2018-04-01

Force and Motion Based Methods for Planar Human-Robot Co-manipulation of Extended Objects

Erich Allen Mielke
Brigham Young University

Force and Motion Based Methods for Planar Human-Robot Co-Manipulation of Extended Objects

Erich Allen Mielke

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Marc Killpack, Chair
Mark Colton
Mike Goodrich

Department of Mechanical Engineering

Brigham Young University

ABSTRACT

Force and Motion Based Methods for Planar Human-Robot
Co-Manipulation of Extended Objects

Erich Allen Mielke
Department of Mechanical Engineering, BYU
Master of Science

As robots become more common operating in close proximity to people, new opportunities arise for physical human-robot interaction, such as co-manipulation of extended objects. Co-manipulation involves physical interaction between two partners where an object held by both is manipulated in tandem. There is a dearth of viable high degree-of-freedom co-manipulation controllers, especially for extended objects, as well as a lack of information about how human-human teams perform in high degree-of-freedom tasks. One method for creating co-manipulation controllers is to pattern them off of human data. This thesis uses this technique by exploring a previously completed experimental study. The study involved human-human dyads in leader-follower format performing co-manipulation tasks with an extended object in 6 degrees of freedom. Two important tasks performed in this experiment were lateral translation and planar rotation tasks. This thesis focuses on these two tasks because they represent planar motion. Most previous control methods are for 1 or 2 degrees-of-freedom.

The study provided information about how human-human dyads perform planar tasks. Most notably, planar tasks generally adhere to minimum-jerk trajectories, and do not minimize interaction forces between users. The study also helped solve the translation versus rotation problem. From the experimental data, torque patterns were discovered at the beginning of the trial that defined intent to translate or rotate. From these patterns, a new method of planar co-manipulation control was developed, called Extended Variable Impedance Control. This is a novel 3 degree-of-freedom method that is applicable to a variety of planar co-manipulation scenarios.

Additionally, the data was fed through a Recursive Neural Network. The network takes in a series of motion data and predicts the next step in the series. The predicted data was used as an intent estimate in another novel 3 degree of freedom method called Neural Network Prediction Control. This method is capable of generalizing to 6 degrees of freedom, but is limited in this thesis for comparison with the other method.

An experiment, involving 16 participants, was developed to test the capabilities of both controllers for planar tasks. A dual manipulator robot with an omnidirectional base was used in the experiment. The results from the study show that both the Neural Network Prediction Control and Extended Variable Impedance Control controllers performed comparably to blindfolded human-human dyads. A survey given to participants informed us they preferred to use the Extended Variable Impedance Control. These two unique controllers are the major results of this work.

Keywords: physical human-robot interaction, co-manipulation, human-intent estimation

ACKNOWLEDGMENTS

I would like to thank my wife who has lovingly supported all my decisions in our marriage. During my academic career, she has faced many difficult situations, yet has been a constant source of support when I needed it, and has made my time in graduate school a joy. Though she doesn't share my level of enthusiasm for robotics, she has listened to me explain my technical problems, helped me find people for user studies, and brought me treats when she thought I needed a break.

I also would like to thank my parents for all they have done for me. They are examples of true scholars, and seem to never tire in helping their children. They taught me to do my best and to achieve my goals, and have provided advice and knowledge at every stage in my life. I can never repay them, and only hope that I can do the same for my children.

My thanks also goes to the Army Research Laboratory Robotics Collaborative Technology Alliance who funded a majority of this work. I also want to thank my graduate committee members, Mark Colton and Mike Goodrich, for all their help when I have come to them with questions, and for serving as members of my committee.

I am thankful for my fellow peers in the RaDLab. Gaining knowledge from them and growing my own capabilities has been extremely beneficial. Our friendship has grown as we have collaborated and interacted in the lab, and I will miss my daily interactions with them. I am grateful for all those who helped in pilot studies, or lent a hand when I was developing controllers. The experimental studies would not have been possible without the help of Eric Townsend, Nick Walton, Chris Gustafson, and Alex Jensen.

Finally, I would especially like to thank Marc Killpack, who has made this possible. He has provided me with much needed advice on all aspects of research and professional goals, and has given me room to grow as a student and as a person.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Problem Description

As robots become more common operating in close proximity to people, new opportunities arise for physical human-robot interaction (pHRI). One such application of pHRI is collaborative manipulation (or co-manipulation) of rigid extended objects (see Fig. 1.1). For the remainder of this thesis, co-manipulation is defined as the physical interaction between two or more partners in manipulating a jointly-held object. Extended objects are defined in this thesis as objects with significant length, width, and height, such that they are not easily moved by a single person. Also, all extended objects are assumed rigid unless specified otherwise. This area of robotics, pHRI, takes advantage of the strengths of both members of the human-robot team: strength and execution from the robot and intelligence and experience from the human. Co-manipulation can include anywhere from simple 1 degree-of-freedom (DOF) tasks up to complex 6 DOF tasks. Dyadic co-manipulation–which is co-manipulation between a pair of users–can involve either a leader-follower or a shared-leader setup.

A leader-follower setup involves one user making decisions for the dyad and initiating movement, whereas a shared-leader setup involves both users taking initiative in decisions and movement. For this thesis, the leader-follower setup is used, with the human leading in the human-robot dyad. Because humans have more expertise in the tasks they wish to perform–given state-of-the-art technology–beginning with a leader-follower setup is an intuitive starting point for a control basis. A future goal for this project is to have both shared-leader and leader-follower modes. The goal of human-human and human-robot teams, in any format, is to enhance the capabilities of the team in order to complete some task. The key problem with the human-robot dyad is how to ensure that this goal is met, and this goal can be split into two sub-goals: determining human intent and determining a control method for high DOF co-manipulation of extended objects. Human intent is defined in this thesis as *the consensus velocity of the object manipulated by a dyad*. Velocity

Figure 1.1: Person and a robot co-manipulating an extended object.

was chosen as it allows intent to be captured for tasks where target position is unknown to the robot. Additionally, current co-manipulation methods incorporate velocity as intent [3, 4]. Determining intent in co-manipulation tasks can be difficult, even in low DOF tasks, where uncertainty and ambiguity can exist when tasks include manipulating an extended object that may need to be translated, rotated, or both. When an extended object is included in co-manipulation tasks, forces applied in a lateral direction could indicate either intent to translate laterally, or intent to rotate the object in the plane. Similar problems with ambiguity exist in higher DOF tasks also. In order to be effective, a robot performing co-manipulation of extended objects must be able to distinguish between rotation and translation intent in multiple DOF.

Determining a control method for co-manipulation is also a difficult problem. As mentioned previously, the goal of the dyad is to enhance the capabilities of the team [5]. The control method, therefore, should be designed such that the team's capabilities are indeed enhanced. Control methods that are too complex to use, too narrow in scope, or unable to adapt to disturbances and trajectory changes, do not enhance the capabilities of the dyad, and the human partner might be better off trying to complete the task alone.

One method of solving these two problems is to learn from human behaviors [6]. Including humans in control loops with robots can be dangerous, ineffective, or both, because humans are complex and unpredictable. However, if humans exhibit trends in their behavior, or, in other words, their behavior is captured by certain repeatable patterns or actions, their inclusion in human-robot dyads is much more reasonable and safe [7]. For these reasons, this thesis explores basing human-intent estimation and co-manipulation control models on human behavior.

1.2 Problem Motivation

Extended objects are objects that are heavy or unwieldy, and necessitate two or more people to carry them. A robot capable of replacing a human in these teams would help in situations like search and rescue where current high-payload robots are too heavy and dangerous to relocate and operate. Robots that can safely interact with a human could help lift and remove rubble from disaster areas or take a victim on a stretcher to safety. These robots would allow fewer people to complete the same amount of work, or allow more teams to operate and reach more people in need of help. Other applications include using robots to help load and unload moving vans, using robots to help move objects around warehouses, and any other co-manipulation applications where human-human teams are currently needed (see Fig. 1.2).

While past work on co-manipulation shows that collaboration is possible in some tasks, it is not clear that the algorithms and intent-estimators developed will work in less-defined scenarios. In order for a robot to work effectively with humans in co-manipulation, the robot needs to be able

(a) Workers sort through rubble at disaster site [8]. (b) Wounded man is carried on a stretcher [9].

Figure 1.2: Motivational scenarios for co-manipulation of extended objects by dyads.

to respond in complex situations involving object movement in 6 dimensions, 3 translation and 3 rotation. In fact, the current state-of-the-art controllers do not meet this criteria, and are often built for specific 1 or 2 DOF motions, involving planar arm movements only. Real-life situations will require high DOF movements involving whole-body motion and bi-manual manipulation by the participants, rather than planar arm movements only.

Controllers that can be used in co-manipulation by a human-robot dyad, in 6 DOF, and based on intent provided by the human will cause scenarios like those shown in Fig. 1.2 to become much safer and more efficient for the people who work in that environment. This is especially true as more robots enter the human workspace. In the past few years, the field of soft robotics has evolved as the demand has increased for robots that can work not only intelligently and intuitively, but also safely with humans. It is not too hard to envision a soft robot, such as the one shown in Fig. 1.3, working with a human to perform all manner of daily tasks with little risk of human injury. As robots become more technologically advanced, algorithms for co-manipulation will become a necessity, rather than a novelty.

Figure 1.3: A pneumatically actuated robot with fabric-based limb and body structure.

1.3 Specific Contributions

The contributions of this thesis are:

1. Observations and analysis of metrics from 6 DOF human-human physical co-manipulation study
2. Development of Intent Estimators
   (a) Force and Torque Pattern Recognition
   (b) Recursive Neural Network Prediction
3. Closed Loop Controllers
   (a) Extended Variable Impedance Control
   (b) Neural Network Prediction Control
4. Human-robot physical co-manipulation experiment to compare developed controllers to each other and to human-human dyads
5. Observations and analysis of human-robot physical co-manipulation study

Additionally, work was done in tandem with another graduate student on the human-human physical co-manipulation study. Table 1.1 shows a delineation of work done previously and novel work done specifically for this thesis.

Table 1.1: Human-human physical co-manipulation study contributions.

Previously Done [2]	Thesis Contribution
Development of Study	Analysis of Data
Collection of Data	Metric Development
Initial Neural Network Formulation	Formulation of Force-Pattern Intent Estimation
Analysis of Neural Network Prediction	Lateral/Rotational Velocity Patterns

1.4 Thesis Overview

Chapter 2 explores previous literature and methods to solving this problem. This includes reviewing previous human-human co-manipulation studies, force-based co-manipulation methods, motion-based co-manipulation methods, performance metrics for co-manipulation, and various human-intent estimation strategies.

Chapter 3 describes the analysis of a 6 DOF human-human physical co-manipulation study on co-manipulation of an extended object. This includes a summary of the study, an investigation of various metrics, and characterization of lateral translation and planar rotation tasks. From this study, an intent estimator and a new control method of planar co-manipulation for extended objects was developed, and is formulated in Chapter 4. This control method is known as Extended Variable Impedance Control.

Another intent estimation method and controller are described in Chapter 5. This control method is called Neural Network Prediction Control. This chapter also includes the architecture of the neural network. The predictions from the neural network are applicable to 6 DOF, but are only implemented in planar motion in this thesis. In Chapter 6, both the Extended Variable Impedance Control and Neural Network Prediction Control are implemented on hardware and compared based on metrics discussed in Chapter 3. Conclusions and discussions on possible future research in this area are located in Chapter 7.

CHAPTER 2. RELATED WORKS

Researchers have been studying pHRI co-manipulation for many years. The efforts of researchers can be grouped into a few different categories: studies about co-manipulation or human behaviors, force-based and motion-based co-manipulation methods, determining performance of human-robot dyads through metrics, and human intent estimation. This thesis deals with all these categories, so they will all be explored more in depth.

2.1 Co-Manipulation Studies

There are many studies that explore the characteristics of human movement. We are especially interested in the studies that look at co-manipulation between human-human dyads or human-robot dyads, because a co-manipulation controller based on human-human interaction (HHI) is likely to be more effective and intuitive than one that is not based on HHI.

There are studies that explore human-arm reaching movement characteristics, that give insight into how humans move. One of the most widely used studies was performed by Flash and Hogan [10], which illustrates the tendency of upper-arm reaching movements to resemble minimum-jerk trajectories. The equation for determining the trajectory is given by Eq. 2.1.

$$x(t) = x_0 + (x_f - x_0)\left\{6\left(\frac{t}{t_f}\right)^5 - 15\left(\frac{t}{t_f}\right)^4 + 10\left(\frac{t}{t_f}\right)^3\right\} \tag{2.1}$$

Another fundamental study of human characteristics was performed by Rahman et al. [11]. In this study, they performed a 1 DOF translation co-manipulation experiment between two human users. One of their objectives in this experiment was to measure the impedance of the users while performing this task. The human arm can be considered a mass-spring-damper system, and thus is defined by inertia, stiffness, and damping parameters. They showed that in these tasks, one user often performed most of the work, and showed that the stiffness and damping parameters changed

over the course of the task, which is known as variable impedance. They also showed that for 1 DOF translation tasks, the achieved trajectory corresponded well with the minimum-jerk trajectory.

There were also a number of studies investigating how humans cooperate through forces and haptic channels. Reed et al. [12] performed an experiment testing their hypothesis that humans can cooperate by specializing their forces. The test involved a 1 DOF crank with handles on each side, and the task was to rotate the crank to a specified position. They found that human-human dyads were able to perform simple tasks significantly faster than they did when working alone. Ganesh et al.'s study [5] also showed that a physical connection between two individuals improves the dyad's performance. In a different study by Wel et al. [13], this result was confirmed, as subjects performing a 1 DOF task benefited from the added haptic channel communication. However, when Reed et al. included a robot, this advantage disappeared. Additionally, when a partition was placed between the user and the robot, they performed better, indicating there may be a lack of trust, either conscious or subconscious, in humans when working with robots.

One of the only studies done with a human-human dyad carrying an extended object was done by Bussy et al. [14]. In this experiment, they had dyads move a beam in 1 DOF, forward and backward. They explored three scenarios within this experiment, one with no role assignment, one with subject 1 as leader, and one with subject 2 as leader. What they found was that by looking at normalized velocities over the whole task, the trajectory can be decomposed into constant velocity phases, as well as constant acceleration phases. They used these observations to create a state machine, where state transitions are triggered by the object velocity. The states indicated whether to stay at a certain velocity or to accelerate to a different velocity.

Along with the works mentioned, there are other studies that show that not only does a haptic channel improve performance of a human-human dyad under normal conditions, but that a haptic channel can be used as the only source of information exchange between partners. Sawers et al. [15] performed an experiment where participants performed a series of dance steps with a partner. Each individual held on to an apparatus with each hand. This apparatus collected force data throughout the experiment. One individual was selected as leader, the other as follower, and both were then blindfolded and were only permitted to communicate via the haptic channel, i.e. through the apparatus. These dyads were able to complete the prescribed dance patterns while only pushing and pulling on the apparatus. Another study, done by Mojtahedi et al. [16], also

showed that interaction forces may communicate movement goals between human-human dyads in cooperative physical interactions. While Sawers et al.'s study used a leader-follower setup, other researchers have explored dominance and role-switching between partners [17–20]. These are listed here, but are not explored further, as this thesis explores only the leader-follower setup.

2.2 Control Methods for Co-Manipulation

2.2.1 Force-Based Co-Manipulation Methods

Many different control methods have been used in the last 20 years for human-robot co-manipulation. Much of the work in this area has been very successful, and is used in this thesis as a basis of knowledge for intent estimation and robot control. However, it is not clear how many of the control methods apply to large-scale, extended objects, or to movements requiring more than 1 or 2 DOF. One of the first controllers for cooperative manipulation of an object by robots and humans was an impedance controller developed by Ikeura et al. [21, 22]. They also developed strategies for situations that required using direction of force and change in magnitude of force. This type of control technique is known as variable-impedance control [23, 24]. The defining characteristic of this method is measuring Cartesian-coordinate forces at the end effector to determine motion intent in certain Cartesian directions. A desired velocity for the end effector of a robot is calculated using Eq. 2.2, based on the measured force inputs. Tsumugiwa et al. [4] showed that varying the impedance with the $\alpha\dot{\mathbf{F}}\dot{\mathbf{p}}$ term in Eq. 2.2 allows for increased performance of human-robot interaction in calligraphy. They had the human operator write the Japanese Kanji character "Dou" and found that allowing the impedance to vary allowed for a more stable interaction.

$$\mathbf{F} = m\ddot{\mathbf{p}} + c\dot{\mathbf{p}} - \alpha\dot{\mathbf{F}}\dot{\mathbf{p}} \tag{2.2}$$

This model is very successful in predicting Cartesian movements, as was shown in other studies as well, [3,25], but it does not generalize to include rotational movements. It also is heavily dependent on human force input, meaning the robot does not proactively contribute to moving the object being manipulated, and the human partner must exert more force than would be required in a human-human dyad. Since a co-manipulation controller based on HHI should be intuitive

and effective, any technique used should be easy–mentally and physically–for the human user to employ and should be versatile in application.

The initial work in variable impedance control (VIC), however, provided a basis for using haptic information in future pHRI controllers. One such controller was implemented by Ranatuga et al. [26] The objective of their work was to create an intuitive and robust controller that did not depend on a trajectory objective, much like VIC. They designed an inner-loop controller to behave like an impedance controller, with an outer-loop controller that incorporated the human dynamics and task performance requirements. With this, they were able to perform 1 DOF point-to-point motion tasks. This was important as it showed a way to provide accurate, intuitive trajectory control without previous knowledge of the trajectory, which is necessary for situations such as search and rescue. However, the work assumed direct contact between human and robot, i.e. no extended object co-manipulation, and was limited in DOF. In fact, there is an inherent problem with VIC, and other methods, such as Leica et al.'s method for moving extended objects [27], that limits how many DOF are viable. This is known as the translation versus rotation (TvR) problem. The TvR problem is described in Fig. 2.1. In a simple planar task, the leader has the option of moving the extended object by either translating forward/backward, translating laterally, or rotating the board. The problem arises when the leader wishes to move laterally, and so applies a force in that direction, as Eq. 2.2 would mandate. However, the follower, who lies some distance away from the applied force, perceives the force as a torque, and begins to rotate the board. There is information missing in VIC to deal with the TvR problem.

Two approaches to solve this problem were made by Karayiannidis et al. and Nguyen [28, 29]. Karayiannidis et al. used the direction and magnitude of the applied force to an extended object to create a state machine which switches between translation and rotation modes. The state machine, however, fails to transition between states correctly when moving at different speeds than described in their experiment. In other words, their method works for slow rotations and fast translations. If the leader wants to perform a fast rotation or a slow translation, the algorithm is not able to correctly predict this. Nguyen improved upon this by using Hidden Markov Models. His method showed that it is possible to predict human behavior in co-manipulation tasks. His algorithm allowed for different speeds of rotation and translation, but ultimately performed worse than Karayiannidis et al.'s method. Notably, the performance metric for Nguyen's experiment was

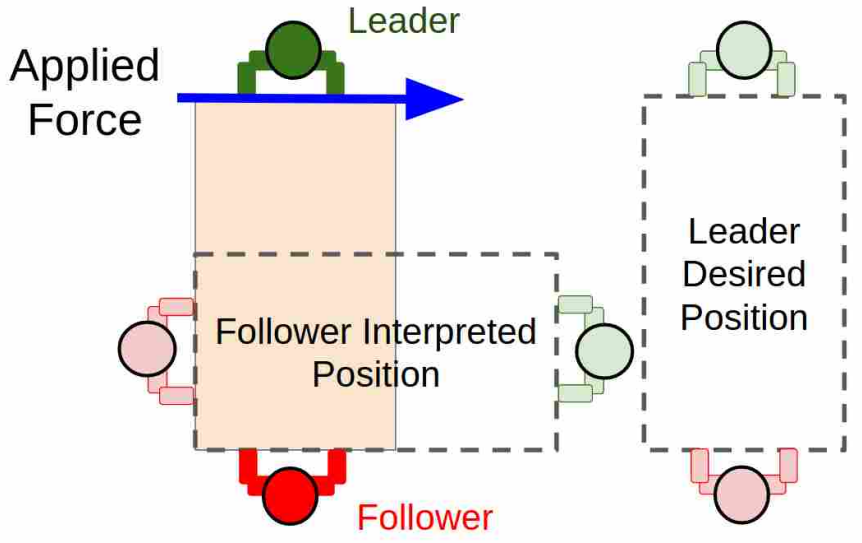


Figure 2.1: Illustration of the TvR Problem; the leader applies a lateral force, desiring to translate in that direction, while the follower may interpret this as intent to rotate, due to the distance at which the force is applied.

the mode accuracy, or whether or not the robot was translating or rotating correctly. They did not compare to any of the metrics established by other researchers. Due to the nature of Hidden Markov Models, it is also unclear what human behavior was captured that provided the information that solved the TvR problem.

One of the only attempts at bi-manual, planar co-manipulation we saw was developed by Bussy et al. [30]. Their method relies on force inputs to a trajectory-based control law, where the trajectories are then decomposed into a finite state machine to determine the desired velocities. We attempted to replicate this method, but were unsuccessful, as it appeared some necessary information was omitted from the paper. This research was successful in at least anterior translation coupled with planar rotation, and theoretically generalizes to include lateral translation. However, the paper makes no mention of attempts to move in lateral translation, and a video of the controller shows only anterior translation with planar rotation. Indeed, it is not clear how this paper deals with the TvR problem.

Other work has been done by Peternel et al. [31] fusing force feedback with EMG sensors and optical markers attached to the human's arm. The myoelectric activity recorded by the EMG sensors is included in the control law to provide more information about the stiffness the human was applying in a 1 DOF sawing task. Using EMG in this way was successful in enhancing

the adaptability of the controller, and although we did not consider using EMG sensor data for this thesis, it should be considered in future work. Additionally, Peternel et al., in a different work [32], showed how robots can adapt to human fatigue in pHRI. They showed that robots are capable of adapting parameters to maintain appropriate force levels in co-manipulation, although this work did not explore an extended object. While human fatigue is not a concern in this thesis, it should be considered in any future work involving prolonged pHRI.

2.2.2 Motion-Based Co-Manipulation Methods

In addition to force-based methods, many insights into human-robot interaction have been gained from studying motion-based intent. One of the common methods of motion-based co-manipulation is using a minimum-jerk basis. Corteville et al. [33], did so for a 1 DOF point-to-point experiment. Also, Maeda et al. showed a way to use minimum-jerk trajectories to predict human intent for proactive robot behavior. Their controller estimates the final time and position that the human is attempting to reach based on the past few steps and uses position control to follow that trajectory along with impedance control to adapt to errors [34]. This strategy reduced the amount of effort a human partner needs to exert in co-manipulation tasks, which is one of the problems with variable impedance control.

However, there are certain restrictions on using minimum-jerk as a controller basis. The trajectory start and end points must be known beforehand, minimum-jerk based methods are not robust to disturbances, and they do not take into consideration the haptic channels, which have been shown to increase performance. Additionally, Thobbi et al. [35] showed that there are some human movements that are not minimum-jerk movements. They approached solving the TvR problem using a Q-learning algorithm, and creating a reactive controller to observe human actions and predict the future position of the human. Rather than assume the human (when leading) was always following a minimum-jerk trajectory, the learned model allowed the robot to account for a larger variation in trajectory. However, they do not consider higher DOF, nor do they incorporate haptic inputs. Miossec and Kheddar [36] also explored non-minimum jerk-based trajectories. They continued the work done by Bussy et al. [14], where dyad motions are longer and include walking and not just arm movement. Their experiment indicated that the interaction forces are too difficult to use as a basis for control, and instead use velocity thresholds to change state in a finite state

machine. Their state machine controller, with constant velocity and constant acceleration phases, was able to perform a transportation task of an extended object [30]. However, because their experiment only looked at 1 DOF, it is not clear how their method would extend to higher DOF.

Ge et al. [37] showed that machine learning can be a useful tool in HRI. Their research used supervised learning to predict the motion of the human limb in order to learn a desired trajectory for pHRI without an extended object. They found their method performed better than using an impedance controller. While their work, along with that shown by Thobbi et al. [35], shows that human performance can be learned and applied to pHRI controllers, they did not account for co-manipulation of an extended object. Additionally, they indicated that the interaction force between a human-robot dyad should be minimized, which may not be an ideal objective for some co-manipulation tasks. Another use of machine learning was demonstrated by Berger et al. [38]. They used machine learning to take in accelerometer and pressure sensor information to learn a statistical model to guide the robot's behavior. This study was unique in that they did not use force sensors at the wrist, or at any interaction point between the human and the robot, but instead calculated the center of mass of the robot, as well as used pressure sensors on the feet of the robot. The implementation of this control method involved humans applying forces either forward/backward or laterally to a shared object, to indicate the desired direction of movement. However, they did not explore the TvR problem, and it is not clear how well this method performs in comparison to human-human dyads.

A different approach, using shared-leader control, was taken by Medina et al., who demonstrated an anticipatory model that took into account expected human trajectory as well as human force variability [39]. These sources of uncertainty were then incorporated into a cost function along with an impedance model of the robot. They showed that incorporating uncertainty led to a higher perceived helpfulness in the robot partner of a human-robot dyad. Their model did not use minimum-jerk assumptions, and was still able to be effective in a trajectory-based situation. However, the cost function used in this research incorporated a term for human force minimization, which may not necessarily be a goal in co-manipulation. It also was implemented in 2 DOF with a single-arm virtual haptic interface, so it is unclear whether it would extend to bi-manual co-manipulation with an extended object in higher DOF. Because this is a shared-leader approach,

it was not considered for comparison in this thesis, but should be considered for comparison in future work involving shared-leader methods.

2.3 Performance Metrics

An issue in co-manipulation studies and methods is determining what constitutes a successful dyad. For example, in Sawers et al.'s study, two dyads may complete the dance steps, but successful completion probably is not the only important result from the study. One dyad might take longer than the other, or a dyad might also have more variability in motion than another dyad. Therefore, there needs to be performance metrics that allow for comparison between dyads.

Haptic information has been shown to be a viable communication method, and some researchers have suggested this information is used by dyads to minimize certain criteria. Groten [7] described a number of these, including minimizing interaction forces and root-mean-square error, and maximizing time on target. Interaction forces are described using Fig. 2.2. These forces are the forces that are applied to the object that are absorbed by the users, and do not directly contribute to the motion of the object. In other words, interaction forces are the forces applied by one person that are countered by their partner, and therefore cause no acceleration of the object. Root-mean-square error is a measure of how close the object is staying to a desired reference trajectory. Time on target involves determining how much time a robot spends in correct behavior during a task. A reference trajectory that is commonly used, such as in Corteville et al. [33] and other previously mentioned studies, is the minimum-jerk trajectory. However, there are also tasks that do not fit well with the minimum-jerk trajectories [35, 36]. Therefore, some alternative trajectories may need to be used if using a root-mean-square error on trajectory.

Ivaldi et al. [6] also described a few other metrics, such as minimizing jerk, torque change, geodesic trajectories, energy, and effort. These are all fairly well explained by their titles, and the objective of minimizing these metrics is to achieve human-like behavior. More metrics not mentioned by Ivaldi et al., but commonly used in other works are minimizing task completion time [3, 36] and position error in trajectory following tasks such as tracing a path through a maze [23, 35]. In reviewing the literature, it is not clear which of these metrics is most important. For example, completion time can capture how quickly a dyad is able to complete a task, but this may not be the most important measure of the dyad. Perhaps their task involves moving as close to the

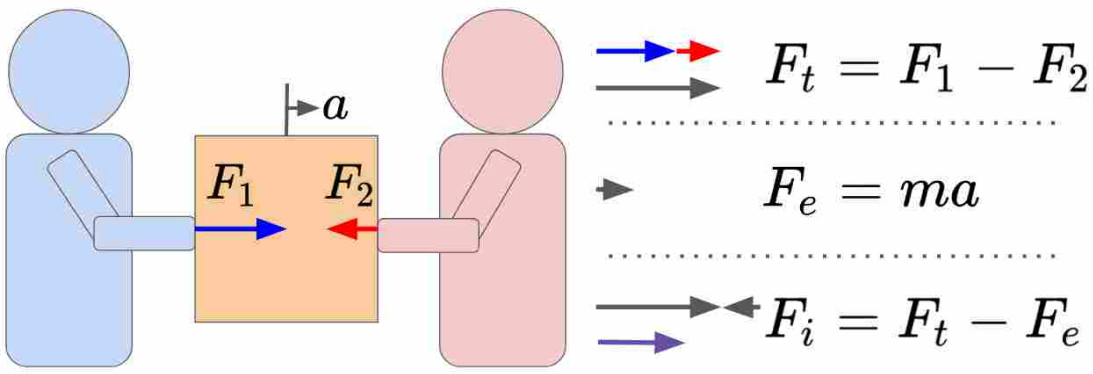


Figure 2.2: Two users co-manipulate an object. The applied forces can be interpreted as a total force, F_t (measured by a force sensor), an external force, F_e, or the difference between these two F_i, where F_i is the interaction force between the users.

desired trajectory as possible, and moving faster causes more errors. This is also not an exhaustive list of metrics, and this thesis will explore some other possible metrics, in addition to using this list as a basis for characterizing performance of co-manipulation methods. Determining what behavior dyads display is essential not only for comparing one dyad to another, but also for comparing one co-manipulation controller to another.

2.4 Human Intent Estimation

As we explored the previous works, we noticed that one of the main hurdles researchers still need to clear is to determine human intent. Many papers have suggested that haptic channels are an appropriate method of communication for human intent [12, 40–42]. This makes sense, as we have seen that human teams can move objects by interacting only through forces applied to the objects, rather than by communicating verbally or otherwise [15, 16]. Many studies have been done to conclude that robots can be controlled by human force input in this manner, but these studies often involve the human acting directly on the robot, and not through any extended object [4, 22, 33, 43].

Some research involved shared virtual-environment loads [44, 45], and others involved upper-arm movements of individuals and dyads [12, 13]. These experiments clarify many aspects of pHRI, including verifying that haptic information aids in co-manipulation tasks, noting some interaction patterns, and combining planning and learning to complete goal-oriented tasks. Ex-

ploring virtual environments and upper-arm movements, however, neglects any effect caused by the users walking, which could cause unforeseen disturbances and vibrations. Using only upper arm-movements also limits the applicability of these control methods.

We also saw that human intent can be captured by the motion of the manipulated object, [14,35], or by the motion of the subject [37]. These methods often produce good results, but ignore the benefits gained by providing haptic feedback to the robot partner. Another method of intent estimation that has been used is programming by demonstration, as in Rozo et al. [46]. Here, intent is compressed into a section of possible motions the human-robot dyad could take. The disadvantage here is that it is not robust to disturbances, or trajectories that have not been modelled previously. Our definition of intent is appropriate for co-manipulation of extended objects because it allows us to capture intent for motion that does not have definite start of end points (as observed by the robot), or motion that involves unforeseen obstacles.

2.5 Summary

The works cited here describe most of the current research being performed in pHRI co-manipulation. As has been shown, there are very few studies that look at co-manipulation of extended objects, and even fewer that look at high DOF co-manipulation. Control methods are varied between force-based and motion-based, but almost all are limited in applicability due to low DOF, or other limitations. We also have not seen a working co-manipulation controller for a human-robot dyad, with at least 3 DOF that can be used in undefined situations or respond to disturbances, in any of the related literature.

CHAPTER 3. ANALYSIS OF 6 DOF HUMAN USER STUDY

3.1 Summary of Exploratory Study

As described in more detail in [1] and [2], we had 21 human-human dyads perform a series of co-manipulation tasks carrying a 10.3 kilogram table. Figure 3.1 shows one task being performed and video of the most complex task can be seen at `https://youtu.be/DAbLRDN20yE`. There were 6 different tasks that were each performed 6 times, 3 times blindfolded and 3 times non-blindfolded. The purpose of this study was to understand how people collaborate moving an object with no constrained degrees of freedom, and to expand on the knowledge gained from Chapter 2. We also expected to identify new principles from human-human data that could be used for control development. The tasks chosen were based on this idea, and represent several different types of common motion. This thesis focuses mainly on two tasks, which are lateral translation and planar rotation, i.e. 3 DOF planar motion.

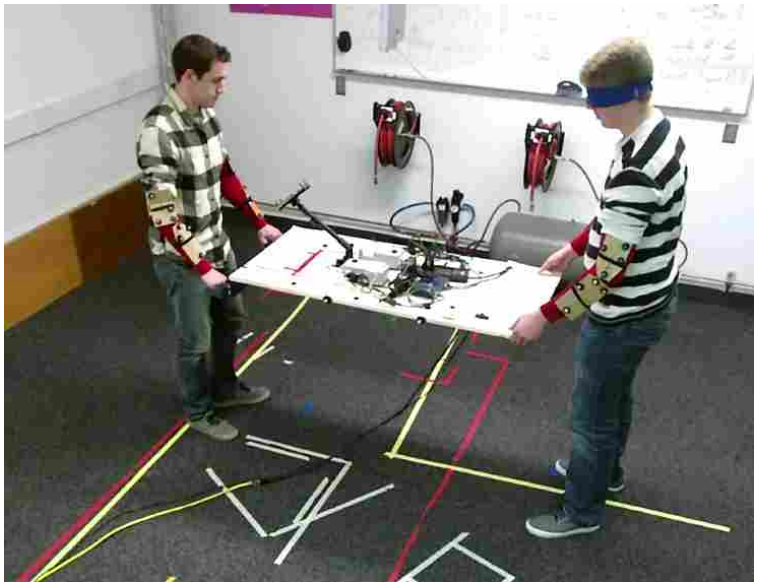


Figure 3.1: A human-human dyad, leader on left with follower blindfolded, performs a co-manipulation trial while holding an extended object.

During the experiment, each partner was assigned a role as leader or follower. The leader was then given instructions on where to move the table. In half of the tasks, the follower was blindfolded and talking was prohibited. This can be thought of as a baseline for any robot controller that only uses haptic information for control. In the other half of the trials, talking was permitted and there was no blindfold. The performance in these trials could be considered the long-term objective for performance and something against which we can benchmark our future human-robot controllers. We recorded the motion of the table during all the tasks, as well as the forces and torques being applied to the table. As discussed in Chapter 2, intent, defined in Chapter 1 as the consensus velocity of the manipulated object, can be captured by force and motion. Fig. 3.2 shows the equipment used for the experiment. The wooden table, measuring 59x122x2 cm, contained all the sensory equipment, including: two ATI Mini45 force/torque sensors fastened to the blue 3D-printed handles, two ATI NET F/T Net Boxes (which passed data over Ethernet to the computer at a rate of 100 Hz), power cables, an ethernet switch, a tablet mount with a tablet, and infrared markers for use with Cortex Motion Capture software with a Motion Analysis Kestrel Digital Realtime System. This motion capture system captured data at a rate of 200 Hz.

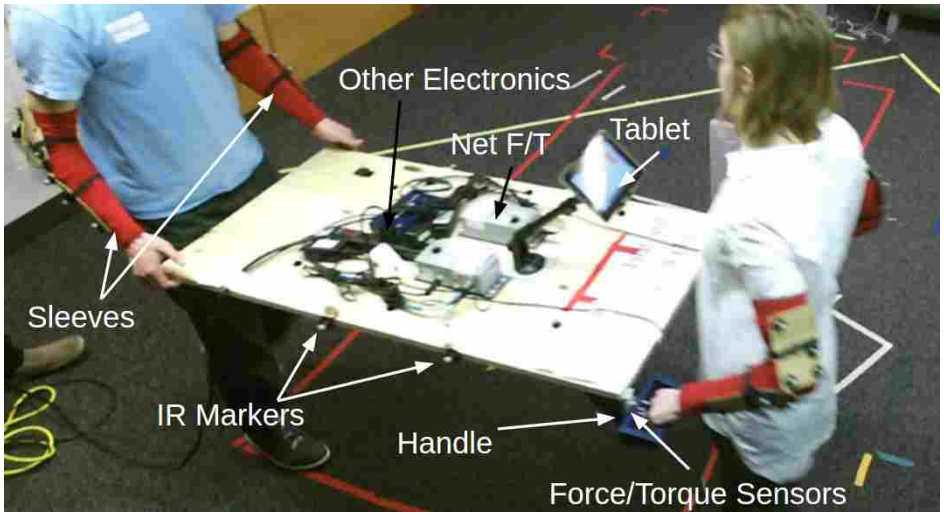


Figure 3.2: Setup for table and during trials.

Subjects were randomly assigned leader and follower positions, which they stayed in for the duration of the experiment. Task order was randomized, using randomized partial counterbalancing, as was whether or not the follower used their blindfold. Instructions were sent to the leader

via the tablet, so the follower was never made aware of which task they were about to perform. Each participant also wore sleeves with infrared markers to capture their arm movement throughout the task, and a Microsoft Kinect 2 was used to record video and point-cloud data for each trial.

After each dyad completed their trials, they were given a survey and asked to rate the performance of their partner on a variety of criteria, such as whether the partner moved at appropriate speeds, moved in the correct direction, whether they trusted their partner, etc. These questions help qualitatively measure performance that we might otherwise miss, and helps provide a baseline for developed robot control methods. Additionally, people may trust and prefer people more than robots, and this data helps determine if there is a preference gap that cannot be captured by objective performance measures. Well designed human-robot co-manipulation controllers should approach both quantitative and qualitative performance of human-human dyads.

Although the experiment involved 6 different tasks, and our future goals include incorporating controllers for 6 DOF translational and rotational movements, this chapter focuses on characterizing a subset of the tasks. We focus mainly on the blind versions of both the pure translation task, as well the pure rotation task. The emphasis was placed on these tasks for a few reasons. First, as discussed in Chapter 2, most research done in this area of pHRI for co-manipulation involved either lateral movement with no extended object, or only anterior direction movements (see Fig. 3.3 for directions reference).

When co-manipulating an extended object, the intent of the leader is complicated by the TvR problem. Therefore, characterizing how humans are able to recognize a desired lateral movement with an extended object and distinguish it from a desired rotational movement is key for successful co-manipulation of extended objects. Second, other tasks include components of translation and rotation. Therefore knowing the defining characteristics of lateral translation and planar rotation helps to recognize these motions in more complex tasks. We also examine only the tasks where users were blindfolded to simplify the analysis, since these tasks involved only haptic communication. The sighted tasks will be used as upper bounds on possible performance, or as a basis for controllers with more input–such as robotic vision–in future work. Last, because there are few co-manipulation methods for low DOF control, developing a planar 3 DOF approach provides a strong foundation to develop higher DOF methods.

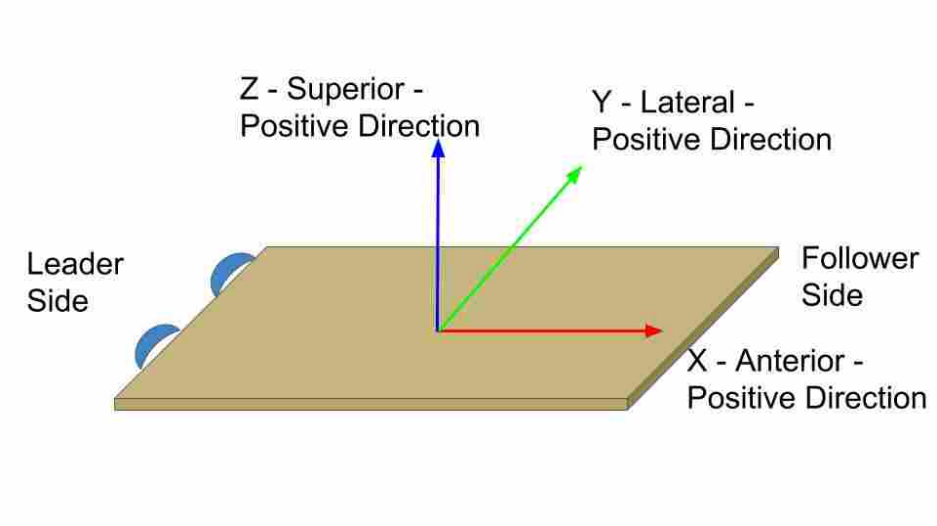


Figure 3.3: Anatomical direction reference with corresponding table axes: x is anterior, y is lateral, and z is superior.

3.2 Metric Investigation

The two-fold objective of the experimental study was to form a foundation for a robot controller that could potentially replace a human in the human-human dyad, and also to determine the performance of human-human dyads in co-manipulation tasks in 6 DOF. In order to be a suitable human replacement in a dyad, a robot controller needs to approach human performance in co-manipulation tasks. Therefore, we needed to identify and define which metrics characterize human-human performance the best.

As discussed previously in Chapter 2, there are difficulties in determining performance metrics from our experiment. First, there are a variety of metrics to choose from, and not all of them are relevant to all types of tasks. Metrics that maybe be a good fit for translation task may not necessarily be a good fit for the complex 3D task. Second, the dyads were given no specific directions other than for the leader to complete the on-screen tasks, and the follower to follow the leader. Therefore, we cannot say a specific dyad was trying to complete the task quickly, or precisely, rather, we can only say that the dyads performed the task in whatever way they preferred. Third, behavior metrics, like time on target from Groten [7], are hard to determine for complex tasks. For a 1 DOF translation task, it might be relatively easy, we can determine when it is translating by looking at times that motion is restricted to a single DOF. However, for a task like

the pick and place task, which is described in [1] as Task 1, there are many different methods to reach the destination, and it is difficult to determine what the *correct* behavior is. This chapter explores possible metrics and determines which are most suitable for creation and evaluation of pHRI co-manipulation controllers.

3.2.1 Variable Relationships

Because our experiment involved 6 DOF, it is often difficult to decouple how a certain force or action affects the environment over time. Instead of looking at the states strictly as time-dependent series, we also wanted to explore how the states are related to each other. This can be done by using phase portraits. Phase portraits are often used to explore things like limit cycles in dynamical systems. We used them to look at non-linearities and explore the relationship between states in a more intuitive manner.

There are a total of 18 variables considered in this dynamical system: 12 motion states (position, velocity, orientation, and angular velocity in x, y, and z) and 6 inputs (force and torque in x, y, and z). This means there are 171 unique variable relationships that could be drawn to compare these variables. Because we had to visually inspect all of these relationships, it was beneficial to eliminate some of the possible combinations.

Some of the variable relationships can be classified as phase portraits, since they are in the state space of the system. A few of these are less interesting to us, for instance comparing p_x, p_y, and p_z might reveal some interesting paths that the board takes, but don't provide much other information. Therefore, we did not include those states in our analysis. Since we are focusing on the planar co-manipulation tasks in this chapter, it makes sense to consider only the velocity states that are in the plane, which allows us to narrow the focus to v_x, v_y, and ω_z states. We did not include comparing the states to themselves, and also omitted v_x compared to v_y for readability of the plot.

Other interesting combinations, although not phase portraits, are the input/output relationships, which may give information about how the board is reacting to the inputs. Due to the nature of the planar co-manipulation tasks this thesis analyzes, we chose to focus on looking at only the effects of inputs F_x, F_y, and τ_z on the velocity states, reducing the number of variable relationships to 11.

We looked at the rotation and translation task phase portraits for these 11 combinations. A representative set can be seen in Fig. 3.4, which was generated for the rotation task. An interesting plot here is the comparison of ω_z and τ_z. For this particular plot, it appears there is a linear relationship during the middle of the task, where a decrease in torque corresponds to an increase in angular velocity. Similarly, in the comparison of v_y and τ_z, a decrease in torque leads to a decrease in velocity.

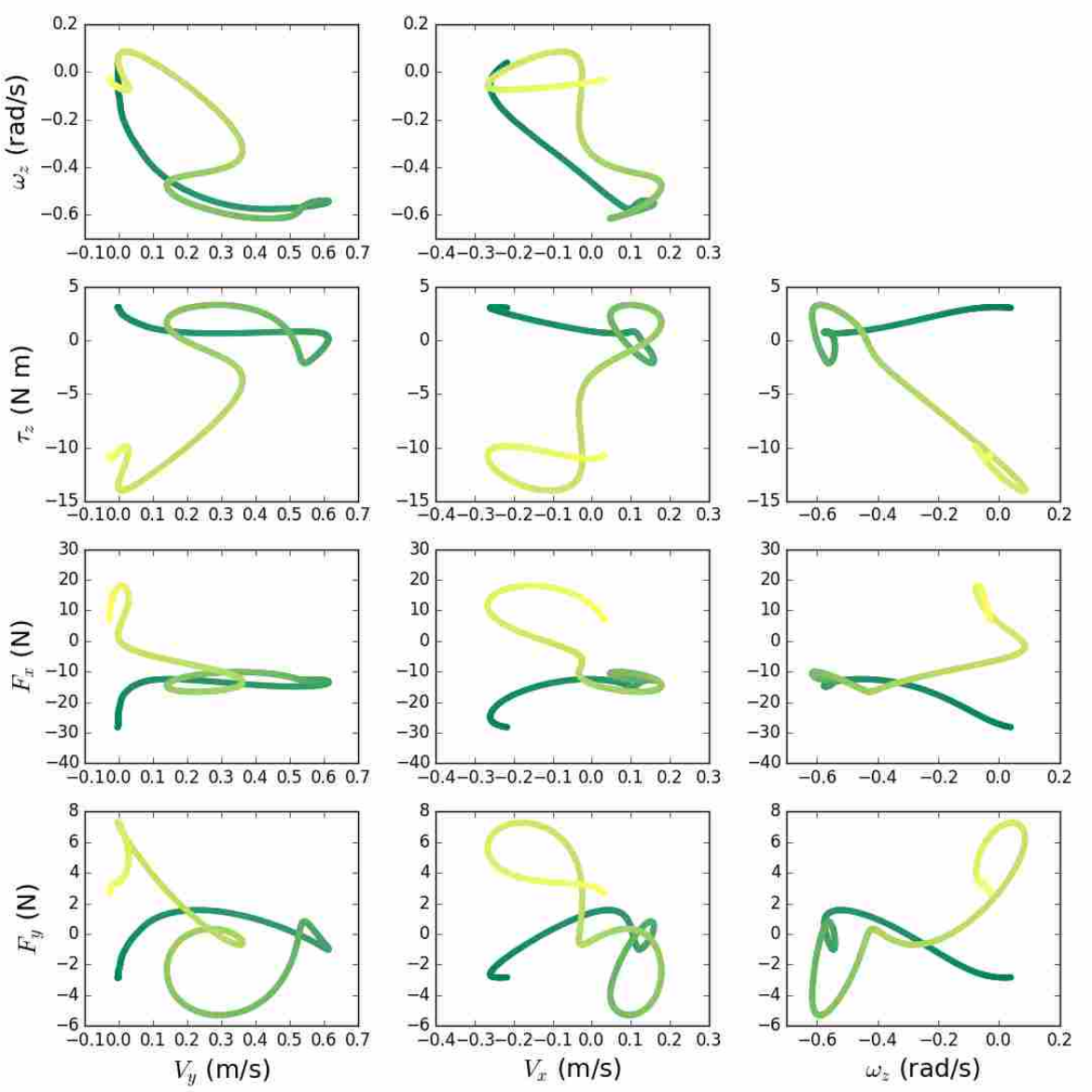


Figure 3.4: Variable relationships for a single trial of rotational task using v_x, v_y, ω_z, F_x, F_y, and τ_z as variables for comparison, with color shift indicating time dimension–task start is green and task end is yellow.

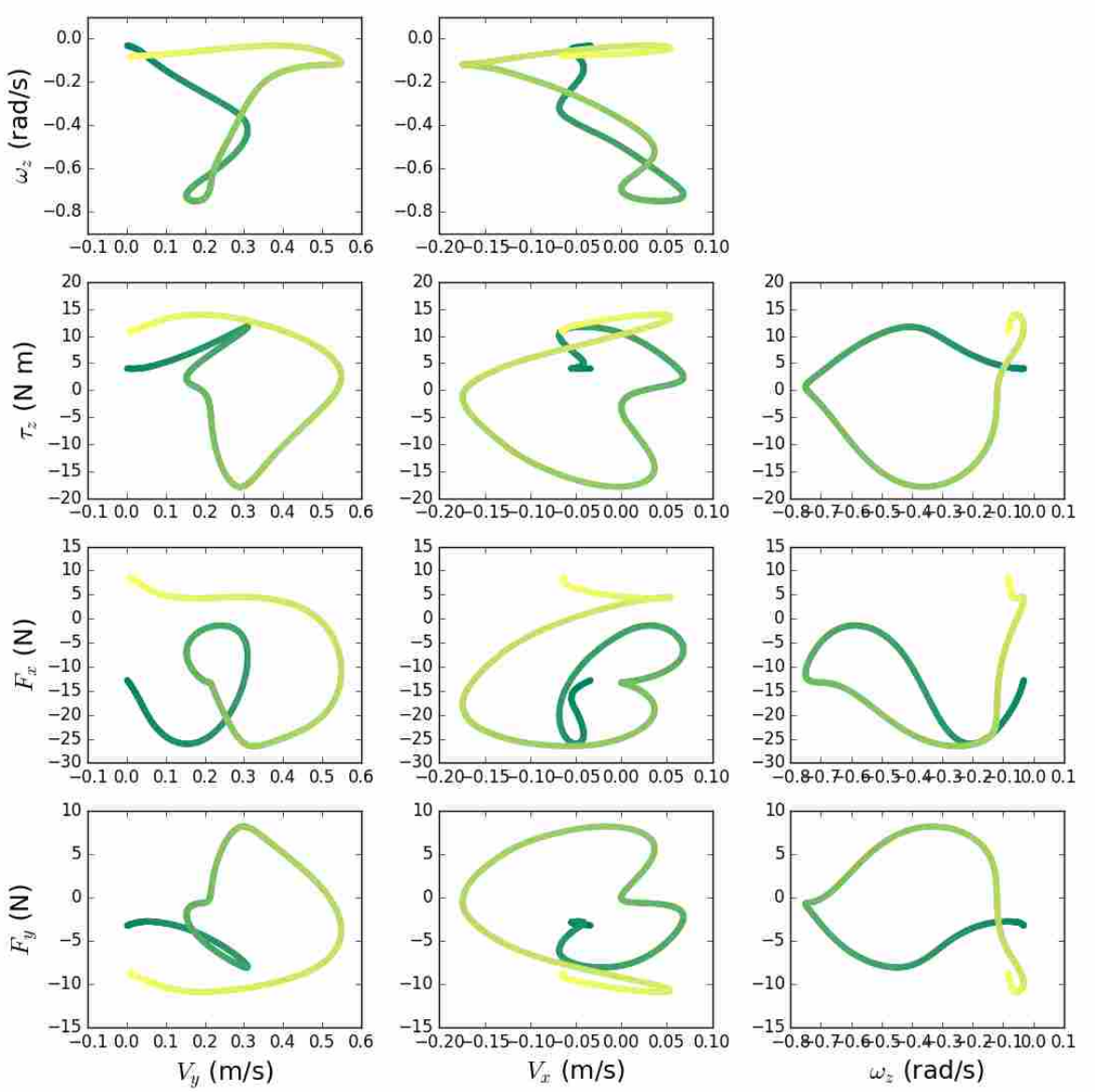


Figure 3.5: Variable relationships for a different trial of rotational task than Fig. 3.4.

These patterns, however, do not hold for all dyads, and sometimes do not even hold for other trials by the same dyad. A different trial, performed by the same dyad as Fig. 3.4, is shown in Fig. 3.5. Here, we can see that those linear events are replaced by circular ones. In fact, there did not appear to be a cohesive pattern that we could draw from these phase portraits. The phase portraits indicate that each dyad, and possibly each individual trial, can be performed in a different manner. Therefore, it is necessary to find metrics that are applicable across all dyads. There are

more efficient ways to explore and extract salient information from other variable relationships, and to determine other useful trends in the phase space, but this is outside the scope of this thesis.

3.2.2 Interaction Forces

Interaction forces are the forces that do not directly relate to motion, i.e. the forces applied by each participant that do not accelerate the object. Noohi et al. [41] suggested that interaction forces could be used as a source of communication. Other researchers have suggested (see Chapter 2) that the interaction forces should be minimized, as they represent wasted energy, which can sometimes be an objective in human movement. We sought to explore this metric further as it seemed to be a promising metric.

In our study, the force/torque sensors could not discern between external forces – forces that accelerate the object – and interaction forces, but rather measured the total force applied, so we calculated the interaction force after the experiment ended. An illustration showing how these forces are defined is shown in Fig. 3.6. Both users apply forces to the table, and the sensor measures the total force F_t. F_e, the external force, can be calculated using our knowledge of the table's motion and Newton's Laws, and then we can calculate the interaction force, F_i. Eq. 3.1 shows the combined forces that contribute to the total force, or the force measured by the sensors, and how we calculated the external force. No other external forces were modelled, such as air friction or damping, as we assumed they were negligible. The motion capture data described the pose of the table over time, and was differentiated twice to acquire the acceleration data. With a known mass of the table and acceleration, the external force was estimated (Eq. 3.1), and removed from the total force to give us the interaction force for the task.

$$\begin{aligned} F_t &= F_i + F_e \\ F_e &= ma \end{aligned} \tag{3.1}$$

For the anterior, x, and lateral, y, directions, the only external force being applied is the force applied from the participants, whereas in the vertical z direction, gravity was also applied. For all calculations and analysis in this paper, the forces were low-pass filtered with 20 Hz cutoff

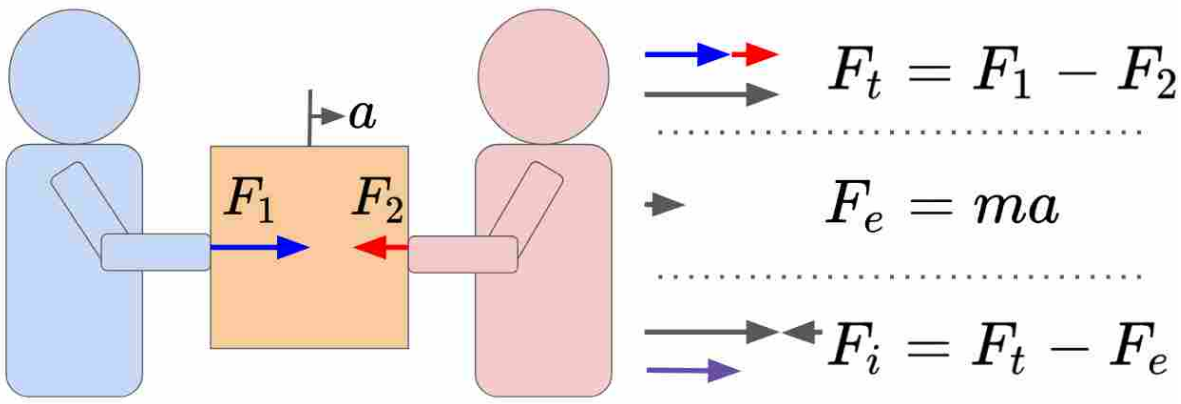


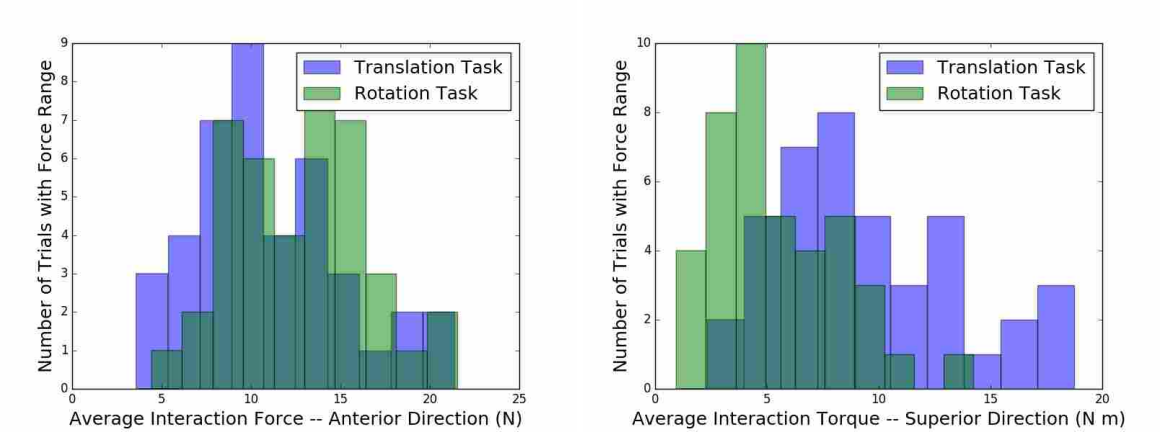
Figure 3.6: Fig. 2.2 reproduced here for reader's convenience.

frequency to represent human response ranges. The muscle response of humans can reach up to 100 Hz for brief, forceful efforts, but often lies within the 10-30 Hz range [47].

Minimizing interaction forces has been used in other works as a basis for control, including variable-impedance control [3, 23]. Our study, however, showed that human-human dyads do not necessarily minimize interaction forces. For both lateral and rotational movements, we calculated the average interaction force in the anterior direction and the average interaction torque about the superior direction. Histograms showing the distribution of average force and torque over the duration of trial, for all translation and rotation trials, are shown in Figs. 3.7a and 3.7b. These plots show a histogram of average interaction force and torque at the center of the table. As can be seen, the interaction force was almost always non-zero for both lateral and rotational movements in the anterior direction (see Fig. 3.3 for clarity on directions). Additionally, we considered the ratio of average interaction to external forces, $F_{i,avg}/F_{e,avg}$, which–when averaged over all the trials–gave a magnitude of 20, indicating the forces used for acceleration of the object were 20 times smaller than those not used for acceleration. It is not clear why the average interaction force was so substantial, but our hypotheses include:

1. These forces were used for object and human stability
2. These forces were used to better communicate intent

We will conduct future studies to explore the hypothesis on stability, but the hypothesis on communicating intent is discussed to some extent in Chapters 3.3 and 4. This result is significant because it implies that lateral collaborative movements may rely on interaction forces



(a) Force along lateral axis.

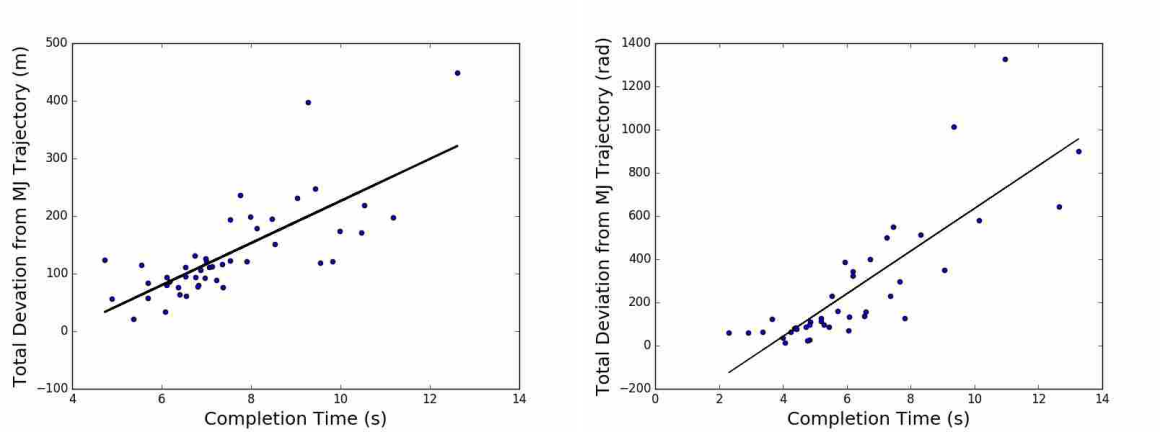
(b) Torque about superior axis.

Figure 3.7: Histograms showing average force or torque (at center of table) for rotation and translation tasks.

along the anterior direction, which is not seen in many state-of-the-art pHRI controllers. Additionally, minimizing interaction forces may not yield results easily understood by human partners in co-manipulation tasks, since it is now evident that humans are not necessarily minimizing these forces. We also notice from Fig. 3.7b that the average torque about the superior axis is generally greater in magnitude for translation tasks than it is for rotation tasks, which is an indicator of how human-human dyads solve the TvR problem. This is also discussed further in Chapters 3.3 and 4.

3.2.3 Minimum-Jerk Exploration

The minimum-jerk (MJ) movement is well-documented as a basis for human arm movements, especially in point-to-point movements, and is used in some cases as a useful metric for human performance. We did not expect to see MJ trajectories in these trials, since one participant was blindfolded and unaware of the task specifications, and the dyads used whole-body motion rather than arm-only motion. Another interesting finding from our study was that both the lateral movement tasks and the planar rotation tasks resembled a MJ movement in lateral position and angular position respectively. This was especially true for the dyads that completed the task more quickly. Figs. 3.8a and 3.8b show a positive correlation between deviation from MJ trajectories and time to complete the task. Deviation from MJ trajectories, or MJ Error, is simply the sum of the residuals between the actual trajectory and ideal MJ trajectory. The slower dyads often had a



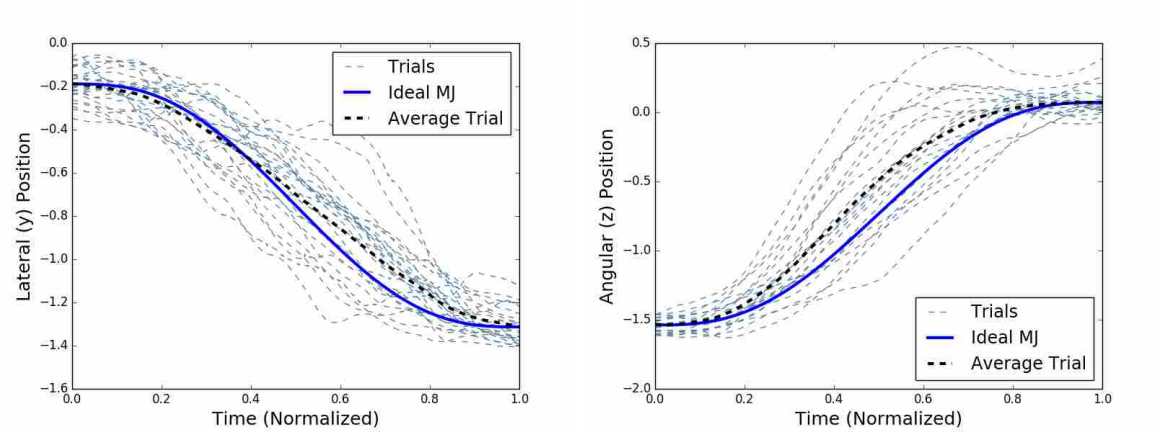
(a) Lateral task. (b) Rotation task.

Figure 3.8: Comparison of completion time to deviation from MJ trajectory with trend line.

larger error between their position and the ideal MJ position, whereas the quicker dyads generally had a smaller error with respect to a MJ trajectory. This could be due to the follower becoming confused, and taking more time to determine the intent of the leader. Therefore, completion time or deviation from a MJ trajectory could be an indication of how much confusion was present in the dyad.

Overall, though, the lateral position and angular position stayed close to the MJ trajectory, and adhering to a similar trajectory over all trials corroborates the results of similar 1-dimensional studies [14]. Figs. 3.9a and 3.9b show the position of the tasks over a normalized time, since each trial took a different amount of time. The gray dotted lines show each individual task, the black dotted line is the average position of all the tasks, and the blue line is the ideal MJ trajectory given an average start and stop position. As we can see, even though the follower did not know the end position, they managed to remain fairly close to the MJ trajectory both for translation and rotation tasks.

Despite the evidence presented for planar co-manipulation tasks, it is not clear that MJ trajectories encompass general co-manipulation for 6D tasks. Dyads may have confusion about intent, may encounter obstacles, or may move for indefinite amounts of time. These situations can lead to non-MJ trajectories, as we have seen in our study (see Chapter 3.3), and we agree with previous research, that a MJ basis for control may be too restrictive [35]. However, we also



(a) Lateral task trajectories.

(b) Rotation task trajectories.

Figure 3.9: Individual trial trajectories, ideal MJ trajectory, and average trial trajectory, all scaled to normalized time.

conclude that MJ trajectories can be useful for describing task metrics (discussed more in Chapter 3.2.4), as was proposed in other work [6].

3.2.4 Metric Selection

From the analysis of dyads in Chapter 3.2 we narrowed down the list of useful metrics to incorporate the behaviors and trends we found among the dyads. Some of these metrics may be of general use, but they were selected for their application to planar scenarios. We believe these trends could be used to indicate default or nominal human performance. We also desired to examine metrics that have been shown in other works to be reasonable metrics for describing human behavior in co-manipulation tasks. We explored these metrics as averages over all dyads, and our reasoning for this is to quantify what people do on average. Each dyad performed slightly differently, and analyzing average human behavior will help to make a controller generalize to multiple users. Although useful in making a robot intuitive, those nominal human-human metrics may actually contradict other effective performance metrics for improving things like energy expended or completion time.

With this information in mind, we analyzed our data for metrics to both classify the nominal performance of the dyads and provide a basis for control methods. The metrics considered were:

- Deviation from MJ trajectory of table
- Task completion time
- Average velocity of table
- Average angular velocity of table
- Torque change

We evaluated more than just these metrics, but chose a representative set for our analysis here. Some metrics were chosen for their intuitive relationships, such as completion time and average velocity, and others were chosen for correlation values. The chosen set was compared using the Pearson correlation coefficient, and the results are summarized in Table 3.1 for the translation task and in Table 3.2 for the rotation task. The metrics used in these tables are MJ error, defined in Chapter 3.2.3, completion time, average lateral velocity, average angular velocity, and torque change. Average velocity and average angular velocity were included as we anticipated high correlations between these values and completion time, and also wanted to see whether there was any unforeseen relationships between these metrics and others. For instance, we expected there might be a high correlation between average angular velocity and minimum-jerk error for the translation task. Velocity metrics, however, are somewhat limited, as they do not incorporate the time-dependent nature of velocity, and may therefore occlude some information since we are taking their average value. Torque change was defined in [6], and is calculated using Eq. 3.2. In this equation, $\dot{\tau}_t$ is the time derivative of torque at time t and $\dot{\tau}_{t+1}$ is the time derivative of torque at time $t+1$. This measure, which is a bit of a misnomer, determines how much the time-derivative of torque changes over the course of the task, with higher values indicating more fluctuation in torque.

$$\Delta\dot{\tau} = \sum_{t=0}^{T-1} {\dot{\tau}_t}^2 + \dot{\tau}_{t+1}^2 \tag{3.2}$$

We expected there to be some correlation between most of these measurements since most of the metrics were proposed in previous research [6, 7, 35, 36]. Surprisingly, some intuitively related metrics offered very little correlation. The most relevant expected metrics–for the lateral

Table 3.1: Correlation between metrics for translation task.

	MJ Error	t_c	$v_{y,avg}$	$\omega_{z,avg}$	$\Delta\dot{\tau}_z$
MJ Error	-	0.63	-0.42	0.05	-0.02
t_c	0.63	-	-0.66	-0.40	0.02
$v_{y,avg}$	-0.42	-0.66	-	0.30	-0.07
$\omega_{z,avg}$	0.05	-0.40	0.30	-	0.04
$\Delta\dot{\tau}_z$	-0.02	0.02	-0.07	0.04	-

Table 3.2: Correlation between metrics for rotation task.

	MJ Error	t_c	$v_{y,avg}$	$\omega_{z,avg}$	$\Delta\dot{\tau}_z$
MJ Error	-	0.83	-0.50	-0.56	0.71
t_c	0.83	-	-0.51	-0.78	0.52
$v_{y,avg}$	-0.50	-0.51	-	0.81	-0.36
$\omega_{z,avg}$	-0.56	-0.78	0.81	-	-0.42
$\Delta\dot{\tau}_z$	0.71	0.52	-0.36	-0.42	-

translation task–were average/max angular velocity and deviation from the MJ trajectory. We expected dyads performing these tasks to minimize the average angular velocity about the z axis, and stay relatively close to the MJ trajectory, but the correlation coefficient between these two metrics was 0.05. In fact, there was a much stronger correlation between deviation from MJ trajectory and completion time and average lateral velocity – being 0.63 and -0.42 respectively. Intuitively, minimizing angular velocity would be an ideal metric for this task, since a perfect lateral translation would involve no angular velocity at all. This supports our hypothesis that some metrics may be ideal from an effective point of view, but may not be nominal behavior for human users in a dyad, and we need to consider these non-intuitive relationships between metrics if we want pHRI controllers to be intuitive to humans.

Another noteworthy observation is the lack of correlation between total torque change and the other metrics. We would expect that torque change would indicate some measure of perfor-

mance, but it appears to be more related, at least in the translation case, to some other aspect of the movement. The lack of correlation may be related to our findings from Chapter 3.2.2, where there were interaction forces in the anterior direction for lateral movements, and may indicate that the non-correlated metrics may affect stability, communication, or something related to preference.

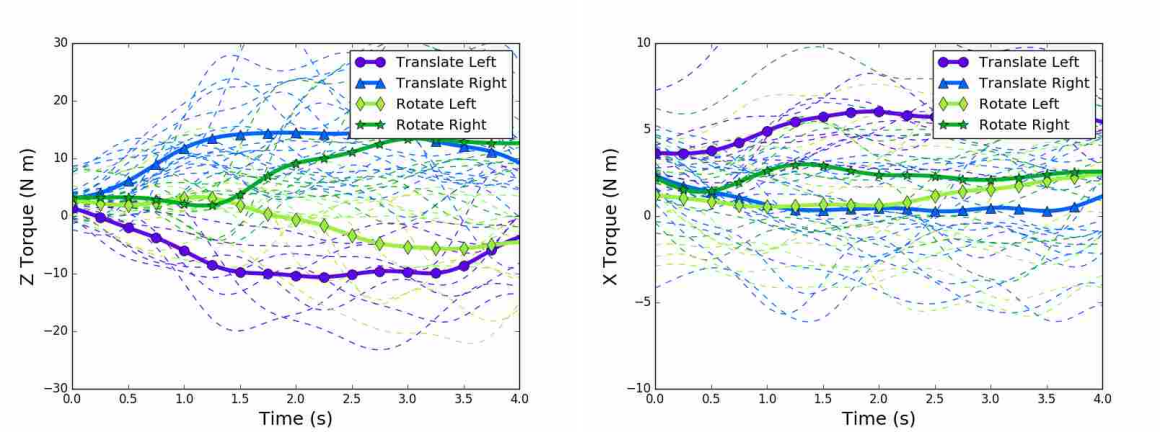
For rotational tasks, the same metric set was applied as can be seen in Table 3.2. However, we were again surprised by the results. We expected task completion time and MJ error to be related, since we had seen that previously. What we did not expect was torque change to be so heavily correlated to the other metrics, where it was not in the translation case. This finding was significant to us, and indicated that there is some fundamental difference between τ_z in these two tasks. We explore this phenomenon further in Chapter 3.3.

The surprising correlations between metrics show we do not fully understand how human-human dyads operate. A reasonable conclusion, however, is that humans may sacrifice performance on certain metrics, such as minimizing interaction forces, in order to improve the performance in another metric, or to affect the nominal performance of the dyad. While our metric observations provide some insight to this topic, our work on metrics is still an open question that is necessary to explore in order to better characterize and design performance of human-robot co-manipulation controllers. However, the remainder of this thesis mainly uses the metrics described in this chapter as a basis for performance of both human-human and human-robot dyads.

3.3 Lateral and Rotation Movement Characteristics

While looking into performance metrics, we recognized some patterns in how the dyads behaved in lateral translation tasks. Studying the videos of the lateral motion task, we saw that the follower often guessed the leader's intent incorrectly, and began to rotate when the leader started their movement. When this happened, the leader would flex their arm on one side of the table, causing a torque on the table about the z-axis, and the follower would then commence moving in the correct manner. With this video evidence, as well as the torque change information from Tables 3.1 and 3.2, we looked for in-task patterns of applied torques which could indicate the leader's intent to start either a translation or rotation task.

In order to see in-task relationships, we decided to take each task and look at the time-series torque for the beginning of the tasks. As we looked at the torque data, we noticed two groups ap-



(a) z-axis torque patterns.

(b) x-axis torque patterns.

Figure 3.10: First 4 seconds of trials showing torque trends for rotation and translation tasks for both directions of motion: dashed lines are individual trials, and bold lines indicate trial averages.

pearing. These two groups represented the torque values for the direction of the rotation task, since the dyads were assigned to randomly rotate either clockwise or counterclockwise for each rotation task performed. We then tried looking at the same z-torque time-series data for the translation tasks, and noticed that two more groups appeared, indicating that there was some pattern showing the difference between translation and rotation tasks, as well as a difference depending on which direction the table was travelling. We then took an average of z-torque for each of the 4 distinct groups: translation left, translation right, rotation clockwise (left), and rotation counterclockwise (right). We noticed there appeared 4 groupings of average z-torque for the entire time series. These findings are summarized in Fig. 3.10a, where bold lines are averages of task type.

As can be seen, translation tasks tend to increase in z-torque more quickly, whereas the rotation tasks hover around the same value for over 1 second before diverging. It is evident from this plot that there is a clear difference in torque patterns between the translation and rotation trials, and also the direction of travel. The intent can be classified as either translation left, or translation right, depending on the z-torque value achieved. However, there is not a difference between z-torque patterns for the first second of left and right rotations. Both directions have an approximately constant torque value for this time segment. This is an important time segment, since it is during this interval that decisions about whether to rotate or translate are made by the follower, and there needs to be some indication given by the leader to signal which direction to go.

For this reason, we also looked into what other signals might be given by the leader to indicate which direction to travel, and to clarify whether to rotate or translate. In watching video of the trial, we also noticed some dyads tended to rotate the board about the anterior (x) axis while performing the tasks. So we did a similar analysis with x-torque values to what we did with the z-torque values. The results can be seen in Fig. 3.10b. Here, the difference is not quite as pronounced, but generally, rotating left and rotating right can be separated due to the x-torque value at the beginning of the trial. It also provides additional insight for determining translation task direction, as these are quite distinct in this analysis.

With the analysis on torque triggers providing background for a control method for extended object co-manipulation, we determined what the velocity profile should look like for these tasks. For the translation tasks, we assumed it would follow the bell-shaped velocity profile from a MJ trajectory, however, we wanted to see how the velocity profile looked when translating over a large distance. Bussy et al. [14] showed that humans often accelerate an object to a steady velocity while translating an object. We wanted to verify this, and also determine what velocity most dyads chose as the steady state velocity. To do this, we looked at our 3D Complex Task data. This task involved a large translation portion, followed by changes in direction and rotation of the board to avoid obstacles. Fig. 3.11 shows the first portion of a typical complex task, which is a translation for over 2 meters. We notice from this data that the results seen in Bussy et al. can be verified, and also that the steady velocity achieved is around -0.35 m/s. Since we want our robot controller to work in an undefined situation, we would expect it to be able to perform a translation task indefinitely, if needed. This observation helps inform us that a dyad's desired velocity is some steady state value.

3.4 Human-Human Performance Summary

The information gathered from this study provides us with some more understanding of human-human co-manipulation. First, we know that each dyad potentially performs the task differently than another dyad. In some cases, the dyad performs the task differently each time it performs the same task. Second, we have verified some aspects of human movement for full body motion, such as following a minimum-jerk trajectory. We also saw that dyads adhering to this trajectory corresponded with how quickly a dyad completed the task. Third, the assumption that

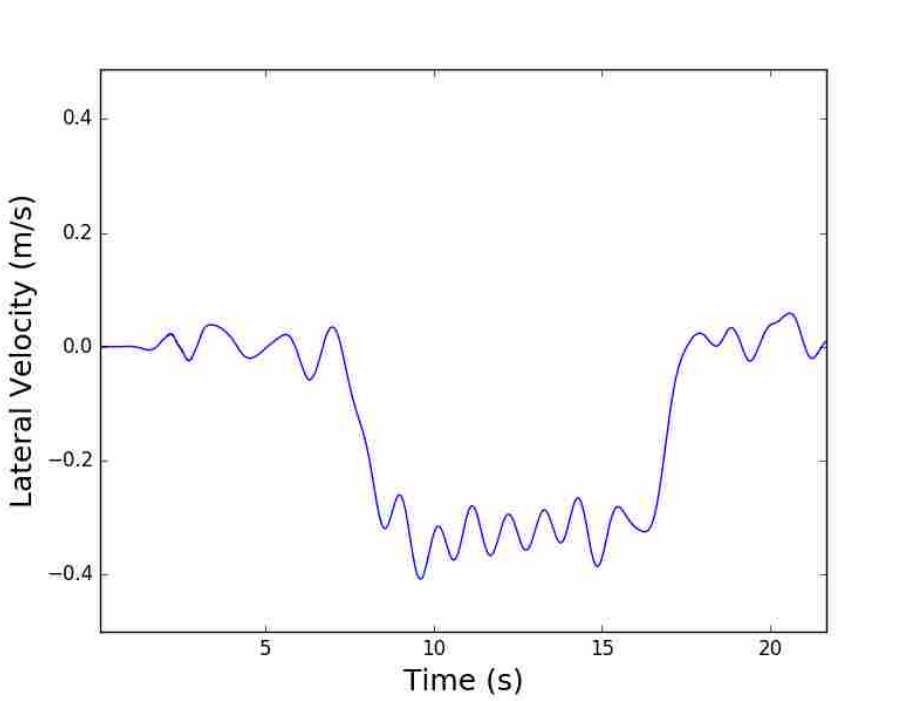


Figure 3.11: Plot showing lateral velocity profile for beginning of Task 5, a 3D Complex Task avoiding obstacles: this portion of the task is a lateral translation for over 2 meters.

all human-human interaction has the goal of minimizing forces was shown to not necessarily be true, as we saw that dyads applied forces in directions orthogonal to motion. Fourth, we explored the correlation between certain performance metrics, and found that performance on some metrics may differ depending on task type, and also performance on certain metrics may be sacrificed in favor of other behavior. Last, new information about torque profiles of translation and rotation tasks provided a unique perspective on how force can be used to solve the problem of co-manipulation of extend objects. Using the tendencies of humans in applied torque patterns from the leader, we can distinguish between task type and task direction intent for the follower to use.

We still are unclear on many aspects of the metrics we have found. We know that for rotation and translation tasks, these metrics are useful in comparing performance to humans, but it is not clear whether these metrics will be applicable to more complex tasks. We also have not exhaustively explored all state relationships, nor all possible metrics, and it is possible there are metrics that can more completely describe human-human dyad performance.

CHAPTER 4. PLANAR EXTENSION OF VARIABLE IMPEDANCE CONTROL

4.1 Motivation and Formulation

In order to verify that the torque patterns we saw in Section 3.3 would be applicable in an extended object co-manipulation scenario, and also to show that current variable impedance co-manipulation techniques are not adequate for extended objects, we built an extension for a variable impedance controller. Variable impedance control (VIC) is a possible solution to undefined or indefinite scenarios, since it is not based on a trajectory, but rather on force inputs which determine robot velocity. What we noticed in practice, though, is that VIC has issues when dealing with bi-manual control and extended object control. We implemented a VIC based on Duchaine and Gosselin's work, [3]. This work was chosen as it most closely represented state-of-the-art in human-robot co-manipulation, and also is capable of adapting trajectory and responding to disturbances. The controller from this work was implemented on our robot platform, which is a Baxter Research Robot mounted onto an AMP-I Mobile Base (see Fig. 4.1). There are force/torque sensors on Baxter's wrists, and the base is equipped with Mecanum wheels making it omnidirectional.

VIC involves the control loop seen in Fig. 4.2. The human communicates their intent to the robot through force sensors, and the VIC model determines a desired velocity based on the applied force, and how the force is changing in relation to the robot's velocity. The general model is shown in Eq. 4.1, and is split into direction specific equations in the figure. Here, F and $\dot{F}$ are applied force and time derivative of force, respectively, $\dot{p}$ and $\ddot{p}$ are velocity and acceleration of the robot, and m, c, and α serve as virtual mass, damping and weighting parameters, which are tuned to define the impedance. The model can be discretized and implemented as a discrete system, solving for the desired velocity at each time step. The state transition matrix and state control matrix are shown in Eqs. 4.2 and 4.3. It is notable that these matrices are dependent on $\dot{F}$, and are therefore changing at each time step. Additionally, c, m and α can be defined separately for each DOF, although in our implementation, we used one variable for all DOF. The discretized

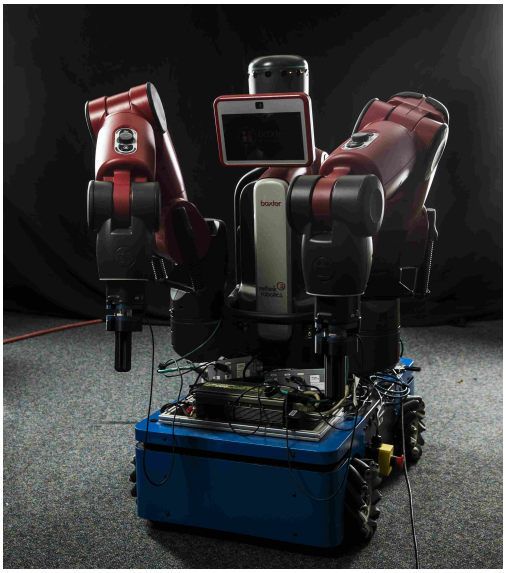


Figure 4.1: Baxter Research Robot from Rethink Robotics mounted onto an AMP-I Mobile Base from HStartTech, with ATI Mini-45 force/torque sensors on each wrist.

state space equation is formulated, and can be seen in Eq. 4.4. To determine matrices **G** and **H**, assuming matrix **A** is non-singular, equations 4.5 and 4.6 are used. The discretized state space equation, Eq. 4.4, can then be used to determine the desired velocity state based on the current state and inputs.

$$F = m\ddot{p} + c\dot{p} - \alpha\dot{F}\dot{p} \tag{4.1}$$

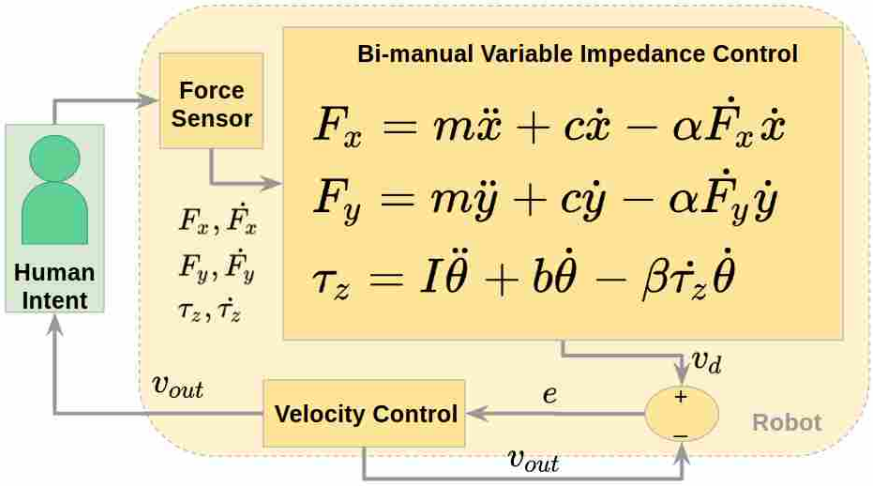


Figure 4.2: Control loop for BMVIC.

$$\mathbf{A} = \begin{bmatrix} \frac{c-\alpha\dot{F}}{m} & 0 & 0 \\ 0 & \frac{c-\alpha\dot{F}}{m} & 0 \\ 0 & 0 & \frac{c-\alpha\dot{F}}{m} \end{bmatrix} \tag{4.2}$$

$$\mathbf{B} = \begin{bmatrix} \frac{1}{m} & 0 & 0 \\ 0 & \frac{1}{m} & 0 \\ 0 & 0 & \frac{1}{m} \end{bmatrix} \tag{4.3}$$

$$\mathbf{x}((k+1)T) = \mathbf{G}(T)\mathbf{x}(kT) + \mathbf{H}(T)\mathbf{u}(kT) \tag{4.4}$$

$$\mathbf{G}(T) = e^{\mathbf{A}T} \tag{4.5}$$

$$\mathbf{H}(T) = (e^{\mathbf{A}T} - \mathbf{I})\mathbf{A}^{-1}\mathbf{B} \tag{4.6}$$

Given the output from the algorithm, we send those velocities–v_x, v_y, etc.–as commanded velocities to the base. We treated the arms as low-impedance devices, as we set the PD gains on a joint torque controller to be low (see Section 4.2). This model was made for single arm manipulation, so we implemented a VIC for each arm independently in order to achieve bi-manual manipulation. However, this is not an ideal method for bi-manual control. Pushing one arm forward and one arm backward could apply zero net force, causing the robot to remain stationary, rather than rotate, as we hoped. To deal with this, we added a torque model to Duchaine and Gosselin's VIC model, as seen in Eq. 4.7. Here, τ and $\dot{\tau}$ are applied torque and time derivative of torque, respectively, with $\dot{\theta}$ and $\ddot{\theta}$ as angular velocity and acceleration, while I, b, and β serve as virtual inertia, damping and weighting parameters. All forces and torques referenced here are with respect to the center of the table. This equation is solved similarly to Eq. 4.1, but the **A** and **B** matrices are defined as seen in Eqs. 4.8 and 4.9.

$$\tau = I\ddot{\theta} + b\dot{\theta} - \beta\dot{\tau}\dot{\theta} \tag{4.7}$$

$$\mathbf{A} = \begin{bmatrix} \frac{b-\beta\dot{\tau}}{I} & 0 & 0 \\ 0 & \frac{b-\beta\dot{\tau}}{I} & 0 \\ 0 & 0 & \frac{b-\beta\dot{\tau}}{I} \end{bmatrix} \tag{4.8}$$

$$\mathbf{B} = \begin{bmatrix} \frac{1}{I} & 0 & 0 \\ 0 & \frac{1}{I} & 0 \\ 0 & 0 & \frac{1}{I} \end{bmatrix} \tag{4.9}$$

The bi-manual torque-based model theoretically allows VIC to be extended to planar motion, where pushing one arm forward and one arm backward will provide a net torque, indicating a desired angular velocity (in the plane only), in addition to any desired cartesian velocities calculated by the original model. In summary, at each time step, Eqs. 4.1 and 4.7 are solved to determine desired velocity and angular velocity to send to the velocity controller. We will refer to the bi-manual torque-based model as bi-manual VIC or BMVIC.

We also extended VIC in a new way, using our results from Section 3.3. We used Eq. 4.1 as a base controller for anterior/posterior translation, and added torque-based triggers for lateral translation and planar rotation. The logic of this extended variable impedance control (EVIC) is shown in Algorithm 1. Torque threshold values were determined by using Figs. 3.10a and 3.10b as a starting reference, then refined to appropriate levels by trial and error. We centered the thresholds around zero for ease of implementation. If none of the torque threshold conditions are met, the algorithm commands no lateral translation or rotation about the superior axis. As mentioned previously, if the torque threshold conditions are met, the robot accelerates until it reaches a specified steady state velocity. The robot acceleration was limited to the capabilities of our robot platform. A control loop showing how this algorithm is implemented is shown in Fig. 4.3. The main difference between EVIC and BMVIC is that EVIC uses Algorithm 1 to determine the desired lateral and angular velocities, whereas BMVIC relies on Eqs. 4.1 and 4.7 to calculate the desired lateral and angular velocities.

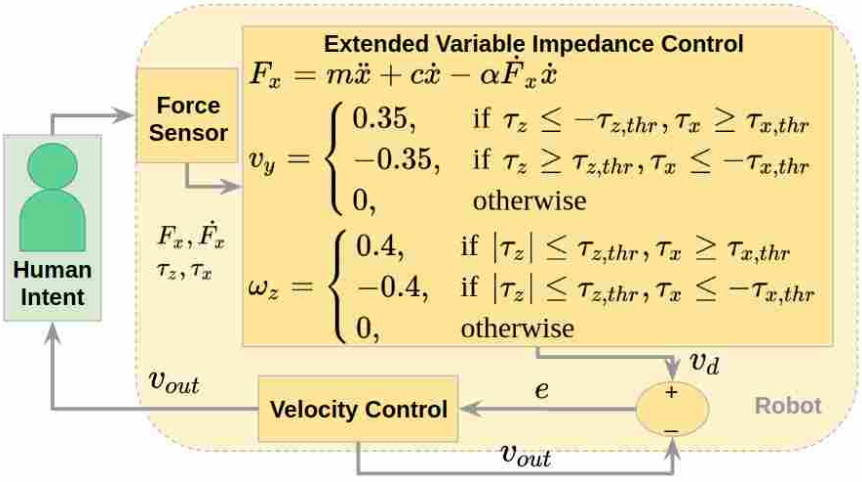


Figure 4.3: Control loop for EVIC.

Algorithm 1 Extended Variable Impedance Control

Require: τ_z, τ_x
  if $\tau_z \leq -\tau_{z,threshold}$ and $\tau_x \geq \tau_{x,threshold}$ **then**
    Left Translation
  else if $|\tau_z| \leq \tau_{z,threshold}$ and $\tau_x \geq \tau_{x,threshold}$ **then**
    Right Rotation
  else if $|\tau_z| \leq \tau_{z,threshold}$ and $\tau_x \leq -\tau_{x,threshold}$ **then**
    Left Rotation
  else if $\tau_z \geq \tau_{z,threshold}$ and $\tau_x \leq -\tau_{x,threshold}$ **then**
    Right Translation
  else
    Stop

4.2 Extended Object Co-manipulation Implementation

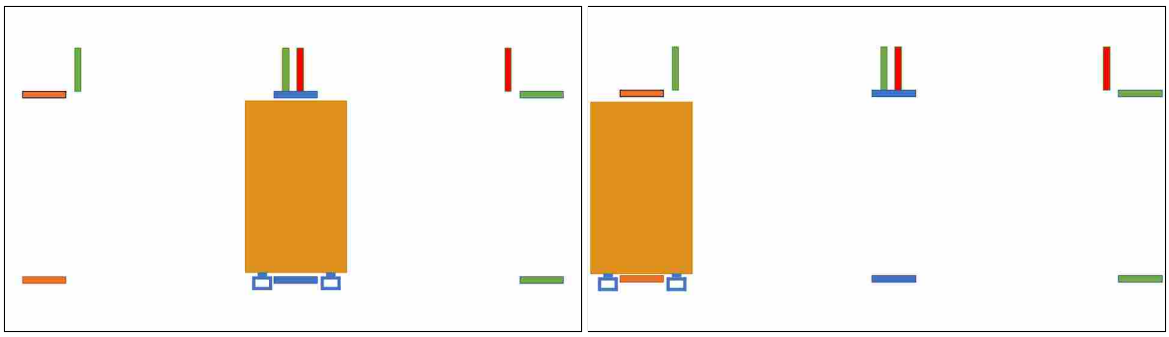
We implemented both BMVIC, as well as the EVIC on our robot platform, shown in Fig. 4.1. A video showing EVIC running can be seen at `https://youtu.be/5vicqv788dI`. Our purpose in implementing these controllers was to determine feasibility of the controllers, and also to get initial data quantifying performance against the blindfolded human-human dyads. We ran both controllers testing the capability of planar co-manipulation of an extended object on the following criteria: lateral translation and planar rotation, or rotation about the superior axis. The Baxter arms were also running an impedance controller to maintain their position, while giving them a more human-like (low impedance) behavior. Humans have been shown to exhibit low impedance be-

havior for certain tasks [47]. The impedance control law, given in Eq. 4.10, used K_p and K_d gains of [40, 120, 40, 16, 8, 10, 12] and [7, 8, 4, 7, 1.5, 1.5, 1] respectively. The desired angles, q_{cmd}, used were [0, -0.84, -1.27, 2.26, -0.34, -1.22, -2.25] radians and [0, -0.84, 1.27, 2.26, 0.34, -1.22, 2.25] radians for left and right arms respectively. We ran the controller at a rate of 500 Hz, manipulating the same table from our human-human dyad experiment (see Fig. 3.2.) For determining performance of the controllers, we compared the completion time and MJ error for both lateral and rotational tasks. We also had a qualitative metric: whether BMVIC, EVIC, or neither controller was preferred.

$$\tau_{cmd} = K_p(q_{cmd} - q) - K_d\dot{q} \tag{4.10}$$

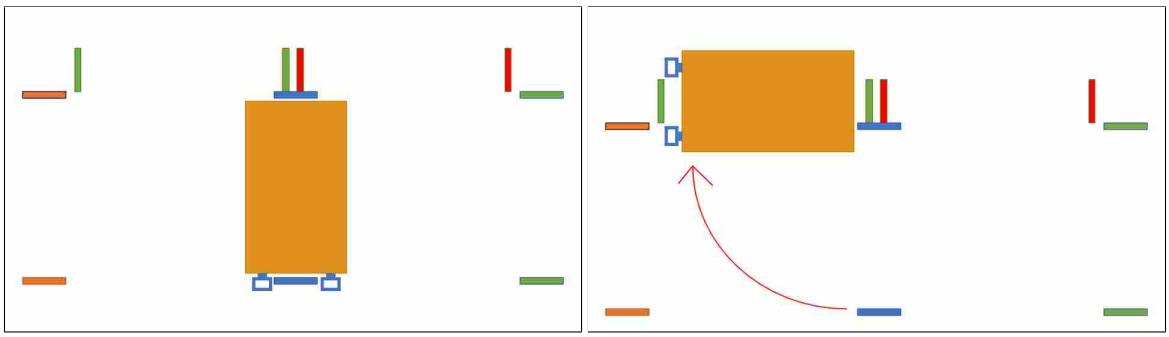
We invited 4 participants to test EVIC. The participants, all with some experience with robots and familiarity with the research, were trained to a competency level, then performed Tasks 3 (lateral translation) and 4 (planar rotation) from our previous experiment, (see Figs. 4.4 and 4.5 for task instructions). The tasks were performed 6 times each. We also collected the same data–force, motion, and point cloud–as in our human-human experiment. We then had the participants test BMVIC, using the same process as with EVIC. Unfortunately, when the first participant was being trained on BMVIC, the robot became unstable and sheared an internal component in the arm. We replaced the component that broke, and ran testing with the next subject. While the second subject was performing a rotation task, the same internal component on the other arm sheared. We recognize this does not allow for a detailed comparison of performance between BMVIC and EVIC, however we can make two important points. The first is that forces induced on the robot arms by the BMVIC method were unsafe and detrimental to the robot. Since EVIC is based on human-human data, we can reasonably conclude that BMVIC requires forces outside of normal human comfort or intuition. The second point is that the two participants who were able to run both control methods indicated the preferred (pre-robot failure) control method was EVIC.

While this was an unfortunate circumstance, we were able to obtain results for EVIC with all 4 participants, as can be seen in Table 4.1. This table explores some of the metrics formulated in Chapter 3.2.4. EVIC performs reasonably well, and is within 30% completion time of humans for both the translation and rotation tasks. However, we recognized that there is some learning curve



(a) Task 3 (translation task) starting position. (b) Task 3 (translation task) ending position.

Figure 4.4: Top view of task 3 (lateral translation task) beginning and ending positions, as given to the leader via tablet. Red, orange, green, and blue lines correspond to tape lines on the floor of the arena, and help the leader navigate the task.



(a) Task 4 (rotation task) starting position. (b) Task 4 (rotation task) ending position.

Figure 4.5: Top view of task 4 (rotation task) beginning and ending positions, as given to the leader via tablet. Red, orange, green, and blue lines serve same purpose as in Fig. 4.4.

for the participants, so we also looked at only the last 2 trials (out of 6) for each participant. Looking at the last 2 trials, we can see a significant increase in user performance. The completion time for translation trials went down to 15% difference between EVIC and human-human blind-folded trials, whereas the rotation trials surpassed human-human dyad performance. In both translation and rotation cases, EVIC also managed to outperform the human-human dyads in MJ error.

While we recognize this is a small population sample, we see these initial results as significant since this control method is able to perform near human-human levels with a minimal amount of user training. In the future, we plan on running more participants through the tasks with EVIC to see how these results differ with a larger sample size. Another result from this pilot study is evidence that EVIC can perform both anterior and lateral translation, and can smoothly navigate a

Table 4.1: Comparison of EVIC and blindfolded human-human dyads.

	Human-Human	EVIC	EVIC Last 2
Task Time (s) – Tran.	7.05	9.09	8.12
MJ Error (m) – Tran.	215.29	65.61	43.02
Task Time (s)– Rot.	7.58	9.32	7.38
MJ Error (rad) – Rot.	368.85	135.62	126.15

rotation task, which is a significant step forward in pHRI co-manipulation. Another notable outcome from the pilot study is that our naive use of arm and base coordination did not invalidate the control model. This means the EVIC approach is still viable, and should improve with more explicit arm and base coordination, as well as with improvements to the robot platform, such as a legged base, or a base with quicker acceleration.

CHAPTER 5. NEURAL NETWORK PREDICTION CONTROL

5.1 Motivation

While EVIC, described in Algorithm 1, was a useful method for planar co-manipulation, it took a considerable amount of time searching for patterns by hand to find the necessary information. This is especially true if we want to find patterns for higher DOF tasks, where even more forces and torques may be coupled to produce specific movements. We decided to, rather than sift through the data by hand, use machine learning to look through the HHI data to predict human intent.

Fig. 5.1 describes a control loop for pHRI. Because our data considered the interaction between a human leader and a human follower, the input x could be considered what the leader did–applied forces or moved the object–to indicate their intent to the follower. The follower then deciphered the intent, x_d, and moved as they believed appropriate, x_r. The leader then reacted to this motion. Essentially, the leader put in a stream of data that indicated intent to move the object in a certain manner, which the follower then deciphered and acted on the intent. An estimator created through machine learning should follow this basic input data to velocity command outline.

There are a variety of neural network structures that could be considered for this purpose. Previous research done by Yan and Li [48] used RNNs to model the inverse dynamics of a 2 joint robot. Also, Wu and Song [49] showed that a feedforward disturbance term could be modelled by an RNN for a legged robot. These works indicated that robot controllers can be described with RNNs. Other work, done by Martens et al. [50], showed how Recurrent Neural Networks (RNNs) are used in predicting data, specifically text. A sequence of characters can be fed into the RNN as an input, and the RNN will predict the next character in the sequence. In considering what form our neural network should take, the RNN model proved to be the most applicable to our work. From our exploratory study, we had sequences of forces applied to and motion of a table that could be used as inputs to an RNN. We determined that we could feed the force and motion data, as a

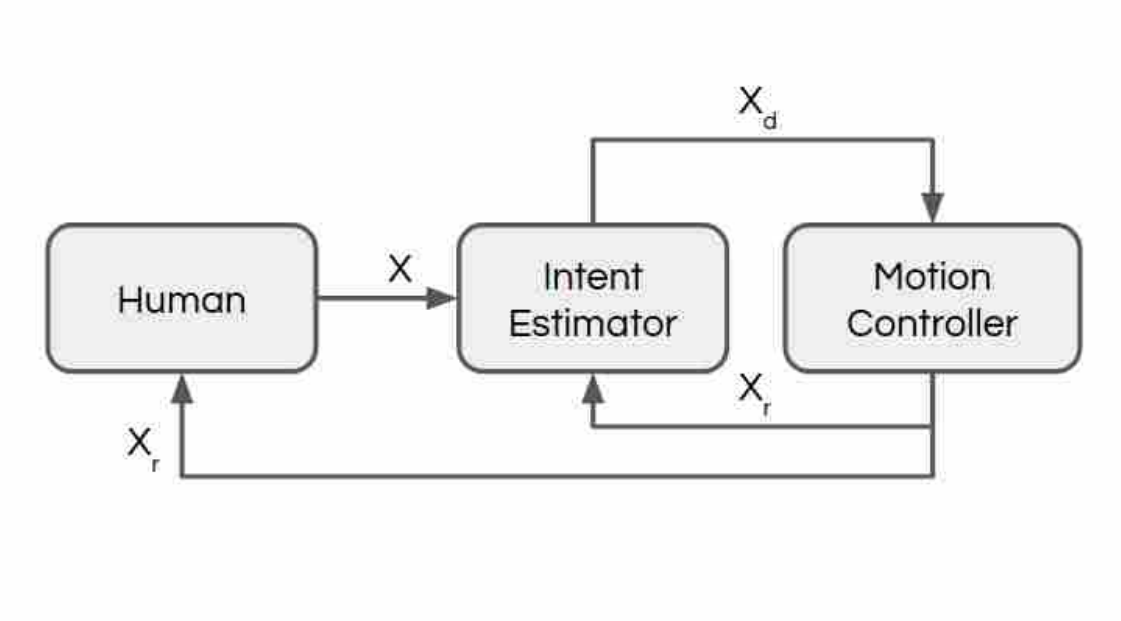


Figure 5.1: Basic control loop structure of intent estimation in co-manipulation. The human moves the co-manipulated object, and the motion of the object, x, is fed into an intent estimator, which determines a desired motion of the robot, x_d. The robot's motion, x_r then influences the object motion, as well as the human's.

substitute for characters in other RNNs, and receive a motion prediction as an output, similar to how RNNs are used for predicting text. This prediction encapsulates the human intent, encoded in the desired velocity of the object in the object frame, and provides a goal for the robot to achieve.

5.2 Architecture

Because humans do not instantaneously react to stimuli, it is likely that the table movement is dependent on multiple past states and inputs. Since there was precedent for using RNNs as portions of control loops, and as predictions of future states, we decided on using the RNN structure. To implement this, we used a number of previous time steps of motion data (velocity and angular velocity) of the table as inputs. For this thesis, the measured and desired velocities for the neural network are all taken in the object's reference frame. While we had both force and motion data at our disposal from the experiment, only motion data was included in our neural network. Our reasoning for not including force data is that including it caused our neural network solution to have poor convergence. We know that forces play an important part in any dynamical system, therefore

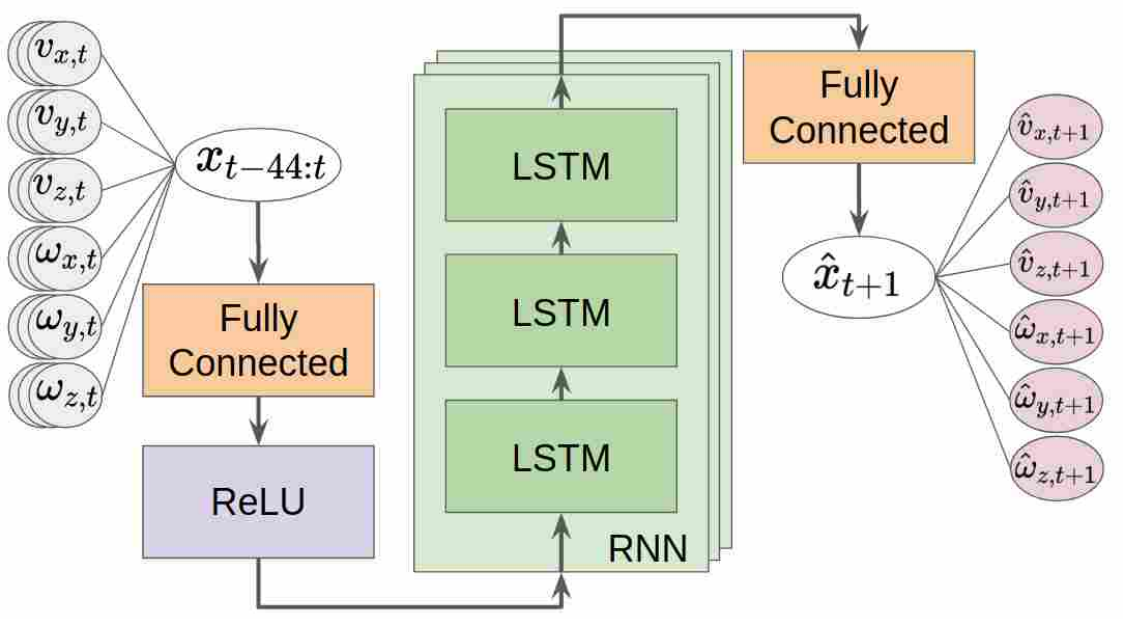


Figure 5.2: Basic neural network structure; time-series motion data inputs (left) enter the network, are sent through a fully connected layer, a ReLU layer, through an LSTM Cell RNN, and another fully connected layer, then finally predicted velocities are given as outputs (right).

we believe that including a model where forces can be used should be a priority in future work. Additional changes that could improve our model include: changing the RNN structure, optimizing number of time steps to consider, or using a different structure of neural network. However, this thesis does not consider these options, but leave them for future work, as we obtained an accurate prediction using our current network with motion data inputs. The basic structure of the neural network is shown in Fig. 5.2. Initial development of the network architecture was done in [2], but some changes were implemented to help improve the structure and predictions. Specifically, the RNN structure with LSTM cells was implemented in this thesis, as well as combining translation and rotation predictions into a single architecture, and removing acceleration as an input.

Our final network consists of 3 LSTM layers each with 120 hidden states. The process of choosing a neural network structure and other parameters was not exhaustive and it is possible that better structure and parameters could be obtained. However, the process for choosing neural net parameters and training took over 10 hours. With more experience and more efficient computation, we may have been able to find better structures, but this is also left to future work. Methods other than neural networks may exist, but our purpose in this thesis is to show that human intent

estimation is possible based on the data collected from our user study, and neural networks allowed us to achieve this goal.

Additionally, it was shown by Chipalkatty et al. that more complex predictions of future movement can actually decrease performance if they do not agree with what the human is trying to do. This is because humans cannot be completely modeled due to their unpredictability. Chipalkatty et al. found that it was more important that the human understand what the robot will do next, meaning that our controller should be intuitive for a human partner in a human-robot dyad [51]. In addition to being intuitive, the prediction should also be accurate and repeatable. The inputs to the neural network, as seen in Fig. 5.2, are 45 past steps of velocity and angular velocity of the table in the x, y, and z directions, $\{x_{t-44}, x_{t-43}..., x_{t-1}, x_t\}$. Since motion data was collected at 200 Hz, each time step represents 0.005 s, and 45 time steps is 0.225 s. The outputs are the predicted velocity and angular velocity of the table in the x, y, and z directions for 1 time step into the future, $\hat{x}_{t+1}$, where $\hat{x}$ indicates a predicted value.

Our neural net formulation also uses what Engel et al. describe as iterated prediction [52]. The neural network itself only predicts 1 time step into the future. Then, the prediction, $\hat{x}_{t+1}$, is appended to the input to give $\{x_{t-44}, x_{t-43}..., x_{t-1}, x_t, \hat{x}_{t+1}\}$. The first step of the input is dropped to obtain a new input of past motions for the neural net, $\{x_{t-43}, x_{t-42}..., x_t, \hat{x}_{t+1}\}$. The new data is input into the neural net which outputs a prediction 1 step forward, but 2 total steps into the future, $\hat{x}_{t+2}$. This is then appended to the input. The process is repeated 10 times to obtain a prediction of 10 steps, $\{\hat{x}_{t+1}, \hat{x}_{t+2}..., \hat{x}_{t+9}, \hat{x}_{t+10}\}$. This process is represented by Fig. 5.3. Because the outputs of each prediction step become the inputs for the next, the inputs and outputs must be the same variables.

5.3 Neural Network Creation

5.3.1 Training

We pre-processed the data for the neural net to improve the results. The velocity data was scaled to have 0 mean and standard deviation of 1 over the entire set of data. This was then inverted on the output to show the results in their proper units. This same scaling can be used on new data as long as the mean and standard deviation are similar to the training data. This is the case in

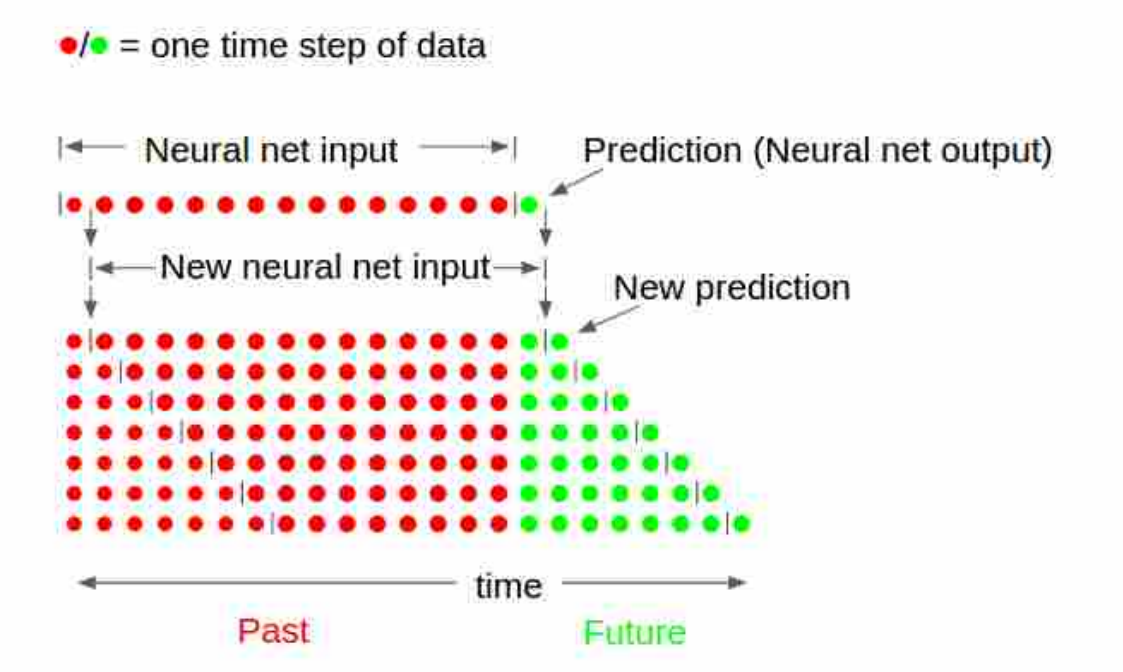


Figure 5.3: Illustration of iterated prediction, where previous time steps are used to obtain one future prediction of states, which is then appended to previous time steps. The first time step is removed, and the network is run again in order to achieve multiple future predictions. Included from [2] with permission.

our experiment, as velocity values fall into the average adult human range. The entire set of data consists of 2.5 million time steps for each variable. Data was split into to training and validation sets. 75% of the data was assigned to the training set and the other 25% to the validation set.

Training was performed by creating batches of data that randomly pulled in 32 sets of 45 steps of data from the entire training set. Sets of 45 steps were also created from the validation set. The neural net was trained on new training batches for a number of iterations, until the cost function was below a threshold we chose. We used the mean squared error (MSE) for the cost function, as shown in Eq. 5.1.

$$MSE = \sum_{n=1}^{32} (\hat{x}_{n,t+1} - x_{n,t+1})^2 \tag{5.1}$$

Once this threshold was reached, each training batch took in the original data, created a prediction, appended the prediction to the end of the time-series, and removed the first step, following the pattern shown in Fig. 5.3. This process was repeated until the desired number of

future predictions was reached. By training the neural network on data that includes predictions, the stability of the prediction is improved.

Another benefit of iterated prediction is the inclusion of predicted velocities each training step reduces the amount of overfitting, since new data is essentially being introduced each iteration. By training this neural network, we have created an intent estimator to be used in a human-robot dyad performing co-manipulation of an extended object. The object's motion data can be fed as an input at each time step, and the trained neural network will output a predicted velocity for the object. The robot must then use this estimation to calculate its own trajectory to make the object follow the desired behavior.

5.3.2 Validation

After training, the validation set was used to evaluate the prediction. This was done in a similar process to [2], where the validation set was run through the trained neural network. The predicted velocity and angular velocity values were compared to those from the HHI data as shown Fig. 5.4. A magnified look a the last 3 seconds of predictions can also be seen in Fig. 5.5. These figures show a single translation task with v_x, v_y, and w_z as states. Thin lines in the plot are actual velocities and bold lines are predictions. Note that the predictions are shown at 10 second intervals to increase readability of the plot. As can be seen, the predicted velocities are fairly close to the actual velocities from the task. While the prediction deteriorates as we move farther along the iterated prediction, this is acceptable, as only one of the first few predictions will be used for control, and then a new prediction will be generated. This validation, along with the preliminary results with a slightly different neural net architecture from [2], confirm that the neural network is accurately predicting human intent in the form of desired velocity. Acceleration data can also be included, but we found it is not necessary for accurate predictions.

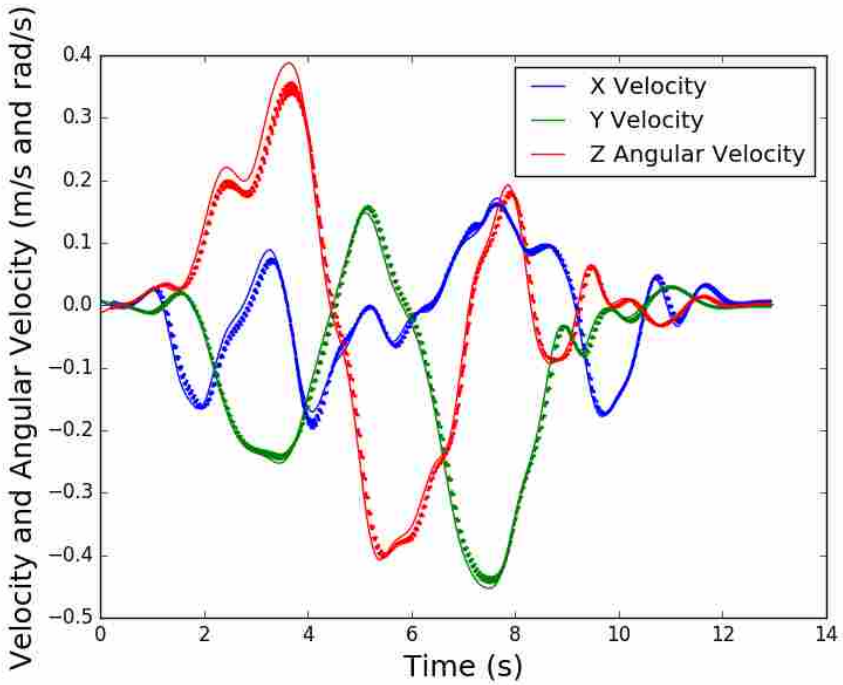


Figure 5.4: Validation of neural network for a lateral translation task, thin lines are actual velocities and bold lines are predictions for future time steps.

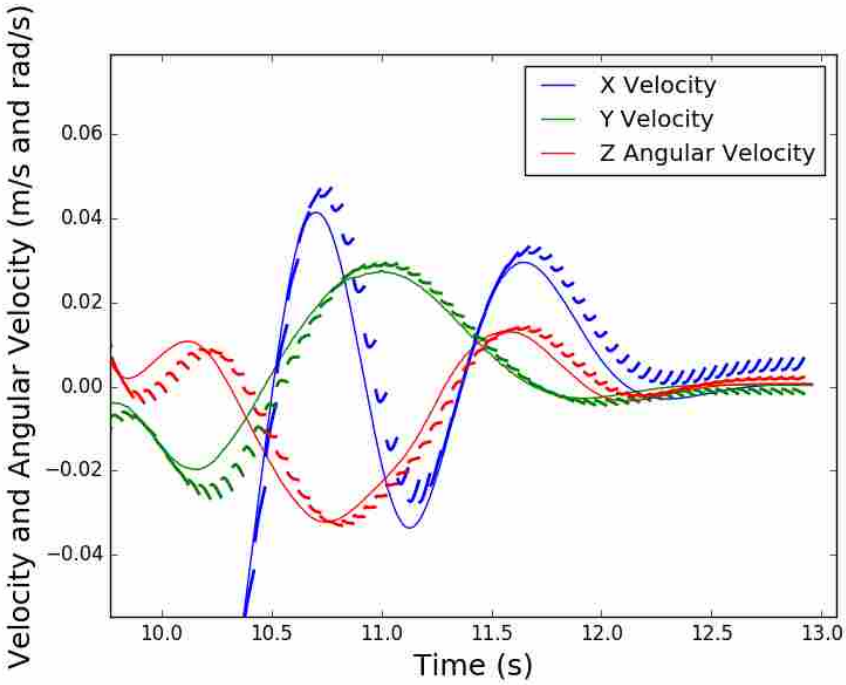


Figure 5.5: Last 3 seconds of validation plot (Fig. 5.4).

5.4 Neural Network Prediction Control

As discussed previously, and displayed visually in Fig. 5.1, predicting human intent is only one portion of the puzzle. We also need to tell the robot what to do with the prediction with a motion controller. The neural network outputs a predicted velocity and angular velocity of the center of mass of the table. Along with the prediction of the velocity of the center of mass, we can know the velocity of other known points on the table, such as the velocity of the leader's edge and the follower's edge, using kinematic relationships. For our motion controller, however, we simply used the predicted velocity of the center of mass, again in the object's reference frame. The robot controller then needs to account for the distance from its center of mass to the table's, which can be done using the transport theorem, shown in Eq. 5.2. Here, $\vec{v}_r$ is the robot's calculated velocity in its reference frame, with $\vec{p}$ as the distance from the table frame to the robot frame, and $\vec{\omega}$ as the table's angular velocity in the table frame. Also, $\vec{v}_{rel}$ is the table's velocity in its frame. For reference, Fig. 5.6 shows the reference frames of the table and robot, as well as how p and ω are related to these frames. Notice that the table frame and robot frame are rotated 180 degrees from each other. If we assume that the robot frame does not rotate independently from the table frame, we can then rotate the velocities into the table frame.

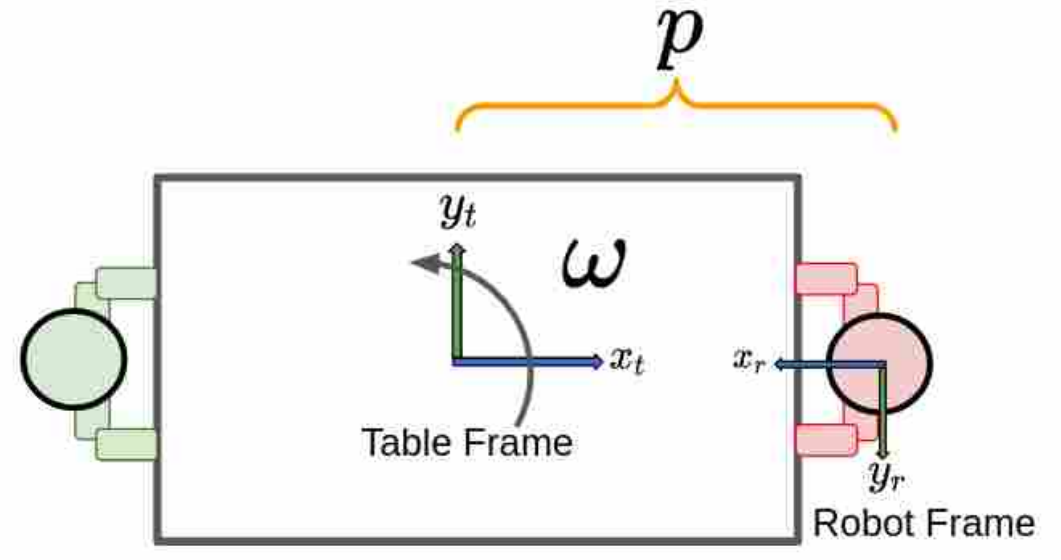


Figure 5.6: Frames of table and robot where p is the distance from the table to robot frame.

$$\vec{v}_r = \vec{v}_{rel} + (\vec{\omega} \times \vec{p}) \tag{5.2}$$

We now have the tools to complete the control loop shown in Fig. 5.1. The intent estimator is replaced with the neural network model. The motion controller is replaced with Eq. 5.2, and is subsequently fed into the low level control of the robot, which sends voltages down to the wheels to match the desired velocity. The achieved velocity, x_a, is then what the human interacts with, completing the loop. Also, x_a is estimated using numerical differentiation and a 2nd order low-pass filter of the pose information coming from the motion capture. This loop is shown in Fig: 5.7. We call this control method Neural Network Prediction Control (NNPC). A notable feature of this method is that the commanded velocity, x_r, is a continuous variable on $[-v_{max}, v_{max}]$, where v_{max} is determined empirically for each DOF. This means the human user has control of the speed of the interaction, so if the response x_a is not suitable to the human, they can adjust their inputs to move faster or slower.

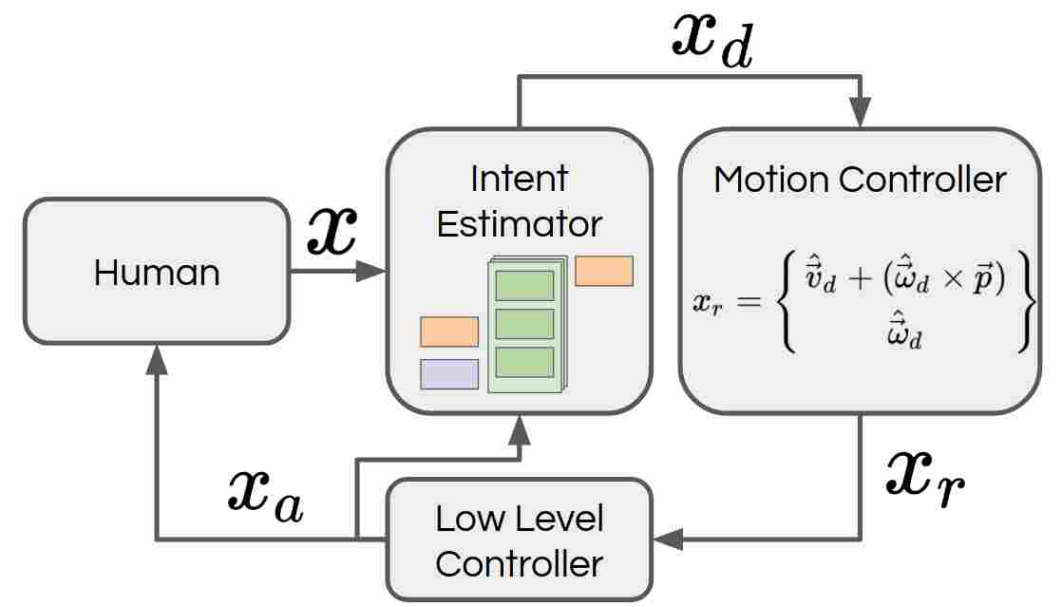


Figure 5.7: Neural Network Prediction Control loop.

CHAPTER 6. PHRI CO-MANIPULATION EXPERIMENTAL STUDY

6.1 Experiment Description

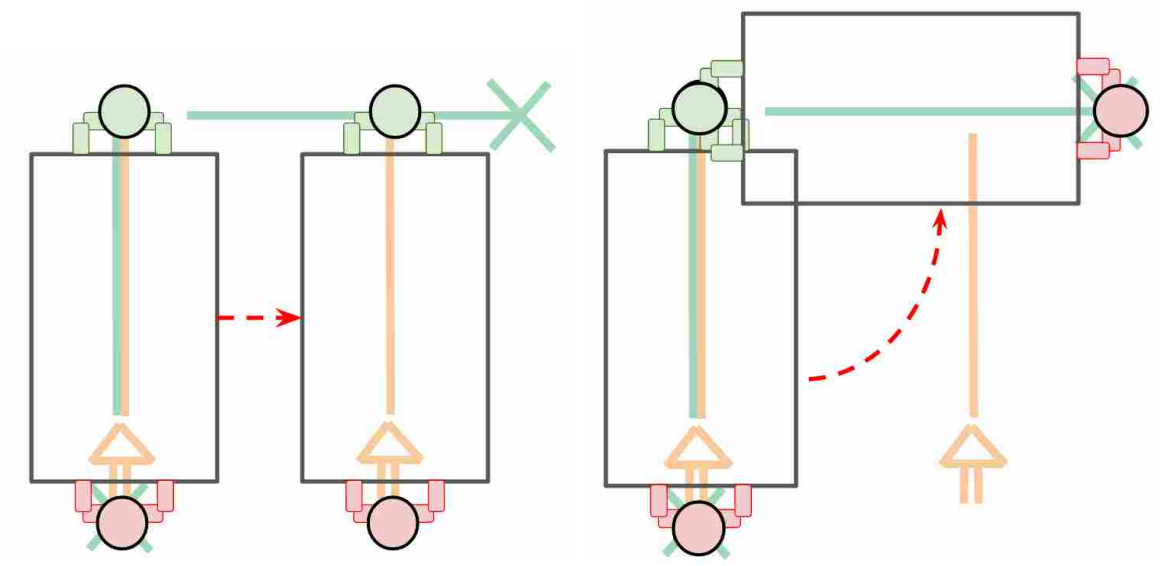
As mentioned previously in Chapter 4 and [53], EVIC works only for 3 DOF planar control–anterior and lateral translation and rotation in the plane–so we developed an experiment to compare a planar implementation of NNPC and EVIC. NNPC theoretically can be applied to control 6 DOF, but we have left this for future work. This experiment was designed to be as close as possible to the lateral translation and planar rotation tasks from our 6 DOF HHI experiment [1]. The same robot platform setup was used for NNPC as was used for EVIC in Chapter 4.2.

6.1.1 Tasks

Because both controllers are planar co-manipulation methods, we used only planar tasks from the previous HHI experiment in [1]. As can be seen in Fig. 6.1, each participant performed two tasks: translation and rotation. In this diagram, the human is represented by the bottom person. The translation task consisted of the subject moving from one orange arrow to the next, with tape lines extending on the ground to help the user align the board correctly. Rotation tasks were similar, except with the participant moving from one green x to the next. Tasks could be run starting at either arrow or x, and the direction was randomized throughout the trial. Additionally, Fig. 6.2 shows a participant running a translation trial in the arena.

6.1.2 Equipment

The position of the board was tracked via Cortex Motion Capture software with a Motion Analysis Kestrel Digital Realtime System. A total of 8 Kestrel cameras were used to track 8 infrared markers placed on the board. Using a static global frame established by the motion capture system, the position and orientation of the board could be tracked over time, and we transformed



(a) Translation task diagram. (b) Rotation task diagram.

Figure 6.1: Task motion for the translation and rotation tasks, start and stop points for leader delineated by arrows for translation and Xs for rotation.

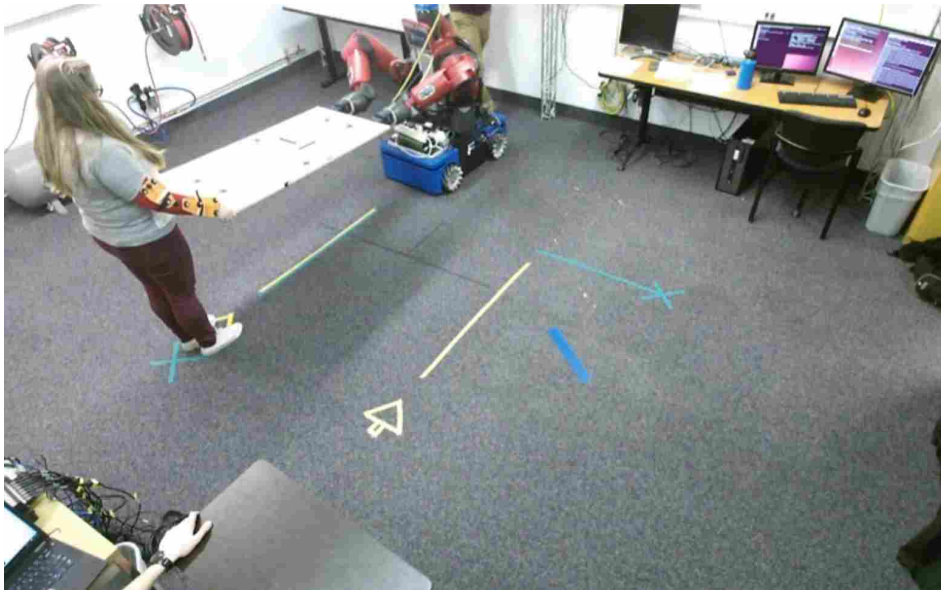


Figure 6.2: An image of a participant co-manipulating the table during a translation task.

the data into the robot's frame for use in the neural network. The motion capture data was collected at a rate of 200 Hz. In order to run NNPC, we need a method of telling the controller the object velocity. We used a 2nd order low-pass filter and numerical differentiation on position and orientation data to get the object velocity. Additionally, participants wore sleeves with infrared markers

to track the position of their arms during the trial. This data was not used during analysis, but was collected to match similar data collected during the experiment in [1].

The object the teams moved was a 59x122x2 cm wooden board – meant to simulate an object (like a table) that is difficult for one person to maneuver, which weighs 10.3 kg. Attached to the robot end of the board were a pair of ABS 3D-printed handles, to which two ATI Mini45 force/torque sensors were fastened. The sensors transmitted data via ATI NET F/T Net Boxes, which passed data over Ethernet to the computer at a rate of 100 Hz. The sensor is attached to ABS 3D-printed wrist adapters on the other side, which fasten to Baxter's wrists.

The test arena was a volume measuring 490x510x250 cm. The arena was also equipped with a video capturing device, with which Fig. 6.2 was captured. The device we used was a Microsoft Kinect 2, which allowed us to capture 3D point cloud data, as well as color video of each trial. Although we did not use the point cloud data for analysis in this thesis, the data may be useful in future work.

6.1.3 Subjects and Procedure

Subjects for this study were male and female students from Brigham Young University in Provo, UT. There were a total of 16 students–4 female and 12 male–ranging from 18-23 years of age, with an average age of 20. Students were from a variety of majors, with STEM majors making up a majority. Participants were asked to rate their familiarity with robots on a scale from 1-5, with 5 being the most familiar, and the average rating was 2. IRB approval was obtained for this experimental study.

Participants entered the Robotics and Dynamics Lab and filled out the appropriate waivers. They were then briefed on the purpose of the research and given an introduction to what data would be collected, and what would be expected of them. Sleeves were then placed on the participants arms in order to track their arm motion during the trial. Subjects were then given basic operating instructions for both EVIC and NNPC controllers. This instruction included how to translate in the anterior and lateral directions, and how to rotate the board for each controller. A controller was randomly selected, and each participant practiced with that controller until they were able to complete a competency task, they moved on to the other controller, and repeated the competency training. The competency task consisted of aligning the board with the tape lines on the ground,

starting from a translated and rotated position. The practice assured us that each participant would have at least enough familiarity to complete the translation and rotation tasks.

Once competency training was completed, a controller was selected at random to be the first controller for data collection. The randomization of controllers was counterbalanced. Participants knew the controllers only as A (NNPC) and B (EVIC). They were not given any specific details about the formulation of the controllers, other than the basic operating instructions in the competency task. The subjects then ran a series of translation and rotation tasks with the selected controller. Tasks were randomized (counterbalanced) in order between translation and rotation. Once a type of task, either rotation or translation, was selected, the participant ran that task type one direction, and then ran the same task type, but in the other direction. The direction order was not randomized, and the first direction is indicated in Fig. 6.1. Due to the nature of the controller, the robot was not able to lift the table from the ground, so the table was laid on a rest stand between trials. A single trial consisted of the subject lifting the table off of the rest, then a researcher would remove the rest from below the table. Once the rest was completely out of the way, the subject then performed the specified task. Participants indicated they were finished by verbally communicating completion. Once they indicated they had completed the task, a researcher would replace the rest underneath the table, and the participant would lower the table back onto the rest. Each task was repeated 6 times, 3 one direction and 3 the other direction, for each controller. Once trials were completed for one controller, the participants were given a survey, and asked to rate the controller on certain qualitative characteristics. Once completed, they moved on to the other controller.

A video showing a representation of tasks can be seen at `https://youtu.be/4b-wxn9_gFQ`. This video was taken after the participant had completed all trials, so we could use a higher definition camera to record video, and what is seen is a good representation of the skill level of the human-robot dyad post-experiment.

6.2 Results and Discussion

6.2.1 Metrics

There are a number of metrics that can be used to quantify performance of the controllers. A good summary of these metrics is found in [6], as well as our previous discussion in Chapter 3.

Among these metrics are a few that are especially applicable to the tasks and control methods we have developed. These are: minimum jerk, completion time, and minimum torque change. While none of these metrics alone can completely store all the information of each controller, together they give us a fairly good indication of how each controller was performing in relation to HHI data from [1].

Minimum jerk error (MJE), or deviation from a minimum-jerk trajectory, is a measure of how close the actual trajectory was to a minimum-jerk trajectory in meters, and is calculated using Eq. 6.1. This measure accounts for human tendency to match these trajectories. Completion time is simply the time from the start of the task to the end of the task. The start of the task is calculated by determining when the object has first moved beyond 5% of the distance between the starting and ending y or θ_z positions, for translation and rotation respectively. The end of the task is calculated by determining when the object settles into 95% of the distance between starting and ending y or θ_z positions. A buffer of 0.5 s is added to the total time to account for the missed motion. This measure accounts for how quickly a dyad performed the task. Although quick task completion is not always a direct objective of dyads, this measure can help compare the capabilities of two dyads. Minimum-torque measure (MTM) is a measure of how much the time-derivative of torque changes over the course of the task. In instances where the follower predicted incorrectly, there was an unforeseen obstacle, or some other disturbance, MTM can account for human tendency to reduce the amount of force or torque required to move. It is calculated using Eq. 6.2.

$$MJE = \sum_{t=0}^{T} x_{mj,t} - x_{a,t} \tag{6.1}$$

$$MTM = \sum_{t=0}^{T-1} \dot{\tau}_t^2 + \dot{\tau}_{t+1}^2 \tag{6.2}$$

6.2.2 Quantitative Results

While each task type was performed 6 times for each controller, we only consider the data from the last 2 trials performed for each task type. Our reasoning behind this is that each participant learned throughout the experiment, so we want to take the best representation of the control method. This is acceptable to us, because if a human-robot team were to be deployed in

Table 6.1: Metrics of EVIC and NNPC for rotation and translation tasks, also compared with blindfolded HHI and sighted HHI data from [1].

Metric and Task Type	Blind HHI	EVIC	NNPC	Sighted HHI
Completion Time (s)–Rotation	7.08	8.25	8.26	6.58
Completion Time (s)–Translation	7.18	7.91	7.75	4.93
MJE (rads)–Rotation	392.71	96.44	87.38	344.70
MJE (m)–Translation	149.91	50.24	48.51	98.92
MTM ($N^2 \cdot m^2/s^2$)–Rotation	488454.38	65602.60	12770.75	341253.43
MTM ($N^2 \cdot m^2/s^2$)–Translation	387937.56	48191.90	15220.89	151758.83

real life, the human would undoubtedly be trained on working with the robot. If we consider the practice time and time taken to perform the first 4 trials of each type, the total training time for our participants was approximately 30 minutes.

The results from the experiment based on the metrics in the previous section can be seen in Table 6.1. This table compares EVIC and NNPC to each other, as well as to the lower (blind HHI) and upper (sighted HHI) bounds of human performance. As can be seen, NNPC performed best in most of the metrics. It was able to get closer to blind HHI performance in completion time, and outperformed EVIC, as well as both blind and sighted HHI performance in both MJE and MTM, where lower numbers indicate more efficient performance in that task. EVIC, while not quite as good, still outperformed blind and sighted HHI in most of the metrics, except for completion time. It is notable that the blind HHI performance captured here is for a human-human leader-follower dyad, where the follower was blindfolded, and communication was limited to haptic communication only, whereas sighted HHI allowed for communication in any form desired by the dyad.

We also determined the statistical significance of these quantitative results. The mean values of the metrics for both the EVIC and NNPC controllers were compared with each other, and also with the blind and sighted human-human dyads, based on the 3 metrics used in this section. We ran an unpaired t-test to determine p values, and also determined Cohen's d to calculate an effect size. From the p values, we were able to see which groups were statistically likely to have the same mean, and therefore see whether there was a statistical difference between the groups, based on a standard $p < 0.05$ criteria. With the effect size, we were able to determine the strength

Table 6.2: Statistical significance of quantitative metrics.

	p			Cohen's d Effect Size		
Comparison Groups	Comp. Time	MJE	MTM	Comp. Time	MJE	MTM
EVIC vs NNPC Trans.	0.73	0.89	0.017	Small	Medium	Large
EVIC vs NNPC Rot.	0.98	0.70	0.14	Very Small	Small	Medium
EVIC vs Blind Trans.	0.07	0.00	0.00	Medium	Huge	Huge
EVIC vs Blind Rot.	0.06	0.00	0.00	Medium	Very Large	Very Large
NNPC vs Blind Trans.	0.16	0.00	0.00	Medium	Huge	Huge
NNPC vs Blind Rot.	0.05	0.00	0.00	Medium	Very Large	Very Large
EVIC vs Sighted Trans.	0.00	0.00	0.00	Huge	Large	Very Large
EVIC vs Sighted Rot.	0.01	0.00	0.00	Large	Large	Large
NNPC vs Sighted Trans.	0.00	0.00	0.00	Huge	Large	Huge
NNPC vs Sighted Rot.	0.01	0.00	0.00	Large	Large	Very Large

of each comparison. Effect sizes were calculated, and then categorized into very small, small, medium, large, very large, or huge categories, based on Sawilowsky's work [54]. These statistics are summarized in Table 6.2.

There are a few results which are important to recognize from this analysis. First, EVIC and NNPC are not statistically different in terms of completion time or MJE, but do seem to differ in MTM, which additionally has a fairly large effect size. Second, both EVIC and NNPC are not statistically different from the blind human-human dyads in terms of completion time. Last, EVIC and NNPC are statistically different from both blind and sighted human-human dyads in terms of minimum-jerk error and MTM, and these comparisons are all categorized as large or higher. A more in-depth discussion of these results is found in Chapter 6.2.4, but as a short summary, the statistics show that these controllers have approached a level comparable to blind human-human dyads with respect to the completion time metric, but are still quite distinguishable in terms of MJE and MTM metrics.

Another noteworthy observation is that both EVIC and NNPC, while capable, have difficulties with fine-motor adjustments. Throughout the trials, participants occasionally overshot or undershot their desired position, and had to make fine motor adjustments to achieve the desired position. An example of undershooting is shown in Fig. 6.3. The dyad is able to complete 90% of

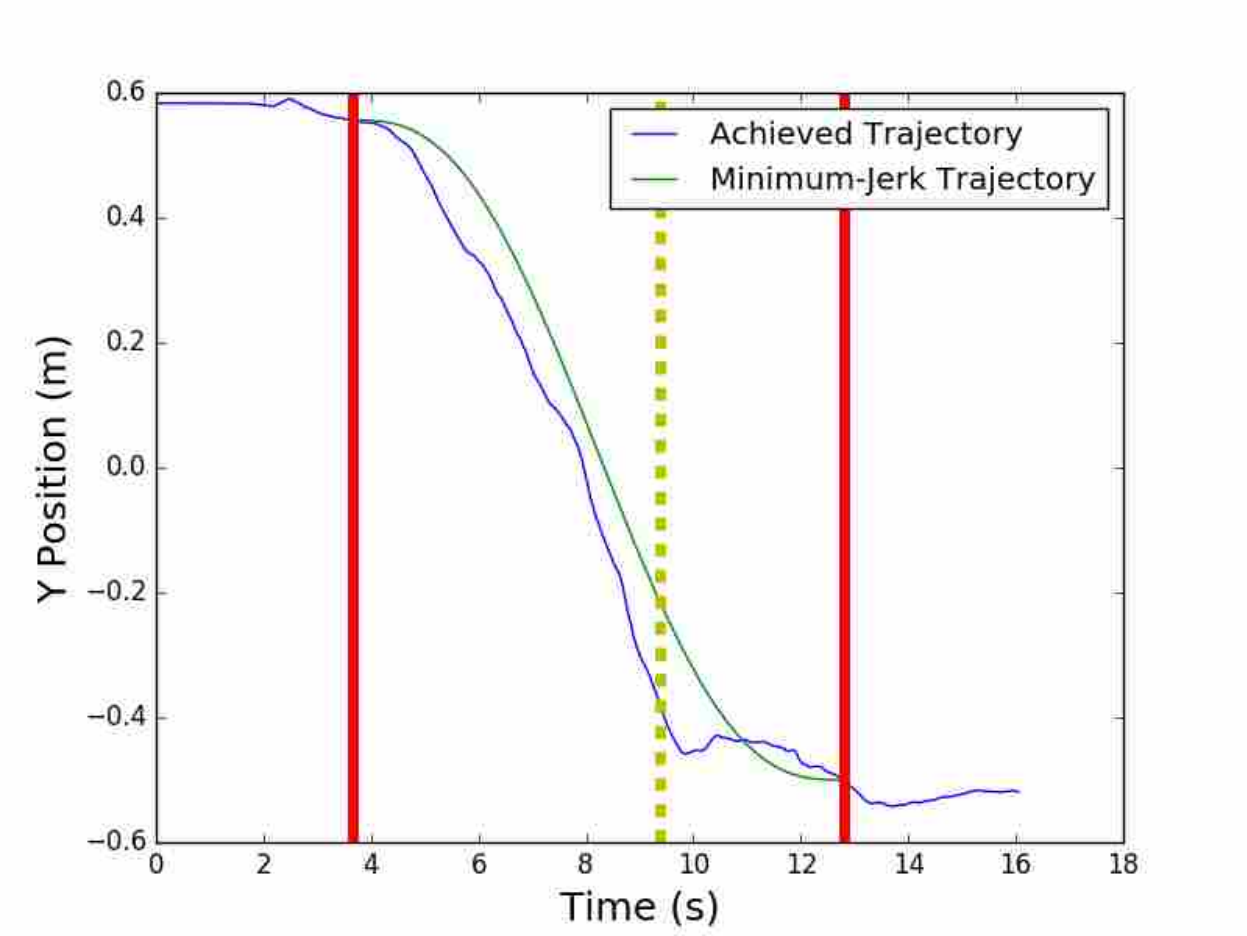


Figure 6.3: Undershooting behavior of a human robot dyad for a translation task, where bold, vertical lines indicate start/stop points, and dashed vertical line indicates 90% completion point– movement after this point is considered a fine motor adjustment.

the task, represented by the dashed vertical line, in just under 6 seconds, but spends approximately 3 seconds trying to complete the remaining 10%, which amounts to about 10 cm of movement, with more fine adjustments.

On average, the remaining time at 90% completion was 2.40s for EVIC and 2.55s for NNPC. We also compared the achieved trajectories with an ideal minimum-jerk trajectory. The time to 90% task completion on a minimum-jerk trajectory was, on average, 2.20s for EVIC and 2.12s for NNPC. From this data, it would appear that EVIC is slightly better at fine-motor adjustments than NNPC, since EVIC had a smaller discrepancy between achieved and minimum-jerk 90% completion time.

To determine if a few under-performing dyads skewed the average, we also took the median 90% completion time. For achieved and minimum-jerk trajectories, respectively, with EVIC, this gave values of 1.98s and 2.11s. Similarly for NNPC, it gave values of 1.97s and 2.02s. From these results, we conclude that this data is positive skewed, and only a small number of dyads had trouble with fine-motor adjustments, causing the higher mean values. Therefore, we can conclude the lack of fine motor skills in the controllers did not significantly hamper their ability to complete the tasks, but should be addressed in future work to help improve the performance of those dyads who struggled with undershooting or overshooting.

6.2.3 Qualitative Results

In addition to the metric performance, we also wanted to know what each person thought of the controllers subjectively. As mentioned in Chapter 6.1.3, we gave each participant a survey after they had performed all the tasks with one controller, and then the same survey after they finished the other controller. The questions we asked dealt with how they thought their partner, a robot in this case, performed. They were asked to rate the categories from 1–Strongly Disagree–to 5–Strongly Agree. The questions were:

- My partner was helpful
- My partner moved quick enough
- My partner moved too slow
- There was confusion between me and my partner
- I trusted my partner to perform the task correctly
- I felt safe completing the task
- I trusted my partner to move at correct speeds
- I trusted my partner to move in the correct direction
- My partner did not push or pull too hard

- My partner moved predictably
- My partner helped me to do task better than I could alone
- My partner shared the task equally

The average for each controller rating is given in the first 2 columns of Table 6.3. The controller that performed better in each category is in bold. For some categories, like *Too Slow*, a lower number is desired, whereas for others, like *Safe*, a higher number is desired. The same survey, except for the *Correct Direction* question, was given after the HHI study, and the results are shown in the third column of Table 6.3. We have included only the responses of the human designated as the leader from the human-human dyads.

Similar to our analysis of the quantitative results, we determined the statistical significance of the survey results, in order to determine if there was a statistical difference between the responses about the EVIC and NNPC controllers. For each question, we ran an unpaired t-test to calculate a p value and determined Cohen's d to find the effect size. These statistics are found in Table 6.4. What we found is that only the *Good Force Amount* question obtained a p value of less than 0.05, indicating it was the lone statistically significant answer. However, this question, as well as a number of others, had a medium effect size.

6.2.4 Discussion

The first thing we noticed from Table 6.3 is that people still clearly prefer working with a human partner over a robot partner. One reason for this may be that humans do not trust robots entirely, as is evidenced by the 5th, 7th, and 8th questions in the survey, which all ask about trust in the partner. Perhaps the same pHRI experiment, done instead with a blindfold and earmuffs on the human would have returned more favorable ratings for the robot controllers. This is something that needs to be explored further, as our experiments do not definitely prove that humans do not trust robots.

As was mentioned above, NNPC was the more capable controller in terms of the performance metrics. It also was preferred in the survey questions that dealt with these metrics. NNPC was superior in rating for the *Fast Enough, Too Slow, Correct Speed, Correct Direction,* and *Better*

Table 6.3: Ratings of survey questions, with 5 as strongly agree and 1 as strongly disagree. Bold numbers indicate preference between EVIC and NNPC for the specified category.

	EVIC	NNPC	HHI
Helpful	**3.88**	**3.88**	4.52
Fast Enough	3.38	**3.63**	4.38
Too Slow	3.31	**2.94**	2.09
Confusing	**3.38**	3.5	2.09
Correct Task	**3.94**	3.75	4.42
Safe	**4.5**	4.44	4.71
Correct Speed	3.44	**3.56**	4.47
Correct Direction	3.5	**3.56**	N/A
Good Force Amount	**3.44**	2.81	4.47
Predictable	**3.63**	3.5	4.28
Better than Alone	3.5	**3.56**	4.38
Equal Share	**3.75**	3.5	4.23

Table 6.4: Statistical significance of qualitative metrics.

	EVIC vs NNPC	
Survey Question	p	Cohen's d Effect Size
Helpful	0.5	Very Small
Fast Enough	0.19	Medium
Too Slow	0.15	Medium
Confusing	0.34	Small
Correct Task	0.28	Medium
Safe	0.36	Small
Correct Speed	0.37	Small
Correct Direction	0.41	Small
Good Force Amount	0.04	Medium
Predictable	0.35	Small
Better than Alone	0.38	Small
Equal Share	0.15	Medium

than Alone metrics. The superior rating in these metrics corroborates what we saw quantitatively, indicating that NNPC was a more efficient controller, at least according to the metrics we have seen.

However, participants also indicated they preferred using EVIC. They thought it was more helpful, applied more appropriate levels of force, and saw it as the more predictable controller. While NNPC outperformed EVIC in the TMT metric, this may not be a good thing. From our observations here, as well as in our previous experiment [1], haptic communication is a large part of how humans perform co-manipulation tasks. NNPC users experienced less force overall, based on the TMT metric, but the survey indicated that EVIC applied more appropriate forces. Additionally, EVIC and NNPC were only statistically different in terms of the MTM metric and the *Good Force Amount* question. From these results, we can conclude that NNPC is not applying appropriate forces, and is therefore considered more difficult and less intuitive to use by the participants. In fact, our results here agree with what was said by Chipalkatty et al. [51], who indicated that training a controller to be the most efficient or best performing controller may cause it to be a less preferable controller to humans. Because EVIC and NNPC were statistically different from both blind and sighted human-human dyads in terms of MJE and MTM, we can say that minimizing forces and torques, and also deviation from a minimum-jerk trajectory, may not be a satisfactory goal of co-manipulation controllers. So while NNPC is the better performing controller, EVIC might be a more intuitive and appropriate controller for real-world applications, since it applies more appropriate forces.

Looking at completion time in Table 6.2, we can see that both EVIC and NNPC were not statistically distinguishable from the blind human-human dyads. This means that according to this metric, both controllers perform up to the standard of the blind dyads. While this is an encouraging result, We know that there is some missing information in our model. Although similar in completion time, our controllers performed much different than both types of human-human teams in the TMT and MJE metrics. Perhaps there is a complex cost function that humans are using to determine how much weight to put on adhering to minimum-jerk trajectories, minimizing torque change, and moving in a timely manner. These considerations should be made in future controller development.

We have also seen (see Fig. 6.3) that these controllers are not well suited for fine-motor adjustments. We believe that because our robot arms were not helping to reach position, but were simply holding an equilibrium position, the dyad was not able to account for small errors in the position of the board. An algorithm that is able to coordinate arm and base motion would be more suited to fix these small errors, and could possibly bring our pHRI controllers for co-manipulation closer to human-human sighted performance.

CHAPTER 7. CONCLUSION

7.1 Summary of Results

This thesis has presented observations and analysis of metrics from a 6 DOF co-manipulation human-human interaction experiment. From these observations, we saw that, on average, humans tend to move in minimum-jerk trajectories, even in co-manipulation situations. An exception to the minimum-jerk trajectory relationships is in prolonged or indefinite trajectories, where a constant velocity profile is more appropriate. We also saw that interaction forces and torques can be used to communicate intent from leader to follower. These observations were used to define an extension to Variable Impedance Control called Extended Variable Impedance Control. This controller is a planar co-manipulation method, where the dyad is in leader-follower setup, with the robot as the follower.

Additionally, another controller, known as Neural Network Prediction Control, was defined. This controller takes previous motion of the shared object and predicts future motion, and commands the robot accordingly. This method is also in leader-follower format. The advantage of this method is that it is capable of extension into 6 DOF, although was only implemented in planar 3 DOF in this thesis. These two controllers are, to our knowledge, two of the only bi-manual, planar, extended object co-manipulation controllers. One of the only other control methods we have seen for similar situations was shown in [30], and we discussed its limitations in Chapter 2.

An experiment of human-robot dyad co-manipulation of extended objects was performed, using both of the controllers described in this thesis. The dyads performed lateral translation and planar rotation tasks 6 times for each controller. The direction of motion (left/right and clockwise/counter-clockwise) was alternated for each task. Participants did not know the specific details of the controllers, but were trained on each to a level of competency before data was recorded. The experiment showed a number of results. Both control methods performed at a level comparable to humans in task completion time. However, they both performed at a higher level

than humans in both minimum-jerk trajectory error, as well as minimum-torque measurement. A survey was also given to participants where they qualitatively rated the controllers. While the Neural Network Prediction Control performed better–according to the metrics used–participants indicated their preferred method was the Extended Variable Impedance Control, claiming it applied more appropriate levels of force, was more intuitive, and was more helpful.

7.2 Limitations and Possible Extensions

A big limitation of the Extended Variable Impedance Control is its restriction to planar motion. Most real-world tasks that we envision for human-robot dyads will require at least two more DOF–z-axis translation (lifting) and x-axis rotation. It is not clear what force/torque patterns contribute to these two movement types, and how they would be distinguished from the other patterns described in this thesis. We recommend that future work explore the human-human physical comanipulation data tasks involving lifting and rotation about the x-axis. Lifting motion may also be attainable by using the variable impedance model, although this presents problems with translation versus rotation about the y-axis.

Another limitation, involving both control methods, is the disconnect between base motion and arm motion. In watching videos of the human-human dyads, it is clear that coordinated motion between legs and arms is essential for better performance of the tasks, and is a necessity for complex motions such as rotating about various axes while translating. Both controllers were capable of performing slightly above the level of blind-follower human-human dyads, but are below sighted-follower dyads. In order to bridge this gap, they will need more coordinated control of arms and base. A logical first step is to incorporate the z-axis translation prediction into the Neural Network Prediction Control. The impedance controller being run on the arms to mimic human behavior could be kept for x and y motion, but z-axis translation could be commanded to increase this controller to 4 DOF motion. An experiment could then be run to determine the performance of human-robot dyad lifting capabilities.

We also recommend that more work be done in analyzing the human-human physical comanipulation data for other possible physical human-robot interaction theories. For example, it is possible that humans have learned to ignore certain forces or torques when performing certain motions. The only tasks that have been thoroughly analyzed are the lateral translation and planar

rotation tasks. In order to have a better picture of how humans perform general physical co-manipulation tasks, and how best to program a robot partner, it is necessary to explore the other tasks from our data set described in Chapter 3. While the Neural Network Prediction Control does this in part, it still does not fully incorporate nor interpret the applied forces and torques. A possible reason that force data caused divergence, that was not explored in this work, is that the force data was filtered. The filtering causes a phase shift that may cause some issues when it is implemented in the neural network. Looking into the phase shift problem, or exploring other architectures–such as an RNN classifier where the classes are constant velocity or acceleration values–should be pursued. Additionally, unexplored alternatives to the neural network, such as autoregressive–moving-average models, could be used to incorporate force into the prediction.

7.3 Closing Remarks

This thesis has provided two unique physical co-manipulation controllers that can be applied to human-robot dyads. We have observed that these controllers are capable of operating to an acceptable level with human partners, in 3 DOF. Human intent has also been characterized in terms of both force patterns and motion patterns. While there are limitations to these controllers, they are significant steps toward having 6 DOF human-robot physical co-manipulation in unstructured environments.

REFERENCES

[1] Mielke, E. A., Townsend, E. C., and Killpack, M. D., 2017. "Analysis of Rigid Extended Object Co-Manipulation by Human Dyads: Lateral Movement Characterization.". vi, 17, 21, 52, 54, 56, 57, 63

[2] Townsend, E. C., 2017. "Estimating Short-Term Human Intent for Physical Human-Robot Co-Manipulation." Master's thesis, Brigham Young University. viii, 5, 17, 45, 47, 48

[3] Duchaine, V., and Gosselin, C. M., 2007. "General model of human-robot cooperation using a novel velocity based variable impedance control." *Proceedings - Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics 2007*, pp. 445–451. 2, 9, 14, 25, 35

[4] TSUMUGIWA, T., YOKOGAWA, R., and HAM, K., 2002. "Variable Impedance Control Based on Estimation of Human Arm Stiffness for Human-Robot Cooperative Calligraphic Task." pp. 644–650. 2, 9, 15

[5] Ganesh, G., Takagi, a., Osu, R., Yoshioka, T., Kawato, M., and Burdet, E., 2014. "Two is better than one: physical interactions improve motor performance in humans.." *Scientific reports,* **4**, p. 3824. 2, 8

[6] Ivaldi, S., Sigaud, O., Berret, B., and Nori, F., 2012. "From Humans to Humanoids: the Optimal Control Framework." *Paladyn, Journal of Behavioral Robotics,* **3**(2). 3, 14, 28, 29, 55

[7] Groten, R. K., 2011. "Haptic Human-Robot Collaboration : How to Learn from Human Dyads.". 3, 14, 20, 29

[8] Nepalese soldiers clear the debris after an earthquake at the kaalmochan temple in kathmandu, nepal, may 2, 2015. [Online; accessed March 20, 2018]. 3

[9] Mh-60s seahawk (hc-5) - banda aceh, indonesia - february 2005 [Online; accessed March 20, 2018]. 3

[10] Flash, T., and Hogan, N., 1985. "The coordination of arm movements: an experimentally confirmed mathematical model.." *The Journal of neuroscience,* **5**(7), pp. 1688–1703. 7

[11] Rahman, M. M., Ikeura, R., and Mizutani, K., 2000. "Control characteristics of two humans in cooperative task and its application to robot Control." *IECON Proceedings (Industrial Electronics Conference),* **1**, pp. 1773–1778. 7

[12] Reed, K. B., Patton, J., and Peshkin, M., 2007. "Replicating Human-Human Physical Interaction." pp. 10–14. 8, 15

[13] Wel, R. P. R. D. V. D., Knoblich, G., and Sebanz, N., 2011. "Let the Force Be With Us : Dyads Exploit Haptic Coupling for Coordination." pp. 1420–1431. 8, 15

[14] Bussy, A., Kheddar, A., Crosnier, A., and Keith, F., 2012. "Human-humanoid haptic joint object transportation case study." *IEEE International Conference on Intelligent Robots and Systems*, pp. 3633–3638. 8, 12, 16, 27, 33

[15] Sawers, A., Bhattacharjee, T., McKay, J. L., Hackney, M. E., Kemp, C. C., and Ting, L. H., 2017. "Small forces that differ with prior motor experience can communicate movement goals during human-human physical interaction." *Journal of NeuroEngineering and Rehabilitation,* **14**(1), p. 8. 8, 15

[16] Mojtahedi, K., Whitsell, B., Artemiadis, P., Santello, M., Gregg, R., Killpack, M. D., and Santello, M., 2017. "Communication and Inference of Intended Movement Direction during HumanHuman Physical Interaction." pp. 1–12. 8, 15

[17] Groten, R., Feth, D., Goshy, H., Peer, A., Kenny, D. A., and Buss, M., 2009. "Experimental analysis of dominance in haptic collaboration." *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 723–729. 9

[18] Evrard, P., and Kheddar, A., 2009. "Homotopy switching model for dyad haptic interaction in physical collaborative tasks." *Proceedings - 3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics 2009*, pp. 45–50. 9

[19] Tatti, F., and Baud-Bovy, G., 2015. "Force sharing strategies in a collaborative force detection task." *IEEE World Haptics Conference, WHC 2015*, pp. 463–468. 9

[20] Mörtl, A., Lawitzky, M., Kucukyilmaz, a., Sezgin, M., Basdogan, C., and Hirche, S., 2012. "The role of roles: Physical cooperation between humans and robots." *The International Journal of Robotics Research,* **31**(13), pp. 1656–1674. 9

[21] Ikeura, R., Monden, H., and Inooka, H., 2002. "Cooperative Motion Control of a Robot and a Human." pp. 2–3. 9

[22] Rahman, M. M., Ikeura, R., and Mizutani, K., 2002. "Impedance characteristic of human arm for cooperative robot." pp. 1455–1460. 9, 15

[23] Ikeura, R., and Inooka, H., 1995. "Variable Impedance Control of a Robot for Cooperation with a Human." pp. 3097–3102. 9, 14, 25

[24] Dimeas, F., and Aspragathos, N., 2015. "Reinforcement learning of variable admittance control for human-robot co-manipulation." *IEEE International Conference on Intelligent Robots and Systems,* **2015-Decem**, pp. 1011–1016. 9

[25] Ficuciello, F., Villani, L., and Siciliano, B., 2015. "Variable Impedance Control of Redundant Manipulators for Intuitive Human-Robot Physical Interaction." *IEEE Transactions on Robotics,* **31**(4), pp. 850–863. 9

[26] Ranatunga, I., Lewis, F., Popa, D. O., and Tousif, S. M., 2016. "Adaptive Admittance Control for Human-Robot Interaction Using Model Reference Design and Adaptive Inverse Filtering." *IEEE Trans. Control Systems Technology,* **00**(0), pp. 1–10. 10

[27] Leica, P., Toibero, J. M., Roberti, F., and Carelli, R., 2013. "Bilateral human-robot interaction with physical contact." *2013 16th International Conference on Advanced Robotics, ICAR 2013*. 10

[28] Karayiannidis, Y., Smith, C., and Kragic, D., 2014. "Mapping human intentions to robot motions via physical interaction through a jointly-held object." *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication,* **2014-Octob**(October), pp. 391–397. 10

[29] Nguyen, V. Q., 2016. "Intent Recognition Of Rotation Versus Translation Movements In Human-Robot Collaborative Manipulation Tasks.". 10

[30] Bussy, A., Gergondet, P., Kheddar, A., Keith, F., and Crosnier, A., 2012. "Proactive behavior of a humanoid robot in a haptic transportation task with a human partner." *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*(2), pp. 962–967. 11, 13, 65

[31] Peternel, L., Tsagarakis, N., and Ajoudani, A., 2017. "A human-robot co-manipulation approach based on human sensorimotor information." *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* **25**(7), pp. 811–822. 11

[32] Peternel, L., Tsagarakis, N., Caldwell, D., and Ajoudani, A., 2017. "Robot adaptation to human physical fatigue in humanrobot co-manipulation." *Autonomous Robots*, pp. 1–11. 12

[33] Corteville, B., Aertbelien, E., Bruyninckx, H., De Schutter, J., and Van Brussel, H., 2007. "Human-inspired robot assistant for fast point-to-point movements." *Proceedings - IEEE International Conference on Robotics and Automation*(April), pp. 3639–3644. 12, 14, 15

[34] Maeda, Y., Hara, T., and Arai, T., 2001. "Human-robot cooperative manipulation with motion estimation." *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No.01CH37180),* **4**, pp. 2240–2245. 12

[35] Thobbi, A., Gu, Y., and Sheng, W., 2011. "Using human motion estimation for human-robot cooperative manipulation." *IEEE International Conference on Intelligent Robots and Systems*, pp. 2873–2878. 12, 13, 14, 16, 27, 29

[36] Miossec, S., and Kheddar, A., 2008. "Human motion in cooperative tasks: Moving object case study." *2008 IEEE International Conference on Robotics and Biomimetics, ROBIO 2008*(i), pp. 1509–1514. 12, 14, 29

[37] Ge, S. S., Li, Y., and He, H., 2011. "Neural-network-based human intention estimation for physical human-robot interaction." *URAI 2011 - 2011 8th International Conference on Ubiquitous Robots and Ambient Intelligence*, pp. 390–395. 13, 16

[38] Berger, E., Vogt, D., Haji-Ghassemi, N., Jung, B., and Amor, H. B., 2015. "Inferring guidance information in cooperative human-robot tasks." *IEEE-RAS International Conference on Humanoid Robots,* **2015-Febru**(February), pp. 124–129. 13

[39] Medina, J. R., Lorenz, T., and Hirche, S., 2015. "Synthesizing anticipatory haptic assistance considering human behavior uncertainty." *IEEE Transactions on Robotics,* **31**(1), pp. 180–190. 13

[40] Basdogan, C., Ho, C.-h., Srinivasan, M. A., and Slater, M. E. L., 2001. "An Experimental Study on the Role of Touch in Shared Virtual Environments." pp. 443–460. 15

[41] Noohi, E., Zefran, M., and Patton, J. L., 2016. "A Model for Human-Human Collaborative Object Manipulation and Its Application to Human-Robot Interaction." *IEEE Transactions on Robotics,* **32**(4), pp. 880–896. 15, 24

[42] Groten, R., Feth, D., Klatzky, R. L., and Peer, A., 2013. "The role of haptic feedback for the integration of intentions in shared task execution." *IEEE Transactions on Haptics,* **6**(1), pp. 94–105. 15

[43] Ikeura, R., Morita, A., and Mizutani, K., 1997. "Variable damping characteristics in carrying an object by two humans." *Proceedings 6th IEEE International Workshop on Robot and Human Communication. RO-MAN'97 SENDAI*, pp. 130–134. 15

[44] Madan, C. E., Kucukyilmaz, A., Sezgin, T. M., and Basdogan, C., 2015. "Recognition of Haptic Interaction Patterns in Dyadic Joint Object Manipulation." *IEEE Transactions on Haptics,* **8**(1), pp. 54–66. 15

[45] Lawitzky, M., Medina, J. R., Lee, D., and Hirche, S., 2012. "Feedback motion planning and learning from demonstration in physical robotic assistance: Differences and synergies." *IEEE International Conference on Intelligent Robots and Systems*, pp. 3646–3652. 15

[46] Rozo, L., Calinon, S., Caldwell, D. G., Jimenez, P., and Torras, C., 2016. "Learning Physical Collaborative Robot Behaviors From Human Demonstrations." *IEEE Transactions on Robotics,* **32**(3), pp. 513–527. 16

[47] Etienne Burdet, D. W. F., and Milner, T. E., 2013. *Human Robotics: Neuromechanics and Motor Control*. Massachusetts Institute of Technology, 55 Hayward Street, Cambridge, MA 02142. 25, 40

[48] Yan, L., and Li, C. J., 1997. "Robot learning control based on recurrent neural network inverse model." *Journal of Robotic Systems,* **14**(3), pp. 199–212. 43

[49] Wu, Y., Song, Q., and Yang, X., 2007. "Robust Recurrent Neural Network Control of Biped Robot." *Journal of Intelligent and Robotic Systems,* **49**(2), pp. 151–169. 43

[50] Martens, J., 2011. "Generating Text with Recurrent Neural Networks." *Neural Networks,* **131**(1), pp. 1017–1024. 43

[51] Chipalkatty, R., and Droge, G., 2013. "Less Is More : Mixed Initiative Model Predictive Control With Human Inputs." *IEEE Transactions on Robotics,* **29**(3), pp. 1–9. 46, 63

[52] Engel, Y., Mannor, S., and Meir, R., 2004. "The kernel recursive least squares algorithm." *IEEE Transactions on Signal Processing,* **52**(8), pp. 2275–2285. 46

[53] Mielke, E. A., Townsend, E. C., and Killpack, M. D. "Planar Human-robot Co-manipulation of Extended Objects : Data-driven Models and Control from Analysis of Human-human Dyads." *Submitted to Springer Autonomous Robotics*. 52

[54] Sawilowsky, S. S., 2009. "New Effect Size Rules of Thumb." *Journal of Modern Applied Statistical Methods,* **8**(2), pp. 597–599. 58