2013 IEEE International Conference on Robotics and Automation (ICRA)
Karlsruhe, Germany, May 6-10, 2013

Risk-sensitive Interaction Control in Uncertain Manipulation Tasks

José Ramón Medina, Dominik Sieber and Sandra Hirche

Abstract— **Manipulation tasks are a great challenge for robots due to the uncertainty arising from unstructured environments. In this paper we propose a novel control scheme for contact tasks based on risk-sensitive optimal feedback control. It provides a systematic approach to adjust the trade-off between motion and force control under uncertainty. Following a previously acquired task model, the proposed approach provides both a variable stiffness solution and a motion reference adaptation. This control scheme achieves increased adaptability under previously unseen environmental variability. An implementation on a robotic manipulator validates the applicability and adaptability of the proposed control approach in two different manipulation tasks.**

I. Introduction

The field of robotic manipulation in unstructured environments imposes interesting challenges on several levels. The acquisition of a task plan for complex manipulation task is difficult. Furthermore, the execution of such task plans, which typically imply both, position/velocity and force goals, require online adaptability due to potential variability of real environments. The adjustment of such motion and force goals depending on the environment situation as well as the adaptation of the manipulator compliance to unexpected situations are crucial for the successful execution of manipulation tasks in changing environments and is the focus of this paper.

Robotic manipulation covers a wide field of applications, such as grasping or door opening tasks [1]. Due to the complexity of analytical task representation, a recurrent choice in recent literature for this topic is the extraction of a task plan based on exemplary human demonstrations. The programming by demonstration paradigm [2] provides efficient method for rapid skill transfer to robots, applicable to manipulation tasks through teleoperation [3] or physical coaching [4]. Manipulators can be easily programmed following this principle extracting motion-based generalized plans [5]. Motion-based approaches successfully execute manipulation tasks if reproduced in the same environment [6], but fail in uncertain environments. The inclusion of contact forces into the control scheme facilitates a desired compliant interaction of the manipulator and thereby improves its behavior in contact [7]. Task plans based on demonstrations including contact forces provide better generalization capabilities than pure position models [8]. However, such an approach implies the challenge of controlling two different references at the same time, i.e. position and force. Due to environmental variability, additional adaptation of the controlled behavior is required during the reproduction phase. One option is the adaptation of the position reference depending on the observed force error using force feedback [9], while other methods follow hybrid control [10] or parallel force and position control [11]. A valid alternative relies on varying the stiffness of the robot depending on the observed position error during kinesthetic teaching [12]. However, a systematic approach to adapt both, the compliance and the reference of the robotic manipulator based on observed force errors does not exist to the best of the authors knowledge.

In this paper we will explicitly address the problem of environment uncertainty during the reproduction of a manipulation task, i.e. we assume a mismatch between the planned position and force trajectories and the current environment condition. We propose a novel adaptive motion control scheme based on risk-sensitive optimal feedback control to adapt the trade-off between motion and force control based on the uncertainty during task reproduction. Risk-sensitive control [13] is successfully applied in our earlier works on physical human-robot interaction [14], [15], where the robot control needs to adapt to unmodeled and unexpected human behavior. In this paper we show that this concept can be similarly applied to adapt the trade-off between motion and force control based on the uncertainty during task reproduction. We model the manipulation tasks as a cooperation problem for two cooperating agents given by a motion and a force controller. Modeling the unexpected behavior of the force controller as a process noise input for the motion control and following a Model Predictive Control Scheme (MPC), an online variable stiffness as well as a reference adaptation is achieved by performing a risk-sensitive optimization. The proposed control scheme provides high adaptability in different environmental situations as validated in experiments for an implementation on a KUKA lightweight robot performing two different tasks.

The remainder of this paper is structured as follows. Section II describes formally the problem considered in this work. The proposed controller is presented in Section III.

Notation: Bold symbols denote vectors. A multivariate normal distribution centered at $\boldsymbol{u}$ with covariance matrix Σ is denoted $\mathcal{N}(\boldsymbol{u}, \boldsymbol{\Sigma})$. A trajectory of length T of normal distributions, i.e. $\{\mathcal{N}(\boldsymbol{u}_0, \boldsymbol{\Sigma}_0), \mathcal{N}(\boldsymbol{u}_1, \boldsymbol{\Sigma}_1), ..., \mathcal{N}(\boldsymbol{u}_{T-1}, \boldsymbol{\Sigma}_{T-1})\}$ is abbreviated as $\{\boldsymbol{u}, \Sigma\}$.

II. Problem Formulation

The task considered in this paper consist of a robotic manipulator moving from an initial configuration to a final goal following a trajectory of desired position/velocity and desired force. This desired trajectory describes a manipulation task

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

978-1-4673-5643-5/13/$31.00 ©2013 IEEE

including contact with the environment. We assume that the manipulator measures contact forces at its end-effector. Its movement can be represented directly in joint or task space. For simplicity we consider that the robot's motion is represented in task space and is governed by an admittance-type control law given by

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

where $\boldsymbol{x}$ is the position of the end-effector and the admittance parameters are given by the rendered mass $\boldsymbol{M}_r$ and friction $\boldsymbol{D}_r$; $\boldsymbol{u_p}$ represents the motion control input which corrects the error in position and velocity space and $\boldsymbol{u_f}$ the force control input that accounts for errors in force space. We denote the robot's motion state by $\boldsymbol{\xi} = (\boldsymbol{x}\ \dot{\boldsymbol{x}})^T$. The right side of (1) reflects the coupling between two controllers, which can be interpreted as a cooperation only if a common final goal is assumed , i.e. both, the desired force and the desired motion trajectory can be tracked without the motion and force control counteracting each other. This is achievable only if the task is executed in the same environment as during demonstrations. Otherwise, and this is the more general case that includes environment variability from demonstration to reproduction, it is interpreted as a competition.

We assume that the desired motion/force trajectories are encoded in a a task model λ, which is extracted from previous human demonstrations and is represented as a trajectory of normal distributions of position/velocity $\{\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\}$ and force $\{\hat{\boldsymbol{\mu}}_{\boldsymbol{f}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{f}}\}$. We also consider the possibility of environmental variability during the reproduction of the task, which might not be reflected in the task plan.

A control scheme that reproduces a learned task given by λ would minimize the tracking error with respect to the expected trajectories of both position/velocity $\boldsymbol{\xi_d} = (\boldsymbol{x_d}\ \dot{\boldsymbol{x}}_{\boldsymbol{d}})^T$, and force, $\boldsymbol{f_d}$. From (1) we can write

$$\boldsymbol{M}_r\ddot{\boldsymbol{x}} + \boldsymbol{D}_r\dot{\boldsymbol{x}} = (\boldsymbol{D}(\dot{\boldsymbol{x}}_{\boldsymbol{d}} - \dot{\boldsymbol{x}}) + \boldsymbol{K}(\boldsymbol{x_d} - \boldsymbol{x})) + K_f(\boldsymbol{f_d} - \boldsymbol{f}) \ , \tag{2}$$

where $\boldsymbol{f}$ is the measured force at the end-effector and $\boldsymbol{x_d}$ and $\dot{\boldsymbol{x}}_{\boldsymbol{d}}$ are the desired position and velocity respectively; $\boldsymbol{K}$, $\boldsymbol{D}$ are the stiffness and damping constituting the motion control and K_f is the proportional gain for the force control.

In order to reproduce the expected forces at the end-effector, the force control scheme tracks the means of the expected force trajectory, i.e., $\boldsymbol{f_d} = \hat{\boldsymbol{\mu}}_{\boldsymbol{f}}$. Note that when K_f differs from the identity matrix, its effect is equivalent to modifying the rest of the admittance parameters as

$$\boldsymbol{M}'_r = \boldsymbol{M}_r K_f^{-1} \quad \boldsymbol{D}'_r = \boldsymbol{D}_r K_f^{-1}$$
$$\boldsymbol{D}' = \boldsymbol{D} K_f^{-1} \quad \boldsymbol{K}' = \boldsymbol{K} K_f^{-1}.$$

For simplicity, we assume K_f as the identity matrix.

Given this problem setting, the focus of this paper lies on the design of the motion control scheme that generates $\boldsymbol{u_p}$, i.e. on the design of the possibly time-varying matrices $\boldsymbol{K}$ and $\boldsymbol{D}$ and the position/velocity reference $\boldsymbol{\xi_d}$.

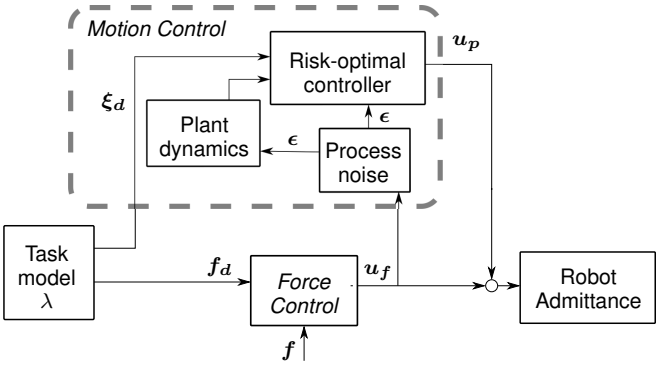

Fig. 1. General control scheme

III. Risk-Sensitive Position Control

Manipulation tasks usually have different phases depending on the existence of contact with the environment. During pure unconstrained motion phases no contact forces are present and tracking the desired motion trajectory is the only control goal. However, when forces must be exerted in contact with the environment, their role is typically critical for the task success and need to be prioritized over the motion trajectory following. Let us consider the grasping of an object as an exemplary task: during the grasping phase forces are crucial and tracking the desired motion becomes a secondary goal. If the object to be grasped is placed in a slightly different position than the expected one (assuming no visual feedback is available), the only way to successfully perform the grasping task is to rely on force feedback, making sure that the motion tracking does not disturb.

From a control point of view, a model-based motion control scheme can interpret the activity of the force control as an error in its task model. If the motion-based model was ideal, i.e. the contact location and the environment impedance exactly the same as during the recording of the task, there would be no need for a force controller (except for contact with rigid environments). However, force tracking is necessary when such an exact model is not available. The correction coming from the desired force tracking is potentially produced by environmental variability/uncertainty in contact location and/or changed environment impedance during the reproduction of the task. Within the field of learning the proper reproduction of the manipulation task despite uncertainty relates to the desired generalization property, i.e. executing a learned task in environmental situations differing from the learned ones. In such cases, the force controller corrections can not be computed in advance. In consequence, from (1), we model the corresponding force input as process noise $\boldsymbol{u_f} = \epsilon$ in the plant dynamics of the motion controller

$$\boldsymbol{M}_r\ddot{\boldsymbol{x}} + \boldsymbol{D}_r\dot{\boldsymbol{x}} = \boldsymbol{u_p} + \epsilon \ . \tag{3}$$

We assume that the process noise ϵ is normally distributed, i.e. $\epsilon = \mathcal{N}(\boldsymbol{u_0}, \boldsymbol{\Sigma_u})$. It represents the *model error* of the motion task model. An schematic overview of the control scheme is depicted in Fig. 1.

Due to the discrete time nature of the implementation, we discretize the system from (3) with a sampling time

interval Δt yielding a discretized plant dynamics in the form $\boldsymbol{\xi}_{k+1} = A\boldsymbol{\xi}_k + B\boldsymbol{u}_{\boldsymbol{p}_k}$ given by

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

where $\boldsymbol{x}_k$, $\boldsymbol{v}_k$ and $\boldsymbol{u}_{\boldsymbol{p}_k} \in \mathbb{R}^3$ are the discrete time position, velocity and motion control input at time k in Cartesian space. The model error or process noise ϵ_k is constantly updated, depending on the observed force input $\boldsymbol{u}_{\boldsymbol{f}}$ from (1). The process noise characteristics are calculated online as the first and second order moments of $\boldsymbol{u}_{\boldsymbol{f}}$ over the past W observed samples, leading to the normal distribution given by $\hat{\epsilon}_k = \mathcal{N}(\boldsymbol{u}_{\boldsymbol{0}k}, \boldsymbol{\Sigma}_{\boldsymbol{u}k})$. [1]

Given the plant dynamics from (4) and a desired trajectory to follow $\{\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\}$, the aim of the controller is to generate the corresponding control input that tracks the given trajectory considering the observed model error.

A. Reference Adaptation

The motion controller tracks the trajectory given by $\{\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\}$. The desired reference $\boldsymbol{\xi}_{\boldsymbol{d}}$ is then given by the mean of the demonstrated trajectories

$$\boldsymbol{\xi}_{\boldsymbol{d}k} = \hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}_k}. \tag{5}$$

However, depending on the process noise model, the desired motion reference can be adapted in order to accommodate for diverging forces.

Depending on the interpretation of the influence of the force control on the motion control, the process noise is modeled as *biased* or *unbiased*. If we assume that observed deviations produced by the force control do not lead to a hypothetical goal divergence, we can optimistically consider cooperation and therefore an unbiased noise model $\mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$. In contrast, if the possibility of divergence is assumed, the noise model becomes biased, i.e. $\mathcal{N}(\boldsymbol{u}_{\boldsymbol{0}}, \boldsymbol{\Sigma}_{\boldsymbol{u}})$. For the latter case, the dynamics from (4) are expressed as

$$\boldsymbol{\xi}_{k+1} = A\boldsymbol{\xi}_k + B(\boldsymbol{u}_{\boldsymbol{p}k} + \boldsymbol{u}_{\boldsymbol{0}k} + \epsilon_k)\ . \tag{6}$$

with unbiased process noise $\epsilon_k = \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}k})$. An additive reference adaptation $\boldsymbol{\xi}_{\boldsymbol{a}}$ is calculated as the motion difference produced by the bias $\boldsymbol{u}_{\boldsymbol{0}}$ applying the system dynamics

$$\boldsymbol{\xi}_{\boldsymbol{a}k+1} = A\boldsymbol{\xi}_{\boldsymbol{a}k} + B\boldsymbol{u}_{\boldsymbol{0}k}$$

If we then add it to the desired tracking reference from (5), the resulting desired trajectory is adapted to the observed force controller divergence as

$$\boldsymbol{\xi}_{\boldsymbol{d}k} = \hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}_k} + \boldsymbol{\xi}_{\boldsymbol{a}k}\ . \tag{7}$$

With this modeling, the dynamics from (6) take again the same form as in (4) with unbiased process noise, as the bias's effect is already modeled adapting the reference.

[1]Note that the process noise can also be modeled accounting for the *expected model error* captured by the task model λ and given by a zero mean normal distribution with covariance $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{f}}$ [15].

B. Risk-Sensitive Optimization

With the dynamics from (4) and the desired reference (7) the computation of the motion control input $\boldsymbol{u}_{\boldsymbol{p}}$ can be formulated from an optimality point of view as the minimization of the distance to the desired trajectory. Due to the continuous reestimation of the process noise present in the plant dynamics, an MPC scheme must be adopted and the optimization must be constantly recalculated as the problem parameters change. A quadratic cost function at sample time k for this problem takes the form

$$J_k = \sum_{i=k}^{k+T} \|(\boldsymbol{\xi}_{\boldsymbol{d}i} - \boldsymbol{\xi}_i)\|_Q^2 + \|\boldsymbol{u}_{\boldsymbol{p}i}\|_R^2\ , \tag{8}$$

where T is the time horizon, $\|x\|_Q^2$ stands for the quadratic form $x^{\mathrm{T}}Qx$ and Q and R are weighting factors that allow a trade-off between control cost and tracking error minimization. Note that the weighting factor Q can be also chosen proportional to the inverse of the expected motion covariance $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}$ in order to account for potential task constraints encoded into the variability between the learned trials [15].

The minimization of the expectation $\mathbb{E}[J_k]$ of the cost function for the dynamics (4) leads to a feedback solution that provides optimal tracking. However, the influence of the process noise ϵ is ignored, i.e. the variability of the force error does not influence the motion control. In contrast, a risk-sensitive controller directly considers the process noise in the dynamics, adapting the manipulator compliance $\boldsymbol{K}$, $\boldsymbol{D}$ depending on a risk-sensitivity parameter θ. In this case the cost function takes the form

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

The exponential form of this cost function makes the optimization sensitive not only to the expected cost $\mathbb{E}[J_k]$ but also to higher order moments of it, which are directly related to the process noise. If $\theta = 0$ the controller is risk-neutral and the process noise has no influence on the resulting gain. For $\theta < 0$ and $\theta > 0$ the controller becomes risk-averse and risk-seeking, respectively.

Solving this optimization problem with linear dynamics leads to a an optimal feedback control law in the form

$$\boldsymbol{u}_{\boldsymbol{p}k} = -V_i(\boldsymbol{\xi}_{\boldsymbol{d}i} - \boldsymbol{\xi}_i)\ , \tag{9}$$

where V_i is the feedback matrix given by a modified form of the Ricatti recursion [16]

$$V_i = -R^{-1}B'(BR^{-1}B' + \theta\boldsymbol{\Sigma}_{\boldsymbol{u}k} + \Pi_{i+1}^{-1})^{-1}A\ ,$$

and

$$\Pi_i = Q_i + A'(BR^{-1}B' + \theta\boldsymbol{\Sigma}_{\boldsymbol{u}k} + \Pi_{i+1}^{-1})^{-1}A\ ,$$

with $\Pi_T = Q_T$. The resulting feedback gain V_i represents both the damping $\boldsymbol{D}$ and stiffness $\boldsymbol{K}$ parameters from (2) as

$$V_i = \begin{pmatrix}0 & 0\\ \boldsymbol{K}_i & \boldsymbol{D}_i\end{pmatrix}.$$

Note that, due to the MPC scheme, while the recursion for the optimization is calculated for a time horizon T, the

only feedback matrix applied in the motion control from (9) is the one for the simulated step $i = k$.

As a result, 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. Applied to our scenario, a risk-seeking policy arises as the most intuitive solution. In this case, the stiffness decreases when the force control needs to correct the desired force trajectory in order to avoid disturbances. The risk-averse case is a suitable option for pure motion tasks as the stiffness rises under force control corrections, increasing the tracking precision. Note that while the presented solution only considers linear dynamics, the approach is straightforwardly extendable to nonlinear dynamics using iterative optimization methods [17].

IV. Experiments

In order to evaluate the performance and demonstrate the applicability of the proposed control scheme, a full-scale experiment is conducted in our laboratory. Using a robotic manipulator two different tasks involving contact with the environment are performed for four different motion control strategies. The presented results show the advantages of the proposed approach testing the adaptation capabilities to different environmental variations.

A. Experimental Setup

The robotic platform employed for this experiment is shown in Fig. 2; it consists of a four-wheeled omni-directional mobile platform (here only used for repositioning the robot) and two identical commercially available KUKA LWR (light-weight robot) 4+ manipulators. Only one of the two manipulators is used in all the experiments driven in the Cartesian impedance control mode. For measuring resultant forces independent of the configuration and of the human guiding force during kinesthetic teaching of the manipulation task, a force/torque sensor (*JR3*) is attached to the wrist. As end-effector, a Schunk PG70 two-finger parallel gripper is used for simple grasping and object pushing with the fingers. A detailed description of this robot can be found in [18].

In order to operate the KUKA LWRs, *Fast Research Interface* (FRI), which allows user control and status monitoring of the manipulators based on UDP protocol is used. This software package is integrated in our real-time robot control framework [19]. The robot control loop runs at a frequency of 1 kHz. For evaluating the proposed controllers, rotational motions are ignored so that the system dynamics remain decoupled linear dynamics for each Cartesian DoF.

In order to acquire a generalized model of a task given a set of exemplary demonstrations, a probabilistic task model λ given by a time-based HMM is trained using the EM algorithm. In a similar way as trajectory HMMs, time-based HMMs can generalize the given observations as a trajectory of expected normal distributions. However, in contrast to trajectory HMMs, time-based HMMs preserve the standard HMMs flexibility of representing sequences with different lengths due to their discretized state space. This property avoids the necessity of aligning observations due to their different lengths using time warping methods like dynamic time warping. The generalized expected Gaussians trajectory of both position/velocity $\{\hat{\boldsymbol{\mu}}_{\boldsymbol{\xi}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}}\}$ and forces $\{\hat{\boldsymbol{\mu}}_{\boldsymbol{f}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{f}}\}$ is calculated using Gaussian Mixture Regression over the time domain. See [4] for a detailed description of the method.

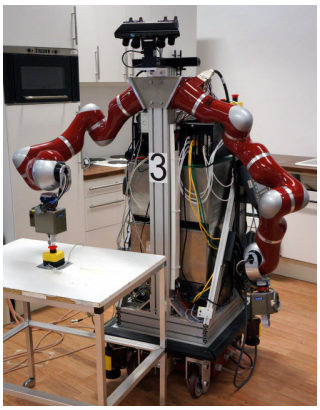

Fig. 2. Task consisting of pushing a button followed by pointing out text written on a whiteboard.

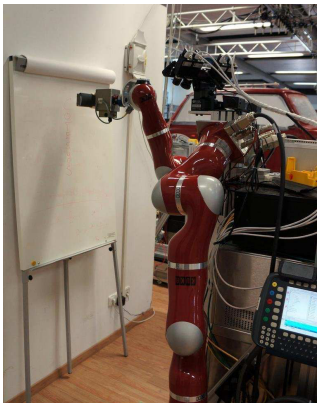

Fig. 3. Task consisting of cleaning a whiteboard with an eraser grasped with a two-finger gripper.

Task model λ	Pushing Button and Pointing			Erasing Whiteboard	
Controller	Initial Distance	Button	Pointing Distance	Initial Distance	Whiteboard
(a)	0 cm	○	0 cm	-5 cm	○
	5 cm	×	0 cm	0 cm	●
	10 cm	×	0 cm	+5 cm	●
(b)	0 cm	○	0 cm	-5 cm	●
	5 cm	○	0 cm	0 cm	●
	10 cm	○	0 cm	+5 cm	●
(c)	0 cm	●	8.5 cm	-5 cm	○
	5 cm	●	9.5 cm	0 cm	○
	10 cm	●	7.8 cm	+5 cm	○
(d)	0 cm	●	8.5 cm	-5 cm	○
	5 cm	●	9.5 cm	0 cm	○
	10 cm	●	7.8 cm	+5 cm	○

TABLE I

Experimental results for pushing button and pointing task and for whiteboard erasing task.

● success, ○ partial success, × failure

For evaluating combinations of risk-sensitivity and motion reference, we tested four different motion controllers:

(a) Unbiased process noise $\epsilon = \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ with risk-neutral optimization, i.e. $\theta = 0$ and motion reference given by (5).

(b) Unbiased process noise $\epsilon = \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ with risk-seeking optimization, i.e. $\theta = 10^{-5}$ and motion reference given by (5).

(c) Biased process noise $\epsilon = \mathcal{N}(\boldsymbol{u_0}, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ with risk-neutral optimization, i.e. $\theta = 0$ and motion reference given by (7).

(d) Biased process noise $\epsilon = \mathcal{N}(\boldsymbol{u_0}, \boldsymbol{\Sigma_u})$) with risk-seeking optimization, i.e. $\theta = 10^{-5}$ and motion reference given by (7).

Note that controllers a) and c) are not influenced by the variability of the force control $\boldsymbol{\Sigma_u}$ in contrast to b) and d), that adopt a risk-seeking optimization. Similarly, controllers b) and d) adapt their motion reference depending on the bias $\boldsymbol{u_0}$ while a) and c) do not modify the learned references.

For our experiments, the parameters of Eq. (1) are

$$M_r = \text{ diag } (10 \text{ kg}, 10 \text{ kg}, 10 \text{ kg })$$
$$D_r = \text{ diag } (80 \text{ Ns/m}, 80 \text{ Ns/m}, 80 \text{ Ns/m })$$

and the window size W for the process noise calculation is chosen in order to capture the last 0.3s. For simplicity of presentation, the tracking error precision Q in (8) is chosen such as errors with respect to the expected velocity trajectory are ignored and therefore only position errors are considered resulting in a stiffness adaptation, i.e. $\boldsymbol{D} = 0$ in (2).

B. Experimental Design

Two tasks are considered for the reproduction performance evaluation. A model of each task is acquired after teaching 12 different demonstrations of each task keeping the same environmental conditions.

The first task, shown in Fig. 2, consists of pushing a button followed by pointing out text written on a whiteboard. Three different environmental conditions are tested. The button height is decreased with respect to the teaching environmental conditions for 0, 5 and 10 cm respectively. A task execution is considered successful when the button is correctly pushed and the predefined place at the board is exactly pointed out. The button pressing procedure is considered a success when the button is correctly pushed and clicked, a half success when contact is partially pushed but did not clicked and other cases are considered as failure.

The second task, shown in Fig. 3, consist of cleaning a whiteboard with an eraser which is assumed to be already grasped. Three different environmental conditions are tested, including an initial end-effector position differing in -5, 0 and 5 cm to the teaching environmental conditions. The task is considered successfully executed when the text written on the whiteboard is erased. A half success is considered when the text is only partially erased.

C. Results: Pushing and pointing out

As shown in Fig. 4 for the task reproduction placed 10 cm lower with respect to the initial conditions, the resulting stiffness depicted by the solid blue line is adapted depending on the force control input represented by the black dashed line for the risk-seeking controllers, i.e. (b). The variable stiffness profile from (a) is only a result from the optimization and are not affected by the force controller activity. As no reference adaptation is considered in controller (b), the stiffness decreases more with respect to (a) as it needs a stronger correction from the force control. As a result of the adaptive stiffness with respect to force errors produced by the risk-seeking optimization, the force controller receives less disturbances from the motion controller and can achieve a better force tracking performance which translates in a better task performance: as shown in Table I, the risk-seeking controller (b) performs better than the risk-neutral one (a) when pushing the button.

The reference adaptation for controller (c) and (d) for the 5 cm case is shown in Fig. 5. The adapted reference depicted by the dashed-dotted red line deviates from the task model's desired trajectory represented by the solid green line accommodating for the force controller input represented by the black dashed line. As a consequence, a better task performance for the button pressing phase is achieved as shown in Table I for controllers (c) and (d). However, a drift from the original reference is present at the end of the task execution producing an undesired tracking error effect for the pointing out motion. This effect is not present in the other two controllers that do not adapt their references.

D. Results: Erasing

The results for the erasing task are shown in Table I. For the two controllers not adapting their motion reference, the risk-seeking controller (b) arises again as a better alternative as it correctly erases the whiteboard in all conditions, in contrast to the risk-neutral case (a) that ignores the force control input. While this task suggests that an adaptation of the reference of the motion controller is beneficial if an environmental offset is present, the results show worse performance for the controllers adapting their reference, i.e. (c) and (d). As shown by Fig. 5, the adapted reference for controller (c) depicted by the dashed-dotted red line overcompensates the force controller contribution. This undesired effect may be caused by a too fast accommodation of the observed bias, suggesting that an increased window size for process noise estimation could improve the controller performance by capturing only very clear bias trends.

In summary, the risk-seeking optimization arises a suitable option as it provides variable stiffness that adapts to force corrections providing better task reproduction generalization for unknown environmental variability. Similarly, the reference adaptation also improves the task success performance accommodating diverging forces by adapting the motion reference. Depending on the task, the resulting drift of the reference trajectory w.r.t the desired one increases or decreases the generalization capabilities of the control scheme.

V. Conclusions

In this paper we propose a novel control approach for manipulation tasks based on risk-sensitive optimal feedback control. A variable compliance is achieved by modeling force corrections as process noise in the plant dynamics for the adaptive motion controller and minimizing a risk-sensitive cost function. An additional motion adaptation from the initial planned reference results from the force controller divergence. The benefits of the approach in terms of high adaptability to uncertain and variable environments are convincingly validated in experiments with a robotic manipulator, where position reference adaptation and stiffness adaptation are performed.

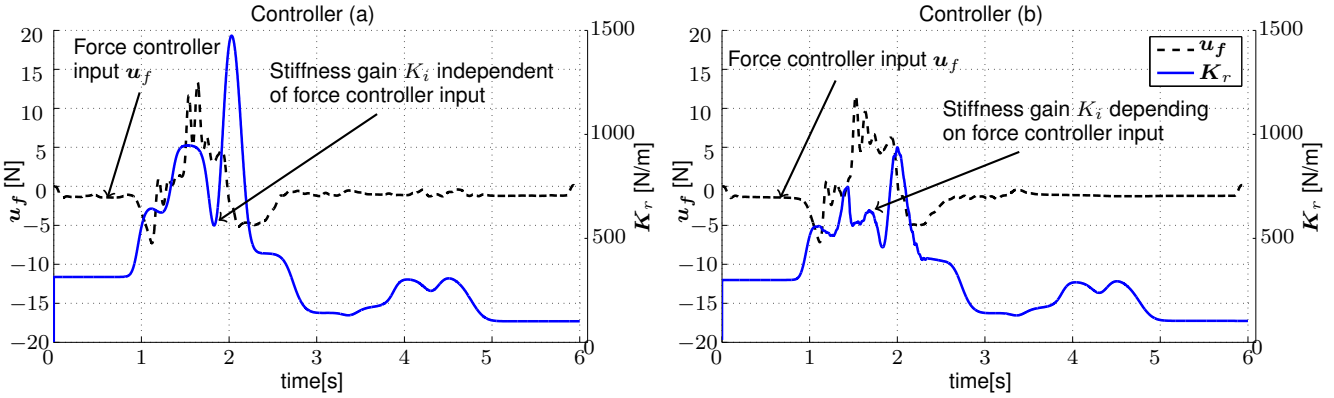

Fig. 4. Control gains independent of the force error (a) and dependent of the force error (c) together with the force controller input during the button pushing and pointing out task for 10cm.

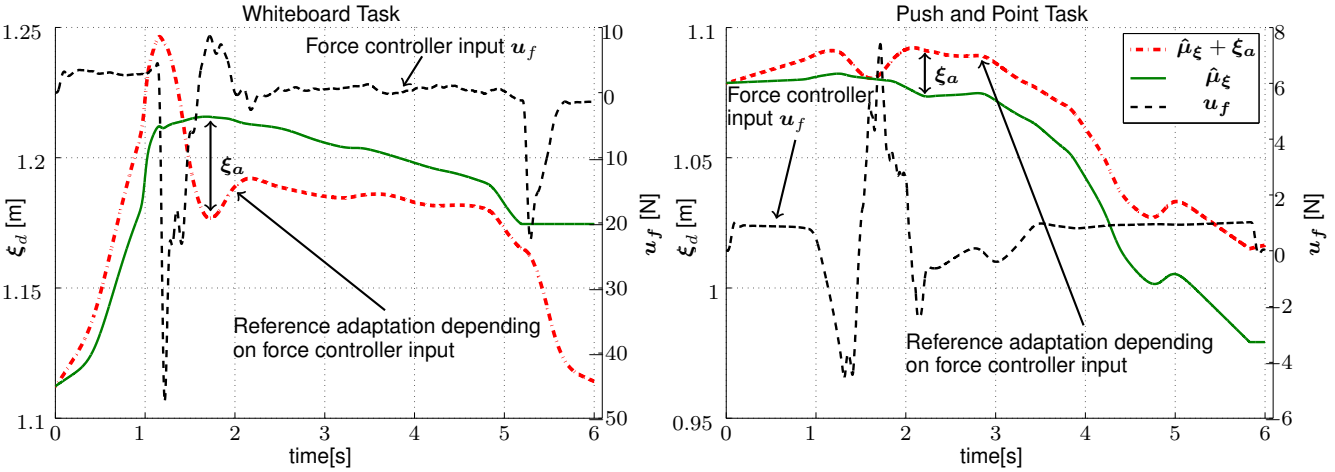

Fig. 5. Force control input, desired trajectory and adapted desired trajectory with motion reference adaptation depending on force controller input (c) for both tasks for 5cm initial distance.

A 6-dimensional solution, studying the trade-off between variable stiffness and reference adaptation and correcting the adapted reference are the matter of our future work.

ACKNOWLEDGMENTS

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

REFERENCES

[1] A. Jain and C. Kemp, "Pulling open doors and drawers: Coordinating an omni-directional base and a compliant arm with equilibrium point control," in *Proc. IEEE ICRA*, 2010, pp. 1807 –1814.

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

[3] M. Rambow, S. Hirche, and M. Buss, "Autonomous manipulation of deformable objects based on teleoperated demonstrations," in *Proc. IEEE/RSJ IROS*, 2012.

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

[5] E. Gribovskaya, S. Khansari-Zadeh, and A. Billard, "Learning Non-linear Multivariate Dynamics of Motion in Robotic Manipulators," *Int. J. Rob. Res.*, 2010.

[6] A. Saxena, J. Driemeyer, and A. Ng, "Robotic grasping of novel objects using vision," *Int. J. Robot. Res.*, vol. 27, no. 2, pp. 157–173, 2008.

[7] N. Hogan, "Impedance control: An approach to manipulation," in *Proc. ACC*, 1984, pp. 304 –313.

[8] P. Pastor, M. Kalakrishnan, S. Chitta, E. Theodorou, and S. Schaal, "Skill learning and task outcome prediction for manipulation," in *Proc. IEEE ICRA*, 2011, pp. 3828 –3834.

[9] P. Pastor, L. Righetti, M. Kalakrishnan, and S. Schaal, "Online movement adaptation based on previous sensor experiences," in *Proc. IEEE/RSJ IROS*, 2011, pp. 365 –371.

[10] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 103, no. 2, pp. 126–133, 1981.

[11] S. Chiaverini and L. Sciavicco, "The parallel approach to force/position control of robotic manipulators," *Robotics and Automation, IEEE Transactions on*, vol. 9, no. 4, pp. 361 –373, aug 1993.

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

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

[14] 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.

[15] J. Medina, D. Lee, and S. Hirche, "Risk Sensitive Optimal Feedback Control for Haptic Assistance," 2012, pp. 1025–1031.

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

[17] E. Todorov and W. Li, "A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems," in *Proc. ACC*, 2005, pp. 300–306.

[18] M. Lawitzky, J. Hernandez, and S. Hirche, "Rapid prototyping of planning, learning and control in physical human-robot interaction," in *Proc. ISER*, 2012, pp. 819–824.

[19] D. Althoff, O. Kourakos, M. Lawitzky, A. Mörtl, M. Rambow, F. Rohrmüller, D. Brščić, D. Wollherr, S. Hirche, and M. Buss, "An architecture for real-time control in multi-robot systems," *Human Centered Robot Systems*, pp. 43–52, 2009.