

Risk-Sensitive Optimal Feedback Control for Haptic Assistance

José Ramón Medina, Dongheui Lee and Sandra Hirche

Abstract— While human behavior prediction can increase the capability of a robotic partner to generate anticipatory behavior during physical human robot interaction (pHRI), predictions in uncertain situations can lead to large disturbances for the human if they do not match the human intentions. In this paper we present a novel control concept in which the assistive control parameters are adapted to the uncertainty in the sense that a the robot takes a more or less active role depending on its confidence in the human behavior prediction. The approach is based on risk-sensitive optimal feedback control. The human behavior is modeled using probabilistic learning methods and any unexpected disturbance is considered as a source of noise. The proposed approach is validated in situations with different uncertainties, process noise and risk-sensitivities in a tow-Degree-of-Freedom virtual reality experiment.

I. INTRODUCTION

Daily-life tasks involving physical coupling between collaborating partners are very common for example in workshops, households, rehabilitation, and assistance to elderly. One of the major challenges is to replace one human partner in such joint physical tasks by an intuitively acting robotic assistant. It is well-known that accurate human behavior prediction is required for intuitive interaction as it enables anticipatory movements of the robot. However, a prediction far away from the real human intention can be more of a hindrance than a help. Estimating in advance how certain the predictions are in order to adapt the physical contribution of the robot can therefore improve its assistance. The idea of this work is to explicitly consider the uncertainty of a learned human behavior model within the interaction control design.

Physical human-robot interaction (pHRI) schemes are investigated for example in [1], [2], where multiple robots assist one or multiple humans carrying a bulky object by passively reacting to their human partners’s input. Such a follower strategy might not be sufficient though in constrained scenarios, where an active task contribution from all participants is necessary [3]. Also, it is well-known that robot assistance based on the anticipation of the human partner’s action can reduce the human effort [4], [5], and increase its transparency [6]. For the prediction of the human action in a specified task a model is necessary. Our previous work presented an incremental learning and prediction framework for multiple tasks and high dimensional data based on Hidden Markov Models (HMMs), where the prediction serves as target configuration to generate an anticipatory movement of the assisting robot using a position-based impedance controller [7]. However, due to the human variability large

deviations of the real human behavior from the predicted one can occur. At the moment, prediction uncertainties are not explicitly taken into account in the robot control. This may easily lead to a robot behavior insisting on a motion along a ”wrongly” predicted path resulting in non-intuitive behavior and even disturbance to the joint task execution. It would be much more desirable, if the robot could adapt its role to the uncertainty of the current situation. Here the role is defined by the parameter of the impedance control which determine how actively the robot tries to assist based on the predicted human motion.

In order to solve this open issue of uncertainty-adaptive robot behavior we propose a novel control approach for pHRI based on stochastic optimal control (SOC) principles, in particular risk-sensitive control. Our approach is inspired by recent findings in human sensorimotor control [8], [9], which state that human sensorimotor behavior is the result of minimizing a cost function while considering process noise and uncertainty in the dynamics. Very promising in this context are works on the combination of learning techniques and SOC in human sensorimotor control, which achieve better accuracy than classical SOC models under noisy conditions [10], [11]. These ideas are also applied to manipulation [12] and pHRI scenarios [13]. Still, the variability does not explicitly influence the control parameters. Risk-sensitive control [14], [15] is an extension to classical SOC methods which addresses this issue as it also considers the variability of the cost, i.e. cost variance. Recent works successfully employ this concept to model human sensorimotor behavior [16], [17]. To the best of our knowledge, no other previous works have investigated the potential of risk-sensitive control in pHRI.

In this work, we investigate the combination of stochastic optimal control principles including risk-sensitivity with learning techniques for application in physical human-robot interaction. We consider a probabilistic dynamic model in terms of a HMM for the representation of the human motion behavior to solve a task. We further interpret any other human behavior deviating from this model as process noise in the system. Therefore, we first extend the classical SOC approach by the Mahalanobis distance as distance metric enabling a robotic partner which dynamically adapts its impedance, i.e. its role, depending on the level of uncertainty and process noise in the dynamics. Furthermore, we investigate a risk-sensitive optimal control approach for haptic assistance in order to achieve different adaptation behavior. The control approach is validated in a virtual environment experiment using a two-Degree-of-Freedom (DoF) setup.

The remainder of this paper is structured as follows: in

Authors are with the Institute of Automatic Control Engineering, Technische Universität München, D-80290 Munich, Germany. {medina, dhlee, hirche}@tum.de

Section II the problem considered and the general architecture of the proposed control scheme is introduced. The risk-sensitive optimal feedback assistive controller is explained in Section III. The dynamic model acquisition is presented in Section IV. Finally, an evaluation of the proposed controller is carried out in Section V.

II. GENERAL ARCHITECTURE

In this paper we consider the prototypical task of a human and a robot moving together from a starting pose to a final common goal while being physically coupled. This representation is suitable for many types of assistive tasks such as joint object transport/manipulation, mobility assistance to human, and physical rehabilitation. In order to accommodate unexpected inputs from the human, a robot should exhibit a compliant behavior towards the force input by a human partner. This *reactive behavior* can be achieved by an admittance control law. However, a purely reactive robotic partner becomes a heavier load rather than a help due to its reaction time and electromechanical limitations. A better alternative is a robot that does not only react, but also generate an *assistive behavior* based on human motion prediction resulting in an anticipatory virtual force input added to the human's input for the impedance control. Depending on the application, the interaction between both partners can be through an object, as in cooperative manipulation tasks, or directly at the end-effector as in movement assistance for elderly or disabled or exoskeletons. For simplicity of illustration, in our derivations we consider a common interaction contact point between the robot and its human partner at its end-effector. The general assistive control architecture is presented in Fig. 1. The reactive behavior of the robot is generated through a position-based admittance control. Therefore we can assume that the robot behaves like a linear time-invariant mass-damper system reacting with a motion to the input force \mathbf{u} consisting of the human force input \mathbf{u}_h and the assistive robot force input \mathbf{u}_r

$$\mathbf{u} = \mathbf{u}_h + \mathbf{u}_r = \mathbf{M}_r \ddot{\mathbf{x}} + \mathbf{D}_r \dot{\mathbf{x}}, \quad (1)$$

with a rendered inertia \mathbf{M}_r , a rendered viscous friction \mathbf{D}_r , and $\boldsymbol{\xi} = (\mathbf{x} \ \dot{\mathbf{x}})^T$ the state of the system, where \mathbf{x} is the position of the contact point, see also Fig. 1. The assistive robot force is computed within the assistive control block based on the deviation of the predicted motion $\hat{\boldsymbol{\xi}}$ and the current robot state $\boldsymbol{\xi}$. The motion prediction is provided by the task model λ which is assumed to be learned from previous executions with the human, i.e. it includes the human motion behavior model. Here the main focus is on the design of the assistive control law.

The goal of the assistive control is to minimize the human contribution to the task $\|\mathbf{u}_h\|$. We assume the human intends to follow a desired trajectory $\boldsymbol{\xi}_d$ which requires \mathbf{u}_d as a force input for the given dynamics. In that case, minimizing the human contribution means to minimize

$$\|\mathbf{u}_h\| = \|\mathbf{u}_d - \mathbf{u}_r\| ,$$

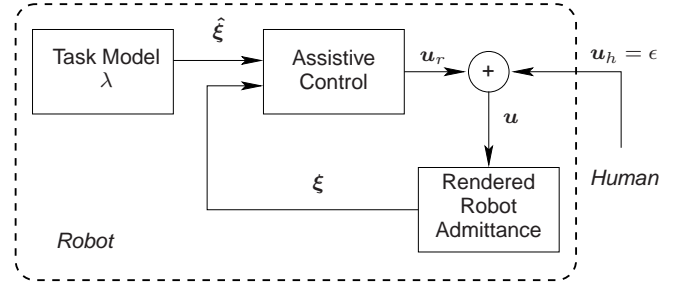


Fig. 1. Overall control scheme: A task model provides predictions for the assistive controller to generate an anticipatory control input that is added to the human force acting on the admittance-type reactive behavior of the robot.

which obviously will be minimal if the robot would exactly apply the desired force \mathbf{u}_d as assistive force. Note that $\mathbf{u}_d(t)$ can be computed perfectly if the desired trajectory $\boldsymbol{\xi}_d(t)$ as well as the current state $\boldsymbol{\xi}(t)$ are exactly known. However, the exact desired trajectory of the human is typically not known. In this paper we assume that λ is a probabilistic model of the task providing an estimation in terms of means and covariances $\hat{\boldsymbol{\xi}} = \{\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\}$ of $\boldsymbol{\xi}_d$, see also Section IV. If the robot tracks this predicted trajectory, it will generate an approximation $\tilde{\mathbf{u}}_d$ of \mathbf{u}_d . The required human contribution to follow the desired trajectory $\boldsymbol{\xi}_d$ is then

$$\mathbf{u}_h = \mathbf{u}_d - \tilde{\mathbf{u}}_d .$$

As the true value \mathbf{u}_d can not be derived and instead an approximate value based on the uncertain prediction is employed, the human contribution \mathbf{u}_h can be modeled as an independent noise input to the system, i.e. $\mathbf{u}_h = \epsilon$. As a result we can rewrite the force input to the rendered admittance (1) as $\mathbf{u} = \mathbf{u}_r + \epsilon$.

The optimal assistive controller will be developed in discrete time. Accordingly, the system (1) is discretized with a sampling time interval Δt . The discrete state of the system at time k is given by $\boldsymbol{\xi}_k = (\mathbf{x}_k \ \mathbf{v}_k)^T$, where \mathbf{x}_k , \mathbf{v}_k are the position and the velocity of the dyad, \mathbf{u}_{rk} the assistive control force of the robot, and \mathbf{u}_k the force acting on the reactive behavior of the robot at time k . The discrete time dynamics of the system, written in the form

$$\boldsymbol{\xi}_{k+1} = \mathbf{A}\boldsymbol{\xi}_k + \mathbf{B}\mathbf{u}_k , \quad (2)$$

is then given by

$$\begin{pmatrix} \mathbf{x}_{k+1} \\ \mathbf{v}_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 - \mathbf{M}_r^{-1} \mathbf{D}_r \Delta t \end{pmatrix} \begin{pmatrix} \mathbf{x}_k \\ \mathbf{v}_k \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{M}_r^{-1} \Delta t \end{pmatrix} (\mathbf{u}_{rk} + \epsilon_k) , \quad (3)$$

with the initial condition $\boldsymbol{\xi}_0 = (\mathbf{x}_0 \ \mathbf{v}_0)^T$. Considering this dynamics and the task model λ , the proposed assistive controller produces the necessary control input to make the dyad fulfill the task, as explained in the following section.

III. ASSISTIVE CONTROL

The goal of assistance is to track the predicted state $\hat{\boldsymbol{\xi}}_{1..T}$ until a given time horizon T . The assistive controller must consequently generate the corresponding robot's control input sequence, $\mathbf{u}_{r1..T}$. However, unexpected disturbances

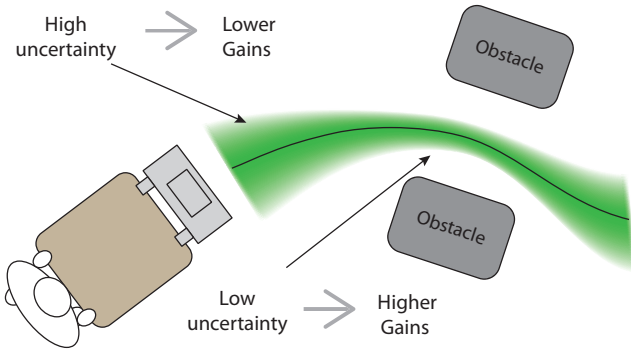


Fig. 2. Exemplary representation of the predicted state mean and variance $\hat{\xi} = \{\hat{\mu}_\xi, \hat{\Sigma}_\xi\}$ and the effect of the Mahalanobis distance. Smaller variances can represent constrained regions meaning higher gains. Similarly, higher variances imply unconstrained regions where lower gains are applied.

from the human side might still occur during the movement and the assistance should still react in order to keep on reaching the desired state. In such a scenario, optimal feedback control arises as a suitable method, as it can plan the optimal control input sequence given a cost function in the form of a feedback control law that keeps tracking the desired trajectory. For a finite horizon task, the cost function of an optimal control problem is

$$J(\mathbf{u}_{r1...T}) = h(\xi_T) + \sum_{k=1}^{T-1} c(\xi_k, \mathbf{u}_{rk}), \quad (4)$$

where $c(\xi_k, \mathbf{u}_{rk})$ is the per-stage cost at time k for ξ_k and \mathbf{u}_{rk} , and $h(\xi_T)$ is the final cost. In order to make the robot track the predicted configuration trajectory $\hat{\xi}_{1...T}$, the cost function could be designed penalizing the (weighted) Euclidean distance to the predicted state trajectory while minimizing control cost. If we assume complete state observation and no sensor delays, this problem can be formulated under the Linear Quadratic Regulator (LQR) framework.

A. Linear Quadratic Regulator

A finite time horizon LQR problem is given by a system with linear dynamics in the form of (2) and a cost function given by

$$J(\mathbf{u}_{r1...T}) = \mathbf{z}_T^T Q_T \mathbf{z}_T + \sum_{k=1}^{T-1} (\mathbf{z}_k^T Q_k \mathbf{z}_k + \mathbf{u}_{rk}^T R_k \mathbf{u}_{rk}), \quad (5)$$

where \mathbf{z}_k is the state and Q and R are positive semi-definite and positive definite weighting matrices, respectively, which allow a trade-off between required energy and aggressiveness of control. This standard cost function implies a zero equilibrium state. Our case is a tracking problem with the reference given by the mean $\hat{\mu}_{\xi_{1...T}}$. A straightforward option would be to define \mathbf{z}_k as the tracking error $\mathbf{z}_k = \xi_k - \hat{\mu}_{\xi_k}$ and use the (weighted) Euclidean distance in the cost function. However, in contrast to classical LQR here the desired trajectory is a sequence of multivariate normal distributions representing the uncertain human motion behavior prediction. The Euclidean distance is no longer appropriate to reflect the

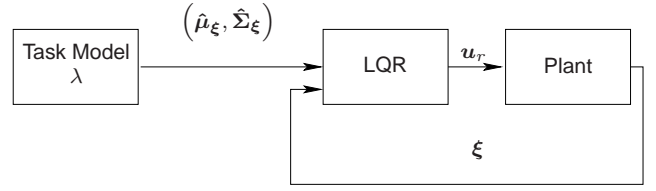


Fig. 3. Assistive control scheme for the standard stochastic optimal control case. The learned task model provides a target set point for the controller, generating the robot's control input based on the control law from (7).

distance to such a distribution. Instead the (weighted) Mahalanobis distance is chosen here as a suitable alternative as it is capable of incorporating the covariance of the prediction into the distance metric. The cost function then becomes

$$J = \mathbf{z}_T^T \hat{\Sigma}_{\xi,T}^{-\frac{1}{2}} Q \hat{\Sigma}_{\xi,T}^{-\frac{1}{2}} \mathbf{z}_T + \sum_{k=1}^{T-1} (\mathbf{z}_k^T \hat{\Sigma}_{\xi,k}^{-\frac{1}{2}} Q \hat{\Sigma}_{\xi,k}^{-\frac{1}{2}} \mathbf{z}_k + \mathbf{u}_{rk}^T R \mathbf{u}_{rk}), \quad (6)$$

where $\hat{\Sigma}_{\xi,k}$ represents the covariance of the desired trajectory. Note that the state weighting matrix in the original LQR cost function (5) is now represented by $Q_k = \hat{\Sigma}_{\xi,k}^{-\frac{1}{2}} Q \hat{\Sigma}_{\xi,k}^{-\frac{1}{2}}$ and in the end-term $Q_T = \hat{\Sigma}_{\xi,T}^{-\frac{1}{2}} Q \hat{\Sigma}_{\xi,T}^{-\frac{1}{2}}$. Also, for the following derivations we will assume a constant control weighting matrix R . The inclusion of the Mahalanobis distance allows the following intuitive interpretation: When the prediction variance $\hat{\Sigma}_\xi$ is high, i.e. the robot is very unsure about the desired trajectory, the Euclidean distance to the current point of the desired trajectory will be less relevant and the control cost $\sum_{k=1}^{T-1} \|\mathbf{u}_{rk}\|^2$ will become more significant, i.e. the optimization will tend to reduce the robot control input. On the other side, lower variances imply more certainty about the desired trajectory and a higher contribution of the robot to track it, as depicted in Fig. 2.

As the system is stochastic 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 such a LQR problem is given by

$$\mathbf{u}_{rk} = -K_k \mathbf{z}_k, \quad (7)$$

where K_k is the feedback matrix given by the Riccati recursion

$$K_k = -R^{-1} B' (B R^{-1} B' + \Pi_{k+1}^{-1})^{-1} A \quad (8)$$

and

$$\Pi_k = Q_k + A' (B R^{-1} B' + \Pi_{k+1}^{-1})^{-1} A, \quad (9)$$

being $\Pi_T = Q_T$.

This assistive control strategy, shown in Fig. 3, tracks the predicted trajectory considering its covariance.

B. Risk-Sensitive Optimal Control

While the LQR solution considers the variance $\hat{\Sigma}_\xi$ as a measure of how precise the trajectory should be, the process noise ϵ has no influence on the control law given by (7). A risk-sensitive controller considers the noise of the control input when calculating the optimal control law and adapts the

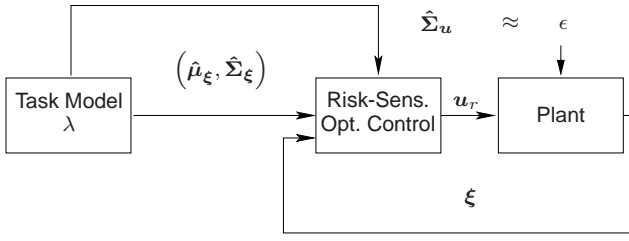


Fig. 4. Assistive control scheme for the risk-sensitive optimal control case. Based on the scheme from Fig. 3, the unexpected behavior of the human is additionally modeled as a process noise ϵ and estimated as the expected variance of the control input $\hat{\Sigma}_u$ that the risk-sensitive controller takes into account.

feedback matrix depending on a risk-sensitivity parameter θ minimizing the cost function

$$\gamma(\theta) = -2\theta^{-1} \ln \mathbb{E}[\exp^{-\frac{1}{2}\theta J}] . \quad (10)$$

Considering the Taylor series expansion of $\gamma(\theta)$,

$$\gamma(\theta) = \mathbb{E}[J] - \frac{1}{4}\theta \text{Var}[J] + \dots , \quad (11)$$

it is observable that not only the expected cost is incorporated, as in the classical LQR case, but also higher order moments of it. If $\theta = 0$, then the controller is risk-neutral and corresponds to the LQR case explained in Sec. III-A. 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 [18]

$$K_k = -R^{-1}B'(BR^{-1}B' + \theta\hat{\Sigma}_u + \Pi_{k+1}^{-1})^{-1}A , \quad (12)$$

and

$$\Pi_k = Q_k + A'(BR^{-1}B' + \theta\hat{\Sigma}_u + \Pi_{k+1}^{-1})^{-1}A , \quad (13)$$

with $\Pi_T = Q_T$ and $\hat{\Sigma}_u$ the variance of the normally distributed zero mean noise term $B\epsilon$ in (3).

Note that the solutions of (12) and (13) are the same as the ones from (8) and (9), except for the term $\theta\hat{\Sigma}_u$. Consequently, for risk-averse behavior $\theta < 0$, the feedback gain will become higher adopting a pessimist attitude as if the noise term $B\epsilon$ was directing the state in the wrong direction. For the opposite case $\theta > 0$, a risk-seeking behavior, the controller will react in an optimistic manner, applying a smaller gain, as the process noise is assumed to be acting in the right direction.

The process noise ϵ can be interpreted as the unexpected behavior of the human partner. In fact, it could be interpreted as the *level of disagreement* between the partners as depicted in Fig. 5. Solving the optimization problem requires an estimation of the noise level. If the probabilistic model λ represents not only the state, but also the resulting control input u , the process noise can be estimated as the expected variance of u , i.e. a normally distributed noise with zero mean and covariance matrix given by $\hat{\Sigma}_u$, as shown in Fig. 4. A high variance $\hat{\Sigma}_u$ implies that an unexpected behavior from the human side could happen and a higher noise is

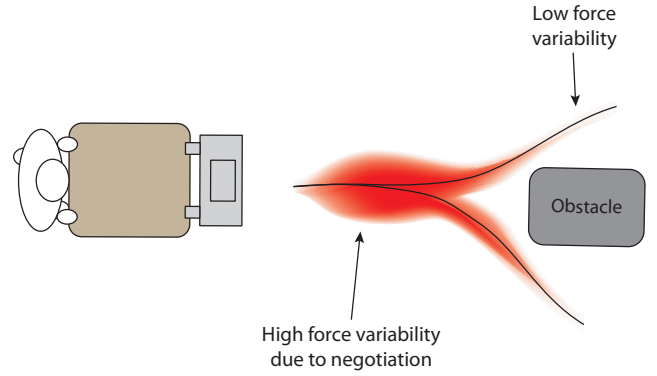


Fig. 5. Exemplary representation of the expected level of disagreement in terms of $\hat{\Sigma}_u$ for a binary path following decision problem. Before a divergence point, the expected disagreement level is higher, as the partners might not agree.

considered in the dynamics. A low control input variance implies a low noise in the system and the expected behavior is closer to the noise-free dynamics.

IV. TASK MODEL ACQUISITION

Acquiring a model of a joint task with a human involved, suggests the use of learning by demonstration techniques. Based on observations of human behavior, a model of the desired task can be obtained as the generalization of multiple executions. In the pHRI case, the fact that both partners are physically coupled makes the determination of the human intention more difficult as it might be influenced by the robot. Furthermore, the control method as explained in Sec. III, already requires a task model, which is also supposed to represent the intentions of the human while interacting with its partner. In order to provide a preliminary model to the assistive robot control, an execution of the task without assistance, i.e. $u_r = 0$ is taken here as a rough representation of the task. This rough model is incrementally improved in several executions where the robot now already tries to assist the human partner. As a result a task model of the human interacting with an active robot is generated. This model reflects different trajectories as well as conflicts and disagreements in terms of higher variances of the control input.

To encode and generalize the observations $O = \{\xi, u\}$ we use an HMM as a probabilistic model of the task. HMMs are discrete state space models that can successfully represent time-series data [19] including force and motion data [7]. Their discretization is already a generalization of the represented observations and, by the methods explained in [20], a smooth generalized trajectory and heteroscedatic variances can also be obtained using Gaussian Mixture Regression.

V. EXPERIMENTS

In order to evaluate the performance of the controller we chose a virtual environment where a heavy virtual object has to be moved along a given trajectory in two dimensions.

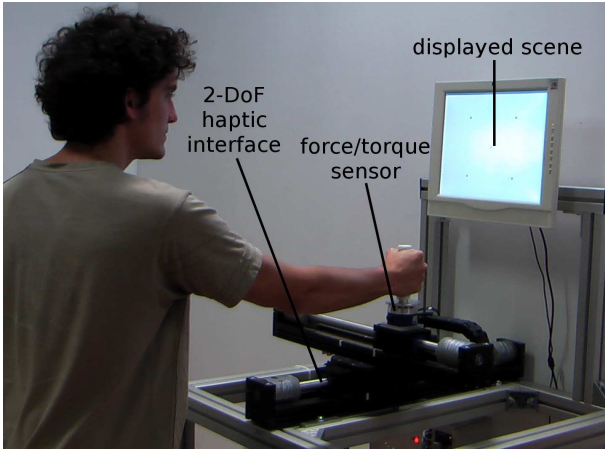


Fig. 6. Experimental setup: 2-DoF haptic device and a virtual environment.

After building a model of the task based on the human behavior during the interaction with the robot, the performance of four different controllers with different risk-sensitivities are compared in terms of their generated feedback matrices and their generated force profile.

A. Experimental Setup

The human partner applies forces to a haptic interface in order to move a virtual object along a given trajectory, see Fig. 6 for the experimental setup. The haptic interface consists of a 2DoF linear-actuated device (*ThrustTube*) having a free-spinning handle at the grasping point. Attached to the handle, a force/torque sensor (*JR3*) measures the human force input. The virtual scene is visually represented on a display placed on top of the interface. The task to be executed was to follow a path describing the letter 'p'.

The control scheme implemented in *MATLAB/Simulink* is executed on a personal computer on the *Linux Real-Time Application Interface (RTAI)* using *Matlab's Real-Time Workshop*. The shared object is physically rendered as a point mass with an inertia of $M_r = 40$ kg and damped by a viscous friction of $D_r = 90$ Ns/m, emulating a heavy object. The control scheme runs at a sampling rate of 1 kHz.

The model of the task is acquired from two different executions of the task, depicted in Fig. 8. In the first run the human leads the task and no robot assistance $\mathbf{u}_r = 0$ is provided. The resulting motion observations are used as the target trajectory for the risk-neutral controller during the second run, when the robot is already assisting. During this second execution, the human provokes some disagreements trying to follow a slightly different path and velocity profiles in order to induce relevant variances. The observations are encoded into a HMM with 40 states. The HMM library is implemented in C++. For simplicity of illustration, we only consider the diagonal of the covariance matrices in the further control computation.

We test four different assistive controllers depending on its risk-sensitivity and its cost function:

- (a) LQR with standard Euclidean distance: $\theta = 0$ and using the cost function from (5).

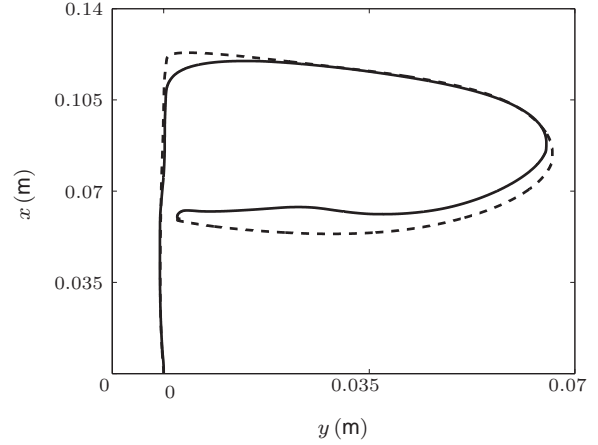


Fig. 8. Position trajectories of the two executions of the task.

- (b) Risk-neutral with Mahalanobis distance: $\theta = 0$ and using the cost function from (10) with J as in (6).
- (c) Risk-averse with Mahalanobis distance: $\theta = -\beta$ and using the cost function from (10) with J as in (6).
- (d) Risk-seeking with Mahalanobis distance: $\theta = \beta$ and using the cost function from (10) with J as in (6).

In our experiments we choose $\beta = 2 \cdot 10^{-7}$, $R = I$ the identity matrix. For (a), we choose a diagonal $Q_k = Q_T = \text{diag}\{\omega_p, \omega_v\}$ with $\omega_p = 10^5$ and $\omega_v = 10^4$ the position and velocity weightings, respectively. For (b)-(d) we set $Q = \text{diag}\{\omega_p, \omega_v\}$ with $\omega_p = 10^{-2}$ and $\omega_v = 10^{-3}$ because of the low values of the variances in the order of 10^{-5} . The general properties of the results are not affected by that.

In the two-dimensional case, the state of the system at time k is given by $\xi_k = (x_k \ y_k \ v_{x,k} \ v_{y,k})^T$. For this scenario, the optimal control law from (7) becomes

$$\begin{pmatrix} 0 \\ 0 \\ u_{rx,k} \\ u_{ry,k} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ K_{vx,k} & 0 & 0 & 0 \\ 0 & K_{vy,k} & 0 & 0 \end{pmatrix} \begin{pmatrix} x_k - \hat{\mu}_{xk} \\ y_k - \hat{\mu}_{yk} \\ v_{x,k} - \hat{\mu}_{v_{xk}} \\ v_{y,k} - \hat{\mu}_{v_{yk}} \end{pmatrix} \quad (14)$$

The four different controllers are evaluated in term of their feedback gains $\{K_x, K_y, K_{v_x}, K_{v_y}\}$, with respect to the variances $\{\hat{\sigma}_x^2, \hat{\sigma}_y^2, \hat{\sigma}_{v_x}^2, \hat{\sigma}_{v_y}^2\}$, and the process noise $\{\hat{\sigma}_{u_x}^2, \hat{\sigma}_{u_y}^2\}$.

B. Experimental Results

The resulting control gains for the settings (a)-(d) are shown in Fig. 7 in rows 1 and 3, the variances in rows 2 and 4, and the noise in row 5. The first and the second column represent the x-components and the y-components, respectively.

The gain profiles of the classical LQR with standard Euclidean distance controller (a) behaves as expected: The gains remain largely constant. This is qualitatively different to the gains for (b)-(d) which vary over time.

For the settings (b)-(d) the controller gains exhibit similar characteristics in terms of their profiles even if they differ

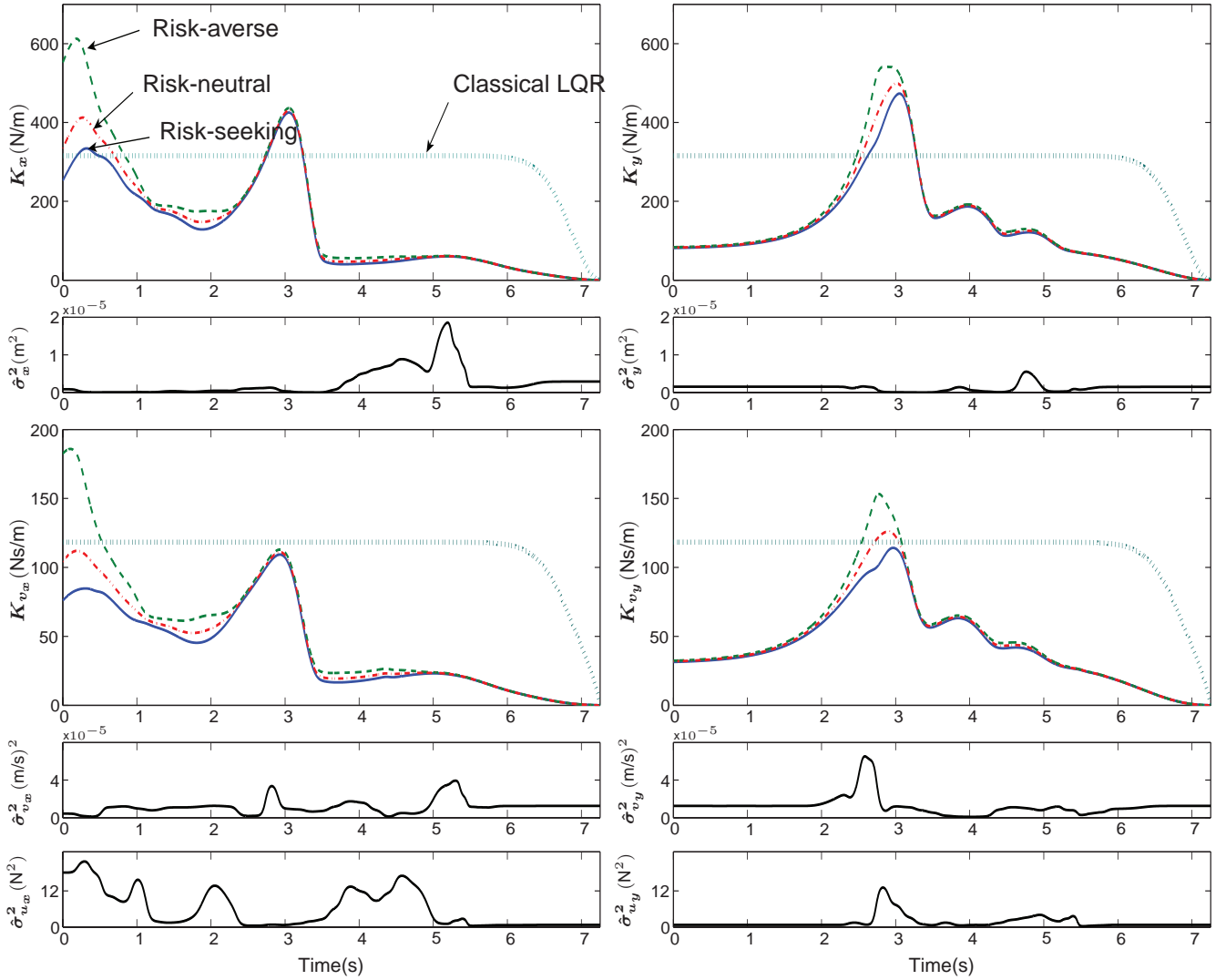


Fig. 7. Control gains for the different controllers together with the variances for the predictions and process noise. The first row and the third row show the position and the velocity gain respectively, while the second and the fourth rows show the predicted position and the velocity variance. The fifth row shows the process noise in terms of the control input variance. The first column of each row refer to the x-component and the second one to the y-component. In rows 1 and 3, the dotted cyan line corresponds to (a) LQR with standard Euclidean distance, the red dash-dotted line to (b) risk-neutral $\theta = 0$ with Mahalanobis distance, the green dashed line to (c) risk-averse $\theta < 0$ with Mahalanobis distance, and the blue solid line to (d) risk-seeking $\theta < 0$ with Mahalanobis distance.

in value. The variances of the tracking error for the position (row 2) and the velocity (row 4) are considered in the Mahalanobis distance for the cost function and clearly influence the results. Lower variance periods imply higher gains making the robot adopt a more rigid behavior, while higher variances produce lower gains, i.e. less stiffness/damping, tolerating higher deviations from the expected state trajectory and letting the human partner more flexibility, as shown in Fig. 9. This effect becomes more evident in the position gains (row 1) than in the velocity gains (row 3) because the position accuracy factor ω_p is set higher than the velocity accuracy factor ω_v .

The setting (b)-(d) differ in their risk-sensitivity, i.e. how the noise / level of disagreement (row 5) affects the controller as depicted in Fig. 10. The risk-neutral controller (b) depicted by the red dash dotted line does not consider the process noise as $\theta = 0$. However, the risk-averse controller (c)

considers the process noise as if it was leading the dyad to a wrong direction and therefore the gains during periods with a higher noise become higher in order to track the reference more dominantly. On the other hand, the risk-seeking controller (d) assumes that the process noise is leading the system to the right direction and consequently the gains during periods with a higher noise become smaller, as the noise is assumed to be already doing part of the work.

These results show that the risk-averse controller becomes stiffer adopting higher gains in case of higher noise. In case of disagreement, the robot tries to impose its path to the human partner becoming more dominant. On the other hand, the risk-seeking controller becomes less stiff, suggesting its partner to decide the way to follow.

In summary, the extension of the classical LQR framework to include the uncertainty of the human behavior prediction and disagreement through the Mahalanobis distance and

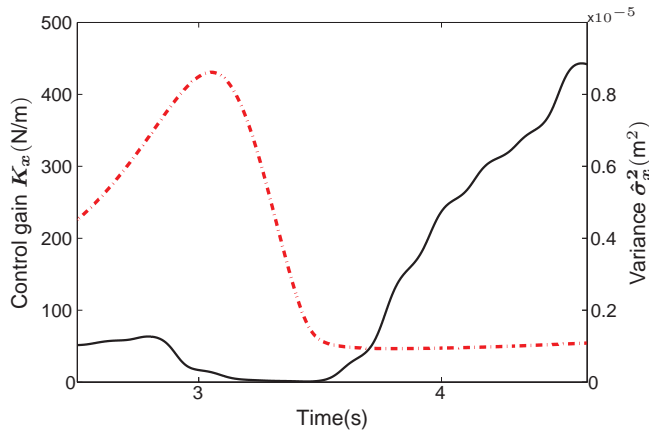


Fig. 9. Experiment excerpt showing the the x-component of the position gain for the risk-neutral controller (b), red dash-dotted line, together with the position variance, black line. Due to the inclusion of the Mahalanobis distance, when high variances are expected, the controller gain falls accordingly and vice versa.

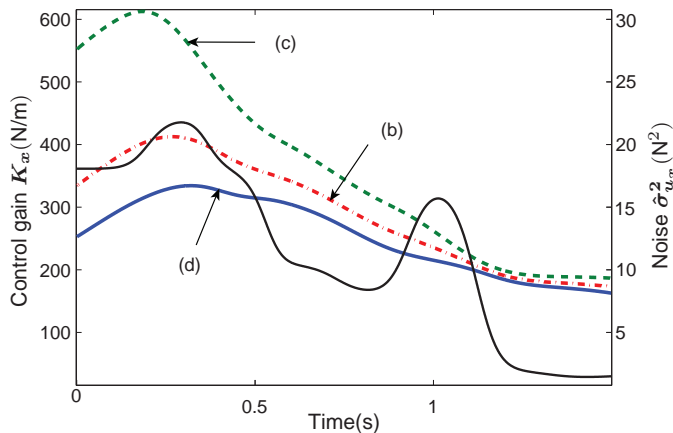


Fig. 10. Experiment excerpt showing the x-component of the position gain for the risk-sensitive controllers together with the expected noise. The blue, red and green lines represent the controllers (b),(c) and (d) respectively. The black line represent the expected noise level. Under high noise levels, the controllers adapt their gains depending on their risk-sensitivity. Low noise levels lead to similar gains for the three controllers.

the concept of risk-sensitivity provides meaningful results concerning a confidence-based role allocation to the robot. It remains to show in future human user studies that this behavior is indeed beneficial for haptic assistance.

VI. CONCLUSIONS

In this paper we present a novel control approach for haptic assistance. The approach is based on the concept of risk-sensitivity and the replacement of the Euclidean distance metrics by the Mahalanobis distance in tracking control problems which are based on uncertain reference trajectories. The uncertain reference trajectory stems from a task/ human behavior model which is inevitable uncertain and which we assume to be represented by probabilistic models (HMMs). With the novel approach the assistive control feedback gains and therewith the robot role adapt to the uncertainties of the human behavior prediction. Moreover, depending on the risk-sensitivity the robot manages disagreement between the partners measured in terms of the force variance in a

more or less dominant way. The approach is successfully evaluated in a 2-DoF experiment with a virtual environment. Studying human preferences regarding the risk-sensitivity and applying this control scheme to more complex tasks is the matter of our future work.

ACKNOWLEDGMENTS

This research is partly supported by the DFG excellence initiative research cluster “Cognition for Technical Systems CoTeSys” and the Caja Madrid Foundation.

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.
- [2] K. Yokoyama, H. Handa, T. Isozumi, Y. Fukase, K. Kaneko, F. Kanehiro, Y. Kawai, F. Tomita, and H. Hirukawa, “Cooperative Works by a Human and a Humanoid Robot,” in *Proc. IEEE ICRA*, 2003, pp. 2985–2991.
- [3] M. Lawitzky, A. Mörtl, and S. Hirche, “Load Sharing in Human-Robot Cooperative Manipulation,” in *Proc. IEEE Ro-Man*, 2010, pp. 185–191.
- [4] B. Corteveille, 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] P. Evrard and A. Kheddar, “Homotopy Switching Model for Dyad Haptic Interaction in Physical Collaborative Tasks,” in *Proc. EHS EuroHaptics*, 2009, pp. 45–50.
- [6] N. Jarrasse, J. Paik, V. Pasqui, and G. Morel, “How can human motion prediction increase transparency?” in *Proc. IEEE ICRA*, 2008, pp. 2134–2139.
- [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] E. Todorov, “Optimality principles in sensorimotor control,” *Nature Neuroscience*, vol. 7, pp. 907–15, 2004.
- [9] —, “Optimal control theory,” 2006.
- [10] D. Mitrovic, S. Klanke, R. Osu, M. Kawato, and S. Vijayakumar, “A computational model of limb impedance control based on principles of internal model uncertainty,” *PLoS ONE*, vol. 5, no. 10, 2010.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] P. Whittle, “Risk-sensitive linear/quadratic/gaussian control,” *Advances in Applied Probability*, vol. 13, no. 4, pp. pp. 764–777, 1981.
- [15] 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.
- [16] 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.
- [17] D. A. Braun, A. J. Nagengast, and D. M. Wolpert, “Risk-sensitivity in sensorimotor control,” *Frontiers in Human Neuroscience*, vol. 5, no. 1, 2011.
- [18] A. Shaiju and I. Petersen, “Formulas for discrete time LQR, LQG LEQG and minimax LQG optimal control,” in *Proc. IFAC*, 2008.
- [19] D. Kulic, D. Lee, C. Ott, and Y. Nakamura, “Incremental learning of full body motion primitives for humanoid robots,” in *IEEE Humanoids*, 2008, pp. 326–332.
- [20] D. Lee and C. Ott, “Incremental kinesthetic teaching of motion primitives using the motion refinement tube,” *Auton. Robot.*, pp. 1–17, 2011.