Synthesizing Anticipatory Haptic Assistance Considering Human Behavior Uncertainty

José Ramón Medina, *Member, IEEE,* Tamara Lorenz, *Member, IEEE,* and Sandra Hirche, *Senior Member, IEEE*

Abstract

Intuitive and effective physical assistance is an essential requirement for robots sharing their workspace with humans. Application domains reach from manufacturing and service robotics via rehabilitation and mobility aids to education and training. In this context, assistance based on human behavior anticipation have shown superior performance in terms of human effort minimization. However, when robot's expectations mismatch human intentions undesired interaction forces appear incurring safety risks and discomfort. Human behavior prediction is therefore a crucial issue: it enables effective anticipation but potentially produces disagreements when prediction errors occur. In this article we present a novel control scheme for anticipatory haptic assistance where robot behavior adapts to prediction uncertainty. Following a data-driven stochastic modeling approach, robot assistance is synthesized solving a risk-sensitive optimal control problem, where the cost function and plant dynamics are affected by model uncertainty. The proposed approach is objectively and subjectively evaluated in an experiment with human users. Results indicate that our method outperforms other assistive control approaches in terms of perceived helpfulness and human effort minimization. [1]

Index Terms

Physical human-robot interaction, haptic assistance, stochastic optimal control, risk-sensitive control, learning by demonstration.

Authors are with the Institute for Information-Oriented Control, Department of Electrical Engineering and Information Technology, Technische Universität München, D-80290 Munich, Germany. {medina, lorenz, hirche}@tum.de

Manuscript received February 3, 2014; revised August 19, 2014; accepted November 20, 2014.

[1]Previous versions of this work has been published in [1] and [2]. This article have been rewritten containing additional analysis and experimental results.

I. INTRODUCTION

Physical human-robot interaction (pHRI) has wide and relevant applications covering, among other areas, manufacturing, service and care robotics, rehabilitation, medical robotics, education and training. In tasks such as object transport or movement aids where the interaction involves physical coupling in direct contact or through an object, a robotic assistant is expected to behave in an intuitive manner while minimizing the human workload. Traditional interaction control schemes mainly focus on the generation of compliant robot behavior, which is indispensable for human safety [3]. In addition, cognitive robots are expected to exhibit also goal-oriented behavior. By anticipating human actions, these *proactive* control schemes outperform purely reactive ones for tasks such as movement assistance [4] and in many cases are even necessary when environmental constraints are involved [5]. However, due to the continuous physical coupling between agents, when robot's expectations mismatch human intentions undesired interaction forces appear incurring safety risks and discomfort. Human behavior prediction is therefore a crucial issue: it enables effective anticipation but potentially produces disagreements when prediction errors occur.

In order to predict human actions proactive anticipatory control schemes rely on feed-forward human *behavior models* of motion, force or both of them. Due to the complexity of human decision-making, analytical models of human behavior in arbitrary tasks are rarely available. In this context, learning is an effective alternative. By means of probabilistic techniques, generalized human task-solving behavior is acquired in a data-driven manner. However, learned models (as well as potential analytical models) are usually far from perfect: prediction errors frequently occur, inducing counteracting forces during haptic interaction thereby significantly diminishing assistive performance. Therefore, an a priori estimation of potential errors or, equivalently, the expected prediction accuracy enriches the prospective capabilities of the robot's assistance and potentially enhances interaction. This estimation is given by the prediction uncertainty in probabilistic models or similarly, by the expected prediction variance considering statistical models. Although usually neglected in control design for pHRI, model uncertainty plays a key role in a decision-making process in cooperative tasks when an interacting partner is involved [6].

The contribution of this article is a novel anticipatory model-based haptic assistance scheme that considers model uncertainty in the robot interaction control. Formulated as an optimal control

problem, the robotic assistant adapts its control depending on two sources of uncertainty encoded in a previously learned human behavior model. First, predicted motion uncertainty is included in the optimization criterion by considering the Mahalanobis distance to the expected human-desired trajectory. Second, an estimation of the expected and observed human force variability caused by disagreements with the robot affects the robot dynamics as an additive stochastic input. In order to explicitly take this variability into account, we propose a risk-sensitive control approach. The solutions for this optimal control problem depend on a risk-sensitivity parameter, which defines the attitude towards the partner in case of disagreement. The assistive performance of the proposed interaction control scheme is evaluated in a psychological experiment with naive human users. Not only objective measures such as human effort, power or disagreement are evaluated but also a subjective evaluation of the perceived helpfulness. Results indicate better performance in terms of human effort minimization and perceived helpfulness when uncertainty is considered.

A. Related Work

The design of physical assistants using proactive control schemes is a very recent topic and only a limited body of work is found in the literature. Concerning the acquisition of human behavior models for proactive assistance, the application of learning techniques in pHRI tasks by applying the PbD paradigm [7] have recently gained remarkable attention. Acquiring demonstrations through teleoperation, human-like interactive behavior is reproduced considering learned motions and interaction forces [8]–[10]. Interpreting the interaction as a continuous teaching process, the robot may evolve from a passive follower into an active contributor using incremental learning techniques [11], [12]. As an alternative to learned models, planning approaches also provide suitable task solving behavior when both the common goal and the environment are known. Especially compelling are feedback motion planning methods [13], [14], which continuously adapt their plans when human deviations occur [15]. Analytical behavior models have also been explored in simple point to point settings, such as the well-known minimum jerk principle [16]. During cooperative manipulation tasks, a polynomial extrapolation becomes more suitable [17]. However, all the aforementioned methods neglect prediction uncertainty or interaction force variability due to disagreements in their assistive control schemes.

Works implementing a proactive assistant considering uncertainty are limited to [18], where

the robotic role is adapted depending on prediction confidence following the framework presented in [19]. A deeper insight in uncertainty-dependent decision-making is found in other application domains, especially in the stochastic optimal control literature [20], [21]. Following this idea human sensorimotor control have been successfully modeled considering uncertainty as an additive term of the optimization criteria [22] or by means of a risk-sensitive optimization [23], [24].

Limited amount of work is found in the literature concerning disagreements with the human partner and their effects in robot control. A haptic negotiation adaptation between two possible roles is proposed in [25]. Depending on human force, the role of the robot switches between dominant and recessive, achieving better performance than equally shared control, as indicated in a human user study [26]. Another option smoothly adapts the role of the robot depending on the human force magnitude while following a predefined trajectory [27]. A large-scale human user study shows the advantages of the dynamic role allocation scheme. Approaching the design of a robotic assistant from an optimality point of view, a suitable and general approach, together with a taxonomy of interactive motor behaviors, is presented in [28]. The optimal controller minimizes a cost function penalizing the euclidean distance to the human desired trajectory, the robot desired trajectory in addition to quadratic norms of both partners' force.

B. Notation

By convention, bold characters are used for vectors and capital letters denote matrices. The expression $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ describes a normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. $\|\cdot\|$ denotes the Euclidean norm. $\mathsf{E}[x]$ and $\mathsf{Var}[x]$ denote the expected value and the variance of x respectively.

The remainder of this article is structured as follows. Section II describes the interaction control scheme. The proposed anticipatory assistive control is explained in Section III. An implementation of the 2-dimensional translational case is detailed in Section IV. Results of simulations and of the experimental user study are presented in Section V.

II. Modeling physical human-robot interaction

In this work we consider the class of physical human-robot interaction tasks where the robot is physically coupled to a human partner. The goal is to reach a goal configuration starting from an initial configuration. This prototypical setting is representative for many different tasks such as mobility assistance to humans or joint object transport/manipulation. Depending on

the application, the interaction can be through an external object, as in cooperative manipulation tasks, or at the end-effector as in movement assistance for elderly or disabled or in exoskeletons. In case of the interaction through an object, its geometry together with the grasping points lead to a decomposition into redundant and non-redundant forces [27]. Redundant components can be exerted by either the robot or the human and are instrumental for the haptic negotiation process. In this work, we focus our attention into interaction forces which only appear in redundant degrees of freedom. As a representative case for this setting, we consider a common interaction contact point between the robot and its human partner at the end-effector. For simplicity of exposition we assume that the nonlinear robot dynamics is feedback-linearized and an impedance control law renders the robot reactive behavior [29] to

$$M_r\ddot{\boldsymbol{x}}(t) + D_r\dot{\boldsymbol{x}}(t) = \boldsymbol{u}(t) = \boldsymbol{u_h}(t) + \boldsymbol{u_r}(t) \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^6$ is the pose of the end-effector, $\boldsymbol{u_h} \in \mathbb{R}^6$ is the applied wrench by the human, $\boldsymbol{u_r} \in \mathbb{R}^6$ is the assistive control input of the robot, and $M_r, D_r \in \mathbb{R}^{6\times 6}$ are positive definite matrices representing the rendered inertia and viscous friction respectively. For later convenience, we discretize the system from (1) with a sampling time interval Δ, yielding impedance controlled dynamics

$$\boldsymbol{\xi}_{k+1} = \boldsymbol{f}(\boldsymbol{\xi}_k, \boldsymbol{u}_{\boldsymbol{h},k} + \boldsymbol{u}_{\boldsymbol{r},k}) \ , \tag{2}$$

where k is the time index such that $t = k\Delta$ and $\boldsymbol{\xi}_k = \ [\ \boldsymbol{x}_k\ \dot{\boldsymbol{x}}_k\ ]^\mathsf{T}$ the state of the system [2].

Dynamics (1) compensates for small deviations between the human and robot intended motions. In addition to this *reactive behavior* represented by the compliance control, the robot implements an anticipatory *proactive behavior* given by $\boldsymbol{u_r}$. To this end, the robot relies on a behavior model λ, which provides predictions of the next human desired state and human wrench in terms of normal distributions $\hat{\boldsymbol{\xi}}_{\boldsymbol{d},k} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},k}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k})$ and $\hat{\boldsymbol{u}}_{\boldsymbol{h},k} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{u_h},k}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u_h},k})$. The behavior model represents expected trajectories based on multiple previously performed trials. The resulting normally distributed state space predictions represent the trial to trial statistical mean and variance of human desired motion. In addition, the predicted human wrench variability also captures counteracting forces when the human and the robot disagree. This effect is not observable in state space trajectories. Consider as an example a binary decision problem while

[2]The time-discretized system dynamics will be specified in Section IV for a two-dimensional example.

approaching an obstacle, where both agents disagree on which way to go, either left or right. We further assume that both partners have a common goal and deviations from the predicted trajectory induced by the human are small and do not express a divergence or a tendency towards a different goal. Given this problem setting, the focus of this work lies on the design of a possibly time-varying assistive control law synthesizing $\boldsymbol{u_r}$ taking uncertainties into account.

III. Anticipatory Control

In this work, the objective governing haptic assistance is human force minimization. From an optimality point of view, this concept is expressed by the minimization of a cost function in the form

$$J = \sum_{k=0}^{T-1} \|\boldsymbol{u_{h},_k}\|^2 \tag{3}$$

where T is the time horizon for the optimization [3]. The human has an urge towards a goal as predicted by the task model λ but relying on a human force estimation of T steps potentially accumulates high tracking errors in task space. A more suitable alternative considers the minimization problem in task space based on the tracking error between task model predictions $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}} = \{\hat{\boldsymbol{\xi}}_{\boldsymbol{d},0} \cdots \hat{\boldsymbol{\xi}}_{\boldsymbol{d},T}\}$ and the current state $\boldsymbol{\xi}$, i.e.

$$J = \sum_{k=0}^{T} \|(\hat{\boldsymbol{\xi}}_{\boldsymbol{d},k} - \boldsymbol{\xi}_k)\|^2 \ .$$

The solution to this optimization problem leads to controller which does not consider any energy expenditure of the robot nor any torque constraints of the robot actuators. In addition, disturbances to the tracking may result in rather abrupt behavior of the robot, which might not be desirable when targeting intuitive assistance. In order to keep the computational complexity low to achieve real-time computation, we add a soft penalty on the energy expenditure in terms of a quadratic term. As a result the cost function takes the form as in the classical linear quadratic optimal control problem

$$J = \|(\hat{\boldsymbol{\xi}}_{\boldsymbol{d},T} - \boldsymbol{\xi}_T)\|^2_{Q_T} + \sum_{k=0}^{T-1} \|(\hat{\boldsymbol{\xi}}_{\boldsymbol{d},k} - \boldsymbol{\xi}_k)\|^2_{Q_k} + \|\boldsymbol{u_{h},_k}\|^2_{R_k} \tag{4}$$

[3]Note that we will later employ a receding horizon control scheme, i.e. without fixed final time T. For simplicity of presentation we consider for now the finite horizon control problem (3)

where $\|x\|_Q^2$ stands for the quadratic form $x^{\mathrm{T}}Qx$, Q_T the final cost weight and Q_k and R_k are weighting factors that define the trade-off between tracking performance and robot contribution. Observe that two factors challenge the equivalence between (3) and human intentions: i) the prediction error of $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$; ii) the weighting factors Q_k and R_k may differ from human preferences. As perfect predictions and weighting factors design is usually not feasible when dealing with humans, corrections/disagreements from the human partner are expected. This source of variability, together with the normally distributed nature of predictions define the uncertainty that the robotic assistance needs to cope with. In the following we analyze the consequences of uncertainty in the robot behavior and synthesize their effect as variations in the cost function to optimize (4) and in the predicted dyad dynamics.

A. Considering motion prediction uncertainty

Cost function (4) depends on the sequence of multivariate normal distributions $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$. A classical realization of the cost function for optimal tracking control measures the Euclidean distance to the reference's mean, i.e.

$$J = \|(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},T} - \boldsymbol{\xi}_T)\|_{Q_T}^2 + \sum_{k=0}^{T-1} \|(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},k} - \boldsymbol{\xi}_k)\|_{Q_k}^2 + \|\boldsymbol{u}_{\boldsymbol{h},k}\|_{R_k}^2 \ . \tag{5}$$

However, in order to accordingly measure the distance to a multivariate Gaussian, the weighted *Mahalanobis distance* is a suitable option as it includes the covariance of the prediction into the distance metric. In this case, the cost function becomes

$$J = \|(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},T} - \boldsymbol{\xi}_T)\|_{\hat{Q}_T}^2 + \sum_{k=0}^{T-1} \|(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},k} - \boldsymbol{\xi}_k)\|_{\hat{Q}_k}^2 + \|\boldsymbol{u}_{\boldsymbol{h},k}\|_{R_k}^2 \ , \tag{6}$$

where $\hat{Q}_k = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}^{-\frac{1}{2}} Q_k \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}^{-\frac{1}{2}}$ and $\hat{Q}_T = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{-\frac{1}{2}} Q_T \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{-\frac{1}{2}}$ and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},i}$ is per definition a symmetric positive-definite matrix $\forall i \ \ 0 \leq i \leq T$. This formulation diminishes the cost of tracking errors in directions with high prediction variance while it increases the penalty of deviations in regions with low variance. This is a desirable behavior in most applications where a a low variance of the learned desired trajectory over repeated trials indicates some importance to keep track of it. As an example consider the situation depicted in Fig. 1. Low variance directions might be caused for example by environmental constraints while high variance regions imply unconstrained areas. This suggests flexibility in unconstrained directions and strict tracking through narrow passages,

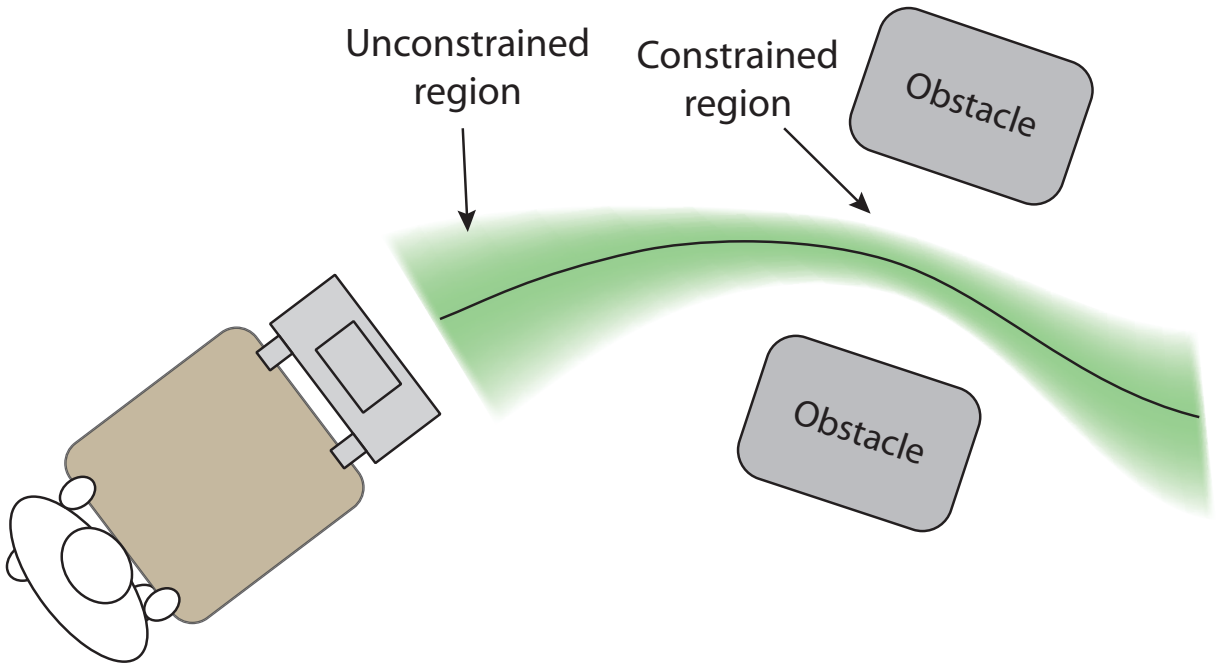


Fig. 1. Exemplary representation of the predicted state mean and variance $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}} \sim \{\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\}$. Low variance directions suggest the possibility of constrained passages, while high variances areas imply unconstrained areas.

a behavior in accordance with the minimization of the Mahalanobis distance as in (6). Note that the opposite effect is achieved using the inverse of the covariance matrix, i.e. $\hat{Q}_k = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}^{\frac{1}{2}} Q_k \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}^{\frac{1}{2}}$ and $\hat{Q}_T = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{\frac{1}{2}} Q_T \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{\frac{1}{2}}$. In this case the cost increases inversely with the variance: directions with low predicted variance diminish their cost with respect to high variance regions, which increase their penalty. The idea of employing the measure of variance as a way to inform the system on how systematically it should follow a reference trajectory was also applied in [30] for an impedance control scheme and in [31] for kinematic control.

B. Considering disagreement

In addition to motion prediction, the task model λ may also provide human wrench predictions $\hat{\boldsymbol{u}}_{\boldsymbol{h}}$. We assume that there exists a human desired nominal motion trajectory. In consequence, if the task model λ contains sufficiently many trials, the mean of the predicted trajectory $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$ approaches the human desired nominal motion trajectory. If the robot exerts the necessary wrench to track this trajectory, we expect the human wrench's statistical mean tending to 0, i.e. $\mathsf{E}[\boldsymbol{u}_{\boldsymbol{h}}] = 0$. However, the human wrench's covariance encodes valuable information: regions

with high variance are caused by significant or recurrent inter-trial corrections, meaning a conflictive area, see Fig. 2 for an illustration. To model this unbiased variability, we denote the *expected disagreement* with the human partner $\boldsymbol{\varepsilon_e} \sim \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u_h}})$. Applying the same idea to current observations of $\boldsymbol{u_h}$, the *current disagreement* is denoted as $\boldsymbol{\varepsilon_c} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u_h}})$. An estimate of the current disagreement is computed empirically at time k as the sample covariance over a window of W samples

$$\boldsymbol{\Sigma}_{\boldsymbol{u_h}} = \frac{1}{W} \sum_{i=k-W+1}^{k} \boldsymbol{u}_{\boldsymbol{h},i} \boldsymbol{u}_{\boldsymbol{h},i}^{\mathsf{T}} \ . \tag{7}$$

To consider both estimations, see Fig. 2 for visualization, the maximum of both disagreements

$$\boldsymbol{\varepsilon} = \max\left(\boldsymbol{\varepsilon_e}, \boldsymbol{\varepsilon_c}\right) \ , \tag{8}$$

is considered a reasonable solution. Note that keeping the zero mean assumption for the current disagreement is a conservative approach. In fact, the intra-trial mean of the human wrench along a window of W samples may differ from zero. Interpreting human wrench corrections as unbiased results in higher uncertainty compared to an interpretation as a biased variable. The consequences of this design will be more apparent in the analysis of the optimal control solution.

Disagreements result in variability of the dyad's motion due to conflicts with the human partner. Expressing this idea in mathematical terms, dynamics (2) contains an additional additive disagreement vector $\boldsymbol{\varepsilon}$, yielding the stochastic dyad dynamics

$$\boldsymbol{\xi}_{k+1} = \boldsymbol{f}(\boldsymbol{\xi}_k, \boldsymbol{u}_{\boldsymbol{r},k} + \boldsymbol{\varepsilon}_k) \ . \tag{9}$$

Given initial state $\boldsymbol{\xi}_0$, the optimal robotic assistance results from minimizing cost function (6), i.e.

$$\min_{\boldsymbol{u}_{r,0} \cdots \boldsymbol{u}_{r,T-1}} J(\boldsymbol{\xi}_0, \boldsymbol{u}_{r,0} \cdots \boldsymbol{u}_{r,T-1}) \ , \tag{10}$$

constrained to the dyad dynamics (9). As the dynamics are constantly updated with the current disagreement from (7), a recalculation of the solution is necessary. To this end, we adopt a Model Predictive Control (MPC) scheme [32], where the dynamics (9) are considered stationary for the optimization horizon T and the optimal control problem is solved at every time step. Cost function (6) at time step k formulated in a receding horizon fashion becomes

$$J_k = \|(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},k+T} - \boldsymbol{\xi}_{k+T})\|^2_{\hat{Q}_T} + \sum_{i=k}^{k+T-1} \|(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},i} - \boldsymbol{\xi}_i)\|^2_{\hat{Q}_i} + \|\boldsymbol{u}_{\boldsymbol{r},i}\|^2_{R_k} \ . \tag{11}$$

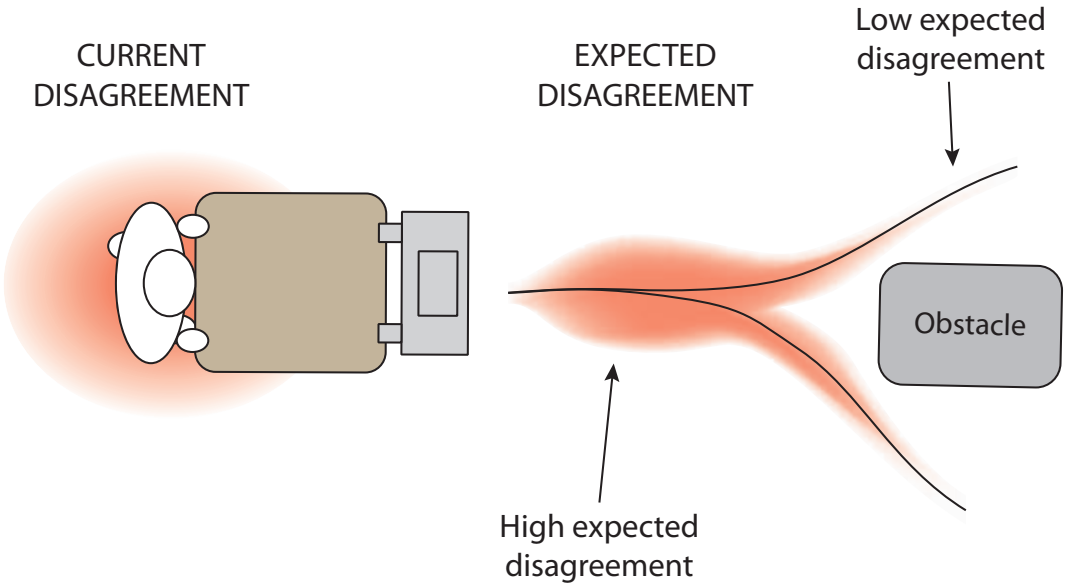


Fig. 2. Schematic representation of the expected disagreement ε_e and the current disagreement ε_c.

Although the optimization is solved for a horizon of T steps, only the solution for the first step is applied as assistive robot control.

C. Risk-sensitive optimization

Problem (10) is a stochastic optimization problem due to the stochastic nature of the dynamics (9). Traditionally it is solved by minimizing the *expected cost*, $\mathsf{E}[J]$. However, the uncertainty induced by the human behavior represented in ε is not characterized by its expected value but its higher order moments. In order to capture the uncertain human influence in the dynamics here we propose to employ a risk-sensitive optimization [33], [34]. The risk-sensitive solution considers all infinite cumulants of the cost distribution [35]. The risk-sensitive optimization criterion considers a cost function in the form

$$\gamma_k(\theta) = -2\theta^{-1} \ln \mathsf{E}[\exp^{-\frac{1}{2}\theta J_k}] \ , \tag{12}$$

where $\theta \in \mathbb{R}$ is the risk-sensitivity parameter and J_k is given in (11). The Taylor series expansion of $\gamma(\theta)$ around $\theta = 0$,

$$\gamma(\theta) = \mathsf{E}[J] - \frac{1}{4}\theta \, \mathsf{Var}[J] + ... \ ,$$

shows the influence of higher order moments in the solution, parameterized by θ. If $\theta = 0$ the controller is risk-neutral and corresponds to the classical optimization of the expected cost. For $\theta < 0$ and $\theta > 0$ the controller becomes risk-averse and risk-seeking, respectively. In our robot assistance scenario, the variance of the cost is determined by previously experienced and current disagreements with the human partner (8), and θ determines its assessment. In the risk-seeking case, disagreement is considered beneficial and decreases the resulting cost. In contrast, a risk-averse controller considers disagreements as a detrimental influence and the cost increases.

IV. The 2-dimensional Translational Case

The risk-sensitive optimization problem can be solved efficiently through the Riccatti equation for linear dynamics. Realtime computations are fundamental for intuitive haptic assistance; the physical coupling between human and robot requires immediate adaptation. To satisfy this prerequisite, we limit our study to a linear implementation. We consider now dynamics (1) in the plane, i.e. $\boldsymbol{x} \in \mathbb{R}^2$. In this case, the discrete-time dynamics from (9), written in the form $\boldsymbol{\xi}_{k+1} = A\boldsymbol{\xi}_k + B\boldsymbol{u}_k$ are given by

$$\begin{bmatrix} \boldsymbol{x}_{k+1} \\ \dot{\boldsymbol{x}}_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta \\ 0 & 1 - M_r^{-1} D_r \Delta \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_k \\ \dot{\boldsymbol{x}}_k \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & M_r^{-1}\Delta \end{bmatrix} (\boldsymbol{u}_{\boldsymbol{r}k} + \boldsymbol{\varepsilon}_k) \ . \tag{13}$$

where $M_r, D_r \in \mathbb{R}^{2\times 2}$.

The disagreement $\boldsymbol{\varepsilon}_k$ from (8), given by the maximum of the expected and current disagreement is approximated by $\mathcal{N}(0, \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{u},k})$, $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{u},k}$ being the Löwner-John hyperellipsoid [36]. This Gaussian approximation calculates the minimum volume hyperellipsoid around the set defined by $\boldsymbol{\varepsilon}_{\boldsymbol{e},k} \sim \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u_h},k})$ and $\boldsymbol{\varepsilon}_{\boldsymbol{c},k} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u_h},k})$.

For this particular problem setting and for comparison purposes we present two different solutions: a classical linear quadratic regulator (LQR), which considers the reference prediction uncertainty (6) but discards the variability of the disagreement in its optimal solution and, in contrast, a linear exponential quadratic regulator (LEQR) which considers both uncertainty sources in its risk-sensitive optimization (12).

A. Linear Quadratic Regulator Solution

The minimization of the expected cost $\mathsf{E}[J_k]$ with J_k as in (11) and dynamics (13) considering only the regulation of the desired trajectory is an instance of an LQR problem. The solution to this problem at sample time k for a receding horizon of T steps yields a feedback control law given by

$$\boldsymbol{u_{r},}_{i} = -K_i(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi},i} - \boldsymbol{\xi}_i) \ , \tag{14}$$

where K_i is the feedback matrix of the Ricatti recursion

$$K_i = -R_i^{-1}B^{\mathsf{T}}(BR_i^{-1}B^{\mathsf{T}} + \Pi_{i+1}^{-1})^{-1}A \tag{15}$$

and the cost-to-go is given by

$$\Pi_i = \hat{Q}_i + A^{\mathsf{T}}(BR_i^{-1}B^{\mathsf{T}} + \Pi_{i+1}^{-1})^{-1}A \ , \tag{16}$$

being $\Pi_T = \hat{Q}_T$.

Due to the receding horizon nature, only the feedback matrix for time step $i = k$, K_k, is applied in the system. It is in the form

$$K_k = \begin{bmatrix} 0 & 0 \\ K_{x,k} & K_{v,k} \end{bmatrix} , \tag{17}$$

where $K_{x,k}$ and $K_{v,k}$ are the position and velocity gains respectively.

The resulting assistive tracking control law synthesizes a variable impedance control law with its gains depending on prediction uncertainty. This becomes more obvious when applying the control law for the limit $\Delta \to 0$ in the system dynamics (1)

$$M_r\ddot{\boldsymbol{x}} + D_r\dot{\boldsymbol{x}} = \boldsymbol{u_h} - (K_x(t)(\boldsymbol{x_d} - \boldsymbol{x}) + K_v(t)(\dot{\boldsymbol{x}}_{\boldsymbol{d}} - \dot{\boldsymbol{x}})) \ , \tag{18}$$

where $\boldsymbol{x_d}$, $\dot{\boldsymbol{x}}_{\boldsymbol{d}}$, $K_x(t)$ and $K_v(t)$ are the position and velocity components of $\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}$ and the feedback gains $K_{x,k}$ and $K_{v,k}$ in the limit $\Delta \to 0$ respectively.

This control strategy tracks the predicted trajectory considering its variance by means of the Mahalanobis distance present in $\hat{Q}_i$ and $\hat{Q}_T$. As a result, the robot gets stiffer in the directions where the motion prediction has low variance and becomes more compliant for high variance. However, the variability induced by disagreements with the human partner is neglected.

B. Linear-Exponential Quadratic Regulator

The optimization problem considers now cost function (12) with J_k as in (11). In this case, the resulting feedback law as in (14) given by a modified form of the Ricatti recursion [37]

$$K_i = -R_i^{-1}B^{\mathsf{T}}(BR_i^{-1}B^{\mathsf{T}} + \theta B\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{u},i}B^{\mathsf{T}} + \Pi_{i+1}^{-1})^{-1}A \ , \tag{19}$$

and

$$\Pi_i = \hat{Q}_i + A^{\mathsf{T}}(BR_i^{-1}B^{\mathsf{T}} + \theta B\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{u},i}B^{\mathsf{T}} + \Pi_{i+1}^{-1})^{-1}A \ , \tag{20}$$

with $\Pi_T = \hat{Q}_T$.

As in the LQR case, motion prediction uncertainty is equivalently taken into account by means of the Mahalanobis considered in $\hat{Q}_i$ and $\hat{Q}_T$. However, the additive term $\theta B\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{u},i}B'$ present in the risk-sensitive solution leads to different results. The risk-neutral case $\theta = 0$ yields the same solution as for the LQR canceling the additional term. For a risk-seeking optimization $\theta > 0$ the expected variability $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{u},i}$ is assumed to be collaborating as if it was doing part of the work towards fulfilling the task. Accordingly, the resulting gain from (19) decreases adopting a more compliant behavior. Understanding the risk-sensitivity parameter as the robot's attitude, a *recessive* attitude is achieved implementing a risk-seeking controller: the robot becomes compliant under disagreement with its partner. In contrast, the risk-averse solution $\theta < 0$ increases the overall cost as if variability were directing the system towards an undesired state. As a result, gains are increased and the robot becomes stiffer. This case corresponds to a *dominant* attitude by generating an aggressive response to disagreements.

V. Experiments

In order to illustrate the peculiarities and advantages of the proposed controller, we first observe the influence of both prediction uncertainty and negotiation variability on a one-dimensional simulation. For evaluation purposes, we further study human preferences performing a user study where 19 participants interact with 7 different assistive controllers in a virtual scenario using a haptic interface.

A. One-dimensional simulation

As an exemplary scenario, we study a one-dimensional mass-damper system with dynamics equivalent to (13), $M = 1$ kg and $D = 1$ Ns/m. We first analyze the influence of motion prediction uncertainty using the Mahalanobis distance on the robot assistive behavior following

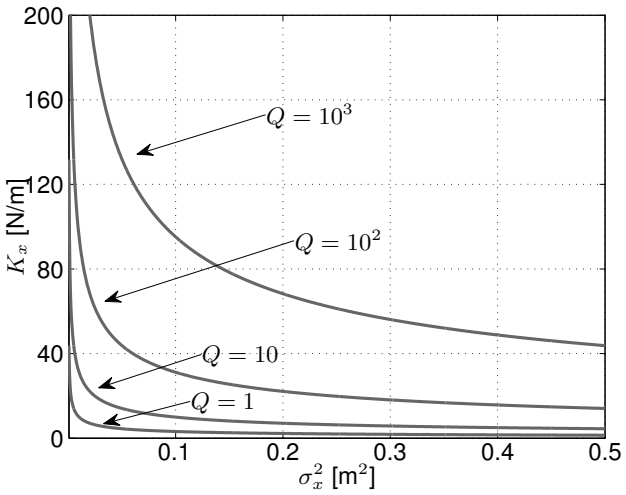


Fig. 3. Infinite horizon position gains w.r.t position prediction variance for several precision parameters Q with $R = I$.

the control scheme proposed in Section IV-A. We assume constant weighting matrices $Q = Q_k$ and $R = R_k$. The optimal position tracking gains for an LQR control optimizing the cost (6) for the infinite horizon $T \to \infty$ w.r.t. prediction variance σ_x^2 are depicted in Fig. 3. The infinite horizon solution provides a stationary feedback gain [38]. The optimal gains decrease with increasing prediction variances, i.e. low prediction variance produces a stiffer robot assistive behavior while high uncertainty leads to a more compliant assistive behavior. Note that, when variances tend to zero, gains tend to infinity. This degenerate case is easily avoided adding a regularization quantity, ensuring positive definiteness. For the case of the identity matrix, the weighting matrices of (6) become $\hat{Q}_k = (I + \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k})^{-\frac{1}{2}} Q (I + \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k})^{-\frac{1}{2}}$ and $\hat{Q}_T = (I + \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T})^{-\frac{1}{2}} Q_T (I + \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T})^{-\frac{1}{2}}$. As a result, when variances tend to 0, the cost generalizes to (5).

The influence of disagreement σ_u^2 on the optimal gains of a risk-sensitive controller from Section IV-B is shown in Fig. 4 for several values of θ. While the risk-neutral solution $\theta = 0$ ignores disagreement, risk-seeking solutions $\theta > 0$ decrease gains as the disagreement increases. In contrast, risk-averse solutions $\theta < 0$ increase tracking gains with increasing disagreement. An intuitive explanation in our scenario follows. A risk-seeking assistant assumes disagreement accounts for part of the tracking task and reduces its gains adopting a recessive attitude. In contrast, a risk-averse assistant understands disagreement as counter-productive and increases its gains in order to perform the tracking task while correcting expected deviations.

In order to illustrate how disagreement affects the dyad behavior during negotiation, we

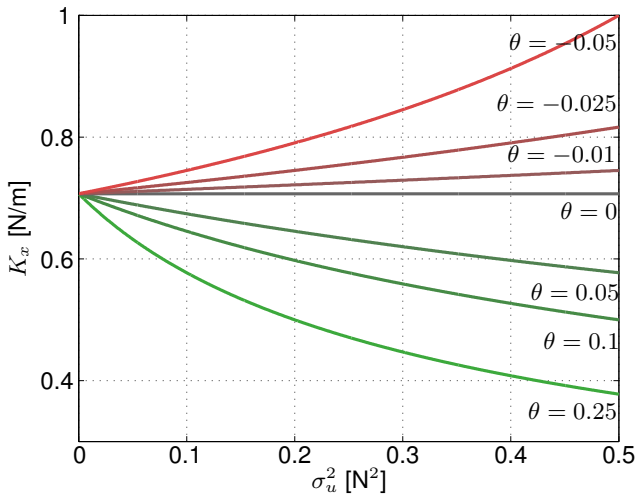


Fig. 4. Infinite horizon position gains w.r.t disagreement variance for several risk-sensitivities in an optimization with $Q = R = I$.

simulate a coupled dyad consisting of agents a and b, with control inputs $\boldsymbol{u_a}$ and $\boldsymbol{u_b}$ respectively. Consider agent b behaves as a PD-controller with constant gain, i.e.

$$\boldsymbol{u_{b},}_{k} = K_b(\boldsymbol{\xi_b} - \boldsymbol{\xi}_k) \ . \tag{21}$$

In contrast, agent a implements a risk-sensitive control strategy and tracks a constantly conflicting reference $\boldsymbol{\xi_a} = -\boldsymbol{\xi_b}$. We focus now on the effects that the current disagreement, computed according to (7) with $W = 1$, produces on agent a. For simplicity of illustration we neglect the expected disagreement. Note that the expected disagreement or the maximum of both disagreements produce equivalent effects on the control gains. The simulation results for a risk-neutral, risk-seeking and risk-averse implementation are depicted in Fig. 5. As shown in Fig. 5(a), the risk-seeking control remains almost inactive in comparison with the risk-neutral controller depicted in Fig. 5(b). In contrast, the risk-averse controller reacts more aggressively, as shown in Fig. 5(c). These three different attitudes yield different disagreement levels, depicted as gray regions. As the risk-seeking controller ceases tracking its conflicting reference, the disagreement is significantly reduced, while the risk-averse implementation provokes even higher disagreement levels than the risk-neutral solution. This behavior is clearly explained by the resulting gains and positions for all three controllers, as depicted in Fig. 6. While the risk-neutral solution ignores disagreement producing constant gains, the risk-seeking controller's almost cancels its feedback

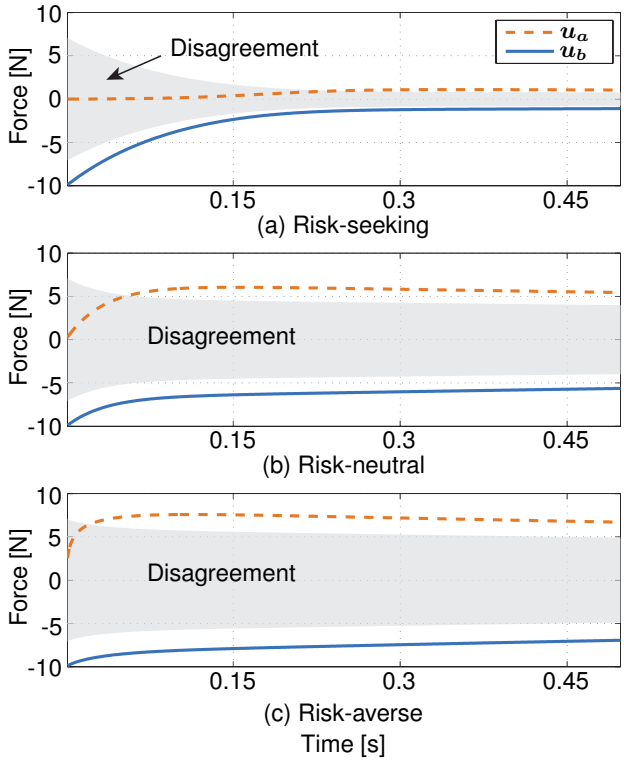


Fig. 5. Simulation results for the one-dimensional linear mass-damper system example with initial state $\boldsymbol{\xi_0} = [0\ 0]^\mathsf{T}$, references $\boldsymbol{\xi_b} = -\boldsymbol{\xi_a} = [-1\ 0]^\mathsf{T}$, mass $m = 1$ kg, damping $d = 1$ Ns/m with two agents, agent b following the control law (21) with gain $K_a = 1$ N/m and agent a applying the proposed control scheme in three variants. All three controllers are calculated with parameters $Q = I$ and $R = 5 \cdot 10^{-4} \cdot I$ and sample time of 1 kHz . Fig. 5(b) shows the simulation result for the risk-neutral control, Fig. 5(a) for a risk-seeking control with $\theta = 10^6$ and Fig. 5(c) a risk-averse control with $\theta = -4.5 \cdot 10^2$. Grey regions represents the variance of disagreement considered by agent a.

gain due to the initial conflicts, adopting a recessive attitude. In contrast, the risk-averse controller increases its gain substantially reacting in a dominant fashion.

B. User Study

To evaluate the proposed assistive control schemes and the preferences of humans interacting with it, we designed an experiment in which a human subject together with a virtual assistive robot transport a virtual object from an initial position towards a defined goal position.

1) Human Behavior Model Acquisition: A model representing the pHRI task must reflect both the human task execution preferences and the interaction with the robotic partner. These requirements suggest the use of learning by demonstration techniques. In our setting, a preliminary model of the task was acquired by initially letting the human lead, i.e., $\boldsymbol{u_r} = 0$. With this

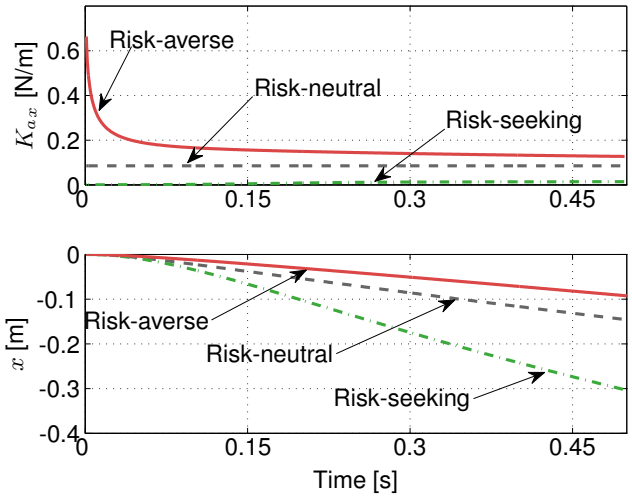


Fig. 6. Resulting tracking gains (a) and trajectories (b) for agent a in the simulated example from Fig. 5.

first rough representation, the robot actively assists and additionally observes interaction patterns during further task trials. Observed motion and human wrench trajectories are the training set for the model learning method yielding λ. Aiming for a safe and conservative interaction, we use a time-based left-to-right Hidden Markov Model applying regression on time domain instead of state space domain, as explained in [39]. As a result, the model provides a generalized trajectory of normal distributions of human-desired motion $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$ and applied wrenches $\hat{\boldsymbol{u}}_{\boldsymbol{h}}$. Although restricted to this trajectory, predictions always belong to the same homotopy class as demonstrated trajectories. Considering a window over last observations, we use the Viterbi algorithm as a filter that estimates the human state as a time index of the generalized trajectory. See [11] for a detailed explanation of the applied method.

2) Experimental Setup: The human subject applies forces to a haptic interface in order to move the virtual object, as shown in Fig. 7. It consists of a two degrees-of-freedom linear-actuated device (ThrustTube) which has a free-spinning handle at the grasping point, i.e. $\boldsymbol{x} \in \mathbb{R}^2$. Attached to the handle, a force/torque sensor (JR3) measures the human force input. The control scheme running at 1 kHz and implemented in *MATLAB/Simulink* is executed on a personal computer with *Linux PREEMPT Real-Time kernel* using *Matlab's Real-Time Workshop*. The shared object from dynamics (13) is physically rendered with a mass of $M_r = diag\{m, m\}$ with $m = 90$ kg and damped by a viscous friction of $D_r = diag\{d, d\}$ with $d = 200$ Ns/m, emulating a heavy object. All virtual obstacles and walls are haptically rendered in order to provide a haptic feedback of

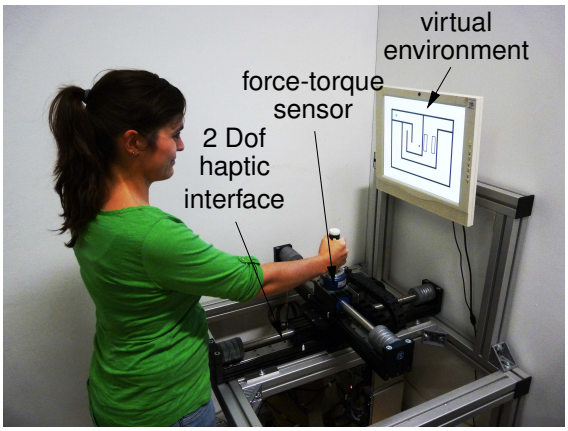


Fig. 7. Experimental setup.

the environment to the user. The HMM used to encode observations has 40 states and predictions are updated with a rate of 50Hz.

On top of the interface, a virtual maze is presented on a screen, see Fig. 8. On the right side the maze includes two obstacles moving horizontally in order to potentially provoke disagreements between the human and the robot. Note that the human behavior model does not consider the obstacle's state, which results in an increased predicted variance in the region perturbed by the obstacle's motion. This condition aims for evaluating the assistive performance when the model is not accurate enough and variability becomes more relevant. On the left side, a navigation task requires high movement precision caused by a narrow and curved path.

3) Task and Procedure: In total 19 persons (5 female) participated in the experiment. They were between 23 and 31 years old ($M = 26.84$ years). The experimental task was to move the virtual object represented by a red dot from the upper right corner (start position) to the upper left corner of the maze (green dot at target position) without hitting any object or wall, see Fig. 8. Participants were instructed to move at their comfortable speed and to finish the task even if they might hit the wall or an object.

Overall the experiment was divided into 7 parts corresponding to 7 different control laws of 3 trials each. In the first trial, the human lead together with a passive robot, i.e. $\boldsymbol{u_r} = 0$, and the resulting observations were encoded into the task model. This model was used during the second trial, where the robot actively assisted its human partner. This second task execution already captures negotiation forces between the human and the robot. Observations of the first

and second trials were used to train the task model for the third trial, providing already an estimation of the expected trajectory, both in state and force space, $\left\{\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}, \hat{\boldsymbol{u}}_{\boldsymbol{h}}\right\}$. The third trial was the only one considered for further evaluations. We tested 7 different control laws for each participant:

(a) No active assistance: $\boldsymbol{u_r} = 0$.

(b) LQR control using cost function (5).

(c) Risk-neutral control with Mahalanobis distance: $\theta = 0$ using cost function (12).

(d) Risk-averse control with Mahalanobis distance and expected disagreement: $\theta = -\alpha$, using cost function (12) and $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon_e}$.

(e) Risk-averse control with Mahalanobis distance considering both expected and current disagreement : $\theta = -\alpha$, using cost function (12) and $\boldsymbol{\varepsilon} = \max(\boldsymbol{\varepsilon_e}, \boldsymbol{\varepsilon_c})$.

(f) Risk-seeking control with Mahalanobis distance and expected disagreement: $\theta = \beta$, using cost function (12) and $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon_e}$.

(g) Risk-seeking control with Mahalanobis distance considering both expected and current disagreement: $\theta = \beta$, using cost function (12) and $\boldsymbol{\varepsilon} = \max(\boldsymbol{\varepsilon_e}, \boldsymbol{\varepsilon_c})$.

All implementations rely on the solutions presented in Section IV. In our experiments $\beta = 8.1 \cdot 10^4$ and $\alpha = 8.1 \cdot 10^2$, $R = I$ and $Q_k = Q_T = diag\{\omega_p, \omega_v\}$, being ω_p and ω_v the position and velocity weightings. For (b), we chose $\omega_p = 10^{10}$ and $\omega_v = 10^7$. For (c)-(g) we set $\omega_p = 10^5$ and $\omega_v = 10$, due to the low variance values, in the order of 10^{-5}. The receding horizon for the optimization was $T = 0.2$s and to estimate the current noise we used a window W corresponding to 0.05s.

The experimental procedure was as follows: participants were asked to face the haptic device and grasp the handle, as shown in Fig. 7. Next, the experimenter initialized the control algorithm and told the participant to start moving. After reaching the green target, participants were asked to free the handle which was moved back to the initial position automatically. Before every third trial participants were verbally informed that this was going to be the trial they had to rate. Subsequently they had to rate the perceived helpfulness when

- passing a *moving obstacle* (subtask MO).
- navigating through a *narrow channel* (subtask NC).

Ratings were done on a 6-point scale from 1 (counter productive) to 6 (helpful) and resulted

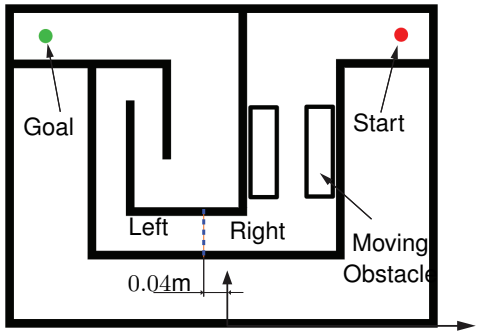


Fig. 8. Virtual environment (maze). With the handle of the 2-DoF haptic interface participants were able to move the virtual object (red dot) to the goal position (green dot).

Cont.	M_h	$M_{\boldsymbol{P_h}}$[W]	$M_{\|\boldsymbol{u_D}\|}$[N]	$M_{\|\boldsymbol{u_h}\|}$[N]	$M_{\|\boldsymbol{u_c}\|}$[s]
(a)	3.90	0.61	-	12.92	1.32
(b)	3.42	0.58	3.15	12.73	2.42
(c)	3.58	0.52	2.91	12.30	2.02
(d)	3.47	0.55	2.28	11.85	2.29
(e)	3.92	**0.39**	2.72	**10.46**	1.38
(f)	3.76	0.47	2.17	11.19	**1.03**
(g)	**4.11**	0.54	**1.29**	11.50	1.12

TABLE I

AVERAGE VALUES FOR PERCEIVED HELPFULNESS M_h, HUMAN POWER $M_{\boldsymbol{P_h}}$, DISAGREEMENT $M_{\|\boldsymbol{u_D}\|}$, HUMAN FORCE $M_{\|\boldsymbol{u_h}\|}$ AND COLLISION FORCES$M_{\|\boldsymbol{u_c}\|}$

in the explicit measure *perceived helpfulness*. Every participant performed 3 x 7 = 21 trials of which 7 trials were rated.

Regarding implicit measures we evaluated:

- the mean power exerted by the human $M_{\boldsymbol{P_h}}$, defined as $\frac{1}{t}\int_0^t \boldsymbol{u_h}^\mathsf{T}\dot{\boldsymbol{x}}d\tau$.
- the mean disagreement $M_{\|\boldsymbol{u_D}\|}$ between both agents, defined as $\frac{1}{t}\int_0^t\|\boldsymbol{u_D}\|d\tau$ where

$$\boldsymbol{u_D} = \begin{cases} \dfrac{-\boldsymbol{u_h}}{\|\boldsymbol{u_h}\|}\cdot\boldsymbol{u_r}, & \text{if } -\boldsymbol{u_h}\cdot\boldsymbol{u_r} > 0 \\ & \qquad \wedge\, \boldsymbol{u_h} \neq \boldsymbol{0} \\ 0, & \text{otherwise.} \end{cases}$$

- the mean human applied force $M_{\|\boldsymbol{u_h}\|}$, defined as $\frac{1}{t}\int_0^t\|\boldsymbol{u_h}\|d\tau$.

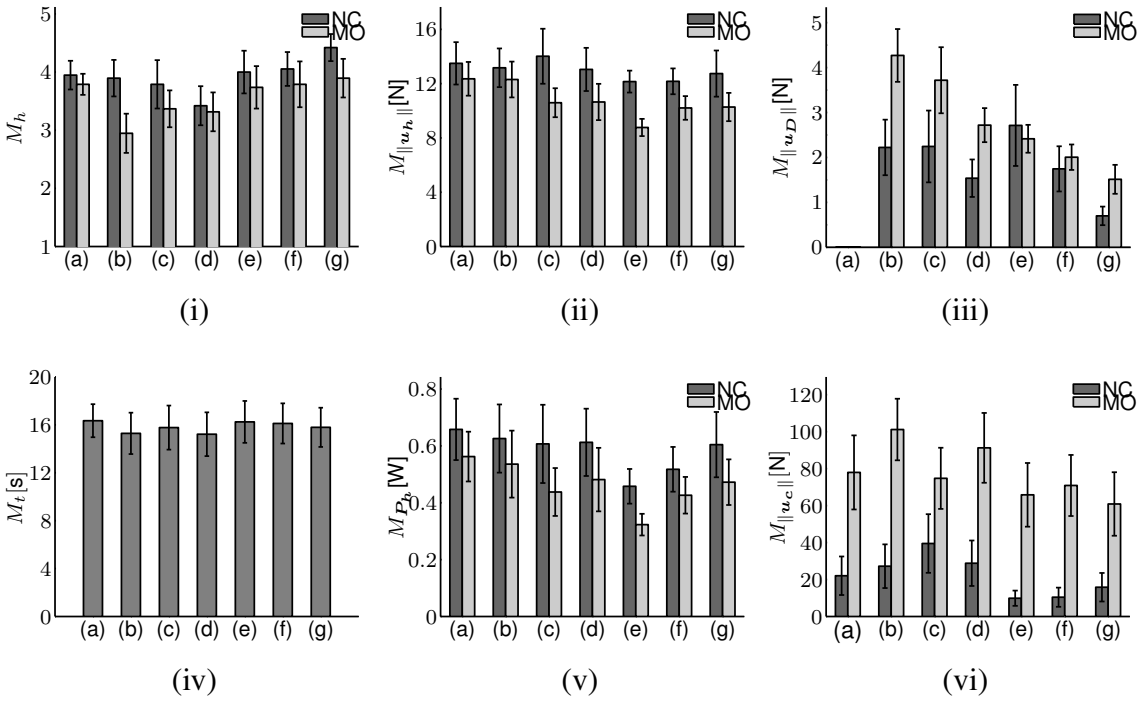


Fig. 9. Experimental results. Each of the figures shows the mean and the standard error of the different evaluated measures for the seven control laws (a)-(g). Fig. 9(i) shows the mean perceived helpfulness M_h, Fig. 9(ii) depicts the human exerted force $M_{\|\boldsymbol{u_h}\|}$, Fig. 9(iii) reflects the disagreement $M_{\|\boldsymbol{u_D}\|}$, Fig. 9(iv) shows the execution times M_t, Fig. 9(v) depicts the human power $M_{\boldsymbol{P_h}}$ and Fig. 9(vi) shows the collision forces $M_{\|\boldsymbol{u_c}\|}$.

- the mean contact forces during collisions with the virtual environment $M_{\|\boldsymbol{u_c}\|}$, defined as $\frac{1}{t}\int_0^t \|\boldsymbol{u_c}\| d\tau$, where $\boldsymbol{u_c}$ is the force applied on virtual obstacles.
- the mean execution time $M_{\boldsymbol{t}}$.

Means were taken over all participants for the respective controller. To gain knowledge on the human perception and response during different kind of tasks the presented maze was divided into two subtasks. For measures passing the *moving obstacle*, data was calculated until the participant was passing the turning point indicated by the dashed line shown in Fig. 8, placed at 0.04 m to the left of the origin. The rest of the task execution corresponded to subtask *narrow channel*. Data analysis was done Matlab and statistical analysis was carried out with SPSS (Statistical Package for the Social Sciences).

4) Results and Discussion: To access the rating of *perceived helpfulness*, see Fig. 9(i), a 2 x 7 repeated measures ANOVA was performed with the between-subject factors subtask (MO vs. NC)

Perceived Helpfulness M_h			Power $M_{\boldsymbol{P_h}}$[W]			Disagreement $M_{\|\boldsymbol{u_D}\|}$[N]		
Comparison	$F(1,18)$	p	Comparison	$F(1,18)$	p	Comparison	$F(1,18)$	p
(a)>(b)	7.15	$< .05$	(a)>(e)	11.35	$< .01$	-	-	-
(e)>(b)	5.90	$< .05$	(b)>(e)	4.54	$< .05$	(b)>(g)	32.47	$< .001$
(g)>(b)	7.08	$< .05$	(c)>(e)	2.58	n.s.	(c)>(g)	16.86	$< .01$
(a)>(d)	5.00	$< .05$	(d)>(e)	2.91	n.s.	(d)>(g)	11.37	$< .01$
(e)>(d)	5.50	$< .05$	(f)>(e)	2.23	n.s.	(e)>(g)	10.26	$< .01$
(g)>(d)	7.21	$< .05$	(g)>(e)	4.46	$< .05$	(f)>(g)	6.08	$< .05$

TABLE II

OVERVIEW OF RESULTS ON PLANNED COMPARISONS FOR PERCEIVED HELPFULNESS, HUMAN POWER AND DISAGREEMENT

and control method (a)-(g). Marginal differences were observed between subtasks, $F(1,18) = 3.95$, $p = .062$, which shows equally perceived helpfulness of the systems response on both sides. Numerically higher ratings for NC($M = 3.96$) compared to MO($M = 3.51$) reflect the fact that the task model is not aware of the moving obstacle and is not as accurate as for NC. Regarding control schemes there was a significant main effect, $F(6,108) = 2.46$, $p < .05$. Comparisons show that control laws (a), (e) and (g) resulted in significantly higher ratings compared to controllers (b) and (d), see Table I and Table II. This shows that control laws (e) and (g), both risk-sensitive and considering the current disagreement, were perceived as more helpful than method (b), the classical LQR control. It is also remarkable, that numerically all control methods considering the Mahalanobis distance (c)-(g) were rated higher than the classical LQR control (b). Furthermore, only control laws considering the current disagreement in the dynamics, (e) and (g), were numerically rated higher than the pure passive follower (a).

For all implicit measures 2x7 repeated measures ANOVAs were carried out with the between-subject factors subtask(NC vs. MO) and control method (a-g). If the sphericity criterion was not met, Greenhouse-Geisser correction was applied.

Regarding subtasks there was a higher human power applied on NC($M = 0.58$W) compared to MO($M = 0.46$W), $F(1,18) = 15.59$, $p < .01$, see Fig. 9(v). After correction, the differences between applied power for the control methods only marginally reached significance, $F(3.0, 54.8) = 2.44$, $p = .074$ Concerning control schemes, a tendency towards higher human power applied in control laws (a), (b) and (g) is identified compared to control law (e), see Table II. The risk-averse control law (e)

becomes stiffer under the presence of disagreement and, as it also considers the current disagreements with its partner, it is the most agressive partner. The observed differences suggest that this control method reduces the power applied by the human as the robot constantly takes a leading role.

Disagreement, see Fig. 9(iii), was marginally higher on MO($M = 2.77$ N) than on NC($M = 1.86$ N), $F(1, 18) = 4.22$, $p = .055$, in accordance with the results from perceived helpfulness. The behavior models lacks obstacle awareness and provides low prediction performance for MO thereby yielding disagreements. Looking at the results for control law (g), a highly significant main effect was observed, $F(5, 90) = 5.97$, $p < .001$. Planned comparisons show that all control methods cause a higher disagreement than (g), see Table II. As (g) is risk-seeking, it consequently tends to avoid disagreements adopting a more passive role. The fact that it also considers the current disagreement boosts this effect.

Applied forces, Fig. 9(ii), were higher on NC($M = 12.97$N) than on MO($M = 10.74$N), $F(1, 18) = 15.76$, $p <$ Although differences between control methods were not significant after correction, $p > .1$, the risk-averse with current disagreement, control law (e), numerically required the lowest applied force.

On NC, the contact force from collisions, see Fig. 9(vi), was smaller ($M = 0.65$ N) compared to MO($M = 2.67$ N), $F(1, 18) = 28.73$, $p < .001$.

No significant differences were found between movement times of the different control methods, see Fig. 9(iv), $F(6, 108) = 0.46$, $p > .8$.

In summary, the consideration of uncertainty in the optimization is shown beneficial for haptic assistance, as exhibited by the superior performance of all proposed control schemes w.r.t a naive LQR control ignoring any variability. The inclusion of the Mahalanobis distance in the optimization criterion leads to higher perceived helpfulness as it considers the learned trajectory's uncertainty. The benefits of a risk-sensitive optimization are only significant when an estimation of current disagreement is additionally considered. The risk-seeking control law achieves minimal disagreement as it adopts a recessive attitude during negotiation with its partner. This is especially noticeable passing through the obstacle: as the behavior model lacks any obstacle awareness, counteracting forces arise frequently due to poor prediction performance. In terms of human power, the risk-averse policy is more effective. However, the risk-seeking variant is preferred by human users, favoring disagreement avoidance to effort minimization. If the

behavior model is accurate risk-aversion effectively minimizes human effort but the complexity of human behavior makes fulfilling this prerequisite highly improbable. Nevertheless, other pHRI application domains may take advantage of risk-averse policies. For instance, consider a kinesthetic teaching scenario where the robot has the role of the teacher. In this case the robot aims for minimizing variations in human performance, a condition in accordance with a risk-averse control scheme.

VI. Conclusion

In this article we present a novel stochastic optimal control scheme for the synthesis of anticipatory haptic assistants. We explicitly consider human behavior uncertainty in terms of uncertain motion predictions and force variability due to disagreements. On the one hand, the Mahalanobis distance increases robot compliance under high motion prediction uncertainty. On the other hand, the risk-sensitive optimization governing the robot's behavior synthesizes a spectrum of attitudes towards disagreements, from recessive to dominant. A psychological study with human subjects highlights the advantages of the proposed approach: the consideration of motion prediction uncertainties and disagreement variability leads to higher performance in terms of perceived helpfulness and human effort minimization respectively.

The application of the proposed approach in non-linear systems is the matter of our future work. Further issues to explore are, as suggested by our results, the design of physical assistants with dynamic attitude in order to exploit the advantages of the different policies together with the addition of human-related constraints in the control design.

ACKNOWLEDGMENTS

This work is supported by the ERC Starting Grant "Control based on Human Models (con-humo)" under grant agreement no. 337654.

References

[1] J. Medina, D. Lee, and S. Hirche, "Risk Sensitive Optimal Feedback Control for Haptic Assistance," in *Proc. IEEE ICRA*, 2012.

[2] J. Medina, T. Lorenz, D. Lee, and S. Hirche, "Disagreement-Aware Physical Assistance Through Risk-Sensitive Optimal Feedback Control," in *Proc. IEEE/RSJ IROS*, 2012.

[3] T. Flash and N. Hogan, "The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model," *J. Neurosci.*, vol. 5, pp. 1688–1703, 1985.

[4] N. Jarrassé, J. Paik, V. Pasqui, and G. Morel, "How can human motion prediction increase transparency?" in *Proc. IEEE ICRA*, 2008, pp. 2134–2139.

[5] M. Lawitzky, A. Mörtl, and S. Hirche, "Load Sharing in Human-Robot Cooperative Manipulation," in *Proc. IEEE Ro-Man*, 2010, pp. 185–191.

[6] J. Grau-Moya, E. Hez, G. Pezzulo, and D. Braun, "The Effect of Model Uncertainty on Cooperation in Sensorimotor Interactions."

[7] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot Programming by Demonstration," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, 2008, pp. 1371–1394.

[8] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching Physical Collaborative Tasks: Object-Lifting Case Study with a Humanoid," in *Proc. IEEE Humanoids*, 2009, pp. 399–404.

[9] S. Calinon, P. Evrard, E. Gribovskaya, A. Billard, and A. Kheddar, "Learning collaborative manipulation tasks by demonstration using a haptic interface," in *ICAR*, 2009, pp. 1–6.

[10] E. Gribovskaya, A. Kheddar, and A. Billard, "Motion Learning and Adaptive Impedance for Robot Control during Physical Interaction with Humans," in *Proc. IEEE ICRA*, 2011, pp. 4326–4332.

[11] J. Medina, M. Lawitzky, A. Mörtl, D. Lee, and S. Hirche, "An Experience-Driven Robotic Assistant Acquiring Human Knowledge to Improve Haptic Cooperation," in *Proc. IEEE/RSJ IROS*, 2011, pp. 2416–2422.

[12] S. Ikemoto, H. Amor, T. Minato, B. Jung, and H. Ishiguro, "Physical human-robot interaction," *IEEE Robot. Autom. Mag.*, 2012.

[13] J. Rosell, C. Vázquez, A. Pérez, and P. Iñiguez, "Motion Planning for Haptic Guidance," *J. Intell. Robotic. Syst.*, vol. 53, no. 3, pp. 223–245, 2008.

[14] C. Vázquez, J. Rosell, L. Chirinos, and O. Domínguez, "Haptic Primitives Guidance Based on the Kautham Path Planner," in *Proc. IEEE/RSJ IROS*, 2010, pp. 4686–4691.

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

[16] B. Corteville, E. Aertbelien, H. Bruyninckx, J. D. Schutter, and H. V. Brussel, "Human-inspired Robot Assistant for Fast Point-to-point Movements," in *Proc. IEEE ICRA*, 2007, pp. 3639–3644.

[17] S. Miossec and A. Kheddar, "Human motion in cooperative tasks: Moving object case study," in *Proc. IEEE ROBIO*, 2008, pp. 1509–1514.

[18] A. Thobbi, Y. Gu, and W. Sheng, "Using Human Motion Estimation for Human-Robot Cooperative Manipulation," in *Proc. IEEE/RSJ IROS*, 2011, pp. 2873–2878.

[19] P. Evrard and A. Kheddar, "Homotopy Switching Model for Dyad Haptic Interaction in Physical Collaborative Tasks," in *Proc. EHS EuroHaptics*, 2009, pp. 45–50.

[20] E. Todorov, "Optimality principles in sensorimotor control." *Nature Neuroscience*, vol. 7, pp. 907–15, 2004.

[21] ——, "Optimal control theory," 2006.

[22] S. K. D. Mitrovic and S. Vijayakumar, "Learning impedance control of antagonistic systems based on stochastic optimization principles," *The International Journal of Robotics Research*, no. 5, pp. 556–573, 2010.

[23] A. J. Nagengast, D. A. Braun, and D. M. Wolpert, "Risk-sensitive optimal feedback control accounts for sensorimotor behavior under uncertainty," *PLoS Computational Biology*, vol. 6, no. 7, 2010.

[24] D. A. Braun, A. J. Nagengast, and D. M. Wolpert, "Risk-sensitivity in sensorimotor control," *Frontiers in Human Neuroscience*, vol. 5, no. 1, 2011.

[25] O. Oguz, A. Kucukyilmaz, T. Sezgin, and C. Basdogan, "Haptic Negotiation and Role Exchange for Collaboration in Virtual Environments," in *Proc. IEEE HAPTICS*, 2010, pp. 371–378.

[26] A. Kucukyilmaz, M. Sezgin, and C. Basdogan, "Conveying Intentions through Haptics in Human-Computer Collaboration," in *IEEE WHC*, June 2011, pp. 421–426.

[27] A. Mörtl, M. Lawitzky, A. Küçükyılmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The Role of Roles: Physical Cooperation between Humans and Robots," *Int. J. Robot. Res.*, 2012.

[28] N. Jarrassé, T. Charalambous, and E. Burdet, "A framework to describe, analyze and generate interactive motor behaviors," *PLoS ONE*, vol. 7, no. 11, p. e49945, 11 2012.

[29] M. Spong, S. Hutchinson, and M. Vidyasagar, *Robot modeling and control*. Wiley New York, 2006, vol. 3.

[30] K. Kronander and A. Billard, "Online learning of varying stiffness through physical human-robot interaction," in *Proc. IEEE ICRA*, 2012.

[31] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Trans. Syst., Man, Cybern. B*, vol. 37, no. 2, pp. 286 –298, 2007.

[32] C. E. Garcia, D. M. Prett, and M. Morari, "Model predictive control: theory and practice–a survey," *Automatica*, vol. 25, no. 3, pp. 335–348, 1989.

[33] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *Automatic Control, IEEE Transactions on*, vol. 18, no. 2, pp. 124 – 131, apr 1973.

[34] P. Whittle, "Risk-Sensitive Linear/Quadratic/Gaussian Control," *Advances in Applied Probability*, vol. 13, no. 4, pp. pp. 764–777, 1981.

[35] C. Won, "Cost distribution shaping: the relation between bode integral, entropy, risk-sensitivity, and cost cumulant control," in *Proc. ACC*, vol. 3, 2004, pp. 2160–2165 vol.3.

[36] S. Boyd, L. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, 1994.

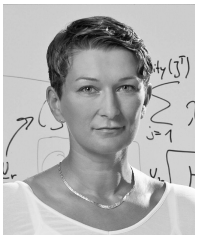
[37] A. Shaiju and I. Petersen, "Formulas for discrete time LQR, LQG LEQG and minimax LQG optimal control," in *Proc. IFAC*, 2008.

[38] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. Wiley-interscience New York, 1972, vol. 1.

[39] D. Lee and C. Ott, "Incremental kinesthetic teaching of motion primitives using the motion refinement tube," *Auton. Robot.*, pp. 1–17, 2011.

José Ramón Medina received his degree on Ingeniero Superior en Informática from the University of Seville, Spain, in 2009. Since November 2009 he is working towards the Ph.D. degree at the Chair of Information-Oriented Control, Department of Electrical Engineering and Information Technology, Technische Universität München (TUM), Germany. His research interests include machine learning and robot control with applications in physical human-robot interaction.

Tamara Lorenz Since 2011, Tamara Lorenz is a research assistant and PhD student at the Institute of Information-Oriented Control at Technische Universität München. Her research focus is on the dynamics and rules that underlie human-human interaction. She is also interested in how human interaction can be modeled to improve the interaction between humans and robots. As her work is very interdisciplinary, Tamara Lorenz was additionally affiliated with the Experimental Psychology Department at Ludwig-Maximilians Universität (LMU) from 2009 to 2013 and the Graduate School of Systemic Neurosciences, LMU, since 2009. Tamara Lorenz holds a diploma (MSc. equivalent) in Mechanical Engineering from Technische Universität München (TUM), Germany, specializing in human-machine interaction and medical engineering.



Sandra Hirche received the Diplom degree in Mechanical Engineering and Transport Systems from the Technical University Berlin, Germany, in 2002 and the Doctor of Engineering degree in Electrical Engineering and Information Technology from the Technische Universität Müchen, Munich, Germany, in 2005. From 2005 to 2007 she was awarded a Postdoc scholarship from the Japanese Society for the Promotion of Science at the Fujita Laboratory, Tokyo Institute of Technology, Tokyo, Japan. From 2008 to 2012 she has been an associate professor at Technische Universität Müchen. Since 2013 she is TUM Liesel Beckmann Distinguished Professor and has the Chair of Information-oriented Control in the Department of Electrical Engineering and Information Technology at Technische Universität München. Her main research interests include cooperative, distributed and networked control with applications in human-robot interaction, multi-robot systems, and general robotics. She has published more than 150 papers in international journals, books and refereed conferences. Dr. Hirche has served on the Editorial Boards of the IEEE Transactions on Control Systems Technology and the IEEE Transactions on Haptics. She has received multiple awards such as the Rohde & Schwarz Award for her PhD thesis in 2005, the IFAC World Congress Best Poster Award in 2005 and together with students Best Paper Awards of IEEE Worldhaptics and IFAC Conference of Manoeuvring and Control of Marine Craft in 2009.