Imitation Learning of Positional and Force Skills Demonstrated via Kinesthetic Teaching and Haptic Input

Petar Kormushev, Sylvain Calinon and Darwin G. Caldwell

Advanced Robotics Department, Italian Institute of Technology (IIT), 16163 Genova, Italy

{petar.kormushev, sylvain.calinon, darwin.caldwell}@iit.it

Abstract

A method to learn and reproduce robot force interactions in a Human-Robot Interaction setting is proposed. The method allows a robotic manipulator to learn to perform tasks which require exerting forces on external objects by interacting with a human operator in an unstructured environment. This is achieved by learning two aspects of a task: positional and force profiles. The positional profile is obtained from task demonstrations via kinesthetic teaching. The force profile is obtained from additional demonstrations via a haptic device. A human teacher uses the haptic device to input the desired forces which the robot should exert on external objects during the task execution. The two profiles are encoded as a mixture of dynamical systems, which is used to reproduce the task satisfying both the positional and force profiles. An active control strategy based on task-space control with variable stiffness is then proposed to reproduce the skill. The method is demonstrated with two experiments in which the robot learns an ironing task and a door opening task.

keywords: imitation learning, kinesthetic teaching, programming by demonstration, pHRI, haptics

1 INTRODUCTION

Integrating robots into human-populated unstructured environments requires them to be able to easily learn new skills. Similarly to humans, a robot should be able to acquire new skills using various forms of learning.

Some skills can be successfully transferred to a robot using imitation strategies [1] [2]. Others can be learned very efficiently by the robot alone using reinforcement learning [3]. In both cases, the robot should be able to construct a task representation which reflects all aspects of the task, e.g. positional constraints, time constraints, force constraints, movement coordination, task variability, and so on. In this paper we propose a user-friendly methodology to transfer skills to a robot including both positional and force aspects of the task.

To transfer new skills to robots, several directions in Human-Robot Interaction (HRI) have been adopted which try to benefit from the shape, human likeness and interaction capabilities of the robots. A

review of possible interactions through speech, gestures, imitation, observational learning, instructional scaffolding, as well as physical interaction such as molding is given in [4].

Thomaz et al. in [5] presented HRI experiments in which the user can guide the robot's understanding of the objects in its environment through facial expression, vocal analysis, detection of shared attention and by using an affective memory system. Rohlfing et al. in [6] highlighted the importance of having multi-modal cues to reduce the complexity of human-robot skill transfer and considered multi-modal information as an essential element to structure the demonstrated tasks.

The physical contact between robots, humans, and the environment can take various forms. A broad spectrum of issues regarding the problem of safe and dependable physical HRI is given in [7].

When direct physical contact is possible, *kinesthetic teaching* offers a user-friendly and intuitive method to demonstrate new skills to a robot by manually guiding the robot's arms through the motion [1]. Different than standard teleoperation techniques, it allows the user to feel firsthand the limitations of the robot's body (e.g. small number of DOF, joint limits, short limbs, singularities, too big stiffness, too slow movement, self-collision, etc.), which is especially important on highly-dynamic tasks [8]. Teleoperation techniques tend to eliminate this "firsthand perception" of the robot's body limitations and the remote user often remains unaware of the commands the robot can or cannot execute. Kinesthetic teaching offers a simple way to solve this, as it uses directly the robot in its environment as interface. The kinesthetic teaching process shares similarities with the molding behaviors observed in caregiver-child interaction (e.g. when teaching an infant to assemble pop beads [9]), or in teacher-learner interaction (e.g. when teaching a beginner to execute a golf swing [10, 11]). Kinesthetic teaching also relates to the concept of scaffolding applied to human-robot teaching where a human teacher guides a robot to new competence levels by exploiting and extending its existing competencies [12].

Several approaches have been proposed to perform kinesthetic teaching on robots. The most simple approach consists of using backdrivable motors and encoders to collect joint angle information while the user moves the robot, with the motors being switched off when recording a gesture. This approach only works for small robots (or weak motors), as it requires the user to manually move the motors' mechanical components. This method also has the disadvantage that it tends to decorrelate the gesture. The main reason is that it is easier for the user to move the motors one by one during demonstration than executing a natural coordinated movement [13].

Active control is thus preferred in many applications with heavier and/or more powerful robots. A possible solution consists of providing the robot with force sensors or pressure sensors placed on various parts of its body, and to actively guide the robot depending on the sensed force/pressure input, see e.g. [14]. This approach has the advantage to be generically applicable to various kinds of robotic platforms, as the attachment of sensors does not directly interfere with the mechanical and motor control design of the robot. One disadvantage is that it requires a large set of sensors to cover different parts of the body (if one wants to move each limb individually). For this reason, a common practical approach consist of utilizing a single force sensor mounted at the level of the end-effector, and move the robot through the task by displacing manually the end-effector in task space. From a control perspective, the

process is simplified, but it can only transfer kinematics information.

Lightweight robots offer new perspectives to augment the personal capabilities of the robot through kinesthetic teaching, where control at a torque level can be used to let the user move the robot as if it had no weight, and as if no motors were in its articulations. Through a gravity compensation actuation, the user can concentrate on the task to demonstrate by grasping (and re-grasping) the robot in ways that are natural and efficient for him/her to execute the task. Successful applications of this torque-based kinesthetic teaching process have been recently developed for the Barrett WAM and KUKA/DLR robotic arms [8, 15].

In this paper, we refer to the kinematic redundancy of the robot when the robot possesses an infinite number of generalized inverse control strategies, see e.g. [16, 17]. We refer to task redundancy when the task can be achieved through an infinite number of solutions, see e.g. [18]. We take the perspective that both the robot and task redundancies can be exploited to regulate the dynamics of the movement and the stiffness of the robot during reproduction.

After having observed several demonstrations of a similar task, the robot creates a compact model of the skill, by taking into account the variations and correlations observed along the movement. If a part of the movement was consistent across the different trials, this part of the task should probably be reproduced in this specific manner. On the other hand, if a large variability was observed among the different demonstrations, reproducing a specific reference trajectory may be irrelevant to fulfill the task requirements.

During reproduction, the robot is using this information to set an adequate stiffness that will fulfill the task constraints, which allows to simultaneously consider other constraints. There are several situations where the interaction can benefit from the variability and correlations of the task: (i) to let the user physically move the robot while reproducing the task; (ii) to let the robot modify the generalized trajectory to adopt gestures that are safer for a user who is close to the robot.

2 PROPOSED APPROACH

The proposed methodology for demonstrating the skill is based on the joint use of a gravity-compensated lightweight robot arm for kinesthetic teaching, and an additional haptic device to transfer the required force skills. A flowchart of the proposed method is shown in Fig. 1.

The positional and force constraints of the demonstrated skill are represented as a mixture of dynamical systems encoding robustly position and force trajectories. The *Dynamic Movement Primitives* (DMP) framework originally proposed by Ijspeert *et al* [19], and further extended in [20, 21] (see [22] for a discussion on the similarities of the proposed controller with DMP) is extended by considering a full matrix $K_i^{\mathcal{P}}$ associated with each of the K primitives (or states) instead of a fixed $\kappa^{\mathcal{P}}$ gain. This allows us to take into consideration variability and correlation information along the movement for learning and reproduction.

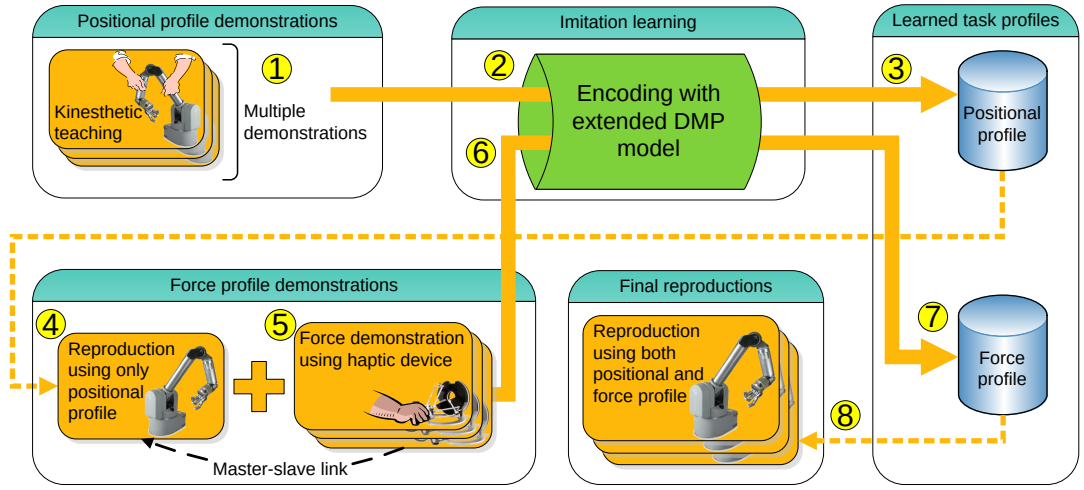

Figure 1: Flowchart of the proposed method. (1) Multiple demonstrations of the positional profile of the task are performed via kinesthetic teaching. (2) The positional data is encoded in a mixture of Dynamical Systems and (3) stored as a positional profile of the task. (4) Using only the learned positional profile, the task is executed on the robot while (5) simultaneously performing multiple force profile demonstrations using a haptic device, which is coupled with the robot in a master-slave link. (6) The multiple demonstrations of the force data are encoded a mixture of Dynamical Systems and (7) stored as a force profile of the task. (8) Finally, a stand-alone reproduction of the task is performed using the learned both positional and force profiles of the task.

2.1 Encoding the positional profile of the skill

M examples of a skill are demonstrated to the robot in slightly different situations. Each demonstration $m \in \{1, \dots, M\}$ consists of a set of T_m positions x, velocities $\dot{x}$ and accelerations $\ddot{x}$ of the end-effector in Cartesian space, where each position x has $D = 3$ dimensions. A dataset is formed by concatenating the $N = \sum_{m=1}^{M} T_m$ datapoints $\{\{x_j, \dot{x}_j, \ddot{x}_j\}_{j=1}^{T_m}\}_{m=1}^{M}$. A desired acceleration is computed based on a mixture of K proportional-derivative systems

$$\hat{\ddot{x}} = \sum_{i=1}^{K} h_i(t) \Big[K_i^{\mathcal{P}} (\mu_i^X - x) - \kappa^{\mathcal{V}} \dot{x} \Big]. \tag{1}$$

Parts of the movement where the variations across the different demonstrations are large indicate that the reference trajectory does not need to be tracked precisely. By using this information, the controller can focus on the other constraints of the task such as collision avoidance. On the other hand, parts of the movement exhibiting strong invariance across the demonstrations should be tracked precisely, i.e., the stiffness used to track the position errors needs to be high.

The superposition of basis vector fields is determined in (1) by an implicit time dependency, but other approaches using spatial and/or sequential information could also be used [23, 24]. Similarly to DMP, a decay term defined by a canonical system $\dot{s} = -\alpha s$ is used to create an implicit time dependency $t = -\frac{ln(s)}{\alpha}$, where s is initialized with $s = 1$ and converges to zero. We define a set of Gaussians $\mathcal{N}(\mu_i^T, \Sigma_i^T)$ in time space τ, with centers μ_i^T equally distributed in time, and variance parameters Σ_i^T set to a constant value inversely proportional to the number of states. α is initially fixed depending on the

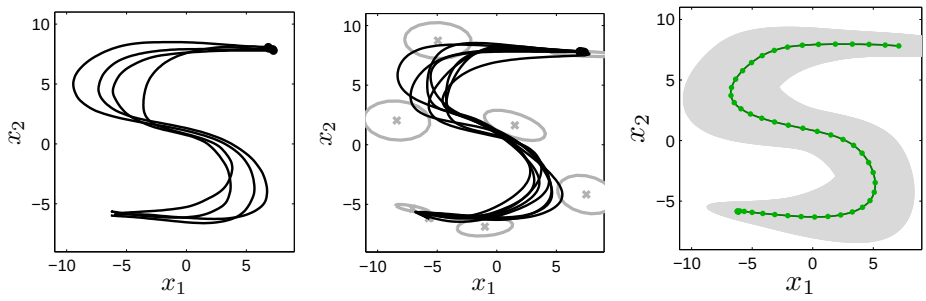

Figure 2: Illustration of the learning and retrieval processes. *Left:* Four examples of the task provided as demonstrations. *Center:* Learned model represented by Gaussians $\mathcal{N}(\mu_i^X, \Sigma_i^X)$, and multiple reproduction attempts. *Right:* The trajectory in green lines shows a generalized reproduction attempt. The points show positions at constant time intervals.

duration of the demonstrations. The weights $h_i(t)$ are defined by

$$h_i(t) = \frac{\mathcal{N}(t;\ \mu_i^T, \Sigma_i^T)}{\sum_{k=1}^{K} \mathcal{N}(t;\ \mu_k^T, \Sigma_k^T)}. \tag{2}$$

By determining the weights through the decay term s, the system will sequentially converge to the set of attractors in Cartesian space defined by μ_i^X. The centers μ_i^X in task space and stiffness matrices $K_i^{\mathcal{P}}$ are learned from the observed data, either incrementally or in a batch mode (through least-squares regression). For example, parts of the movement where the variations between the demonstrations are high indicate that the reference trajectory does not need to be tracked precisely. By using this information, the controller can focus on the other constraints of the task such as moving away from the user. On the other hand, parts of the movement exhibiting strong invariance among the demonstrations should be tracked precisely, i.e., the stiffness used to track the position errors needs in this case to be high.

In a batch mode, by concatenating the training examples in a matrix $Y = [\ddot{x}\frac{1}{\kappa^{\mathcal{P}}} + \dot{x}\frac{\kappa^{\mathcal{V}}}{\kappa^{\mathcal{P}}} + x] \in \mathbb{R}^{N\times D}$, and by concatenating the corresponding weights computed with (2) in a matrix $H \in \mathbb{R}^{N\times K}$, we can write the linear equation $Y = H\mu^X$, with $\mu^X \in \mathbb{R}^{K\times D}$ representing the concatenated attractor centers μ_i^X. The least-squares solution to estimate the attractor centers is then given by $\mu^X = H^{\dagger}Y$, where $H^{\dagger} = (H^{\top}H)^{-1}H^{\top}$ is the pseudoinverse of H. By defining a desired range of stiffness values $[\kappa^{\mathcal{P}}_{\min}, \kappa^{\mathcal{P}}_{\max}]$, we define the stiffness and damping gains for the estimation of the parameters as $\kappa^{\mathcal{P}} = \kappa^{\mathcal{P}}_{\min} + \frac{\kappa^{\mathcal{P}}_{\max} - \kappa^{\mathcal{P}}_{\min}}{2}$ and $\kappa^{\mathcal{V}} = 2\sqrt{\kappa^{\mathcal{P}}}$.

To take into account variability and correlation along the movement and among the different demonstrations, we compute for each state $i \in \{1, \ldots, K\}$ the residual errors of the least-squares estimation, in the form of covariance matrices

$$\Sigma_i^X = \frac{1}{N}\sum_{j=1}^{N}(Y'_{j,i} - \bar{Y}'_i)(Y'_{j,i} - \bar{Y}'_i)^{\top} \quad \forall i \in \{1, \ldots, K\}, \qquad \text{where } Y'_{j,i} = H_{j,i}(Y_j - \mu_i^X). \tag{3}$$

In the above equation, $\bar{Y}'_i$ is the mean of Y'_i over the N datapoints. $\mathcal{N}(\mu_i^X, \Sigma_i^X)$ thus describes a Gaussian in Cartesian space $\mathcal{X}$. The set of K Gaussians defines the sequence of virtual attractor points in Cartesian space that the system will try to reach, where each attractor encapsulates variability and

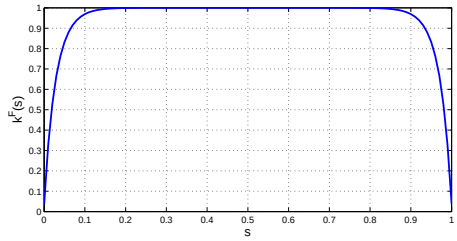

Figure 3: Force envelope used to smoothly limit the exerted force at the beginning and towards the end of the task.

correlation information. The residuals terms of the regression process are then used to estimate the stiffness matrices $K_i^{\mathcal{P}}$ in Eq. (1) through eigencomponents decomposition

$$K_i^{\mathcal{P}} = V_i D_i V_i^{-1}, \qquad \text{with } D_i = \kappa_{\min}^{\mathcal{P}} + (\kappa_{\max}^{\mathcal{P}} - \kappa_{\min}^{\mathcal{P}}) \frac{\lambda_i - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}. \tag{4}$$

In the above equation, λ_i and V_i are the concatenated eigenvalues and eigenvectors of the inverse covariance matrix $(\Sigma_i^X)^{-1}$. The basic idea is to determine a stiffness matrix proportional to the inverse of the observed covariance. For example, if high variability is observed, stiffness will become low as the tracking does not need to be precise. If D_i in (4) is set to λ_i, the eigencomponents decomposition gives $K_i^{\mathcal{P}} = (\Sigma_i^X)^{-1}$. We rescale D_i to obtain stiffnesses in the desired range $[\kappa_{\min}^{\mathcal{P}}, \kappa_{\max}^{\mathcal{P}}]$ (determined by the user and hardware's limitation) based on the initial range of eigenvalues $[\lambda_{\min}, \lambda_{\max}]$ (determined by the variability within the motion and among several demonstrations). The minimal stiffness limit allows us to set a degree of compliance that still allows the system to move (e.g. to limit friction effects), while the maximum stiffness limit can be fixed depending on the robot capabilities, or based on desired safety limits.

Fig. 2 illustrates the approach with a 2-dimensional example. In the top-right graph, we see that the trajectories reproduced stochastically from the learned model show different levels of variability along the trajectory. This variability shows similar characteristics to the one of the training set. We also see that the reproduced trajectories share similar smoothness to the demonstrated trajectories. The trajectories have been represented with dots showing the position of the robot at fixed time intervals.

Source codes for the algorithms in this section are available at: `http://programming-by-demonstration.org/`.

2.2 Encoding the force profile of the skill

The demonstrations of force interactions required by the proposed method are performed via a haptic input device. The device is put in gravity-compensation mode and a controller based on virtual mass-spring-damper model is implemented. The attractor point of the controller is set to be the geometrical center $\bar{x}_h$ of the haptic device workspace. Any force which the user applies to the haptic device handle results in a displacement away from the center. Based on this displacement and the current speed of the handle, the controller calculates the haptic force as

$$f = \kappa^{\mathcal{PH}}(\bar{x}_h - x_h) - \kappa^{\mathcal{VH}}\dot{x}_h, \tag{5}$$

where x_h is the current Cartesian position of the haptic device handle, and $\bar{x}_h$ is the geometrical center of the working area of the haptic device. The stiffness gain $\kappa^{\mathcal{PH}}$ and damping gain $\kappa^{\mathcal{VH}}$ are determined empirically according to the expected magnitude and changing speed of the forces during the force profile task demonstrations. The gains are determined to have an appropriate mapping between the real forces required to execute the task and the force applied to the haptic system, by considering a desired range of skills and the haptic device characteristics.

The exerted forces $f \in \mathbb{R}^3$ on the haptic device are recorded during the demonstrations and used to encode the force profile of the task. We use a similar approach to the one for encoding the positional profile described in Section 2.1. In this case the DMP for the force takes the following form

$$\hat{\dot{f}} = \sum_{i=1}^{K} h_i(t) \Big[K_i^{\mathcal{PF}} (\mu_i^F - f) - \kappa^{\mathcal{VF}} \dot{f} \Big]. \tag{6}$$

The synchronization between the two profiles (positional and force profiles) is achieved through the shared canonical variable s altering $h_i(t)$. Examples of learned force profiles are given in Section 3.

During reproduction, in most cases, it is desirable that there is no sudden large force exerted by the robot at the beginning and at the end of the task. To guarantee smooth start-stop transitions during reproduction, we limit the exerted force using a *force envelope* defined by the multiplicative gain

$$k^F(s) = \beta^f \Big[1 - \big| arctanh(k^L(s - 0.5)) \big|^{k^C} \Big], \tag{7}$$

which modulates the force $f^F = k^F(s)\ f$ applied to the end-effector. The parameters k^L and k^C determine the curvature of the force envelope. For the experiments described in this paper, we used values $k^L = 1.52$ and $k^C = 10$, which produces the force envelope shown in Fig. 3. β^F is a constant force scaling parameter, necessary when the required task forces are much larger than the force capabilities of the haptic device. For the experiments in this paper we used $\beta^F = 3$, i.e. the robot exerted forces three times larger than the corresponding input forces from the haptic device.

2.3 Controller

To control the robot, we exploit the torque-feedback properties of the manipulator, where the robot remains actively compliant for the degrees of freedom that are not relevant for the task. We control the n degrees of freedom (DOFs) robot through inverse dynamics solved with recursive Newton Euler algorithm [25]. The joint forces f_i at each joint $i \in \{1, \ldots, n\}$ are therefore calculated as

$$f_i = f_i^a - f_i^e + \sum_{j \in c(i)} f_j, \tag{8}$$

where f_i^a is the net force acting on link i, f_j with $j \in c(i)$ are the forces transmitted by the child $c(i)$ of link i, and f_i^e are the external forces defined as

$$f_i^e = F^P + F^F + F^G. \tag{9}$$

In the above equation, $F^P = [f^P, M^P]^\top \in \mathbb{R}^6$ is the vector of force and momentum requested to accomplish the positional constraints (only applied at the end-effector for $i = n$), $F^F = [f^F, 0]^\top \in \mathbb{R}^6$

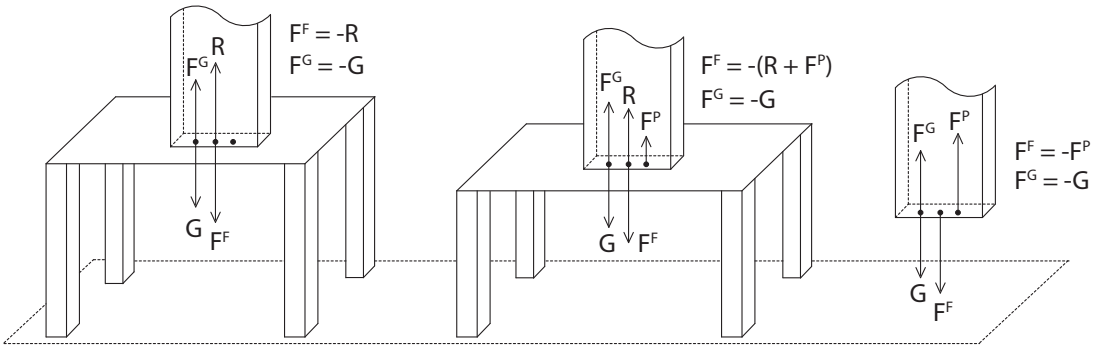

Figure 4: Three different examples: *(left)* with high table, *(middle)* with low table, *(right)* without table. The first case is the normal situation when the table has the same height as during the demonstrations. The second case is when the table is lowered, and the third case is when the table is completely removed. The forces acting on the end-effector of the robot in each of the three cases are shown. The force R is the reaction force of the table on the end-effector. In the second and the third case, the positional force F^P is trying to return the end-effector to the desired position, and is counteracting the exerted force F^F.

is the force exertion profile (only applied at the end-effector for $i = n$), and $F^G = [f^G, 0]^\top \in \mathbb{R}^6$ is the gravity compensation force. Tracking of a desired path in Cartesian space is ensured by the force command $f^P = m^P \hat{\ddot{x}}$, where m^P is a virtual mass and $\hat{\ddot{x}}$ is a desired acceleration command (described in next subsection).

An advantage of the proposed controller, regarding physical human-robot interaction, is that the robot remains gravity compensated not only during the teaching but also during the reproduction of the movement. After the task is learned, the robot is reproducing it based on the learned positional and force profiles for the end-effector only, which means that the particular robot joint configuration does not matter, as long as the positional constraint for the end-effector is satisfied. This allows the user to move freely the *redundant* DOFs during the reproduction, as demonstrated in [26]. The user can exploit both the robot's kinematic redundancy and the task's redundancy if necessary (e.g. to avoid a collision) by manually moving the robot's redundant DOFs (e.g. the elbow in the case of the Barrett WAM robot). This advantage does not relate to or affect the teaching, but matters only during the reproduction of the task.

The proposed controller is able to adapt to some changes in the environment, e.g. different heights of the object on which it exerts the force, such as a table. In such cases, the exerted force differs slightly from the demonstrated force. More specifically, if the new table is lower than the table used during the demonstration, then the system automatically adapts by lowering down the end-effector until it touches the table. In this case, the exerted force will be slightly lower than the demonstrated one, because the positional error will produce an upward force which will reduce the effect of the downward force used to apply a pressure on the table. Three example cases are shown in Fig. 4. The first case is the normal situation when the table has the same height as during the demonstrations. The second case is when the table is lower, and the third case is when the table is completely removed. Experimental results for these three cases are given in Section 3.2.

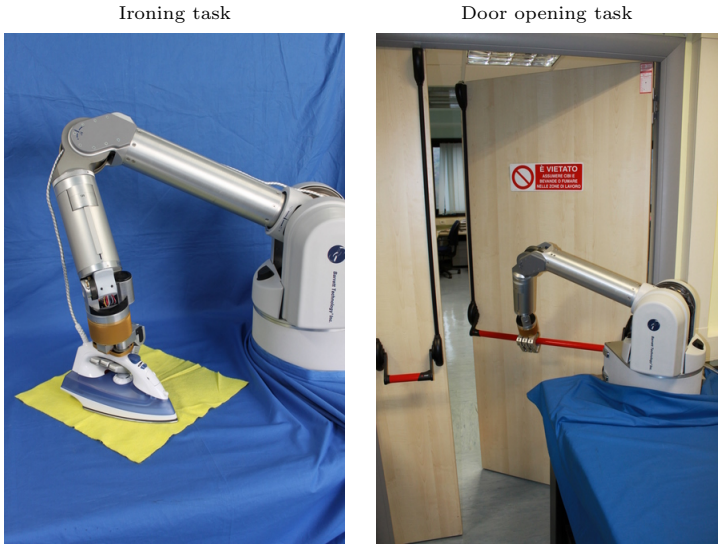

Figure 5: Experimental setup for the two tasks.

3 EXPERIMENTS

3.1 Experimental setup

The proposed method is evaluated on two human-robot interaction experiments, which aim to teach the robot to perform two tasks: *ironing task* and *door opening task*. The experimental setup for each of them is shown in Fig. 5. All experiments are conducted using a torque-controlled 7-DOF *Barrett WAM* robotic arm with 3-finger Barrett Hand attached, and a 7-DOF *Force Dimension Omega 7* haptic device.

The demonstrations needed for learning the positional profile of each task are recorded via kinesthetic teaching, i.e. a human demonstrator is holding the arm of the robot and manually guiding the robot to execute the task. During this, the WAM robot is put in a gravity-compensation mode which allows the user to move it effortlessly. The recorded trajectories are then encoded with the approach described in Section 2.1 to form the positional profile of the task.

After that, the learned task is reproduced on the robot based only on the positional profile. During this reproduction, a second demonstration phase takes place. The human demonstrator is holding the handle of the haptic device and exerting forces with variable magnitude and direction during the reproduction. The haptic device and the robot are linked in a master-slave system, in the sense that the force which the human exerts on the haptic device is being relayed to the robot's end-effector with the appropriate scaling defined in Section 2.2. In this way, the robot can execute the task semi-autonomously by following the already learned positional profile and combining it with the force input coming from the haptic device held by the human. The obtained set of force demonstrations is used to learn the force profile of the task. At this point, an autonomous reproduction of the task can be attempted by the robot using the extracted positional and force constraints.

Setup for measuring the exerted force

Numerical results

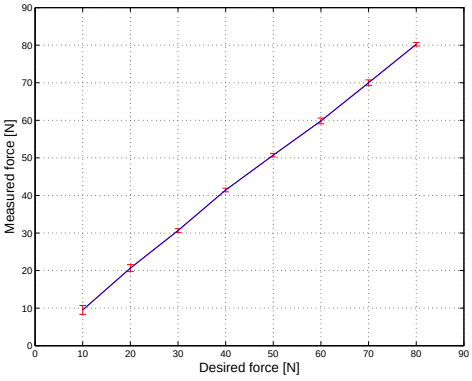

Figure 6: Measuring the actual exerted force when providing a desired force command. *Left:* Experimental setup. *Right:* Results from the measurements. In red, the error bars showing the standard deviation of the actual exerted force.

The two demonstration phases of the teaching process are shown in Fig. 8 for the ironing task, and in Fig. 12 for the door opening task. Each task is encoded using the proposed model, by fixing the number of states (or primitives) empirically with respect to the length of the demonstrations. For the ironing task, 6 demonstrations were provided for the positional profile, and 3 for the force profile. For the door opening task, 5 demonstrations were provided for the positional profile, and 5 for the force profile.

3.2 Force/position controller evaluation

Two sets of preliminary experiments are conducted to evaluate the proposed force/position controller introduced in Section 2.3. The first experiment aims to evaluate the precision of force exertion in the static case, while the second set of experiments evaluates the dynamic tracking of a force profile in time.

For the first experiment, the robot is instructed to push down with its hand vertically on top of weighting scales as shown in Fig. 6. This is done using the same described controller by manually setting the force f^F in Eq. (9) to be a constant vector equal to the desired force. Then, the produced force is measured by the scales and compared with the desired force. The experimental results show that the measured forces match very closely the desired forces. The small errors are caused by inaccuracies in the gravity compensation controller (due to friction or precision of motor torque commands) but they do not affect the successful task execution. This experiment demonstrates the fitness of the proposed control strategy to produce a force with a desired magnitude using a torque-controlled robot without force sensing.

For the second set of experiments, a dynamic reference force profile is used, defined as:

$$F_3 = -A\Big[1 + e^{-t/6} sin(\pi t - \pi/2)\Big], \tag{10}$$

where t is the time, and the amplitude $A = 25$ N. This reference force profile is shown in Fig. 7, and compared to the measured actual exerted force by the robot. To record the dynamically changing

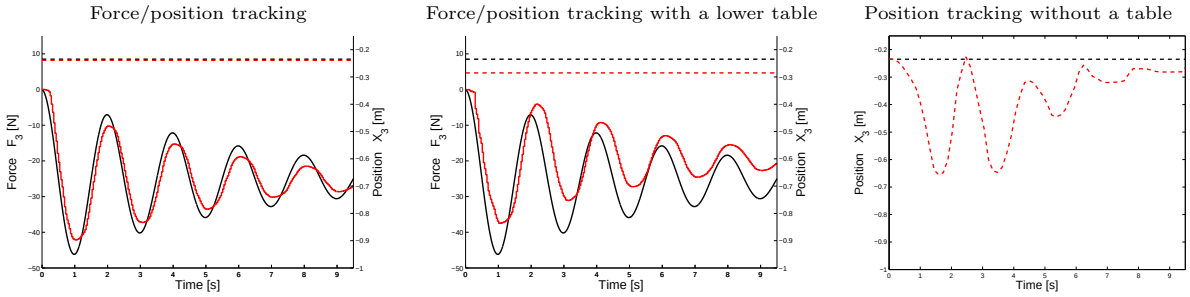

Figure 7: Evaluation of dynamic force/position tracking of the proposed controller. The three plots correspond to the three cases depicted in Fig. 4. The reference force profile is shown with a black solid line, and the measured actual exerted force - with red lines. Similarly, the positional profile and the measured actual positions are shown with black and red dashed lines. For all three cases, the reference vertical position is fixed (at $X_3 = -0.235$ m), but the actual height of the table is changing. *Left:* In this case, the table height is the same as the desired position. Thus, the calculated force $F^P = 0$, which allows the controller to track the force profile very closely to the desired profile. *Middle:* In this case, the table is lower than the desired vertical position with 5 cm. As a result, a force $F^P > 0$ is calculated to compensate for the positional error, and this affects the force tracking by a slight reduction in the measured actual exerted force. This shows how the proposed controller adapts to small uncertainties and changes in the environment. *Right:* In this extreme case, the table is completely removed. There are no force measurements, because the robot's end-effector is not in contact with any object. Instead, the position of the end-effector is shown, demonstrating the influence of the force tracking to the position tracking. Due to the mixed influence of both positional and force contraints to the controller, the end-effector does not collide with the floor, but is withheld by the upward force F^P. This shows how the proposed controller deals with a drastic change in the environment, as well as with conflicting positional and force profiles.

exerted force, an external force/torque sensor ATI Mini45 SI-580-20 is used. The F/T sensor is placed on the surface of the table and the robot is exerting a variable force on top of it during the experiments. Three different cases were evaluated, matching the three cases depicted in Fig. 4. The purpose of these experiments is two-fold: first, to evaluate the dynamic force tracking capabilities of the controller, and second, to assess the adaptation capabilities of the mixed force/position controller in presence of changes in the environment. The experiments also demonstrate how the controller deals with conflicting position- and force-constraints, in cases when the two profiles are not consistent with each other. The experimental results from the three cases are shown and explained in Fig. 7.

This adaptability is achieved without having to explicitly define priorities of position- and force-tracking. Instead, the input data collected during the demonstrations is used to calculate variable stiffness which later regulates the relative priority of position- and force-tracking. For example, if all the demonstrations show consistency in the positional profile, this leads to high stiffness for the position tracking, and gives higher influence of position over force tracking. On the other hand, if the demonstrations exhibit highly-variable positional profile, the resulting low positional stiffness yields bigger influence to the force tracking. This allows the human teacher to implicitly control the relative priority of position- and force-tracking for the controller.

3.3 Ironing task

The ironing task poses a wide variety of challenges for robotics. These challenges range from the detection, grasping and folding of clothes [27, 28], to position and orientation analysis of the iron depending

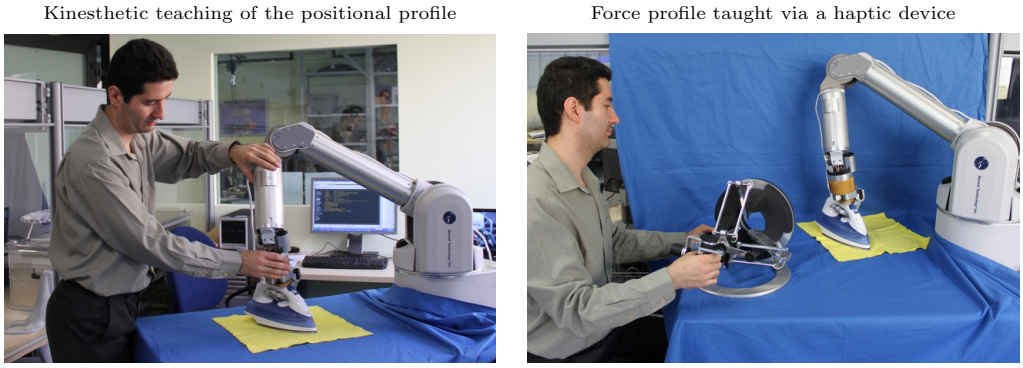

Figure 8: Teaching the positional and the force profiles of the ironing task.

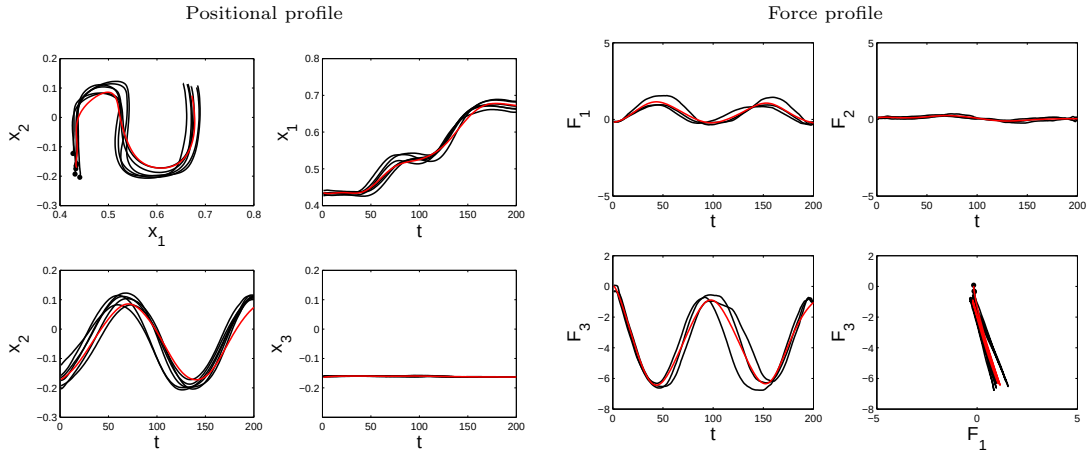

Figure 9: Positional and force profiles of the ironing task. Demonstrations (in black) and separate reproductions (in red) based on the learned individual profiles. The position axis x_1 and x_2 are in the horizontal plane, and x_3 is the vertical axis. Similarly, the force axis F_1 and F_2 are in the horizontal plane, and F_3 is the vertical axis.

on the regions of garment and on the proximity to an operator [29]. Here we concentrate on the simultaneous consideration of force and position information for the ironing skill. For example, while the iron is being moved, it should be pushed down towards the table top, to press the clothes well. In addition, the exerted force on the iron should be stronger when the tip of the iron is moving forward, and weaker when the iron is moving backward, to avoid creating wrinkles on the clothes.

3.3.1 Learning the task profiles

The ironing skill imposes varying positional constraints throughout the execution of the task. For example, the path that the iron should follow is more constrained in the vertical axis than in the horizontal plane because it is more important to have the iron in contact with the table than to follow a specific path on the table. These varying positional constraints are reflected by the collected data from the demonstrations and incorporated in the task profile.

Regarding the force profile, we would like the robot to push down on the iron with varying magnitude during different parts of the S-shaped ironing trajectory. This pushing behavior cannot be emulated

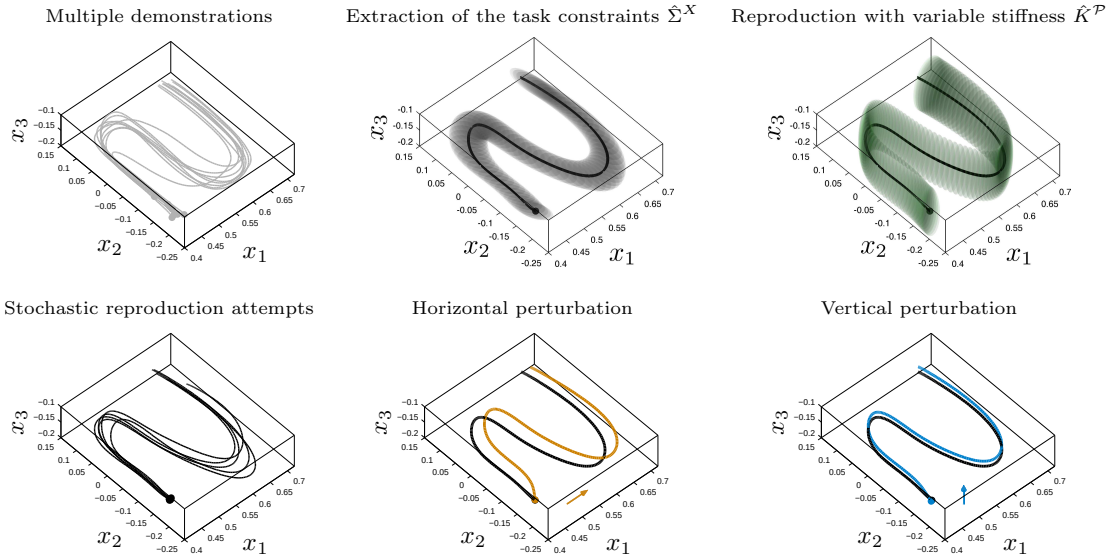

Figure 10: Learned positional profile for the ironing task. *Top:* Extracting task constraints from demonstrations through the residuals of the regression process. Adaptive stiffness gain matrix computed from the residuals information. The covariance matrices $\hat{\Sigma}^X = \sum_{i=1}^{K} h_i(t)\Sigma_i^X$ and stiffness gain matrices $\hat{K}^{\mathcal{P}} = \sum_{i=1}^{K} h_i(t)K_i^{\mathcal{P}}$ are represented with grey and green ellipsoids respectively. *Bottom:* Stochastic reproductions of the movement. Two perturbed reproduction attempts: one with applying a constant force parallel to the table at the end-effector of the robot, and another with applying a constant force with the same amplitude, but perpendicular to the table.

only with position control, because during the pushing action there is no positional change (the robot's end-effector is staying at the same height above the ironing board), and only the force exerted by the robot on top of the surface changes.

Fig. 9 shows the obtained positional and force profiles of the ironing task after the teaching process. Fig. 10 demonstrates the ability of the proposed method to encode the variable positional task constraints and to use adaptive stiffness gain matrix to perform adequate task reproductions. In particular, the learned model shows that it is more important to track the movement in the vertical direction than in the other two directions of the horizontal plane (nearly flat ellipsoids in the second graph). As a result, the stiffness matrices have an elongated shape reflecting this property.

The two simulated perturbation experiments demonstrate that the robot learned to keep its end-effector stiff in vertical direction and compliant (soft) in horizontal direction. As a result, when applying a constant force to the end-effector during reproduction, the deformation is stronger if the force is parallel to the table than if the force is vertical. The covariance matrices $\hat{\Sigma}^X$ and stiffness gain matrices $K^{\mathcal{P}}$ are respectively depicted with grey and green ellipsoids in Fig. 10. The variability of the demonstrations reflects the important characteristics of the task. The model correctly encapsulates these variations through the set of covariance matrices $\hat{\Sigma}^X$ estimated through the residuals of the regression process. A variable stiffness gain matrix $\hat{K}^{\mathcal{P}}$ is automatically calculated in order to fulfill the learned task constraints.

The combined positional and force profiles for the ironing task are shown in Fig. 11. The lengths of the arrows correspond to the magnitude of the exerted forces. We see that the robot correctly reproduces

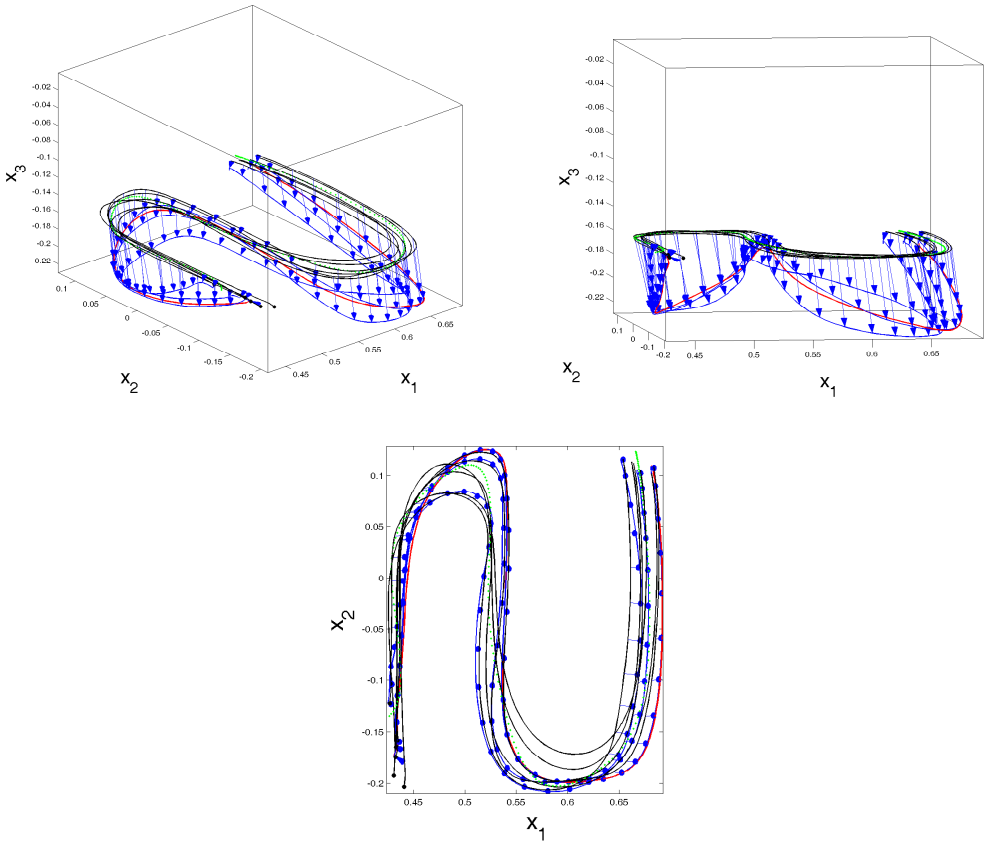

Figure 11: Combined positional and force profiles for the ironing task. Positional demonstrations (with black lines), force demonstrations (with blue arrows), and positional reproduction (with green dotted lines). For better visibility, blue lines show the trajectories swept by the blue arrow tips (i.e. the demonstrated force along the trajectory). The red line shows the same trajectory as reproduced by the combined model.

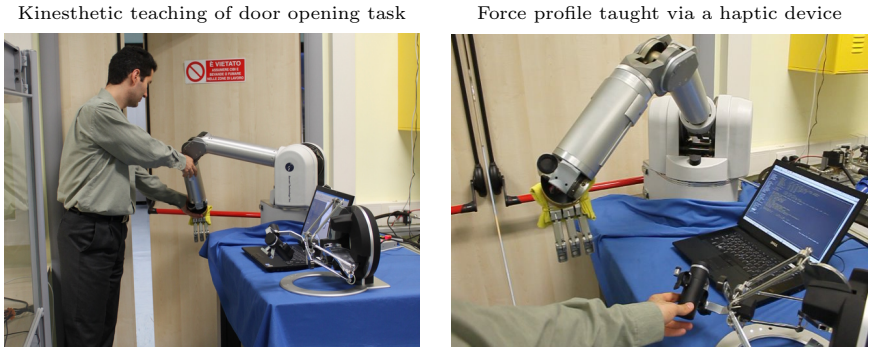

Figure 12: Learning positional and force profiles for the door opening task.

the skill by applying vertical forces of varying amplitude during the movement while following the path. In order to quantify the forces exerted by the robot, we conducted an additional experiment.

3.4 Door opening task

The second task consists of opening a door which has a horizontal bar that needs to be pushed with a larger force than a standard door handle to open it, as shown in Fig. 5. In this task, the robot first has to push with its hand the horizontal bar downwards (which unlocks the door) and then to push forward and to the right in an arc-shaped trajectory to open the door wider.

By using the proposed compliant control scheme, if f^F is removed from Eq. (9), the results are not satisfactory, because the variable stiffness gains extracted during the first phase of the demonstration do not allow to open the door with the stiffness required to do so. Instead, it makes the robot pause during the movement when it reaches the bar, and delays it until enough force is accumulated due to the increasing positional error, after which the rest of the motion is performed hastily to catch up with the movement. This problem is solved by the proposed method through the use of the force profile.

3.4.1 Learning the task profiles

The teaching process for the door opening task is shown in Fig. 12. The individual positional and force profiles obtained from the demonstrations are shown in Fig. 13.

Regarding the positional profile of the task, the position of the hand needs to be more constrained when it pushes on the handle than during its approach, in order to bring it to the desired position of the horizontal bar. These varying positional constraints are reflected by the collected data and learned by the proposed method.

Regarding the force profile of the task, there is one noticeable peak in the force exerted when the hand of the robot is pushing the door handle. There is a strong correlation between the vertical component of the force (f_1, parallel to the door) and the forward horizontal component (f_3, perpendicular to the closed door), which is required to push the handle (see bottom-right plot in Fig. 13).

Fig. 14 shows the combined positional and force profile for the door opening task. Fig. 15 shows

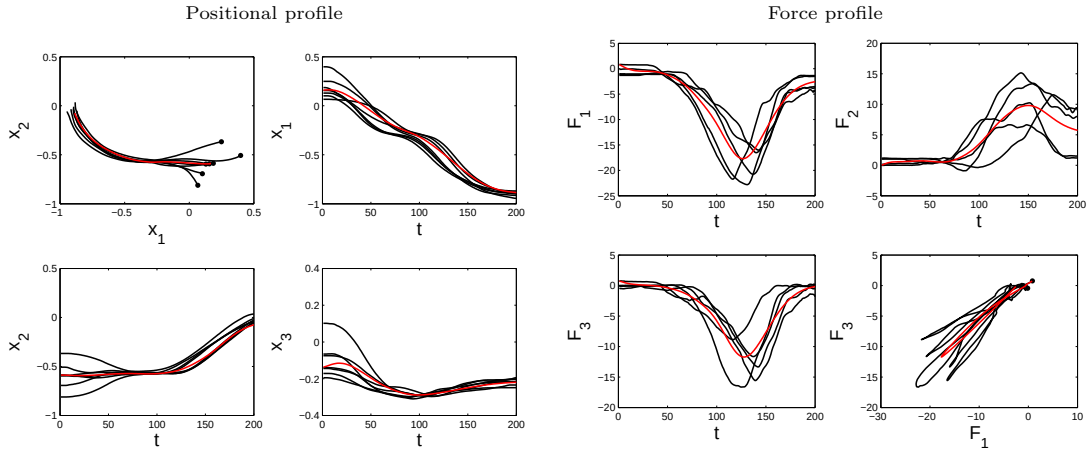

Figure 13: Positional and force profiles of the door opening task. Demonstrations (in black) and reproductions (in red) from the learned model. The dots show the starting positions of the trajectories.

snapshots from the autonomous reproduction of the task.

4 DISCUSSION

The two presented experiments demonstrate the ability of the proposed method to successfully encode and reproduce both positional and force profiles of tasks posing quite different requirements. For example, the ironing task requires variable stiffness in vertical and horizontal direction during the reproduction. Also, it requires exerting a variable force with constant direction on an object which does not move, regardless of the exerted force (the table top in this case). The door opening task, on the other hand, focuses on exerting a large force at a particular part of the trajectory, and also varying the direction of the force during the task reproduction. In both cases, the tasks are difficult to realize using only position information, which shows the advantage of incorporating force information in the task model.

Combining positional and force profiles through separate sequential demonstrations is a first milestone towards learning complex haptic skills from demonstration. To increase the user-friendliness of such HRI interfaces, we plan to extend in future work the proposed framework to the use of demonstrations conveying simultaneously position and force information. One possible way to achieve this is to equip the robot's arm with force/torque sensors which can sense the exerted forces simultaneously with the kinesthetic teaching for the positional profile. This would offer a more intuitive and user-friendly way to teach the force profile compared to using an additional external haptic interface.

The robot remains gravity-compensated throughout all the interactions presented in the paper for both demonstration and reproduction phases. During reproduction, this allows the user to physically move the robot using the redundancy of the robot's structure and the redundancy of the task. The same property can be exploited to improve the safety of the human-robot interaction by using the available redundancy to avoid hitting obstacles or other people close to the robot.

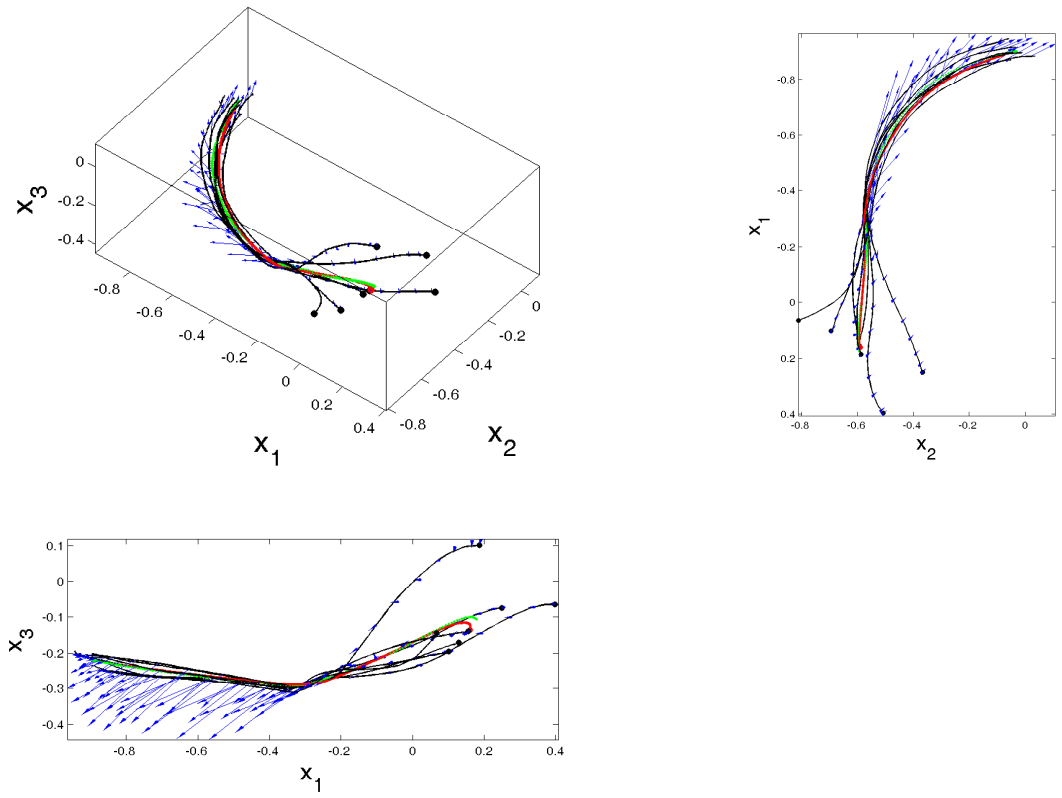

Figure 14: Combined positional and force profile for the door opening task. The black lines show the original demonstrations. The blue arrows represent the demonstrated exerted force for the end-effector along the trajectories. The green line represents the reproduction of the positional profile during the force profile teaching phase. The red line shows the final positional reproduction from the learned combined model.

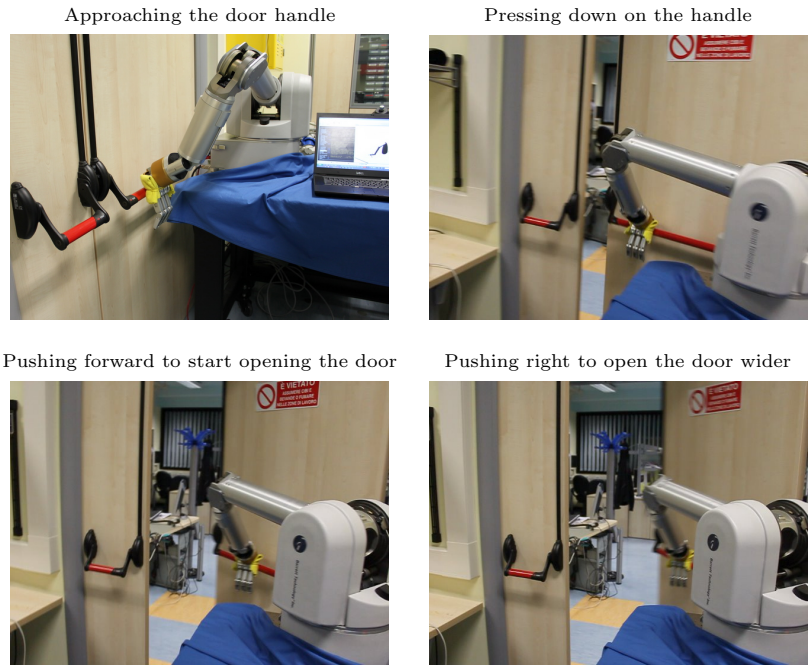

Figure 15: Standalone reproduction attempt of the door opening task.

In the presented experiments, during the teaching process and the reproduction process, the orientation of the end-effector was kept constant. The proposed method, however, can be extended to also encode the orientation of the end-effector. We plan to use this in future research to create orientational profiles of tasks which require rotation of the manipulated objects. Similarly, we plan to extend our method to the demonstration of moment of force to allow the robot to execute tasks such as unscrewing a bolt.

5 CONCLUSION

The presented method allows a person to transfer skills requiring simultaneous consideration of positional and force information to a robot.

We proposed a two-steps procedure where position information is first learned through kinesthetic teaching, and further enhanced by force information learned through the use of a haptic device. Positional and force profiles are encoded in a compact model based on a mixture of dynamical systems, which encapsulates variation and correlation information.

For reproduction, we proposed an active control strategy based on task-space control with variable stiffness, and a force exertion strategy allowing the robot to exert forces on external objects. The user teaches the robot how to perform a task by providing multiple demonstrations of the skill he/she wishes to transfer. After extraction of the task constraints, the robot reproduces the task by automatically selecting an adequate level of compliance to reproduce the essential positional and force characteristics of the skill. We demonstrated the feasibility of the approach with an ironing task and a door opening task.

REFERENCES

[1] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Secaucus, NJ, USA: Springer, 2008, pp. 1371–1394.

[2] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robot. Auton. Syst.*, vol. 57, no. 5, pp. 469–483, 2009.

[3] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, ser. Adaptive computation and machine learning. Cambridge, MA, USA: MIT Press, 1998.

[4] T. Salter, K. Dautenhahn, and R. te Boekhorst, "Learning about natural human-robot interaction styles," *Robotics and Autonomous Systems*, vol. 54, no. 2, pp. 127–134, 2006.

[5] A. L. Thomaz, M. Berlin, and C. Breazeal, "An embodied computational model of social referencing," in *Proc. IEEE Intl Workshop on Robot and Human Interactive Communication (Ro-Man)*, Nashville, TN, USA, August 2005, pp. 591–598.

[6] K. J. Rohlfing, J. Fritsch, B. Wrede, and T. Jungmann, "How can multimodal cues from child-directed interaction reduce learning complexity in robots?" *Advanced Robotics*, vol. 20, no. 10, pp. 1183–1199, 2006.

[7] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, "An atlas of physical human-robot interaction," *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, 2008.

[8] P. Kormushev, S. Calinon, and D. G. Caldwell, "Robot motor skill coordination with EM-based reinforcement learning," in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, October 2010, pp. 3232–3237.

[9] P. Zukow-Goldring, "Caregivers and the education of the mirror system," in *Proc. Intl Conf. on Development and Learning (ICDL)*, La Jolla, California, USA, 2004, pp. 96–103.

[10] N. J. Hodges and I. M. Franks, "Modelling coaching practice: the role of instruction and demonstration," *Sports Sciences*, vol. 20, no. 10, pp. 793–811, 2002.

[11] R. Horn and A. M. Williams, "Observational learning: Is it time we took another look?" in *Skill Acquisition in Sport: Research, Theory and Practice*, A. M. Williams and N. J. Hodges, Eds. London, UK: Routledge, 2004, pp. 175–206.

[12] N. Otero, J. Saunders, K. Dautenhahn, and C. L. Nehaniv, "Teaching robot companions: the role of scaffolding and event structuring," *Connection Science*, vol. 20, no. 2-3, pp. 111–134, 2008.

[13] S. Calinon and A. Billard, "Statistical learning by imitation of competing constraints in joint space and task space," *Advanced Robotics*, vol. 23, no. 15, pp. 2059–2076, 2009.

[14] P. Evrard, E. Gribovskaya, S. Calinon, A. Billard, and A. Kheddar, "Teaching a collaborative task using teleoperation," in *Proc. IEEE-RAS Intl Conf. on Humanoid Robots (Humanoids)*, Paris, France, December 2009, pp. 399–404.

[15] A. Albu-Schaeffer, S. Haddadin, C. Ott, A. Stemmer, T. Wimboeck, and G. Hirzinger, "The DLR lightweight robot - Design and control concepts in human environments," *Industrial Robot*, vol. 34, no. 5, pp. 376–385, 2007.

[16] A. De Luca and L. Ferrajoli, "Exploiting robot redundancy in collision detection and reaction," in *IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, Nice, France, September 2008.

[17] J.-O. Kim, M. Wayne, and P. K. Khosla, "Exploiting redundancy to reduce impact force," *Journal of Intelligent and Robotic Systems*, vol. 9, no. 3, pp. 273–290, 1994.

[18] M. Howard, S. Klanke, M. Gienger, C. Goerick, and S. Vijayakumar, "Methods for learning control policies from variable-constraint demonstrations," in *From Motor Learning to Interaction Learning in Robots*, O. Sigaud and J. Peters, Eds. Springer Berlin / Heidelberg, 2010, pp. 253–291.

[19] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Trajectory formation for imitation with nonlinear dynamical systems," in *Proc. IEEE Intl Conf. on Intelligent Robots and Systems (IROS)*, Maui, USA, 2001, pp. 752–757.

[20] S. Schaal, P. Mohajerian, and A. J. Ijspeert, "Dynamics systems vs. optimal control a unifying view," *Progress in Brain Research*, vol. 165, pp. 425–445, 2007.

[21] H. Hoffmann, P. Pastor, D. H. Park, and S. Schaal, "Biologically-inspired dynamical systems for movement generation: automatic real-time goal adaptation and obstacle avoidance," in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, Kobe, Japan, May 2009, pp. 2587–2592.

[22] S. Calinon, F. D'halluin, D. G. Caldwell, and A. G. Billard, "Handling of multiple constraints and motion alternatives in a robot programming by demonstration framework," in *Proc. IEEE-RAS Intl Conf. on Humanoid Robots (Humanoids)*, Paris, France, December 2009, pp. 582–588.

[23] S. Calinon, F. D'halluin, E. L. Sauser, D. G. Caldwell, and A. G. Billard, "Learning and reproduction of gestures by imitation: An approach based on hidden Markov model and Gaussian mixture regression," *IEEE Robotics and Automation Magazine*, vol. 17, no. 2, pp. 44–54, 2010.

[24] M. Khansari and A. G. Billard, "BM: An iterative method to learn stable non-linear dynamical systems with Gaussian mixture models," in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, Anchorage, Alaska, USA, May 2010, pp. 2381–2388.

[25] R. Featherstone and D. E. Orin, "Dynamics," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Secaucus, NJ, USA: Springer, 2008, pp. 35–65.

[26] S. Calinon, I. Sardellitti, and D. G. Caldwell, "Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies," in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, October 2010, pp. 249–254.

[27] J. S. Dai, P. M. Taylor, H. Liu, and H. Lin, "Folding algorithms and mechanisms synthesis for robotic ironing," *International Journal of Clothing Science and Technology*, vol. 16, pp. 204–214, 2004.

[28] J. Maitin-Shepard, M. Cusumano-Towner, J. Lei, and P. Abbeel, "Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding," in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, Anchorage, Alaska, USA, May 2010, pp. 2308–2315.

[29] J. S. Dai, P. M. Taylor, P. Sanguanpiyapan, and H. Lin, "Trajectory and orientation analysis of the ironing process for robotic automation," *Intl Journal of Clothing Science and Technology*, vol. 16, no. 1–2, pp. 215–226, 2004.