Dynamic Role Adaptation with Risk-Sensitive Optimal Feedback Control in Physical Human-Robot Interaction

Masao Saida[1], José Ramón Medina[2], and Sandra Hirche[2]

Abstract— **Anticipatory behavior based on the human behavior prediction enables the robot to improve the quality of its assistance in physical human-robot interaction (pHRI). However, predictions are partly afflicted with high uncertainties originating from the intrinsic variability in human behavior and the influence of the environment which requires the negotiation among the partners. In this paper, we propose a novel control approach that dynamically adapts the robot's role to uncertainties in real time facilitating the negotiation between the human and the robot. The approach is based on risk-sensitive optimal feedback control. The negotiation between the human and the robot is realized through a dynamical changing risk-sensitivity parameter. The proposed approach is experimentally validated in a cooperative transport scenario in a two-dimensional visuo-haptic virtual environment.**

I. Introduction

Robots of the future are expected to enter daily-life such as domestic and medical/welfare environments. Physical human-robot interaction is one of the major challenges because of the physical coupling between the human and the robot requiring real-time decision making capabilities of the robot. For intuitive interaction with the human and an effective contribution to the task, the robot should predict the human's behavior and assist the human proactively. However, a robot's active behavior based on an incorrect prediction could be more of a hindrance than a help. Therefore, it is very important to consider uncertainties arising from the intrinsic variability in human behavior. An additional source of uncertainty stems from the (dynamic) environment. Both aspects, require the robot to interpret haptic disagreement signals online and to negotiate the motion in a seamless and intuitive way.

Physical human-robot interaction (pHRI) issues are studied for example in cooperative load transport [1]-[3]. The robot's anticipative assistance based on the prediction of the human's motion can reduce the human's effort in the task [4]. Active robotic contribution based on the prediction of the human behavior is investigated in point to point movements [5] , [6], and full-scale transport tasks in constrained environments as in our previous work [7]. In these works, however, the prediction uncertainties are not taken into account in the proactive assistance.

Following the conceptual idea that the robot should behave less pro-active in situations where it is very uncertain about the human motion intention, we propose a novel interaction control concept based on risk-sensitive optimal feedback control [8]. Within this control approach we consider the human behavior being represented by a probabilistic model learned in previous executions. Modeling the human's unexpected behavior as process noise, the proposed control makes the human's variability or *disagreement* directly influence the robot's behavior. Instrumental for that is the risk-sensitivity parameter, which determines how strongly the disagreement is considered in the robot control policy and therefore the level of pro-activity in the robot behavior, synthesizing its *role* w.r.t. the human. However, this model considers only previous executions, i.e. the robot considers the *expected* process noise in feedforward fashion. The control is not adapted to any unexpected behavior not reflected in the model. Therefore, we also propose an improved control scheme to accommodate for unmodeled human execution variability [9]. This approach estimates the *current* process noise, which is derived from the current force input applied by the human in a feedback fashion. Nevertheless, the risk-sensitivity parameter is so far kept constant during the task execution. Our psychological experiment in [9] shows that, depending on the situation, an active robotic assistant should dynamically adapt its risk-sensitivity, adopting a passive or dominant role.

The contribution of this paper is a scheme to dynamically change the risk-sensitivity parameter adapting the robot's role during the interaction. The continuous negotiation process with the human partner reflected in a variable risk-sensitivity is synthesized considering both the predicted disagreement level and environmental constraints. In addition, the proposed approach is experimentally validated in a human user study in a cooperative transport scenario in a two-dimensional visuo-haptic virtual environment.

The remainder of this paper is organized as follows: Section II describes the general architecture of the proposed control scheme. The assistive control with the human behavior learning and the process noise estimation is explained in Section III. We propose the dynamic role adaptation through dynamic risk-sensitivity in Section IV, and experiments for evaluation of the proposed approach are presented and discussed in Section V.

II. General Architecture

In this paper, as an exemplary representation of pHRI, we consider a joint object transport task: A human and a robot move an object in cooperation from a start to a goal configuration while being physically coupled. For simplicity,

[1]M. Saida is with Derartment of Bioengineering and Robotics, Tohoku University, 6-6-01, Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan. saida@irs.mech.tohoku.ac.jp

[2]J. Medina and S. Hirche are with the Institute of Automatic Control Engineering, Technische Universität München, D-80290 Munich, Germany. {medina, hirche}@tum.de

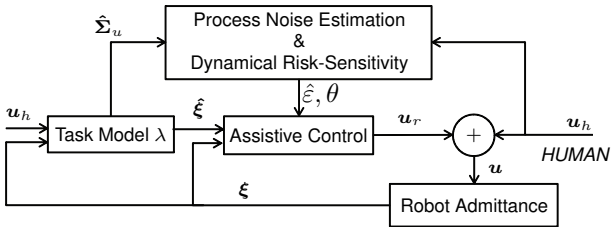


Fig. 1. General framework of the control scheme

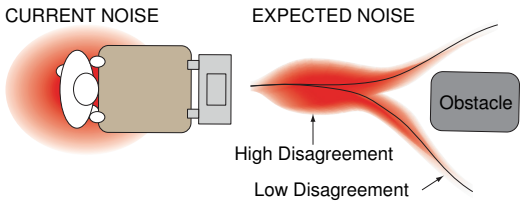


Fig. 2. Current process noise and expected process noise

in our derivations we assume that an interaction contact point between the human and the robot is the representative control point, i.e. we consider the object to be represented by a point mass. The extension to arbitrary object geometries involving geometrically distributed grasp points is straightforward using the formulation from [3]. Also for simplicity, we consider the planar case. The general architecture of the robot control is shown in Fig. 1. We implement an admittance control law that considers human force input $\boldsymbol{u_h} \in \mathbb{R}^{2\times 1}$ for a compliant reactive behavior of the robot and an assistive control input $\boldsymbol{u_r} \in \mathbb{R}^{2\times 1}$ based on the human motion prediction. The system, considering decoupled dynamics, can be expressed by

$$\boldsymbol{M}_r\ddot{\boldsymbol{x}} + \boldsymbol{D}_r\dot{\boldsymbol{x}} = \boldsymbol{u_h} + \boldsymbol{u_r}, \tag{1}$$

where $\boldsymbol{M}_r \in \mathbb{R}^{2\times 2}$ and $\boldsymbol{D}_r \in \mathbb{R}^{2\times 2}$ are the rendered diagonal inertia matrix and damping coefficient matrix, respectively, $\boldsymbol{\xi} = [\boldsymbol{x} \quad \dot{\boldsymbol{x}}]^T$ is the state of the system, where $\boldsymbol{x} \in \mathbb{R}^{2\times 1}$ is the position of the control point.

The assistive control input $\boldsymbol{u_r}$ is generated by the assistive control of risk-sensitive optimal feedback control, as explained in Section III-A. A probabilistic model is learned in previous executions, as explained in Section III-B. Furthermore, the process noise estimation $\hat{\varepsilon}$ considers both the *expected* and the *current* process noise, as explained in Section III-C, as shown in Fig. 2. The former is based on previous observations and the latter is derived from the current human input $\boldsymbol{u}_h$. In addition, the dynamic risk-sensitivity θ is defined, as explained in Section IV. As a result, the robot role is adapted to uncertainties in real-time.

We assume that the human has its own motion plan represented by a desired trajectory $\boldsymbol{\xi_d}$ and generates the required force input $\boldsymbol{u_d}$ to track the desired trajectory $\boldsymbol{\xi_d}$. In order to track the predicted trajectory $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$, the robot generates an estimate $\hat{\boldsymbol{u}}_{\boldsymbol{d}}$ of $\boldsymbol{u_d}$. The discrepancy between both generated control inputs leads to a corrective force from the human side given by $\boldsymbol{u_h} = \boldsymbol{u_d} - \hat{\boldsymbol{u}}_{\boldsymbol{d}}$. Hence, we model the human force input as a process noise in the system dynamics from (1), i.e. $\boldsymbol{u_h} = \varepsilon$. Here we assume it normally distributed with zero mean as we assume that the demonstrated trajectories estimated by $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$ reflect the way, the human would like to perform the task. Therefore $\hat{\boldsymbol{u}}_{\boldsymbol{d}}$ is considered as an unbiased estimate of $\boldsymbol{u_d}$. The discretized version of the system dynamics (1) with a sampling time interval Δt in the form $\boldsymbol{\xi}_{k+1} = \boldsymbol{A}\boldsymbol{\xi}_k + \boldsymbol{B}\boldsymbol{u}_k$ is then given by

$$\begin{bmatrix}\boldsymbol{x}_{k+1}\\ \dot{\boldsymbol{x}}_{k+1}\end{bmatrix} = \begin{bmatrix}1 & \Delta t\\ 0 & 1 - \boldsymbol{M}_r^{-1}\boldsymbol{D}_r\Delta t\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_k\\ \dot{\boldsymbol{x}}_k\end{bmatrix} + \begin{bmatrix}0\\ \boldsymbol{M}_r^{-1}\Delta t\end{bmatrix}[\boldsymbol{u_r}_k + \varepsilon_k], \tag{2}$$

where k stands for the time instance. The assistive control input is generated based on the dynamics (2).

III. Assistive Control Design

The goal of the assistance control is to dynamically adapt robot effort to solve the task depending on uncertainties while tracking the predicted state $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$ until a given time horizon T. The dynamic role adaptation enables the robot to reduce the disagreement by negotiating with the human. In order to address this issue we propose a risk-sensitive optimal feedback control approach [8].

A. Risk-Sensitive Optimal Feedback Control

With the reference trajectory given by the probabilistic approximation $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$, we define the tracking error estimate as $\boldsymbol{z}_k = \hat{\boldsymbol{\xi}}_{\boldsymbol{d}k} - \boldsymbol{\xi}_k$. However, the reference state $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}} \simeq \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}})$ is a sequence of multivariate normal distributions. Therefore, the cost function in the minimization problem with the dynamics expressed as (2) is represented in the classical Linear Quadratic Control (LQR) framework with the (weighted) Mahalanobis distance as follows

$$J = \boldsymbol{z}_T^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{-\frac{1}{2}}Q\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},T}^{-\frac{1}{2}}\boldsymbol{z}_T + \sum_{k=1}^{T-1}(\boldsymbol{z}_k^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}^{-\frac{1}{2}}Q\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}^{-\frac{1}{2}}\boldsymbol{z}_k + \boldsymbol{u_r}_k^{\mathrm{T}}R\boldsymbol{u_r}_k), \tag{3}$$

where $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi},k}$ represents the covariance of the desired trajectory, and Q and R are positive definite weighting matrices that allow a trade-off between control input cost and human contribution minimization.

Next, in order to consider the process noise ε in the optimal control law calculation, we adopt a risk-sensitive optimization [10], [11]. In this case, the cost function is defined by

$$\gamma(\theta) = -2\theta^{-1}\ln\mathbb{E}[\exp^{-\frac{1}{2}\theta J}]. \tag{4}$$

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 is given by

$$\boldsymbol{u_r}_k = K_k\boldsymbol{z}_k = K_k(\hat{\boldsymbol{\xi}}_{\boldsymbol{d}k} - \boldsymbol{\xi}_k), \tag{5}$$

where K_k is the feedback matrix given by a modified form of the Ricatti recursion [12] as follows

$$K_k = -R^{-1}B^{\mathrm{T}}(BR^{-1}B^{\mathrm{T}} + \theta\varepsilon + \Pi_{k+1}^{-1})^{-1}A, \tag{6}$$

and

$$\Pi_k = Q_k + A^{\mathrm{T}}(BR^{-1}B^{\mathrm{T}} + \theta\varepsilon + \Pi_{k+1}^{-1})^{-1}A, \quad (7)$$

with $\Pi_T = Q_T$ and the process noise ε is estimated as normal distribution with zero mean and covariance given by $\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$ in the task model λ. The term $\theta\varepsilon$ produces the following effect in the robot behavior: If $\theta = 0$ the controller is risk-neutral and corresponds to the LQR case. For $\theta < 0$ and $\theta > 0$ the control becomes risk-averse and risk-seeking, respectively.

B. Learning and Prediction of Human Behavior

In our approach, a preliminary model of the task is acquired by an initial execution of the task without assistance, i.e. $\boldsymbol{u_r} = 0$. Using this first rough representation, the robot actively assists and incrementally observes interaction with the human during additional task executions. Modeling the state and control input, $\{\boldsymbol{\xi}, \boldsymbol{u}\} \simeq \{\mathcal{N}(\boldsymbol{\mu_\xi}, \boldsymbol{\Sigma_\xi}), \mathcal{N}(\boldsymbol{u}, \boldsymbol{\Sigma_u})\}$, the robot acquires a representation of the task that represents both the desired state trajectory $\boldsymbol{\xi_d}$ of the human and the level of expected disagreement between partners.

For realizing a safe interaction, we use a time based HMM, applying regression in time domain as explained in [13]. This provides a generalized trajectory of the task in terms of means and heteroscedastic variances. Using the Viterbi Algorithm in a window over the last observations and estimating the current sample of the generalized trajectory, the next state $\hat{\boldsymbol{\xi}}_{\boldsymbol{d}}$ is predicted, see for details [7].

C. Process Noise Estimation

The expected and the current process noise are a measure of the expected and current *level of disagreement* respectively, as shown in Fig. 2. They are estimated as the variance of the human control input $\boldsymbol{u_h}$, i.e. a normally distributed noise with zero mean and covariance matrix given by $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}$ and $\boldsymbol{\Sigma_u}$, respectively. In order to consider both, the current and the expected process noise in the optimization, we approximate the process noise level $\hat{\varepsilon} = \mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\varepsilon}}})$ as $\max(\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}), \mathcal{N}(0, \boldsymbol{\Sigma_u}))$. Specifically, we perform a Gaussian approximation $\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\varepsilon}}})$ for this problem, where $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\varepsilon}}}$ is defined by 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 $\mathcal{N}(0, \boldsymbol{\Sigma_u})$, as explained in detail in [9].

D. Relation between Robot Gains and Risk-Sensitivity

The *aggressiveness* of the robot behavior is strongly related to the control gains, e.g. high gains mean more aggressive behavior following the prediction. As mentioned in Section III-A, the risk-sensitivity parameter influences the control gain, i.e. $K(\hat{\varepsilon}, \theta)$ from (6), in the sense of how strongly a disagreement influences the gain. Therefore, the proposed control changes its behavior depending on the parameter even for the same noise. For $\theta > 0$ the gains become lower, i.e. the robot behaves more passively. On the other hand, for $\theta < 0$ higher gains make the robot behave more dominantly. It should be noted that a different selection of the matrices Q_k and R in (6) and (7) may have induce a similar behavior of the control. The concept of risk-sensitivity provides a systematic approach to vary the control gains based on the prediction uncertainty.

IV. Dynamic Role Adaptation in pHRI

The benefit of continuously adapting roles of the robot in pHRI is demonstrated in [5], where the robot role is changed between the follower and leader behaviors learned in the previous observations depending on the current human input. In this work, we consider a dynamic role adaptation depending on both the disagreement level and the environmental situation.

A. Dynamical Risk-Sensitivity Parameter

The proposed risk-sensitivity adaptation seeks a reduction of human effort and an adaptive attitude design by combining two intuitive concepts. On one side, the assumption that the human agrees with the robot contribution when lower disagreement levels are observed w.r.t expected ones. On the other side, the imposition of a dominant or passive role following a given attitude design law. In this work we synthesize this concepts with the *adaptation function* and the *attitude function* respectively.

1) Adaptation Function f: The adaptation function adapts to lower levels of disagreement from the human side when a higher levels are expected, understanding them as an increased level of trust from the human side w.r.t previous experiences. The process noise estimation, explained in Section III-C, considers both the expected and the current disagreement. Under higher current disagreement, the robot can adapt its behavior by updating the process noise as $\hat{\varepsilon} = \mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\varepsilon}}})$ even for the constant risk-sensitivity parameter. However, under lower current disagreement, i.e. $\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\varepsilon}}}) = \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$, the constant risk-sensitivity exaggerates the robot's behavior more than necessary because the gains only adapt to the expected disagreement. From the point of view of robot contribution to the task, the robot should behave more dominantly as the human exhibits a lower level of disagreement than the robot's expectation.

Based on this, the adaptation function f lets the risk-sensitivity parameter assign greater importance to the current than the expected disagreement. Specifically, the function f is defined as the proportion of the current to the expected disagreement $(0 \leq f \leq 1)$, i.e. how low the current one is compared to the expected one, using the Mahalanobis distance as follows

$$f = \begin{cases} \sqrt{\boldsymbol{u_h}^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}^{-1} \boldsymbol{u_h}} & (\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\varepsilon}}}) = \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})) \\ 1 & (\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\varepsilon}}}) > \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})) \end{cases}. \quad (8)$$

2) Attitude Function g: The attitude function defines the disposition of the robot to act in a more passive or dominant way $(-1 \leq g \leq 1)$. In this work we define the attitude function depending strictly on environmental constraints. However, other higher level aspects such as artificial personality or dominance could be also modeled. We consider two different cases

Control Policy	Process Noise ϵ	Parameter θ	Gain
Dominant	$\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\epsilon}}}) > \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$	$-\beta$	$\nearrow\nearrow$
	$\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\epsilon}}}) = \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$	$-\beta < \theta < 0$	$\nearrow$
Neutral	$\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\epsilon}}}) = \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$	0	$-$
Passive	$\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\epsilon}}}) = \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$	$0 < \theta < \alpha$	$\searrow$
	$\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\epsilon}}}) > \mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$	α	$\searrow\searrow$

Note: $\nearrow$ and $\searrow$ show that the gains become higher and lower than the risk-neutral one respectively, $\nearrow\nearrow$ and $\searrow\searrow$ show much higher and much lower respectively, and the approximation $\mathcal{N}(0, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\epsilon}}})$ indicates the minimum volume hyperellipsoid around the set defined by the expected $\mathcal{N}(0, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}})$ and the current $\mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$.

TABLE I

CONCEPT OF DYNAMIC ROLE ADAPTATION

- *Following Role*: defining the attitude function as $g = 1$ and assumming a pure passive behavior. In this work this case is given by unconstrained and safe situations.
- *Leading Role*: changing the attitude function g from 1, purely passive, to -1, purely dominant, in a continuous manner following a certain condition. In this work, we define it by evaluating the potential safety of the current configuration.

Merging both functions f and g, we define the risk-sensitivity parameter θ as

$$\theta = \begin{cases} \alpha \cdot f \cdot g & (0 \leq g \leq 1) \\ \beta \cdot f \cdot g & (-1 \leq g \leq 0) \end{cases}, \tag{9}$$

where α and β are positive constants, and θ is upper-bounded by α in passive behavior or lower-bounded by $-\beta$ in dominant behavior. The bounded values for θ express how strongly the process noise influences the gain. We distinguish α from β from the fact that the order of the parameter in passive behavior is different from the order in dominant behavior.

As a result, the magnitude of θ changes depending on the level of the current process noise and the sign changes based following the given attitude law. This concept enables the robot's behavior to adapt to both the process noise and the environmental constraint *dynamically* and *continuously*. Table I shows the proposed dynamic role adaptation. Under unconstrained situations $g = 1$ the robot follows a more or less passive attitude depending on the adaptation function f. However, if potentially unsafe situations are detected $g \leq 1$, the robot adopts a leading role as it assumes that the previously learned trajectories are safe and tends to track them in a more agreessive manner.

B. Negotiation Model in pHRI based on Game Theory

In the Section IV-A, while the robot role is usually a complementary one adapting the human preference, it tries to lead the human, i.e. to request his/her to perform a complementary role in spite of his/her intention under the environmental constraints. However, if the robot tries to lead the human but the human still disagrees with its intention, the robot might have to change its role policy to avoid disagreement. In pHRI, both partners need to decide their roles between passive and dominant manners negotiating with each other and consequently realize the cooperative task without disagreement. Therefore, this is an *attitude negotiation* issue in pHRI, which is considered as the next step.

Bargaining Problem	$\Leftrightarrow$	**Attitude Negotiation in pHRI**
Two Players	$\Leftrightarrow$	Human and Robot
Mixed Strategies	$\Leftrightarrow$	Roles between leader and follower
Payoff Function	$\Leftrightarrow$	Role differences between two partners

TABLE II

RELATION BETWEEN GAME THEORY AND NEGOTIATION IN PHRI

In this scenario, bargaining problem [15] based on Game theory arises as a suitable modeling method of the attitude negotiation between two partners in pHRI. The strategic form of two-player game [16] consists of a set of players, a set of strategy profiles based on a set of actions, and a set of payoff functions. In the cooperative two-player bargaining game [16], two players, i.e. the human and the robot, employ certain *mixed strategies*, letting them choose a probability distribution over possible actions in cooperation. In pHRI, their roles, which express their infinitely intermediate roles between passive and dominant roles, are interpreted as the actions in the strategies. In addition, the role differences between two partners depending on the situation are regarded as the payoff functions. Based on their strategies and the payoff functions, they negotiate and cooperate to optimize their own *payoffs* following some bargaining solution. As a result, their roles are decided as a result of the attitude negotiation. Table II summarizes the relation between the bargaining problem and the attitude negotiation in pHRI.

Defining the dynamical role differences during task execution as the payoff functions and choosing the proper optimization policy, the bargaining problem enables us to decide the robot role for the attitude negotiation. We consider that the combination of dynamic role adaptation with risk-sensitive optimal feedback control and the attitude negotiation model based on the bargaining game theory is a good solution to realize the intuitive negotiation in pHRI as a future work.

V. EXPERIMENTS

For evaluating the performance of the proposed approach, we conducted two experiments. For simplicity, a human operator transports a heavy virtual object from a starting point to a goal point in cooperation with a haptic interface in a two-dimensional virtual environment. First, we confirm that the robot's role dynamically changes depending on the current disagreement with the human and the surrounding environment as a preliminary experiment. Second, we compare the performance of the four different control approaches with different risk-sensitivities in order to confirm the validity of the proposed algorithm.

A. Experimental Setup

The virtual environment is designed as the experimental setup as shown in Fig. 3. The human applies forces to a

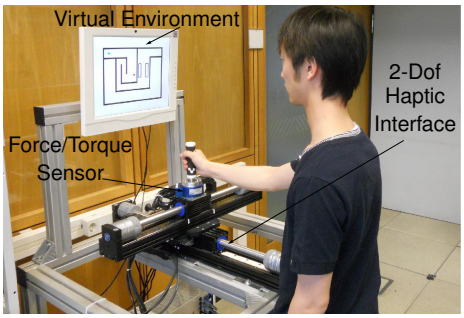


Fig. 3. Experimental setup

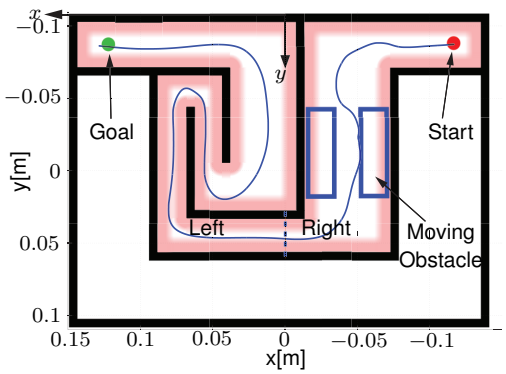


Fig. 4. Virtual environment (maze) and path of shared object (exp.1)

haptic interface in order to move the virtual object. This haptic interface consists of two-degree-of-freedom linear-actuated device (*ThrustTube*) that has a freely-jointed handle at the grasping point. In addition, a force/torque (*JR3*) is attached to the handle in order to measure the force input applied by the human. A virtual maze is visually presented on a display placed on top of the interface, as shown in Fig. 4. On the right side of the maze, two horizontally moving obstacles are set in order to intentionally induce disagreements between the human and the robot. On the left side, the maze includes a narrow path, where both partners are required a high movement precision.

The control scheme 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 an inertia matrix $\boldsymbol{M}_r = diag(m, m)$ with $m = 90$kg and a damping coefficient matrix $\boldsymbol{D}_r = diag(d, d)$ with $d = 200$Ns/m, which imitates a heavy object. All rendered virtual obstacles and walls provide a haptic feel of the environment to the user via the handle. The control scheme runs at 1kHz. The HMM used to encode the observations has 40 states and predictions are updated with a rate of 50Hz.

With reference to the influence of the surrounding environment to the control policy, i.e. the attitude function g, we only consider the distance between the shared object and the walls of the maze for simplicity. In other words, we suppose that the robot has the information of the map, without the moving obstacles, and tries to behave more dominantly near the wall in order to secure the human's safety. Specifically, we define the attitude function g based on the distance between the shared object and the wall as follows

$$g = \begin{cases} -1 & (\rho(\boldsymbol{x}) < \rho_o) \\ -\cos\pi\dfrac{\rho(\boldsymbol{x}) - \rho_o}{\rho_1 - \rho_o} & (\rho_o \le \rho(\boldsymbol{x}) \le \rho_1)\ , \\ 1 & (\rho_1 < \rho(\boldsymbol{x})) \end{cases} \tag{10}$$

where $\rho(\boldsymbol{x})$ is the shortest distance from the current position of the shared object to the wall, ρ_o is the boundary region within which the attitude function influences the risk-sensitivity maximally, and ρ_1 is the limit distance of the attitude function influence, as shown by the pink area near the wall in Fig. 4.

B. Experimental Method

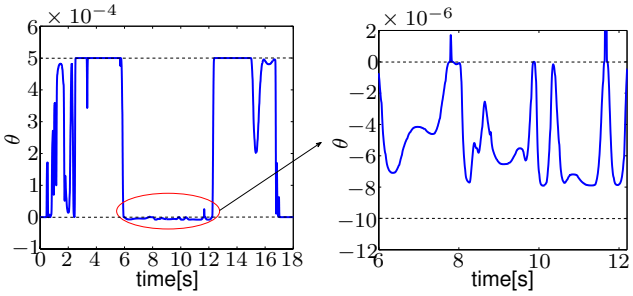
In these experiments, the human carried the object represented by the red point in cooperation with the robot from the start point (upper right-hand corner) to the goal represented by the green point (upper left-hand corner) without hitting any object or wall. The following was the procedure of the experiment. In the first trial, the human lead the task without robot assistance, $\boldsymbol{u_r} = 0$, and the rough model of the task was acquired from the resulting observation. This model was used during the second trial where the robot actively assisted the human partner. The observation of the second trial, as well as the first one, was learned together. As a result, we acquired the model of the task including a model of the interaction between the human and the robot. The third trial used the model learned from the two previous executions, and the third execution of the task was only used for the evaluation of the proposed control as following two experiments.

1) Preliminary Experiment: In the first experiment, we confirmed whether or not the risk-sensitivity parameter dynamically changed depending on the current disagreement and the surrounding environment during the task. We also confirmed the robot's role also changed depending on the process noise and the risk-sensitivity properly. In this experiment, one participant performed 3 trials using the dynamic risk-sensitivity. We set $\alpha = 5.0 \times 10^{-4}$ and $\beta = 1.0 \times 10^{-5}$.

2) Evaluation of Proposed Approach: Second, we tested the four different assistive control approaches depending on its risk-sensitivity:

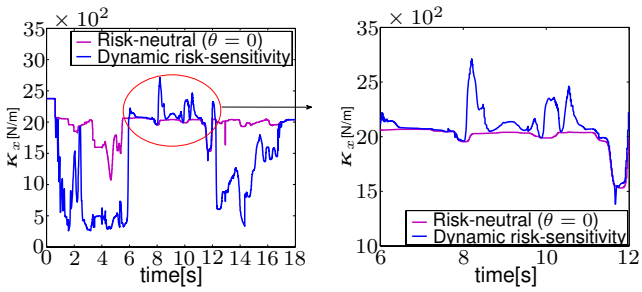
(a) No active assistance: $\boldsymbol{u_r} = 0$.
(b) Passive considering current process noise: $\theta = \alpha$.
(c) Dominant considering current process noise : $\theta = -\beta$.
(d) Dynamic Risk-sensitivity considering current process noise: θ as in (9)-(10).

In this experiment, each participant performed 3 (trials) $\times$ 4 (approaches) $= 12$ trials in all. The order of 4 approaches was selected randomly and the participants were not given any information about it. Here, $\alpha = 5.0 \times 10^{-5}$ and $\beta = 5.0 \times 10^{-7}$, $R = I$ and $Q_k = Q_T = diag(\omega_p, \omega_v)$, being ω_p and ω_v the position and velocity weightings. For (b)-(d), we set $\omega_p = 10^5$ and $\omega_v = 10$. The receding horizon for the optimization was $T = 0.2$s and we used a window



(a) Whole time execution (b) On narrow path

Fig. 5. Risk-sensitive parameter θ (Exp. 1)



(a) Whole time execution (b) On narrow path

Fig. 6. Position gain $\boldsymbol{K}_x$ (Exp. 1)

of $N = 0.05$s for estimating the current noise. Moreover, we chose $\rho_o = 5.6 \times 10^{-3}$m and $\rho_1 = 16.8 \times 10^{-3}$m for the environmental function $\mathcal{H}$. Here, we determined all the parameters empirically. Therefore, we need to evaluate how they interact in more detail as future works.

The experimental procedure was as follows: before the experiment participants were instructed to move at their comfortable speed and complete the task even if they might hit the object or the wall. Next, the participants were asked to face the haptic device and grasp the handle. In addition, 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.

C. Measures for Evaluation

For evaluation, we calculated the following measures:

- the mean force applied by the human $M_{\|\boldsymbol{u_h}\|}$.
- the mean disagreement $M_{\|\boldsymbol{u_D}\|}$ between both partners:

$$\boldsymbol{u_D} = \begin{cases} -\dfrac{(\boldsymbol{u_h})}{\|\boldsymbol{u_h}\|} \cdot \boldsymbol{u_r}, & \left(\boldsymbol{u_h} \cdot \boldsymbol{u_r} < 0\right) \wedge \left(\boldsymbol{u_h} \neq \boldsymbol{0}\right) \\ 0, & \text{otherwise} \end{cases}.$$

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

The maze was divided into two parts in order to evaluate the performance of the dynamically changing risk-sensitivity in different situations. For measures on the right side, the data was calculated until the participant passed the turning point indicated by the dashed line shown in Fig. 4.

D. Experimental Results

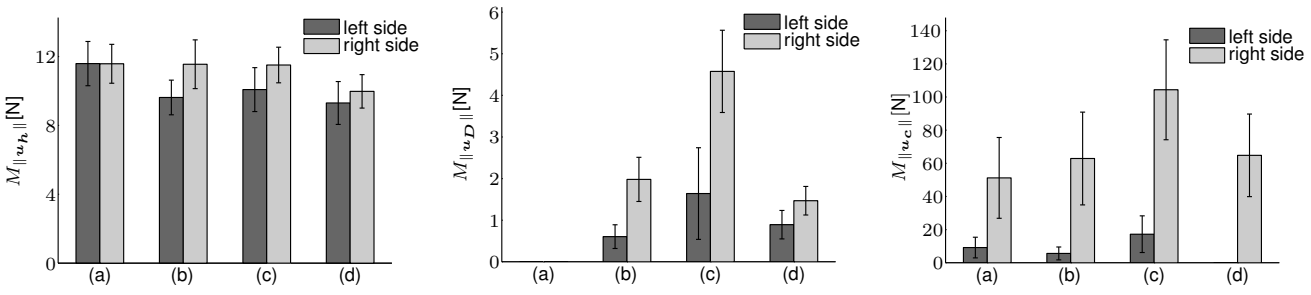
1) Preliminary Experiment: The experimental results are illustrated in Fig. 4-6. Figure 4 shows the path of the shared object on x-y-plane. Figure 5 expresses the risk-sensitivity parameter θ for the dynamical risk-sensitive and risk-neutral control with respect to time. Figure 6 expresses the x-component of the position gain $\boldsymbol{K}_x$, as an example in the feedback gain matrix shown in (5) with respect to time. Especially, Figs. 5(b) and 6(b) show the results on the narrow path of the maze's left side.

From Figs. 4 and 5(a), on the right side and the second half of the left side in the maze, the risk-sensitivity parameter θ changes in positive way depending on the current noise. In addition, from Fig. 6(a), the position gain $\boldsymbol{K}_x$ becomes lower than the risk-neutral one based on the parameter θ. On the other hand, from Figs. 5(b) and 6(b), the parameter θ alters in negative way based on the distance between the object's position and the wall, and the gain becomes higher. Therefore, it is obvious that the robot realizes the dynamic change of the risk-sensitivity parameter and its own role properly depending on both the process noise and the environment.

2) Evaluation of Proposed Approach: In total, 9 persons, including 4 females, participated in the experiment. They were between 25 and 28 years old ($M = 27.11$ years). All participants reported not to suffer any motorical restrictions in their arms.

The calculated measures for evaluation are shown in Fig. 7. With regard to the results on the right side, from Fig. 7(b), the disagreement for dominant approach (c) is highest because the approach (c) becomes stiffer despite the presence of the high current noise induced by the moving obstacles. Consequently, from Fig. 7(a), the force applied by the human for (c) is almost the same as the passive approach (b) although the dominant approach tries to contribute to the task in more dominant manner. On the other hand, both the disagreement and the human force for the approach (d) become lower than (b) because, under low noise situation, the approach (d) behaves more dominantly by decreasing the risk-sensitivity to zero, $\theta \to +0$ and the human agrees with it. In Fig. 7(c), the performance in the collision by approach (d) is almost the same as by (b) and lower than (c). These results indicate that the robot can assist the human more effectively while adapting human's current intention in the presence of no environmental constraints.

Next, regarding the left side of the maze, from Fig. 7(b), the disagreement for (d) is placed between (b) and (c) because the robot's behavior only changes from passively to dominantly when the object approaches the wall. Figure 7(c) shows an important result that no collision between the object and the wall for the approach (d) exists, i.e. any participant did not hit the wall at all on the left side. The reason is because: the approach (d) behaves more and more dominantly depending on the wall approach causing the feedback gains to increase gradually. On the other hand,



(a) Human applied force ($M_{\|\boldsymbol{u_h}\|}$) (b) Disagreement ($M_{\|\boldsymbol{u_D}\|}$) (c) Collision forces ($M_{\|\boldsymbol{u_c}\|}$)

Fig. 7. Experimental results (Exp. 2)

the human understands its role and intention by sensing the change in the gains and agrees with the robot. Consequently, it is considered that all participants can avoid hitting the wall. Also, the human force in (d) became slightly lower than in (b).

In conclusion, the proposed approach exhibits better performance than the constant risk-sensitivity in both cases; left and right side of the maze. Changing the risk-sensitivity dynamically, the robot cannot only adapt human current behavior more flexibly but also communicate its intention with the human via the interaction force based on the resulting feedback gains. However, if the human disagrees with robot's intention, the latter may have to change its policy to avoid disagreement. This is a *negotiation* issue in pHRI which is considered as future work.

VI. Conclusions

In this paper, we propose the dynamic role adaptation using a risk-sensitive optimal feedback control in physical human-robot interaction. We define the dynamical risk-sensitivity parameter based on the current process noise level, i.e. the disagreement level, and environmental constraint which requires the negotiation between the human and the robot. Depending on the parameter, the robot can dynamically change its role between a passive and a dominant manner. As a result, the robot realizes the negotiation with the human communicating its role with him/her via physical interaction force which changes depending on the risk-sensitivity. Finally, the proposed approach is experimentally validated in a cooperative transport scenario in a two-dimensional visuo-haptic virtual environment.

While results indicate a clear benefit of the robot's adaptive role change, the tuning of the parameters is challenging. A systematic approach for tuning the parameters is subject to future work. In addition, we are planning to extend the *dynamic role adaptation* to the *negotiation* in more complex situations.

Acknowledgments

This research is partly supported by the DFG excellence initiative research cluster "Cognition for Technical Systems CoTeSys".

References

[1] Y. Hirata, Y. Kume, Z.-D. Wang, and K. Kosuge, "Handling of a Single Object by Multiple Mobile Manipulators in Cooperation with Human Based on Virtual 3-D Caster Dynamics", JSME International Journal (C), 48(4), pp.613-619, 2005.

[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", Proc. of IEEE International Conference on Robotics and Automation, pp. 2985-2991, 2003.

[3] M. Lawitzky, A. Mörtl, and S. Hirche, "Load Sharing in Human-Robot Cooperative Manipulation", Proc. of IEEE International Symposium on Robot and Human Interactive Communication, pp. 185-191, 2010.

[4] B. Corteville, E. Aertbelien, H. Bruyninckx, J. D. Schutter, and H. V. Brussel, "Human-inspired Robot Assistant for Fast Point-to-Point Movements", Proc. of IEEE International Conference on Robotics and Automation, pp. 3639-3644, 2007.

[5] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching Physical Collaborative Tasks: Object-Lifting Case Study with a Humanoid", Proc. of 9th IEEE-RAS International Conference on Humanoid Robots, pp. 399-404, 2009.

[6] A. Thobbi, Y. Gu, and W. Sheng, "Using Human Motion Estimation for Human-Robot Cooperative Manipulation", Proc. of IEEE/RSJ International Conference on Intelligent Robotics and Systems, pp. 2873-2878, 2011.

[7] J. Medina, M. Lawitzky, A. Mörtl, D. Lee, and S. Hirche, "An Experience-Driven Robotic Assistant Acquiring Human Knowledge to Improve Haptic Cooperation", Proc. of IEEE/RSJ International Conference on Intelligent Robotics and Systems, pp. 2416-2422, 2011.

[8] J. Medina, D. Lee, and S. Hirche, "Risk Sensitive Optimal Feedback Control for Haptic Assistance", Proc. of IEEE International Conference on Robotics and Automation, pp. 1025-1031, 2012.

[9] J. Medina, T. Lorenz, D. Lee, and S. Hirche, "Disagreement-Aware Physical Assistance Through Risk-Sensitive Optimal Feedback Control", Proc. of IEEE/RSJ International Conference on Intelligent Robotics and Systems, 2012.

[10] P. Whittle, "Risk-sensitive linear/quadratic/gaussian control", Advances in Applied Probability, vol. 13, no. 4, pp. 764-777, 1981.

[11] 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, 1973.

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

[13] D. Lee and C. Ott, "Incremental kinesthetic teaching of motion primitives using the motion refinement tube", Autonomous Robots, vol. 31, no. 2-3, pp. 115-131, 2011.

[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] John F. Nash, Jr., "The Bargaining Problem", Econometrica, vol. 18, no. 2, pp. 155-162, 1950.

[16] K. Brown and Y. Shoham, "Essentials of Game Theory: A Concise, Multidisciplinary Introduction", Morgan and Claypool Publishers, San Rafael, CA, 2008.