Disagreement-Aware Physical Assistance Through Risk-Sensitive Optimal Feedback Control

José Ramón Medina, Tamara Lorenz, Dongheui Lee and Sandra Hirche

Abstract— **Proactive physical robotic assistance in the presence of human prediction uncertainty is a very challenging control problem. In this paper we propose a risk-sensitive optimal feedback controller for physical assistance that autonomously adapts the robot's behavior even during unknown situations. Using a probabilistic model to represent the cooperative task execution behavior and modeling the human as a source of process noise in the system, the proposed assistive controller proactively contributes to the task anticipating the human motion. Estimating online the current level of disagreement and prediction uncertainty, the assistive controller consequently calculates the optimal task contribution providing higher adaptability. A psychological evaluation compares different assistive control strategies in a virtual scenario using a two-Degree-of-Freedom haptic experimental setup. Results show that considering the current level of disagreement enhances the performance of the controller in terms of helpfulness and human effort minimization.**

I. INTRODUCTION

Physical interaction plays a key role for robotic assistants coexisting with humans. Many daily life tasks include physical coupling between partners pursuing a common goal. Rather than just passively follow the human partner, physical robotic assistants should also proactively contribute to such cooperative tasks anticipating human behavior. However, a proactive contribution can disturb a smooth interaction if the robot's predicted human behavior is different from the true human's intentions resulting in a disagreement. In our view the explicit consideration of both prediction uncertainties and disagreement with the human partner are key aspects for control design in order to enhance proactive physical assistants.

While reactive physical robotic assistants can successfully follow human partners [1], some situations require a proactive contributor in order to reach a desired configuration [2]. Regarding the force input of the partners we distinguish between the non-redundant and redundant degrees of freedom of the task. In this paper we focus on the physical assistance in the redundant task directions, where an input from either partner has the same effect on the motion, i.e. they share the physical effort to solve the task. In such directions it is desirable that the robotic partner anticipates the human motion during the task execution and adapts its contribution depending on the partner prediction's uncertainty. It is known that in physical assistance, human motion prediction reduces

the human's effort and increases the robot's transparency [3]. Proactive robotic assistants following this principle can be found in the literature for point to point movements [4], [5] and more complex tasks [6], [7]. However, in these works the robot's proactive action does not directly consider the uncertainty of the prediction model nor the disagreement with the human partner.

Recently, we proposed a novel concept for the control of physical robotic assistants [8] based on risk-sensitive optimal feedback control [9]. Interestingly, human sensorimotor behavior is found to follow risk-sensitivity principles [10]. Complex tasks such as object manipulation [11] and cooperative transportation [12] exhibit improved performance combining such principles with probabilistic learning techniques. In our recent work [8], in order to explicitly consider the variability of human behavior, we model the uncertain human force contribution as process noise, which directly influences the aggressiveness of control imposing a confidence-based role allocation to the robot. The process noise intuitively reflects the level of disagreement with the human partner. A probabilistic dynamic model represents the human-robot cooperative behavior and estimates both the expected task execution behavior and the level of process noise. However, this approach, being a model-based feedforward control, considers only previous task executions, which prevents the robot from reacting to any unexpected behavior not reflected in the cooperative task model. Furthermore, its suitability in real interaction with humans is yet to be proven as no psychological evaluation is presented.

In this paper we introduce a risk-sensitive optimal feedback controller for physical assistance, estimating the current level of disagreement in order to accommodate for *unmodeled* human execution variability. For the control design we follow a Model Predictive Control (MPC)-type strategy [13] providing continuous adaptation during the interaction. We further investigate both human preferences and assistive performance of different proactive controllers physically interacting with humans in a psychological experiment. A total of 19 participants interacted with a robot using seven different controllers. Results show that our proposed method outperforms both classical stochastic optimal control methods and our previous approach [8] in terms of perceived helpfulness and human effort minimization.

The remainder of this paper is structured as follows. Section II describes the architecture of the proposed controller. The assistive control is explained in Section III. A human behavior model is acquired as explained in Section IV. An experimental user study is presented in Section V.

All authors are with the Department of Electrical Engineering and Information Technology, Technical University of Munich, D-80290 Munich, Germany. {medina, lorenz, dhlee, hirche}@tum.de T. Lorenz is also with the Experimental Psychology Institute, Ludwig-Maximilians University, Munich, Germany.
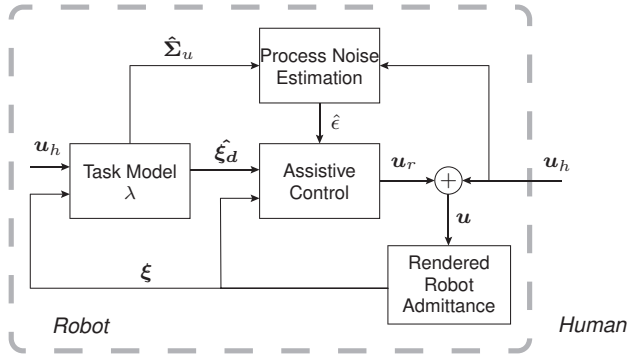
Fig. 1. General control scheme. The assistive controller generates an active task contribution added to the human force input for an admittance type controller. A task model provides state dependent predictions, that, together with the estimated process noise serve as input for the assistive control

II. GENERAL ARCHITECTURE

In this work we consider the class of physical human-robot interaction tasks where the robot is physically coupled to the human partner through an object. The goal is to bring the object from an initial configuration to a goal configuration. This prototypical task is representative for many different assistive tasks such as mobility assistance to humans, joint object transport/manipulation or physical rehabilitation. Depending on the application, the interaction can be through the object, as in cooperative manipulation tasks, or at the endeffector as in movement assistance for elderly or disabled or for exoskeletons. For simplicity, in our derivations we consider a common interaction contact point between the robot and its human partner at its end-effector. The general architecture of the robot control is depicted in Fig. 1. A compliant *reactive behavior* providing an intuitive reaction to the human force input $\boldsymbol{u_h}$ is achieved implementing an admittance control scheme. In order to generate an anticipatory *proactive behavior*, an assistive input $\boldsymbol{u_r}$ is added. The system dynamics are then given by

$$\boldsymbol{u} = \boldsymbol{u_h} + \boldsymbol{u_r} = \boldsymbol{M}_r \ddot{\boldsymbol{x}} + \boldsymbol{D}_r \dot{\boldsymbol{x}} , \qquad (1)$$

with rendered mass $\boldsymbol{M}_r$ and rendered viscous friction $\boldsymbol{D}_r$, $\boldsymbol{u_h}$ the applied force by the human, $\boldsymbol{u_r}$ the assistive control input of the robot, and $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{x} \\ \dot{\boldsymbol{x}} \end{pmatrix}$ the state of the system, where $\boldsymbol{x}$ is the position of the dyad.

The focus of this work lies on the design of the assistive control law that generates the assistive control input $\boldsymbol{u_r}$. As the real human desired trajectory $\boldsymbol{\xi_d}$ is not known to the robot, a probabilistic model of the task λ based on previous task executions provides an estimation $\hat{\boldsymbol{\xi}}_d$ of human motion as a trajectory of expected normal distributions $\hat{\boldsymbol{\xi}}_d = \mathcal{N}(\boldsymbol{\mu}_{\hat{\boldsymbol{\xi}}}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\xi}}})$. Deviations from this reference produced by the human partner are interpreted as disagreement, reflected as additive process noise ϵ in the system dynamics. Due to the coupling with the human and the potential error of the probabilistic approximation λ, a continuous reestimation of $\hat{\boldsymbol{\xi}}_d$ and the estimated level of process noise $\hat{\epsilon}$ is needed. Both variables define the problem setting for an optimal feedback controller that consequently recalculates the optimal robot contribution following an MPC strategy. However,

instead of optimizing a classical quadratic cost function, we apply a risk-sensitive optimization due to the presence of additive process noise, as explained in Section III. The proposed process noise estimation considers both the expected disagreement, based on previous observations, and the current disagreement, derived online from the current human input $\boldsymbol{u_h}$. This allows the robot to adapt its behavior even on previously unseen situations adding an additional feedback adaptation, as explained in Section III-A.

III. ASSISTIVE CONTROL

The goal of the proposed physical robotic assistant is to minimize the human force contribution while performing a cooperative haptic task by proactively following the anticipated human trajectory. For a receding horizon optimization problem, this implies minimizing a cost function in the form

$$J_k = \sum_{i=k}^{k+T} \|(\hat{\boldsymbol{\xi}}_{d_i} - \boldsymbol{\xi}_i)\|_Q^2 + \|\boldsymbol{u}_{r i}\|_R^2, \qquad (2)$$

where k is the sample time, T is the time horizon, $\|x\|_Q^2$ stands for the quadratic form $x^{\mathrm{T}} Q x$ and Q and R are weighting factors that allow a trade-off between control cost and tracking error minimization. This cost function has the following interesting interpretation. Assuming that nominally[1] the human behaves as a PD-controller

$$\boldsymbol{u_{h k}} = \boldsymbol{K_h}(\boldsymbol{\xi_{d k}} - \boldsymbol{\xi_k}),$$

where $\boldsymbol{K_h}$ is the tracking gain of the human and $\boldsymbol{\xi_d}$ is the human desired trajectory that is unknown to the robot, we can write the cost function (2) as

$$J_k = \sum_{i=k}^{k+T} \|\boldsymbol{u_{h i}}\|_{\hat{Q}}^2 + \|\boldsymbol{u_{r i}}\|_R^2,$$

where $\hat{Q} = \boldsymbol{K_h}^{\mathrm{T}} Q \boldsymbol{K_h}$. This expression formally describes the goal of an active robotic assistant with the intuitive concepts of human contribution minimization and the trade-off between two cooperative agents via the weighting factors $\hat{Q}$ and R.

While, ideally, no human force is required to fulfill the task as the robot anticipates the predicted human motion, prediction errors may produce disagreement with the human. Assuming no robot contribution to the task, the human applies the force $\boldsymbol{u} = \boldsymbol{u_d}$ in order to track the desired trajectory $\boldsymbol{\xi_d}$. An assistive controller that minimizes the cost function (2) generates an approximation $\boldsymbol{u_r} = \widetilde{\boldsymbol{u_d}}$ of $\boldsymbol{u_d}$ in order to minimize the human contribution. However, a difference between both control inputs is expected mainly due to the potential inherent error of the probabilistic estimation $\hat{\boldsymbol{\xi}}_d$. This discrepancy leads to a corrective human force given by

$$\boldsymbol{u_{h k}} = \boldsymbol{u_{d k}} - \widetilde{\boldsymbol{u_{d k}}}. \qquad (3)$$

As $\boldsymbol{\xi_d}$ is not known to the robot and $\boldsymbol{u_d}$ can not be computed in advance, we consequently model the human control input as *process noise* in the system dynamics from (1),

---

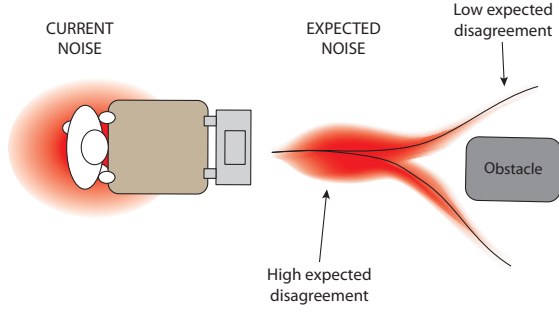[1]Deviations from the nominal model are captured in noise terms.

Fig. 2. Exemplary representation of both the expected process noise level $\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$ for a binary path following decision problem and the current process noise $\mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$.

i.e. $\boldsymbol{u_h} = \epsilon$. Here we assume it normally distributed with zero mean. It also represents the *level of disagreement* with the human partner.

Due to the discrete time nature of the implementation, we discretize the system from (1) with a sampling time interval Δt yielding a discretized plant dynamics in the form $\boldsymbol{\xi}_{k+1} = A\boldsymbol{\xi}_k + B\boldsymbol{u}_k$ given by

$$\begin{pmatrix} \boldsymbol{x}_{k+1} \\ \boldsymbol{v}_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 - \boldsymbol{M}_r^{-1}\boldsymbol{D}_r\Delta t \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_k \\ \boldsymbol{v}_k \end{pmatrix} + \begin{pmatrix} 0 \\ \boldsymbol{M}_r^{-1}\Delta t \end{pmatrix} (\boldsymbol{u_r}_k + \epsilon_k) . \tag{4}$$

The assistive control input is calculated solving the minimization problem with cost function (2) under the constraint of the plant dynamics (4).

A. Process noise estimation

Solving a stochastic optimal control problem requires a prior estimation of the process noise level present in the plant dynamics. In our previous work [8], the task model λ provided a prediction of ϵ, being normally distributed with zero mean and variance $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}$. While this *expected process noise* considers the previous task executions captured in the task model λ, it does represent any current unmodeled behavior from the human side. A human partner can always behave in an unexpected manner and such variability should be consequently reflected in the robot control.

From the assumption in (4), $\boldsymbol{u_h} = \epsilon$ is normally distributed with zero mean, i.e. the assistive controller expects no force from the human side and any other behavior deviating from this mean determines the covariance of the process noise. We therefore estimate the *current process noise* as normally distributed, with zero mean and covariance given by the sample covariance of the last W observations of $\boldsymbol{u_h}$

$$\boldsymbol{\Sigma}_u = \sum_{i=k-W}^{k} (\boldsymbol{u_h}_i - \boldsymbol{\mu_{u_h}})(\boldsymbol{u_h}_i - \boldsymbol{\mu_{u_h}})^{\mathrm{T}}, \tag{5}$$

where $\boldsymbol{\mu_{u_h}}$ is the expected mean of the human which is assumed to be 0. Due to the adopted MPC structure of the controller, the current process noise is constantly reestimated and works as an additional feedback that directly considers the human input. This provides reaction capability to unexpected human behavior as it captures current deviations from the human partner.
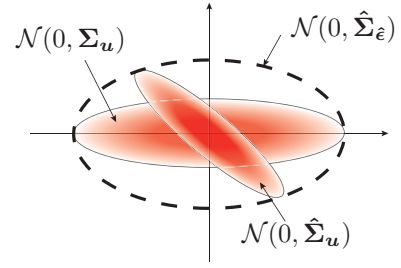

Fig. 3. Gaussian approximation to the $\max(\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}), \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}}))$ in a two-dimensional example. The dashed ellipse represents the Löwner-John hyperellipsoid.

The expected and the current process noise are a measure of the expected and current level of disagreement respectively, as shown in Fig. 2. In order to consider both noise sources in the optimization, we estimate the overall process noise level $\hat{\epsilon} = \mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\epsilon}})$ as the density function given by $\max(\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}), \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}}))$. However, as the maximum of two normal distributions is not normally distributed, we perform a Gaussian approximation to this problem $\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\hat{\epsilon}})$, being $\hat{\boldsymbol{\Sigma}}_{\hat{\epsilon}}$ the Löwner-John hyperellipsoid [14]. This approximation calculates the minimum volume hyperellipsoid around the set defined by $\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$ and $\mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ as shown in Fig. 3, allowing the controller to consider both process noise sources in the optimization.

B. Linear Quadratic Regulator Solution

The minimization problem with cost function (2) and the dynamics (4) can be represented in the classical Linear Quadratic Control (LQR) framework. For simplicity, we consider w.l.o.g. a constant starting sample time of $k = 1$. Note that the following derivations can be straightforwardly extended to the receding horizon case. As the reference trajectory is given by $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$, we can define $\boldsymbol{z}_k$ as the tracking error $\boldsymbol{z}_k = \boldsymbol{\xi}_k - \hat{\boldsymbol{\xi}}_{\boldsymbol{d}_k}$. However, in contrast to classical LQR

$$J(\boldsymbol{u_r}_{1\ldots T}) = \|\boldsymbol{z}_T^2\|_{Q_T} + \sum_{i=1}^{T-1} (\|\boldsymbol{z}_i\|_{Q_i}^2 + \|\boldsymbol{u_r}_i\|_{R_i}^2) , \tag{6}$$

the reference $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}} = \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}})$ is a sequence of multivariate normal distributions. In order to accordingly measure the distance to such an uncertain reference, we choose the (weighted) Mahalanobis distance as it considers the covariance of the prediction into the distance metric. Defining $\boldsymbol{z}_k = \boldsymbol{\xi}_k - \boldsymbol{\mu}_{\hat{\boldsymbol{\xi}}_k}$, the cost function becomes

$$J = \boldsymbol{z}_T^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{-\frac{1}{2}}Q\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{-\frac{1}{2}}\boldsymbol{z}_T + \sum_{i=1}^{T-1} (\boldsymbol{z}_i^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},i}^{-\frac{1}{2}}Q\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},i}^{-\frac{1}{2}}\boldsymbol{z}_i + \boldsymbol{u_r}_i^{\mathrm{T}}R\boldsymbol{u_r}_i) , \tag{7}$$

where $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},i}$ represents the covariance of the desired trajectory.

Due to the process noise ϵ in the dynamics, the solution to the control problem is calculated minimizing the *expected cost*, $\mathbb{E}[J]$. The optimal feedback control law for a LQR problem is given by

$$\boldsymbol{u_r}_i = -K_i\boldsymbol{z}_i , \tag{8}$$

where K_i is the feedback matrix of the Ricatti recursion

$$K_i = -R^{-1}B'(BR^{-1}B' + \Pi_{i+1}^{-1})^{-1}A \qquad (9)$$

and

$$\Pi_i = Q_i + A'(BR^{-1}B' + \Pi_{i+1}^{-1})^{-1}A , \qquad (10)$$

being $\Pi_T = Q_T$.

This assistive control strategy, shown in Fig. 4, tracks the predicted trajectory considering its covariance.

C. Risk-Sensitive Solution

In contrast to the LQR solution where the process noise ϵ has no influence on the feedback gain (8), a risk-sensitive controller directly considers the process noise in the dynamics, adapting the feedback matrix depending on a risk-sensitivity parameter θ. In this case the cost function takes the form

$$\gamma(\theta) = -2\theta^{-1}\ln \mathbb{E}[\exp^{-\frac{1}{2}\theta J}] . \qquad (11)$$

If $\theta = 0$ the controller is risk-neutral and corresponds to the LQR case explained in Sec. III-B. For $\theta < 0$ and $\theta > 0$ the controller becomes risk-averse and risk-seeking, respectively.

Solving this optimization problem leads to a modified form of the Ricatti recursion

$$K_i = -R^{-1}B'(BR^{-1}B' + \theta\hat{\epsilon} + \Pi_{i+1}^{-1})^{-1}A , \qquad (12)$$

and

$$\Pi_i = Q_i + A'(BR^{-1}B' + \theta\hat{\epsilon} + \Pi_{i+1}^{-1})^{-1}A , \qquad (13)$$

with $\Pi_T = Q_T$.

This solution is the same as those from (9) and (10), except for the term $\theta\hat{\epsilon}$. Consequently, in the risk-averse case, $\theta < 0$, the feedback gain becomes higher interpreting the noise in a pessimist manner as it if was directing the state in the wrong direction. For the risk-seeking case, $\theta > 0$, the feedback gain becomes lower adopting an optimist attitude as it assumes that the noise is already doing part of the job and therefore directing the state in the right direction.

IV. HUMAN BEHAVIOR LEARNING AND PREDICTION

A model representing a 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 can be acquired by initially letting the human lead, i.e., $\boldsymbol{u_r} = 0$. With this first rough representation, the robot actively assists and additionally observes interaction patterns during further task trials. Modeling state and control input trajectories, $\{\boldsymbol{\xi}, \boldsymbol{u}\}$, the robot acquires a task model that represents both the human desired state trajectory and the expected disagreement levels.

While many probabilistic approaches, such as HMMs and Gaussian Mixture Models as dynamical systems or Gaussian Process based models, can achieve a rich representation of the encoded behavior, using sparse demonstration sets, the predicted dynamics can be very far away from the real ones. To ensure a safe interaction, we use a time based HMM, applying regression on the time domain as explained in [15].
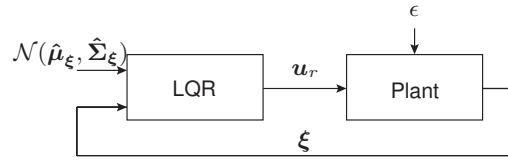

Fig. 4. Assistive control scheme for the standard stochastic optimal control case. The controller generates the robot's control input based on the control law from (8) given the uncertain trajectory $\hat{\boldsymbol{\xi}} = \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}})$.
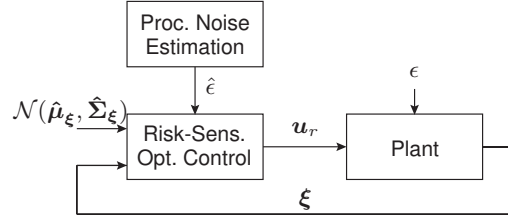

Fig. 5. Assistive control scheme for the risk-sensitive optimal control case. In addition to the scheme from Fig. 4, unexpected behavior of the human is estimated as process noise $\hat{\epsilon}$ and directly considered in the optimal feedback matrix calculation.

This provides a generalized trajectory of the task in terms of means and heteroscedatic variances. While the predicted dynamics are restricted to the generalized trajectory, predictions will always be in previously experienced configurations of the state space and therefore potentially safe in static scenarios. Using the Viterbi Algorithm in a window over the last observations and estimating the state in time domain, the next state $\hat{\boldsymbol{\xi}}_d$ is predicted. See [7] for a detailed explanation of the applied method.

V. EXPERIMENTS

To evaluate the proposed controllers and the preferences of humans interacting with it, we designed an experiment in which a human actor has to transport a virtual object from an initial position towards a defined goal position.

A. Experimental Setup

The human actor applies forces to a haptic interface in order to move the virtual object, as shown in Fig. 6. It consists of a two degrees-of-freedom linear-actuated device (ThrustTube) which has a free-spinning handle at the grasping point, i.e. $\boldsymbol{x} \in R^2$. Attached to the handle, a force/torque sensor (JR3) measures the human force input. On top of the interface, a virtual maze is presented, see Fig. 7. On the right side the maze includes two obstacles moving horizontally in order to potentially provoke disagreements between the human and the robot. On the left side, a navigation task requires high movement precision caused by a narrow folded path.

The control scheme running at 1 kHz and implemented in *MATLAB/Simulink* is executed on a personal computer with *Linux PREEMT Real-Time kernel* using *Matlab's Real-Time Workshop*. The shared object is physically rendered with a mass of $\boldsymbol{M_r} = diag\{m, m\}$ with $m = 90$ kg and damped by a viscous friction of $\boldsymbol{D_r} = diag\{d, d\}$ with $d = 200$ Ns/m, emulating a heavy object. All virtual obstacles and walls are haptically rendered in order to provide

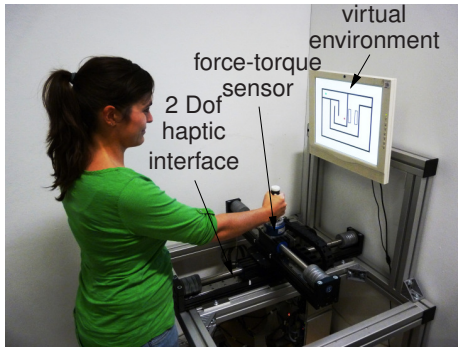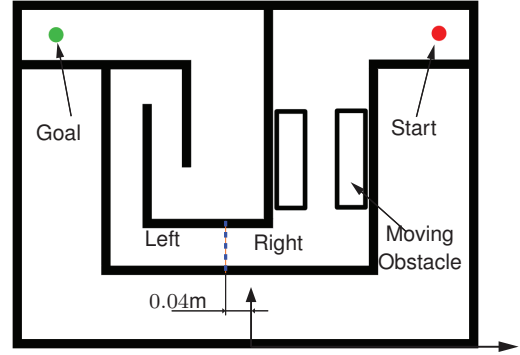Fig. 6.   Experimental setup.

Fig. 7.   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).

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

B. Task and Procedure

The experimental task was to carry the 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. 7. 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 controllers of 3 trials each. In the first trial, the human lead, $u_r = 0$, and the resulting observations were encoded into the task model used during the second trial where the robot actively assisted its partner. This second task execution behavior was also learned together with the first one in order to acquire a model of the interaction. The third trial was the only one considered for further evaluations and it used the resulting learned model from the two previous executions. We tested 7 different controllers for each participant:

(a) No active assistance: $u_r = 0$.
(b) LQR using the cost function from (6).
(c) Risk-neutral with Mahalanobis distance: $\theta = 0$ and using the cost function from (11) with J as in (7).
(d) Risk-averse with Mahalanobis distance: $\theta = -\alpha$, using the cost function from (11) with J as in (7) and $\hat{\epsilon} = \mathcal{N}(0, \hat{\Sigma}_u)$.
(e) Risk-averse with Mahalanobis distance considering current process noise : $\theta = -\alpha$, using the cost function from (11) with J as in (7) and $\hat{\epsilon} = \mathcal{N}(0, \hat{\Sigma}_{\hat{\epsilon}})$.
(f) Risk-seeking with Mahalanobis distance: $\theta = \beta$, using the cost function from (11) with J as in (7) and $\hat{\epsilon} = \mathcal{N}(0, \hat{\Sigma}_u)$.
(g) Risk-seeking with Mahalanobis distance considering current process noise: $\theta = \beta$, using the cost function from (11) with J as in (7) and $\hat{\epsilon} = \mathcal{N}(0, \hat{\Sigma}_{\hat{\epsilon}})$.

In our experiments $\beta = 10^{-5}$ and $\alpha = 10^{-7}$, $R = I$ 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

Cont.	M_h	M_{P_h}[W]	$M_{\|u_D\|}$[N]	$M_{\|u_h\|}$[N]	$M_{\|u_c\|}$[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_{P_h}), DISAGREEMENT ($M_{\|u_D\|}$), HUMAN FORCE ($M_{\|u_h\|}$) AND COLLISION FORCES($M_{\|u_c\|}$)

was $T = 0.2$s and to estimate the current noise we used a window of $W = 0.05$s.

The experimental procedure was as follows: participants were asked to face the haptic device and grasp the handle, as shown in Fig. 6. 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 help through the system

- on the *right side* for passing the moving obstacle.
- on the *left side* for navigating through the narrow channel.

Ratings were done on a 6-point scale from 1 (counter productive) to 6 (helpful) and resulted 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_{P_h}.
- the mean disagreement $M_{\|u_D\|}$ between both agents. We calculate the disagreement between partners as:

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

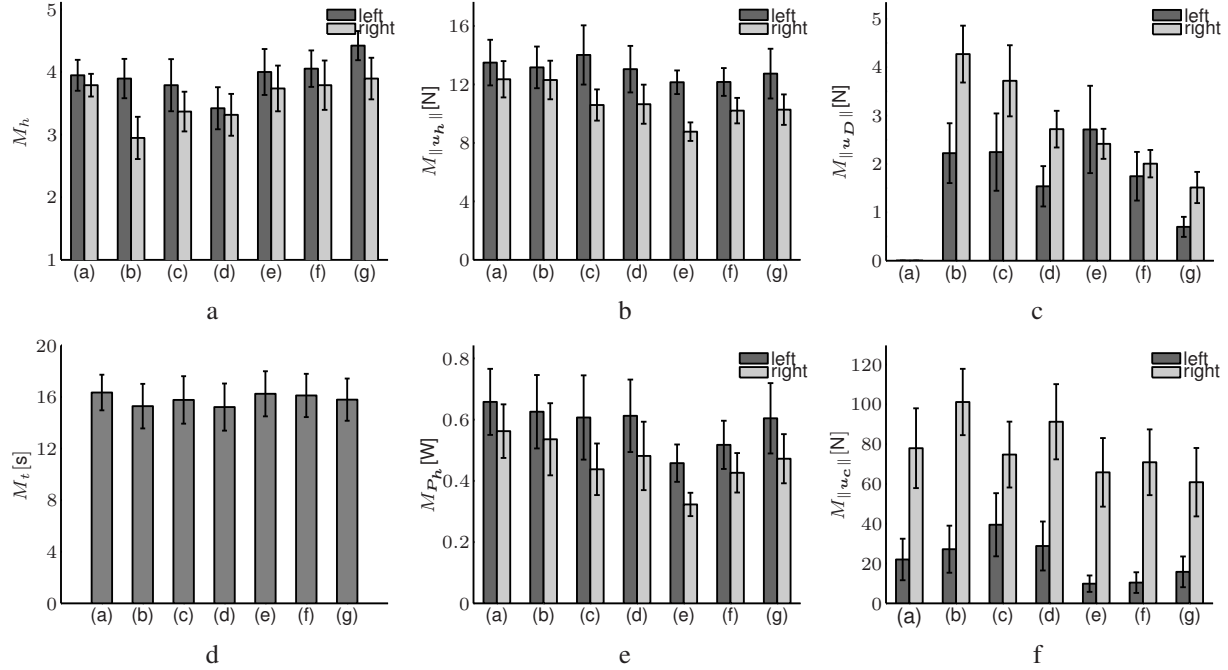- the mean human applied force $M_{\|u_h\|}$.

Fig. 8. Experimental results. Each of the figures shows the mean and the standard error of the different evaluated measures for the seven controllers (a)-(g). Fig. 8a shows the mean perceived helpfulness (M_h), Fig. 8b depicts the human exerted force ($M_{\|u_h\|}$), Fig. 8c reflects the disagreement ($M_{\|u_D\|}$), Fig. 8d shows the execution times (M_t), Fig. 8e depicts the human power(M_{P_h}) and Fig. 8f shows the collision forces ($M_{\|u_c\|}$).

Perceived Helpfulness(M_h)			Power(M_{P_h}[W])			Disagreement($M_{\|u_D\|}$[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

- the mean contact forces during collisions with the virtual environment $M_{\|u_c\|}$.
- the mean execution time M_t.

Means were taken over all participants for the repective controller. To gain knowledge on the human perception and response during different kind of tasks the presented maze was divided into two parts. For measures on the *right side*, data was calculated until the participant was passing the turning point indicated by the dashed line shown in Fig. 7, placed at 0.04 m to the left of the origin. Data Analysis was done in MS Excel (explicit measure) and Matlab (implicit measures). Statistical Analysis was carried out with SPSS.

C. Results and Discussion

In total 19 persons (5 female) participated in the experiment. They were between 23 and 31 years old ($M = 26.84$ years).

To access the rating of *perceived helpfulness*, see Fig. 8a, a 2 x 7 repeated measures ANOVA was performed with the between-subject factors maze side (left vs. right) and control method (a-g). Marginal differences were observed between maze sides, $F(1,18) = 3.95$, $p = .062$, which shows equally perceived helpfulness of the systems response on both sides. Numerically higher ratings for the left side ($M = 3.96$) compared to the right side ($M = 3.51$) might indicate a bigger need for support during navigation while during a task which requires fast estimations for clearing a moving obstacle one rather relies on one's own capabilities. Regarding the controllers there was a significant main effect, $F(6,108) = 2.46$, $p < .05$. Planned comparisons show that controllers (a), (e) and (g) resulted in significantly higher ratings compared to controllers (b) and (d), see Table I and Table II. Taken together this shows that controllers (e) and (g), both risk-sensitive and considering the current process noise, were perceived as more helpful than method (b), the classical LQR. It is remarkable, that numerically all controllers considering the Mahalanobis distance (c)-(g) were rated higher than the classical LQR (b). Furthermore, only controllers considering the current process noise 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 maze side

(left vs. right) and control method (a-g). If the sphericity criterion was not met, Greenhouse-Geisser correction was applied.

Regarding maze side there was a higher human power applied on the left side ($M = 0.58$W) compared to the right side ($M = 0.46$W), $F(1, 18) = 15.59$, $p < .01$, see Fig. 8e. After correction, the differences between applied power for the control methods only marginally reached significance, $F(3.0, 54.8) = 2.44$, $p = .074$. In this context, contrasts might depict a tendency towards higher power applied in controllers (a), (b) and (g) compared to controller (e), see Table II. The risk-averse controller (e) becomes stiffer under the presence of process noise and, as it also considers the current disagreements with its partner, it is the most agressive controller. The observed tendential differences suggest that this control method reduces the power applied by the human as the robot constantly takes a leading role.

Disagreement, Fig. 8c, was marginally higher on the right side ($M = 2.77$ N) than on the left side ($M = 1.86$ N) , $F(1, 18) = 4.22$, $p = .055$, supporting the statement from perceived helpfulness that participants rather accepted support through the system during navigation while they preferred to rely on their own estimations when a moving obstacle has to be cleared. Looking at the results for controller (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 reduce disagreement under the presence of process noise adopting a more passive role. The fact that it also considers the current process noise boosts this effect.

Applied forces, Fig. 8b, were higher on the left side ($M = 12.97$N) than on the right side of the maze ($M = 10.74$N), $F(1, 18) = 15.76$, $p < .01$. Although differences between control methods were not significant after correction, $p > .1$, the risk-averse with current noise estimation controller (e) numerically required the lowest applied force.

On the left side, the contact force from collisions, see Fig. 8f, was smaller ($M = 0.65$ N) compared to the right side ($M = 2.67$ N), $F(1, 18) = 28.73$, $p < .001$.

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

In summary, depending on the situation and measure, the proposed assistive controller performed better both in the risk-seeking and the risk-averse case than our previous approach and the classical LQR. In terms of disagreement, the risk-seeking controller considering the current process noise performs better as it adopts a more passive behavior under disagreement with its partner. In terms of human power, the opposite risk-averse policy achieves a better assistance as it reduces the human effort. These results show, on one side, that considering the current process noise for the assistive controller greatly enhances its performance. On the other side, they suggest that, depending on the situation, an active robotic assistant should adapt its risk-sensitivity to a risk-seeking or a risk-averse policy.

VI. Conclusions

In this paper we presented a risk-sensitive optimal feedback controller for proactive physical assistance that adapts the robot's task contribution depending on both the current and the expected disagreement with the human partner. When the human behavior deviates from the predictions given by a probabilistic task model, the proposed controller consequently adapts the controller gains in an online fashion. Our psychological experiment shows that the proposed approach provides better performance in terms of flexibility for the risk-seeking case and in terms of human effort minimization for the risk-averse one. The combination of such strategies depending on environmental situations together with a larger experiment in more complex tasks are the matter of our future work.

Acknowledgments

References

[1] Y. Hirata, Y. Kume, Z.-D. Wang, and K. Kosuge, "Decentralized Control of Multiple Mobile Manipulators based on Virtual 3-D Caster Motion for Handling an Object in Cooperation with a Human," in *Proc. IEEE ICRA*, vol. 1, 2003, pp. 938 – 943 vol.1.

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

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

[4] 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.

[5] 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.

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

[7] 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.

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

[9] 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.

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

[11] G. Ganesh, A. Albu-Schaffer, M. Haruno, M. Kawato, and E. Burdet, "Biomimetic motor behavior for simultaneous adaptation of force, impedance and trajectory in interaction tasks," in *Proc. IEEE ICRA*, 2010, pp. 2705 –2711.

[12] A. K. E. Gribovskaya and A. Billard, "Motion learning and adaptive impedance for robot control during physical interaction with humans," in *Proc. IEEE ICRA*, 2011, pp. 4326–4332.

[13] 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.

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

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