Personalized Modeling for Human-Robot Collaborative Manipulation

Aaron Bestick

Electrical Engineering and Computer Sciences
University of California at Berkeley

December 1, 2019

Personalized Modeling for Human-Robot Collaborative Manipulation

by

Aaron Matthew Bestick

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ruzena Bajcsy, Chair
Professor Ronald Fearing
Professor Masayoshi Tomizuka

Fall 2017

Personalized Modeling for Human-Robot Collaborative Manipulation

Abstract

Personalized Modeling for Human-Robot Collaborative Manipulation

by

Aaron Matthew Bestick

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ruzena Bajcsy, Chair

When two humans perform a collaborative manipulation task, they leverage an intuitive understanding of which motions are natural and safe for their interaction partner. Intuition lets human collaborators predict both the *feasibility* of an action, as well as their partner's *ergonomic preference* for one feasible action over another. This mutual understanding in human-human teams allows for comfortable and efficient collaboration. However, in human-robot teams, robots typically lack the models which would give them this same understanding of human action choice. This problem is compounded by that fact that humans are unique. A model that accurately predicts the actions of one human collaborator may be incorrect for another.

This dissertation explores how personalized mechanical models and ergonomic cost functions for humans can endow collaborative robots with an understanding of human action feasibility and ergonomic preference analogous to that possessed by human teammates. Specifically, we focus on the task of robot-to-human object handoffs. We show that planning handoffs with knowledge of a human's arm kinematics yields improved ergonomics and safety, and is preferred by human collaborators. Next, we demonstrate how ergonomic cost functions which predict a human's preference for different feasible actions can allow co-robots to shape human action choices, helping humans choose globally ergonomically optimal actions in multi-part tasks. Finally, we show how these ergonomic cost functions can be learned online, allowing co-robots to quickly adapt to the preferences of an individual human partner.

To my parents,
who gave me the freedom and encouragement to find my own passions,
and taught me to never stop learning

Contents

List of Figures

List of Tables

Acknowledgments

Any project that takes six years is bound to be the work of many people, and this PhD dissertation is no different.

I first have to thank my advisor, Ruzena Bajcsy, for giving me the freedom to explore ideas and find the ones that interested me, providing just the right amount of structure to keep me on the right trajectory, and for being an advocate, confidant, and even more than that, a constant personal inspiration through all the inevitable ups and downs that come with grad school.

I also want to thank Anca Dragan, who became a second advisor of sorts after she arrived at Berkeley midway through my PhD as a new faculty member. Along with being a consistently positive and supportive friend, Anca was exactly the mentor and collaborator I needed to help focus my work and drive it toward completion. I'm skeptical I would have made it to this point without her help.

Moving back a few years more to when I first arrived at Berkeley, I have to thank Sam Burden and Lillian Ratliff. It's important to have role models in any new environment, and Sam and Lily were exactly the type of scientists and engineers I aspired to be when I made it to this point in my grad school career. They both generously volunteered their time to help me get started on research that I, frankly, hadn't the faintest idea how to approach, and they set me on the trajectory that eventually led here.

I was lucky to have the other members of Ruzena's awesome HART Lab to lean on too, particularly Katie Driggs-Campbell, my friend and cube-mate who provided moral support while putting up with my constant distractions, Rob Matthew, Daniel Aranki, and Gregorij Kurillo, for always being willing to offer consultation and advice, and Ravi Pandya, for being a fantastic research collaborator. My good friends Austin Buchan, Roy Dong, Dan Calderone, and others were also an essential source of balance, and helped me maintain my sanity when when things got frustrating.

Beyond just my lab, Berkeley EECS has an exceptionally collaborative, open, and socially conscientious culture. My forays into the real world have only highlighted how unique and special a place Berkeley is. It truly is an environment that helps people to do their best work, and I was very fortunate to benefit from it. The administrative staff in EECS are a key part of this culture too, particularly Shirley Salanio and Jessica Gamble, who are a big factor in many students' surviving the Berkeley PhD program with their mental health intact.

Chapter 1

Introduction

1.1 Background

Robots have made impressive progress toward automating many types of physical labor traditionally performed by humans. Speculation about the effect on the labor market and our broader society of human workers displaced en masse by automation have been impossible to avoid in recent years in both academia [1–3] and the popular press [4, 5].

This concern has been particularly acute for manufacturing and service jobs which presently rely on human workers to assemble products, move and organize items, and perform other physical labor [1–3]. This concern is valid, timely, and important, but it can overshadow another important fact: many physical tasks that on the surface seem ripe for automation have elements which still require human perception and decision making, and will continue to in the future, as far as we can predict.

This dissertation will address two broad classes of problems which arise in systems where humans and robots physically collaborate:

Problem 1: Humans are Different from One Another

Individual humans have different physical abilities and preferences due to natural variation, age, disability, disease and other factors. A person with a broken arm immobilized in a sling will have a different reachable workspace than an "average" person (Chapter 2). Two people with similar body dimensions and capabilities may choose different grasp configurations when a robot hands them identical objects at identical handoff poses (Chapters 3 and 4).

This variation between individuals can limit the utility of generic, one-size-fits-all human models learned across entire human populations (Chapter 2). It requires adjusting our human models to specific, individual human collaborators, preferably online during an interaction (Chapter 4).

Problem 2: Human Actions are Nondeterministic

Even when dealing with a single human collaborator, we can't assume that the human will reliably choose the same action when presented with the same situation repeatedly. For example, when a robot hands a human a coffee mug at the same pose multiple times, the human will exhibit a strong preference for a small subset of the feasible grasp configurations on the mug. This preference is noisy though: the chosen configuration will always vary slightly from the last, and may be significantly different once in a while (e.g. grasping the top of the mug rather than its side).

Modeling humans as deterministic can produce suboptimal collaborative behavior when human action choices vary from the prediction. Using probabilistic models of human action choice allows us to model this noise explicitly and ensure optimal interactions even in the face of human nondeterminism (Chapters 3 and 4).

Automation-Resistant Tasks

To motivate the idea of physical human-robot collaboration, let's examine several example tasks that are in many ways ideal for automation, but are nonetheless still done by human workers. What are the aspects of these tasks which make them difficult to fully automate? We can observe two broad themes:

- One-off or ill-posed task specifications

- Perceptual or modeling uncertainty in environment or task

Consider a few selected examples:

Aircraft Interior Assembly

Ductwork, luggage bins, interior panels, and wiring are installed completely by hand, despite heavy weight and awkward bending. Sometimes multiple people are needed to install a single large assembly. Why? Installation often requires precise alignment of semi-flexible parts (ducts, interior panels) which are then secured with a variety of fasteners (cable ties, screws, bolts, electrical connectors). The precise dimensions of the parts vary throughout the aircraft, and fit and finish are often not precisely controlled, requiring human intervention to align screw holes, bend panels to achieve an acceptable fit, etc.

Building Construction

Drywall, plywood, and other building materials are measured, cut, lifted into position, and secured by teams of human workers. Simple mechanical fixtures and hoists are used for selected pieces (ceilings), but most work is performed by hand, with multiple people climbing on ladders while maneuvering heavy and unwieldy pieces of material. Why? Similar to airplane assembly, installation requires constant fine alignment, cutting of material to fit

Figure 1.1: Automotive final assembly still relies largely on human workers, in contrast to the body assembly and paint portions of the assembly process. Flexible materials like wiring, upholstery, and plastic interior panels are difficult for robots to manipulate without elaborate fixtures, and human dexterity is required to feed wires and align parts. Simple gravity compensation tooling (as shown) is sometimes used to lift heavy pieces, but provides little additional assistance to the human workers. (Photo credit: Ford Motor Company, https://flic.kr/p/5sqHgk)

around obstacles like light fixtures and electrical sockets, and precise maneuvering to fit pieces together with an acceptable final appearance.

Automobile Final Assembly

In contrast to the highly automated body assembly and paint portions of a car manufacturing line, final assembly still largely requires humans to do the work (as shown in Figure 1.1) [6,7]. Simple mechanical fixtures help to manipulate unwieldy items like seats and engines, but humans still climb in and out of vehicles and bend in awkward positions to install wiring, dashboards, interior paneling, and other components by hand. Why? As before, the materials are too flexible to be easily modeled, the assembly processes require dexterity to feed wires and align parts, which precludes large, rigid fixturing.

Furniture Moving

Human movers lift heavy and unergonomic furniture and items manually, despite the risks of back and limb injury. Why? The task specification, where the human wants the furniture placed, is unclear, and communicating it to the robot explicitly would be tedious and time-consuming. In addition, its impractical for a robot to have the detailed physical models of objects and environment that would be needed to manipulate furniture precisely without damaging collisions.

Patient Transfer in Medical Facilities

Nurses and caregivers frequently slide, lift, and turn patients with limited mobility to help them in and out of bed or reposition them to avoid sores. The caregivers frequently work alone, the patients they move are heavy, and the location (often in a bed) precludes good body mechanics for lifting, pushing and pulling. Why? Again, the task specification is ill-defined and difficult to communicate to a robot explicitly, and motions that may be feasible for one patient may cause pain or injury for another.

Interestingly, all of these examples have attributes which make automated assistance desirable. They require heavy lifting, awkward bending, and repetitive motions. This combination makes injuries particularly prevalent, as shown in Table 1.1

Table 1.1: Workplace injury rates for selected automation-resistant tasks in 2015. All the example tasks discussed above have injury rates equal to or greater than the average across all occupations. (Data from the US Bureau of Labor Statistics [8]). [*]Injury rate is the number of reportable injuries or illnesses per 100 full-time equivalent workers.

Occupation	Injury Rate[*]
All Occupations	*3.3*
Aircraft Manufacturing	3.3
Residential Building Construction	3.9
Motor Vehicle Manufacturing	6.6
Household and Office Goods Moving	5.6
Nursing and Residential Care Facilities	6.8

Its easy to imagine a robotic assistant which could mitigate many of these risks. The assistant would handle the heavy lifting and rough alignment portions of a task, while leaving the reasoning and precise manipulation to a human collaborator.

Communication in Collaborative Tasks

Human teams collaborate in similar ways all the time when they lift a heavy object together or manipulate furniture through a crowded apartment. But what types of communication or modeling are required for a robot to replicate a simple version of this task. As in many human-robot interaction scenarios, its instructive to examine an equivalent human-human interaction in more detail. Consider two humans manipulating a piece of furniture (Figure 1.2). They may use:

- Explicit verbal communication: talking to each other to coordinate actions

- Implicit visual communication: eye contact and gaze direction help to indicate understanding or motion intent

- Interaction forces: the forces transmitted though the object itself signal each human's motion intent

Figure 1.2: Humans rely on a variety of communication modes when performing collaborative manipulation tasks. These include explicit verbal communication, implicit visual signaling, interaction forces, and intuition about their interaction partners' ergonomic preferences and safety constraints. This thesis considers the last of these modes, and investigates how personalized human mechanical models and cost functions can approximate this human intuition for specific tasks. (Photo credit: https://flic.kr/p/pJFYxR)

- **Intuition about human ergonomic costs and safety constraints**: humans intuitively understand that some motions are dangerous or unergonomic for their collaborators

While all of these modes are potentially useful and relevant in human-robot collaborative manipulation, this dissertation will focus on the last one. Specifically, we'll investigate how personalized human mechanical models and cost functions can approximate this intuition for the specific task of object handovers. We'd like our models to capture the limits on a specific human collaborator's motion imposed by safety and joint ranges of motion, as well as the person's ergonomic preference for some action choices over others (as mentioned in Problem 1: Humans are Different from One Another). We'll investigate how co-robots can use such models to generate safe and comfortable collaborative actions and even influence humans' behavior to yield a more ergonomically optimal outcome in situations where humans are noisy and short-sighted (Problem 2: Human Actions are Nondeterministic).

Comparison With Other Machine Learning Approaches

We choose to model humans using highly structured mechanical models, but one could also apply more general machine learning methods. Recent approaches from the literature follow the imitation learning or inverse reinforcement learning paradigms, and represent variability in human mechanical structure and motion preferences as deep neural networks [9,10], mixtures of multivariate Gaussians [11,12], or cost functions with more general, non-mechanically derived parameterizations. Given these options, why would one choose to use structured mechanical models?
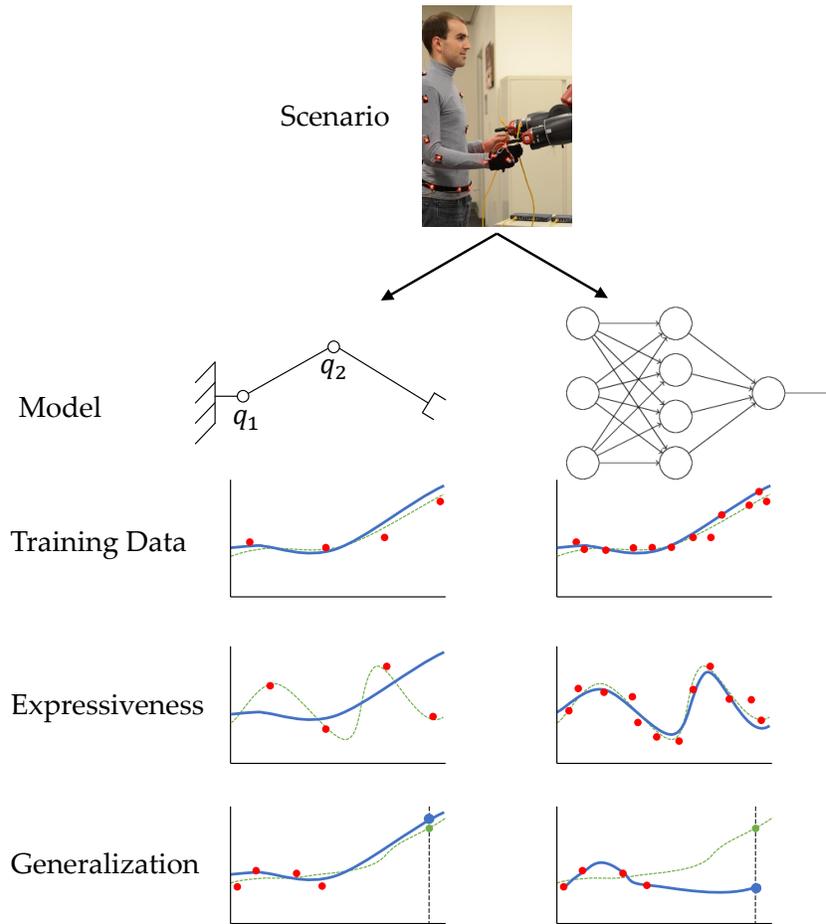
Figure 1.3: Conceptual comparison of structured mechanical models with more generic learned models (e.g. neural networks). Structured models have far fewer model degrees of freedom, and thus require much less training data. Their generalization to unseen scenarios can also be superior if the model class of the structured model fits the model class of the actual system closely. However, if the true class of the task model is unknown, it can be wise to use a more expressive, less constrained model. For the human-robot collaborative tasks we consider in this dissertation, rigid kinematic models are a reasonable approximation of the human skeleton, making them a good choice when compared with other models which are more complex and impose less *a priori* structure. (In the plots above, green represents the true model, blue the learned model, and red the training data).

Human motion is governed by a person's mechanical structure in much the same way as a robot's motion. By pre-specifying a model class with less degrees of freedom that "bakes in" some of this mechanical structure (e.g. a kinematic model parameterized by twists), we can obtain a personalized model that:

1. Requires less data to train (Chapter 2)

2. Generalizes more accurately to scenarios not seen previously (Chapters 2 and 4)

3. Lets us directly apply mathematical tools from robotics for kinematic/dynamic analysis and planning (as in [13], used in all subsequent chapters)

This assumes, of course, that the actual human-robot system conforms to the assumptions of the model (rigid kinematics, motion governed by Newtonian dynamics, etc.). Using lower complexity models sacrifices expressiveness. In cases where a system's true structure is hard to discern (e.g. some human mental state/cognitive models) [14, 15], one would be well advised to use less structured models which avoid incorporating possibly incorrect assumptions.

The basic conceptual differences between structured mechanical models and more generic learned models are illustrated in Figure 1.3.

Incremental Learning

It would be both unnecessary and impractical to learn a completely new human preference model from scratch for every person a co-robot encounters. Instead, we'd like a more incremental process, where a robot starts with a high-uncertainty generic model of human preferences, then refines that model as it interacts with a new person and observes their behavior. The generic model, while suboptimal, will allow acceptable task performance while the robot interacts with the human and learns a personalized, high-precision model which will allow optimal task performance.

We investigate this scenario in Chapter 4, where we learn the parameters of a simple human arm configuration ergonomic model online during a handoff task. This learning takes place quickly – a personalized model is learned after only 5-10 handoffs, which illustrates the benefits of a structured model with a low dimensional parameter space. This low model dimensionality also facilitates the use of active learning approaches which would be computationally infeasible when paired with very complex models.

1.2 Example Application: Object Handoffs

Many real-world applications can get by with unsophisticated models of human capabilities and preferences. Examples include scenarios where just replaying the same action over and over, or applying a simple adaptation such as maintaining a constant pose relative to a person's torso (e.g. the 'Relative' scheme in Chapter 2) are adequate. It's easy to find cases where these ad-hoc models are inadequate though. Examples include humans whose capabilities differ from average due to natural variation, age, disease, or disability, or humans whose characteristics change over time due to fatigue, familiarity with a task, or other factors.

Throughout this thesis, we'll consider physical human-robot interaction (pHRI) problems where we seek to minimize a human ergonomic cost function $C_{ergo}(u_R, u_H)$ defined for a given task, where u_R is the action chosen by the robot, and u_H is the action chosen by the human. The core challenge in this problem is that while we can control u_R, we cannot control the

human's choice of u_H, leaving us uncertain what the cost produced by a particular robot action will be.

While the theory and concepts discussed in this thesis are quite general, for the sake of concreteness, we'll focus on the example task of a robot-to-human object handoff. This deceptively simple task crisply illustrates many of the issues mentioned above, including nondeterministic human actions (Chapters 3 and 4), humans optimizing greedily/myopically (Chapter 3), and variation in human range of motion and ergonomic preferences between individuals (Chapters 2 and 4).

1.3 Organization and Contributions

This dissertation is divided into three chapters, which address the planning of ergonomically optimal handoffs using different formulations. A central question addressed in each chapter is: How do the human and the robot select their actions $u_H{}^*$ and $u_R{}^*$, respectively. Each of the following three chapters makes a set of assumptions about how these actions are chosen, then investigates the implications for the resulting collaborative task.

The contributions of this dissertation are as follows:

Chapter 2: Human Kinematic Models for Safety and Feasibility

In Chapter 2, we present a methodology for fitting personalized kinematic models to people that capture the feasibility of different arm configurations and object grasps. In contrast to the existing literature, this modeling framework estimates kinematic model parameters such as joint axes, joint ranges of motion, and limb lengths directly from a specific individual's training data, without the use of *a priori* assumptions on joint and limb parameters. This kinematic formalism explicitly encodes the natural kinematic constraints of the human skeleton, such as joints with limited degrees of freedom and range of motion, given minimal training data. It can be quickly adapted to human collaborators whose motion may be limited by disability, age, or injury. A user study demonstrates that planning with personalized kinematic models yields less unergonomic bending and twisting during object handoff tasks than generic approaches, when considering human collaborators with limited upper body mobility.

Human/Robot Model Assumptions

This chapter assumes the human collaborator selects an action according to a known, deterministic model:

$$u_H{}^* = \underset{u_H}{\arg\min}\, C_{ergo}(u_R, u_H) \tag{1.1}$$

and that the robot need only choose the optimal corresponding action:

$$u_R{}^* = \underset{u_R}{\arg\min}\left(\min_{u_H} C_{ergo}(u_R, u_H)\right) \tag{1.2}$$

Chapter 3: Proactive Assistance for Collaborative Tasks

In Chapter 3, we consider that humans don't plan their actions based solely on feasibility. Instead, they have preferences for some feasible actions over others. Humans' preferences may also be greedy or myopic, leading to globally suboptimal action choices in some tasks. We develop a method for modeling these preferences by assuming humans are approximately rational agents that optimize some task or ergonomic cost function $C_{human}(u_R, u_H)$ which is known by the robot, but may be different from the global ergonomic cost function $C_{ergo}(u_R, u_H)$. We develop an *expected cost* minimization algorithm which leads even greedy humans to select globally optimal actions. Our user and simulation studies demonstrate that this algorithm produces better ergonomic performance than existing state-of-the-art methods in multi-part handoff tasks, is preferred by users, and is robust to inaccuracy in our human models.

Human/Robot Model Assumptions

In this chapter, we assume a more empirically realistic model where the human chooses their action nondeterministically according to a known distribution:

$$u_H{}^* \sim P(u_H|u_R), \quad P(u_H|u_R) \propto e^{-\beta C_{human}(u_R, u_H)} \tag{1.3}$$

and the robot chooses an action which yields a minimum human ergonomic cost *in expectation* given the likelihood of different human responses:

$$u_R{}^* = \arg\min_{u_R} \left(\underset{u_H \sim P(u_H|u_R)}{E} [C_{ergo}(u_R, u_H)] \right) \tag{1.4}$$

Chapter 4: Active Learning of Human Ergonomic Preferences

In Chapter 4, we consider that, because individual human preferences and capabilities differ widely, a co-robot won't know the $P(u_H)$ model for an unknown human it encounters. We'll model this uncertainty by assuming the C_{human} cost function from Chapter 3 is parameterized by a vector λ with an unknown value. We can expect to have only an uncertain *prior belief* $P_0(\lambda)$, which we can update $(P_1(\lambda), P_2(\lambda), \ldots)$ as we observe the human's action choices at future time steps.

We show that certain robot actions u_R ("queries") elicit informative human responses which decrease the entropy of our belief $H(P(\lambda))$ more quickly than other less informative queries. Using this observation, we develop an active perception/learning [16,17] approach to choose u_R's which most precisely and rapidly identify a given human's λ parameters online.

Human/Robot Model Assumptions

We assume the human chooses their actions according to the same model as in Equation 1.3, but now parameterized by an uncertain λ:

$$u_H{}^* \sim P(u_H|u_R, \lambda), \quad P(u_H|u_R, \lambda) \propto e^{-\beta C_{human}(u_R, u_H; \lambda)} \tag{1.5}$$

and we choose robot actions to balance quickly learning the human's λ with minimizing expected ergonomic cost in the current task (an *exploration vs. exploitation* tradeoff):

$$u_R{}^* = \arg\min_{u_R} \left(\mathop{E}_{\lambda \sim P_t(\lambda)} \left[\mathop{E}_{u_H \sim P(u_H | u_R, \lambda)} \left[-\Delta H(P_t(\lambda), P_{t+1}(\lambda)) + C_{human}(u_R, u_H; \lambda) \right] \right] \right) \quad (1.6)$$

Chapter 2

Human Kinematic Models for Safety and Feasibility

2.1   Introduction

Throughout daily life, humans collaborate with one another to manipulate objects in the world. Whether rearranging furniture or transferring tools, each person has a sense of their partner's preferential limb and body configurations. We prefer to grasp objects: close to the body rather than out of reach; between waist and chest height rather than overhead or underfoot; and within view rather than out of sight. Endowing co–robots with this collaborative manipulation capability remains an obstacle to ubiquitous deployment of autonomous service robots.

One subproblem in this domain which has received attention recently is the selection of "handoff configurations" or "object transfer points" (Figure 2.1). This problem arises in tasks where a robot must pass an object to or from a human collaborator. Previous approaches construct a cost function over the possible handoff configurations with respect to the human collaborator, then select an optimal configuration with respect to this cost function from the set of all configurations which are feasible for the robot. Previous authors have designed these functions to capture desirable qualities of handoff configurations, such as safety, visibility, and comfort [19], or usability, naturalness, and appropriateness [20, 21]. Cost functions can also capture ergonomic risk factors for injury [22–26].

When computing these functions, many methods in the literature use a generic human kinematic model to capture the comfort or ergonomics of handoff configurations. These models typically use kinematic parameters determined *a priori* from average morphomet-
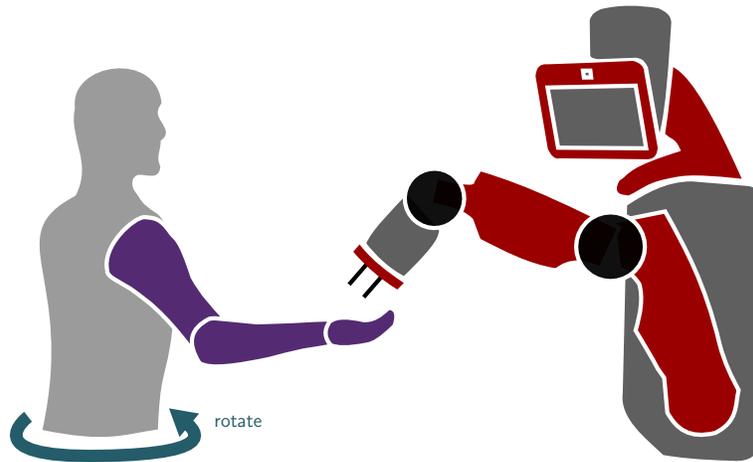
Figure 2.1: It is easier for a robot to hand off an object when it knows the kinematics of its human partner. If it neglects these kinematics, the person can compensate by *rotating* their torso, but this can lead to injury [22–26]. Thus human–robot collaborative manipulation is aided by providing the robot with a personalized kinematic model.

rics [19–21]. Though appropriate for a segment of the population, it is clear that injury, disability, fatigue, and natural variation in body size and shape can make predictions from any particular kinematic model inapplicable to many individuals. Simple adjustments such as scaling kinematics with height may improve predictions, but a *personalized* model fit to an individual's kinematics should give superior performance (at the expense of additional effort invested in model calibration and validation). Indeed, previous authors have noted that customization to individual human partners is needed [20].

Algorithms for fitting a kinematic "skeleton" to an individual using motion capture or vision data exist in the literature, but the resulting models often use only three degree-of-freedom spherical joints and fail to capture other kinematic constraints such as limb length and joint range of motion. The Kinect algorithm [27] for real-time skeletal pose estimation is a prominent example. The proprietary software included with many motion capture systems typically enforces fixed limb lengths, but does not model joints with limited degrees of freedom or restricted range of motion [28–30].

Organization and Contributions

In this chapter we present a system for estimating personalized human kinematic models from motion capture data. Our work makes two main contributions: First, kinematic model parameters such as joint axes, joint ranges of motion, and limb lengths are estimated directly from a specific individual's training data, without the use of *a priori* assumptions on joint and limb parameters. Second, our kinematic formalism explicitly encodes the natural kinematic constraints of the human skeleton, such as joints with limited degrees of freedom and range of motion, given minimal training data. We exploit a parsimonious *twist* representation [13] as

in [31] to obtain a minimal intrinsic parameterization for joints. This work complements the existing literature, as the models it generates can be directly incorporated into frameworks for object handoff planning [19–21] and more general human modeling [32–34]. Specifically, our personalized kinematic models are a drop-in replacement for the generic human kinematic models used in [19–21] and our learned model parameters can be used to rescale and calibrate the detailed musculoskeletal models in [32–34] to a specific individual.

We expect that adapting robot behavior using personalized models will confer advantages including safer, more ergonomic interaction with humans of varying physical dimensions and more effective collaboration with humans whose capabilities are restricted by injury or disability. To demonstrate the utility of the proposed framework, we compared three schemes for generating bimanual object handoff locations from a robot (Baxter Research Robot, Rethink Robotics) to a human partner in a motion capture arena. The three handoff schemes differed in the data available at the moment of object transfer:

- **Constant:** A handoff pose constant relative to the robot body frame. This was selected as a naive approach to serve as a control.

- **Relative:** A handoff pose constant relative to the human torso frame. This scheme represents configurations computed from the generic human kinematic models in [19–21], as discussed in Section 2.2.

- **Personal:** A preferred handoff configuration predicted using a personalized human kinematic model. This is the method developed in this paper, and is discussed in Section 2.2.

To evaluate these approaches, we compared the *rotation* (Figure 2.1) in the human's trunk at the moment of object handoff, since this statistic correlates with lower back injury [26, Table 1].

Each subject performed a randomized sequence of handoff experiments with the handoff location generated using each scheme. In addition, each subject repeated the process with two treatments: first unencumbered and subsequently with the dominant arm restricted by a strap. Restricting the dominant arm with a strap was intended to simulate loss of range of motion due to an acute injury, for instance if it is painful to move the limb through its uninjured range of motion or if the limb is physically encumbered by a cast or sling. We expected to observe significantly more trunk *rotation* when the robot only had access to its own reference frame (the *constant* scheme), particularly for subjects whose morphology differed from the generic model used to choose the fixed handoff location. Furthermore, we expected that adjusting the handoff location to account for the human's trunk position and orientation (the *relative* scheme) would nevertheless result in significant *rotation* when the dominant arm was restricted.

2.2 Methods

Kinematic Model

Following conventions set by the International Society of Biomechanics [32,33], we represent the topology of a human using a rooted tree composed of up to five kinematic chains. The tree is represented by (J, E) where J is a set of joints and $E \subset J \times J$ is a set of edges representing rigid links. Each $j \in J$ has a single degree–of–freedom (DOF) specified by a twist $\xi_j \in \text{se}(3)$. Thus displacing joint j by an amount $\theta_j \in \mathbb{R}$ yields a rigid body transformation

$$\exp\left(\widehat{\xi}_j \theta_j\right) \in \text{SE}(3) \tag{2.1}$$

between the *input* and the *output* of j, where the $\widehat{\ }$ operator maps a twist vector $\xi = [\omega | v]^T \in \mathbb{R}^6$ to its equivalent homogeneous representation:

$$\widehat{\xi} = \begin{bmatrix} \widehat{\omega} & v \\ 0 & 0 \end{bmatrix}, \quad \widehat{\omega} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \tag{2.2}$$

Similarly, each edge $(i, j) \in E$ represents a rigid link from the output of joint i to the input of joint j. Thus given a configuration vector $\theta \in \mathbb{R}^n$ specifying displacements for each of the joints in J, the rigid body transformation from the input of the root r to the output of joint $j \in J$ is given by the *product of exponentials* [13]

$$g_j(\xi, \theta) = \left(\prod_{i \in c(j)} \exp\left(\widehat{\xi}_i \theta_i\right) \right), \tag{2.3}$$

where $c(j)$ is the unique, ordered sequence of joints connecting r to j. The world frame position of a *feature* $p_i \in \mathbb{R}^3$ (e.g. a motion capture marker) rigidly affixed to the output of joint j is then given by

$$g_j(\xi, \theta) \begin{bmatrix} p_i \\ 1 \end{bmatrix}. \tag{2.4}$$

This compact representation can model revolute (rotational) in addition to prismatic (linear displacement) joints.

The twist formalism for kinematics has two main advantages for this application: First, its lack of singularities makes the parameter estimation cost function $J(\xi, \theta, p)$ (in Section 2.2) *smooth* with respect to the joint parameters ξ. Second, with only six free parameters per joint, the twist parameterization is *minimal*, which minimizes the amount of training data necessary for the identification algorithm to achieve a specified accuracy.

To complete our model, we define the map $\alpha : \{1, \ldots, m\} \to J$, which specifies which joint's output each feature is rigidly attached to. We refer to the collection $S = (J, E, \alpha, \xi, p)$ of a tree (J, E) with a feature-to-joint mapping α, twists $\xi \in \text{se}(3)^n$, and features $p \in \mathbb{R}^{3 \times m}$ as a *kinematic skeleton*.

Joint j
Parameters:
$\xi_j \in \mathrm{se}(3)$

State:
$\theta_j \in [\theta_{Mj}, \theta_{mj}]$

Feature p_i
Parameters:
$p_i \in \mathbb{R}^3$
$\alpha : \{1, \ldots, m\} \to J$

known   estimated

$$S = (J, E, \alpha, \xi, p)$$

joints   feature positions

rigid links   feature-joint assignments   joint parameters
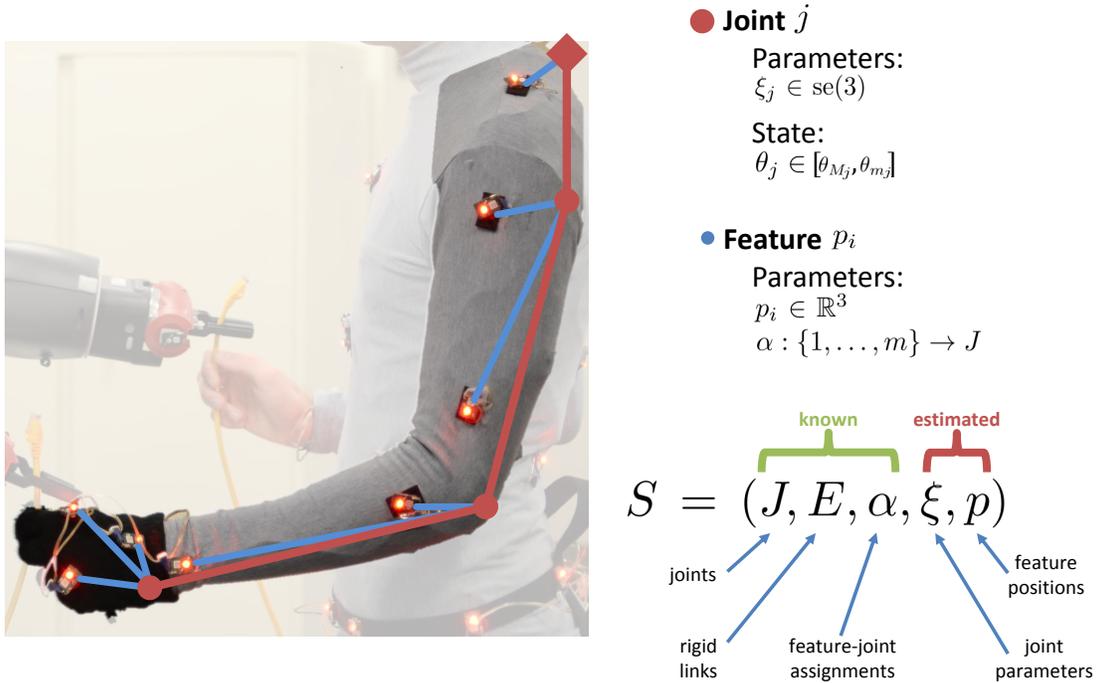
Figure 2.2: The human skeleton model illustrated with a single human arm assumed to have three degrees of freedom. The model consists of three *joints* j_i, each of which is parameterized by a *twist* $\xi_i \in \mathrm{se}(3)$, which specifies the axis of motion of the joint. The joint also has a state θ_i, which specifies the current joint displacement. The state is constrained to lie in the interval $[\theta_{m_i}, \theta_{M_i}]$ (the feasible range of motion for this joint). Attached to each joint as children is a set of zero or more observable *features* (motion capture markers, in this case). Each feature i is parameterized by $p_i \in \mathbb{R}^3$, which specifies the position of the feature relative to the world coordinate system with all of the joint states in the kinematic skeleton set to zero. Finally, a map $\alpha : \{1, \ldots, m\} \to J$ specifies which joint in the tree each feature is a child of.

$$\exp\left(\widehat{\xi}_j \theta_j\right) \in \mathrm{SE}(3) \quad\Longrightarrow\quad g_j(\xi, \theta) = \left(\prod_{i \in c(j)} \exp\left(\widehat{\xi}_i \theta_i\right)\right) \quad\Longrightarrow\quad g_j(\xi, \theta) \begin{bmatrix} p_i \\ 1 \end{bmatrix}$$
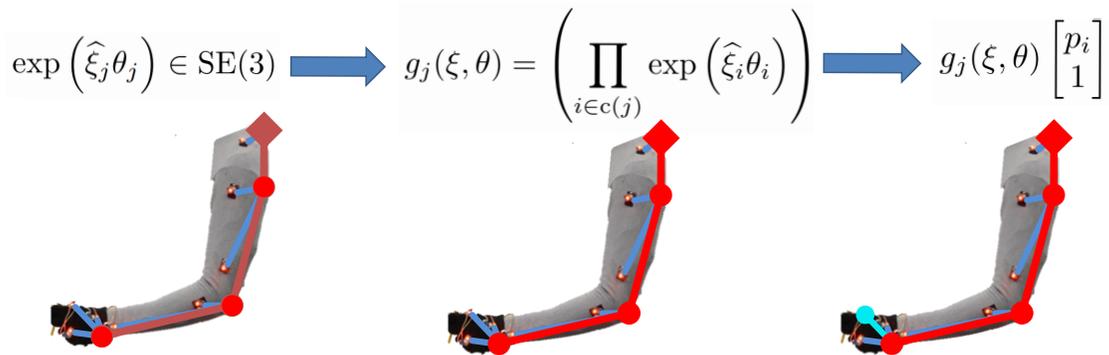
Figure 2.3: The human skeleton model allows easy computation of the predicted position of each feature using standard forward kinematics. First, the homogeneous transformation associated with each joint is computed. Next, the transforms along the path from the selected feature to the tree's root are composed, in order. Finally, the resulting transformation matrix is multiplied by the zero-configuration position of the feature to give its predicted position.

Parameter Identification

To model human motion using the *kinematic skeleton* developed in Section 2.2, we assume the tree structure (J, E) and feature-to-joint mapping α are known but the twists $\xi \in \mathrm{SE}(3)^n$ and feature locations $p \in \mathbb{R}^{3 \times m}$ are unknown. It is difficult to directly measure ξ and p, therefore we estimate these quantities from a training dataset $\eta_{1:N}$, with each $\eta_k \in \mathbb{R}^{3 \times m}$. This dataset conists of N (noisy) observations of the coordinates of m features in the world frame (provided, in our case, by a motion capture system). These are collected while the subject performs some sequence of training motions. For the upper body model used in our experiments, this sequence consisted of moving the shoulder and elbow joints through their full ranges of motion, as shown in Figure 2.4.

Given this training dataset, the skeleton parameters are estimated as in [31] using non-linear least-squares *prediction error minimization* [35] on a collection of error vectors with the form

$$\varepsilon(\xi, \theta_k, p_i) = g_{\alpha(i)}(\xi, \theta_k) \begin{bmatrix} p_i \\ 1 \end{bmatrix} - \begin{bmatrix} \eta_{k,i} \\ 1 \end{bmatrix}. \tag{2.5}$$

Note that in addition to the skeletal parameters ξ and p, the joint displacements θ must also be estimated for each frame in the training dataset to completely specify the prediction error. Thus, the error function which is minimized is

$$J(\xi, \theta, p) = \sum_{k=0}^{N} \sum_{i=0}^{m} \|\varepsilon(\xi, \theta_k, p_i)\|_2^2. \tag{2.6}$$

For all joints $j \in c(\alpha(i))$ that precede $\alpha(i)$, the derivatives of ε with respect to ξ_j, θ_j, and p_i are given by (see also (33) in [36])

$$\begin{aligned} D_{\xi_j}\varepsilon(\xi, \theta, p_i) &= \widehat{A}_j p_i, \\ D_{\theta_j}\varepsilon(\xi, \theta, p_i) &= (\mathrm{Ad}_{g_j(\xi,\theta)} \xi_j)^\wedge p_i, \\ D_{p_i}\varepsilon(\xi, \theta, p_i) &= R_j(\xi, \theta), \end{aligned}$$

where the matrix A_j is given in [36, Eqn. 14] and R_j is the rotational component of $g_j(\xi, \theta)$. (For joints $j \notin c(\alpha(i))$ that do not precede $\alpha(i)$, the derivatives $D_{\xi_j}\varepsilon$, $D_{\theta_j}\varepsilon = 0$.) Prediction error minimization was performed with the SciPy [37] interface to the `lmder` routine in MINPACK [38].

Figure 2.4: Calibration poses for 2–DOF human arm kinematics.

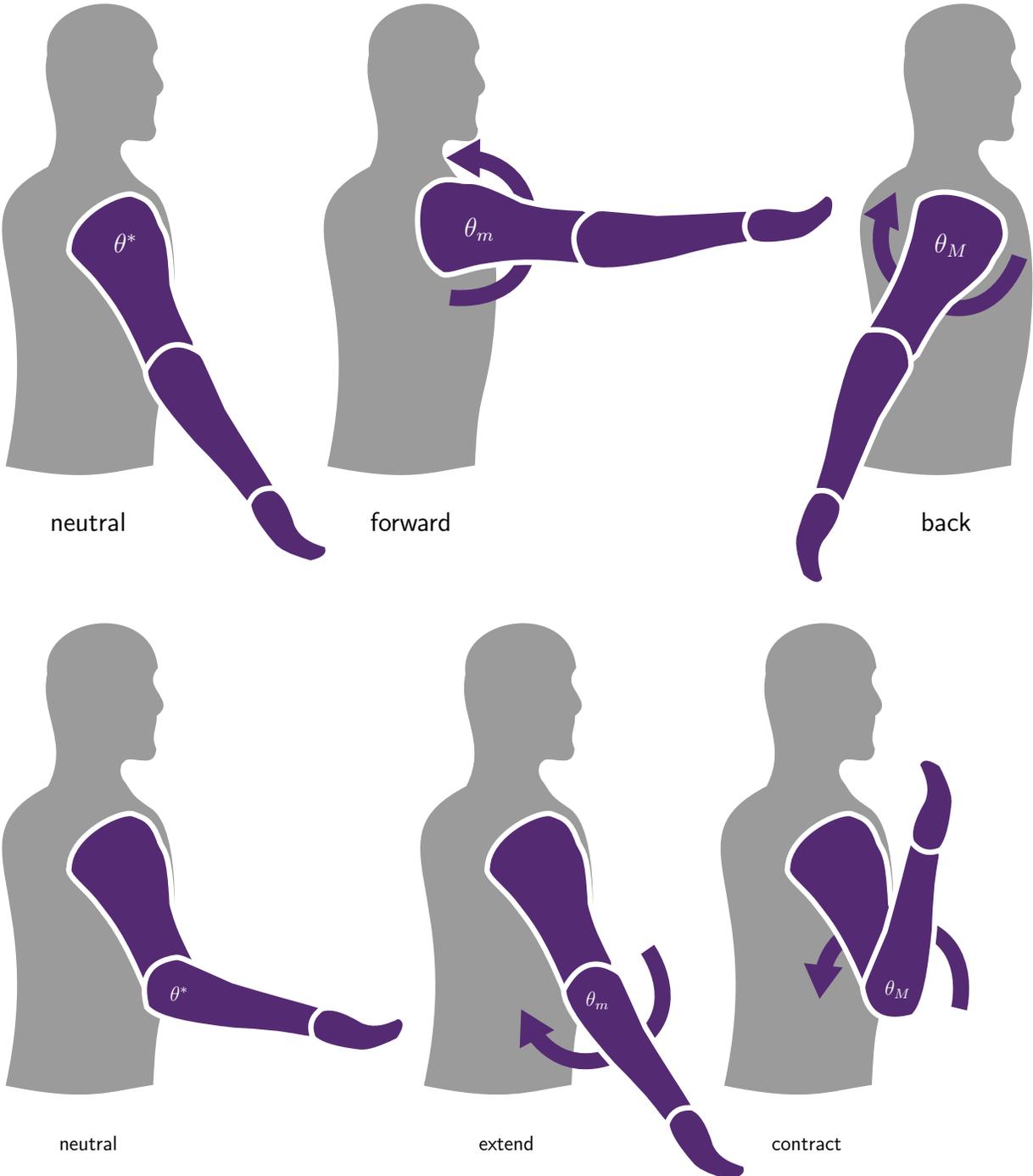## State Estimation

After estimating the geometric parameters of a kinematic skeleton S for a specific individual as in Section 2.2, we can estimate the state of the model (the joint displacements θ) online from a sequence of motion capture feature position measurements. This estimation was performed using an *Unscented Kalman Filter* (UKF) [39]. The filter is applied to discrete–time stochastic processes of the form

$$
\begin{aligned}
x_{k+1} &= f(x_k) + u_k, \ \ u_k \sim \mathcal{N}(0, U_k), \\
y_k &= h(x_k) + v_k, \ \ v_k \sim \mathcal{N}(0, V_k),
\end{aligned}
\tag{2.7}
$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ specifies the deterministic dynamics, $h : \mathbb{R}^n \to \mathbb{R}^m$ the observation function, and u_k, v_k are independent and (respectively) identically distributed Gaussian random variables. The initial state distribution is assumed Gaussian and denoted by $\mathcal{N}(x_0, P_0)$.

For the kinematic model of Section 2.2 the state is given by the generalized joint coordinates $x = \theta$, the dynamics are driven entirely by the process noise u (i.e. $f \equiv \mathrm{id}_{\mathbb{R}^n}$), while the observation of a feature affixed to the output of joint j at a position p is given in (2.4), and $h(\theta)$ is therefore obtained by vertically concatenating the three-dimensional position vectors from all features.

The UKF recursively estimates the first two moments of the state $\mathcal{N}(x_{k+1}, P_{k+1})$ at step $k + 1$ using the estimate $\mathcal{N}(x_k, P_k)$ from the previous step and observations y_{k+1}; for details we refer the interested reader to [39]. First, an array of *sigma points* [39] are assembled. The empirical process and observation covariances and cross-covariances are estimated by propagating these points through the model (2.7). Finally, the innovation step of the classical Kalman filter [40] is performed using the estimated covariances.

Pose Prediction

Given a kinematic tree (J, E) calibrated to a subject as in the preceding section, we now consider the problem of predicting the behavior of the human subject during a collaborative manipulation task. We begin by reviewing the rich scientific literature that aims to address this problem before describing how some of the most popular theories can be incorporated into our framework.

At present, one of the most popular and fruitful theories of motor control is *optimal feedback control theory*, where it is posited that the central nervous system synthesizes motion by minimizing a cost $C \in \mathbb{R}$ that varies as a function of the joint angle θ, torque τ, or end effector x trajectory (and, possibly, their derivatives) over a time interval $[0, T] \subset \mathbb{R}$, either in open–loop or through receding–horizon feedback [41, 42]. For trajectory generation, one of the earliest proposed and oft–cited forms for the cost function is *minimum jerk* [43],

$$
C_{\dddot{x}} = \int_0^T \| \dddot{x}(t) \|^2 \, dt.
\tag{2.8}
$$

However, subsequent studies have shown that other statistics such as *minimum torque change* [44],

$$C_{\dot{\tau}} = \int_0^T \|\dot{\tau}(t)\|^2 \, dt. \tag{2.9}$$

or *minimum motor command* [42] produce better predictions.

In the present setting, we are more concerned with the final pose of the subject than the trajectory adopted to reach that pose. For static posture prediction, a classical law (alternately attributed to Donder or Listing [45]) posits that attributed to each hand pose there exists a unique preferred limb posture. Though appropriate for some experimental settings, [45] found that this "law" yields poor predictions for limb posture in a reaching task, and demonstrated that *minimum work*,

$$C_W = \int_0^T \tau^T(t)\dot{\theta}(t)dt \tag{2.10}$$

provides superior predictions.

We conclude that, depending on the collaborative manipulation task under consideration, a cost function with the form given in either (2.8), (2.9), or (2.10) may provide superior predictions of human behavior. Note that our framework is applicable to any cost function that varies smoothly with respect to joint angle θ, torque τ, or end effector x trajectory, including but not limited to (2.8–2.10). For the handoff experiments in this paper, we elected to choose the simplest cost function that is consistent with our kinematic skeleton model. Specifically, we select a posture θ^* that is merely feasible based on the subject–specific joint limits computed in the calibration step, then minimize the *preferred posture* cost

$$C_{\theta^*} = \|\theta - \theta^*\| \tag{2.11}$$

using the choice $\theta^* = \frac{1}{2}(\theta_M - \theta_m)$, i.e. halfway between the upper and lower joint limits. We emphasize that the cost (2.11) was chosen for its simplicity, and that we expect choosing a more biologically–plausible cost will yield strictly superior results relative to those obtained with (2.11). In particular, any ergonomic benefit conferred by employing the preferred posture cost (2.11) should be enhanced by instead minimizing jerk (2.8), torque change (2.9), or work (2.10).

Finally, we note that each of the costs above (2.8–2.11) only require a personalized *kinematic* model for the subject. It is conceptually straightforward to extend the framework in this paper to accommodate personalized *dynamic* models, but doing so requires conducting a more elaborate set of calibration experiments to estimate inertial parameters [46, 47]. A dynamic model is unnecessary in the present paper; since we focus on demonstrating the value of personalized models using the simplest quasistatic cost (2.11), the inertial parameters of a dynamic model would have no effect on the pose prediction.

Handoff Experiments

We compared three schemes for generating bimanual object handoff locations from a robot (Baxter Research Robot, Rethink Robotics) to a human partner in a motion capture arena. The three handoff schemes differed in the data available at the moment of object transfer:

Constant

Generates a handoff pose in a fixed location relative to the robot body frame. This scheme does not make use of any data about the human's location or kinematic structure. It is a naive scheme which we include as a control.

Relative

Generates a handoff pose that is constant relative to the human torso frame. This scheme represents the handoff locations generated using the generic human kinematic models in prior approaches. For example, in [20, 21], the authors use a generic human kinematic model to evaluate one component (denoted f_{take}) of their handoff cost function, which is maximized at the object pose with the largest number of possible "take" configurations. Similarly, in [19], the authors optimize a handoff cost function which includes an "arm comfort" term. This term is itself a sum of the squared displacement of the human's joints from a resting configuration, plus the gravitational potential energy of the arm's current configuration, and is also evaluated using a generic human kinematic model. Because these approaches use *generic* human models, they will produce the same optimal handoff configurations with respect to the human's body frame, regardless of variations in the limb dimensions or range of motion of a particular human partner. The *relative* scheme represents these approaches by performing the object handoff at a configuration which is constant relative to a frame attached to the human collaborator's torso.

Personal

Predicts a preferred handoff configuration using a personalized kinematic model. This scheme is our approach. It generates a handoff configuration which is optimal with respect to the chosen ergonomic cost function (Section 2.2), for the limb dimensions and range of motion of an individual human collaborator, identified as described in Section 2.2.

Each subject performed a randomized sequence of 3×10 handoff experiments with the handoff location generated 10 times using each scheme. In addition, each subject repeated the process with two treatments: first unencumbered and then with the dominant arm restricted by a strap. Restricting the dominant arm with a strap was intended to simulate loss of range of motion due to acute injury. As an evaluation criteria, we compared the *rotation* in the human's trunk at the moment of object handoff. We expected the following outcome:

H_0 The *constant* and *relative* schemes generate significantly more *rotation* than the *personal* scheme with the subject's arm restricted.
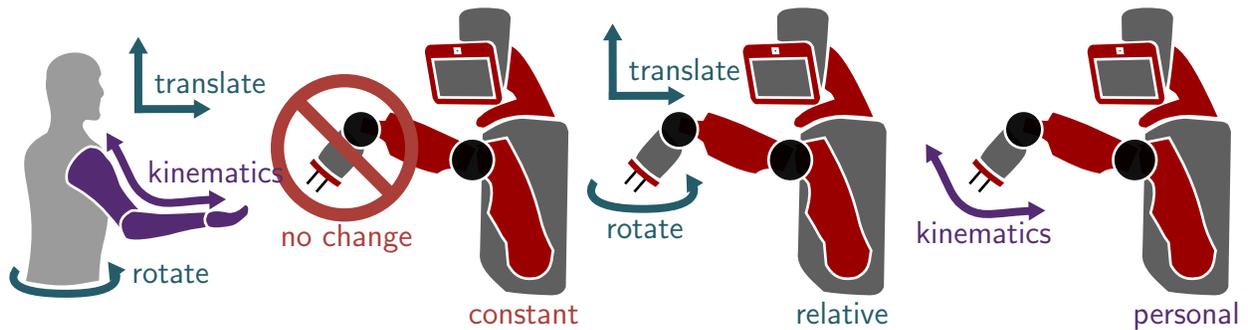
Figure 2.5: Illustration of handoff schemes. Using the *constant* scheme, the robot hands the object off at a fixed location in its frame of reference. With the *relative* scheme, it hands off to a fixed location in the human's reference frame, and therefore compensates for rotation and translation of the human's torso. The *personal* scheme compensates for general changes to kinematics, including rotation, translation, and restriction of range–of–motion (e.g. due to injury, disability, or fatigue).

Human Subjects Protocol

Each test subject was first outfitted in an upper body motion capture suit. A total of 24 markers were attached to the suit, with four on the subject's chest, four on the back, and eight distributed along the length of each arm and hand (Figure 2.6). An active marker motion capture system was used so that each marker was uniquely identifiable in the resulting dataset. Data was collected at 50 Hz.
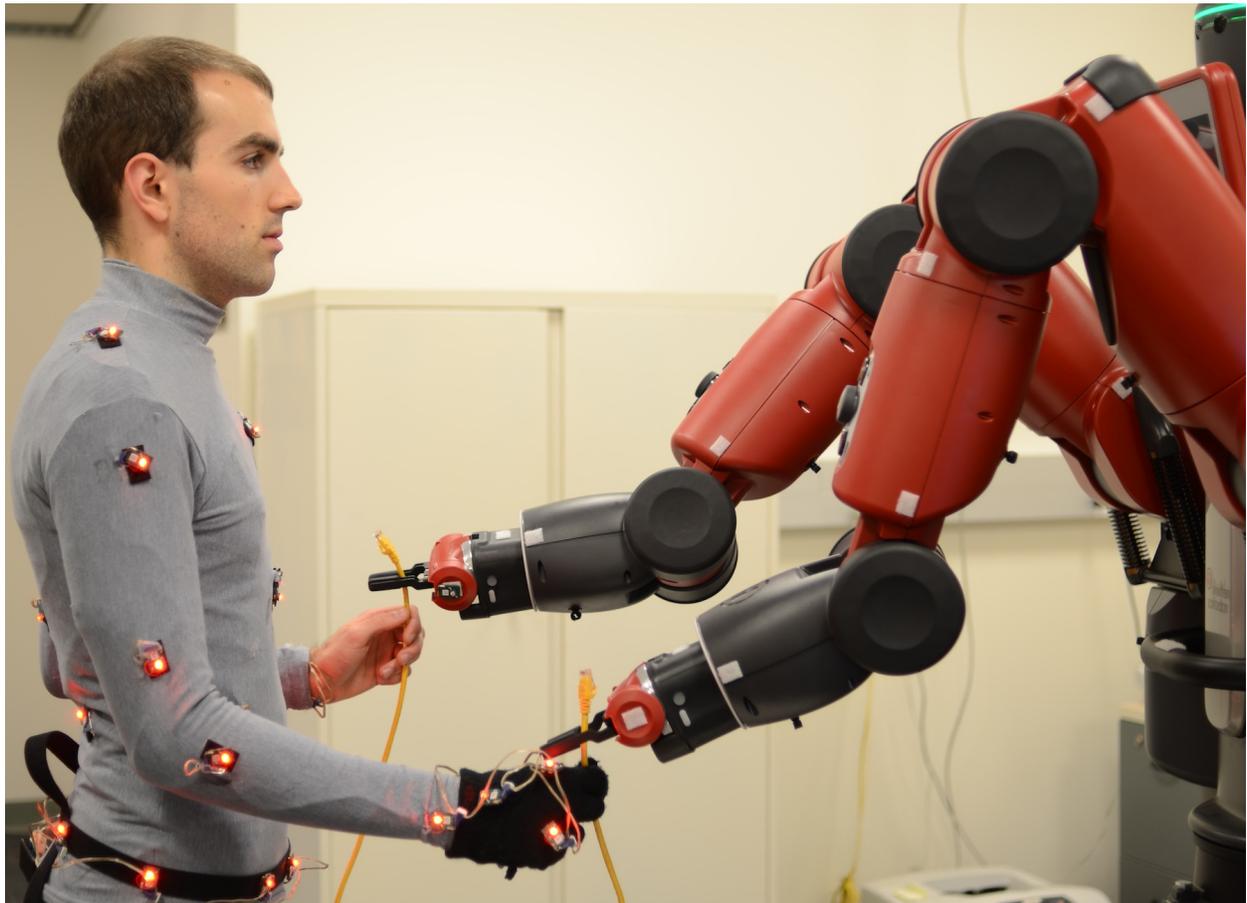
Figure 2.6: A photo of the experimental setup. Each experiment tested 30 randomly ordered handoffs, of which 10 were generated using each of the *constant*, *relative*, and *personal* schemes discussed in Section 2.2. The pose of the human test subject's torso was measured at the moment of the handoff using the eight motion capture markers on the chest and back.

Before beginning the handoff trials, the test subject performed a calibration sequence to fit a personalized kinematic model as described in Section 2.2, and to estimate the subject's preferred handoff pose as in Section 2.2. The subject performed three sets of each of the calibration motions shown in Figure 2.4. The subject was then instructed to stand with their feet at marked locations on the floor and the following test procedure was performed for a total of 30 trials:

1. robot picks up a single cable from a table while displaying "Please wait" on its head display;

2. robot chooses a handoff pose, hands the cable to the subject (Figure 2.6), and displays "Ready?" on its display;

3. after two seconds, robot's grippers open and arms retract;

4. subject takes the cable and plugs it into a piece of network hardware, then waits for the next handoff.

The handoff pose in each trial was randomly chosen as one of constant, relative, or personalized, as defined in Section 2.2 and illustrated in Figure 2.5. The handoff poses were chosen such that 10 handoffs of each type were performed in one session. During each trial, motion capture was used to record the pose of the subject's torso both before the beginning of the trial and at the moment the handoff occurs.

After completing the first set of 30 handoffs with normal arm movement, the test subject's dominant arm was affixed to their torso with a strap to restrict its movement. This was intended to simulate the range of motion observed with one's arm in a sling after an injury. The test subject then completed another calibration sequence. This sequence was used to identify a new kinematic model and predict a new handoff pose given the newly restricted range of motion. A new session of 30 trials was then run using the same protocol as before.

2.3 Results

Personalized Kinematic Models

After fitting a kinematic model to a test subject's calibration sequence, the accuracy with which the model's rigid kinematic structure captured the subject's actual motion was evaluated by computing the reprojection error for each motion capture marker in each frame of the calibration sequence. Across all four test subjects and both the restricted and unrestricted motion trials, the median reprojection error for the 16 arm markers ranged from 4.86 cm to 0.29 cm, with a mean of 1.49 cm. This relatively low error suggests that the kinematic model identified by the parameter fitting algorithm accurately captured the kinematic constraints observed in the subject's actual motion.

Figure 2.7 shows the feasible workspaces computed from the personalized kinematic models of two test subjects. Note that though the workspaces are qualitatively similar, there are significant differences between the two subjects. In particular, the workspaces of the two subjects with restricted arm motion have only a small area of overlap, indicating that a single, generic kinematic model would have had difficulty capturing both subjects' physical constraints simultaneously. Even in the unconstrained case, the portion of the reachable workspace with the person's arm extended rearward is significantly larger for Subject 2 than for Subject 1 (Figure 2.7).
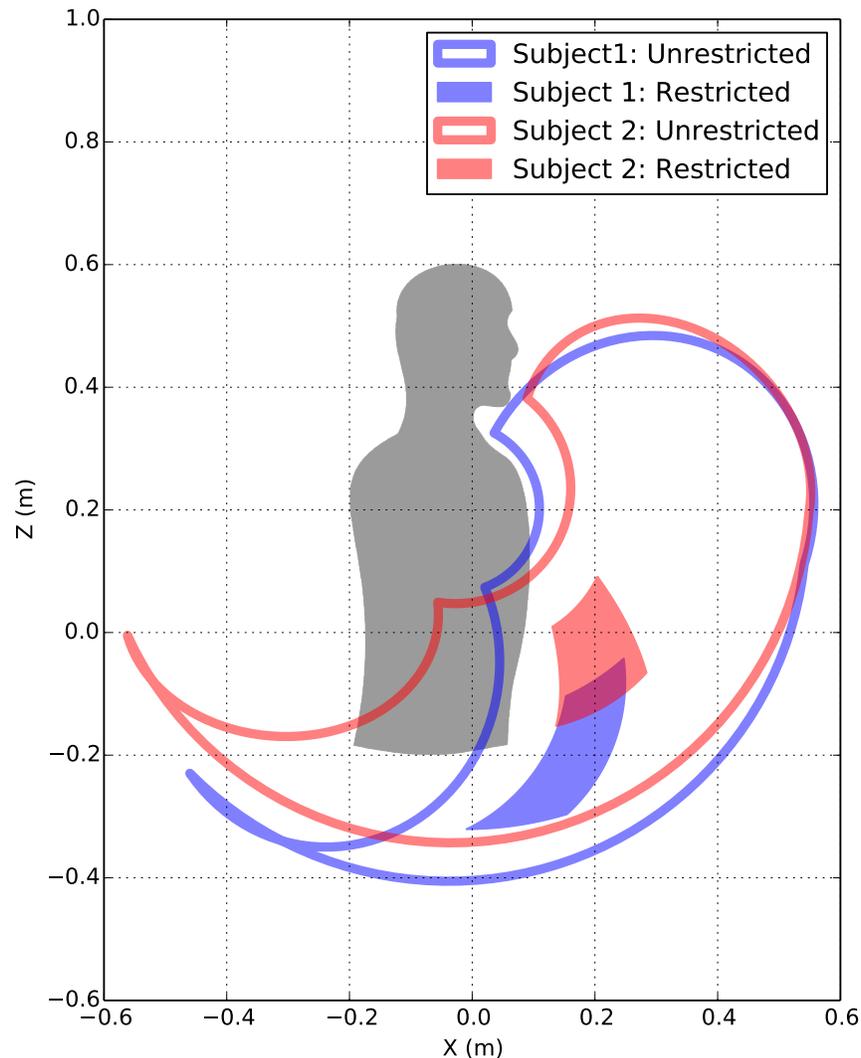
Figure 2.7: Feasible workspaces computed from the personalized kinematic models of two selected test subjects for both the restricted and unrestricted cases. Models were estimated as described in Section 2.2. Note that there is little overlap between the restricted workspaces of the two subjects, and that even for the unrestricted case, the portion of the workspace to the rear of the subject is significantly larger for Subject 2 than for Subject 1.

Handoff Ergonomics

We applied a *pairwise t–test* [48] to assess whether the choice of the constant, relative, or personal handoff schemes produced a statistically significant change in the torso rotation angles measured at the moment of the handoff. This analysis was performed both individually for each subject and on the pooled dataset from all four subjects. The t–test implicitly assumes that the true distributions of rotation angles for a given subject are Gaussian and

have identical variance across all three handoff schemes.

For trials with unrestricted arm movement, there was no statistically significant difference ($p < 0.05$) between the torso rotation angles measured with the three different handoff schemes when tested across the pooled dataset (Figure 2.8). However, the analysis on the pooled dataset for trials with restricted arm movement produced significant differences in rotation angles between the personal and constant as well as the personal and relative handoff schemes.

This analysis suggests that the choice of handoff scheme produced a significant difference in torso rotation angles in subjects with a restricted range of motion.

2.4 Discussion

The similarity of the torso rotation angles measured for all three handoff schemes in the unrestricted case suggests that handoff planning methods based on generic kinematic models perform well when interaction partners have "typical" body dimensions and range of motion (Figure 2.8). However, the performance of these methods degrades when subjects' range of motion is restricted, since this necessitates significant compensatory motion in the torso to adapt to the robot's chosen handoff configurations. The use of handoff configurations generated using a personalized kinematic model significantly improved performance by allowing test subjects to maintain a neutral body posture at the moment of object handoff.

We believe the use of personalized kinematic models shows promise not only for the specific case of object handoffs, but also more broadly for human–robot collaboration. Our present implementation utilizes motion capture data to provide a proof-of-concept demonstration of the personalized kinematic model framework. The framework is easily extensible to real-world applications by using inertial measurements and other data streams such as camera data that can be obtained from devices such as the Kinect [49]. For simplicity we employed the personalized kinematic model only to estimate a *feasible* handoff location; it is straightforward to extend our framework to incorporate other constraints and objective functions [50] to predict human posture and motion from energetic [44, 51] or dynamic [47] principles. This work complements the existing literature, as the models it generates can be easily incorporated into frameworks for object handoff [19–21] and more general human modeling [32–34]. More broadly, it provides a foundation for employing personalized kinematic models in human–robot collaboration.

2.5 Extensions

While the bulk of this chapter investigates personalized kinematic models as applied to humans with limbs composed of one degree-of-freedom joints, other modeling scenarios are of interest as well. Two that we'll investigate further are:
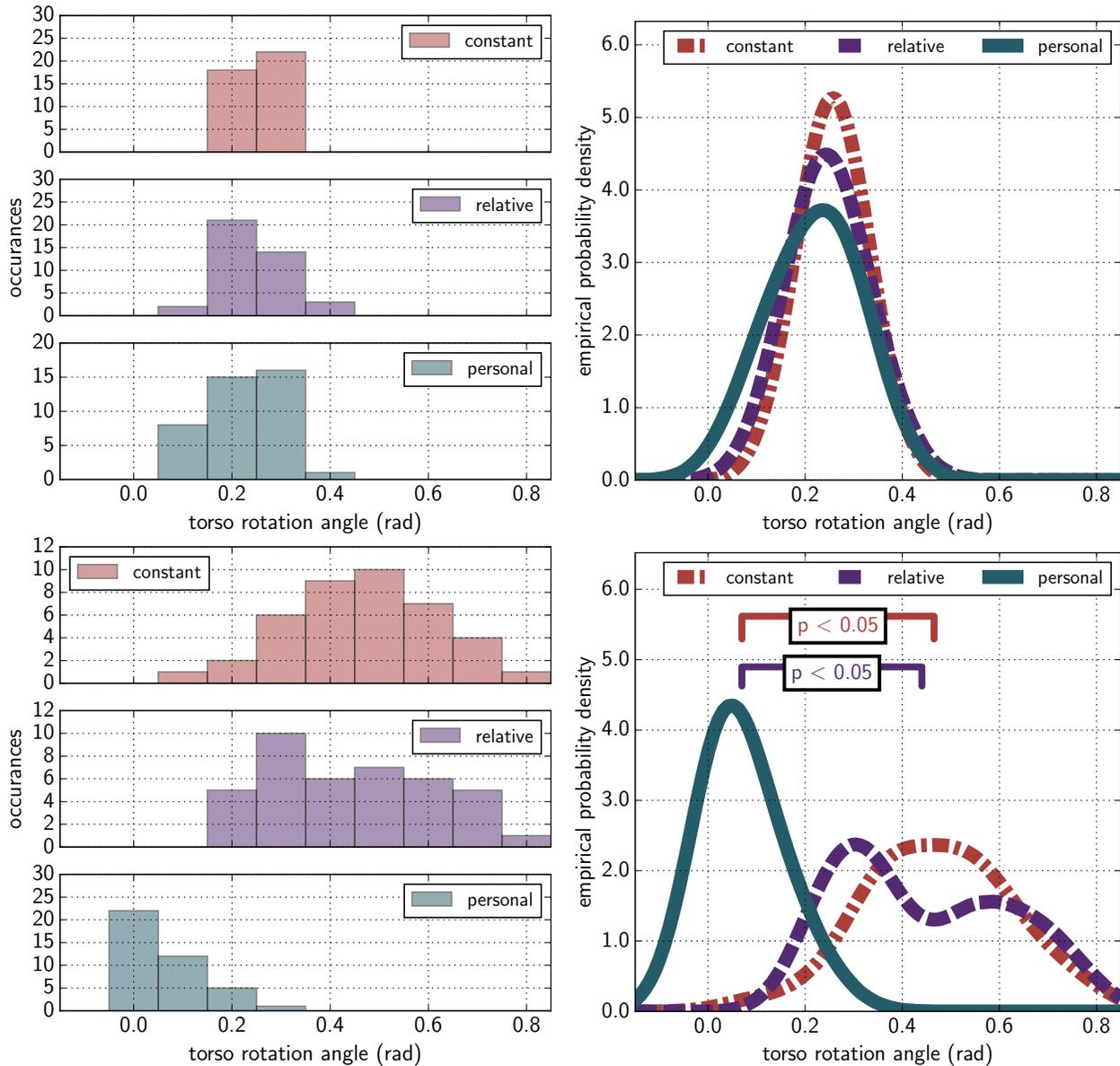
Figure 2.8: Rotation angle distributions of the pooled datasets for *unrestricted* (top) and *restricted* (bottom) kinematics. The distributions of rotation angles overlap substantially for the three handoff schemes with *unrestricted* kinematics. However, the distribution obtained from the *personalized* kinematics is significantly different from that obtained using only the *relative* position of the human or a *constant* handoff location (pairwise t–test, $N = 10$, $p < 0.05$).

- Multiple Degree of Freedom Joints: Some human joints are more accurately modeled not as a series of independent twists, but instead as a single joint with more than one degree of freedom. Examples include the hip and shoulder, which are well approximated by three degree of freedom ball joints.

- Jacobian Computation: The manipulator Jacobian is a useful tool for numerical inverse kinematics, manipulability evaluation, and other applications. Our personalized kinematic models support simple computation of Jacobians for arbitrary base and manipulator frames attached to any location on the kinematic tree.

Multiple Degree-of-Freedom Joints

While the kinematic representation used above, with a tree of one degree of freedom joints, each parameterized by a twist ξ_j, is appropriate for the example situation in our human experiments, it's easy to imagine other scenarios for which this representation is not ideal. Consider a scenario where we require a model of the full range of motion of a human shoulder, for example, as opposed to the simple motion used above. The shoulder is well approximated as a three degree of freedom "ball" joint. We could attempt to model the shoulder using three twists in series, but this creates a number of problems:

- The resulting joint has 18 parameters to be estimated (6 per twist), requiring significantly more training data to fully constrain all model parameters

- The twist axes of the resulting joint are not constrained to intersect at a common point, as we'd like for a ball joint

- The twist axes are not constrained to be orthogonal, as in the typical Euler angle parameterization of a ball joint's state

In fact, without observable features on the intermediate links connecting the twists of the ball joint, the three twist parameterization will be impossible to fully constrain, regardless of the amount or type of calibration data. The inability to force the twist axes to be orthogonal is of particular concern, as this could cause the kinematic state estimation problem to be ill conditioned when twist axes are near-parallel. This problem could be partially addressed by adding a regularization term to the optimization cost function which penalizes non-orthogonal twists, but even with this solution, the unnecessary high dimensionality of the parameter estimation problem remains. The solution is to choose a different, more appropriate joint parameterization for multi-DoF joints.

We'll accomplish this by allowing each joint j_i to have not just one, but multiple twists associated with it, which we'll denote as $\Xi_i = [\xi_{i,1}, \ldots, \xi_{i,N_i}]$. Each twist has an associated displacement, so the state of a single multi-DoF joint is given by the vector $\Theta_i = [\theta_i, 1, \ldots, \theta_i, N_i]$.

Finally, for each multi-DoF joint, we'll specify a vector of parameters $r_i \in \mathbb{R}^{N_{r_i}}$ that together define the feasible motions for the joint, and a map $t_i : r_i \mapsto \Xi_i$ that maps a given parameter vector to the joint's sequence of twists.

Example: Three DoF Ball Joint

Consider the case of the three degree of freedom ball joint we'd use to model a shoulder or hip. The joint should have three orthogonal twists, each of which represents one of its axes of rotation. Thus, $\Xi_i = [\xi_{i,1}, \xi_{i,2}, \xi_{i,3}]$, and the corresponding joint state will be $\Theta_i = [\theta_{i,1}, \theta_{i,2}, \theta_{i,3}]$.

We'll assume here that the three joint axes of rotation are parallel to the world frame x, y, and z axes, so the joint is parametrized by the world frame (x, y, z) coordinates of the intersection point of its three axes, and $r_i \in \mathbb{R}^3$.

Finally, we'll define the map t_i as:

$$t_i(r_i) = \left[\begin{bmatrix} -\omega_1 \times r_i \\ \omega_1 \end{bmatrix}, \begin{bmatrix} -\omega_2 \times r_i \\ \omega_2 \end{bmatrix}, \begin{bmatrix} -\omega_3 \times r_i \\ \omega_3 \end{bmatrix} \right] \tag{2.12}$$

where $\omega_1 = [1, 0, 0], \omega_2 = [0, 1, 0]$, and $\omega_3 = [0, 0, 1]$.

Jacobian Computation

Suppose we have two coordinate frames attached as children of joints j_{base} and j_{manip} on a kinematic tree model. We'll denote the world frame poses of these frames with all joint displacements set to zero as T_{base} and T_{manip}, respectively, where $T_{base}, T_{manip} \in \mathrm{SE}\,3$. Note that neither j_{base} nor j_{manip} need be the tree's root joint. We'd like to compute the spatial frame manipulator Jacobian $J_{bm}(\theta)$ of T_{manip} with respect to T_{base} at a specified configuration θ.

The computation is as follows:

1. Compute the transformed twist associated with each joint at the current configuration as:

$$\xi_i' = \mathrm{Ad}_{\left(e^{\hat{\xi}_1 \theta_1} \dots e^{\hat{\xi}_{i-1} \theta_{i-1}} \right)} \xi_i$$

   where $\mathrm{Ad}_{(\dots)}$ is the adjoint transformation matrix generated by the displacements of all the joints between j_i and the root of the kinematic tree

2. Find the unique joint chains $C_{base} \triangleq c(j_{base})$ and $C_{manip} \triangleq c(j_{manip})$ which connect the base and manipulator joints to the tree's base joint

3. Remove any joints present in both C_{base} and C_{manip} from both sequences to yield the shortest path connecting the two joints

4. Reverse the order of C_{base}

5. Generate sequences of the transformed twists along the path from j_{base} to j_{manip}:

$$\Xi_{base} = [-\xi_i']_{i \in C_{base}}, \qquad \Xi_{manip} = [\xi_i']_{i \in C_{manip}}$$

6. Concatenate the two transformed sequences into a single sequence of twists:

$$\Xi_{chain} = [\Xi_{base} | \Xi_{manip}]$$

7. Use the forward kinematics map to compute the world frame poses of the base and manipulator frames at the current configuration θ, which we'll denote as $T_{base}(\theta)$ and $T_{manip}(\theta)$.

8. Express each $\xi'_{chain,i} \in \Xi_{chain}$ in the base coordinate frame:

$$\xi'_{chain,i} = \mathrm{Ad}_{(T_{base}(\theta)^{-1})} \xi'_{chain,i}$$

9. Compute the motion of $T_{manip}(\theta)$ generated by each twist in Ξ_{chain}, expressed with respect to $T_{base}(\theta)$:

$$J_{bm}(\theta) = \begin{bmatrix} \nu \\ \omega \end{bmatrix}_{\xi'_{chain,i} \in \Xi_{chain}}, \text{ where } \begin{bmatrix} \nu \\ 1 \end{bmatrix} = \widehat{\xi'}_{chain,i} T_{manip}(\theta) \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \text{ and } \begin{bmatrix} - \\ \omega \end{bmatrix} = \xi_{chain,i}$$

Chapter 3

Proactive Assistance for Collaborative Tasks

3.1 Unconstrained Human Degrees of Freedom in pHRI Tasks

A fundamental problem in pHRI is that we can't control what the human does. A robot's actions can impose constraints on certain human degrees of freedom, but in general there will be other DoFs where the human's action is unconstrained. This is problematic because the efficiency, safety, and comfort with which the task can be performed is a function of both the robot and human's actions, and variance due to human action choices is typically large. This makes it difficult for a co-robot to offer effective active assistance.

In this section, we'll give examples of uncertainty (grasp choice, kinematic configuration), show how our individualized human models can quantify this uncertainty as a function of the constraints imposed by the robot and the task, and discuss how predictive cost functions can reduce uncertainty by predicting how humans will respond to different sets of constraints.

Formulation

We can think of a human-robot collaborative task as a two player game with a joint cost function $C(u_R, u_H)$. This function gives the total safety/ergonomic cost of the interaction as a function of the robot action u_R and the human action u_H. We'd like to choose the robot control action u_R that minimizes the cost C.

The material in this chapter is derived from A. Bestick, R. Bajcsy, and A. Dragan, "Implicitly assisting humans to choose good grasps in robot to human handovers," in *International Symposium on Experimental Robotics*, 2016.

However, this cost is also dependent on the action of the human u_H, which is not controlled by us. We know that our choice of robot action can sometimes impose constraints on the human action, so $u_H \in \mathcal{U}_H(u_R)$. The robot's action choice can also change the *probability* with which the human will choose different actions, so that $u_H \sim P(u_H | u_R)$.

This probability distribution allows us to *predict* the human's action as a function of the robot's action choice. Given this prediction, we can choose a robot action choice which minimizes the ergonomic cost of the interaction *in expectation*.

$$\min_{u_R} E_{u_H \sim P(u_H | u_R)} \left[C(u_R, u_H) \right] \tag{3.1}$$

Intuitively, we can think of $P(u_H, u_R)$ as our human model, which captures the way a human collaborator is likely to respond when presented with different interaction scenarios. This distribution can be specified explicitly, or learned from empirical observations. This formulation also nicely models scenarios where a human is performing a *greedy* optimization which would generally not yield a globally optimal value for C. Using our model, this human greediness is modeled by $P(u_H, u_R)$, and the robot's optimization presents the human with options which will incentivize them to make an action choice u_H which yields a globally optimal outcome in expectation.

3.2 Object Handovers

Object handovers are an important component of many collaborative manipulation tasks. When a co-robot helps a human to cook a meal or assemble a device in a factory work cell, it must hand objects to the human in a way which is intuitive, ergonomic, and safe. This seemingly simple task requires the control of many different variables, including handoff pose in $SE(3)$, robot grasp, and robot approach trajectory. The optimal handoff sequence may also depend on a particular human collaborator's physical preferences or capabilities, and on the parts of the task preceding or following the handover. As a result, making robot-to-human object handovers seamless has been an area of growing interest in robotics research [18, 53–61].

Imagine a co-robot helping you to unload a dishwasher. The robot hands you a mug so you can place it in the cupboard. The way the robot presents the mug to you (its pose in $SE(3)$, the grasp already occupied by the robot on the object) gives you a number of options for how to grasp it, some more natural and ergonomic than others. Ultimately, the robot's choice of handoff pose and grasp affects how *comfortable* the handover is for you, as well as how easily you can perform other tasks with the object *after* the handover.

Naturally, the robot can take this into account when planning its handover. Prior work has focused on selecting robot grasping configurations [53, 55, 57–59, 61] or object handover locations [18, 54, 56, 60] that maximize the number or range of feasible human grasps [53, 58, 59] or minimize human ergonomic cost [18, 54–57, 60, 61].

Two wrinkles make existing approaches problematic in many real-world situations though:
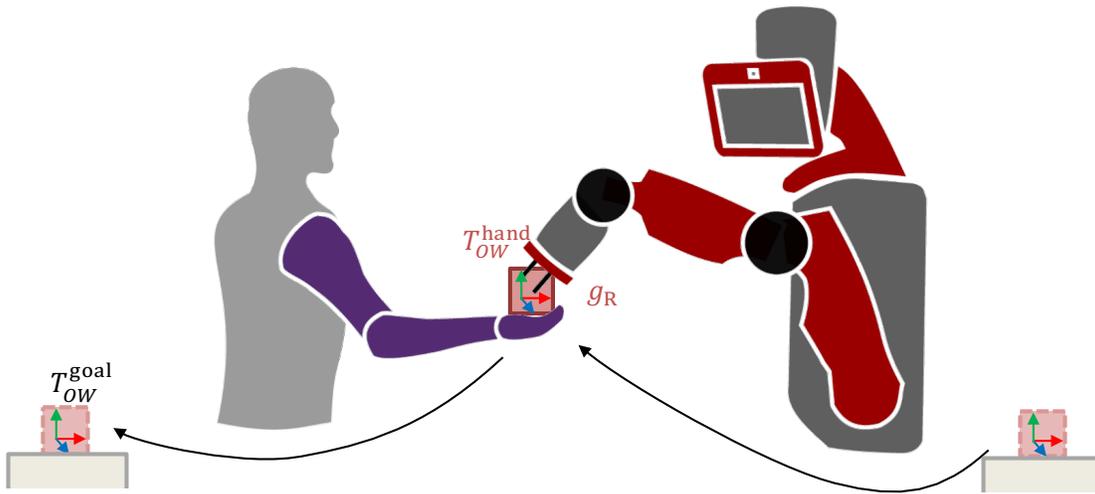
Figure 3.1: **Setup & Summary.** We focus on finding robot handover grasps and object transforms that encourage the human to select good grasps, especially when the human has a next goal for the object. We model how humans select grasping configurations, and leverage that model to minimize expected total ergonomic cost to the human.

- Humans are *myopic*: Even though the human's choice of grasp at handover may affect the ergonomics of subsequent tasks they perform with the object (e.g. placing it in a cupboard), the human will often select a grasp based on immediate comfort at the moment of handover, without regard to the ergonomic cost that choice will incur in subsequent task stages

- Humans are *noisy*: we can't control the human's actions directly, and therefore we can't guarantee that the human will select exactly the lowest cost grasp

This work makes two contributions that address these limitations: First, we select handover configurations which are optimal with respect to *total* (handover + goal) cost, given the our knowledge that the human chooses their grasp myopically. Second, we optimize this cost *in expectation*, given a probability model which assumes humans are less likely to select higher cost grasps.

These contributions have two key implications, which we'll examine in detail throughout this paper:

Encouraging the human to plan ahead

A robot typically hands a person an object so the person can do something with it (like place it at a specified goal pose). However, humans are often not very good at thinking ahead to this eventual goal: they might myopically select a grasp which requires them to contort uncomfortably or even set the object down and regrasp it to reach the goal.

Our first contribution has the effect of implicitly assisting humans to choose globally optimal grasps, by selecting handover configurations at which the most appealing grasp options for a myopic human are also near-globally optimal. Our simulation and user studies demonstrate that this implicit assistance produces more ergonomic interactions and is preferred by human collaborators. We show that giving the robot explicit knowledge of the human's myopic nature significantly improves overall task ergonomics, versus assuming the human is a global optimizer.

Reducing the risk of the human selecting a suboptimal choice

Of course, we can't control or predict exactly which grasp the human will choose in any given handover. Our human grasp selection model is probabilistic: We know only that lower cost grasps are preferred to higher cost ones. Our approach minimizes the chances of the human selecting a grasp with high total cost by maximizing the ergonomic cost incurred at handoff by humans that choose them. We do this by choosing handoff poses where suboptimal grasps require uncomfortable bending for the person to reach them, or by having the robot grasp a suboptimal grasp region itself so as to block the human from choosing it.

This behavior arises naturally from our second contribution: selecting handoff poses which have optimal *expected* total cost. This objective function penalizes handover configurations where high cost grasps are likely to be chosen, and rewards those where grasps with low global cost are likely.

The chapter is organized as follows: In Section 3.4, we describe the mathematical formulation of our problem and the solution methods we use. Section 3.5 presents the results of several simulation experiments which compare our method to existing methods in the literature (Section 3.5), test the importance of properly modeling a human's myopic and noisy characteristics (Section 3.5), and evaluates the consequences if our human model assumptions are incorrect (Section 3.5). Section 3.6 presents two user studies which compare our approach to previous methods from the literature (Section 3.6) and assess the importance of modeling the human's myopic and noisy characteristics (Section 3.6)

Our results suggest that a robot can successfully influence people to take objects in a way that makes it easier for them to achieve their goal, and that modeling people as both myopic and noisy is a "safe" assumption, in that it produces low total ergonomic cost handovers even when a human is, in reality, not myopic or noisy.

3.3 Related Work

The literature on robot-human handover planning considers a variety of approaches. These differ in whether they consider some notion of ergonomic cost, whether they optimize the full pose of the object at handoff or just the position, and whether they explicitly model the human collaborator's myopic/noisy aspects. We discuss this existing work below.

Table 3.1: Prior Handover Planning Approaches

	Feasibility Only		Ergonomic Cost	
	H Only	H + G	H Only	H + G
Position Only			[18], [54], [56], [60]	
Grasp Config.	[53]	[58], [59]	[57], [61], [55]	[52](this work)

Ergonomics and Feasibility

The majority of existing work considers the cost and feasibility of a given handoff plan only at the instant of object handoff. This is particularly true for work which considers the human ergonomic cost of a handoff [18, 53, 54, 56, 57, 60, 61]. As discussed previously, the handoff configuration chosen by the human significantly impacts the ease with which they can perform the remainder of the task after the object handoff. Thus, evaluating ergonomics at the start and goal positions of the object as well as during the handoff should confer significant ergonomic benefits upon humans working with co-robots in such tasks.

Several other existing works take this approach. In [55], the authors assess the ergonomic cost to reach the start, goal and handoff configurations, while in [58, 59] the authors consider the feasibility of the start, goal, and handoff configurations, but don't define any ergonomic cost which would allow them to choose an *optimal* configuration from the set of all which are feasible.

The prior work also differs in whether it considers the full 6 DOF human and robot grasp configurations on the object [53, 55, 57–59, 61], or only the 3 DOF positions at which the handoff will occur [18, 54, 56, 60]. Considering the full 6 DOF grasp configuration allows us to more accurately assess the feasibility and/or cost of a plan, since it's possible that the human or robot could reach a given point in the 3D workspace, but still be unable to grasp the a particular object at that point using an allowable configuration. Other prior work is summarized in Table 3.1, which categorizes the existing literature along three axes:

- Full 6 DoF grasp configuration vs. 3 DoF position only

- Optimization of some ergonomic cost vs. planning only for feasibility

- Planning only for the handover vs. for both the handover and goal

Myopic/Noisy Human Collaborators

As discussed earlier, two features of human collaborators' behavior complicate a co-robot's attempts to plan joint actions: humans are both myopic and noisy. Previous authors address these features differently. Many simply ignore the myopic/noisy nature of humans, and instead assume that humans are perfect global optimizers, always choosing the globally optimal action given the options presented to them by a co-robot. Some of our previous

Table 3.2: Prior Work with Myopic/Noisy Humans

Perfect Human	Myopic Only	Noisy Only	Myopic+Noisy
[18], [55], [56], [58], [60]		[53], [54], [61]	[52], [62], (this work)

work makes these assumptions [18]. The nondeterminism of human action choice is well known though, and some previous authors model this human noisiness explicitly [53,54,61]. Perhaps the most common of these approaches makes no attempt to predict which action choice the human will make, and instead tries only to present the human with the greatest number of possible grasp configuration options, assuming that they will make an acceptable choice if given the opportunity to do so [53,61].

We were unable to find any prior work on object handoffs or human-robot collaborative manipulation more generally that models a myopic, but deterministic human. In [62], the authors consider a robotic collaborator who is both myopic and noisy, and plan actions that lead to optimal task performance. Our earlier work [52], from which this paper is derived, considers human collaborators who are both myopic and noisy.

Existing work is summarized in Table 3.2.

3.4 Methods

To fully specify a handoff configuration, we must select the robot's grasp on the object g_R and the object pose with respect to the world frame T^{hand} at which the object will be presented to the human. We can then compute which grasp options are available to the human collaborator and compute the likelihood of the human choosing each of these options using our grasp selection model. Finally, we use this data to find the expected total handover+goal ergonomic cost incurred by the human for a given handoff configuration (g_R, T^{hand}), and to select a configuration which yields the lowest possible expected cost. The steps of this procedure are explained in detail below.

Handoff Task Model

The object to be handed off allows the human to grasp it at some set of poses $G_H \subset SE(3)$, which we represent as a collection of Task Space Regions [63], and discretize to give a finite set of feasible human grasps, so that $G_H \triangleq \{g_{H1}, ..., g_{Hk}\}$.

To compute the configurations the human can use to reach different possible object grasps, we use a human arm kinematic model combined with an analytical inverse kinematics (IK) solver. The kinematic model uses seven degrees of freedom to model the human arm, with three in the shoulder, one in the elbow, and three in the wrist. Given a handoff grasp and object pose, (g_R, T^{hand}), each possible human grasp g_{Hi} will be reachable with zero or more IK solutions, which we collect into a set $Q^{\text{hand}}_{g_{Hi}}$.

The union of these sets $Q^{\text{hand}}_{(g_R, T^{\text{hand}})} \triangleq \bigcup_{\forall i} Q^{\text{hand}}_{g_{Hi}}$ gives all the available "taking" configurations available to the human given the robot's choice of (g_R, T^{hand}).

A human grasp g_{Hi} also induces IK solutions at the object's goal pose, T^{goal}, which we collect in a set $Q^{\text{goal}}_{g_{Hi}}$.

Human Grasp Selection Model

We model the human ergonomic cost as the squared distance from some ideal nominal resting configuration q^* w.r.t. some metric W:

$$C(q) \triangleq (q - q^*)^\intercal W (q - q^*) \tag{3.2}$$

where $W = \text{diag}(w)$.

While we chose this cost function for its simplicity, it would be easy to substitute any other function which maps human limb configurations q to some scalar ergonomic cost. We would expect superior performance when using cost functions which more accurately capture the human's preferences.

We model the human as approximately-rational, selecting a grasping configuration q at handover time with higher probability when it has lower cost:

$$P(q) \propto e^{-\lambda C(q)} \tag{3.3}$$

$P(q)$ at the time of handover is normalized over all possible grasping configurations $Q^{\text{hand}}_{(g_R, T^{\text{hand}})}$. We can also compute the probability of a grasping configuration given a particular grasp, $P(q^{\text{hand}}|g_H)$, by normalizing over $Q^{\text{hand}}_{g_H}$, and $P(q^{\text{goal}}|g_H)$ at the goal by normalizing over $Q^{\text{goal}}_{g_H}$. Finally, we can compute the probability of a human grasp by summing over all the IK solutions at that grasp: $P(g_H) = \sum_{q \in Q^{\text{hand}}_{g_H}} P(q)$.

Expected Total Cost Computation

To compute the expected ergonomic cost of a given handover configuration, we must specify two different ergonomic cost functions: the actual handover+goal ergonomic cost incurred by the human C_{ergo}, and the *perceived* ergonomic cost C_{choice} the human expects to incur as a function of their grasp choice. The first cost C_{ergo} is what we seek to minimize with our choice of handover configuration, and the second cost C_{choice} is what we assume the human tries to optimize when they make their choice of grasp at handover.

The total ergonomic cost C_{ergo} is simply the sum of the ergonomic costs at the handover and goal:

$$C_{ergo} \triangleq C(q^{hand}) + C(q^{goal}) \tag{3.4}$$

The human's perceived cost considers that that, for a myopic human, the immediate cost incurred at handover may be weighted more heavily than a future cost incurred at the goal. We add a variable weight $\gamma \in [0, 1]$ to capture this possibility:

$$C_{choice} \triangleq \gamma C(q^{hand}) + (1 - \gamma) C(q^{goal}) \tag{3.5}$$

We can represent different assumptions about the human by varying γ. For example:

- $\gamma = 1$: The human is perfectly myopic, and makes action choices based solely on the immediate ergonomic cost incurred at handover

- $\gamma = 0.5$: The human is a perfect global optimizer, and balances ergonomic cost at handover with cost at the goal

- $\gamma = 0$: The human ignores the handover ergonomic cost, and chooses an action based only on the goal ergonomic cost it will result in

Optimization Procedure

When the human does not have a (known) goal, we optimize for expected cost at the handover time:

$$\min_{g_R, T_{OW}^{\text{hand}}} \sum_{q \in Q_{(g_R, T_{OW}^{\text{hand}})}^{\text{hand}}} P(q)C(q) \tag{3.6}$$

When the human does have a goal, we optimize for expected total cost. The expected cost at the goal is based on the probability of each grasp based on what happened at the handover, $P(g_H)$, and the probability of each configuration given that grasp:

$$\min_{g_R, T_{OW}^{\text{hand}}} \sum_{g_H} \left[\sum_{q \in Q_{g_H}^{\text{hand}}} P(q|g_H)P(g_H)C(q) + \sum_{q \in Q_{g_H}^{\text{goal}}} P(q|g_H)P(g_H)C(q) \right] \tag{3.7}$$

3.5 Simulation

Comparison to Heuristics

We evaluated our proposed expected total cost optimization in three simulation studies by comparing the expected total human ergonomic cost it achieved against the cost produced by two other common heuristic optimization criteria used in earlier work. The heuristic criteria were:

- $\max |Q|$: Maximize the *number* of feasible human arm configurations at the moment of handover

- $\min \frac{\sum_{q_H \in Q} C(q_H)}{|Q|}$: Minimize the *average ergonomic cost* of all feasible human arm configurations at handover, assuming that all configurations are equally likely

First we performed two case studies, which compared the expected ergonomic costs produced by our method against those from the two baseline heuristics in an object handover task using a coffee mug. In the first case study, we consider a one part task in which the

human simply takes the mug from the robot. In the second, we add a goal pose at which the human must place the mug after taking it from the robot, and the expected ergonomic cost is the sum of the expected costs at both the handoff and the goal.

Next, we conducted a simulation experiment which compared the two baseline methods' performance with our proposed method over a range of objects and object goal poses. The results of these studies are presented below.

Case Study: Expected Cost at Handover

Figure 3.2 compares optimizing for feasibility, average cost, and expected cost, in a scenario where the PR2 robot is handing over a mug to a human. For each case, we take the robot grasp and object transform that arises from the optimization, and compute: 1) the human grasping configuration of minimum cost; 2) the most "risky" human grasping configuration, that is not high-cost cost enough to be easily discarded by the human; 3) all human grasping configurations available; and 4) the histogram of costs for these configurations.

We find that maximizing the number of feasible options can be dangerous, because it might mean the expected cost is rather high, and the best configuration is not as good. Compared to minimizing average cost, we find that minimizing expected cost will allow more high-cost configurations because there is a very low probability for the human to pick them (marked "unimportant" on the histogram), but will allow fewer configurations that have good cost but not great (marked "problematic" on the histogram). These are configurations for which the probability is high enough that the human might pick them, but they are not as good as the best configurations.

Experimental Insight 1: *A robot that models human handover choices can make it more likely that the person will actually select a comfortable handover grasp.*
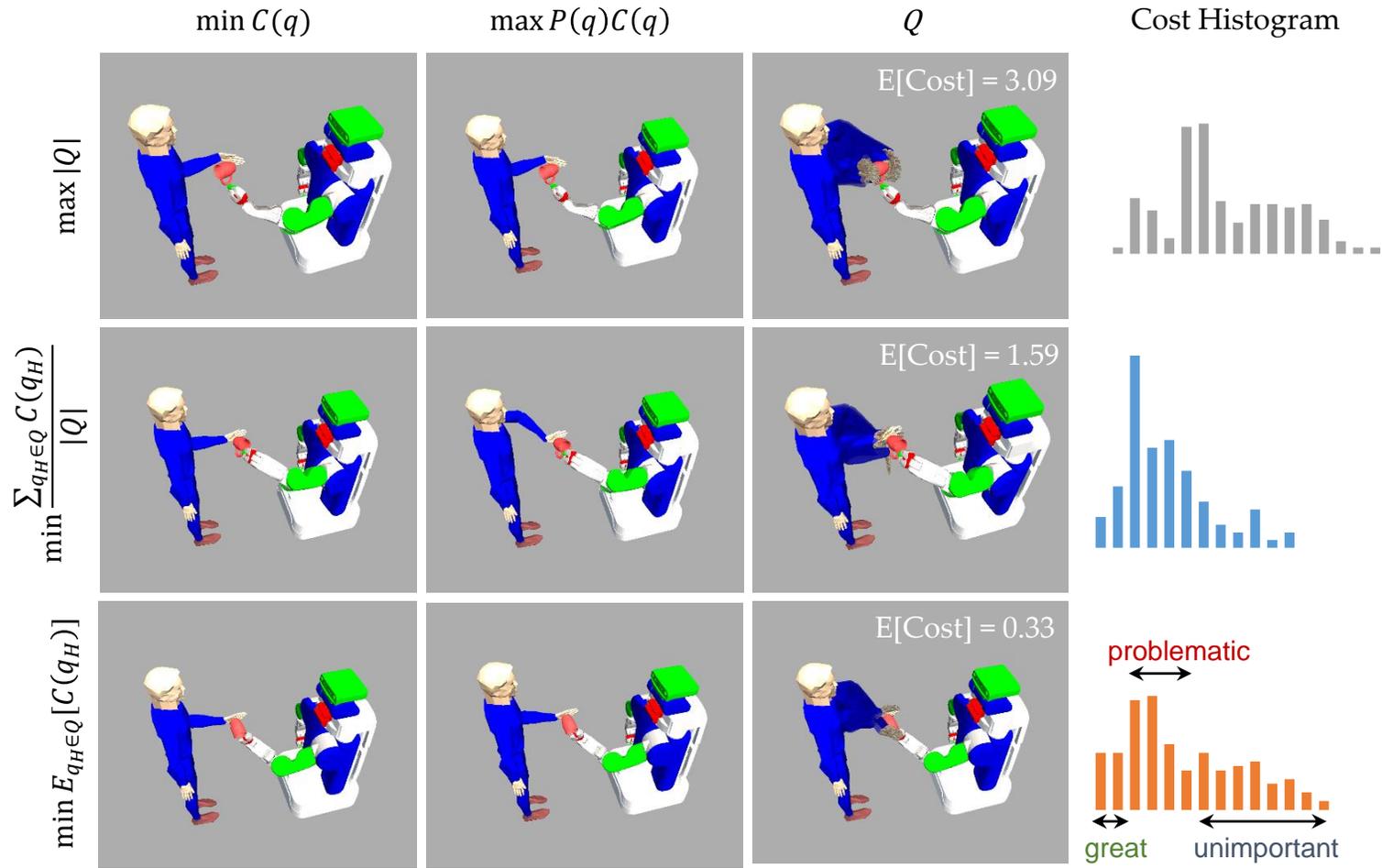
Figure 3.2: **Case Study w/o Goal - Increasing Great Choices, Reducing OK Choices, Disregarding Bad Choices.** A comparison between maximizing the number of feasible human grasping configurations Q (top), minimizing the average ergonomic cost (middle), and minimizing *expected* ergonomic cost (bottom), for the case of a single handover without a known object goal pose. The columns show the most probable human configuration (left), the configuration with the largest contribution to the total cost (middle), and the full space of configurations (right). Our method increases the number of great choices and decreases the number of OK choices which the human might actually pick. It also keeps bad choices if needed, because they have a low probability of being selected anyway.

Case Study: Expected Total Cost (Goal + Handover)

Figure 3.3 compares the three approaches from above when accounting for the human goal. Feasibility here accounts for the number of feasible configurations at both the handover and the goal, average cost accounts for cost at the start and goal, and so does expected cost. For each case, we take the resulting robot grasping configuration and compute 1) the human grasping configuration of minimum handover cost, which is what the human will most likely choose if they are being myopic; 2) given this grasp, the configuration of minimum cost at the goal (assuming no regrasp); and 3) the expected cost at the handover and at the goal for each human grasp.

We find that maximizing feasible options can lead to very poor options at the goal. Compared to minimizing average cost, we find that minimizing expected cost is better at eliminating grasps that have low cost at handover time but only allow for high cost at the goal.

**Experimental Insight 2:** *A robot that models human handover choices can make it more likely that the person will select a handover grasp that also allows for comfortably achieving the goal after the handover.*
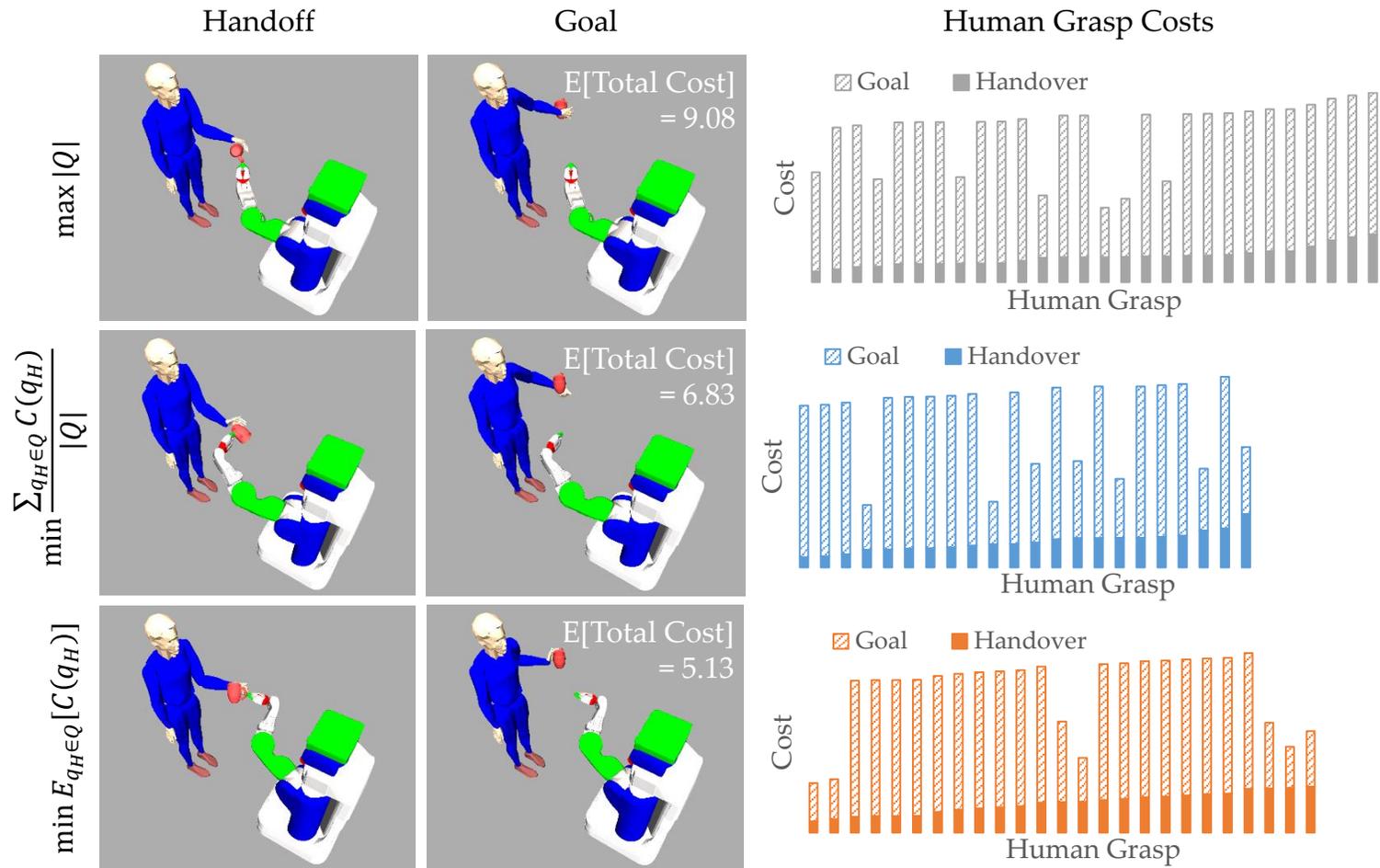
Figure 3.3: **Case Study w. Goal - Reducing Total Cost.** A comparison between maximizing the number of feasible human grasping configurations Q at the handover and goal (top), minimizing the average ergonomic total cost (middle), and minimizing *expected* total cost (bottom). The columns show the most probable human configuration at handover time (left), and at the goal (center), along with a plot of cost for each available grasp to the human. Our method makes it such that the tempting configurations (low cost at handover) also have low cost at the goal.

Simulation Study

Our case studies used a single object and a single goal configuration. Here we expand to an experiment that manipulates both as factors.

Manipulated Factors: We manipulate three factors. The first is the *metric* we optimize, as in the case study: maximizing number of feasible options, minimizing average cost, or our metric, minimizing expected cost. The second is the *object* being handed over by the robot: a mug as before, a glass, a pitcher, and a plate, for a total of 4 objects. These objects have vastly different TSR choices. The third is the *goal pose*, for which we use 5 different poses. This leads to a total of 3(metrics) x 4(objects) x 5(goals) = 60 conditions.

Dependent Measures: As in the case studies, we measure expected total cost.

Hypothesis: Our metric is designed to optimize expected total cost (the dependent measure), so we already know it will perform the best. The question remains whether our metric will be better by a significant margin. Our hypothesis is that it will: *Our metric will result in a significant improvement in expected cost compared to the baselines.*

Figure 3.4: **Optimal Handover for Different Goal Poses.** The different goal poses in our experiment lead to different optimal handover configurations for the robot, each selected to minimize expected total cost at the handover time *and* at that particular goal. The chart averages the expected total (handoff + goal) ergonomic costs for each of the three metrics.

Analysis: We ran an ANOVA with metric as a factor to test differences among the three metrics across objects and goal poses. We found a significant main effect, $F(2, 58) = 1031.07$, $p < .0001$. A post-hoc analysis with Tukey HSD showed that all three metrics were significantly different from each other, with the average cost outperforming maximum feasibility ($p < .0001$) and our metric outperforming average cost ($p < .001$), in line with our hypothesis.

Fig. 3.4 shows how the robot's grasping configuration changes as the goal pose for the human changes. The robot will present the mug so that the person grabs it by the top
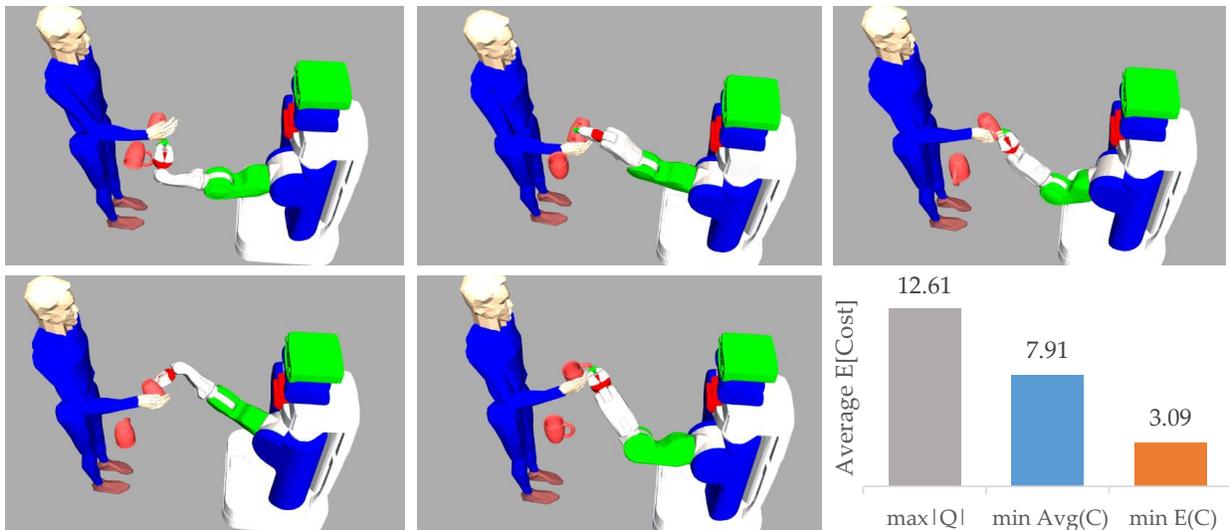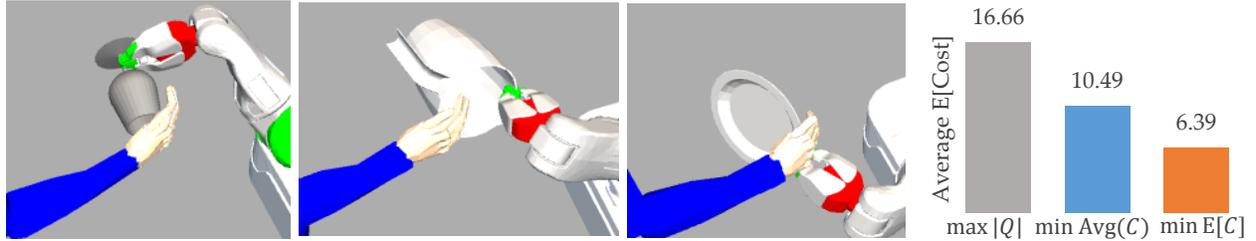
Figure 3.5: **Optimal Handover for Different Objects.** The different objects in our experiment lead to different optimal handover configurations for the robot for a given goal. The chart averages the expected total (handoff + goal) ergonomic costs for each of the three metrics across objects.

when it needs to be placed right side up, by the side when it needs to be placed upside down, etc. In line with our hypothesis, the expected cost was three times lower with our approach compared to the maximum feasibility baseline, and two times lower compared to the minimum cost baseline.

Fig. 3.5 shows how the robot's grasping configuration changes, for a given goal pose, as the object changes. The robot holds the objects in different ways to ensure that the person can easily grasp them by the side and set them down vertically with ease.

Importance of Myopic/Noisy Model

The previous section's results make clear that our expected ergonomic cost minimization produces consistently lower cost interactions then previous heuristic criteria. Let's examine the modeling assumptions made by the expected cost minimization more closely though. If we know that a real human is myopic/noisy, how important is it for our model to capture those features of the person's behavior? What would happen if we made different assumptions, like assuming that the human was, in fact, a global optimizer? We conducted an additional simulation experiment to investigate this question.

Simulation Study

Our simulation study used a single object and varied the object goal pose. Our simulated human was both myopic and noisy, and we tested the expected ergonomic cost incurred by the human when the robot explicitly modeled the human's myopia/noisiness versus when it was unaware of these features.

Manipulated Factors: We manipulate two factors. The first is the goal pose at which the simulated human must place the object after taking it from the robot. The second is the robot's assumption about how the human chooses their actions. The four assumptions tested were:

- Myopic/noisy (MN): Robot's model assumption accurately reflects the true simulated human

- Myopic/perfect (MP): Robot believes the human is myopic, but deterministically selects the action with the lowest perceived cost

- Global/noisy (GN): Robot believes the human is a global optimizer, but accurately models their nondeterminism

- Global/perfect (GP): Robot believes the human deterministically optimizes the total cost for handover+goal (i.e. the human always chooses the true optimal action)

This gives a total of 10(goal poses) x 4(model assumptions) = 40 conditions.

Dependent Measures: As before, we measure expected total cost.

Hypothesis: We know that giving the robot a correct model (MN) of the simulated human will result in the lowest possible expected total cost. The question is how much higher the expected total cost will be if the robot's assumptions are incorrect (MP, GN, or GP). Our hypothesis is that it will be significantly higher: *Correctly modeling the myopic/noisy nature of the human will result in expected ergonomic costs that are significantly lower than those achieved with incorrect model assumptions.*

Analysis: Figure 3.6 shows how the expected ergonomic cost incurred by the human changes as the robot's human model assumptions are varied. The correct model assumption (MN, shown in orange) yields expected ergonomic costs that are significantly lower than with any of the incorrect assumptions. This supports our hypothesis.
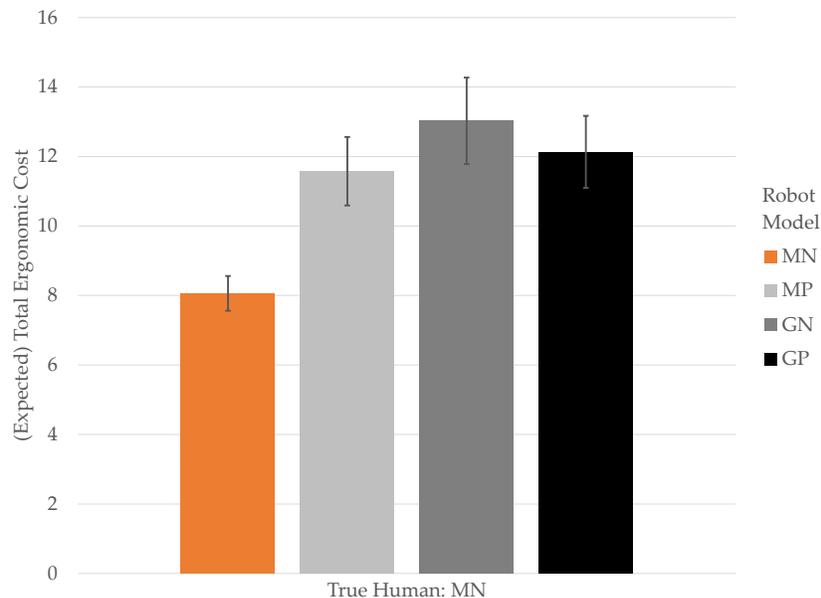
Figure 3.6: **Expected Human Ergonomic Cost vs. Robot Human Model for MN Human:** Empirically, humans display both myopic and noisy characteristics in ergonomic optimization tasks such as handovers. If a human is truly myopic-noisy, then giving the robot a human model which omits one or both of these characteristics (grey and black, respectively) causes the human to incur a significantly larger total ergonomic cost compared to a robot with an accurate human model (orange). Note that this figure shows a selected subset of the data in Figure 3.7.

Losses From Model Inaccuracy

The previous section's results demonstrate that for a myopic/noisy human, using model assumptions that accurately reflect both the myopic and noisy features of the human's behavior is important for minimizing ergonomic cost. What if the true human is not myopic/noisy though? Are the same myopic/noisy modeling assumptions appropriate even if the real human deviates from them? In this section, we measure the loss in ergonomic performance when the true human does not conform to the robot's model.

Simulation Study

Our simulation study used a single object and varied the object goal pose. We varied whether the simulated human was myopic and noisy, and we tested the expected ergonomic cost incurred by the human when the robot's model assumed myopia/noisiness versus when it did not.

Manipulated Factors: We manipulate three factors. The first is the goal pose at which the simulated human must place the object after taking it from the robot. The second is the robot's assumption about how the human chooses their actions. The third are the characteristics of the true human. Both the robot's model assumptions and the true human's characteristics can take one of four values (MN, MP, GN, or GP, as described in the previous section). This gives a total of 10(goal poses) x 4(model assumptions) x 4(true human characteristics) = 160 conditions.

Dependent Measures: As in the previous simulation studies, we measure expected total cost.

Hypothesis: We know that giving the robot a correct model (MN) of the simulated human will result in the lowest possible expected total cost when the human is, in fact, myopic/noisy. Our question here is whether this MN assumption is still a "safe" choice even when the true human is not myopic/noisy. Our hypothesis is that the losses due to model inaccuracy will be small when the robot assumes an MN human: *Modeling the human as myopic/noisy will produce expected ergonomic costs which are near optimal even when the true human is not MN.*

Analysis: Figure 3.7 shows the mean expected ergonomic cost incurred by the human for each of the 16 possible combinations of true human/robot model assumption. Our proposed modeling assumption, myopic/noisy, is shown in orange.

In the case where the human is, in fact, myopic/noisy (leftmost cluster of bars), we see that modeling this myopic/noisy nature explicitly produces significantly lower expected ergonomic costs than any of the other possible model assumptions (grey and black bars).

Next, examine the two cases where the human is actually a global optimizer (GN and GP, the second and fourth clusters of bars, respectively). In both cases, using the correct model assumption (GN and GP, respectively) produces the lowest expected ergonomic cost. This is what we'd expect, since the robot will always be able to optimize most accurately when it has a complete, correct model of the human's decision making process. The differences in ergonomic cost between the correct model assumptions and our proposed MN assumption is statistically significant in both cases, but is also very small in absolute magnitude, particularly when compared to the MN-GN and MN-GP expected cost differences when the the human is truly MN. In effect, there is a large penalty for failing to explicitly model the human's myopia when the human is truly myopic, but only a small penalty for assuming the human is myopic when they are really a global optimizer.

Finally, examine the case where the true human is myopic/perfect (myopic, but noiseless). We see that the correct model assumption, MP, yields an expected ergonomic cost that is significantly lower than any of the other alternatives. Our MN assumption is second best. This suggests that modeling the human's true degree of noisiness is more important for the system's ergonomic performance than it is in the case of myopic, where we can just assume a myopic human and get close to optimal performance even if the human is not perfectly myopic.
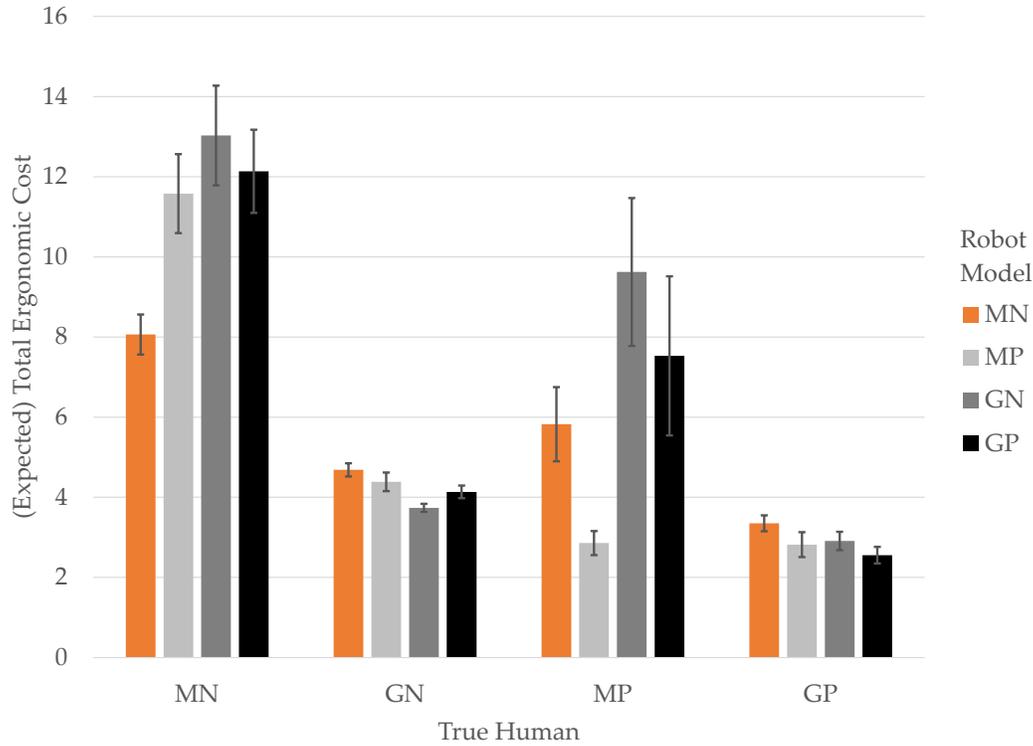
Figure 3.7: **Expected Human Ergonomic Cost vs. Robot Human Model for Different Humans:** It is plausible that, depending on the specific scenario and person, humans may be less myopic or less noisy than we expect. In general, our robot's human model will never match the exact degree of myopia or noisiness displayed by an individual human. It's therefore important to consider what the loss will be due to this inaccuracy, and to choose a "safe" human model which which produces acceptable results for a range of actual human characteristics. Modeling the human as myopic-noisy (orange) produces consistently incurred ergonomic costs for a range of actual human characteristics, which other models (grey/black) produce significantly higher costs for at least one of the four human characteristics we consider.

We can examine the importance of each piece of our myopic/noisy assumption in more detail by separating our analysis into two pieces, the first focusing only on the impact of the myopic assumption, and second only on the noisy assumption.

Figures 3.8 and 3.9 show the same data as in 3.7, but for only one of the model assumptions at once. Figure 3.8 shows the impact of the myopic/global assumption across all trials. We can clearly see that when the human is truly myopic, modeling that myopia produces a large drop in expected ergonomic cost. Conversely, when the human is not myopic, modeling them as myopic anyway produces only a small loss. This suggests that making a blanket

assumption that humans are myopic is well justified for this handoff task.

Figure 3.8: **Expected Human Ergonomic Cost vs. Actual Human/Robot's Model for Myopic/Global Only:** Average ergonomic costs over all trials in which the human/robot were myopic/global, respectively.

We can perform the same analysis for the noisy assumption. Figure 3.9 shows the impact of only the noisy/perfect assumption across all trials. The results here are more ambiguous. The model assumption which matched the true human's noisy/perfect characteristic produced the best performance, but this difference was not significant for either the noisy or the perfect true human. The absolute magnitudes of the ergonomic cost differences were similar, but slightly larger for the case where the true human was perfect. This means there isn't an obvious default choice for the noisy/perfect assumption, in the case where we don't know the human's true characteristics.
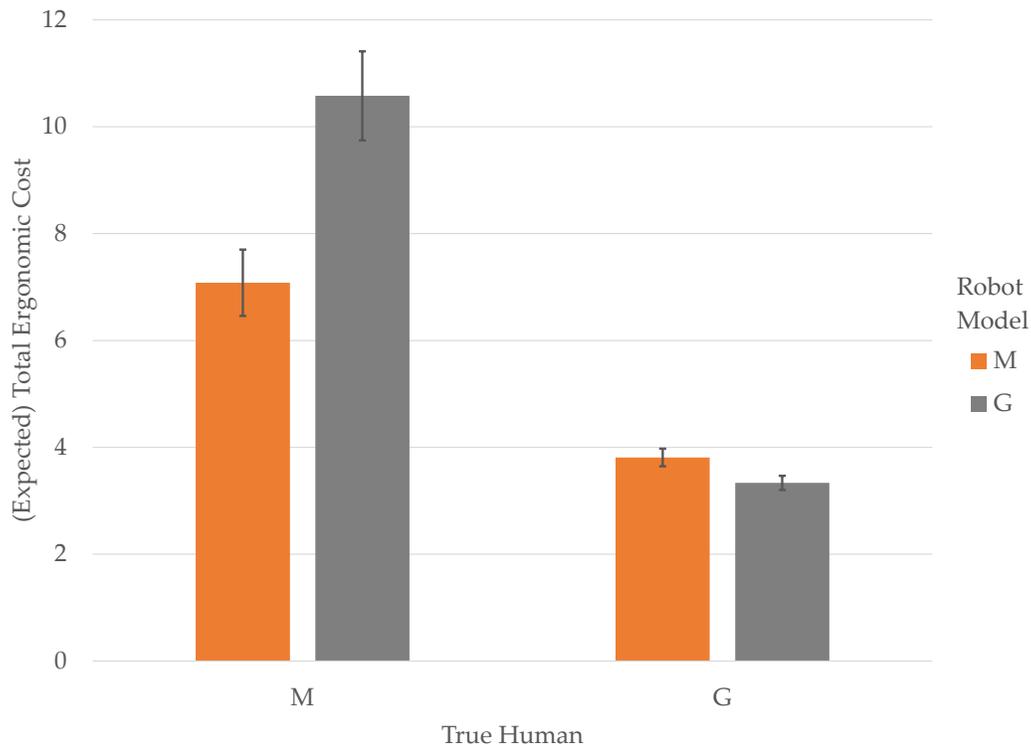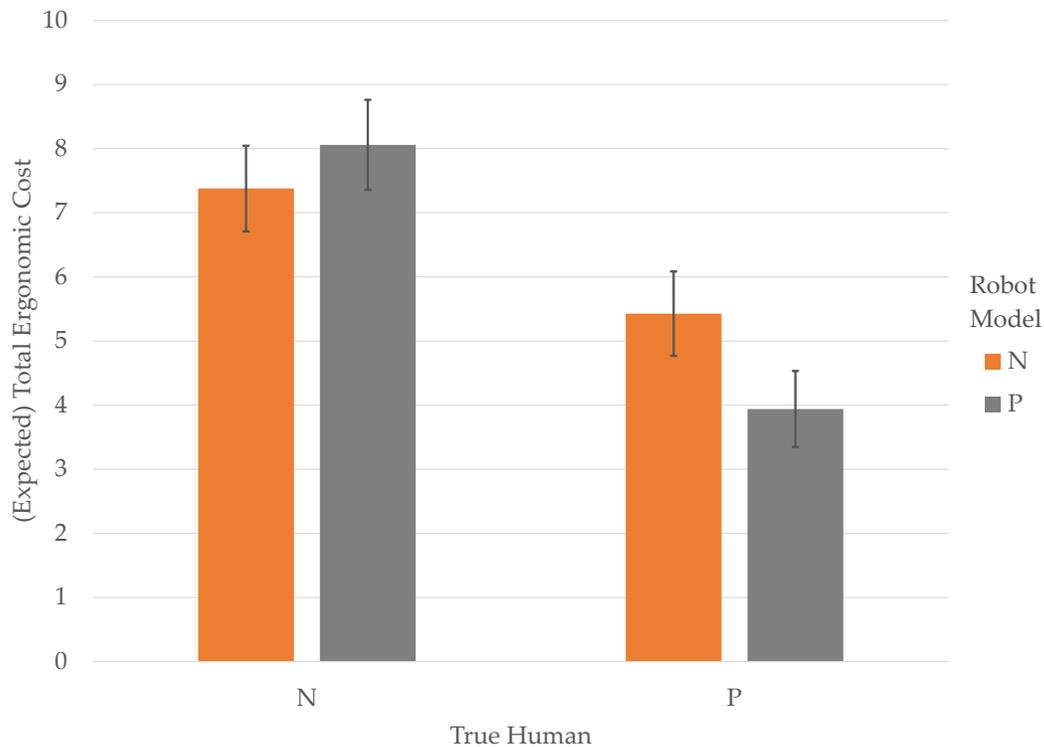
Figure 3.9: **Expected Human Ergonomic Cost vs.   Actual Human/Robot's Model for Noisy/Perfect Only:** Average ergonomic costs over all trials in which the human/robot were noisy/perfect, respectively.

3.6   Human Experiments

Our simulation results above demonstrated that our expected total cost optimization yields improved ergonomic performance versus both the heuristics used in previous handover planning work, and versus alternate model assumptions about human preferences. In this section, we further validate these conclusions.

We present two user studies: The first compared our expected total cost minimization method to the maximum feasible options baseline heuristic from the literature. The second compared the ergonomic performance of our proposed myopic/noisy human model assumption with an alternate set of assumptions (global/perfect), in the expected cost minimization framework.

Our Model vs. Heuristics

The previous sections tested our method in simulation, assuming users who act according to our model. Real people do not. We conducted a user study to test whether the simulation results generalize, and to explore whether users perceive the improvement brought about by our method.

Experimental Design

Manipulated Factors. We manipulated three factors. We manipulated the *metric* the robot used to compute its handover configuration, using our metric based on the user model we proposed, $\min E[C]$, and the maximum feasibility baseline $\min |Q|$.

We used the mug as the handover object for this experiment, and manipulated the *goal pose* using 10 different poses. In these poses, the mug was placed upside down, upright, and to the side to ensure variance.

Finally, we manipulated *whether the user knows the goal* (Fig. 3.10). We did this because we wanted to separate the two assumptions our method is making: that users select grasping configurations based on ergonomic cost, and that users are myopic or greedy in this selection, only accounting for ergonomic cost at handover time but not at the goal. Therefore, manipulating the user's knowledge of the goal enables us to test not only how our method performs overall (in realistic situations in which users have a goal and are aware of it), but also whether our method is influencing the users' grasp choice in the way we expected, assuming users are actually myopic (which in reality might or might not be the case). Altogether, this led to 2(metrics) x 10(goals) x 2(knowledge) = 40 conditions.

Subject Allocation. We recruited 9 users (6 male, 3 female, ages 22-29). All of the factors were *within-subjects*, meaning each user experienced all conditions. We counterbalanced the order of the metrics to avoid order effects, and randomized the order of the goals. We split the experiment in two parts, the first in which the user did not know the goal, and the second in which they did:

In Part 1, the robot handed the object to the person at each of the 20 optimal handover configurations (one for each metric and goal pose), but the user was not told the goal used by the planner. We instructed the user to take the object from the robot and immediately drop it in a box. This ensured that no notion of a goal pose would impact the subject's choice of object grasp. This portion of the experiment evaluated the two algorithms' ability to influence the subjects to select a particular grasp when the subject was not aware of a goal, i.e. when the myopic/greediness assumption holds.

In Part 2, a pictorial marker was placed on a table next to the subject indicating the object's goal pose during each handoff. The subject was told that two different algorithms, "Program 1" and "Program 2," would be used during this part of the experiment. We conducted handovers at the same 20 configurations as before, but this time the subject was instructed to place the object at the indicated goal pose. We told the users before each handover which of Programs 1 and 2 was in use. This portion of the experiment evaluated

the algorithm's ability to influence people to select ergonomically optimal grasps even when they know the goal, i.e. they are not necessarily myopic. Furthermore, it enabled us to ask users to compare the two methods, seeing if their notion of comfort matches ours. If people are actually myopic about the goal when selecting a grasping configuration, then we expect results for this second part to match those from the first part.

In both parts, the subject was instructed to use only their right hand to manipulate the objects, and to not change or reposition their grasp before dropping the object or placing it at the indicated goal position.
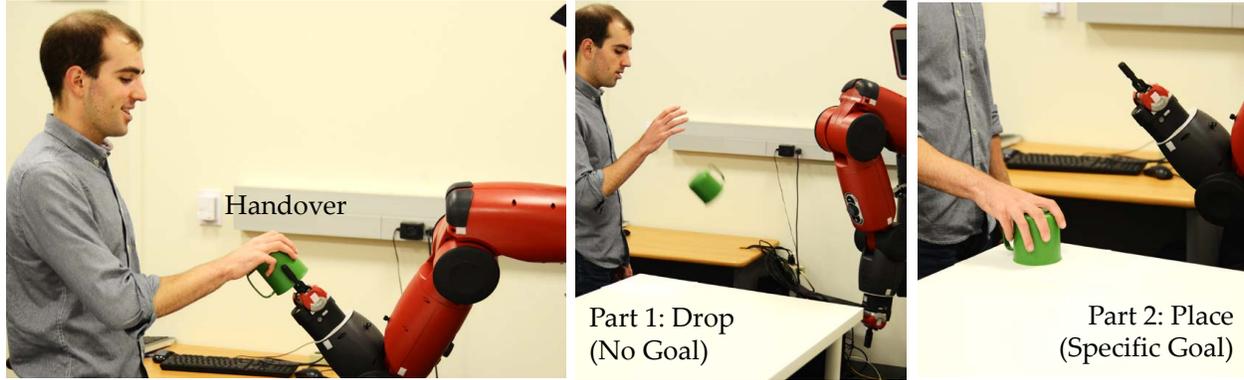

Figure 3.10: **User Study Setup.**

Dependent Measures. We used both objective and subjective measures.

Objective: We annotated for each condition which of the 6 TSRs for the mug the person selected. From this, we computed expected cost over all IK solutions at the goal, for all grasps that were feasible at handover time (i.e. had feasible IK solutions), making 2 assumptions: 1) the person follows our ergonomic model, and 2) we know the human kinematics:

$$\text{OM1: } E[C(q^{goal})], \lambda = 10, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR feas. at handover} \qquad (3.8)$$

To alleviate bias in our results induced by the two assumptions, we introduce 3 additional metrics that break each assumption separately as well as both assumptions together: we break the first assumption by computing average cost (which is the expected cost using a uniform distribution, i.e. $\lambda = 0$) instead of expected ergonomic cost , and we break the second assumption by allowing infeasible grasps that a person with different kinematics might have chosen:

$$\text{OM2: } E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR feas. at handover} \qquad (3.9)$$

$$\text{OM3: } E[C(q^{goal})], \lambda = 10, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR} \qquad (3.10)$$

$$\text{OM4: } E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR} \qquad (3.11)$$

We did not estimate cost at handover time, because we were specifically interested in whether the robot successfully influences users to select grasps that are good at the goal.

Indeed, we might see lower handover time costs for the baseline condition because it restricts the users less.

Subjective. After each complete experiment, the subject answered a series of 1-7 Likert-scale survey questions about which program they preferred, which program made their goal easier to accomplish, and which program inspired the most trust in the robot. These capture each subject's subjective opinion about which metric was more effective at making interaction with the robot comfortable and effective.

Hypotheses.

H1. *IF humans are actually myopic when selecting grasping configurations (e.g. when they are not even aware of the goal), our method successfully influences them to select configurations with lower cost at the goal compared to the baseline.*

H2. *Our method influences people to select configurations with lower cost at the goal compared to the baseline, even when they are aware of the goal.*

H3. *People prefer to work and are more comfortable with a robot using our method compared to the baseline.*

Analysis

H1. We used results for part 1 of the study, when users are not aware of the goal, to test H1. We first computed Cronbach's α for the four objective measures, which was high at .9036. We thus computed an aggregate goal cost using all four measures.

We then ran a repeated-measures factorial ANOVA on this aggregate, with goal and metric as factors. We found a significant main effect for metric, as expected ($F(1, 179) = 377.83$, $p < .0001$), and a significant main effect for goal ($F(9, 171) = 26.79$, $p < .0001$). However, there was also a significant interaction effect, and so we conducted a Tukey HSD post-hoc, comparing all pairs but compensating for multiple comparisons. The analysis revealed that the expected cost (our) metric resulted in significantly lower cost at the goal than the baseline for 7 out of the 10 goals, all with $p < .03$.

This supports our hypothesis H1, but suggests that the benefit of our method does depend on the choice of the goal pose, with the maximum feasibility baseline being sufficient for some goals.

Table 3.3 shows the goal ergonomic costs estimated by each of the four measures, averaged across all nine study participants for this part of the study. It shows that pose optimization with $\min E[C]$ gives consistently lower ergonomic cost at the goal than optimization with $\min |Q|$. This difference is particularly marked for the first two measures, which consider only grasps feasible at the handover. These results suggest that expected ergonomic cost can be used to influence humans to choose grasps with good ergonomic properties even when they are completely unaware of the goal.

Table 3.3: Estimated Human Ergonomic Costs at Goal (Part 1: Users not aware of goal)

Objective Measure	min $|Q|$	min $E[C]$
$E[C(q^{goal})], \lambda = 10, \forall q_{goal} \in \mathrm{IK(g)}, \forall g \in \mathrm{TSR}$ feas. at handover	12.43	6.02
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \mathrm{IK(g)}, \forall g \in \mathrm{TSR}$ feas. at handover	12.41	6.30
$E[C(q^{goal})], \lambda = 10, \forall q_{goal} \in \mathrm{IK(g)}, \forall g \in \mathrm{TSR}$	12.18	11.26
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \mathrm{IK(g)}, \forall g \in \mathrm{TSR}$	12.28	11.45
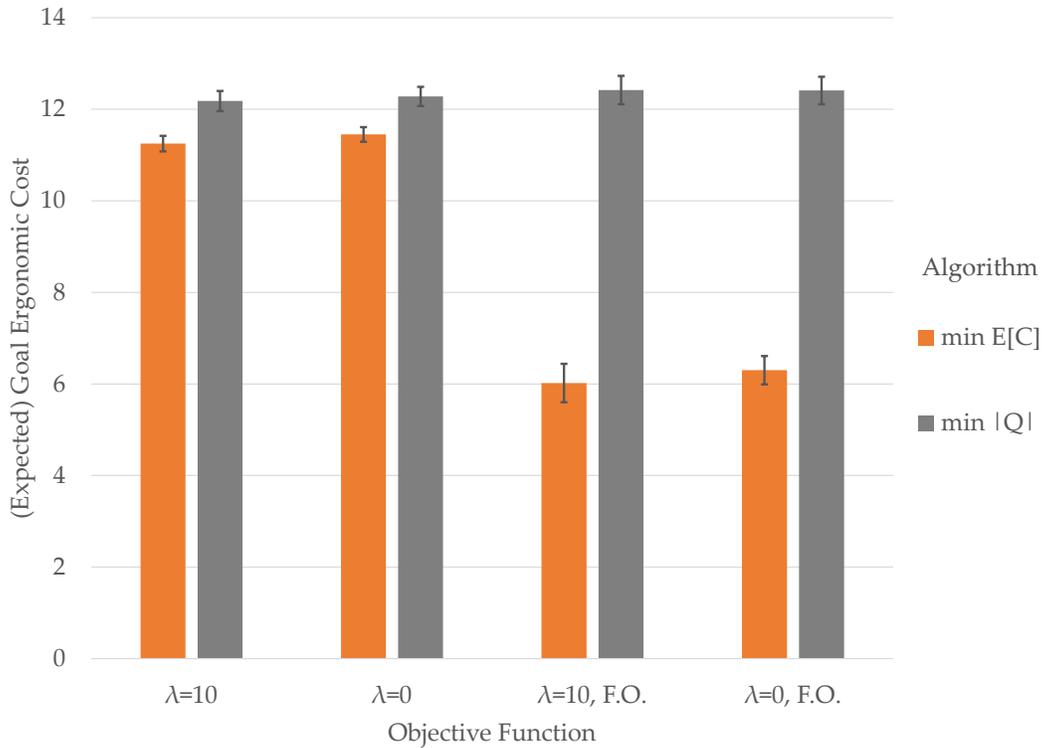
Figure 3.11: **Estimated Goal Ergonomic Costs vs. Algorithm - Part 1:** Expected goal ergonomic costs for trials where the user *was not* aware of the final goal (i.e. a perfectly myopic human). Expected costs were computed according to the four hypothesized human objective measures described in Section 3.6). For all objective functions, planning using our minimum expected cost algorithm (min $E[C]$) yielded significantly reduced goal ergonomic costs relative to planning with the baseline algorithm (min $|Q|$).

H2. For part 2, when users were given specific goals, our objective measures again had high item reliability, Cronbach's $\alpha = .8830$. We again computed an aggregate cost. We again

ran a factorial repeated-measures ANOVA, and the results, as expected, were analogous to the results from part 1. We again saw significant main effects, but also a significant interaction between the factors. As before, a post-hoc with Tukey HSD corrections showed that 7 out of the 10 goals saw significantly lower costs at the goal with our method than with the baseline. The set of these 7 goals was almost identical to the one in part 1, with the exception of one goal no longer showing a significant difference, and one goal starting to show a significant difference.

This supports out hypothesis H2: ***our method does not only help users improve performance when we force them to be myopic by not making them aware of the goal – it helps in realistic situations, when users have a goal that they are aware of***. This suggests that people are indeed myopic in their selections of a grasp configuration.

Table 3.4 shows the goal ergonomic costs estimated by each of the four objective metrics, averaged across all nine study participants for Part 2 of the study, where subjects were instructed to place the object on a pictoral marker at the goal pose after each handover. We see a similar improvement in ergonomic costs when minimizing $E[C]$ versus maximizing $|Q|$.

Here, we found it interesting that the costs dropped slightly across the board. ***This suggests that perhaps when people are aware of the goal they perform slightly better, but that still our method can significantly help them to further improve their performance.***

Table 3.4: Estimated Human Ergonomic Costs at Goal (Part 2: Users aware of the goal)

| Objective Measure	$\min |Q|$	$\min E[C]$
$E[C(q^{goal})], \lambda = 10, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR feas. at handover}$	11.42	5.37
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR feas. at handover}$	11.52	5.61
$E[C(q^{goal})], \lambda = 10, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$	11.72	11.02
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$	11.84	11.23
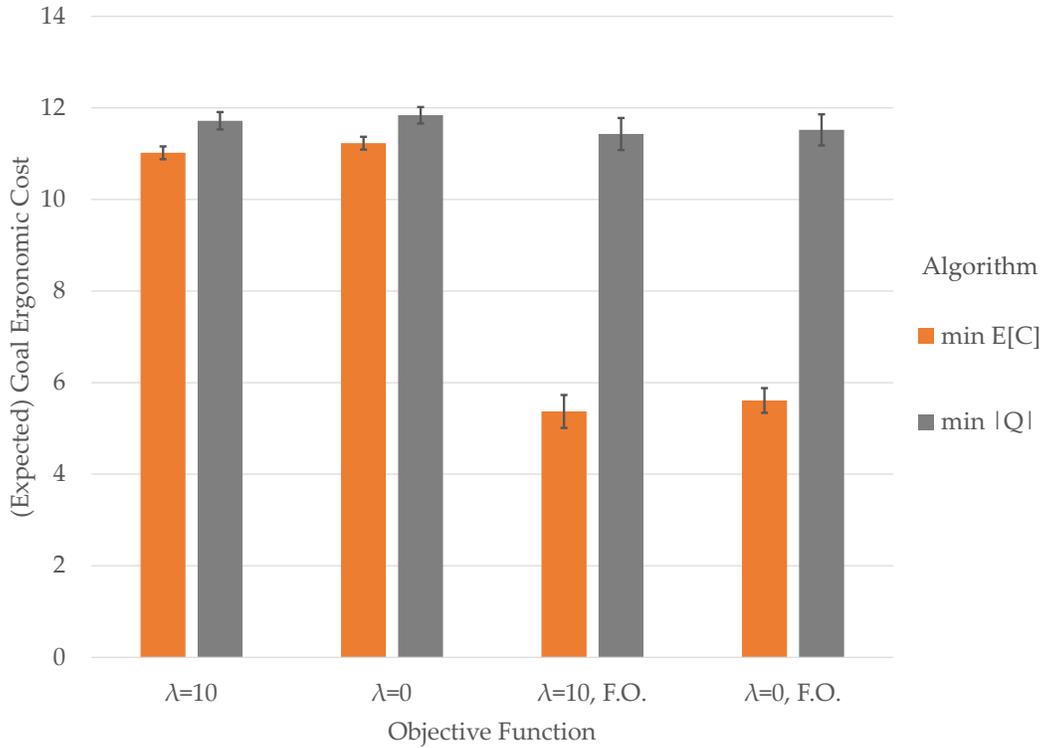
Figure 3.12: **Estimated Goal Ergonomic Costs vs. Algorithm - Part 1:** Expected goal ergonomic costs for trials where the user *was* aware of the final goal (i.e. human not necessarily myopic). Expected costs were computed according to the four hypothesized human objective measures described in Section 3.6). Even when the human was aware of the goal, planning using our proposed min $E[C]$ algorithm produced a significant reduction in expected goal ergonomic cost versus the baseline min $|Q|$ algorithm.

H3. Table 3.5 summarizes users' subjective ratings. *t*-tests showed that **_our method outperformed the baseline in user overall preference, how helpful they thought the robot was, how much they trusted the robot, and how easy it was to do the task. Users thought that the robot understood their goal and that it handed them objects in a way that made their task easier._**

The users' comments were particularly enlightening (here Program 1 refers to the baseline and Program 2 refers to our method):

"With Program 2, I could move straight from grip to the target with a natural motion. With Program 1, I would sometimes have to contort my arm unnaturally to place the mug correctly."; "Program 1 made it easier to pick up objects but harder to achieve the goal.

Table 3.5: Post-Study Survey Results

Statement	min $|Q|$	min $E[C]$	$t(8)$	p
"I prefer Program __"	2.0	6.2	9.73	$<.0001$
"The robot was helpful when running Program __"	3.4	6.4	6.80	$<.0001$
"I trust the robot running Program __"	3.7	6.1	4.4	$<.01$
"The robot understood my goal when running Program __"	2.8	6.4	5.33	$<.001$
"It was physically easy to do the task when the robot was running Program __"	2.8	6.2	6.50	$<.001$
"The robot running Program __ handed me objects in a way that made the task easier"	2.0	6.3	9.19	$<.0001$
"If you had to choose a program you prefer, which would it be?"	0%	100%	-	-

Program 2 sometimes made it more difficult to pick up objects but achieving the goal was easier."; "Program 1 is an a**hole."

Myopic/Noisy vs. Global/Perfect Assumptions

The user study in the previous section demonstrated how our expected total cost minimization algorithm produced more ergonomic handovers than the maximum options heuristic. The expected total cost method relies on a model of human action choice though, and the specific model used is likely to have a large impact on the algorithm's performance. In the previous section we modeled the human as a myopic, noisy optimizer. In the next user study, we compare this myopic/noisy assumption to an alternative global/perfect one.

Experimental Design

The experimental design in this study was similar to the previous one, with two modifications:

- We compared the MN and GP assumptions, both with the min $E[C]$ algorithm, rather than comparing the min $E[C]$ to the max $|Q|$ criterion

- Instead of the goal pose for the mug being on a table, it was in a shelf which made certain grasps on the mug infeasible, as shown in Figure 3.13.

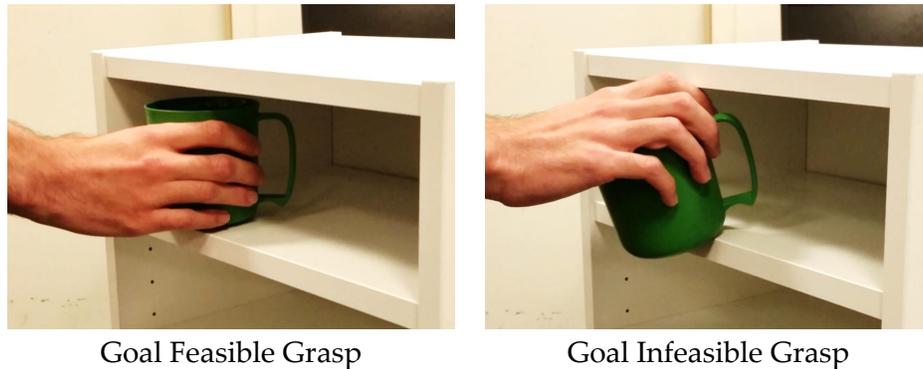Goal Feasible Grasp   Goal Infeasible Grasp

Figure 3.13: The goal poses for the MN vs. GP user study were located in a shelf which restricted the grasps on the mug which were feasible. The picture on the left shows a feasible human grasp, while the right picture shows a grasp which is infeasible at the goal due to the shelf. OM5 counted the total number of goal infeasible grasps chosen by each human test subject.

Manipulated Factors. We manipulated three factors. We manipulated the *model assumptions* the robot used to compute its handover configuration using our min $E[C]$ algorithm.

We used the mug again as the handover object for this experiment, and manipulated the *goal pose* using 10 different poses. As before, the mug was placed upside down, upright, and to the side to ensure a representative range of scenarios. As noted above though, the goal poses were now located in a shelf which restricted the range of feasible grasps available to the human at the goal.

Finally, we manipulated *whether the user knows the goal* (Fig. 3.10). As in the previous user study, this allows us to isolate the portion of the user's action choice that is due to knowledge of the goal from that which is based only on knowledge available at the handoff. Stated differently, we could say that the scenario in which the user is not aware of the goal simulates a perfectly myopic human. Altogether, this led to 2(model assumptions) x 10(goals) x 2(knowledge) = 40 conditions.

Subject Allocation. We recruited 10 users (7 male, 3 female, ages 22-27). All of the factors were *within-subjects*, meaning each user experienced all conditions. We counterbalanced the order of the model assumptions to avoid order effects, and randomized the order of the goals. We split the experiment into two parts, the first in which the user did not know the goal, and the second in which they did:

In Part 1, the robot handed the object to the person at each of the 20 optimal handover configurations (one for each model assumption and goal pose), but the user was not told the goal used by the planner. We instructed the user to take the object from the robot and immediately drop it in a box. This ensured that no notion of a goal pose would impact the subject's choice of object grasp. This portion of the experiment evaluated the two algorithms' ability to influence the subjects to select a particular grasp when the subject was not aware of a goal, i.e. when the myopic/greediness assumption holds.

In Part 2, as before, a visual marker was placed on the shelf to illustrate the goal pose for the mug. The subject was told that two different algorithms, "Program 1" and "Program 2," would be used during this part of the experiment. We conducted handovers at the same 20 configurations as before, but this time the subject was instructed to place the object at the indicated goal pose. We told the users before each handover which of Programs 1 and 2 was in use. This portion of the experiment evaluated the algorithm's ability to influence people to select ergonomically optimal grasps even when they know the goal, i.e. they are not necessarily myopic. Furthermore, it enabled us to ask users to compare the two methods, seeing if their notion of comfort matches ours. If people are actually myopic about the goal when selecting a grasping configuration, then we expect results for this second part to match those from the first part.

In both parts, the subject was instructed to use only their right hand to manipulate the objects, and to not change or reposition their grasp before dropping the object or placing it at the indicated goal position.

Dependent Measures. We used the same objective and subjective measures as in the previous user study, and added one additional objective measure: number of goal infeasible human grasps. We manually annotated each trial with which of the 6 TSRs the human grasped, then estimated the incurred goal ergonomic cost using each of the four objective measures OM1-OM4 used previously.

As before, we estimated ergonomic cost only at the goal, not the handover.

Objective. In addition to OM1-OM4, we added OM5, the *number of goal infeasible grasps.* This is the number of grasps chosen by the human at handover which will be infeasible at the goal (because there's no room to fit a human hand at that location on the mug). If the human was truly a global optimizer, this measure would be equal to zero for all conditions, since the human would foresee and avoid infeasible grasps.

Subjective. The subjects answered the same Likert scale as before, which asked which program they preferred, which program made their goal easier to accomplish, and which program inspired the most trust in the robot.

Hypotheses.

H1. *The myopic/noisy model assumption will result in lower goal ergonomic costs than the global/perfect assumption, even when the human knows the goal.*

H2. *The myopic/noisy model assumption will result in less goal infeasible grasp choices than the global/perfect assumption.*

Analysis

H1. We used the human's estimated goal ergonomic costs to evaluate this hypothesis. Figure 3.14 and Table 3.6 show the estimated costs for the case where the human is unaware of the goal, while Figure 3.15 and Table 3.7 show this same data for the case where the human is aware of the goal. The two figures both show our proposed myopic noisy model in orange. The four pairs of numbers are the ergonomic costs estimated by each of the four objective measures OM1-OM4.

The results from Part 1 (human unaware of goal) show a significant decrease in expected goal ergonomic costs when the myopic/noisy assumption is used, as compared to the global/perfect assumption. This decrease is present for all of the four expected cost objective measures. This supports the first portion of our hypothesis: that humans unaware of the goal will achieve a lower goal ergonomic cost with the MN model assumptions.

This effect was not seen when the humans were aware of the goal though. Figure 3.15 shows that the expected goal ergonomic costs for Part 2 of the study were essentially unchanged between the MN and GP assumptions. This suggests that, for this simple task, humans were able to successfully anticipate and avoid grasps which would be infeasible at the goal.

Table 3.6: Estimated Human Ergonomic Costs at Goal (Part 1: Users not aware of goal)

Objective Measure	GP	MN
$E[C(q^{goal})], \lambda = 1, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR feas. at handover}$	10.10	6.81
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR feas. at handover}$	10.37	7.20
$E[C(q^{goal})], \lambda = 1, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$	10.12	6.82
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$	10.88	7.86
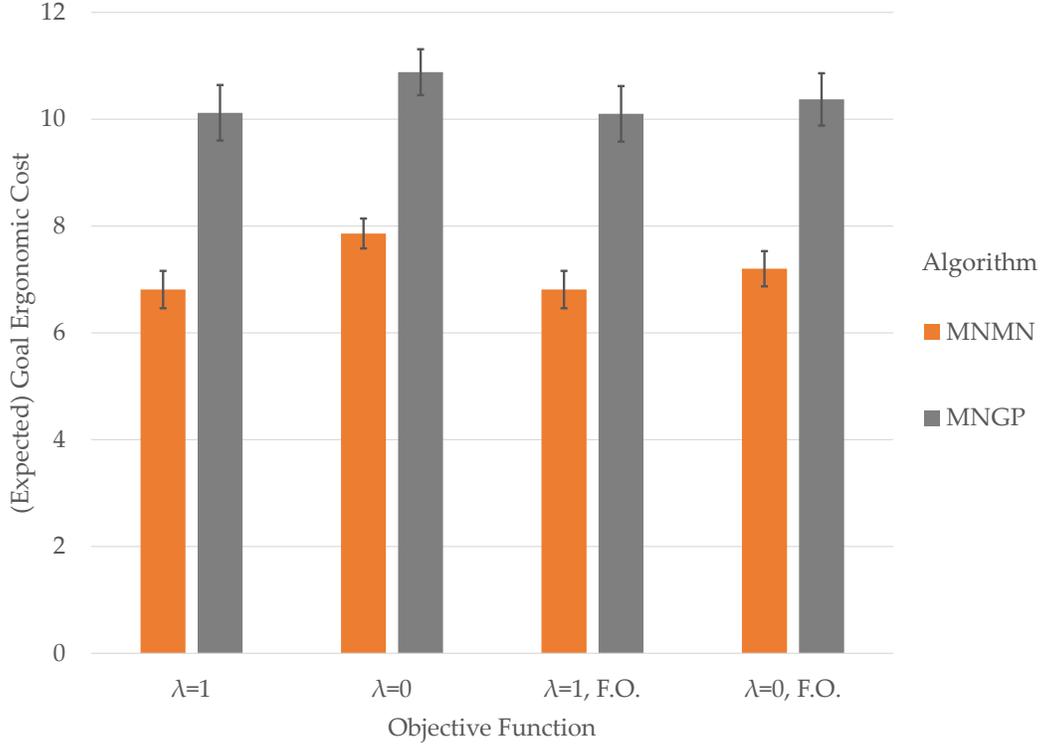
Figure 3.14: **Estimated Goal Ergonomic Costs vs. Model Assumption - Part 1:** Expected goal ergonomic costs for trials where the user *was not* aware of the final goal (i.e. a perfectly myopic human). Expected costs were computed according to the same four hypothesized human objective measures described in Section 3.6). For all objective functions, planning under the assumption that the human was Myopic-Noisy produced anticipated goal ergonomic costs significantly lower than when the human was assumed to be Global-Perfect.

| Objective Measure	GP	MN
$E[C(q^{goal})], \lambda = 1, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$ feas. at handover	5.93	6.01
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$ feas. at handover	6.19	6.52
$E[C(q^{goal})], \lambda = 1, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$	5.95	6.03
$E[C(q^{goal})], \lambda = 0, \forall q_{goal} \in \text{IK(g)}, \forall g \in \text{TSR}$	7.61	7.37

Table 3.7: Estimated Human Ergonomic Costs at Goal (Part 2: Users aware of the goal)
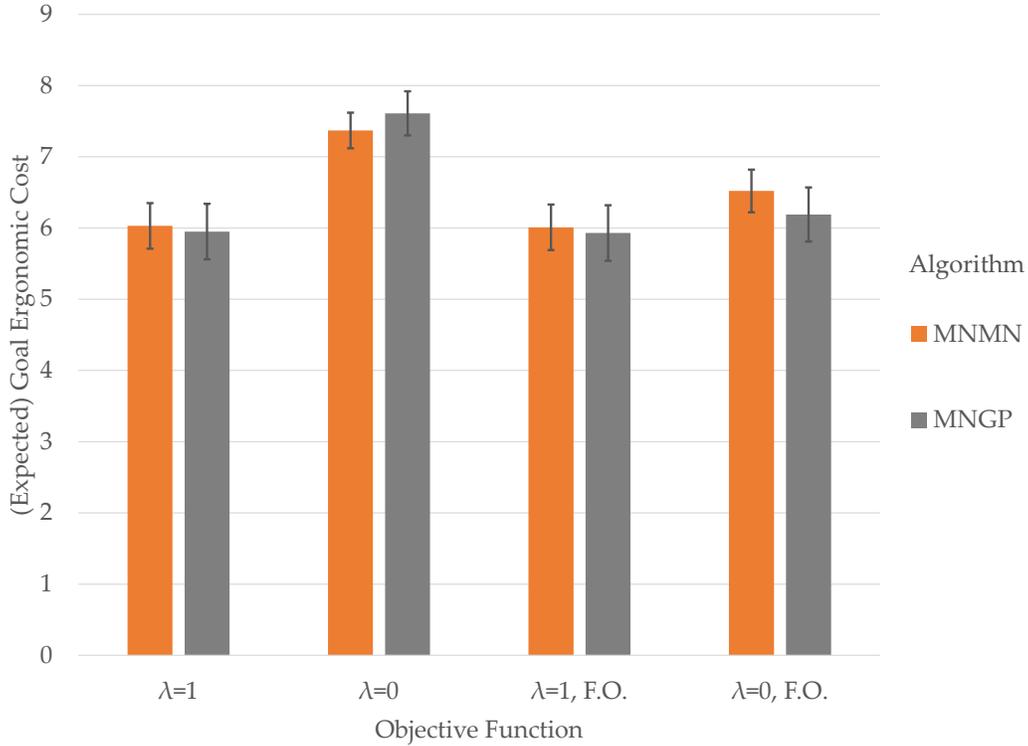
Figure 3.15: **Estimated Goal Ergonomic Costs vs. Algorithm - Part 2:** Expected goal ergonomic costs for trials where the user *was* aware of the final goal (i.e. human not necessarily myopic). Expected costs were computed according to the same four hypothesized human objective measures described in Section 3.6). No significant difference was seen between the goal ergonomic costs under the two different planners, suggesting that for this simple task, humans were usually able to look ahead to the goal and choose a feasible grasp at handoff.

H2. To evaluate our second hypothesis, we'll examine the final object measure, OM5, which is the number of goal infeasible grasps selected at handover by the human. Figure 3.16 shows this data, averaged across the goal/no goal and the MN/GP conditions.

For the case where the human is unaware of the goal, we see the number of goal infeasible grasps drop significantly when using the MN model assumptions. There was also a significant decrease in infeasible grasps when the human was aware of the goal. This suggests that our method does, in fact, help users to avoid grasps which will be infeasible at the goal, and that this is true even when the users are aware of the goal, although the effect size is much smaller.
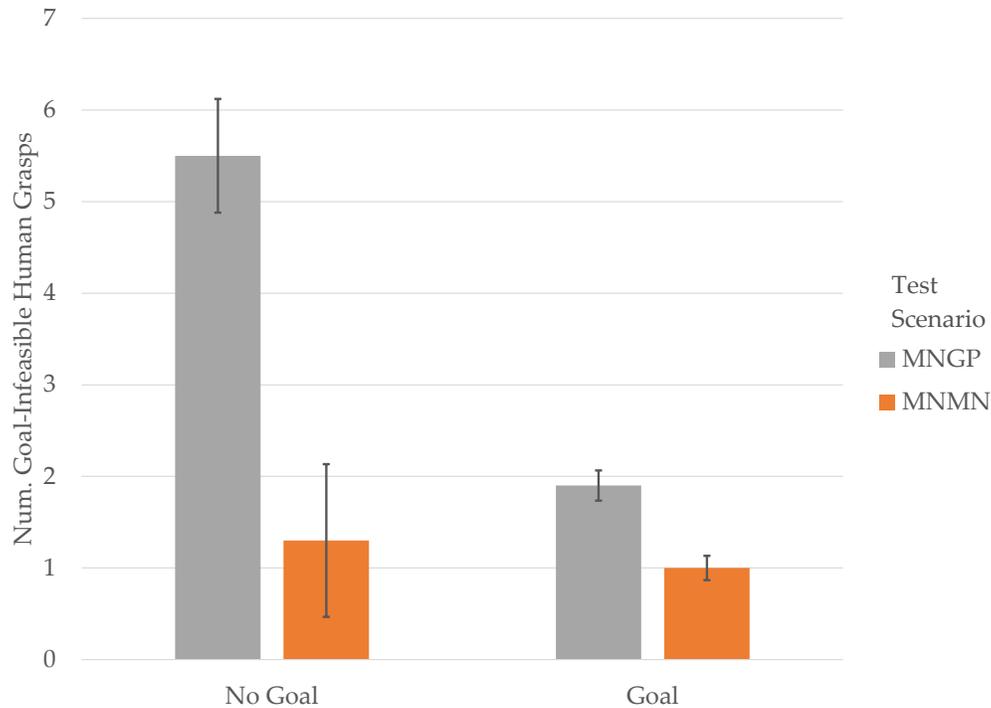
Figure 3.16: **Number of Infeasible Human Grasps**

3.7 Limitations and Future Work

In general, the preliminary results above suggest our algorithm works well. The feasible grasp samples for both the human and robot are intuitively realistic, and provide good coverage of the set of all feasible grasp configurations. Their accuracy could be improved by further tuning the TSRs for the objects to be grasped, or even by building TSRs in a more data-driven way using test data from actual human subjects collected using a motion capture system.

The handoff IK solver also worked reliably. The relative poses of the human and robot in this scenario meant that there were many feasible handoff poses at a variety of configurations. Because of this freedom, most of the optimal configurations identified by the IK solver were very close to the nominal neutral position of the human's arm (at which the handoff ergonomic cost achieves its absolute minimum).

Table 3.8: Post-Study Survey Results

Statement	GP	MN
"I prefer Program __"	3.7	5.0
"The robot was helpful when running Program __"	4.7	5.7
"I trust the robot running Program __"	4.7	5.0
"The robot understood my goal when running Program __"	4.5	5.5
"It was physically easy to do the task when the robot was running Program __"	4.4	5.5
"The robot running Program __ handed me objects in a way that made the task easier"	4.1	5.5
"If you had to choose a program you prefer, which would it be?"	30%	70%

An additional problem is that sampling feasible grasp configurations for both the human and robot and running the IK solver for every possible grasp combination is an inefficient, brute force approach. It should be possible to locally parameterize the space of feasible grasps for the object (perhaps along the degrees of freedom specified already by the TSR), and then include the grasp pose as a decision variable directly in the optimization problem.

Lastly, we'd like to extend this approach to consider the ergonomic cost of the human's trajectory in addition to the costs of the static handoff and goal configurations. An important question here will be tractability of the resulting optimization problem, since we would essentially want to evaluate the gradient of a trajectory cost functional with respect to the grasp and handover configurations, but defining the cost functional itself requires generating an optimal motion plan for the human which makes assumptions about how the human's motion will change in response to the robot's actions.

Chapter 4

Active Learning of Human Ergonomic Preferences

4.1 Introduction

When a robot hands over an object to a person, it has a choice to make – it chooses which specific grasping configuration to use. When the person then takes that object, they too have same choice to make. But depending on what the robot chose, their options might be limited, and they might be forced to twist their arm in some uncomfortable way just to be able to reach the object.

Our goal is to enable robots to choose handover configurations that result in *comfortable* options for the person who is taking the object. But to do that, the robot needs to know what "comfortable" actually means.

In this work, we capture comfort level via an *ergonomic cost*, mapping human grasping configurations to a scalar, based on how comfortable or uncomfortable they are. If the robot had access to this cost function, it could use it to plan its handovers to explicitly make low cost human configurations for taking the object feasible. More interestingly, it could use it as a *predictive model* for how the person will take the object (assuming lower cost configurations are more likely). Having such a model empowers the robot to anticipate human action, and even influence people towards grasps that better suit their ultimate goal for the object [52].

In our previous work, we have simply written down what seemed like a reasonable ergonomic cost function and handed it to the robot [18, 52]. Other works have done the same [56, 60, 65]. But there was always something worrisome about this: how do we know this cost is any good? Even for the average person, we might have gotten it wrong. And fur-
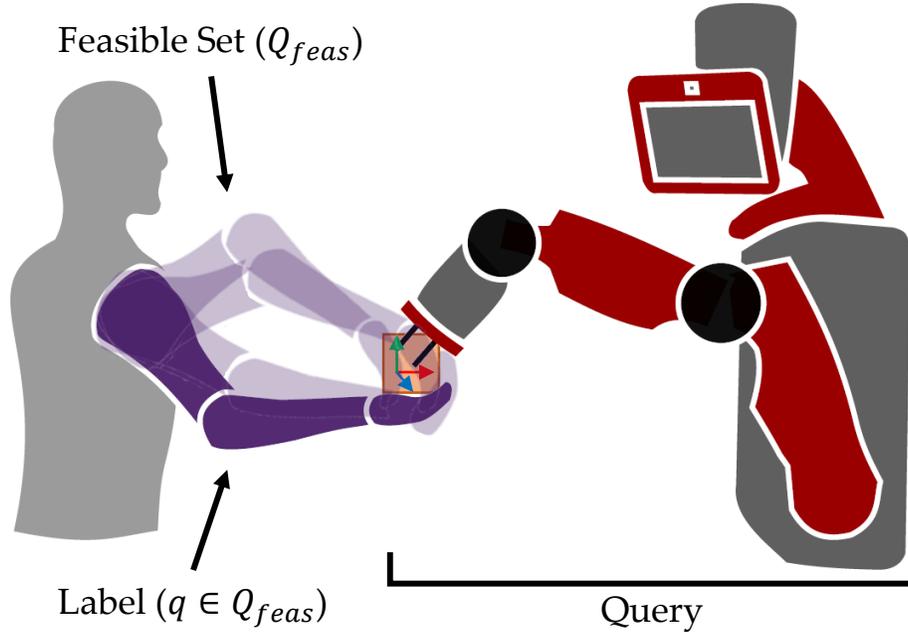
Figure 4.1: The robot learns the human ergonomic preferences by selecting a handover configuration for an object (query), and using the way the person chooses to take the object (label) as an observation about their hidden ergonomic parameters.

ther, not everyone is the average person. We expect to see individual variation in ergonomic preferences, e.g. based on which muscles on their arm happen to be stronger. These differences would be even more pronounced for people whose motion is restricted by age or disability.

In this work, we turn to *learning* a cost function, rather than assuming one and planning with it. The parameters of the cost are a hidden part of the state. Every choice the person makes for a grasp configuration is an *observation* about these parameters, updating the robot's belief via Bayesian inference.

A key aspect of this learning problem is that observations do not happen in isolation from the robot. The robot gets to influence them by selecting its own grasping configuration for offering the object – an *implicit query* that the robot makes to the person (Fig. 4.1). This query induces a *feasible set* of human grasp configurations for taking the object, which results in a *label*, i.e. the person's choice.

Therefore, which queries the robot makes affect its performance. Some queries elicit more information and help the robot learn faster, but sacrifice some human comfort during learning in order to do so.

We thus explore and contrast two natural approaches for selecting queries. In the *passive* approach, we use the principle of separating estimation and control, and the robot always selects the query that leads to the most comfortable configuration for the person according to its current estimate of the ergonomic cost. In the *active* approach, the robot selects queries

that lead to the highest expected information gain.[1]

What we contribute is a formulation of the ergonomic cost learning problem as online estimation via implicit, physical robot queries, and an in-depth analysis of the advantages of disadvantages of the two methods, including how the menu of objects and queries available to the robot impacts their performance.

We do experiments in both simulation (with an easy to visualize 2DOF arm and a more difficult, but realistic 7DOF arm), as well as a user study. We find that online learning of ergonomic cost is feasible, that active learning can be faster, but that this depends on the kinds of objects it has available, and that it comes at small, but non-trivial comfort cost during learning.

4.2   Related Work

Both the active learning paradigm and the optimization of human-robot object handovers have been studied extensively in the robotics/AI literature, respectively.

Learning from Implicit Queries in Robotics

Many robotic learning systems solicit explicit feedback from humans to optimize their own actions. This feedback can include binary preferences for one trajectory over another [66,67], rankings of grasp quality on a continuous scale [68], or class labels applied to images the robot encounters [69].

Intuitively, some queries are more informative than others. The idea of *actively* selecting the most informative queries to present to the human expert has been widely studied and applied [17,70]. These explicit, actively selected queries include binary classification [71–76], ranking [77,78], and labeling [79]. Queries in robot learning can be particularly expensive in both time and effort, often requiring physical motion of the robot. This makes active learning approaches, which minimize the required number of queries to learn a model of a given accuracy, attractive here as well [67–69,80,81].

Sometimes an explicit human response to a query isn't required at all. In these *implicit* learning tasks, the robot simply takes an action, observes the human's response, and uses this response to infer their label, under the assumption that the response represents an optimal execution of a policy based on their true preferences. Recent work has used this framework to learn the characteristics of human drivers in order to predict their future trajectories [82]. This implicit learning approach can be more intuitive in situations where the human doesn't have a conscious, explicit ordering over possible actions, but nonetheless exhibits a preference for some over others, from which this ordering can be reconstructed [70]. We use implicit human feedback. Our approach is conceptually similar to Inverse Reinforcement Learning

---

[1]Note that solving the problem as a POMDP is computationally prohibitive still, but would lead to optimally trading off between exploration and exploitation, i.e. a hybrid between the active and passive approaches.

(IRL) [83], and is essentially Bayesian IRL [84], with the adjustment that, since our belief space is simple the configuration space, maintaining a particle representation of the full belief is tractable.

Ergonomic Handovers

Ideally, a robot would allow you the most comfortable configuration options possible. Existing work focuses on selecting object handover positions [18, 54, 56, 60] or poses [52, 53, 55, 57–59, 61] which accomplish this.

Previous work uses a wide variety of metrics to compare the feasible human grasp configurations allowed by a robot's chosen handover. Examples include predefined ergonomic costs [18, 52, 60], proximity of the object to the human's body [56, 60], visibility of the object [56, 60], manipulability at a given grasp configuration [58, 59] or simply the total number of grasp configurations available to the human given a particular handover pose [61, 69].

4.3   Learning Ergonomic Cost

We pose the ergonomic cost learning problem as one of learning from implicit queries, which results in human demonstrations. Suppose a robot hands an object to a person using a grasp g_R and at a pose T_{hand}. We'll call the tuple (g_R, T_{hand}) the robot's *action*, i.e the *query*. This action, in combination with the kinematic structure of the human's arm and the object's feasible grasp regions, induces a set of feasible human grasp configurations Q_{feas}, which we'll sometimes write as a function of the robot's action $Q_{feas}(g_R, T_{hand})$. When the human is presented this robot action, they choose a single grasp configuration $q_H \in Q_{feas}$, which is our *observation* or *label*. Given multiple action-observation pairs, we train a model of $P(q_H|Q_{feas})$, which is the probability that the human will choose a particular $q_H \in Q_{feas}$ when the robot takes an action (g_R, T_{hand}) that results in the feasible set Q_{feas}.

Feasible Set Computation

We represent the set of all feasible human grasps G_H on the object being handed off as a Task Space Region (TSR) [63], where $G_H \subset SE(3)$. This TSR is discretized to produce a finite set of all feasible grasps $G_H \triangleq \{g_{H1}, \ldots, g_{HN}\}$. Note that while we specify the TSRs for each object manually here, they could just as easily be generated by a simulator which evaluates possible grasps for force closure, a database of feasible object grasps, or many other graspable region representations. TSRs are computationally convenient, and the fact that we can derive the TSRs from many other possible graspable region representations allows us to abstract away the details of the specific representation, and focus on the kinematic constraints induced by the the available grasp regions.

For each grasp $g_{Hi} \in G_H$, we can compute a set of inverse kinematics (IK) solutions $Q_{feas,i}$ which allow the human to reach the specified grasp. We collect all the IK solutions for all

object grasps into a set $Q_{feas} = \bigcup_{i=0}^{N} Q_{feas,i}$. As mentioned above, we'll abstract away this process by simply writing the feasible set as a function of the robot's action $Q_{feas}(g_R, T_{hand})$.

Probabilistic Model

Given a set Q_{feas} of feasible human arm configurations, we seek a model $P(q_H | Q_{feas})$, which gives the probability that a given person will select any of the individual grasps $q_H \in Q_{feas}$.

   We structure this model with an assumption: we assume that the human is approximately rational, and that the likelihood of them selecting a configuration decreases exponentially as the *ergonomic cost* of that configuration increases:

$$P(q_H) \propto e^{-\alpha C_{ergo}(q_H)}. \tag{4.1}$$

   In general, C_{ergo} can be any function which maps configurations to scalar costs. The methods we present could estimate any parametrization of such a function. Nonetheless, to experiment with the methods, we must commit to a parametrization. We choose an intuitive one: we parametrize cost as squared distance from some *neutral* arm configuration q_H^* that captures the most comfortable configuration for the person, but measure distance with respect to an inner product W which is not necessarily Euclidean:

$$C_{ergo}(q_H, \lambda) \triangleq (q_H - q_H^*)^\intercal W (q_H - q_H^*). \tag{4.2}$$

In our experiments, we assume for simplicity (in order to lower the number of parameters we need to estimate) a diagonal *weight* matrix $W = diag(w)$. This captures preferences in moving certain joints away from the neutral configurations more than others (e.g. it might require more effort to displace the shoulder by 0.5 rad than the wrist by the same amount). We collect these parameters into a single parameter vector $\lambda \triangleq [q_H^*, w]$ which the robot needs to estimate by interacting with the human.

   Given a known λ and a set of feasible arm configurations Q_{feas}, the resulting probability of the human choosing a given configuration q_H takes the form of a Boltzmann distribution:

$$P(q_H | Q_{feas}, \lambda) = \frac{e^{-C_{ergo}(q_H; \lambda)}}{\sum_{\widehat{q_H} \in Q_{feas}} e^{-C_{ergo}(\widehat{q_H}; \lambda)}} \tag{4.3}$$

Importantly, this distribution normalized over all other configurations that the person *could have chosen*.

Bayesian Belief Updates

We start with a generic, uncertain belief $P(\lambda)_0$ over the human's cost function parameters. Our goal is to iteratively refine this belief as we collect more training data. Our technique could be used across a population to learn an average ergonomic cost, or for an individual to personalize the ergonomic cost to their preferences.

Because the beliefs over possible cost function parameters produced by our training examples are potentially quite complex, we use a particle filter to perform belief updates. This enables us to represent arbitrarily complex beliefs without being constrained by the form of a parameterized distribution.

To perform a single belief update, the robot chooses an action (g_R, T_{hand}), which induces a set of feasible human configurations $Q_{feas}(g_R, T_{hand})$. We then observe the human's choice q_H. Our complete training example is then the tuple (q_H, Q_{feas}). We update our prior belief $P(\lambda)$ with the training example to give a posterior $P(\lambda | q_H, Q_{feas})$ by applying Bayes' Rule:

$$P(\lambda | q_H, Q_{feas}) \propto P(q_H | Q_{feas}, \lambda) P(\lambda) \tag{4.4}$$

Note that the likelihood function $P(q_H | Q_{feas}, \lambda)$ is equal to the Boltzmann likelihood in (4.3).

We represent the prior belief at each step as a set of N particles $\Lambda = \{\widetilde{\lambda_1}, \ldots, \widetilde{\lambda_N}\}$ and corresponding weights $\Omega^\lambda = \{\omega_1^\lambda, \ldots, \omega_I^\lambda\}$. Given a new training sample (q_h, Q_{feas}), we compute the new particle weights $\Omega^{\lambda'}$ as:

$$\omega_i^{\lambda'} = \omega_i^\lambda \left(\frac{e^{-C_{ergo}(q_H; \widetilde{\lambda_i})}}{\sum_{\widehat{q_H} \in Q_{feas}} e^{-C_{ergo}(\widehat{q_H}; \widetilde{\lambda_i})}} \right) \tag{4.5}$$

Particles are then resampled to produce final Λ' and $\Omega^{\lambda'}$ sets in which all the particle weights are uniform. Note that the normalization constant in (4.5) is different for each value of i, in contrast to a more typical particle filter implementation where the normalization constant is the same for every particle.

Active Query Selection

Suppose we have a menu of N possible robot actions, which induce feasible sets $\{Q_{feas,1}, \ldots, Q_{feas,N}\}$, and our current belief about the cost function parameters is $P(\lambda)$. When we present our human with any of the feasible sets, we'll observe a training datapoint (q_H, Q_{feas}). Intuitively though, some queries elicit human responses which are more informative than others. The active method actively seeks out these queries to present to the human.

If we assume the belief state $P(\lambda)$ is represented using a set of particles p_i with corresponding weights k_i, we can compute the Shannon entropy of the belief state $H(P(\lambda))$ by discretizing the belief space into M discrete beliefs using a grid, then applying the standard definition of entropy:

$$H(P(\lambda)) \triangleq \sum_{i=1}^{M} P(\lambda_i) \log P(\lambda_i) \tag{4.6}$$

We can then compute the expected change in entropy $E[\Delta H(Q_{feas})]$ of our belief (i.e. the information gain) given that we chose a particular query Q_{feas}:

$$E[\Delta H(Q_{feas})] = H(P(\lambda)) - E_{\widehat{\lambda} \sim P(\lambda)} \left[E_{\widehat{q_H} \sim P(q_H | \widehat{\lambda})} \left[H\left(P(\lambda \mid \widehat{q_H}, Q_{feas}) \right) \right] \right] \tag{4.7}$$

To make the calculation more efficient, we approximate by substituting the maximum likelihood value of q_H for the inner expectation:

$$E[\Delta H(Q_{feas})] = H(P(\lambda)) - E_{\widehat{\lambda} \sim P(\lambda)} \left[H \left(P\left(\lambda \mid \arg\max_{q_H} P(q_H \mid \widehat{\lambda}), Q_{feas} \right) \right) \right] \qquad (4.8)$$

After computing the expected information gain for each feasible set in the menu of possibilities, we select the one which produces the greatest expected information gain:

$$Q_{feas}^{active} = \arg\max_{Q_{feas}} E[\Delta H(Q_{feas})] \qquad (4.9)$$

Passive Query Selection

The active query selection algorithm in Sec. 4.3 selects robot actions to maximize information gain from each query. In contrast, a passive approach selects the action that that minimizes expected ergonomic cost to the person at each step, i.e. the expected cost of the grasp configuration the human is most likely to pick, given the robot's current belief about their ergonomic cost parameters. The passive approach still gains information at ever step, but merely as a side effect. This corresponds to separating estimation from control or hindsight optimization [85, 86], i.e. always planning with the current belief as if the ground truth will be revealed at the next step, and updating the belief at every step based on the new observation.

For a given feasible set Q_{feas}, the expected human ergonomic cost is:

$$E[C_{ergo}(Q_{feas})] = E_{\widehat{\lambda} \sim P(\lambda)} \left[\min_{\widehat{q_H} \in Q_{feas}} C_{ergo}(\widehat{q_H}; \widehat{\lambda}) \right] \qquad (4.10)$$

We then select the feasible set which minimizes the expected ergonomic cost incurred by the human:

$$Q_{feas}^{passive} = \arg\min_{Q_{feas}} E[C_{ergo}(Q_{feas})] \qquad (4.11)$$

4.4 Experimental Design

We evaluate our active learning algorithm in three separate scenarios: 1) a simulated 2 DoF planar arm with planar objects, 2) a simulated 7 DoF human arm with real, 3D objects, and 3) a human user study on a real life object handover task. In the first two simulated scenarios, we test the algorithm's ability to recover the known ground truth parameters of our simulated human's ergonomic cost function. In the user study, we evaluate the accuracy with which our learned model can predict the human's grasp in future handoffs.

We test the learning of both the neutral configuration q_H^* and the joint weights w.

Manipulated Variables

In each scenario, we manipulated the query selection algorithm. In the simulation scenarios, we compare *active*, *passive*, and *random* algorithms (i.e. choose an object and a configuration at random). In the human user study, we compare *active* with *passive*. We also manipulated the number of queries allowed, from 1 to 5 for the 2 DoF simulations and the user study, and from 1 to 10 for the 7 DoF simulations.

Objective Measures

We measure *accuracy* and *training cost* – accuracy is most important if we think of this as a training period with the robot, to be the followed by a lifelong interaction; cost is very important if we think of this as a continuous interaction, in which the robot needs to learn without making the person uncomfortable.

   We measure the accuracy of the each algorithm's belief $P(\lambda)$ over the model parameters, at every time step (i.e. # of iterations) using the following objective measures:

- $P(\lambda^*)$: Probability density of the belief at the known ground truth parameter value

- $||\arg\max P(\lambda) - \lambda^*||$: Euclidean distance between the ground truth parameter vector and the mode of our belief

- $\log \mathcal{L}(\arg\max P(\lambda) \mid q_H^{test}, Q_{feas}^{test})$: The log likelihood of the mode of our belief, with respect to a separate test dataset (i.e. test set log likelihood)

   The first two measures require us to know the ground truth ergonomic cost function parameters λ^*, so we evaluate them only for our two simulated experiments.

   We also measure training cost:

- $E[C_{ergo}(q_H)]$: The expected ergonomic cost of a query with respect to the ground truth cost

Subjective Measures

In the user study, we also care about what experience the users prefer, especially in light of the fact that the difference between passive and active is in optimizing information gain vs. (greedily) optimizing user comfort. We ask 4 Likert scale questions and a forced choice (Table 4.1).

Hypothesis

Because the active algorithm is designed to quickly reduce the entropy of the belief, we hypothesize that it will learn the parameters faster (i.e. have higher accuracy for the same number of queries, especially when this number is low). However, we also expect that the

amount of improvement will depend on the set of available queries, and that the passive algorithm will incur lower true ergonomic cost during training.

4.5 Analysis for a Planar Two DoF Arm

We used a simulated planar 2 DoF arm. To create our training set, we randomly selected eight task space object shapes (as shown in Figures 4.5 and 4.6 and enumerated all the IK solutions for a discretized version of the object to yield a menu of eight feasible sets $\{Q_{feas,1}, \ldots, Q_{feas,8}\}$. The three algorithms – active, passive, and random – selected queries from this menu.

   We repeated the training process a total of 50 times. For each trial, we selected a random menu of eight queries to make available to the robot. We selected a random ground truth value q_H^* or w to attempt to learn for each trial.

   Each of the objective measures was evaluated after every iteration. The test set likelihood was evaluated on a set of 300 randomly generated objects. Fig. 4.2 shows the results.

   The active learning algorithm produced faster learning on all three objective measures, for both the neutral configuration q^* and the weights w. This difference was particularly large between the passive and active algorithms when learning the weights, where the passive algorithm failed to converge toward the correct belief even after many iterations.

   On the other hand, active does suffer a loss in true cost, especially after several queries when passive has converged to a decent estimate. How important this is depends on the use case: it is perhaps alright to suffer an initial loss in order to converge to a better cost that the robot will use for many additional interactions; however, users might not tolerate this well in certain tasks.
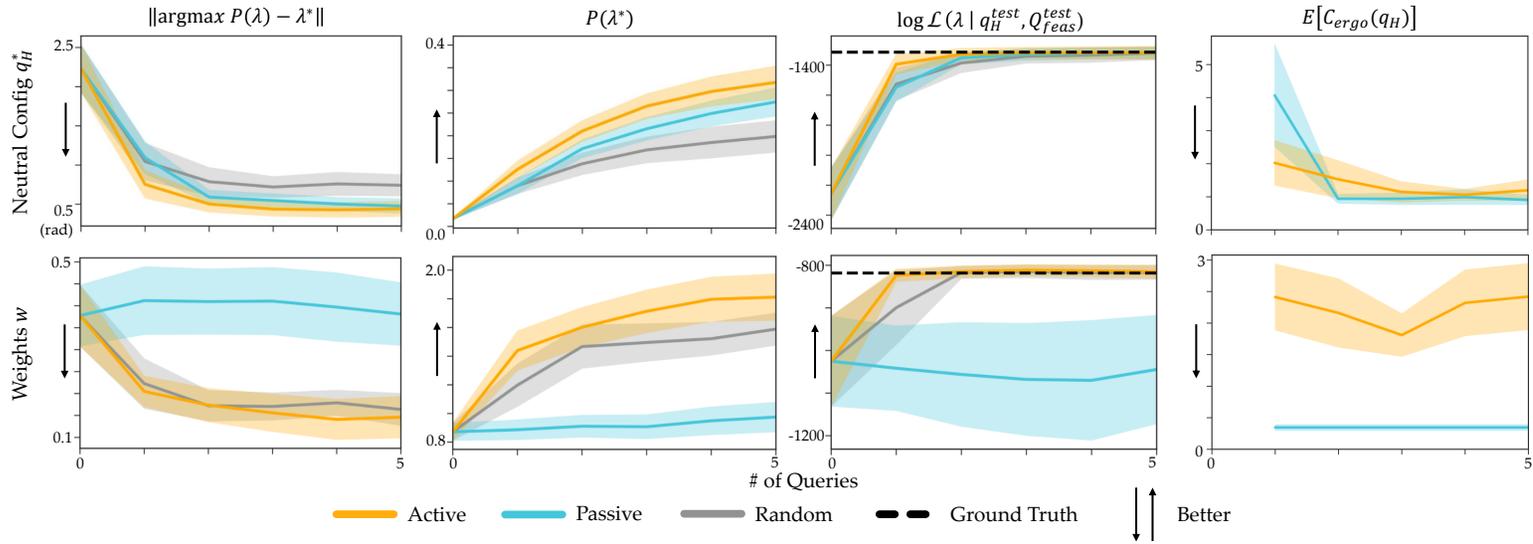
Figure 4.2: The values of each objective measure vs. the number of training handovers, shown for both neutral configuration (q^*) and joint weights (w) learning with a simulated planar 2 DoF arm. The active query selection algorithm yields faster learning than the passive and random algorithms for all three of the objective measures. This difference is particularly marked for learning of the joint weights w (bottom), where the passive algorithm fails to converge towards an accurate belief even after many iterations.
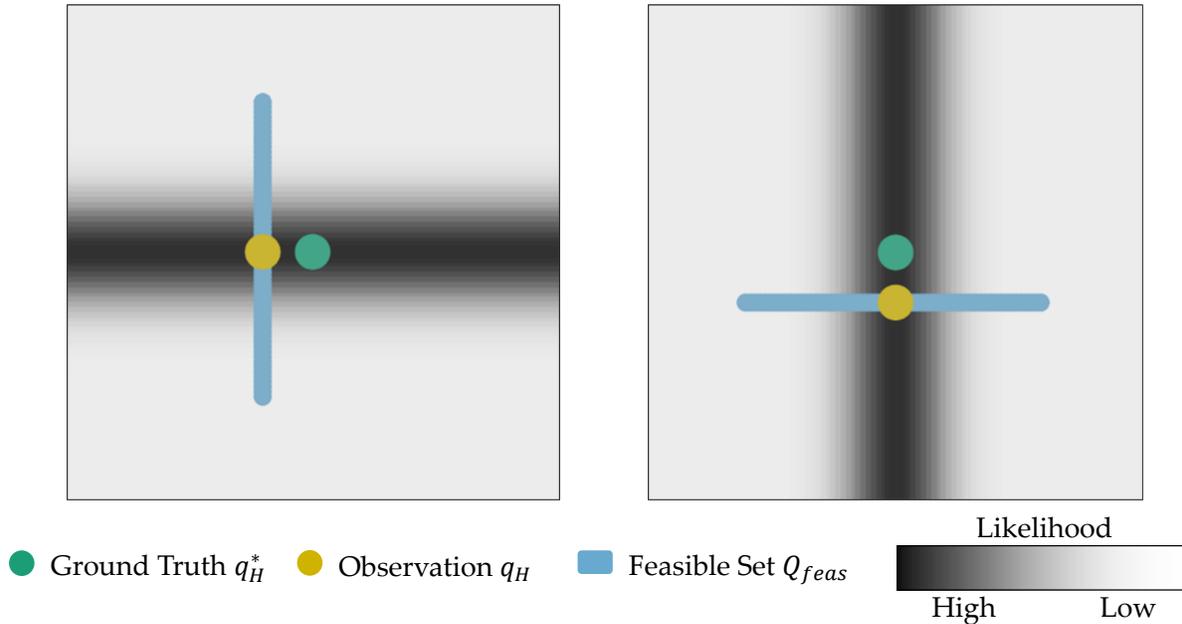
Figure 4.3: Synthetic feasible sets shown with the ground truth optimal configuration (green), the simulated human's choice (yellow), and the resulting likelihood function (grey). Note how the shape of the feasible set of configurations completely changes the resulting likelihood function, even as the human's chosen configuration is relatively similar in both examples.

Artificial Feasible Sets

These initial results suggest that both methods are better than random query selection, with the active method learning faster and the passive method incurring less regret.

Next, we investigate what exactly causes the active learning algorithm to consistently outperform in accuracy. To explore this question, let's examine the active algorithm's decisions on the two simple feasible sets from Fig. 4.3. As the figure makes obvious, the feasible set Q_{feas} of configurations available to the human has a huge impact on the resulting likelihood function $P(q_H^* \mid q_H, Q_{feas})$, even if the human's chosen configuration q_H is similar or identical. When the active query selection algorithm is applied to a menu of queries containing just these two feasible sets, it makes the decisions shown in Fig. 4.4. The probability density of the resulting belief after each iteration is shown in grey. Notice how the active algorithm's desire to reduce the belief's entropy causes it to alternate between the two queries, selecting whichever one will remove the greatest amount of probability mass from the current belief at each iteration.

Object-Derived Feasible Sets

With this insight about the impact of feasible set shape and size on the resulting likelihood used to update our belief, let's return to our original two DoF planar arm scenario. Figures
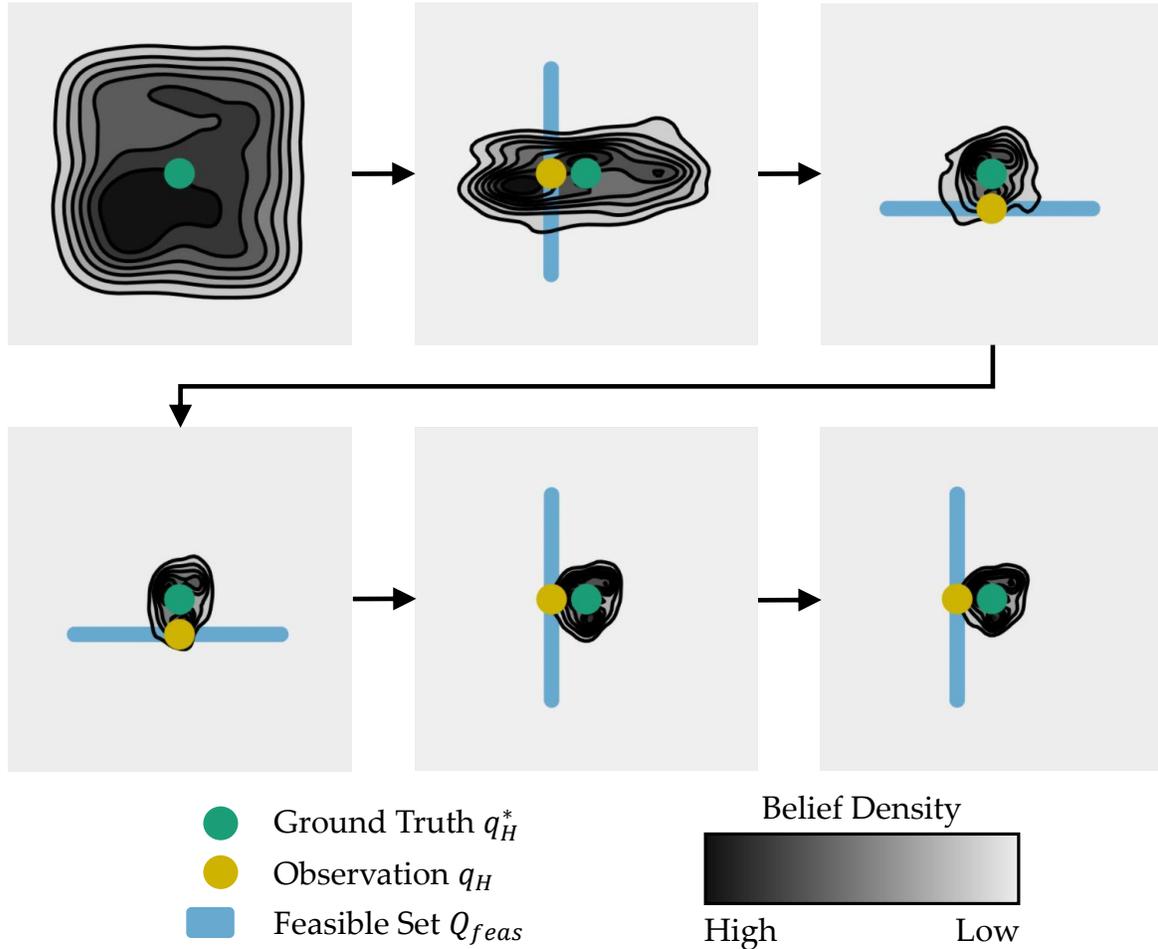
Figure 4.4: Sequence of belief updates when training sets are selected using the active learning algorithm from the set of two possibilities above. Note how the active learning algorithm alternates between the two queries so as to remove the maximum amount of probability mass from the belief at each iteration. Because our model assumes humans are noisy, repeated iterations with the same query will continue to refine our belief past it's initial value after the first iteration.

4.5 and 4.6 show examples of the randomly generated task space objects included in the training sets used to generate the experimental data shown in Fig. 4.2 (bottom row), and the resulting configuration space feasible sets, ground truth q_H^*'s, simulated human's chosen q_H's, and the likelihood functions resulting from each query and observation.

The highly nonlinear nature of the inverse kinematics map means that seemingly similar task space objects can create completely different configuration space feasible sets. In addition, both the shape of the task space objects (Fig. 4.5), and their pose T_{hand} (Fig. 4.6) affect the resulting likelihood function. We can see that, in general, configuration space feasible sets composed of multiple, widely separated disjoint regions produce likelihood functions with sharp gradients which quickly eliminate large pieces of the belief.

What actually happens when we present the active and passive query selection algorithms
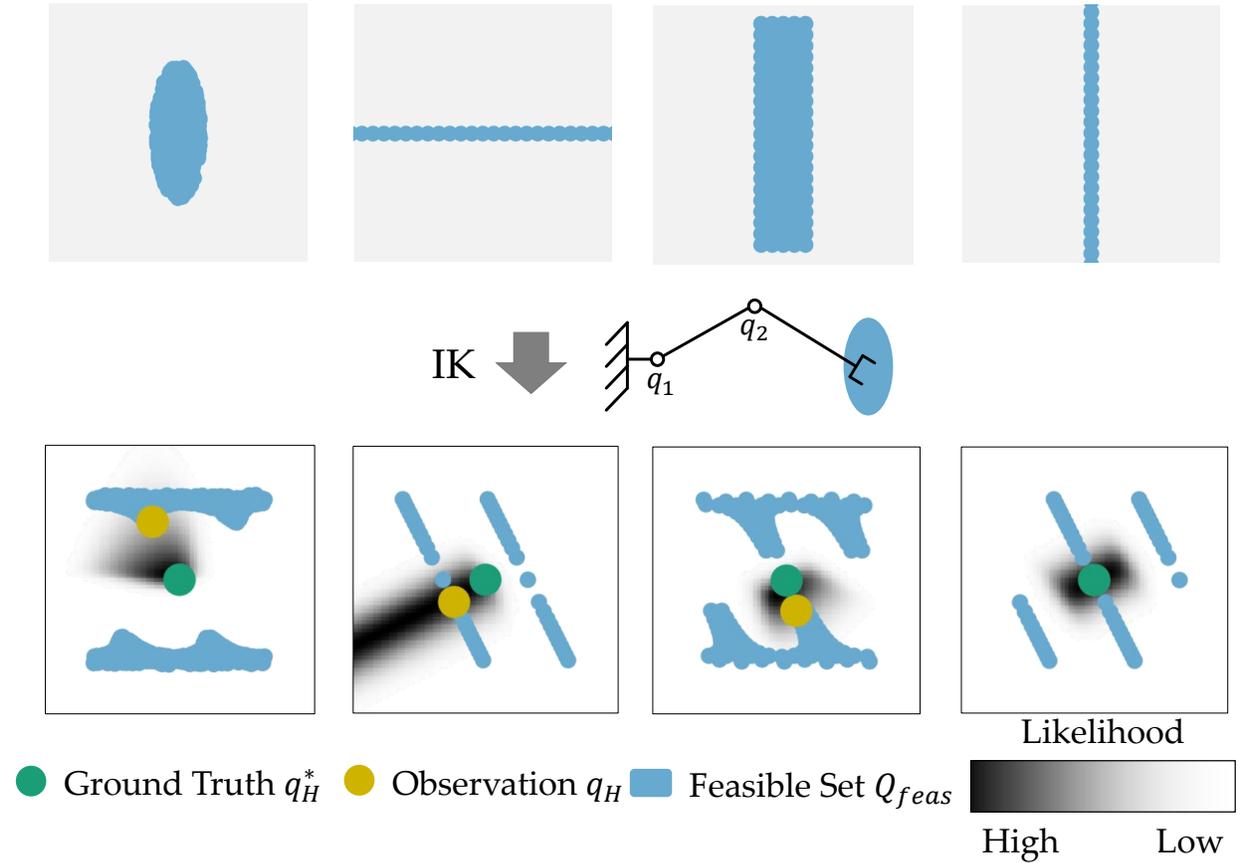
Figure 4.5: A collection of example 2D task space objects and their corresponding configuration space representations. The objects are all centered at the same position, and only their shape varies. Even with this constraint, the resulting configuration space feasible sets and likelihoods vary widely.

with a menu of feasible sets like the one shown in Fig. 4.5? Fig. 4.7 compares active and passive. Notice how the active learning algorithm consistently selects training examples with widely separated, disjoint feasible regions in the configuration space. In contrast, the passive algorithm prefers feasible sets where at least one of the feasible configurations is near the mode of the current belief $P(q_H^*)$. This is reasonable, as the passive algorithm attempts to greedily reduce the human's ergonomic cost, at the expense of slower learning. After the third iteration, the active learning algorithm's belief has converged to the correct ground truth value. In contrast, the passive algorithm's belief is still somewhat uncertain, and the algorithm continues to select a query whose resulting likelihood function will not remove the uncertainty.
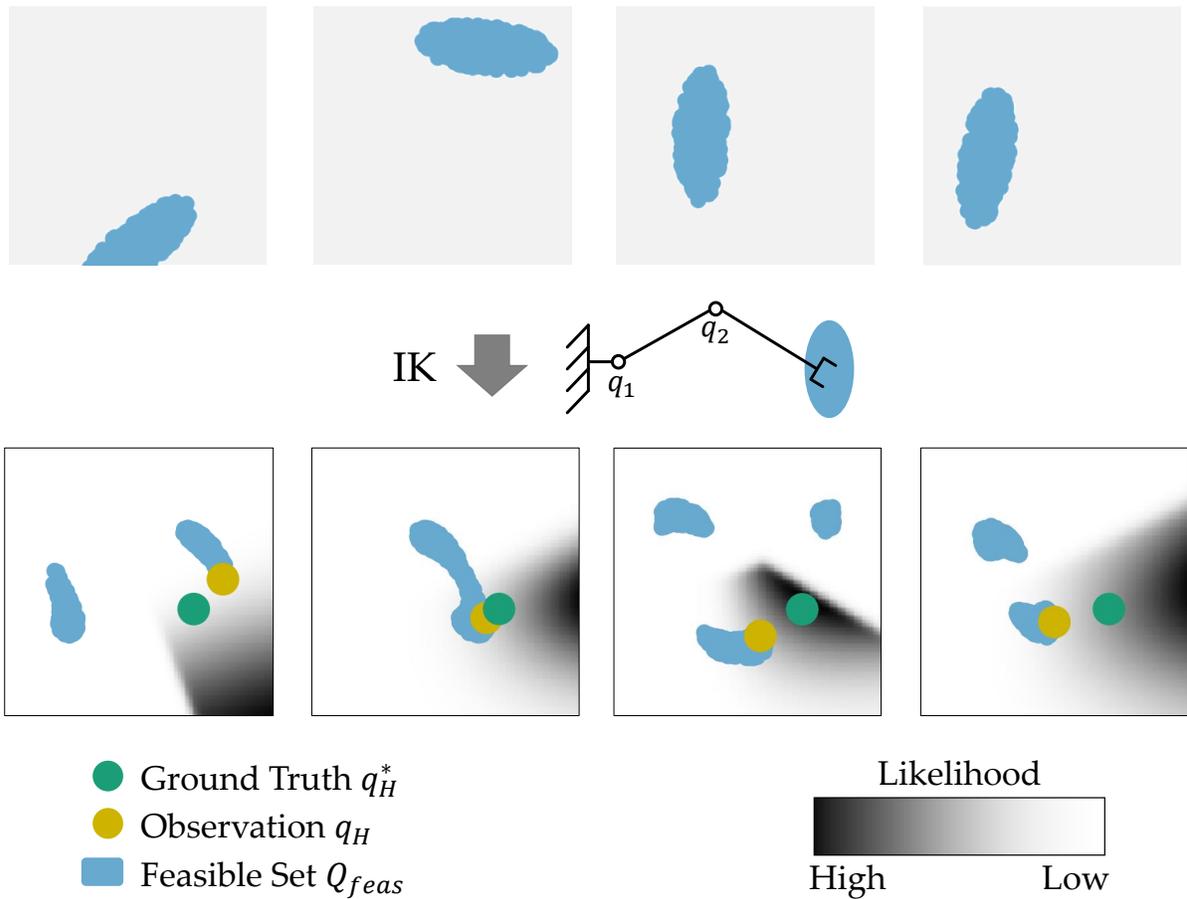
Figure 4.6: A set of example 2D task space objects and their corresponding configuration space representations. Every object has the same shape, but they are each positioned at a different random pose. Just this change in pose creates significant variation between the configuration space feasible sets and likelihoods generated by each object.

4.6 Analysis for a Seven DoF Human Arm

Our second test scenario used a simulated 7 DoF human arm and real objects. Inspired by the advantage of objects that induce separate feasible regions, we use a set of bicycle handlebars, but also a bicycle U-lock which does not have this property.

As in Sec. 4.5, we conducted a total of 50 simulation trials each for the weight and neutral configuration learning portions of the test. For each trial, we supplied the robot with a randomly selected menu of eight object handoff poses and grasps. The ground truth parameters were set to a randomly chosen value, which we attempted to recover. The results are in Fig. 4.8.
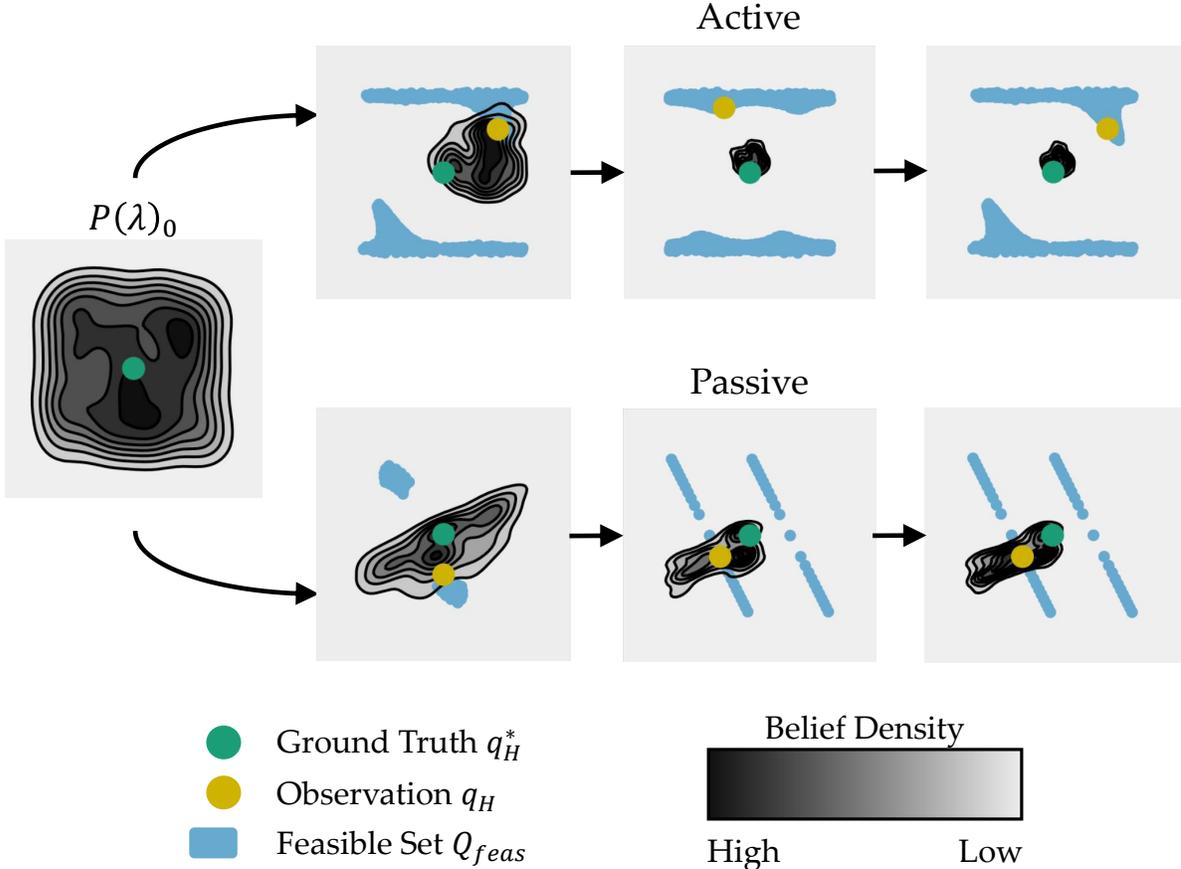
Figure 4.7: Active and passive learning algorithms applied to the same scenario with planar task space objects and the 2 DoF arm. The active algorithm consistently selects queries with widely separated, disjoint feasible regions in the configuration space, while the passive algorithm tends to select queries with at least one feasible configuration close to the ground truth optimal value. The active algorithm's belief converges quickly to the ground truth value, while the passive algorithm allows significant uncertainty to remain after the first three iterations.

Figure 4.8: The values of each objective measure vs. the number of training handovers, shown for neutral configuration (q^*) learning on a simulated handoff task with a 7 DoF human arm model.

As in Sec. 4.5, the active learning algorithm produced consistently faster learning. Particularly notable was the learning of the joint weights, where the passive algorithm produced a test set likelihood which was worse than that of the initial belief, but the active algorithm performed acceptably.
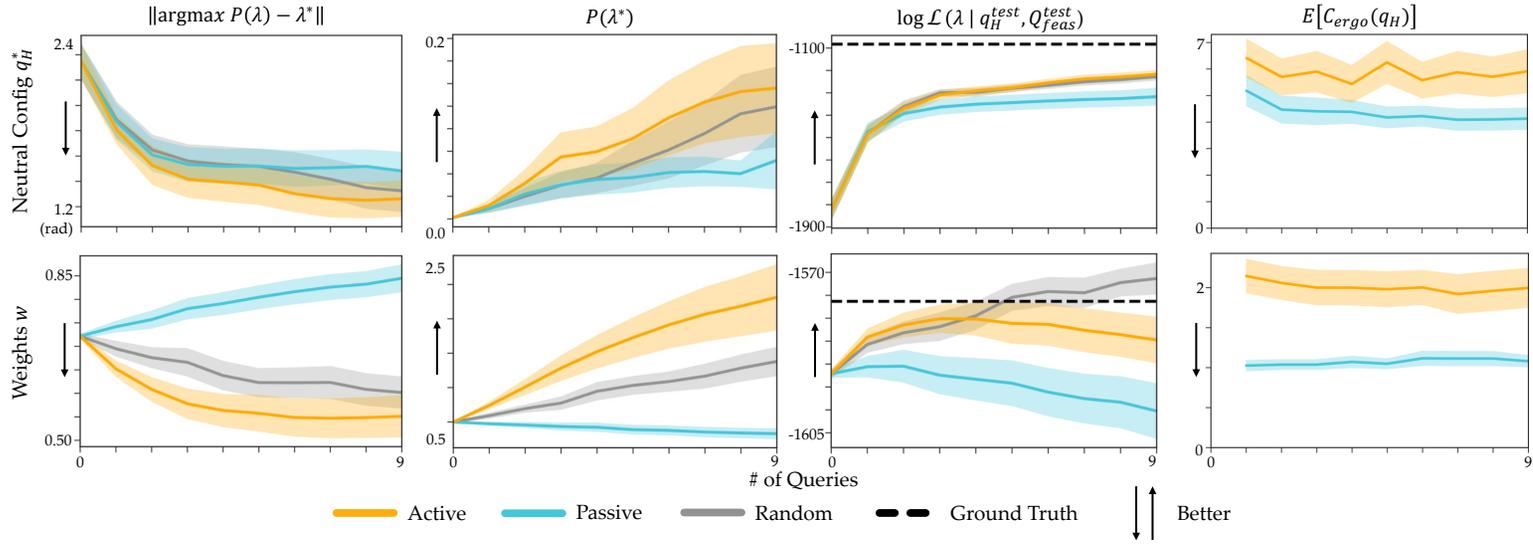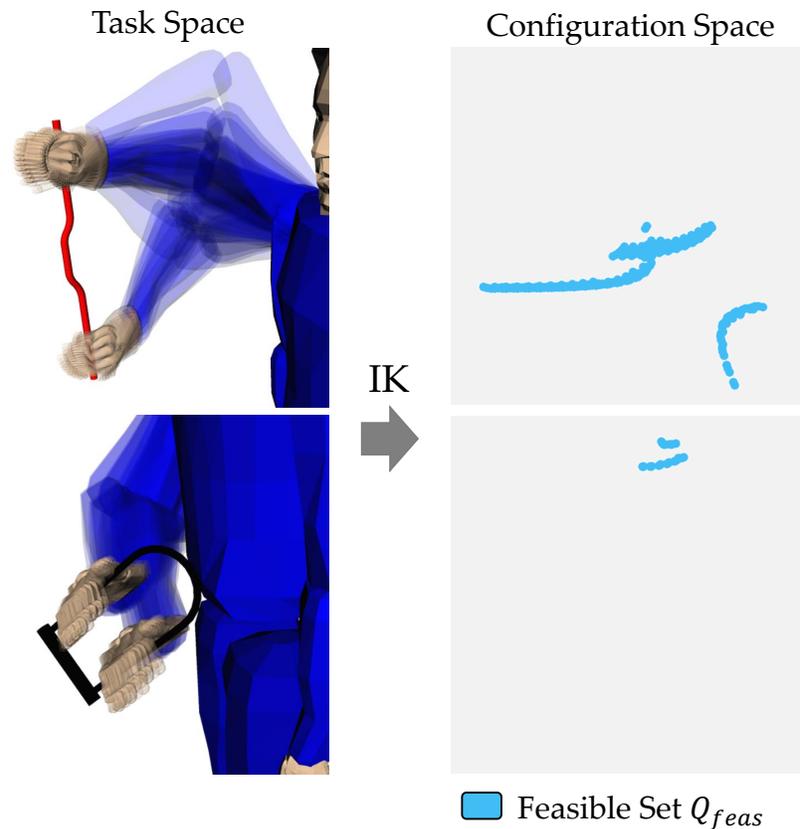

Figure 4.9: Task space sets of human grasp configurations pictured alongside the corresponding configuration space feasible sets Q_{feas}. Notice how the bicycle handlebars produce a large configuration space feasible set with multiple disjoint portions (informative), while the lock produces a much more compact feasible set (uninformative)

To help explain this, examine Fig. 4.9. It shows both the task space feasible sets and corresponding configuration space feasible sets for two selected poses of the bike handlebars and lock. As our earlier trials suggested, the bicycle handlebars, with their two widely separated grip zones produced a much larger, more dispersed configuration space feasible set than the bike lock. Active learning exploits this.

4.7   User Study

Simulation enabled us to tease out interesting aspects of the passive and active approaches, but we still need to study how well these methods perform with real people. We thus conducted a user study with a real handover task.

The handover task used the same bike handlebar and lock objects from the previous section. We selected a menu of robot queries containing four poses of the handlebars and four poses of the lock. We attached rubber grips to both objects to limit the feasible grasps on each to two distinct regions (as shown on the left side of Figure 4.9).

Our study consisted of three phases: two training and one test. In the first phase, the robot conducted five iterations of training using the active learning algorithm. In the second phase, five iterations of the passive algorithm were performed. In the third phase, each of the two objects was presented to the participant at eight different poses, for a total of 16 testing examples.[2]

The performance of the two algorithms is shown in Fig. 4.10. The active learning algorithm consistently yielded a higher likelihood on the test dataset than the passive algorithm for the first two training iterations. After this, the two algorithms were close to identical in performance. This suggests that the active algorithm successfully selected initial training queries which helped it to quickly identify the human's ergonomic model.

Although both training algorithms had access to both the bike handlebars and the bike lock, the active algorithm selected the handlebars exclusively, while the passive algorithm selected only the lock. This is intuitively reasonable: The handlebars induce a large, fragmented configuration space feasible set, which yields informative human responses. The lock produces a small, compact set, which is much less informative, but makes it easy for the robot to guarantee that whichever grasp the human chooses will be reachable comfortably.

As shown in Table 4.1, participants had a slight preference for the passive learning algorithm, but found both just as physically easy.

4.8   Discussion

We formulated ergonomic cost learning as an online estimation problem based on implicit physical queries from the robot. We compared a passive and an active learning approach. In simulation, we found that the active method leads to much faster learning and higher accuracy, but sacrifices user comfort during learning. We also discovered that active learning works best when the object that the robot is handing, combined with the grasping configuration it chooses, leads to feasible choices for the human that have different connected components. With real people, the differences were more subtle: active learning is signif-

---

[2]Since we weren't able to measure the human participants' arm configurations accurately in real time, we instead just provided the active and passive algorithms with the human's chosen grasp region (i.e. which handle they grabbed). The training example q_H was then taken to be the most likely configuration given that grasp, according to a previously specified model.
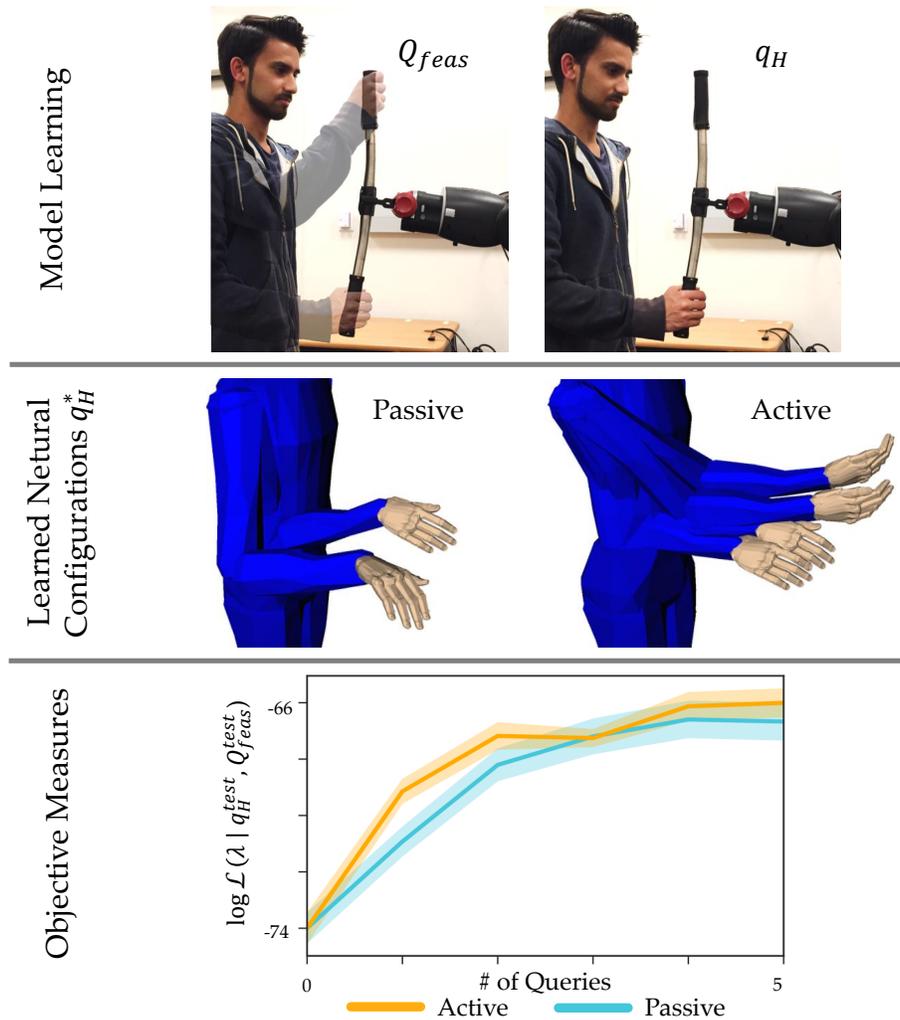
Figure 4.10: The user study asked subjects to select one grasp configuration q_H from a set of two available grasps Q_{feas} in each training query. The passive learning algorithm induced many participants to choose identical labels q_H, and learned only two distinct neutral configurations q_H^* from the set of five subjects. The active algorithm's informative queries elicited a wider variety of behavior, yielding four distinct q_H^*'s. Measured by the test set log likelihood, active learning consistently produced a better model fit after the first one to two iterations, after which the passive algorithm's model had a comparable likelihood. This result is similar to the 7 DoF simulation results in Fig. 4.8. The simulation results suggest that the active vs. passive performance difference may be larger for weight (w) learning.

icantly more accurate in the beginning, but passive converges to a close accuracy. Users choose the comfort of passive, but do not rate it as physically easier to work with.

Above all, what is exciting is that the user study suggests that this online estimation works with real people, in that it improves how well the robot can predict what a real user would do in new situations. An active technique will make it more likely that the robot converges to a better model, but the extent to which that matters in practice remains an

Table 4.1: Post-Study Survey Results

Statement	Active	Passive
"I prefer Program __"	3.4	4.4
"The robot was helpful when running Program __"	4.6	4.6
"It was physically easy to do the task when the robot was running Program __"	4.8	4.8
"The robot running Program __ handed me objects in a way that made the task easier"	4.8	4.6
"If you had to choose a program you prefer, which would it be?"	20%	80%

open question, with passive techniques also performing well overall.

Chapter 5

Discussion and Concluding Remarks

The past three chapters each presented a separate case study examining personalized human modeling as applied to robot-human object handovers. We first examined how personalized human models could be used to plan for *feasibility* (Chapter 2). Next, we added predictive models of human grasp choice to our human kinematic models to generate *ergonomically optimal* actions (Chapter 3). Finally, we investigated how our predictive action choice models could be *learned* online using active information gathering techniques (Chapter 4).

This section summarizes the main conclusions from each of the three case studies, and concludes with a discussion of the significant challenges and limitations which still remain as well as directions for future work.

5.1 Takeaways

Chapter 2: Personalized modeling allows co-robots to seamlessly adapt to human collaborators with different physical abilities

A human's physical abilities may differ from average (due to natural variation, age, injury, or disability) in ways that make a single, generic human model inadequate for planning feasible interactions with the entire range of humans a collaborative robot may encounter.

Supplying the robot with personalized human kinematic models which capture a human's range of motion can allow the robot to automatically adapt its actions to each human's individual constraints. These personalized models can be learned from very little training data.

Chapter 3: Giving co-robots knowledge of humans' noisy, imperfect tendencies allows them to more effectively optimize tasks for human comfort and safety

Simply assuming that human collaborators will always choose a globally ergonomically optimal action if one is made available is problematic for two reasons: First, humans are often not effective global optimizers, and instead make myopic decisions based on immediate comfort

that may uncomfortable or infeasible actions to complete the task later. Second, we can't guarantee that the human will pick any particular action, because human action choice is noisy.

We can address both these problems by having our co-robots optimize for total human ergonomic cost in expectation. Optimizing total cost allows us to nudge even myopic humans toward action choices which result in globally optimal ergonomics in multi part tasks. Doing this optimization in expectation, over a probability distribution of the available human actions, lets us maximize the likelihood of the human choosing a low cost action even if we can't say which exact action they'll pick.

Chapter 4: Models of human action feasibility/preference can be learned online by co-robots

Giving co-robots personalized models of human action feasibility and preference can allow significant increases in safety and comfort for the robots' human collaborators. Manually constructing these personalized models for every individual is cumbersome and impractical though.

We can mitigate this problem by learning human preference models online during interaction with co-robots. This learning makes use of implicit queries, where the robot's action itself is the query, and the human's natural response is the label. This learning can start with an average belief learned over a whole population of people, and refine this belief during interaction to reflect an individual's unique preferences. Active learning techniques can accelerate this learning by selecting robot queries which produce maximally informative human responses.

5.2 Challenges and Future Directions

This work has many limitations, which will necessitate more work before these techniques can be applied to robots in real applications. Examples include:

Our personalized kinematic model estimation and tracking algorithm in Chapter 2 relies on data from a motion capture system, and requires the joint that each marker is attached to be specified manually. Motion capture is impractical in most home, industrial, and other application settings, and manual specification of marker assignments is too burdensome for the average user. The system could be adapted to use data from consumer-grade sensors such as the Kinect though. Even if motion capture is still used, approaches in the literature suggest that the marker/limb assignments could be generated automatically, rather than prespecified [87]. This automatic identification of kinematic structure could likely be extended to the number and location of kinematic degrees of freedom as well.

The ergonomic cost function we use throughout this dissertation – squared weighted configuration space distance – is somewhat ad hoc. It is adequate for differentiating between obviously unergonomic arm postures (arm over head, arm fully outstretched) and more

comfortable postures (arm by side with elbow flexed 90 degrees). For postures which are not so obviously different, however, its predictions are less reliable. Cost functions which more accurately reflect humans' true behavior, whether they simply have more model degrees of freedom or are more carefully derived from biomechanical data, should produce strictly superior results with any of the work here when substituted for our weighted distance cost.

In Chapters 3 and 4, the optimization used to find the minimum expected ergonomic cost uses simple brute force enumeration of every possible human and robot inverse kinematics solution. This is very slow, taking on the order of 15 minutes to solve for one optimal handoff, and precludes the use of these methods in online planning situations without precomputed solutions. In addition, enumerating the possible IK solutions requires us discretize a solution space that is, in general, a continuous manifold containing infinitely many solutions. This limits the precision of our final solution, and also requires us to implicitly choose a metric over the configuration space, in order to specify how far away each discrete IK solution should be from the surrounding solutions. A more elegant solution would pose this problem as a numerical optimization with continuous valued decision variables, in the same fashion as numerical inverse kinematics. This is challenging though, because the expected total cost minimization objective function considers the cost of *all* feasible configurations, not just one.

Aside from these limitations of the methods presented here, we've contemplated two other future extensions of this work seem that seem promising:

While our work models the kinematic structure of humans collaborating with robots, other objects in the environment have kinematic structure as well, like furniture in a home, cardboard packaging in a warehouse, or even patients in a hospital being lifted by nurses rather than moving under their own power. In all of these situations, the difference with our work is that the kinematic chain is *underactuated*. Some or all their degrees of freedom are moved only by forces from the environment, not controlled by individual muscles or motors, as we assumed here. Because the constraints on the motion of the articulated object's links, as well as the forces it exerts on the environment, are dependent upon its kinematic structure and current configuration, our kinematic modeling formalism is a natural fit for analyzing collaborative manipulation of these objects.

Another obvious extension is to consider not just static configurations as we did here, but full trajectories for collaborative motion. Both our analysis of motion constraints using personalized kinematic models and our modeling of human ergonomic cost using configuration cost functions extend nicely to this regime. The manipulator Jacobian of the combined human-robot system specifies the feasible directions of motion at any given configuration, and the static configuration cost can simply be integrated over an entire trajectory to give a running ergonomic cost.

In conclusion though, regardless of which specific applications collaborative robots find use in and which technologies are used to realize those robots, the features of human collaborators that we have explored in this dissertation will remain fundamental: 1)Humans are physically different and co-robots should model and adapt to these differences, and 2)humans are noisy, and co-robots will only ever be able to predict human actions, not directly control them. The work here represents a simple initial investigation into how to address

these fundamental problems in future human-robot collaborative systems.

Bibliography

[1] D. Acemoglu and P. Restrepo, "Robots and Jobs: Evidence from US Labor Markets," Working Paper 23285, National Bureau of Economic Research, Mar 2017.

[2] M. J. Hicks and S. Devaraj, "The myth and reality of manufacturing in America," Tech. Rep. June, 2015.

[3] C. B. Frey and M. A. Osborne, "The Future of Employment: How Susceptible Are Jobs To Computerisation?," pp. 1–72, 2013.

[4] C. C. Miller, "The Long-Term Jobs Killer Is Not China. It's Automation," Dec 2016.

[5] "Automation and Anxiety: Will smarter machines cause mass unemployment?," *The Economist*, Jun 2016.

[6] V. Bolhouse and B. Daugherty, "Chapter 56: Automotive and Transportation Applications," in *Handbook of Industrial Robotics*, Electrical and Electronic Engineering, ch. 56, p. 1069, Wiley, 1999.

[7] J. Mortimer, "BMW Lifts Robot Total to 500 at Plant Oxford," *Assembly Automation*, vol. 28, no. 1, pp. 27–35, 2008.

[8] US Bureau of Labor Statstics, "Incidence rates of nonfatal occupational injuries and illnesses by industry and case types," tech. rep., 2015.

[9] J. Martinez, M. J. Black, and J. Romero, "On human motion prediction using recurrent neural networks," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2891–2900, 2017.

[10] K. Fragkiadaki, S. Levine, P. Felsen, and J. Malik, "Recurrent Network Models for Human Dynamics," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, ICCV '15, (Washington, DC, USA), pp. 4346–4354, IEEE Computer Society, 2015.

[11] J. Mainprice and D. Berenson, "Human-Robot Collaborative Manipulation Planning Using Early Prediction of Human Motion," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 299–306, IEEE, 2013.

[12] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching Physical Collaborative Tasks: Object-Lifting Case Study with a Humanoid," in *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2009.

[13] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1994.

[14] V. A. Shia, Y. Gao, R. Vasudevan, K. D. Campbell, T. Lin, F. Borrelli, and R. Bajcsy, "Semiautonomous vehicular control using driver modeling," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2696–2709, 2014.

[15] K. Driggs-Campbell, R. Dong, S. S. Sastry, and R. Bajcsy, "Robust, Informative Human-in-the-Loop Predictions via Empirical Reachable Sets," *arXiv:1705.00748*, pp. 1–10, 2017.

[16] R. Bajcsy, "Active perception," *Proceedings of the IEEE*, vol. 76, no. 8, pp. 996–1005, 1988.

[17] B. Settles, "Active learning literature survey: Computer sciences technical report 1648," tech. rep., University of Wisconsin, Madison, 2010.

[18] A. Bestick, S. Burden, G. Willits, N. Naikal, S. S. Sastry, and R. Bajcsy, "Personalized kinematics for human-robot collaborative manipulation," in *IEEE International Conference on Intelligent Robots and Systems*, 2015.

[19] E. A. Sisbot and R. Alami, "A human–aware manipulation planner," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1045–1057, 2012.

[20] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, pp. 1986–1993, 2011.

[21] K. W. Strabala, M. K. Lee, A. D. Dragan, J. L. Forlizzi, S. Srinivasa, M. Cakmak, and V. Micelli, "Towards seamless human–robot handovers," *Journal of Human–Robot Interaction*, vol. 2, no. 1, pp. 112–132, 2013.

[22] M. Boussenna, E. N. Corlett, and S. T. Pheasant, "The relation between discomfort and postural loading at the joints," *Ergonomics*, vol. 25, no. 4, pp. 315–322, 1982.

[23] L. Punnett, L. J. Fine, W. M. Keyserling, G. D. Herrin, and D. B. Chaffin, "Back disorders and nonneutral trunk postures of automobile assembly workers," *Scandinavian Journal of Work, Environment & Health*, vol. 17, no. 5, pp. 337–346, 1991.

[24] W. M. Keyserling, M. Brouwer, and B. A. Silverstein, "A checklist for evaluating ergonomic risk factors resulting from awkward postures of the legs, trunk and neck," *International Journal of Industrial Ergonomics*, vol. 9, no. 4, pp. 283–301, 1992.

[25] W. S. Marras, S. A. Lavender, S. E. Leurgans, F. A. Fathallah, S. A. Ferguson, G. W. Allread, and S. L. Rajulu, "Biomechanical risk factors for occupationally related low back disorders," *Ergonomics*, vol. 38, no. 2, pp. 377–410, 1995.

[26] A. Burdorf and G. Sorock, "Positive and negative evidence of risk factors for back disorders," *Scandinavian Journal of Work, Environment & Health*, vol. 23, no. 4, pp. 243–256, 1997.

[27] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognitiom in parts from single depth images," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

[28] Autodesk, "Autodesk MotionBuilder." `http://www.autodesk.com/products/motionbuilder/`. Accessed: 2015-02-20.

[29] PhaseSpace, "PhaseSpace Motion Capture Software." `http://www.phasespace.com/software.html`. Accessed: 2015-02-20.

[30] OptiTrack, "OptiTrack Motive:Body Software." `http://www.optitrack.com/products/motive/body/`. Accessed: 2015-02-20.

[31] A. Ude, C. Man, M. Riley, and C. G. Atkeson, "Automatic generation of kinematic models for the conversion of human motion capture data into humanoid robot motion," in *Proceedings of the IEEE-RAS Conference on Humanoid Robotics*, 2000.

[32] G. Wu, F. C. van der Helm, H. (DirkJan) Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A. R. Karduna, K. McQuade, X. Wang, and et al., "ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand," *Journal of Biomechanics*, vol. 38, no. 5, pp. 981–992, 2005.

[33] G. Wu, S. Siegler, P. Allard, C. Kirtley, A. Leardini, D. Rosenbaum, M. Whittle, D. D. D'Lima, L. Cristofolini, H. Witte, and et al., "ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—Part I: ankle, hip, and spine," *Journal of Biomechanics*, vol. 35, no. 4, pp. 543–548, 2002.

[34] S. Delp and J. Loan, "A computational framework for simulating and analyzing human and animal movement," *Computing in Science Engineering*, vol. 2, no. 5, pp. 46–55, 2000.

[35] L. Ljung, *System Identification: Theory for the User*. Prentice–Hall, 1999.

[36] R. He, Y. Zhao, S. Yang, and S. Yang, "Kinematic–parameter identification for serial–robot calibration based on POE formula," *IEEE Transactions on Robotics*, vol. 26, no. 3, pp. 411–423, 2010.

[37] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," 2001.

[38] J. J. Moré, D. C. Sorenson, B. S. Garbow, and K. E. Hillstrom, "The MINPACK project," *Sources and Development of Mathematical Software*, pp. 88–111, 1984.

[39] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[40] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[41] E. Todorov, "Optimality principles in sensorimotor control," *Nature Neuroscience*, vol. 7, no. 9, pp. 907–915, 2004.

[42] J. Diedrichsen, R. Shadmehr, and R. B. Ivry, "The coordination of movement: optimal feedback control and beyond," *Trends in Cognitive Sciences*, vol. 14, no. 1, pp. 31–39, 2010.

[43] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *The Journal of Neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[44] Y. Uno, M. Kawato, and R. Suzuki, "Formation and control of optimal trajectory in human multijoint arm movement," *Biological Cybernetics*, vol. 61, pp. 89–101, June 1989.

[45] J. F. Soechting, C. A. Buneo, U. Herrmann, and M. Flanders, "Moving effortlessly in three dimensions: does donders' law apply to arm movement?," *The Journal of Neuroscience*, vol. 15, no. 9, pp. 6271–6280, 1995.

[46] G. Rao, D. Amarantini, E. Berton, and D. Favier, "Influence of body segments' parameters estimation models on inverse dynamics solutions during gait," *Journal of Biomechanics*, vol. 39, no. 8, pp. 1531–1536, 2006.

[47] E. P. Hanavan, Jr, "A personalized mathematical model of the human body.," *Journal of Spacecraft and Rockets*, vol. 3, no. 3, pp. 446–448, 1966.

[48] M. Hazewinkel, ed., *Encyclopedia of Mathematics*, ch. Student test. Springer, 2001.

[49] C. Bregler, J. Malik, and K. Pullen, "Twist based acquisition and tracking of animal and human kinematics," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 179–194, 2004.

[50] P. Baerlocher and R. Boulic, "An inverse kinematics architecture enforcing an arbitrary number of strict priority levels," *The Visual Computer*, vol. 20, no. 6, pp. 402–417, 2004.

[51] V. De Sapio, J. Warren, O. Khatib, and S. Delp, "Simulating the task-level control of human motion: a methodology and framework for implementation," *The Visual Computer*, vol. 21, no. 5, pp. 289–302, 2005.

[52] A. Bestick, R. Bajcsy, and A. Dragan, "Implicitly assisting humans to choose good grasps in robot to human handovers," in *International Symposium on Experimental Robotics*, 2016.

[53] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *IEEE International Conference on Intelligent Robots and Systems*, pp. 1986–1993, IEEE, 2011.

[54] C.-m. Huang, M. Cakmak, and B. Mutlu, "Adaptive coordination strategies for human-robot handovers," in *Robotics: Science and Systems (RSS)*, 2011.

[55] J. Kim, J. Park, Y. Hwang, and M. Lee, "Advanced grasp planning for handover operation between human and robot: three handover methods in esteem etiquettes using dual arms and hands of home-service robot," *2nd International Conference on Autonomous Robots and Agents*, pp. 34–39, 2004.

[56] J. Mainprice, E. A. Sisbot, T. Siméon, and R. Alami, "Planning safe and legible handover motions for human-robot interaction," in *IARP Workshop on Technical Challenges for Dependable Robots in Human Environments*, vol. 2, p. 7, 2010.

[57] V. Micelli, K. Strabala, and S. Srinivasa, "Perception and control challenges for effective human-robot handoffs," *RSS 2011 RGB-D Workshop*, 2011.

[58] A. H. Quispe, H. B. Amor, and M. Stilman, "Handover planning for every occasion," *IEEE-RSJ International Conference on Humanoid Robots*, 2014.

[59] A. H. Quispe, H. Ben Amor, H. Christensen, and M. Stilman, "It takes two hands to clap: towards handovers in bimanual manipulation planning," in *Robotics: Science and Systems (RSS)*, 2014.

[60] E. Sisbot and R. Alami, "A human-aware manipulation planner," *IEEE Transactions on Robotics*, vol. 28, pp. 1045–1057, Oct 2012.

[61] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinavasa, M. Cakmak, and V. Micelli, "Towards seamless human-robot handovers," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 112–132, 2012.

[62] B. Hayes and B. Scassellati, "Effective robot teammate behaviors for supporting sequential manipulation tasks," in *IEEE International Conference on Intelligent Robots and Systems*, pp. 6374–6380, 2015.

[63] D. Berenson, S. Srinivasa, and J. Kuffner, "Task Space Regions: A framework for pose-constrained manipulation planning," *The International Journal of Robotics Research*, vol. 30, no. 12, pp. 1435–1460, 2011.

[64] A. Bestick, R. Pandya, R. Bajcsy, and A. Dragan, "Learning human ergonomic preferences for handovers," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018 (in review).

[65] J. Mainprice, M. Gharbi, T. Siméon, and R. Alami, "Sharing effort in planning human-robot handover tasks," in *RO-MAN*, pp. 764–770, IEEE, 2012.

[66] A. Gritsenko and D. Berenson, "Learning Task-Specific Path-Quality Cost Functions From Expert Preferences," in *NIPS Workshop: Autonomously Learning Robots*, 2014.

[67] D. Sadigh, A. D. Dragan, S. Sastry, and S. A. Seshia, "Active Preference-Based Learning of Reward Functions," *Robotics: Science and Systems (RSS)*, 2017.

[68] C. Daniel, M. Viering, J. Metz, O. Kroemer, and J. Peters, "Active Reward Learning," *Robotics: Science and Systems (RSS)*, vol. 10, 2014.

[69] M. Cakmak, C. Chao, and A. Thomaz, "Designing Interactions for Robot Active Learners," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 2, pp. 108–118, 2010.

[70] D. Braziunas, "Computational Approaches to Preference Elicitation," tech. rep., University of Toronto, Department of Computer Science, 2006.

[71] E. Hüllermeier, J. Fürnkranz, W. Cheng, and K. Brinker, "Label ranking by learning pairwise preferences," *Artificial Intelligence*, vol. 172, no. 16-17, pp. 1897–1916, 2008.

[72] F. L. Wauthier, M. I. Jordan, and N. Jojic, "Efficient Ranking from Pairwise Comparisons," *International Conference on Machine Learning*, 2013.

[73] B. Carterette, P. N. Bennett, D. M. Chickering, and S. T. Dumais, "Here or There: Preference Judgments for Relevance," in *European Conference on Advances in Information Retrieval (ECIR)*, 2008.

[74] A. Karbasi, S. Ioannidis, and laurent Massoulie, "Comparison-Based Learning with Rank Nets," *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pp. 855–862, 2012.

[75] A. E. Abbas, "Entropy methods for adaptive utility elicitation," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, vol. 34, no. 2, pp. 169–178, 2004.

[76] C. Boutilier, "A POMDP Formulation of Preference Elicitation Problems," *Proceedings of AAAI Conference on AI*, pp. 239–246, 2002.

[77] J. Fürnkranz, E. Hüllermeier, W. Cheng, and S.-H. Park, "Preference-based reinforcement learning: a formal framework and a policy iteration algorithm," *Machine Learning*, vol. 89, no. 1-2, pp. 123–156, 2012.

[78] R. Holladay, S. Javdani, A. Dragan, and S. Srinivasa, "Active Comparison Based Learning Incorporating User Uncertainty and Noise,"

[79] G. Druck, B. Settles, and A. McCallum, "Active learning by labeling features," *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing Volume 1 - EMNLP '09*, 2009.

[80] J. Kulick, M. Toussaint, T. Lang, and M. Lopes, "Active learning for teaching a robot grounded relational symbols," *International Joint Conference on Artificial Intelligence*, 2013.

[81] C. Dima and M. Hebert, "Active Learning For Outdoor Obstacle Detection," in *Robotics: Science and Systems*, 2005.

[82] D. Sadigh, S. S. Sastry, S. A. Seshia, and A. Dragan, "Information gathering actions over human internal state," *IEEE International Conference on Intelligent Robots and Systems*, 2016.

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

[84] D. Ramachandran and E. Amir, "Bayesian inverse reinforcement learning," in *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, IJCAI'07, pp. 2586–2591, 2007.

[85] S. W. Yoon, A. Fern, R. Givan, and S. Kambhampati, "Probabilistic planning via determinization in hindsight," in *AAAI*, pp. 1010–1016, 2008.

[86] E. K. Chong, R. L. Givan, and H. S. Chang, "A framework for simulation-based network control via hindsight optimization," in *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, vol. 2, pp. 1433–1438, IEEE, 2000.

[87] A. G. Kirk, J. F. O'Brien, and D. A. Forsyth, "Skeletal parameter estimation from optical motion capture data," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 782–788, 2005.