Online Learning of Varying Stiffness Through Physical Human-Robot Interaction

Klas Kronander and Aude Billard

Abstract— **Programming by Demonstration offers an intuitive framework for teaching robots how to perform various tasks without having to preprogram them. It also offers an intuitive way to provide corrections and refine teaching during task execution. Previously, mostly position constraints have been taken into account when teaching tasks from demonstrations. In this work, we tackle the problem of teaching tasks that require or can benefit from varying stiffness. This extension is not trivial, as the teacher needs to have a way of communicating to the robot what stiffness it should use. We propose a method by which the teacher can modulate the stiffness of the robot in any direction through physical interaction. The system is incremental and works online, so that the teacher can instantly feel how the robot learns from the interaction. We validate the proposed approach on two experiments on a 7-Dof Barrett WAM arm.**

I. INTRODUCTION

In order for robots to help humans in everyday task they must be equipped with interfaces that allow their operators to teach them useful tasks without knowledge of robot programming languages. The Programming by Demonstration (PbD) paradigm aims at endowing robots with the capability to learn tasks from demonstrations. Traditionally, PbD has provided means of learning kinematic aspects of a task. It then relied on a stiff position controller to execute the tasks. While many tasks can and have been taught this way, there are tasks which can benefit from or even require control of dynamic interaction with the environment. As an example of simple task belonging to this category, consider the task of transporting a bottle full of liquid towards a glass, and then pouring the contents of the bottle into the glass. Such a task can be performed using only kinematic constraints. However, if sudden perturbations (i.e. a person pushing the robot while reaching for the glass) can be expected, a stiff controller will respond with high forces, which can cause the liquid to spill. It would thus be desirable to control the way that the robot responds to positional perturbations. This can be achieved using impedance control [1], where the objective is to control the mechanical impedance of the robot, i.e. the dynamic relation between positional perturbations and restoring forces. Mechanical impedance is usually specified as a first or second order differential equation, with the objective of making the robot's end effector behave as a mass-spring-damper system[1]. By letting the impedance parameters vary, different behaviors can be achieved to accommodate the requirements in different stages of the task. Thus, the impedance parameters can be made a part of the task constraints. In this framework, a task is specified by a kinematic profile and an impedance profile. In this work we propose a system that allows a teacher to teach varying stiffness profiles for such tasks. The focus is on learning varying stiffness, and the proposed system makes no assumption as to how the kinematic profile is learned. Therefore, the availability of a learned kinematic profile will be assumed throughout this paper.

K. Kronander and A. Billard are with the Learning Algorithms and Systems Laboratory (LASA), School of Engineering, Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland

[1]Using a virtual, specified inertia in the case of a second order differential equation, and the intrinsic inertia of the robot in the case of a first order differential equation.

Fig. 1. The figure shows a human teacher physically interacting with the robot to make it less stiff. This is done by perturbing it around the current equilibrium point.

As the motivation of PbD is to make it easy for users without knowledge of programming to teach tasks to their robots, any interface which is used in the teaching process for PbD should be intuitive. With this in mind, we developed a teaching interface for variable stiffness inspired by the way humans convey such information between each other. In dance and other sports, when the teacher wishes to convey to the student that she should relax, she may say 'relax' and at the same time wiggle the limb that is too stiff. Our PbD approach to teaching varying stiffness is based on the same idea, see Fig. 1. The robot is initially set to move along a desired trajectory with a high stiffness. The teacher intervenes along the trajectory to decrease stiffness when needed. This is done by wiggling the robot's end effector around its equilibrium position, see Fig. 1. The robot learns to become more compliant by computing the difference between the desired and current end-effector position. The larger the amplitude of the wiggling, the more compliant the robot will become.

Online update of the stiffness provides the teacher with a direct haptic feedback and allows her to feel the effect of her interaction. The teacher can then evaluate the extent to which her teaching was successful and decide whether to stop or continue teaching. The wiggling motion can only be used to decrease the stiffness of the robot. However, if the stiffness is decreased to much accidentally, this can be compensated through a second round of teaching during which smaller perturbations are imposed. This is possible because the system incrementally[2] adapts the stiffness according to the provided perturbations.

In the following section, we present an overview of related research. Following that, in Section III we present the problem statement. This is followed by a detailed description of the learning system and the online adaptation of the stiffness in Section IV. Then, we present a validation of the approach on a 7-dof Barrett WAM in a via-point trajectory following task and a pouring task in Section V. We conclude the paper with a discussion of the results and an outlook into future directions of research in Section VI.

II. Related Work

Learning variable impedance control policies has been formulated as an optimal control (OC) problem in [2] and [3]. These works specify the task constraints as a cost function and optimize the control actions subject to the dynamics of the robot. This has the advantage that the impedance profile is tailored to each robotic platform. The cost functions used in the optimization typically include a task performance term and an energy term. Thus, the resulting policies are trade-offs between task performance and energy consumption. [3] further uses inverse optimal control to infer a task-based cost-function in order to transfer variable impedance policies between different systems. In [4] analytical solutions to optimal control of variable stiffness for maximizing link velocity is reported. Closely related to optimal control is reinforcement learning (RL), which has been used for learning variable impedance policies in [5]. In [6], an EM-based reinforcement learning algorithm initialized by human demonstrations is presented. In contrast to these approaches, in our work, the robot relies on an expert that indicates what stiffness should be used by physical interaction. This has the advantage that the potentially cumbersome task of specifying the optimality criteria is avoided. However, finding a teacher that is capable of delivering appropriate instruction may be difficult in some situations, e.g. tasks involving high velocity movements. The OC, RL and PbD approaches are hence complementary.

As stated in the introduction, PbD aims at deriving control policies automatically from a (usually small) set of demonstrated data. The demonstrated data contains examples of desirable state-to-action pairs, and may vary in dimensionality depending on the number of degrees of freedom and the sensory modalities considered. For a detailed review of PbD, refer to [7]. A recent trend in PbD tackles the problem of teaching force-based control policies [8], [9]. In this work, we follow a similar approach, in that the robot is implicitly[3] taught the forces that it should exert in response to perturbations. This is in stark contrast to classical approaches in PbD that usually relied on kinematic information only.

PbD relies on interfaces that the teacher can use to provide demonstration data. One common approach for providing demonstrations is to use the robot's own body and sensors when demonstrating. This can be done e.g. via kinesthetic teaching or teleoperation. In robot teleoperation, master-slave systems for motion control, force control and more recently impedance control [10] have been suggested. The latter has to the best of our knowledge not yet been used for providing demonstrations for learning, and presents an interesting area to explore.

Humans can control the impedance of their limbs through co-contraction of agonist-antagonist muscle pairs. An apparatus for measuring human limb stiffness was presented in [11] and used for examining hypothesis regarding human arm movement control. Similar setups have since then been used in a number of experiments, including [12] which reports that humans modulate stiffness to deal with instabilities due to interactions. A bio-inspired algorithm for concurrent tuning of trajectory, force and impedance was presented in [13]. Our approach uses a human teacher to communicates stiffness to the robot, but *does not attempt to imitate or reproduce human impedance*. The strategies used by humans to modulate impedance are interesting and may provide valuable insight for selecting impedance for robots. However, direct imitation of a human stiffness profile on a robot is inappropriate, due to the significant differences in kinematics and dynamics.

In [14], an approach to use variability demonstrated (kinematic) data to infer a suitable varying stiffness is proposed. The general idea of that approach is: the more variations the less stiff the system should be. Stiffness was hence inversely proportional to the variance along the trajectory. This approach is based on the assumption that if a large variability is demonstrated in a part of the movement, then the accuracy of that part is not likely to be crucial for the task [15]. Thus, a low stiffness can be used for reproducing that part of the task. Conversely, if demonstrations were consistent in a part of the movement, then this part should be followed strictly, i.e. with a high stiffness. Our work uses a similar approach for determining the stiffness based on the covariance of spatial data, but differs in that stiffness is learned from data supplied online. This a significant difference, as the teacher can feel the stiffness assumed by the robot through haptic interaction, and adapt the teaching accordingly.

Incremental learning in PbD endows the robot with the capability to improve its task reproduction by receiving instructions from the teacher while it is performing the task. Tactile guidance was used to refine motion policies in [16]. In [17], an incremental approach for adjusting the motion policy encoded in a Hidden Markov Model is presented. A specialized impedance controller with a so-called refinement

[2]The system is incremental in two senses: 1) within each demonstration, as the teacher can increase the variance of the imposed perturbations until the robot responds with the desired compliance and 2) between different sets of demonstrations, as the pertubations perceived during previous teaching rounds are combined with the current demonstration.

[3]The force response to perturbations follow from the stiffness that the robot is taught. Note that we do not treat teaching of task-based contact forces in this work.

tube is presented, which allows for motion control with partial compliance so that the teacher can refine the motion by physical coaching. Note that while [17] uses impedance control in their system, no learning of task-appropriate impedance is involved. Our work is complementary to these works as we propose a system that allows incremental learning of the stiffness profile.

III. Problem Statement

Let $\boldsymbol{F} \in \mathbb{R}^3$ and $\boldsymbol{\iota} \in \mathbb{R}^3$ denote the linear and angular parts respectively of the total wrench on the robot's end-effector. Furthermore let $\boldsymbol{x} \in \mathbb{R}^3$ denote the position of the robot's end-effector, and $\boldsymbol{r} \in \mathbb{R}^3$ its orientation. Let $\boldsymbol{\zeta} \in \mathbb{R}^6$ and $\boldsymbol{\xi} \in \mathbb{R}^6$ denote the wrench and pose resepectively:

$$\boldsymbol{\zeta} = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{\iota} \end{bmatrix}, \quad \boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{r} \end{bmatrix} \tag{1}$$

In impedance control the goal is to control the dynamic relationship between wrench and the deviation $\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi} - \boldsymbol{\xi}^d$ from the virtual desired pose $\boldsymbol{\xi}^d$. In this paper, we the consider an impedance of the form:

$$\boldsymbol{\zeta} = \boldsymbol{\Phi}\tilde{\boldsymbol{\xi}} - \boldsymbol{\Psi}\dot{\tilde{\boldsymbol{\xi}}} \tag{2}$$

The stiffness $\boldsymbol{\Phi}$ and the damping $\boldsymbol{\Psi}$ are the parameters of this impedance. We assume no coupling across position and orientation and write:

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{K} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}^r \end{bmatrix}, \quad \boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}^r \end{bmatrix} \tag{3}$$

where $\boldsymbol{K}$ and $\boldsymbol{K}^r \in \mathbb{R}^{3\times3}$ represent the translational and the rotational part of the stiffness respectively. In this work, we are concerned with the translational part of the stiffness matrix. The rotational part of the stiffness and damping are not learned, and are set to constant values in the presented experiments. The translational part of the damping is set as a function of the translational stiffness.

As mentioned in the introduction, this paper deals with the learning of tasks with varying stiffness. More specifically, tasks considered in this paper are defined by a time-dependent reference trajectory of desired pose and *translational* stiffness, $\{\boldsymbol{x}_t^d, \boldsymbol{K}_t\}_{t=o}^T$ where T is the duration of the task. We assume in this work that the robot already knows the kinematic profile of the task, and tackle the question of how to teach it a matching stiffness profile.

IV. Demonstration and Learning of Varying Stiffness

Unlike position, which can be demonstrated using e.g. kinesthetic teaching and teleoperation, demonstrating stiffness is non-trivial. The reason for this is that it is not a physical quantity per se, but rather a relationship between deviations from the virtual equilibrium point and the end effector wrench (c.f. Eq. (2)).

A. Direct Demonstration Vs Abstract Communication of Intent

In PbD, regression-based learning algorithms are often used to generalize over the demonstration data set. During the develoment of this work, we explored this approach for learning stiffness. The main difficulty was the collection of the demonstration data. The method we used was inspired by the technique used for measuring human impedance [11]. Examining (2), it is clear that if the system is perturbed in one of the eigendirections of the stiffness matrix, the resulting force response will be in the same direction but with reversed sign. Thus, the stiffness can be measured by applying perturbations in the eigendirections and observing the resulting response. To collect measurements of stiffness, the robot executed the kinematic task profile while the teacher followed along the movement, holding the end-effector. Periodically, the robot paused the motion and asked for a stiffness demonstration by applying positional perturbation to itself. The teacher then physically demonstrated the appropriate restoring force. The eigenvectors of the stiffness matrix represent the principal directions of stiffness, and are hence of great importance to the resulting behavior. A limitation of the above mentioned approach to teach stiffness is that only the eigenvalues can be learned, while the structure of the stiffness matrix must be predefined.

To overcome this limitation, we changed approach and developed a system that allows the teacher to modulate the eigenvalues *and* the eigenvectors of the stiffness matrix. As opposed to generalizing across examples of demonstrated stiffness values, the robot learns stiffness by interpreting perceived positional perturbations of its end-effector as described in the introduction.

B. Stiffness Adjustment Based on Interactions

The way that the robot interprets spatial perturbations for adjusting shares similarities with how variability of a demonstration data set is used to set stiffness in [14]. The basic idea is that if the teacher imposes perturbations with high variance in a direction, the robot should reduce its stiffness in that direction. The symmetric and positive definite stiffness matrix is built around the eigenvectors of the covariance matrix of the perturbation data, with stiffness along each eigenvector set to be inversely proportional to the square root of the corresponding eigenvalue of the covariance of the perturbations.

We introduce the notation $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}^d$ for representing a perturbation data point. Let $\boldsymbol{\Xi} = \{\tilde{\boldsymbol{x}}_j, t_j)\}_{j=0}^J$ denote the set of observed perturbations with their corresponding time stamps, where J is the number of provided perturbation data. At time t, a stiffness matrix is assigned based on the data in $\boldsymbol{\Xi}$ with time stamps in the range $[t-S, t]$. Thus, the stiffness assignment is based on a sliding temporal window-view of length S over the observed perturbation data. Let L_t and U_t define the lower and upper bounds for the indices of data points inside the temporal window:

$$L_t = max\{j \in [1, 2 \dots J] : t_j < t - S\} \tag{4a}$$

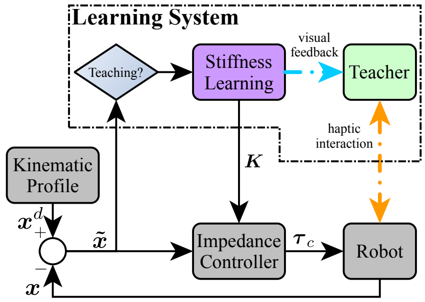


Fig. 2. Overview of the learning system. The stiffness is learned by observing the position deviations from the desired reference point due to the teachers interactions. The robot adapts its stiffness online, so the teacher gets direct haptic feedback of the effect of his interaction. The teacher is further helped by a graphical animation on a screen that represents the current stiffness as an ellipsoid.

$$U_t = min\{j \in [1, 2 \ldots J] : t \leq t_j\} \tag{4b}$$

We denote by $N_t = U_t - L_t$ the number of data points in the sliding window at time t. Let $\boldsymbol{M}_t^1$ and $\boldsymbol{M}_t^2$ denote the first and second empirical moments of the spatial data in the window:

$$\boldsymbol{M}_t^1 = \frac{1}{N_t} \sum_{j=L_t}^{U_t} \tilde{\boldsymbol{x}}_j \tag{5}$$

$$\boldsymbol{M}_t^2 = \frac{1}{N_t} \sum_{j=L_t}^{U_t} \tilde{\boldsymbol{x}}_j \tilde{\boldsymbol{x}}_j^T \tag{6}$$

Furthermore let $\boldsymbol{\Sigma}_t = \boldsymbol{M}_t^2 - \boldsymbol{M}_t^1(\boldsymbol{M}_t^1)^T$ denote the corresponding covariance matrix. This covariance matrix is what determines the stiffness commanded to the robot at time t. The covariance matrix is symmetric and positive definite, so it can be decomposed as $\boldsymbol{\Sigma}_t = \boldsymbol{Q}\Lambda\boldsymbol{Q}^T$, where Λ is a diagonal matrix composed of the eigenvalues $\lambda_t^i > 0, i = 1, 2, 3$, and $\boldsymbol{Q}$ is a matrix containing the orthonormal eigenvectors in its columns. The standard deviation of the data along each eigenvector is given by $\sigma_t^i = \sqrt{\lambda_t^i}, i = 1, 2, 3$. The stiffness matrix $\boldsymbol{K}_t$ is constructed using the same eigenvectors as the covariance matrix :

$$\boldsymbol{K}_t = \boldsymbol{Q}\boldsymbol{\Gamma}\boldsymbol{Q}^T \tag{7a}$$

with

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma(\sigma_t^1) & 0 & 0 \\ 0 & \gamma(\sigma_t^2) & 0 \\ 0 & 0 & \gamma(\sigma_t^3) \end{bmatrix} \tag{7b}$$

where the eigenvalues are set inversely proportional to the square root of the corresponding eigenvalue of the covariance matrix:

$$\gamma(\sigma^i) = \begin{cases} \underline{k} & \overline{\sigma} < \sigma_t^i \\ \overline{k} - (\overline{k} - \underline{k})\frac{\sigma_t^i - \underline{\sigma}}{\overline{\sigma} - \underline{\sigma}} & \underline{\sigma} \leq \sigma_t^i \leq \overline{\sigma} \\ \overline{k} & \sigma_t^i < \underline{\sigma} \end{cases} \tag{7c}$$

for $i = 1, 2, 3$. The admissible values for the stiffness in any direction is bounded below by $\underline{k}$ and above by $\overline{k}$. These, along with the $\underline{\sigma}$ and $\overline{\sigma}$ are tunable parameters of the system.

The teaching process consists of perturbing the robot while it is performing task, i.e. providing a stream of data points which are added to $\boldsymbol{\Xi}$. The data set is sorted in order by increasing time, and new data is simply inserted in the place corresponding to the time at which the perturbation was perceived. The algorithm involves computing empirical covariance of a potentially very large data set. Incremental update of the covariance matrix from time t' to time t follows directly from the additive form of the first and second moments:

$$\boldsymbol{M}_t^1 = \frac{N_{t'}}{N_t}\boldsymbol{M}_{t'}^1 + \frac{1}{N_t}\left(\sum_{i=U_{t'}}^{U_t} \tilde{\boldsymbol{x}}_i - \sum_{i=L_{t'}}^{L_t} \tilde{\boldsymbol{x}}_i\right) \tag{8a}$$

$$\boldsymbol{M}_t^2 = \frac{N_{t'}}{N_t}\boldsymbol{M}_{t'}^2 + \frac{1}{N_t}\left(\sum_{i=U_{t'}}^{U_t} \tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^T - \sum_{i=L_{t'}}^{L_t} \tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^T\right) \tag{8b}$$

and

$$\boldsymbol{\Sigma}_t = \boldsymbol{M}_t^2 - \boldsymbol{M}_t^1(\boldsymbol{M}_t^1)^T \tag{8c}$$

Pseudo-code for the learning procedure is given in Algorthm 1. At line 1, $\boldsymbol{g}(t)$ is introduced to denote the process that at each time instant provides the desired position[4]. As mentioned in section III, the availability of such a process is assumed in this work. In lines 4-7 the update of the data used for stiffness assignment is performed. It is vital that data points are only added to $\boldsymbol{\Xi}$ if teaching is performed, as the covariance would otherwise gradually decrease in the absence of perturbations, with the effect that the robot 'forgets' what it has been taught. In this work, we used no detection of teaching but let the teacher switch between two modes: teaching or not teaching. Lines 8-14 computes the the stiffness matrix based on the current window view of $\boldsymbol{\Xi}$. Then appropriate damping[5]. Refer to Section III for details on the construction of Φ and Ψ. An overview of the complete system is given in Fig. 2.

V. Experiments

Two experiments were conducted to evaluate the proposed system. The first is designed to demonstrate that the system can learn stifness variations both in direction and magnintude, as instructed by the teacher. In the second experiment, we illustrate the usefulness of the system by teaching a stiffness profile for a task of pouring a drink into a glass.

A. Setup

The system for learning stiffness through interaction as described in section IV-B was implemented in a control module for a 7-dof Barrett WAM, using the RobotToolKit (RKT) and ROS software frameworks[6]. For implementing

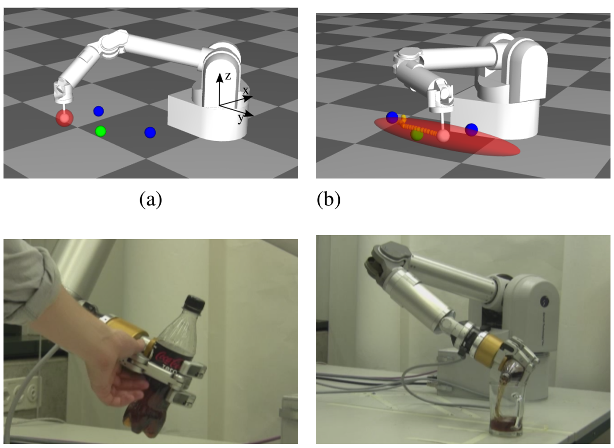
[4]The system can control for orientation at the same time as learning the translational stiffness, but the inclusion of a desired orientation in the reference trajectory is not required for learning translational stiffness.

[5]In this work, we designed the damping matrix to have the same eigenvectors as the stiffness, with eigenvalues $d^i = \sqrt{k^i}$.

[6]RobotToolKit is an open source collection of tools for robot simulation and control developed by Eric Sauser. ROS (Robot Operating System) is open source robot middleware developed by Willow Garage

Algorithm 1 Online Learning of Variable Stiffness

1: Given $\boldsymbol{g}(t) = \boldsymbol{x}_t^d$, S, $\underline{k}$, $\overline{k}$, $\underline{\sigma}$, $\overline{\sigma}$
2: **for** $t < T_f$ **do**
3: sense $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{g}(t)$
4: **if** teaching **then**
5: add deviation to data set: $\boldsymbol{\Xi} \leftarrow \boldsymbol{\Xi} \cup \{\tilde{\boldsymbol{x}}, t\}$
6: sort $\boldsymbol{\Xi}$ in order of increasing time
7: **end if**
8: update L_t and U_t (Eq. 4)
9: update moments based on previous values (Eq. 8)
10: compute Σ_t and its eigenspace $\boldsymbol{Q}, \boldsymbol{\Lambda}$
11: compute $\boldsymbol{\Gamma} \leftarrow diag([\gamma(\sigma_t^1), \gamma(\sigma_t^2), \gamma(\sigma_t^3)])$ (Eq. 7)
12: compute translational stiffness matrix $\boldsymbol{K}_t \leftarrow \boldsymbol{Q}\boldsymbol{\Gamma}\boldsymbol{Q}$
13: compute damping $\boldsymbol{D}_t$ and construct $\boldsymbol{\Phi}_t, \boldsymbol{\Psi}_t$, (Eq. (3))
14: compute wrench $\boldsymbol{\zeta}_t = \Phi_t \tilde{\boldsymbol{\xi}} - \Psi_t \dot{\tilde{\boldsymbol{\xi}}}$
15: **end for**



(a) (b) (c) (d)

Fig. 3. Top-left: The figure shows the layout of the via points for the first task. The points should be taken from left to right, with selective compliance in z-direction at the blue points and in x-direcion at the green point, as indicated by the arrows. Figure top-right shows the graphical aid provided to the teacher while teaching. The simulator mirrors the robot motions while drawing a graphical representation of the current stiffness (the red ellipsoid). Figure bottom-left and bottom-right shows a snapshots of the robot pouring task during teaching and task reproduction.

the control objective (2), the following control law was used for the actuator torques $\boldsymbol{\tau_c}$:

$$\boldsymbol{\tau}_c = \boldsymbol{\tau}_{ff}(\boldsymbol{q}) + \boldsymbol{J}(\boldsymbol{q})^T \boldsymbol{\zeta} \tag{9}$$

where $\boldsymbol{q}$ denotes the joint angles, $\boldsymbol{J}(\boldsymbol{q})$ the manipulator jacobian and $\boldsymbol{\tau}_{ff}(\boldsymbol{q})$ a feedforward gravity compensation term. The rotational stiffness was set to a diagonal matrix with a constant rotational stiffness of 6 Nm/rad around all three axes. The lower and upper bounds for stiffness $\underline{k}$ and $\overline{k}$ where set to 50 and 350 N/m respectively. Empirically, 50 N/m is what the used setup needs to overcome static friction. The upper bound was set as a safety precaution. The parameters $\underline{\sigma}$ and $\overline{\sigma}$ where set to 0.005 and 0.05 respectively. These parameters control the values at which the the stiffness saturate and were set experimentally. The length S of the sliding temporal window was set to 1 second. The teacher could at the beginning of each task reproduction choose if teaching was to be performed or not (cf. line 4 in Alg. 1) by pressing a key on the keyboard.

As the focus of this work is not learning the kinematic task profile, simple record-and-replay was used for generating the desired pose trajectories. To this end, the robot was put in gravity compensation mode, and guided through the different motions by the teacher while the pose trajectory was recorded.

The translational stiffness was fed to the controller at a rate of 10 Hz. The translational damping was set to have the same principal directions as the stiffness (cf. Eq. (7)) and with eigenvalues $d_i^t = 2\sqrt{\gamma^i(t)},\ i \in \{1, 2, 3\}$. The reference point $\boldsymbol{x}^d$ was updated at each iteration of the inner control loop, which runs at 500 Hz.

A RKT simulator was set up to provide graphical aid to the teacher while performing demonstrations by mirroring the robot movement on a screen and drawing a graphical representation of the current stiffness as an ellipsoid, see Fig. 3b. The ellipsoid is shaped inversely to the stiffness, so low stiffness in a direction is represented by the ellipsoid being elongated in that direction.

B. Task 1: Via-point Trajectory

The purpose of this experiment is to illustrate the claim that the proposed system can learn stiffness variations in selective directions. The task consists in following a trajectory passing through three via points. Fig. 3 shows the via points in the robot's workspace. The robot should be maximally compliant in the z-direction at the blue via-points, and maximally compliant in the x-direction in the other directions. This is a quantitative requirement in that the task constraints state specifically that the robot should assume its minimum allowed stiffness $\underline{k}$ for one and one only of the eigenvalues at each of the via-points. Furthermore, the directions of compliance at each via-point are specified to be aligned approximately[7] with the z-axis for the blue via-points and the x-axis for the green via-point. In between the via-points, the robot should stiffen up as quickly as possible[8] to its maximal stiffness in all directions. The reference trajectory moves through the points with approximately constant speed. The total duration of the task is 12 seconds.

The robot executed the task three times while the teacher was providing input. Fig. 4a shows trajectory followed without interaction in the XY-plane of the base coordinate system, with the trajectories resulting from the teachers interaction overlaid. As is clear from the figure, the teacher imposed perturbations in the x-direction at the green via-point. Note also the small amount of variance imposed in y-direction at the blue via-points. These small perturbations are an unintended bi-effect of the larger perturbations imposed in z-direction at the same points, as can be seen in figure 4b.

[7]This requirement is only approximate since it can not be expected by a human teacher to provide perturbations exactly aligned with a given direction.

[8]The rate at which the stiffness can change is limited as an effect of the sliding temporal window, refer to section IV-B

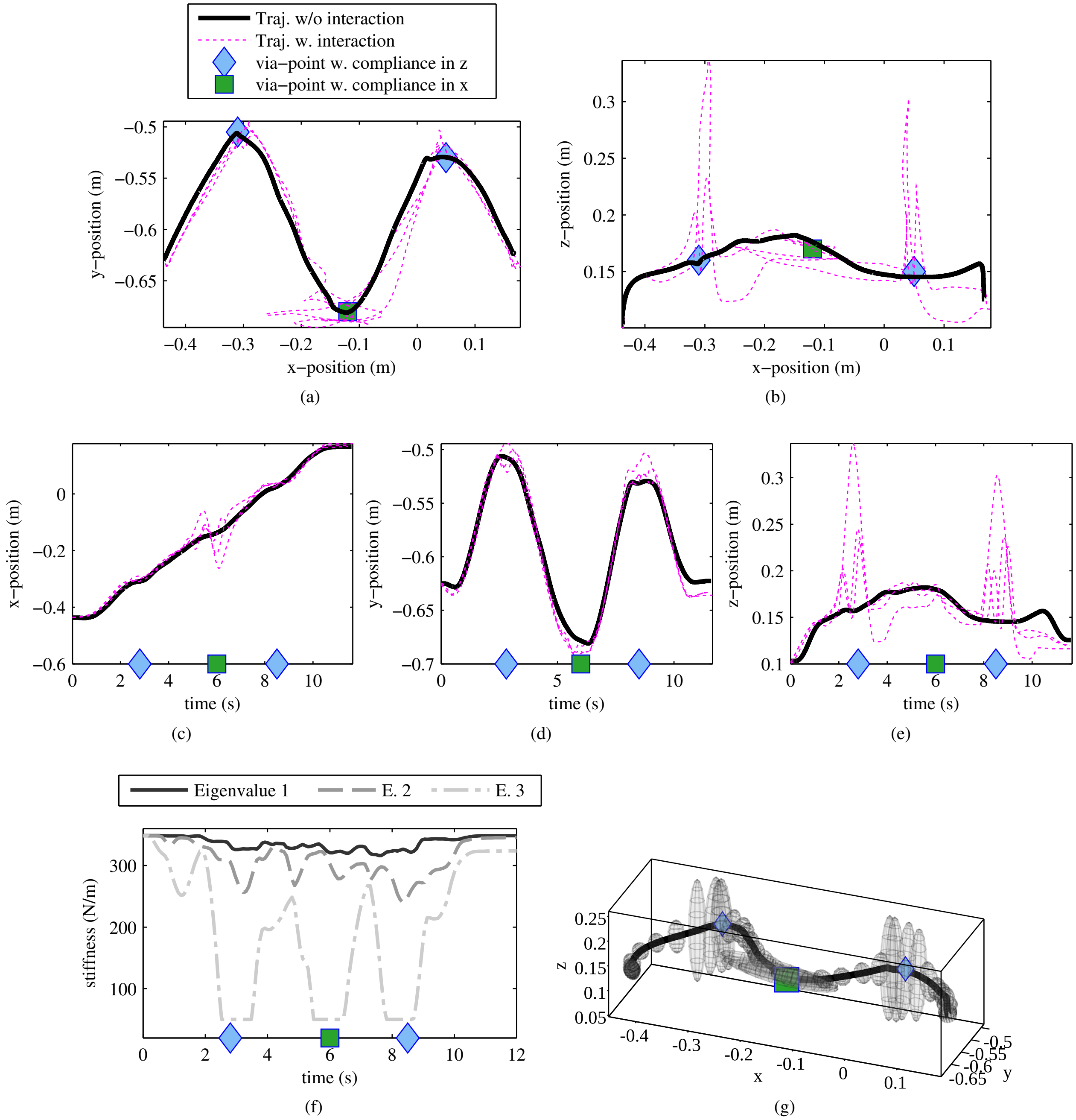


Fig. 4. These figures show data from the via-point trajectory following experiment. (a) and (b): The figures show the reference trajectory with the trajectories resulting from the teachers interactions overlaid in the xy-plane and xz-plane respectively. (c): Trajectory plotted as x-position over time. Note the variance imposed by the teacher around the green via-point. (d): Trajectory along y-direction plotted over time. (e): Trajectory followed in z-direction over time. Note the variance imposed by the teacher around the blue via-points. Figures (f) and (g) show the stiffness trajectories resulting from the teachers interaction showed in (a)-(e). Figure (f) is a plot of the eigenvalues of the stiffness matrix over time.

Figures 4c, 4d and 4e show the trajectories followed along each direction x,y,z of the base coordinate system over time. Even though the via-points were originally defined by space coordinates, they are implicitly anchored in time since the trajectories are time dependent. The times at which the via-points are marked are simply the times at which the reference trajectory pass through these points. As can be seen in these figures, the trajectory from the teaching rounds generally has a bias error when compared to the reference trajectory. This is bias is due to the teacher holding the end-effector and following along even when not imposing perturbations. Note that this bias does not affect the stiffness as only the covariance of the perturbations is used for determining the stiffness (cf. Alg. 1).

In Fig. 4f, the stiffness eigenvalues resulting from the teaching is plotted over time. The plot clearly shows that maximum compliance is only reached in one direction at each of the via points. In Fig. 4f, the stiffness matrix is plotted as an ellipsoid at a series of points along the motion trajectory. The ellipsoids are built around the eigenvectors of the stiffness matrix, with low stiffness along an eigenvector illustrated by the ellipsoid being elongated in that direction. As can be seen in this figure, the direction corresponding to the lowest eigenvalue is approximately z-direction at the blue via-points and approximately x-direction at the green via-point. The human teacher being unable to impose perturbations exactly along the desired directions is the reason for the directions being only approximately correct.

C. *Task 2: Pouring a drink*

This task was chosen to show how the proposed system can be used to teach a realistic task that benefits from a varying stiffness profile. The task consists first transporting a bottle full of soda toward a glass. Once above the glass, the robot was to pour the drink. We state the following desired *qualitative* characteristics for this task:

1) During the reaching phase, the robot should be compliant in all directions, as position errors are not crucial and correcting for position errors with high stiffness can result in high accelerations of the end-effector which spills the drink out of the bottle.
2) In the pouring phase, the robot should stiffen up in all directions, since the drink should be poured into the glass, even if moderately strong perturbations are encountered.
3) In the third stage, when the robot is reaching away from the glass, it is again desirable that a low stiffness is used, for the same reason as mentioned for the reaching stage.

The reference trajectory was acquired using record and replay. The total duration of the task is 25 seconds, and the critical pouring phase starts 10 seconds into the task and ends 6 seconds later.

The refinement of the pouring task consisted in decreasing stiffness in all directions in the reaching phase, letting the robot be stiff while pouring, and again decreasing the stiffness after the pouring phase. Since the stiffness was to be decreased in several directions at each point along the reaching phases, three rounds of teaching were performed. During each round, the teacher concentrated on introducing variance in along the three coordinate-axes x,y and z of the base coordinate system. The x,y,z components of the trajectories from the teaching rounds are swon in figures 5a, 5b and 5c. Note that the teacher did not impose as high variations in x-direction as in y, and z-direction. The reason for this is that the motion of the task was approximately aligned with the x-axis throughout the entire reaching phase. Perturbing the robot heavily along the planned direction of motion makes it hard for the teacher to respect the intrinsic time dependency of the task, thus risking to anchor the teaching in a part of the motion where this was not intended. Respecting the time dependency along other directions, especially those that are orthogonal to the direction of motion, is easier as the teacher can feel the robots desired motion and follow it while perturbing in other directions.

The result of the relatively smaller perturbations in the x-directions are directly visible in Fig. 5d, which shows the stiffness eigenvalues resulting from the three teaching rounds. Clearly, all three eigenvalues drop during the reaching phases, while two of them drop much more than the first. Fig. 5e shows the ellipsoid representation of the stiffness matrices for a subset of the points along the followed trajectory, again making it clear that the stiffness was principally reduced in the XZ-plane. As expected, the robot assumed a high stiffness at the beginning of the task and in the pouring phase, since the teacher provided no interaction there.

VI. Conclusion

We have presented an online, incremental algorithm for learning variable stiffness. The algorithm sets the stiffness inversely proportional to the covariance of perturbation data imposed by the teacher. The data taken into account for determining the stiffness is taken from a sliding temporal window over the set of all provided data. We wish to emphasize that any interface for teaching stiffness should provide haptic feedback to the teacher, as this is the way that humans can evaluate stiffness. Our system offers an efficient solution to this as the manipulator itself is used as haptic display.

Throughout this paper, we have assumed that the robot already knows a kinematic profile for the task. No assumptions have been made as to how this kinematic profile is generated. This means that our system can be used as an add-on to any other system learning robot motions. In this paper, we considered only trajectories with explicit time dependency. We plan to integrate the system in this paper with our previous work for modeling motions as autonomous dynamical systems [18] to allow robust teaching of tasks without time dependency.

The mapping from the standard deviation of the perturbations to stiffness in (7c) was chosen linear to make it easy for the user to identify how the imposed perturbations affects the stiffness. Using other mappings from the perturbations to the stiffness is an approach yet to explore. Such a mapping could be designed e.g. to ignore perturbations with low frequency.

In section IV-B it was explained that the presented system is data driven. The empirical covariance matrix of subset of the collected data is computed at each iteration. For our experimental setup, the computation required was well within the requirements of the 10Hz update frequency used for the stiffness. The real drawback of the data-driven approach is rather that all data points have to be saved in memory. In future work, modeling the data using parametric models of the underlying distribution will be explored, as this could potentially eliminate the need of storing all the perturbation data in memory.

VII. ACKNOWLEDGMENT

This research was supported by the Swiss National Science Foundation through the National Centre of Competence in Research Robotics.

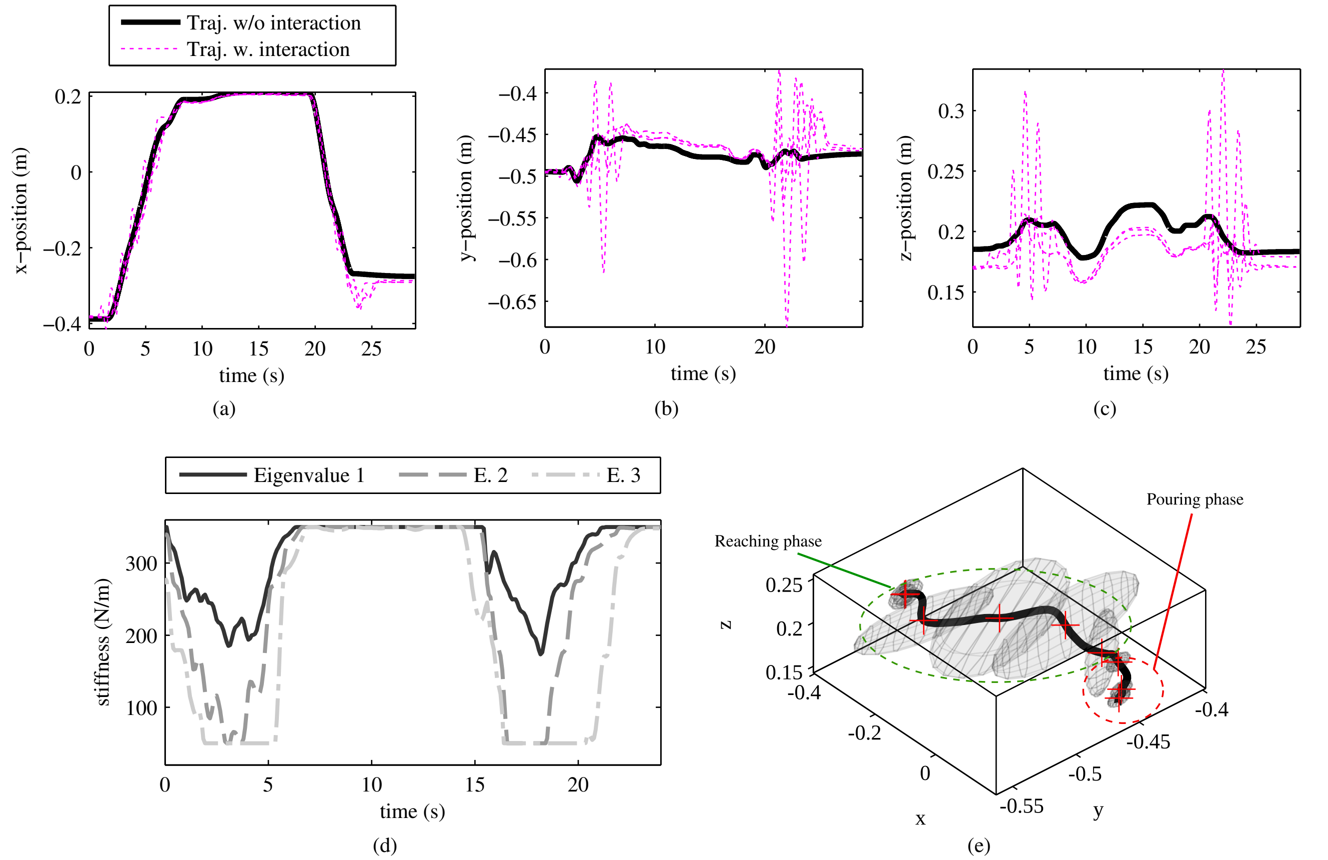


Fig. 5. Figures (a)-(c) shows the trajectory followed by the robot when unperturbed along with the trajectory followed during two teaching rounds overlaid. Note that no perturbations were provided along the x-direction. Figures (d) and (e) show the stiffness resulting from these teaching rounds. Note that all eigenvalues reach their maximum during the pouring phase which takes place at second 10 to second 16. As is seen in figure (e), the principal directions in which the robot has reduced stiffness in the reaching phase lies in the yz-plane, while it remains fairly stiff along the x-axis.

REFERENCES

[1] N. Hogan, "Impedance control: An approach to manipulation," *Journal of Dynamic Systems Measurement and Control (1985)*, no. March, pp. 304–313, 1985.

[2] D. Mitrovic, S. Klanke, and S. Vijayakumar, "Learning Impedance Control of Antagonistic Systems Based on Stochastic Optimization Principles," *The International Journal of Robotics Research*, Dec. 2010.

[3] M. Howard, D. Mitrovic, and S. Vijayakumar, "Transferring impedance control strategies between heterogeneous systems via apprenticeship learning," in *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*. IEEE, 2010, pp. 98–105.

[4] M. Garabini, A. Passaglia, F. Belo, P. Salaris, and A. Bicchi, "Optimality Principles in Variable Stiffness Control : The VSA Hammer," in *International Conference on Intelligent Robots and Systems (IROS)*, 2011.

[5] J. Buchli, E. Theodorou, F. Stulp, and S. Schaal, "Variable impedance control a reinforcement learning approach," *RSS*, vol. 100, no. 1, 2010.

[6] P. Kormushev, S. Calinon, and D. Caldwell, "Robot motor skill coordination with em-based reinforcement learning," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 3232–3237.

[7] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469–483, May 2009.

[8] L. Rozo, P. Jimenez, and C. Torras, "Robot Learning from Demonstration of Force-based Tasks with Multiple Solution Trajectories," in *15th International Conference on Advanced Robotics (ICAR)*, 2011, pp. 124–129.

[9] P. Kormushev, S. Calinon, and D. Caldwell, "Imitation Learning of Positional and Force Skills Demonstrated via Kinesthetic Teaching and Haptic Input," *Advanced Robotics*, pp. 1–20.

[10] D. S. Walker and R. P. Wilson, "User-Controlled Variable Impedance Teleoperation," in *International Conference on Robotics and Automation*, 2010, pp. 5352–5357.

[11] H. Gomi and Kawato, "Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement." *Science (New York, N.Y.)*, vol. 272, no. 5258, pp. 117–20, Apr. 1996.

[12] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance." *Nature*, vol. 414, no. 6862, pp. 446–9, Nov. 2001.

[13] 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 *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 2705–2711.

[14] S. Calinon, I. Sardellitti, and D. Caldwell, "Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 249–254.

[15] a. Billard, "Discovering optimal imitation strategies," *Robotics and Autonomous Systems*, vol. 47, no. 2-3, pp. 69–77, Jun. 2004.

[16] B. Argall, E. Sauser, and A. Billard, "Tactile guidance for policy refinement and reuse," in *Development and Learning (ICDL), 2010 IEEE 9th International Conference on*. IEEE, 2010, pp. 7–12.

[17] D. Lee and C. Ott, "Incremental motion primitive learning by physical coaching using impedance control," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ*, 2010, pp. 4133–4140.

[18] S. Khansari-Zadeh and A. Billard, "Learning stable non-linear dynamical systems with Gaussian Mixture Models," *IEEE Transactions on Robotics*, vol. 27, pp. 1–15, 2011.