algorithms. It can be concluded that both control algorithms differ very little. It seems that the velocity gain and cost weight is estimated with a bit less overshoot. This effect may however be smaller for different tuning of the convergence rate α of the Li et al. [26] algorithm. If this similarity can be mathematically demonstrated, the Li et al. [26] algorithm may perform better in a physical setting than the algorithm presented in this thesis since the Li et al. [26] algorithm does not use acceleration measurements to determine the estimated human gain \hat{L}_h .

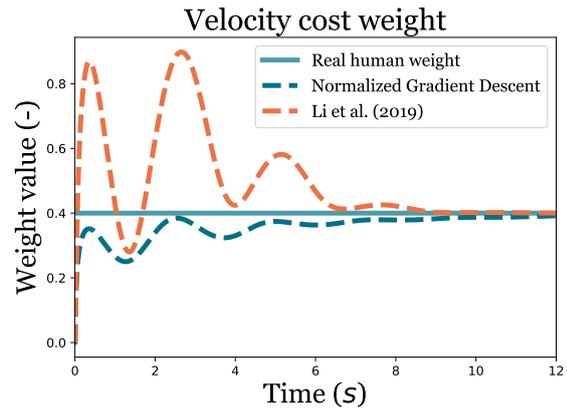
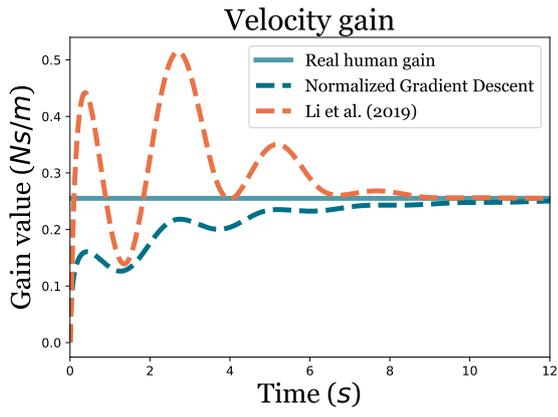


(a) Difference in position gain for the different control algorithms.



(b) Difference in position cost weights for the different control algorithms.

Figure E8: Difference in position gain and cost weights for the different control algorithms.



(a) Difference in velocity gain for the different control algorithms.

(b) Difference in velocity cost weights for the different control algorithms.

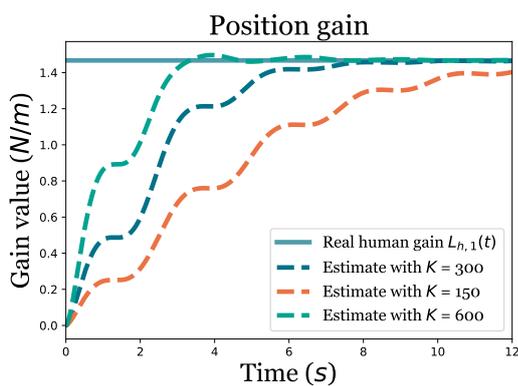
Figure E9: Difference in velocity gain and cost weights for the different control algorithms.

G

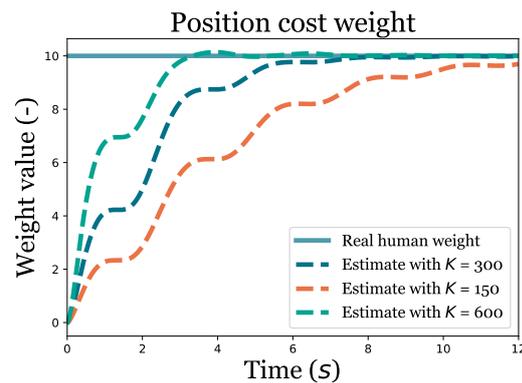
Sensitivity Analysis

This chapter examines the effect of the free design parameters of the *Normalized Gradient Descent Algorithm* for the human controller gains using a state observer controller architecture. These design parameters are the gain matrix K (Appendix G.1), Normalization parameter κ (Appendix G.2) and a controller velocity gain bias $\tilde{L}_{h,2}$ (Appendix G.3). This chapter is limited to examining the effect on the estimation of the cost weight matrix and controller gains.

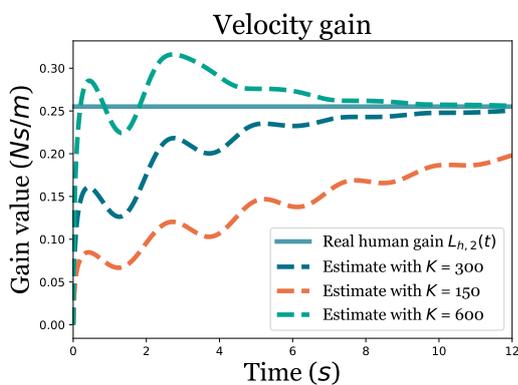
G.1. Gain matrix K



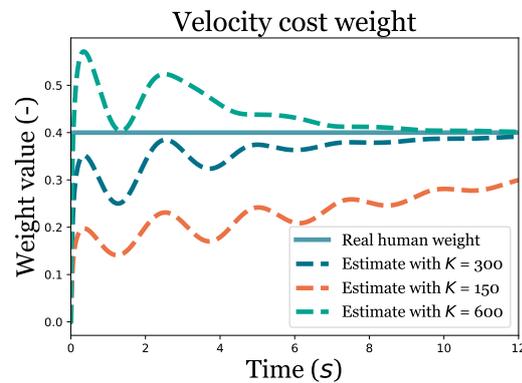
(a) Human position controller gain estimate $\hat{L}_{h,1}$ for different values of K



(b) Human position controller cost weight estimate $\hat{Q}_{h,1,1}$ for different values of K



(c) Human velocity controller gain estimate $\hat{L}_{h,2}$ for different values of K

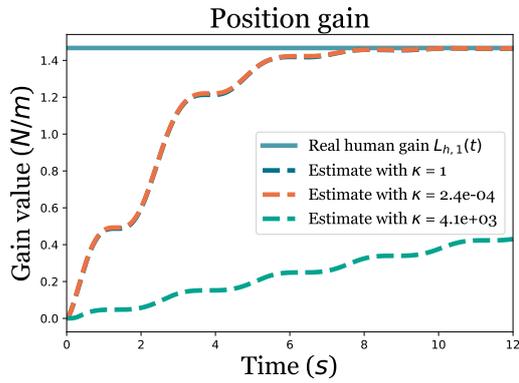


(d) Human velocity controller cost weight estimate $\hat{Q}_{h,2,2}$ for different values of K

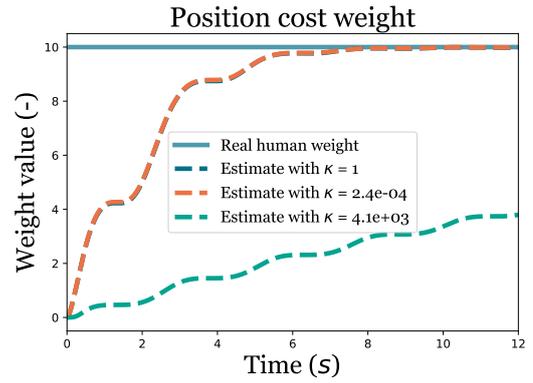
Figure G.1: A higher value of parameter K results in a faster estimation of the human controller gains and cost weights, but results in a higher overshoot.

The gain matrix K determines how fast the algorithm converges. One must however be careful, as a value that is chosen too high, may lead to instability. On the other hand, choosing the value too low results in a very long convergence time. It is advised to choose the gain matrix conservative for physical testing, to guarantee stability. A downside may be that convergence takes some time.

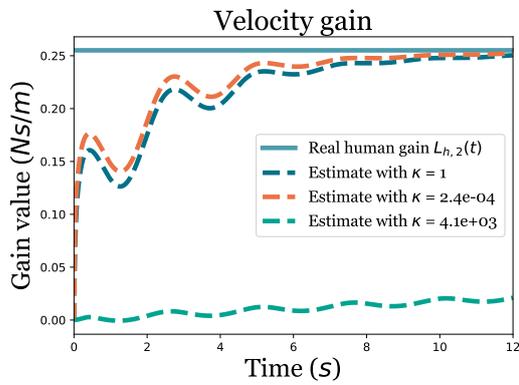
G.2. Normalization parameter κ



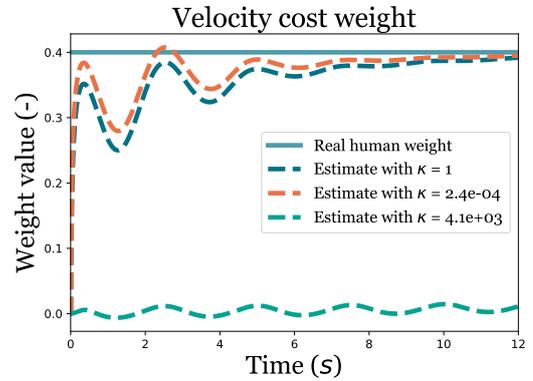
(a) Human position controller gain estimate $\hat{L}_{h,1}$ for different values of κ



(b) Human position controller cost weight estimate $\hat{Q}_{h,1,1}$ for different values of κ



(c) Human velocity controller gain estimate $\hat{L}_{h,2}$ for different values of κ

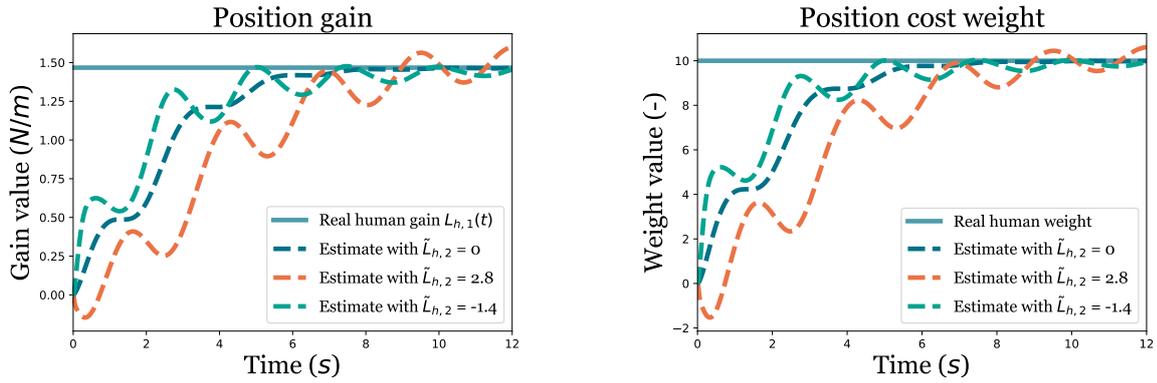


(d) Human velocity controller cost weight estimate $\hat{Q}_{h,2,2}$ for different values of κ

Figure G.2: The effect of parameter κ is limited on the estimation of the human controller gains and cost weights, but choosing a value that is too high distorts the performance of the algorithm.

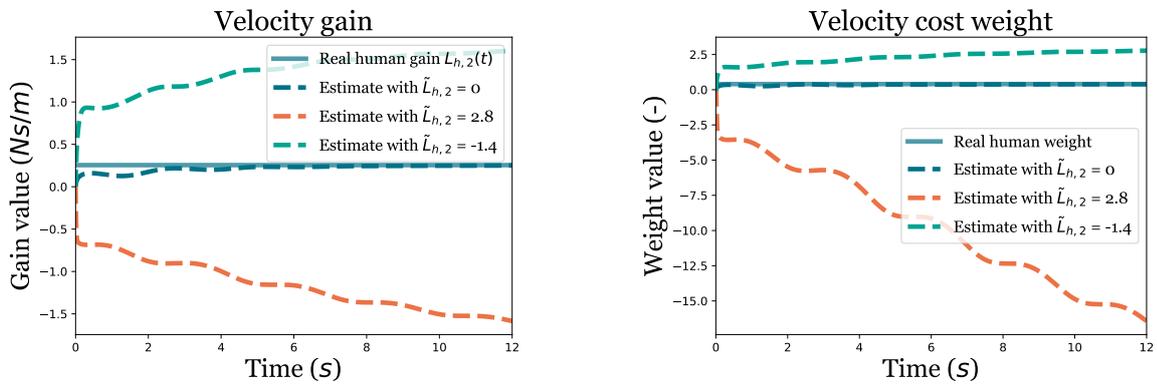
The second parameter in the novel algorithm proposed in this thesis is the normalization parameter. The figures show that this parameter is not of much interest. Only when this parameter is chosen much too high will the performance be affected. A value of 1 will do just fine.

G.3. Velocity gain bias $\tilde{L}_{h,2}$



(a) Human position controller gain estimate $\hat{L}_{h,1}$ for different values of $\tilde{L}_{h,2}$

(b) Human position controller cost weight estimate $\hat{Q}_{h,1,1}$ for different values of $\tilde{L}_{h,2}$

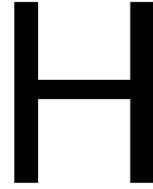


(c) Human velocity controller gain estimate $\hat{L}_{h,2}$ for different values of $\tilde{L}_{h,2}$

(d) Human velocity controller cost weight estimate $\hat{Q}_{h,2,2}$ for different values of $\tilde{L}_{h,2}$

Figure G.3: A bias $\tilde{L}_{h,2}$ for the velocity gain is detrimental on the estimation of the human controller velocity gain and cost weight. This in turn may lead to instability of the controller. The effect on the position gain and cost weight is limited however but still leads to some oscillations.

If the algorithm does not correctly identify the human controller gain, how does that affect the controller? Figure G.3 shows that this may lead to instability. Especially when the gain is estimated to be higher than the real gain, this leads to instabilities. When combining this information with the controller validation in Appendix K, it is justified that the human is only estimated as a position gain and not a velocity gain.



Non-linear Model Identification

One of the assumptions that were made for the control algorithm was that the system that is controlled is linear. This however is not true for the SENSO-Wheel SensoDrive GmbH [40], which shows some clear non-linear behavior. To compensate for these nonlinearities, feedforward compensation of the nonlinearities was performed. To this end, identification of the nonlinear model must first be performed. This was done by doing a grey-box parameter estimation, which assumes a model structure and estimates these parameters using optimization.

H.1. Linear model filtering

Figure H.1 shows the filtered signals. From Figure H.1b it is clear that the steering acceleration becomes fairly distorted through filtering, placing a physical limit on the performance of the control algorithm.

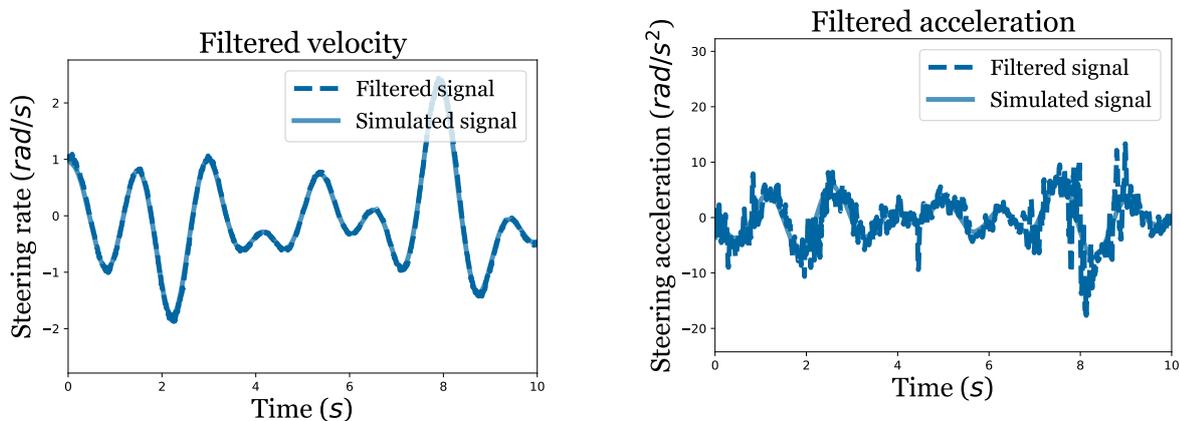
H.2. Grey-box Parameter Estimation

The non-linear model of the steering wheel system can be captured as:

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B(u_r(t) + u_h(t)) + f_{nl}(t) \tag{H.1}$$

The nonlinear component $f_{nl}(t)$ is assumed to be composed of:

- A friction component
- A gravitational component



(a) Comparison filtered steering rate versus simulated steering rate

(b) Comparison filtered steering acceleration versus simulated steering acceleration

Figure H.1: Comparison filtered signals versus simulations

H.2.1. Friction component

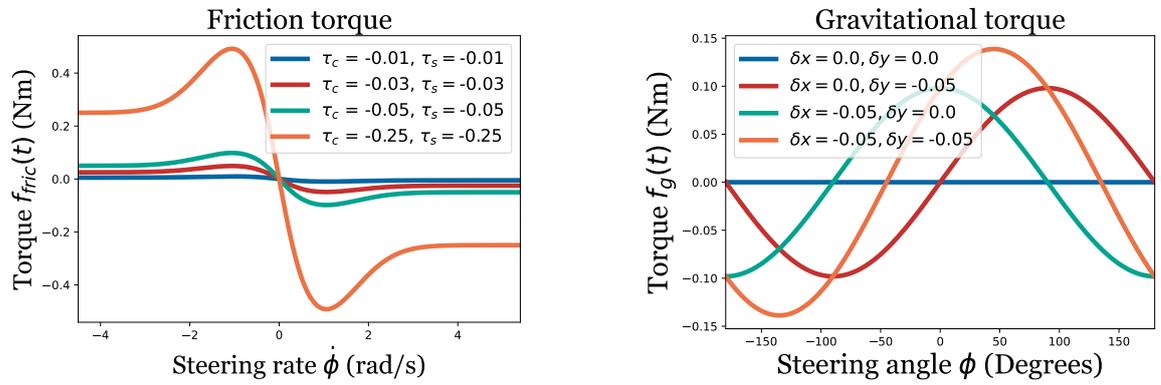
To model the non-linear friction component, the procedure by Specker et al. [41] is partly followed. Here they use a tangens hyperbolicus to model the coulomb friction and a Gaussian bell to model the Stribeck friction. The resulting friction torque $f_{fric}(t)$ is defined as:

$$f_{fric}(t) = \tau_c \tanh\left(\frac{\dot{\phi}(t)}{\dot{\phi}_t}\right) + \tau_s \frac{\dot{\phi}(t)}{\dot{\phi}_{sp}} e^{-\left(\frac{\dot{\phi}(t)}{\sqrt{2}\dot{\phi}_{sp}}\right)^2 + \frac{1}{2}}. \quad (\text{H.2})$$

H.2.2. Gravitational component

The gravitational torque acting on the steering wheel is due to an offset of the center of gravity from the geometrical center of the steering wheel. Both the offset in the horizontal direction δx and in the vertical direction δy are incorporated in the model. The non-linear gravitational torque is defined as

$$f_{grav}(t) = mg(\delta x \cos \phi(t) + \delta y \sin \phi(t)). \quad (\text{H.3})$$



(a) Friction torque for different parameter values

(b) Gravitational torque for different parameter values

Figure H.2: Non-linear components for different sets of parameters

H.2.3. Model definition

The previous sections have shown the definition of the non-linear components of the steering wheel model, based on some parameters θ . These can be combined to:

$$f_{nl}(\phi(t), \dot{\phi}(t), \theta) = \tau_{fric}(\dot{\phi}(t), \theta) + \tau_g(\phi(t), \theta). \quad (\text{H.4})$$

The equations of motion (EoM) can then be defined as

$$I\ddot{\phi}(t) + D\dot{\phi}(t) + K\phi(t) = u_r(t) + u_h(t) + f_{nl}(\phi(t), \dot{\phi}(t), \theta), \quad (\text{H.5})$$

where I , D , and K are the inertia, damping, and stiffness of the steering wheel of which the inertia is unknown and the damping and stiffness can be actively set to certain values. $u_r(t)$ and $u_h(t)$ are the torques applied to the steering system by the robot (motor) and the human respectively.

This system can be rewritten in state-space form using the state $\mathbf{x} = [x_1(t) \ x_2(t)]^T$, where $x_1(t) = \phi(t)$ and $x_2(t) = \dot{\phi}(t)$:

$$\begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} &= \begin{bmatrix} x_2(t) \\ \frac{1}{I}(-Dx_2(t) - Kx_1(t) + u_r(t) + u_h(t) + f_{nl}(x_1(t), x_2(t), \theta)) \end{bmatrix}, \\ \dot{\mathbf{x}}(t) &= A\mathbf{x}(t) + B(u_r(t) + u_h(t) + f_{nl}(\mathbf{x}(t), \theta)) \end{aligned} \quad (\text{H.6})$$

Where $A = \begin{bmatrix} 0 & 1 \\ -\frac{K}{I} & -\frac{D}{I} \end{bmatrix}$ and $B = \begin{bmatrix} 0 \\ \frac{1}{I} \end{bmatrix}$

Since the system is simulated without a human input u_h , the system is simplified to the following system:

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B(u_r(t) + f_{nl}(\mathbf{x}(t), \theta)) \quad (\text{H.7})$$

H.3. Optimization definition

The aim of the optimization is to match the parameters, defined in Appendix H.2.3, as closely as possible to reality. The parameters are grouped into the parameter vector $\theta = [I, m, \delta x, \delta y, \phi_t, \tau_c, \tau_s]$. As in Specker et al. [41], the Stribecker and Coulomb activation velocities can be estimated through the relation $\phi_{sp} \approx 2\phi_t$. Note that the linear component inertia I is also included since this parameter is also unknown. The steering wheel system is then approximated by:

$$\dot{\hat{\mathbf{x}}}(t; \theta) = A(\theta)\hat{\mathbf{x}}(t; \theta) + B(\theta)u_r(t) + B(\theta)f_{nl}(t, \hat{\mathbf{x}}, \theta) \quad (\text{H.8})$$

The goal of the optimization is to minimize the squared (euclidean) distance between the estimated states $\hat{\mathbf{x}}(t; \theta)$ and the measured (real) states $\mathbf{x}(t)$, using the parameter vector θ and constrained to the dynamics states in Equation (H.8).

In order to do so in a comparable manner, multiple measurements are done, using multiple steering wheel settings. The optimized cost function $I(\theta)$ then becomes, for n measurements of m different steering wheel settings (causing different A matrices), $I(\theta) = \sum_{i=1}^n \sum_{j=1}^m \|\hat{\mathbf{x}}(t; \theta) - \mathbf{x}(t)\|_2^2$

The total optimization problem is given as:

$$\begin{aligned} \min_{\theta} \sum_{i=1}^n \sum_{j=1}^m \|\hat{\mathbf{x}}(t; \theta) - \mathbf{x}(t)\|_2^2 \\ \text{s.t. } \dot{\hat{\mathbf{x}}}(t; \theta) = A_j(\theta)\hat{\mathbf{x}}(t; \theta) + B(\theta)u_r(t) + B(\theta)f_{nl}(t, \hat{\mathbf{x}}, \theta) \end{aligned} \quad (\text{H.9})$$

H.3.1. Identification procedure

The identification is performed using a forcing function, which prescribes the input torques that are applied to the steering wheel. This forcing function is described in Appendix I. Two forcing functions were designed, one for identification and one for verification of the grey-box identification. The forcing function has a duration of 33.1s and is repeated twice within the identification and the verification set with a different damping coefficient. For the identification and verification, $B = [0.1, 0.2, 0.3, 0.4] Nm/s$ was used. Different forcing functions were designed. The identified parameters were set to the following initial values:

Estimated parameter	Estimated value
I	0.06 kgm^2
m	0.5 kg
δx	0 m
δy	0 m
ϕ_t	0.3 rad/s
τ_c	0.0 Nm
τ_s	0.0 Nm

Table H.1: Initial parameter values for the optimization

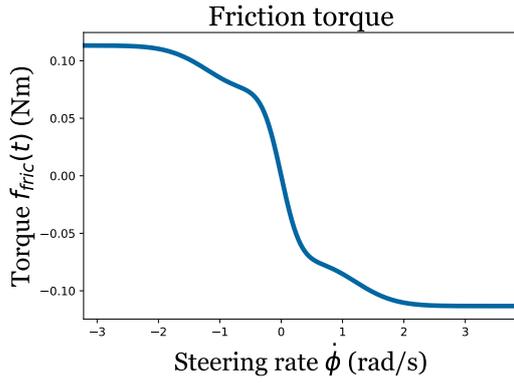
H.4. Results

The final values that were obtained in the grey-box identification are:

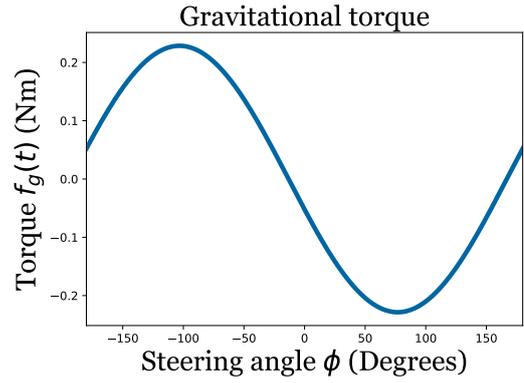
Estimated parameter	Estimated value
I	$4.88 \cdot 10^{-2} kgm^2$
m	0.256 kg
δx	$2.08 \cdot 10^{-2} m$
δy	$8.87 \cdot 10^{-2} m$
ϕ_t	0.358 rad/s
τ_c	$3.12 \cdot 10^{-2} Nm$
τ_s	-0.113 Nm

Table H.2: Obtained estimated parameter values

Figure H.3 demonstrates the dynamics of both nonlinear components.



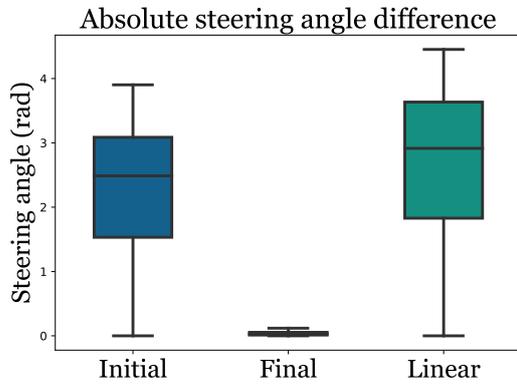
(a) Estimated friction torque curve



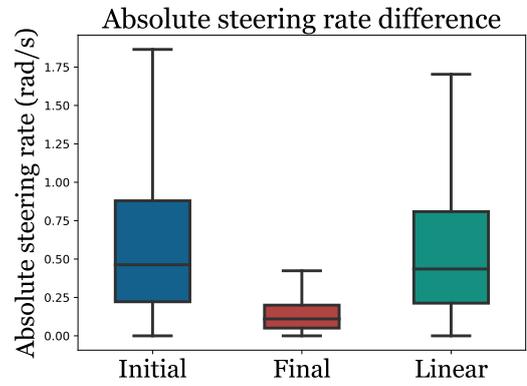
(b) Estimated gravitational torque curve

Figure H.3: Estimated nonlinear components

Figure H.4 and Figure H.5 show how for multiple models, being the initially estimated model, the final model, and a linear approximation, the model fits the measurements.

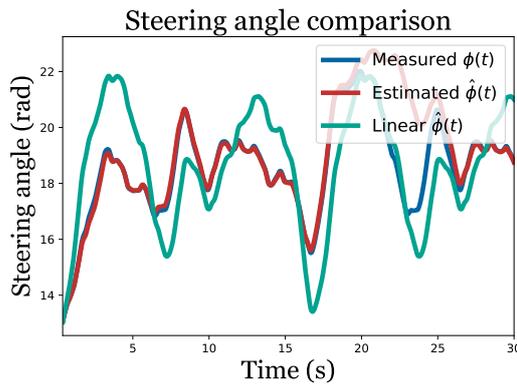


(a) Resulting steering wheel angle fit.

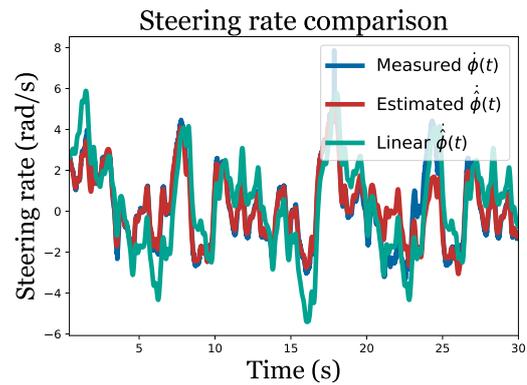


(b) Resulting steering wheel rate fit.

Figure H.4: Boxplots for the fit of the system dynamics using the nonlinear models.



(a) Resulting steering wheel angle fit in time domain.



(b) Resulting steering wheel rate fit in time domain.

Figure H.5: Time domain fit of the system dynamics using the nonlinear models.

Table H.3 shows objectively how well the models capture the dynamics of the steering wheel. A VAF of 71.8% and 93.3% for the steering angle and steering rate model is found in the grey-box identification model.

Estimate	Mean absolute distance		Mean absolute variance	
	Steering angle	Steering rate	Steering angle	Steering rate
Initial	2.29 <i>rad</i>	0.63 <i>rad/s</i>	1.01 <i>rad</i> ²	0.30 <i>rad</i> ² / <i>s</i> ²
Final	0.13 <i>rad</i>	0.19 <i>rad/s</i>	0.28 <i>rad</i> ²	0.11 <i>rad</i> ² / <i>s</i> ²
Linear	2.70 <i>rad</i>	0.60 <i>rad/s</i>	1.33 <i>rad</i> ²	0.28 <i>rad</i> ² / <i>s</i> ²

Table H.3: Metrics, demonstrating the model fit.

Estimate	VAF steering angle	VAF steering rate
Initial	-2.4 %	81.3 %
Final	71.8 %	93.3 %
Linear	-34.9 %	82.6 %

Table H.4: Variance Accounted For (VAF) values

Multi-sine signal design

The amplitudes A_i of the reference signal were designed using a second-order Butterworth filter [28]. Then, to determine the phases, crest function minimization was used [18] to maximize the signal-to-noise ratio. According to Damveld et al. [8], the effect of the phases is negligible, therefore this has been used to randomize the signal to obtain a different set for identification and validation.

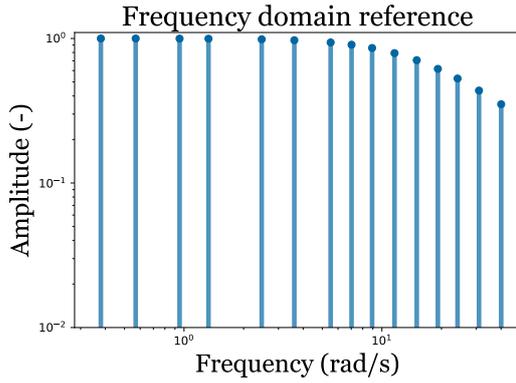
$$u(t) = \sum_i^N A_i \sin(\omega_i t + \theta_i) \quad (I.1)$$

I.1. Nonlinear Model Identification Input Signal

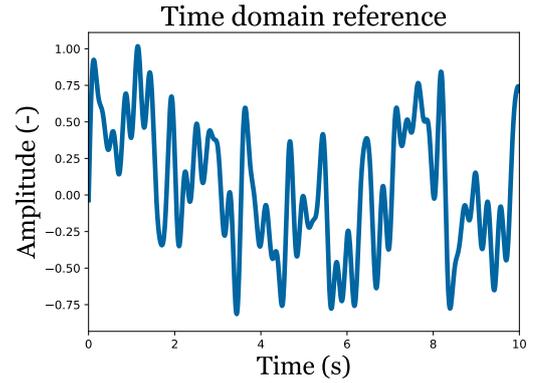
The signal is chosen such that the bandwidth is $\omega_B = 15 \text{ rad/s}$. To guarantee that the system dynamics are captured for a large band of frequencies, fifteen frequencies are used. The designed signals are shown in Table I.1. Two signals were designed, where the two best results of the crest function minimization were selected.

i	A_i	k_i	ω_i	$\theta_{i,id}$	$\theta_{i,ver}$
-	Nm	-	rad/s	rad	rad
1	0.2	2	0.38	-0.98	0.29
2	0.2	3	0.57	1.22	-0.15
3	0.2	5	0.95	0.30	-0.07
4	0.2	7	1.33	-0.49	0.17
5	0.2	13	2.47	0.20	1.33
6	0.19	19	3.61	-1.17	-0.39
7	0.19	29	5.50	0.30	-0.78
8	0.18	37	7.02	-0.87	-0.28
9	0.17	47	8.92	-0.26	-0.39
10	0.16	61	11.6	0.81	0.31
11	0.14	79	15.0	0.00	-0.38
12	0.12	101	19.2	-1.40	-0.32
13	0.11	127	24.1	-0.37	-0.71
14	0.09	163	30.9	-0.67	-0.74
15	0.07	211	40.1	-2.00	-0.48

Table I.1: Forcing function parameters for the identification and verification set.



(a) Frequency domain forcing function.



(b) Time domain forcing function for the identification data-set.

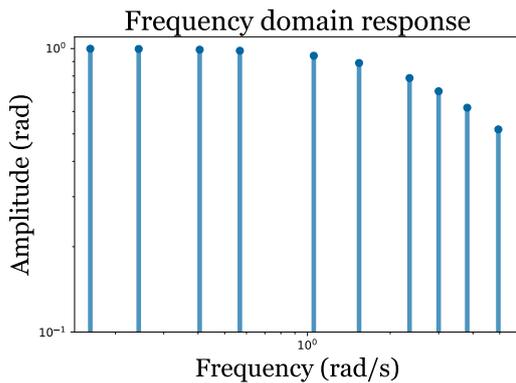
Figure I.1: Visual forcing function representation

I.2. Reference Trajectory Design

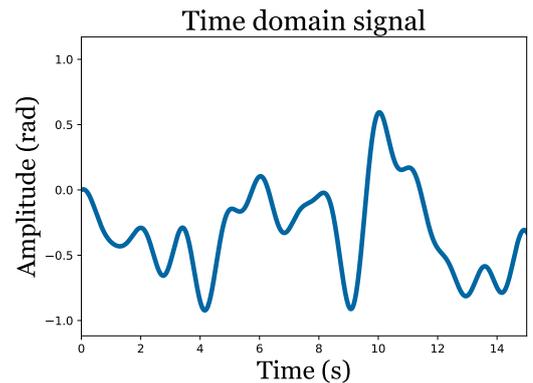
The bandwidth for the reference Trajectory was set to 3rad/s and the amplitude to 0.25rad .

i	A_i	k_i	ω_i	θ_i
-	Nm	-	rad/s	rad
1	0.250	2	1.02	-0.52
2	0.250	3	1.53	1.05
3	0.248	5	2.55	-0.64
4	0.246	7	3.57	-0.14
5	0.236	13	6.62	0.56
6	0.222	19	9.68	-0.77
7	0.197	29	14.8	1.26
8	0.177	37	18.8	0.26
9	0.155	47	23.9	1.58
10	0.130	61	31.1	0.01

Table I.2: Input signal parameters



(a) Frequency domain reference signal



(b) Time domain reference signal

Figure I.2: Visualization of reference signal properties

J

Algorithm

This appendix presents the entire control algorithm that was used in the experiments. The control algorithm consists of two processes that run simultaneously and communicate with one another, being the high-level and low-level controllers. The high-level controller is used to feed the reference state to the low-level controller and to feedback the reference state to the human through a visual interface. The low-level controller uses the reference signal to compute control inputs.

The code is available at <https://github.com/lflipse/gt-state-feedback-observer>

High-level controller (Algorithm)

Algorithm J.1. 1: Initialize experiment
2: Initialize SENSO-Wheel process
3: Initialize visualization
4: **while** $t < t_{end}$ **do**
5: Compute reference signal $r(t)$
6: Send data to low-level controller
7: Receive data from low-level controller
8: Visually feedback reference position with preview and current steering wheel angle $\phi(t)$
9: Save data to output dictionary
10: **end while**
11: Quit SENSO-Wheel process
12: Quite Low-level controller
13: Save output dictionary to .csv file

High-level controller (Algorithm)

Algorithm J.2. 1: Initialize steering rate $\dot{\phi}(t)$ and acceleration $\ddot{\phi}(t)$ filter parameters,
2: Initialize $\hat{L}_h, \hat{Q}_h, L_r, Q_r, \hat{\xi}(t), \Gamma, K, \kappa$ and t_{end} .
3: **while** not quit **do**
4: Measure steering angle $\phi(t)$, compute and filter steering rate $\dot{\phi}(t)$ and acceleration $\ddot{\phi}(t)$,
5: Form error states $\xi(t), \dot{\xi}(t)$ from $r(t)$ and compute $\hat{\xi}(t), \dot{\hat{\xi}}(t)$,
6: Compute $\hat{u}_h(t)$ and $\tilde{u}_h(t)$,
7: Update $\hat{L}_h(t)$ from $\hat{L}_h(t)$ and $\hat{\xi}(t)$ from $\dot{\hat{\xi}}(t)$,
8: Compute $L_r(t)$, compute $u_r(t)$ and solve \hat{Q}_h ,
9: Calculate $f_{nl}(\phi(t))$ and determine $u_{r,t}$.
10: **end while**

K

Controller Validation

The novel control algorithm presented in this thesis estimates the controller gains of a human. Convergence of the estimator guarantees stability, proven in Appendix D. It is therefore essential that the controller gains indeed converge. This is tested in two ways. First using a virtual human, then using a real human. The virtual human is control inputs on the steering wheel of which the control algorithm is unaware, on top of the algorithm itself. These virtual human control inputs are determined using a virtual human controller gain, which is estimated using the control algorithm. Secondly, a real human is tasked to show some driving behaviors, being no steering behavior, weak steering, strong steering, and active countersteering. The real human controller gains are now unknown, but there should be a trend in the estimated values visible, corresponding with the steering behaviors. The validation experiment takes 200s, and the experiment consists of 6 phases. Each phase corresponds to a certain behavior of the human. First, weak steering behavior is performed, followed by no steering behavior. Then, strong steering behavior is performed, again followed by no steering behavior. Then, finally, active countersteering behavior is performed, followed by no steering behavior. In Appendix I, the design of the reference trajectory is provided.

K.1. Virtual Human

The following settings for the controller are used:

$$\Gamma = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}, K = \begin{bmatrix} 75 & 0 \\ 0 & 7.5 \end{bmatrix}, Q_r = \begin{bmatrix} 20 & 0 \\ 0 & 2 \end{bmatrix}, \kappa = 1. \quad (\text{K.1})$$

The virtual human gains are determined as follows:

$$L_{h,vir,(2,4,6)} = \begin{bmatrix} 0 & 0 \end{bmatrix}, L_{h,vir,(1)} = \begin{bmatrix} 1.24 & 0.207 \end{bmatrix}, L_{h,vir,(3)} = \begin{bmatrix} 3.21 & 0.455 \end{bmatrix}, L_{h,vir,(5)} = \begin{bmatrix} -1.24 & -0.207 \end{bmatrix}, \quad (\text{K.2})$$

The corresponding cost weights are then:

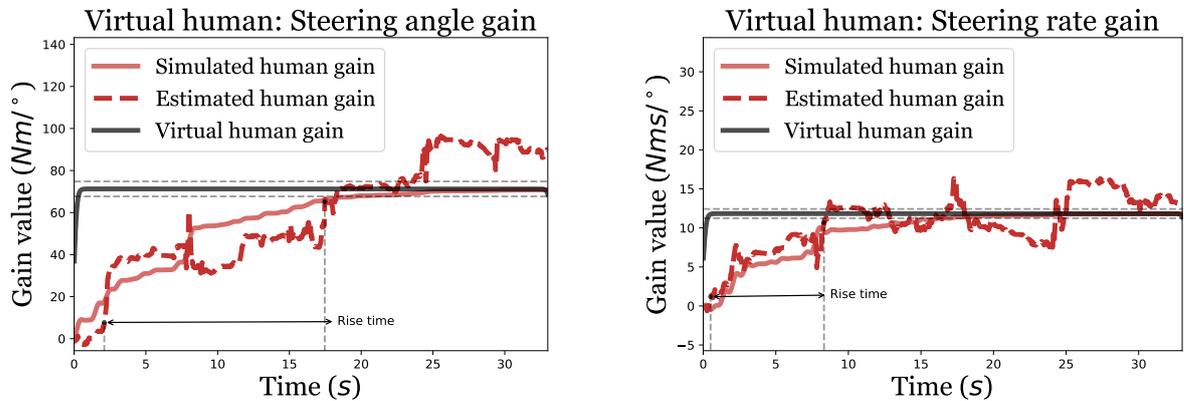
$$Q_{h,vir,(2,4,6)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, Q_{h,vir,(1)} = \begin{bmatrix} 10 & 0 \\ 0 & 0.5 \end{bmatrix}, Q_{h,vir,(3)} = \begin{bmatrix} 25 & 0 \\ 0 & 1 \end{bmatrix}, Q_{h,vir,(5)} = \begin{bmatrix} -13.0 & 0 \\ 0 & -0.585 \end{bmatrix}, \quad (\text{K.3})$$

The following table summarizes, for the time $0s \leq t \leq 33\frac{1}{3}s$, the physical performance of the controller.

	Measurement	Rise time	Settling time	Bias	Noise variance
Angle gain	1	15.3s	-	0.21rad	$3.12 \cdot 10^{-2} rad^2$
	2	21.7s	-	$1.56 \cdot 10^{-2} rad$	$1.12 \cdot 10^{-2} rad^2$
	3	5.87s	-	0.141rad	$1.36 \cdot 10^{-2} rad^2$
	4	7.33s	-	-0.154rad	$2.39 \cdot 10^{-2} rad^2$
	5	13.9s	-	0.148rad	$5.55 \cdot 10^{-3} rad^2$
Average	1-5	12.8s	-	0.0693rad	$1.71 \cdot 10^{-2} rad^2$
Rate gain	1	7.82s	-	$5.68 \cdot 10^{-3} rad/s$	$1.58 \cdot 10^{-3} rad^2/s^2$
	2	7.29s	-	$-1.09 \cdot 10^{-2} rad/s$	$8.78 \cdot 10^{-4} rad^2/s^2$
	3	7.26s	-	$3.50 \cdot 10^{-2} rad/s$	$7.50 \cdot 10^{-4} rad^2/s^2$
	4	10.9s	-	$3.82 \cdot 10^{-2} rad/s$	$7.66 \cdot 10^{-4} rad^2/s^2$
	5	7.35s	-	$3.29 \cdot 10^{-2} rad/s$	$2.18 \cdot 10^{-3} rad^2/s^2$
Average	1-5	8.12s	-	$3.04 \cdot 10^{-2} rad/s$	$1.23 \cdot 10^{-3} rad^2/s^2$

Table K.1: Overview of performance metrics in validating the physical controller

In this table, the bias is calculated as the mean of the difference between the "reference" gain and the estimated virtual human gain from the time the virtual human gain reached 90% (end of the rise time) of the reference value. The noise variance is calculated as the variance of the virtual human gain for this same period. The following figures demonstrate these metrics for the first measurement.

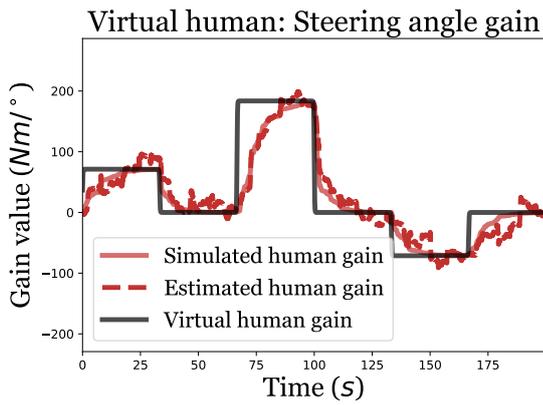


(a) Steering wheel angle gain for robot and virtual human

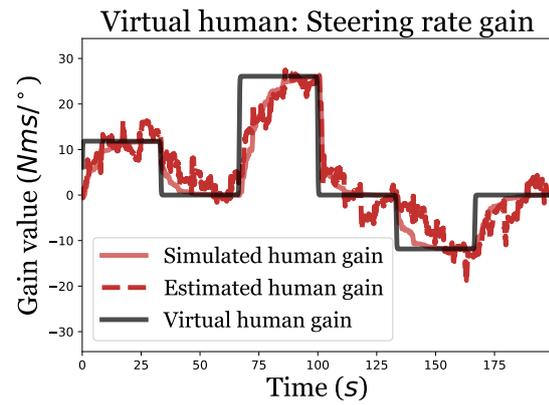
(b) Steering wheel rate gain for robot and virtual human

Figure K.1: These figures visualize the metrics in Table K.1, for both controller gains.

It shows that the algorithm is very noisy. This is due to the physical constraint on filtering of the steering acceleration signal (see Appendix H). This causes that no settling time is reached. Due to the conservative setting of the convergence gain, the algorithm performs fairly slow.

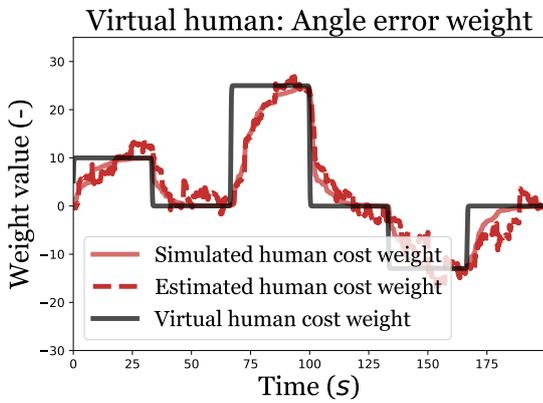


(a) Steering wheel angle gain for robot and virtual human

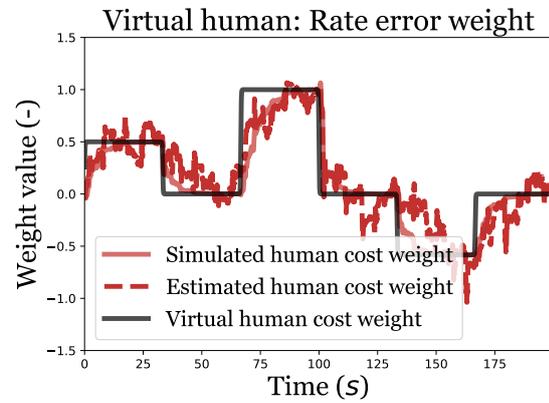


(b) Steering wheel rate gain for robot and virtual human

Figure K.2: Steering wheel gain for robot and virtual human



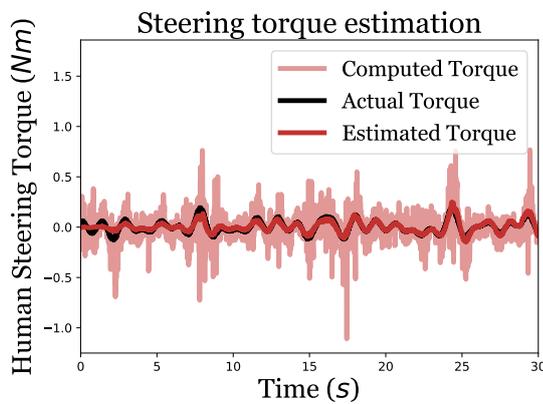
(a) Steering wheel angle cost for robot and virtual human



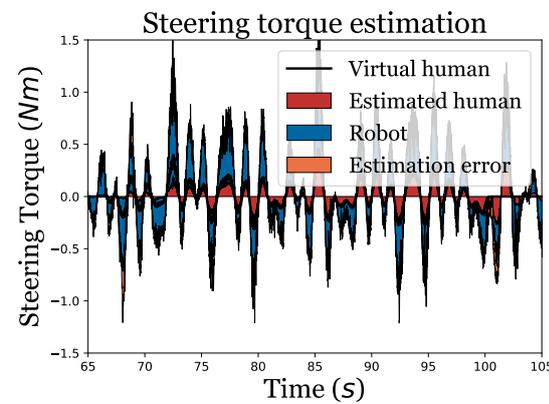
(b) Steering wheel rate cost for robot and virtual human

Figure K.3: Steering wheel gain for robot and virtual human

The figures demonstrate the convergence of the algorithm, albeit somewhat noisy. The following figures show that this leads to the correct identification of the control inputs.



(a) Virtual human estimated control input

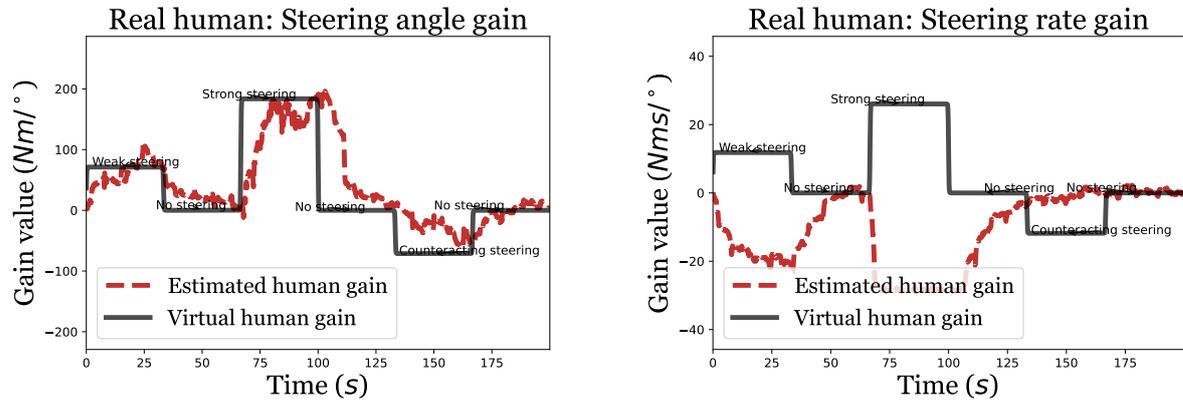


(b) Relation between the (estimated) human and robot control inputs

Figure K.4: These figures demonstrate that the virtual human's control inputs are correctly identified.

K.2. Real Human

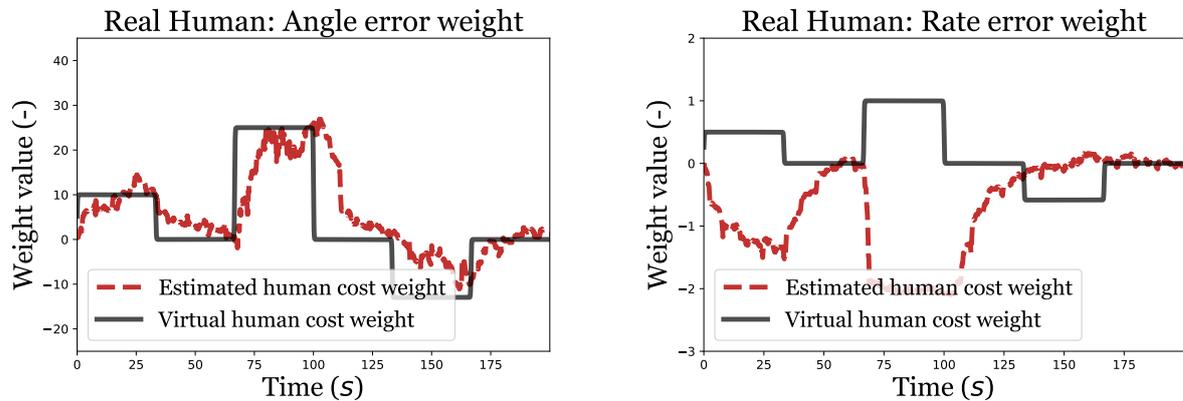
The procedure for the real human validation experiment is similar to that of the virtual human. Now, instead of adding extra control inputs, a real human input torques to the steering wheel. The human is instructed to input weak, strong, and counteracting steering behavior, alternated with no steering behavior. The figures show how the control algorithm estimates human behavior. It shows that the steering angle gain shows a clear trend with the human steering behavior, also showing a close resemblance with the virtual human. Interestingly, the steering rate gains do not. This may be because the human is only shown visual feedback of the position, not the velocity of the target and the controlled system.



(a) Steering wheel angle gain for robot and real human

(b) Steering wheel rate gain for robot and real human

Figure K.5: Steering wheel gain for robot and real human



(a) Steering wheel angle cost for robot and real human

(b) Steering wheel rate cost for robot and real human

Figure K.6: Steering wheel gain for robot and real human

K.3. Conclusion

Based on the results, the following conclusions are drawn:

- The control algorithm correctly identifies a controller's controller gains, leading to correct identification of the controller's inputs;
- In this task, the human behaves like a controller gain on the **position only**.

L

Questionnaire

Questionnaire

Participant ID: _____

Age: _____

Gender: Male Female Prefer not to say

Please fill in the following questions after each 4 trials (1 condition).

Trial 1-4 (Condition 1)	(No)				(Yes)
<i>During the previous 4 trials...</i>	-2	-1	0	1	2
It felt like I was causing the movement of the grey cross.	<input type="radio"/>				
It felt like the steering wheel was causing the movement of the grey cross.	<input type="radio"/>				
It felt like I was contributing to the movement of the grey cross.	<input type="radio"/>				
It felt like the steering wheel was controlling my movement.	<input type="radio"/>				

Trial 5-8 (Condition 2)	(No)				(Yes)
<i>During the previous 4 trials...</i>	-2	-1	0	1	2
It felt like I was causing the movement of the grey cross.	<input type="radio"/>				
It felt like the steering wheel was causing the movement of the grey cross.	<input type="radio"/>				
It felt like I was contributing to the movement of the grey cross.	<input type="radio"/>				
It felt like the steering wheel was controlling my movement.	<input type="radio"/>				

Trial 9-12 (Condition 3)	(No)				(Yes)
<i>During the previous 4 trials...</i>	-2	-1	0	1	2
It felt like I was causing the movement of the grey cross.	<input type="radio"/>				
It felt like the steering wheel was causing the movement of the grey cross.	<input type="radio"/>				
It felt like I was contributing to the movement of the grey cross.	<input type="radio"/>				
It felt like the steering wheel was controlling my movement.	<input type="radio"/>				

M

Informed Consent Form

Informed consent form

Researchers

L.J. Flipse – MSc. Student
E-mail: l.j.flipse@student.tudelft.nl
Tel: +31 (0)6 2715 1094

Prof. dr. ir. D.A. Abbink – Supervisor
E-mail: d.a.abbink@tudelft.nl

Dr. ir. N.W.M. Beckers
E-mail: n.w.m.beckers@tudelft.nl

This document describes the purpose, procedures, benefits, risks and possible discomforts of a driving simulator study. It also describes the right to withdraw from the study at any time in any case. Before agreeing to participate in this study, it is important that the information provided in this document is fully read and understood.

Location of the experiment

TU Delft, Faculty of Mechanical, Maritime and Materials Engineering (3mE)
Cognitive Robotics Lab (F-1-490)
Mekelweg 2, 2628CD Delft

Description of Experiment

In this experiment you will be asked to perform a steering task, in which your task is to follow the target as accurate as possible (see Figure 1 and 2). This task is performed both with manual control (you alone without interaction) and with haptic shared control (together with the steering wheel, which can apply forces on the steering wheel). This research focuses on how different types of haptic shared controllers influence your steering behaviour and performance. The experiment will take approximately 20 minutes to complete.

Confidentiality: The collected data in this experiment is kept confidential and will be used for research purposes only. The data will also be anonymised i.e. you will be identified by a subject number. Your name is not connected to the data in any way. The data will be stored in a data repository and will be copied to OSF when finalized for the purpose of public sharing.

Right to refuse or withdraw: Your participation is strictly voluntary and you may withdraw from or stop this experiment at any time, without consequences.

Questions: If you have any questions regarding this experiment, feel free to contact L. Flipse (contact details are provided at the top of this document).



Figure 1, example of the experiment set-up

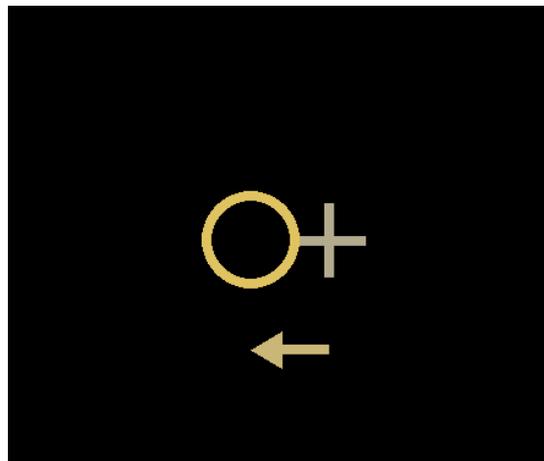


Figure 2, example of task interface

Data: The following personal data will be gathered in the experiment:

- Deidentified demographic information (Gender, age)
- Deidentified driving telemetry data (including steering angle, input torques, estimated control strategy)

Additional information regarding COVID-19: To prevent the spread of the corona virus (in compliance with the university's policy), researchers and participants in the study:

1. don't have any underlying ailments that could be seen as a risk factor for a COVID-19 infection
2. don't have any complaints or symptoms that could be indicative of a COVID-19 infection
3. have not been in contact with a COVID-19 patient at least 14 days before participation in the study
4. take suitable protective measures if a minimum distance of 1.5 meters is not viable
5. are enabled to travel outside of rush hours to and from the research location

Also, any objects or surfaces researchers and participants come into contact will be disinfected prior and after use.

Please confirm the following points before signing

Yes **No**

I understand that I am participating in human factors research;

I understand which personal data is gathered and I consent that the data gathered during the experiment may be used for a MSc thesis and possible future academic research and publications;

I understand that my participation will be anonymous (that is, my name will not be linked with my data) and that all information I provide will remain confidential;

I understand that I will be provided with an explanation of the research in which I participated and will be given the name and e-mail address of an individual(s) to contact if I have questions about the research;

I understand that participation in research is not required, is voluntary, and that I may refuse to participate further without negative consequences;

I adhere to the preventative measures with regards to COVID-19 as explained above;

By signing this form, I am stating that I understand the above information and consent to participate in this study being conducted at TU Delft.

Name: _____

Participant Nr.: _____

Signature: _____

Date: _____

N

Qualitative Results Experiment

N.1. Qualitative Results

Note that all participants were Dutch and all comments have been translated from Dutch to English. At the end of the experiment, each participant was asked what they liked and disliked about the experiment and the adaptation strategies. Additionally, participants were asked whether they noticed a difference between the conditions and in which condition they believed the cooperation with the haptic steering system was the best. For simplicity, the positive reinforcement adaptation strategy is denoted by *pos.* and the negative reinforcement adaptation strategy is denoted by *neg.*. Manual control is denoted by *man.*

N.1.1. Participant 1

- C1 [pos.]: "This was not really my score",
- C2 [neg.]: "I'm just playing catch-up".
- "I noticed a clear difference, the first condition [pos.] helped a little bit, the second [neg.]".
- "I was along for the ride for the second condition [neg.]".
- "I didn't notice the system to change"

N.1.2. Participant 2

- C2 [pos.]: "It felt as if the controller was stopping my movement",
- C3 [neg.]: "The system suddenly helped a lot when there was a big deviation with the reference".
- "I liked C2 [pos.] best, C3 [neg.] was a bit too dominant".
- "I was cooperating most with C2 [pos.], C3 did most of the work [neg.]".

N.1.3. Participant 3

- "I was just steering lightly in C2 [neg.], best cooperation with C1 [pos.]".
- "C2 [neg.] felt a bit unnatural".

N.1.4. Participant 4

- C3 [neg.]: "Annoying to add very few to the steering system".
- "I noticed a clear difference, in C3 [neg.] I was like, "you do it yourself", C2 [pos.] was nicer, better cooperation".
- "I felt like I was contributing negatively in C3 [neg.]".

N.1.5. Participant 5

- "I noticed no clear difference between the conditions".
- "I was helped a bit more for large movements".
- "I didn't feel anything at C3 [neg.]"
- Experimenter note: This participant was by far the best, and his idea of which controller was helping him did not match with what was really presented. This is since the participant outperformed the controller and thus received no feedback in the negative reinforcement adaptation strategy.

N.1.6. Participant 6

- C1 [man.]: "I'm not sure whether the steering wheel was helping me",
- C2 [pos.]: "It [steering system] felt a bit random", "At the end, I was being helped a bit more".
- "Is there like a machine learning thing on this?".
- "I noticed a bit of difference between C2 [pos.] and C3 [neg.], C2 [pos.] worked quite well, felt like I could do nothing at C3 [neg.]".
- "I didn't notice the system to change"

N.1.7. Participant 7

- "Were they different?"
- "I noticed that C2 [neg.] was helping more than C3 [pos.]"
- "I liked C3 [neg.] most, cooperation was best in C2 [pos.]".

N.1.8. Participant 8

- C1 [neg.]: "I couldn't do anything, but later on, I could", "I felt like I was stopping the steering wheel to overshoot",
- "The difference was that one was steering very hard [neg.], the other a bit soft [pos.]".
- "The soft one [pos.] I liked most, I felt like being played by the hard one [neg.], didn't like that one".

N.1.9. Participant 9

- C2 [neg.]: "This is quite fun actually", "I quite like the assistance"
- "C2 [neg.] was a bit too much, C3 [pos.] felt more natural".
- "I felt like I was opposing the forces in C2 [neg.], liked C3 [pos.] better".
- "C2 [neg.] was less cooperative".

N.1.10. Participant 10

- C1 [neg.]: "Quite nice, but what am I adding?"
- C2 [man.]: "I hope that this was manual?"
- "The difference was in the amount of help, I like the strongest one [neg.] best, had to do more in C3 [pos.]".

N.1.11. Participant 11

- C2 [neg.]: "Ah, that's cool", "I am searching what I am doing and what the controller is doing",
- "The difference between neg. and pos. was very subtle".

N.1.12. Participant 12

- C1 [neg.]: *Lets go of the steering wheel*
- C1 [neg.]: "I am only a nuisance to the controller",
- "C1 [neg.] was just forcing me, was heavier. C3 [pos.] was less powerfull".
- "Not really cooperation."

N.1.13. Participant 13

- C1 [neg.]: "I might as well have done nothing",
- C2 [pos.]: "This is different".
- "With the one controller [neg.] you had no influence, with the other [pos.] you had".
- "Best cooperation with pos."

N.1.14. Participant 14

- C1 [pos.]: "Just going along, quite annoying, hard to estimate was to do in large deviations",
- C2 [man.]: "Going much better".
- C3 [neg.]: "Much nicer, much easier. Starting to get this game.".
- "Would the performance be better if I just let go in neg.?".
- "Not such a successful cooperation in C1 [pos.]"
- Experimenter note: This participant was also very good at the task, and its comments on the controllers were quite opposite to that of other participants. This participant also showed signs of meta-adaptation between neg. trials.

N.1.15. Participant 15

- C1 [neg.]: "It felt like I was first opposing the controller and later on started to go along with the controller",
- C2 [pos.]: "Bit better, I have more control now".
- "I noticed a difference between the controller in how much force they delivered."
- "I didn't notice a difference during a trial".
- "Best cooperation in C2 [pos.]".

N.1.16. Participant 16

- C1 [pos.]: "I noticed that my overshoot decreased and my reaction time was increased",
- C2 [man.]: "Much harder this one",
- C3 [neg.]: "I had less authority so I started to do less".
- "Main difference was that C1 [pos.] was more comfortable, and C3 [neg.] was a bit too much"

N.1.17. Participant 17

- C1 [neg.]: "Quite nice that it is helping me",
- C2 [pos.]: "This one is a bit harder, when I steered a lot it was like it was working against me".
- C3 [man.]: "This is way harder.".
- "I like the first one [neg.] more, not noticing that it was helping me, but went a bit smoother".
- "Second one [pos.] was a very much present"
- "Cooperation best with [neg.], not with [pos.] as it worked a bit more against me, but maybe I was working against the controlled system".

N.1.18. Participant 18

- C1 [pos.]: "Quite fun, you start to pay more attention to the haptic after a few trials",
- C2 [man.]: "I did notice a difference, this was manual".
- C3 [neg.]: "Felt like sometimes it was reacting more".
- "Main difference between C1 and C3 was that C3 was more dominant, I was a bit more following".
- "C1 [pos.] yielded the best cooperation, in C3 [neg.] it felt like I was giving the final push".

N.2. Analysis

Participants showed very different behavior and attitudes towards the adaptation strategies.

Often, humans reported that the more subtle controllers were more pleasant. Interestingly, this was irrespective of which type of adaptation strategy was presented to the human. Often, if a participant with a high skill level interacted with the negative reinforcement adaptation strategy controller, the amount of feedback by the controller would be fairly low, resulting in a liking towards this condition. Participant 14, who showed a high skill level, reported being annoyed by the positive reinforcement adaptation strategy controller, whereas it showed a liking to the negative reinforcement adaptation strategy. On the other hand, the participants with a slightly lower skill level clearly showed a liking toward the positive reinforcement adaptation strategy, due to the subtlety of the controller. For these participants, the controllers with a negative reinforcement adaptation strategy were perceived as too dominant.

There exists a clear difference between participants in how much difference they felt between different conditions. Again, this difference is attributed to the skill level of the participants. Participants that naturally performed much lower in the control task usually noticed the negative reinforcement adaptation strategy to be very dominant even to the point that participant 11 let go of the steering wheel as it perceived itself to be a nuisance to the controller. These participants attributed the difference between the controllers to be a difference in the amount of haptic feedback they received. On the other hand, participants that naturally performed fairly well noticed a much lower difference between the conditions.

O

Figures

In this appendix, some additional figures are presented,

O.1. Trial data

The following figures give an overview of the measured signals that are used to determine the metrics to do statistical analysis. In the right column, the distribution of the signals is shown. Since these are usually not normally distributed, the median is taken instead of the mean value. The median is indicated on the y-axis.

The following data is taken from participant 18, who was fairly good in the control task, showing representative data.

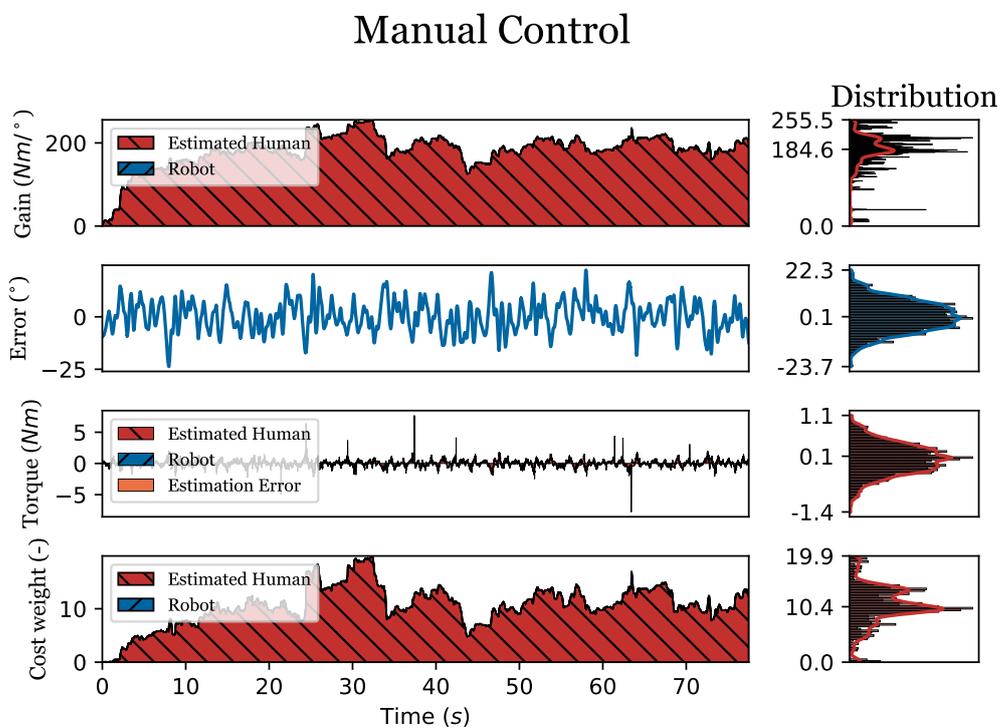


Figure O.1: Overview of metrics for trial 1 of participant 18 in the manual control condition.

Positive Reinforcement

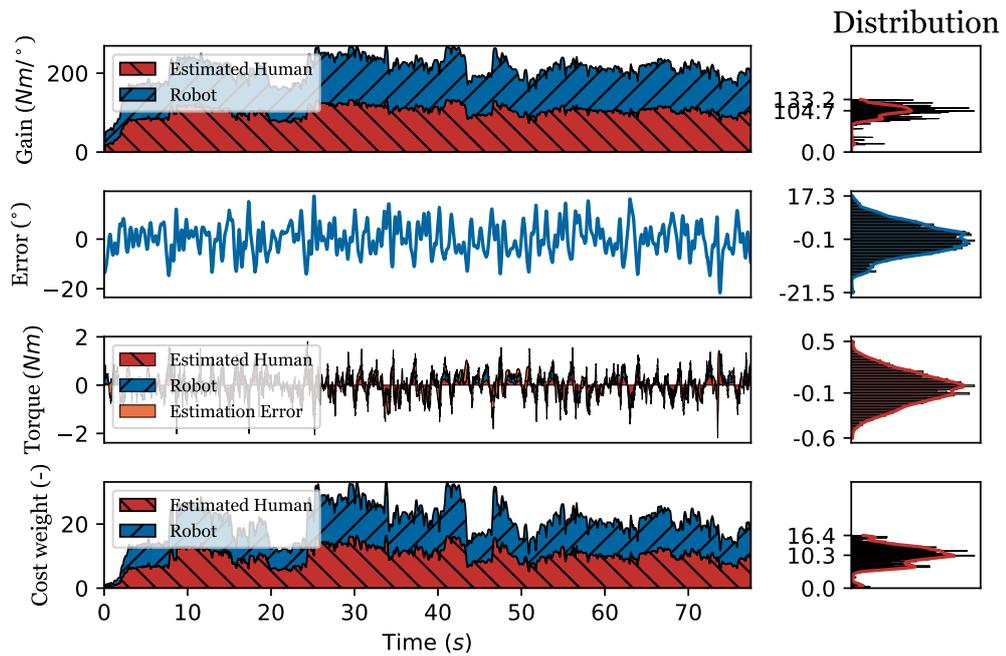


Figure O.2: Overview of metrics for trial 1 of participant 18 in the positive reinforcement condition.

Negative Reinforcement

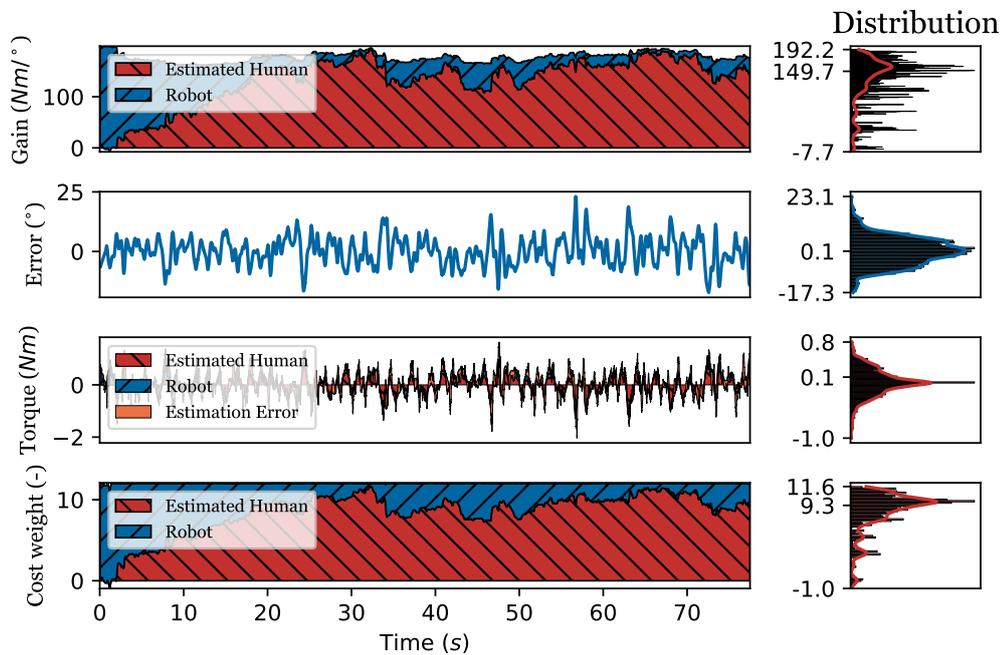


Figure O.3: Overview of metrics for trial 1 of participant 18 in the negative reinforcement condition.

The following data is taken from participant 12, who scored the lowest performance in the manual control task, thus showing how the robot interacted with a participant with low skill level.

Manual Control

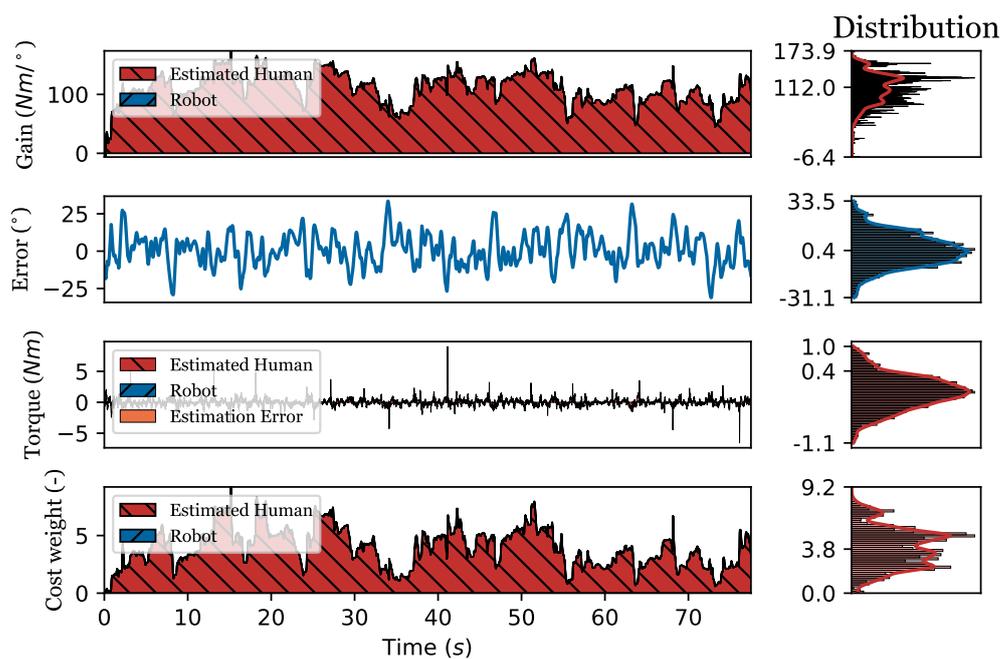


Figure O.4: Overview of metrics for trial 1 of participant 12 in the manual control condition.

Negative Reinforcement

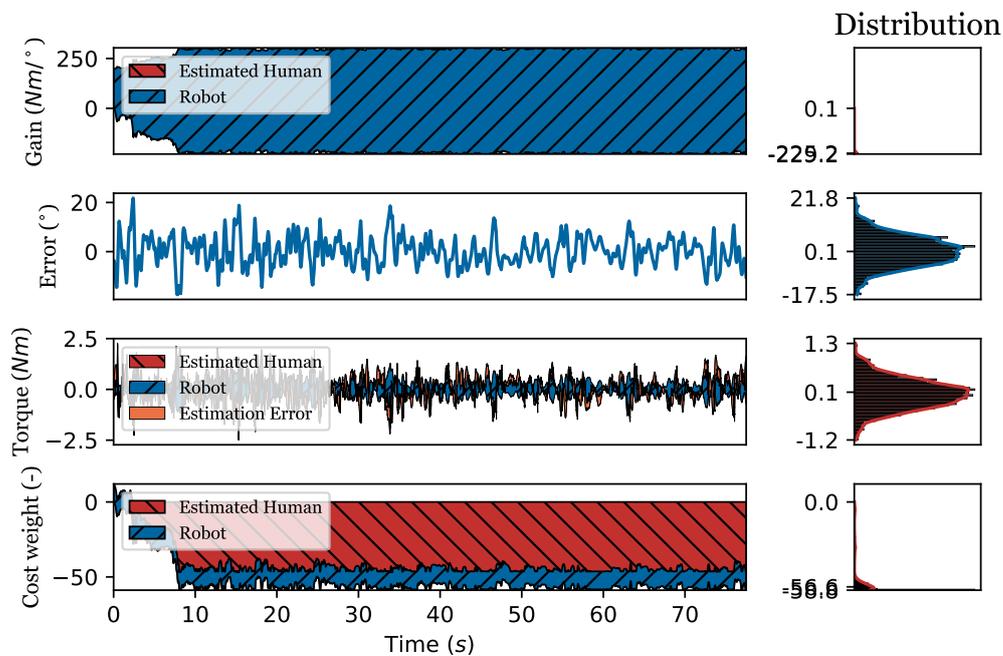


Figure O.5: Overview of metrics for trial 1 of participant 12 in the positive reinforcement condition.

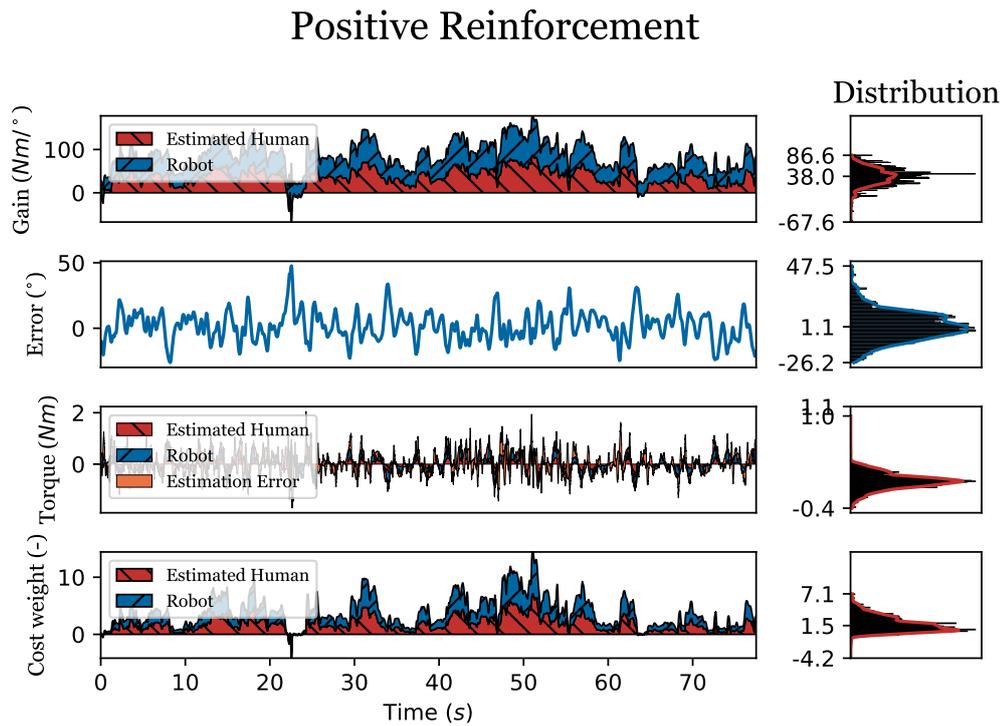


Figure O.6: Overview of metrics for trial 1 of participant 12 in the negative reinforcement condition.

O.2. Participant data

The following data is taken from participants 12, 15, and 18 respectively. Participant 12 illustrates what happens when the skill level of a participant is fairly low. Participant 15 demonstrates trial-by-trial adaptation, demonstrating the slacking hypothesis [29]. Participant 18 demonstrates a fairly high control skill.

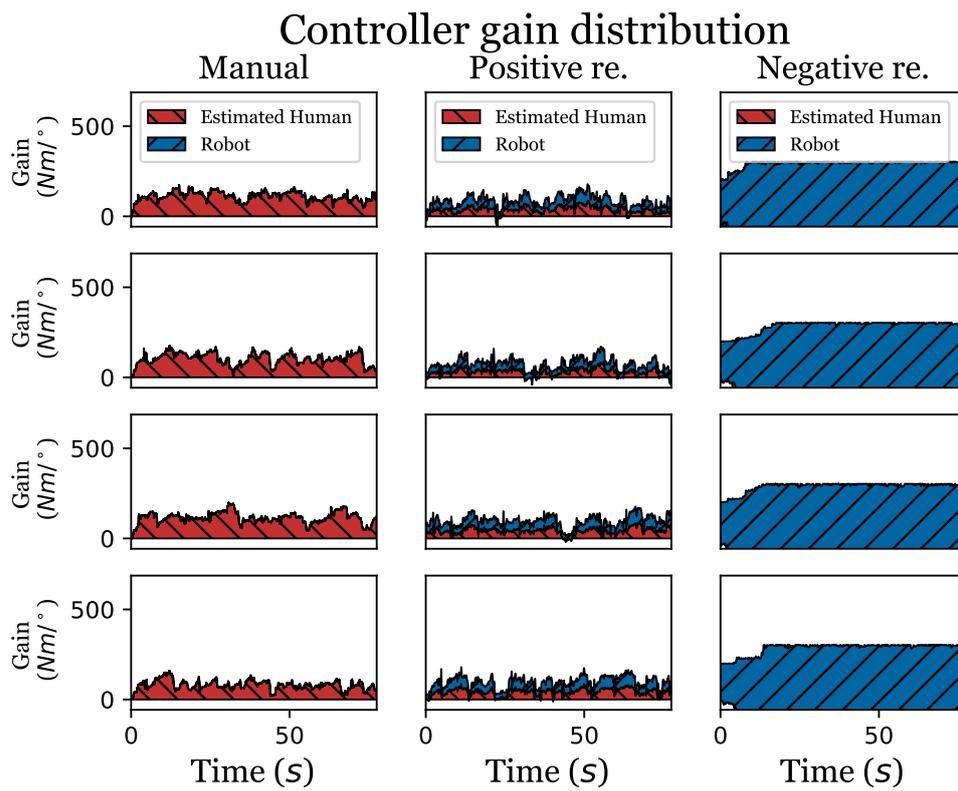


Figure O.7: Controller gain distribution for all trials of participant 12.

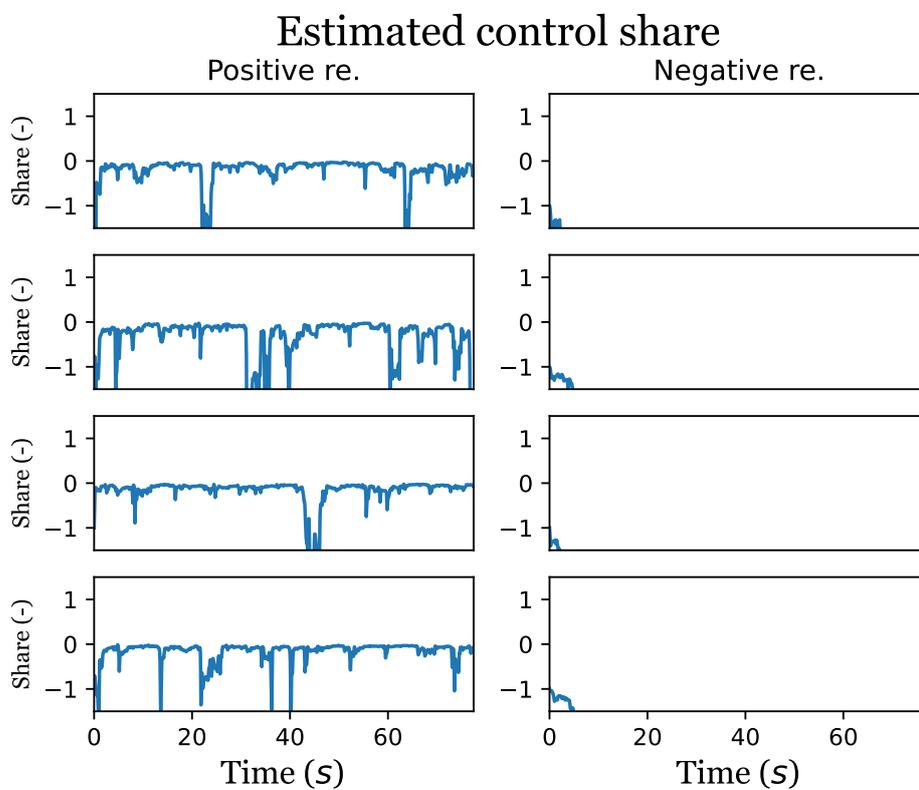


Figure O.8: Control share metric for all trials of participant 12.

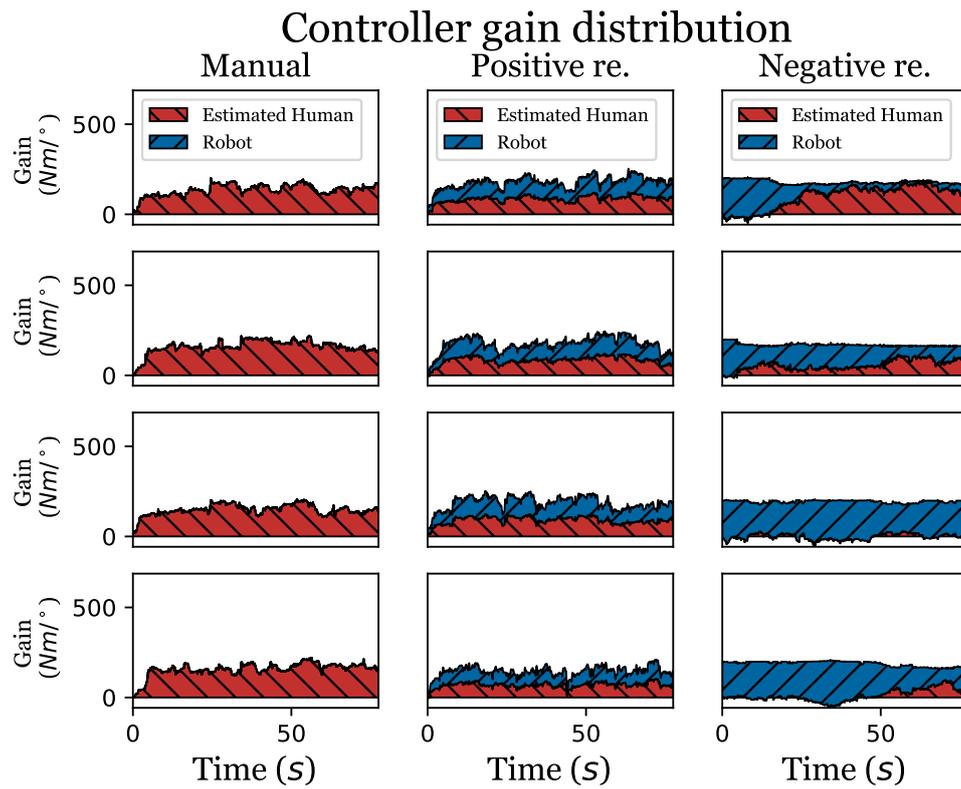


Figure O.9: Controller gain distribution for all trials of participant 15.

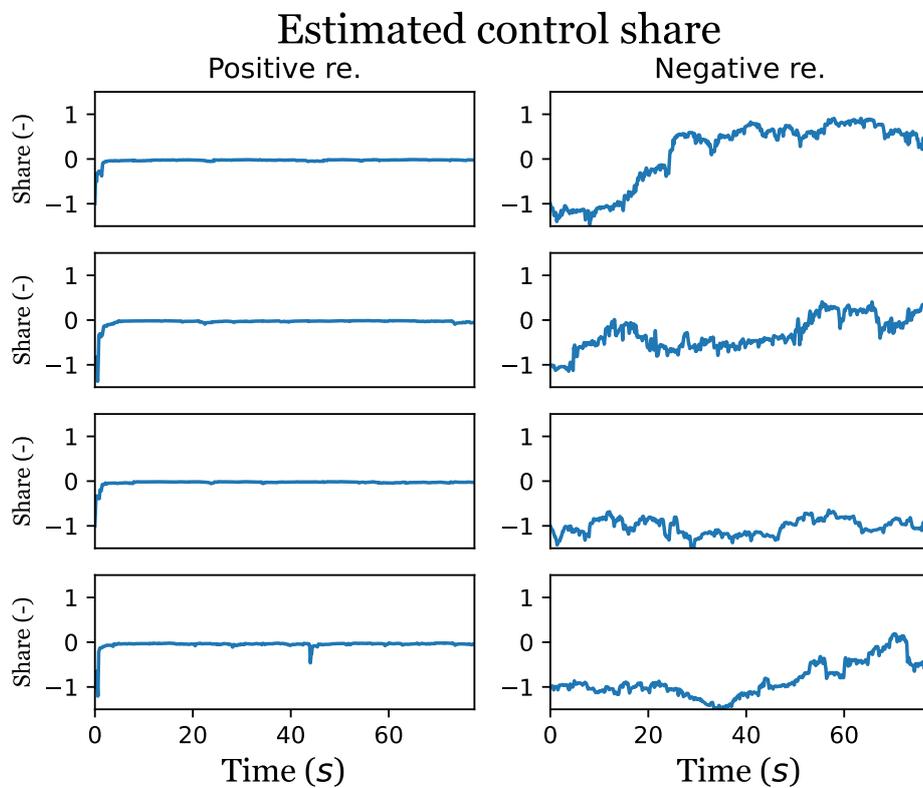


Figure O.10: Control share metric for all trials of participant 15.

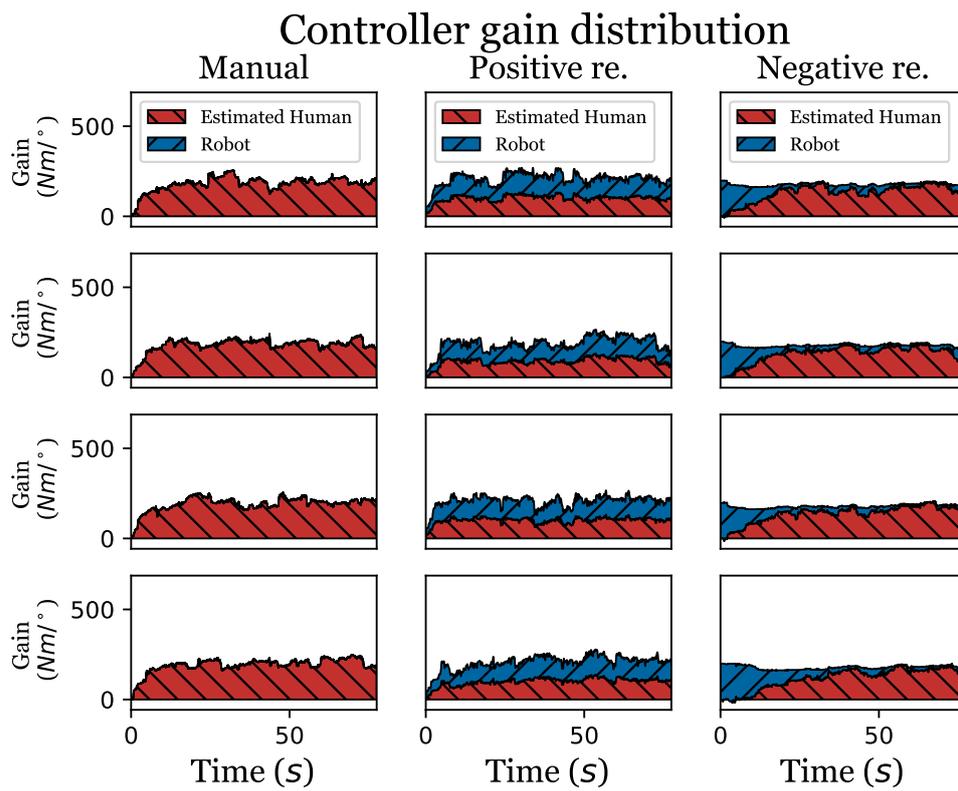


Figure O.11: Controller gain distribution for all trials of participant 18.

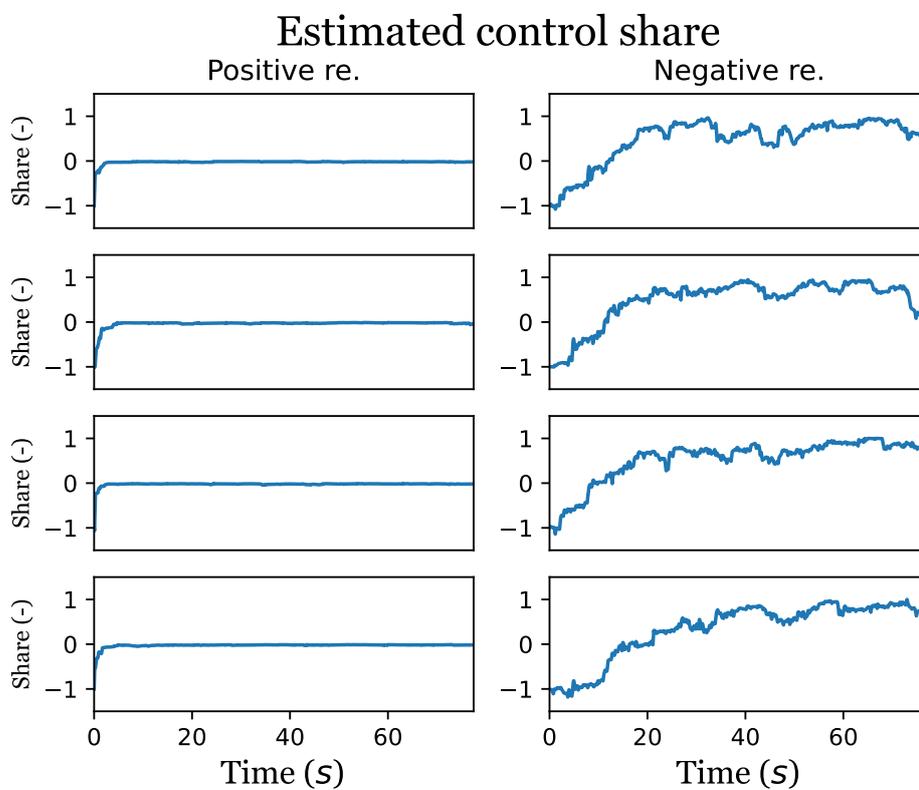


Figure O.12: Control share metric for all trials of participant 18.

O.3. Experiment overview

O.3.1. Box-plots

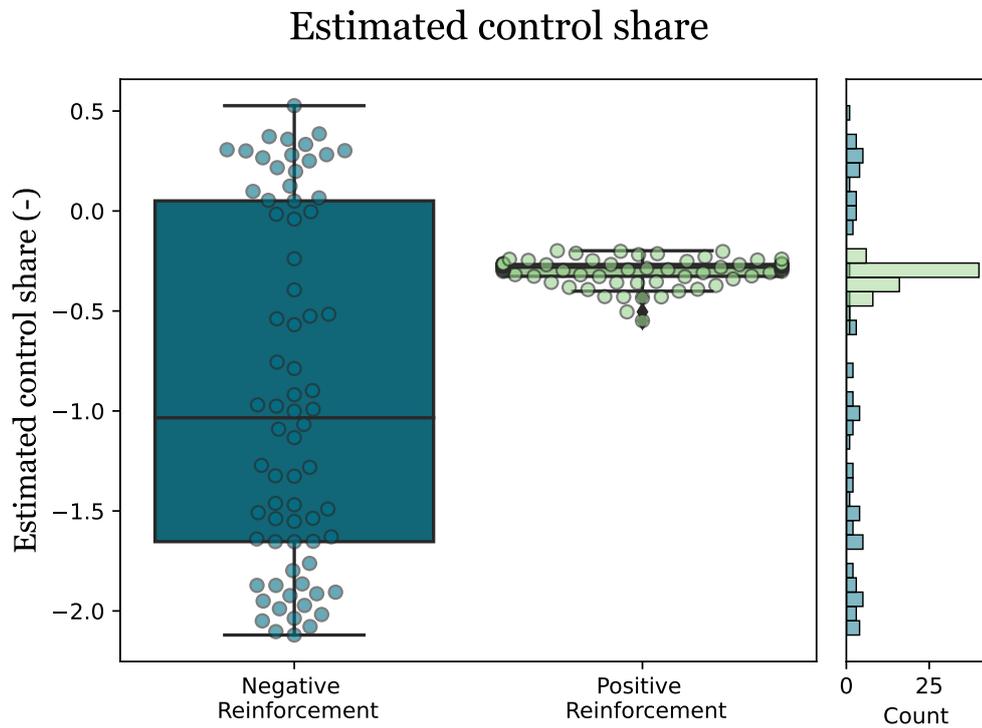


Figure O.13: Box-plot, comparing the median estimated control share metric between conditions.

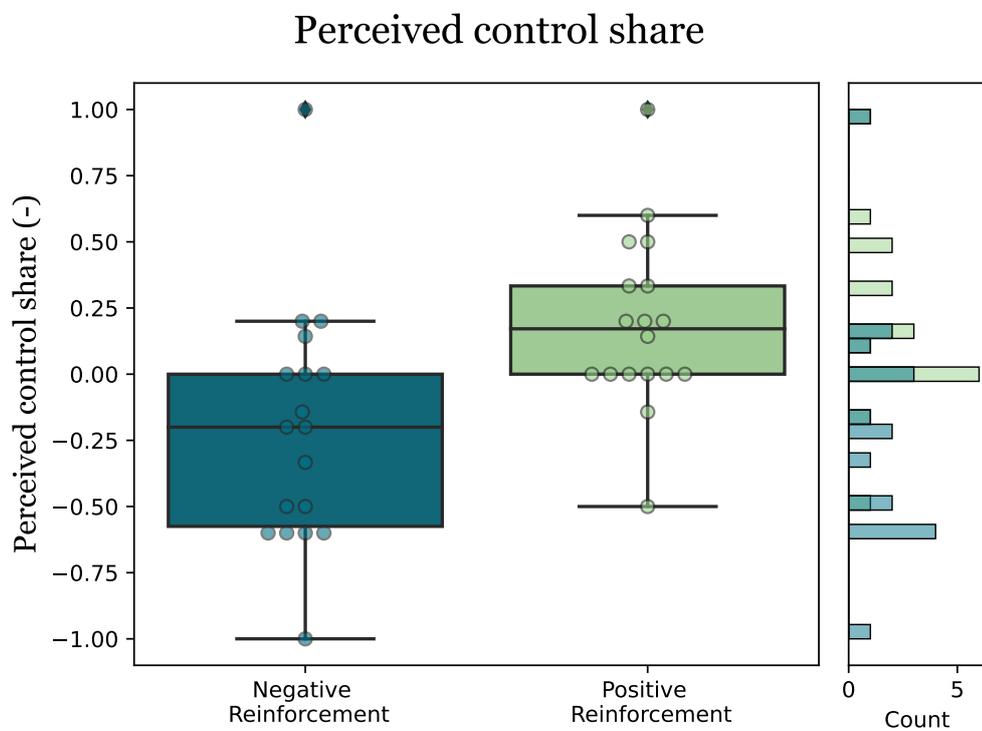


Figure O.14: Box-plot, comparing the perceived control share metric between conditions.

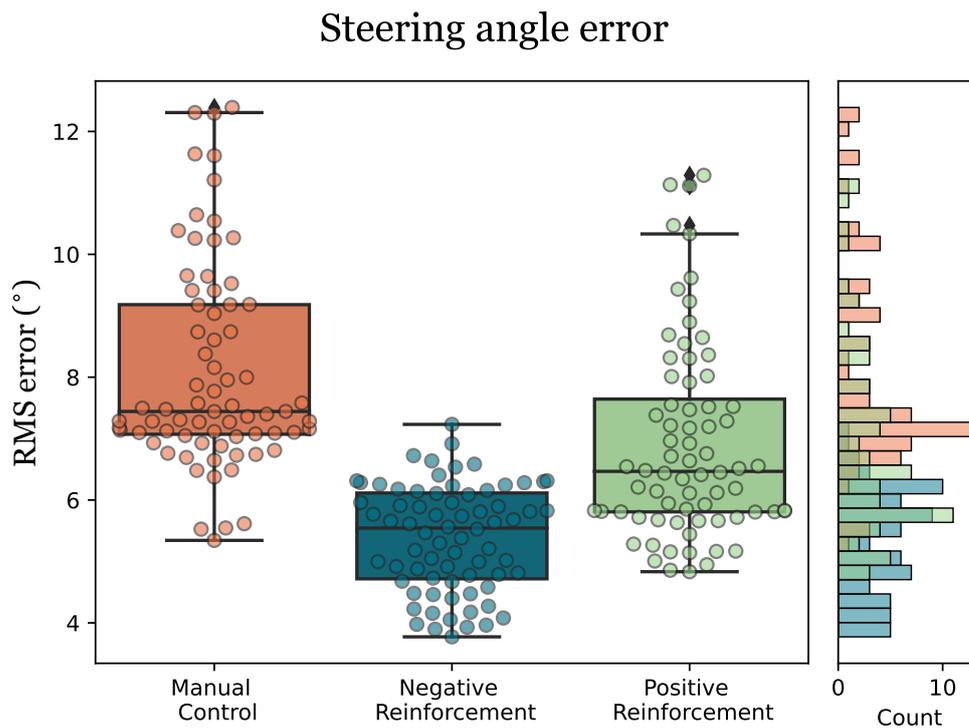


Figure O.15: Box-plot, comparing the root-mean-square steering angle error metric between conditions.

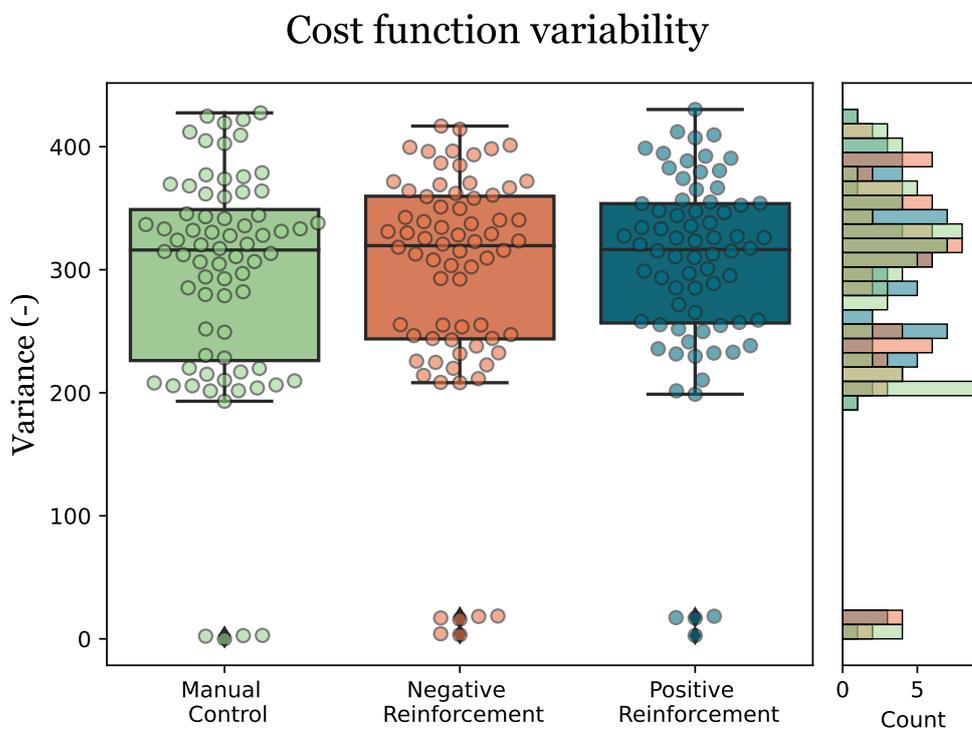


Figure O.16: Box-plot, comparing the cost function variability, within a trial, between conditions.

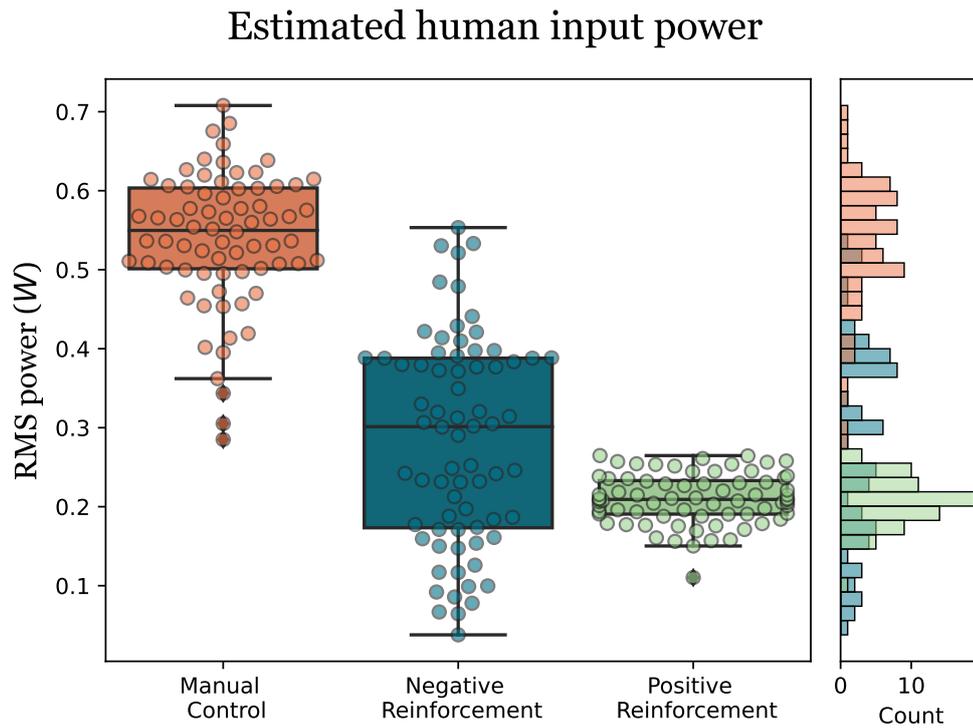


Figure O.17: Box-plot, comparing the root-mean-square estimated human input power metric between conditions.

Note that the estimated human input power may be a somewhat flawed metric, as the estimated human input power might be very low for the negative reinforcement condition due to how the controller estimates the human control inputs. The cost function variability metric shows that the co-adaptation with a trial is not significantly differing from one another, thus showing that convergence of the co-adaptation is not significantly different between conditions.

O.3.2. Interpretation of results

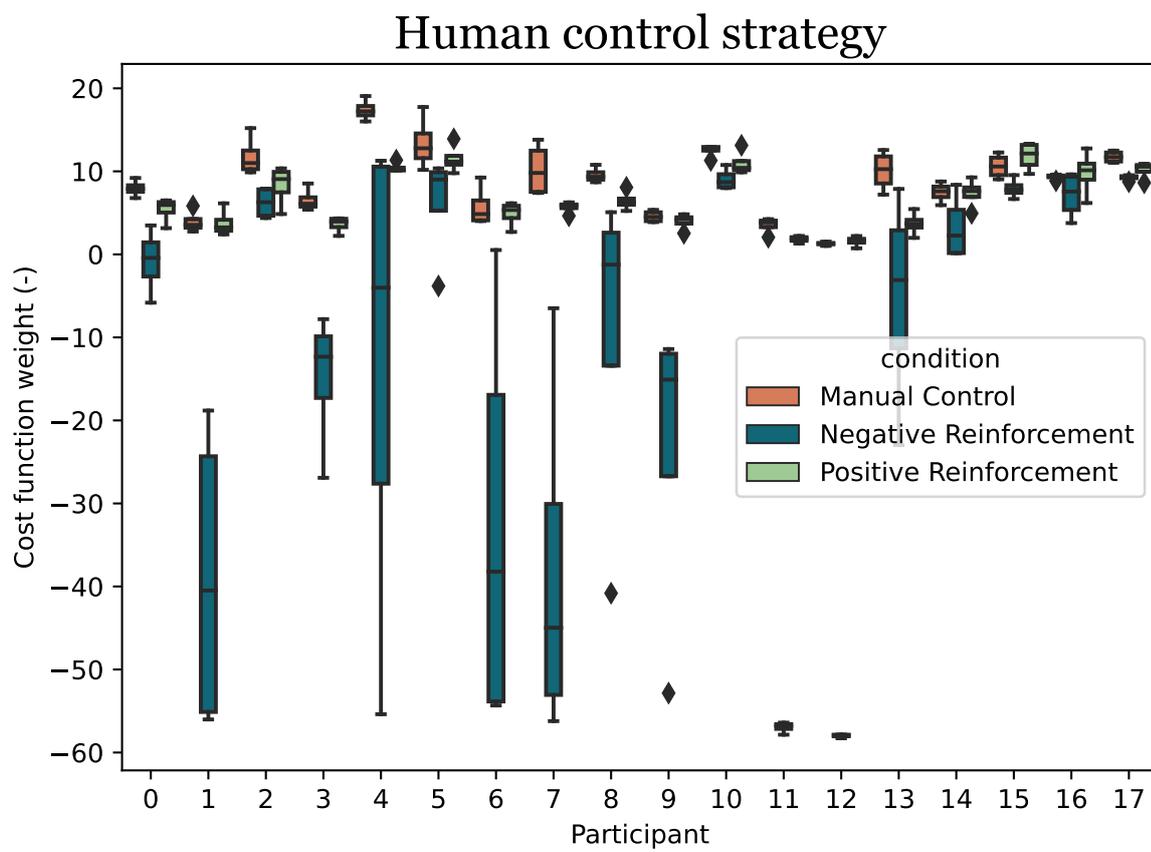
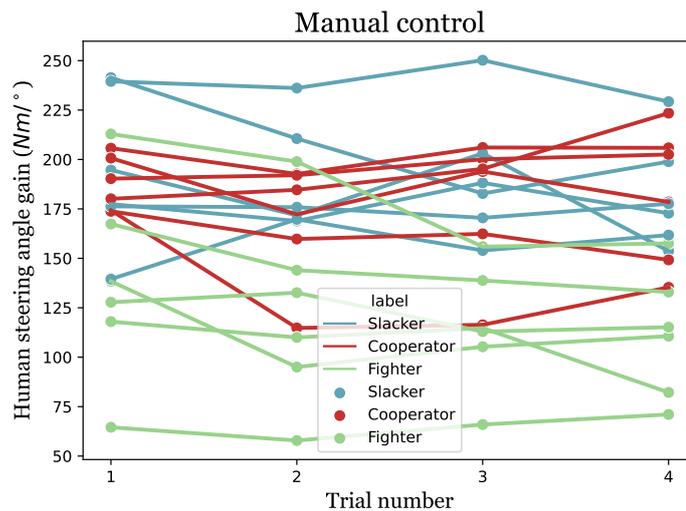
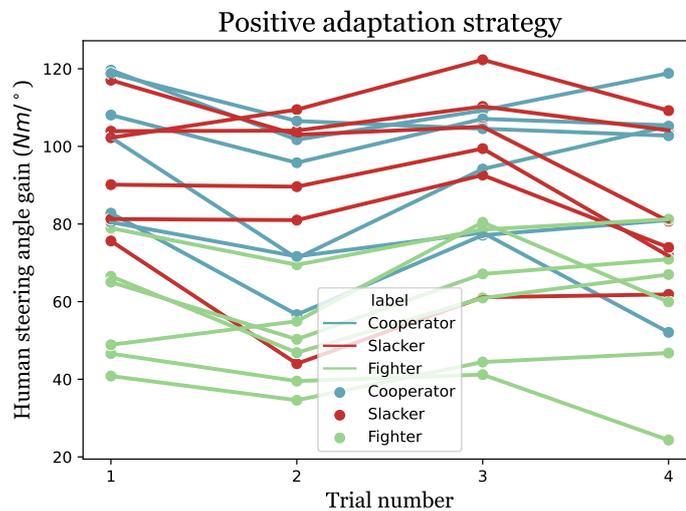


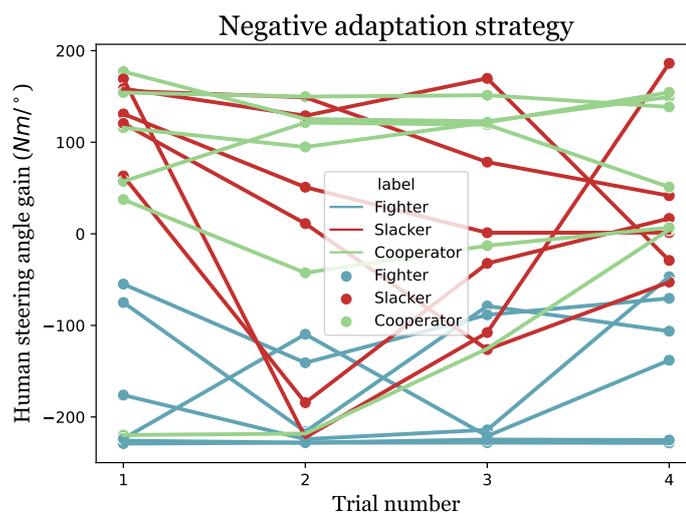
Figure O.18: Overview of the estimated control strategies for each human over all trials. A large variety of behaviors is observed.



(a) This figure demonstrates how for manual control, humans converge to different controller gains over different trials.



(b) This figure demonstrates how for positive reinforcement, humans converge to different controller gains over different trials.



(c) This figure demonstrates how for negative reinforcement, humans converge to different controller gains over different trials.

Figure O.19: The figure demonstrate how for each of the conditions, humans converge to different controller gains over different trials. Using colors, the behaviors that are observed in the negative reinforcement adaptation strategy are illustrated. It shows that a lot more variation is found in controller gains over trials in the negative reinforcement adaptation strategy, compared to the other conditions. Moreover, the behaviors shown using the colors are not observed in the other conditions, demonstrating that this phenomenon is not human-dependent.

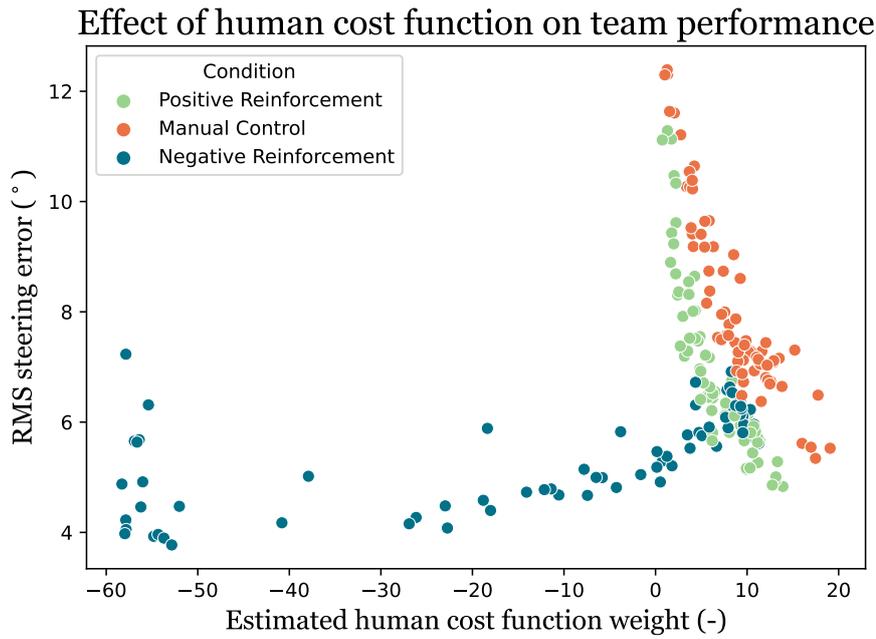


Figure O.20: Reducing control effort, leading to a lower estimated cost function weight, is rewarded with higher control performance in the negative reinforcement adaptation strategy. The positive reinforcement and manual control conditions exhibit the opposite phenomenon.

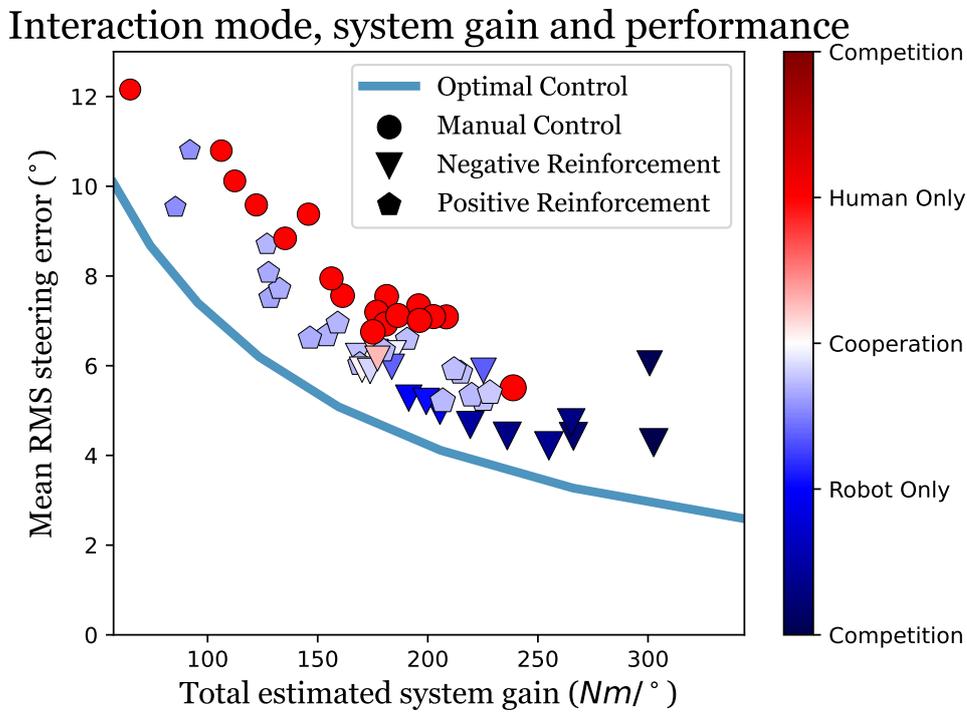


Figure O.21: Competition leads to the highest total system controller gain in the negative reinforcement adaptation strategy, which results in the highest control performance. The control performance and system controller gain in the positive reinforcement adaptation strategy is slightly increased compared to the manual control condition.

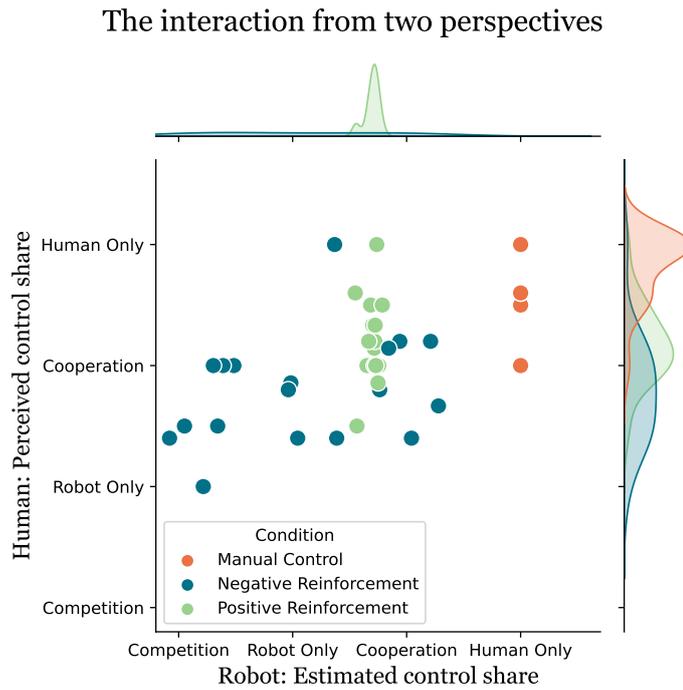


Figure O.23: The perspectives of both the team members in the human-robot team are shown. The perspective of the robot is the *estimated control share* and the perspective of the human is the *perceived control share*. Interestingly, although these perspectives usually coincide, sometimes they do not.

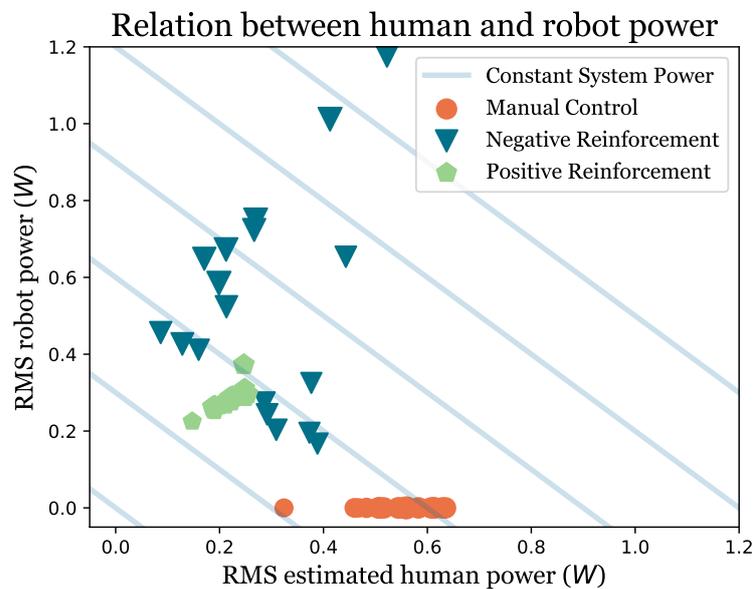


Figure O.22: The manual control and positive reinforcement conditions show a low variability between estimated human and robot power, operating almost all the times within a system power band of 0.3 and 0.7 Watts. However, the negative reinforcement condition shows a large variability, and an asymmetry between estimated human and robot power can be observed.

Again, note that the estimated human power may be a flawed metric. The asymmetry between estimated human and robot power is due to the expectation of the robot to be competing with the human.

Bibliography

- [1] David A. Abbink, Erwin R. Boer, and Mark Mulder. Motivation for continuous haptic gas pedal feedback to support car following. *IEEE Intelligent Vehicles Symposium, Proceedings*, pages 283–290, 2008. doi: 10.1109/IVS.2008.4621325.
- [2] David A. Abbink, Mark Mulder, and Erwin R. Boer. Haptic shared control: Smoothly shifting control authority? *Cognition, Technology and Work*, 14(1):19–28, mar 2012. ISSN 14355558. doi: 10.1007/s10111-011-0192-5.
- [3] Victoria A. Banks, Katherine L. Plant, and Neville A. Stanton. Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016. *Safety Science*, 108(August 2017):278–285, 2018. ISSN 18791042. doi: 10.1016/j.ssci.2017.12.023.
- [4] Richard Bellman and D. S. Mitrinovic. *Elementary Inequalities*, 1965. ISSN 00255718.
- [5] Rolf Boink, Marinus M. Van Paassen, Mark Mulder, and David A. Abbink. Understanding and reducing conflicts between driver and haptic shared control. In *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, volume 2014-Janua, pages 1510–1515. Institute of Electrical and Electronics Engineers Inc., 2014. doi: 10.1109/smc.2014.6974130.
- [6] Barry Brown and Eric Laurier. The trouble with autopilots: Assisted and autonomous driving on the social road. *Conference on Human Factors in Computing Systems - Proceedings*, 2017-May:416–429, 2017. doi: 10.1145/3025453.3025462.
- [7] Vinil T. Chackochan and Vittorio Sanguineti. Incomplete information about the partner affects the development of collaborative strategies in joint action. *PLoS Computational Biology*, 15(12):1–23, 2019. ISSN 15537358. doi: 10.1371/journal.pcbi.1006385.
- [8] H. J. Damveld, G. C. Beerens, M. M. Van Paassen, and M. Mulder. Design of forcing functions for the identification of human control behavior. *Journal of Guidance, Control, and Dynamics*, 33(4):1064–1081, 2010. ISSN 15333884. doi: 10.2514/1.47730.
- [9] J. C.F. De Winter and D. Dodou. Preparing drivers for dangerous situations: A critical reflection on continuous shared control. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, pages 1050–1056, 2011. ISSN 1062922X. doi: 10.1109/ICSMC.2011.6083813.
- [10] Sidney Dekker. The field guide to understanding human error. *The Field Guide to Understanding Human Error*, (August 2000):1–236, 2013. ISSN 0014-0139. doi: 10.1080/00140130701680544.
- [11] Thomas A. Dingus, Feng Guo, Suzie Lee, Jonathan F. Antin, Miguel Perez, Mindy Buchanan-King, and Jonathan Hankey. Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences of the United States of America*, 113(10):2636–2641, 2016. ISSN 10916490. doi: 10.1073/pnas.1513271113.
- [12] Jacob Engwerda. Algorithms for computing Nash equilibria in deterministic LQ games. pages 113–140, 2007. doi: 10.1007/s10287-006-0030-z.
- [13] Michael Flad, Lukas Frohlich, and Soren Hohmann. Cooperative shared control driver assistance systems based on motion primitives and differential games. *IEEE Transactions on Human-Machine Systems*, 47(5):711–722, 2017. ISSN 21682291. doi: 10.1109/THMS.2017.2700435.
- [14] Frank Flemisch, Eugen Altendorf, Yigiterkut Canpolat, Gina Weßel, Marcel Baltzer, Daniel Lopez, Nicolas Daniel Herzberger, Mechthild Irmgard Voß, Maximilian Schwalm, and Paul Schutte. Uncanny and Unsafe Valley of Assistance and Automation: First Sketch and Application of Vehicle Automation. *Advances in Ergonomic Design of Systems, Products and Processes*, (September 2018), 2017. doi: 10.1007/978-3-662-53305-5.

- [15] Luke Fletcher, Seth Teller, Edwin Olson, David Moore, Yoshiaki Kuwata, Jonathan How, John Leonard, Isaac Miller, Mark Campbell, Dan Huttenlocher, Aaron Nathan, and Frank Robert Kline. *The MIT - Cornell collision and why it happened*, volume 56. 2009. ISBN 9783642039904. doi: 10.1007/978-3-642-03991-1_12.
- [16] Amir H. Ghasemi. Game theoretic modeling of a steering operation in a haptic shared control framework. *ASME 2018 Dynamic Systems and Control Conference, DSCC 2018*, 2(September), 2018. doi: 10.1115/DSCC2018-9105.
- [17] Raphaëla Groten, Daniela Feth, Harriet Goshy, Angelika Peer, David A. Kenny, and Martin Buss. Experimental analysis of dominance in haptic collaboration. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pages 723–729, 2009. doi: 10.1109/ROMAN.2009.5326315.
- [18] Patrick Guillaume, Johan Schoukens, Rik Pintelon, and István Kollái. Crest-Factor Minimization Using Nonlinear Chebyshev Approximation Methods. *IEEE Transactions on Instrumentation and Measurement*, 40(6):982–989, 1991. ISSN 15579662. doi: 10.1109/19.119778.
- [19] Yuval Noah Harari. *21 Lessons for the 21st Century*. Random House, 2018.
- [20] Makoto Itoh, Frank Flemisch, and David Abbink. A hierarchical framework to analyze shared control conflicts between human and machine. *IFAC-PapersOnLine*, 49(19):96–101, 2016. ISSN 24058963. doi: 10.1016/j.ifacol.2016.10.468.
- [21] Vahid Izadi, Akshay Bhardwaj, and Amir H. Ghasemi. Impedance Modulation for Negotiating Control Authority in a Haptic Shared Control Paradigm. *Proceedings of the American Control Conference*, 2020-July:2478–2483, 2020. ISSN 07431619. doi: 10.23919/ACC45564.2020.9147984.
- [22] Nathanaël Jarrassé, Vittorio Sanguineti, and Etienne Burdet. Slaves no longer: Review on role assignment for human-robot joint motor action. *Adaptive Behavior*, 22(1):70–82, 2014. ISSN 10597123. doi: 10.1177/1059712313481044.
- [23] Xuewu Ji, Kaiming Yang, Xiaoxiang Na, Chen Lv, and Yahui Liu. Shared Steering Torque Control for Lane Change Assistance: A Stochastic Game-Theoretic Approach. *IEEE Transactions on Industrial Electronics*, 66(4):3093–3105, 2019. ISSN 02780046. doi: 10.1109/TIE.2018.2844784.
- [24] H.K. Khalil. *Nonlinear Systems*. 2002. ISBN 130673897.
- [25] Miltos Kyriakidis, Carlo van de Weijer, Bart van Arem, and Riender Happee. The Deployment of Advanced Driver Assistance Systems in Europe. *SSRN Electronic Journal*, 2012(October):5–9, 2015. ISSN 1556-5068. doi: 10.2139/ssrn.2559034.
- [26] Y. Li, G. Carboni, F. Gonzalez, D. Campolo, and E. Burdet. Differential game theory for versatile physical human-robot interaction. *Nature Machine Intelligence*, 1(1):36–43, jan 2019. ISSN 25225839. doi: 10.1038/s42256-018-0010-3.
- [27] Anders Lindgren and Fang Chen. *State of the Art Analysis: An Overview of Advanced Driver Assistance Systems (ADAS) and Possible Human Factors Issues*. 2007. ISBN 9789173939997.
- [28] Shibendu Mahata, Suman Kumar Saha, Rajib Kar, and Durbadal Mandal. Optimal design of fractional order low pass Butterworth filter with accurate magnitude response. *Digital Signal Processing: A Review Journal*, 72:96–114, 2018. ISSN 10512004. doi: 10.1016/j.dsp.2017.10.001.
- [29] Laura Marchal-Crespo and David J. Reinkensmeyer. Review of control strategies for robotic movement training after neurologic injury. *Journal of NeuroEngineering and Rehabilitation*, 6(1), 2009. ISSN 17430003. doi: 10.1186/1743-0003-6-20.
- [30] Andrii Mironchenko and Christophe Prieur. Input-to-state stability of infinite-dimensional systems: Recent results and open questions. *SIAM Review*, 62(3):529–614, 2020. ISSN 00361445. doi: 10.1137/19M1291248.

- [31] Alexander Mörtl, Martin Lawitzky, Ayse Kucukyilmaz, Metin Sezgin, Cagatay Basdogan, and Sandra Hirche. The role of roles: Physical cooperation between humans and robots. *International Journal of Robotics Research*, 31(13):1656–1674, 2012. ISSN 02783649. doi: 10.1177/0278364912455366.
- [32] Mark Mulder, David A. Abbink, and Erwin R. Boer. Sharing control with haptics: Seamless driver support from manual to automatic control. *Human Factors*, 54(5):786–798, 2012. ISSN 00187208. doi: 10.1177/0018720812443984.
- [33] Max Mulder, Daan M. Pool, David A. Abbink, Erwin R. Boer, Peter M.T. Zaal, Frank M. Drop, Kasper Van Der El, and Marinus M. Van Paassen. Manual Control Cybernetics: State-of-the-Art and Current Trends. *IEEE Transactions on Human-Machine Systems*, 48(5):468–485, 2018. ISSN 21682291. doi: 10.1109/THMS.2017.2761342.
- [34] Xiaoxiang Na and David J. Cole. Game-theoretic modeling of the steering interaction between a human driver and a vehicle collision avoidance controller. *IEEE Transactions on Human-Machine Systems*, 45(1):25–38, 2015. ISSN 21682291. doi: 10.1109/THMS.2014.2363124.
- [35] John F. Nash. The Bargaining Problem. *Econometrica*, 18(2):155, apr 1950. ISSN 00129682. doi: 10.2307/1907266.
- [36] Lior Noy, Erez Dekel, and Uri Alon. The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences of the United States of America*, 108(52):20947–20952, 2011. ISSN 00278424. doi: 10.1073/pnas.1108155108.
- [37] Sebastiaan M. Petermeijer, David A. Abbink, Mark Mulder, and Joost C.F. De Winter. The Effect of Haptic Support Systems on Driver Performance: A Literature Survey. *IEEE Transactions on Haptics*, 8(4):467–479, 2015. ISSN 19391412. doi: 10.1109/TOH.2015.2437871.
- [38] Louay Saleh, Philippe Chevrel, Fabien Claveau, Jean Francois Lafay, and Franck Mars. Shared steering control between a driver and an automation: Stability in the presence of driver behavior uncertainty. *IEEE Transactions on Intelligent Transportation Systems*, 14(2):974–983, 2013. ISSN 15249050. doi: 10.1109/TITS.2013.2248363.
- [39] Wietske Scholtens, Sarah Barendswaard, Daan Pool, Rene Van Paassen, and David Abbink. A New Haptic Shared Controller Reducing Steering Conflicts. *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, pages 2705–2710, 2019. doi: 10.1109/SMC.2018.00462.
- [40] SensoDrive GmbH. Software Manual SENSO-Wheel.
- [41] T. Specker, M. Buchholz, and K. Dietmayer. A new approach of dynamic friction modelling for simulation and observation. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 19:4523–4528, 2014. ISSN 14746670. doi: 10.3182/20140824-6-za-1003.01711.
- [42] Neville A. Stanton and Paul M. Salmon. Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science*, 47(2):227–237, 2009. ISSN 09257535. doi: 10.1016/j.ssci.2008.03.006.
- [43] Gang Toa. *Adaptive control design and analysis*, volume 37. 2003.
- [44] M. Touahmia. Identification of Risk Factors Influencing Road Traffic Accidents. *Engineering, Technology & Applied Science Research*, 8(1):2417–2421, 2018. ISSN 2241-4487. doi: 10.48084/etasr.1615.
- [45] M. M. (René) van Paassen, Rolf Pieter Boink, David A. Abbink, Mark Mulder, and Max Mulder. Four design choices for haptic shared control. In John M. Flach Michael A. Vidulich, Pamela S. Tsang, editor, *Advances in Aviation Psychology, Volume 2*, pages 237–254. Routledge, 2019. doi: 10.4324/9781315565712-12.
- [46] J. E. Vos. Reducing conflicting forces with co-adaptive haptic shared control using online time-varying operator identification. 2018.
- [47] World Health Organization. Global status report on road safety 2015, 2015. ISSN 1475-5785.