

Learning In-Contact Control Strategies from Demonstration

Mattia Racca¹, Joni Pajarinen², Alberto Montebelli³ and Ville Kyrki¹

Abstract—Learning to perform tasks like pulling a door handle or pushing a button, inherently easy for a human, can be surprisingly difficult for a robot. A crucial problem in these kinds of in-contact tasks is the context specificity of pose and force requirements. In this paper, a robot learns in-contact tasks from human kinesthetic demonstrations. To address the need to balance between the position and force constraints, we propose a model based on the hidden semi-Markov model (HSMM) and Cartesian impedance control. The model captures uncertainty over time and space and allows the robot to smoothly satisfy a task’s position and force constraints by online modulation of impedance controller stiffness according to the HSMM state belief. In experiments, a KUKA LWR 4+ robotic arm equipped with a force/torque sensor at the wrist successfully learns from human demonstrations how to pull a door handle and push a button.

I. INTRODUCTION

One of the most alluring practical consequences of allowing robots to learn directly from demonstration is to relax the analytical burden required to transfer physical human skills to robotic platforms. This learning process is known in literature as *programming by demonstration* (PbD), learning from demonstration (LfD), or imitation learning [1], [2]. In this paper, we use the term *in-contact tasks* for tasks that crucially involve the distribution of forces in space and time at the interface between the robot’s end-effector and its surrounding world. In-contact tasks within the domain of human mastery are typically difficult to explicitly encode into declarative terms.

Common interfaces for demonstration in PbD range from direct recording of human actions to immersive teleoperation [3], for example making use of haptic devices capable of providing position and force information. Among them, with *kinesthetic teaching* the human instructor can physically grab and displace the robot’s body parts and operate its tool in space and time. As discussed in [4], the mastered skill can be demonstrated with minimal alteration of the geometric and dynamic properties of the task on an appropriately designed setup.

The approach presented in this paper utilizes kinesthetic teaching, simultaneously collecting position and force profiles of a task as exemplified in Fig. 1. We use a compliant

*This work was supported by Academy of Finland, decisions 264239, 271394 and 292980.

¹M. Racca and V. Kyrki are with School of Electrical Engineering, Aalto University, Finland mattia.racca@aalto.fi, ville.kyrki@aalto.fi

²J. Pajarinen is with the Intelligent Autonomous Systems (IAS) and Computational Learning for Autonomous Systems (CLAS) labs at TU Darmstadt pajarin@ias.tu-darmstadt.de

³A. Montebelli is with Department of Information Technology, University of Skövde, Sweden alberto.montebelli@his.se

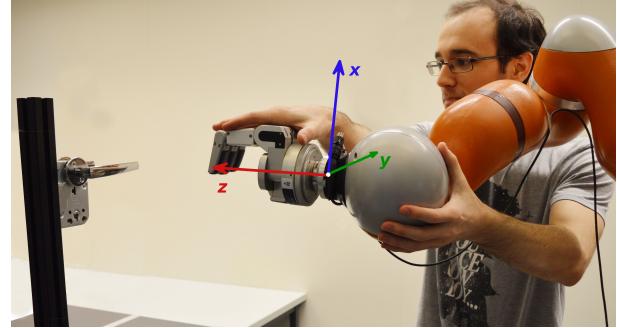


Fig. 1: Learning to pull a door handle from human demonstration with simultaneous teaching of pose and force requirements.

robot with a force-torque sensor placed between the robot flange and the tool [4], [5], as depicted in Fig. 2. This allows a human instructor to perform intuitive and familiar task executions, avoiding separate recordings of position and force. Furthermore, the spatio-temporal correlation between position and force information is preserved and the use of external force recording devices avoided.

In order to model the dynamics of in-contact tasks, we extend the PbD approach of [6], [7], which uses a combination of Hidden Semi-Markov Model (HSMM) and Gaussian Mixture Regression (GMR). In addition to the spatial and temporal constraints, our HSMM encodes the *force profiles* recorded during the demonstrations. The state representation of the HSMM allows us to smoothly modulate the stiffness of a Cartesian impedance controller during the reproduction, based on the belief over the states. The varying impedance of the controller allows the robot to satisfy the position and force constraints observed during the teaching as well as retain safety thanks to the low stiffness interface.

The following Section II presents related work on in-contact PbD. Sect. III describes in detail the mathematical model used for learning the demonstrated task and the stiffness selection mechanism. Next, the experimental setup is presented and the results discussed in Sect. IV. Finally, conclusions are drawn in Sect. V.

II. RELATED WORK

PbD is an effective way for transferring human physical skills to robots. However, when the success of the taught tasks strictly depends on the physical interaction of the robot with the environment, force constraints have to be taken into account. Including the force information in a PbD approach requires modifications to the *teaching*, the *learning* and the *reproduction* phases.

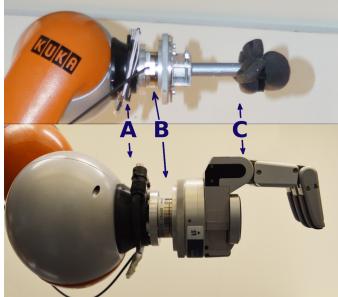


Fig. 2: End-effector configuration: the F/T sensor (B) is mounted between the robot flange (A) and the tool (C).

During the demonstrations, the teacher has to provide a meaningful reproduction of the force profiles. We use *simultaneous kinesthetic teaching* of pose and force constraints [4], [5], based on the end-effector configuration shown in Fig. 2. Rozo et al. [8] exploit a similar configuration to teach an impedance-based robot to adapt its stiffness from demonstration during a collaborative assembly task.

For the modeling of the motion along with the force information of a task, the Dynamic Movement Primitive (DMP) approach is the most widely used [4], [5], [9] although also statistical models have been used [10], [11]. DMP models encode each dimension separately, allowing the learning of force profiles without changes to the basic approach. However, modeling the dimensions independently does not allow to exploit the correlation between the different variables. The model-free Probabilistic Movement Primitive (ProMP) approach presented in [11] captures the correlation of the trajectory, the commands and the sensory signal. However, since the commands are part of the training, this approach requires the robot to be teleoperated during the demonstrations. Kinesthetic teaching, where the unrecorded external forces of the teacher replace the robot commands, is then not directly applicable, narrowing down the choice of tasks.

Calinon et al. [6] propose the combination of Hidden Markov Model (HMM) and GMR to create a probabilistic model of demonstrated trajectories, by modeling a joint-probability density function over the data, and to generalize the taught task via regression. To further encode a duration information for each state, the use of HSMM in PbD is introduced in [7]. HMM-based approaches allow encoding of several motions in the same model [6], a unified way of handling periodic and non-periodic movements (DMP [12] and ProMP [13] require a preliminary choice) and robustness against non-aligned demonstrations. Inspired by these works and the good performance of GMR-based methods in the literature [6], [14], [15], our work extends the HSMM-GMR formulation to include the force information in the learning, bringing then the strengths of HMM-based approaches to the case of in-contact tasks. Morevoer, the choice of HSMM allows us to perform *state-specific control strategies*.

Since interaction with the environment is required for in-contact tasks, compliant control strategies like joint torque

control [9], [10] or impedance control [4], [5] are usually coupled with the models. However, as pointed out in [10] and confirmed by our experiments, simply playing back the learned forces is not sufficient to successfully accomplish tasks. Thus, a decision on which aspect of the motion (the kinematic trajectory or the exertion of forces) to prioritize needs to be made.

Our approach builds on a Cartesian impedance controller whose stiffness parameters are changed during the execution of the motion, based on the force information encapsulated in each HSMM state. Instead of looking at the variance of the demonstrations [9]–[11], we extract the two basic modes of interaction given the recorded force: *in-contact* with the environment and *not in-contact*. Our approach can then choose automatically which aspect of the motion to prioritize, avoiding manually set thresholds [4], [5]. Thanks to the adjustable stiffness interface, the proposed approach can cope with moderate untrained perturbations in the environment as in [5], [9], [10] and allow safe operation close to a human.

III. METHOD

We propose a HSMM-GMR based PbD approach able to model the pose and force information of in-contact tasks from demonstration. We combine the model with a Cartesian impedance controller with stiffness that is dynamically adjusted during the reproduction based on the belief over the HSMM states.

A. Model

Our model is based on the HSMM-GMR model applied to PbD as proposed by Calinon et al. [6], [7]. In order to address the case of in-contact tasks, our model takes into account forces recorded during the demonstrations.

The learning framework uses an HSMM to model the temporal evolution of trajectories with a continuous Gaussian observation probability distribution assigned to each HSMM state. Offline, we use a version of the Baum-Welch algorithm [16] to learn the model parameters from a set of D demonstrated trajectories. Online, we utilize GMR [17] at each timestep to compute the desired twist (positional and rotational velocity) and force.

We will now define the parameters of our N state HSMM and discuss the model and how the parameters are used in more detail. We train the model using a set of D demonstrations. Each demonstration $d \in \{1, \dots, D\}$ consists of a set of T_{max} samples describing the pose and forces sensed at the end-effector of the robot. The N state HSMM is parametrized by

$$\lambda = (\pi, A, \mu, \Sigma, \mu^D, \sigma^D), \quad (1)$$

where π specifies the initial probability of each state, A is the transition probability matrix where $a_{ji} \in A$ is the transition probability from state j to i , $\mu = (\mu_1, \dots, \mu_N)$ and $\Sigma = (\Sigma_1, \dots, \Sigma_N)$ are sets of mean vectors and covariance matrices, respectively, of the N Gaussian joint observation probabilities, and $\mu^D = (\mu_1^D, \dots, \mu_N^D)$ and

$\sigma^D = (\sigma_1^D, \dots, \sigma_N^D)$ are sets of mean values and variances, respectively, of the N Gaussian duration probability distributions. The state duration pdf $p_i^D(t)$ of state i is defined as

$$p_i^D(t) = \mathcal{N}(t; \mu_i^D, \sigma_i^D), \quad (2)$$

with $t = 1, \dots, t_{max}$. In the experiments, we computed the maximum allowed duration of a state t_{max} from the length of the demonstrations by

$$t_{max} = \gamma \left\lceil \frac{T_{max}}{N} \right\rceil, \quad (3)$$

with γ manually set to 2 allowing each state to last up to twice the average duration.

The observation probability at time t for state i is defined as

$$p_i(\mathbf{z}_t) = \mathcal{N}(\mathbf{z}_t; \boldsymbol{\mu}_i, \Sigma_i), \quad (4)$$

where $\mathbf{z}_t = [\mathbf{x}_t^T \dot{\mathbf{x}}_t^T \mathbf{q}_t^T \dot{\mathbf{q}}_t^T \mathbf{f}_t^T]^T$ is the concatenation of the observed variables at time t i.e. position \mathbf{x}_t and velocity $\dot{\mathbf{x}}_t$ of the end-effector, quaternion representation \mathbf{q}_t of the orientation of the tool and its derivative $\dot{\mathbf{q}}_t$ plus the novel addition of the measured force \mathbf{f}_t . The mean vector $\boldsymbol{\mu}_i \in \boldsymbol{\mu}$ and covariance matrix $\Sigma_i \in \Sigma$, characterized as

$$\boldsymbol{\mu}_i = \begin{bmatrix} \boldsymbol{\mu}_i^x \\ \boldsymbol{\mu}_i^{\dot{x}} \\ \boldsymbol{\mu}_i^q \\ \boldsymbol{\mu}_i^{\dot{q}} \\ \boldsymbol{\mu}_i^f \end{bmatrix}, \quad \Sigma_i = \begin{bmatrix} \Sigma_i^{xx} & \Sigma_i^{x\dot{x}} & \Sigma_i^{xq} & \Sigma_i^{x\dot{q}} & \Sigma_i^{xf} \\ \Sigma_i^{\dot{x}x} & \Sigma_i^{\dot{x}\dot{x}} & \Sigma_i^{\dot{x}q} & \Sigma_i^{\dot{x}\dot{q}} & \Sigma_i^{\dot{x}f} \\ \Sigma_i^{qx} & \Sigma_i^{q\dot{x}} & \Sigma_i^{qq} & \Sigma_i^{q\dot{q}} & \Sigma_i^{qf} \\ \Sigma_i^{q\dot{x}} & \Sigma_i^{q\dot{x}} & \Sigma_i^{q\dot{q}} & \Sigma_i^{q\dot{q}} & \Sigma_i^{qf} \\ \Sigma_i^{fx} & \Sigma_i^{f\dot{x}} & \Sigma_i^{fq} & \Sigma_i^{f\dot{q}} & \Sigma_i^{ff} \end{bmatrix}, \quad (5)$$

parameterize the Gaussian observation probability for each state i . We optimize the parameters λ offline over the demonstrations using a version of the Baum-Welch algorithm [16].

During the reproduction of the demonstrated task, we update the belief distribution over HSMM states and use this distribution to compute the desired control parameters: twist and force. The forward variable $\alpha_{i,t}$ specifies the unnormalized probability for the system to be in state i at time t . The belief over states $h_{i,t} = \alpha_{i,t} / \sum_{k=1}^N \alpha_{k,t}$ is a normalized version of the forward variable. The forward variable of state i is initialized as

$$\alpha_{i,1} = \pi_i \mathcal{Q}_i \left(\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{q}_1 \end{bmatrix} \right) \quad (6)$$

and depends on the priors π_i and on the Gaussian probability

$$\mathcal{Q}_i \left(\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{q}_1 \end{bmatrix} \right) = \mathcal{N} \left(\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{q}_1 \end{bmatrix}; \begin{bmatrix} \boldsymbol{\mu}_i^x \\ \boldsymbol{\mu}_i^q \end{bmatrix}, \begin{bmatrix} \Sigma_i^{xx} & \Sigma_i^{xq} \\ \Sigma_i^{qx} & \Sigma_i^{qq} \end{bmatrix} \right) \quad (7)$$

of the starting pose $[\mathbf{x}_1^T \mathbf{q}_1^T]^T$. Note that the forward variable does not directly depend on the observed force or velocities. This means that during online operation the evolution of states depends only on position and orientation,

not force. At each timestep we update the forward variable with

$$\alpha_{i,t} = \sum_{j=1}^N \sum_{d=1}^{\min(t_{max}, t-1)} \alpha_{j,t-d} a_{ji} p_i^D(d) \prod_{s=t-d+1}^t \mathcal{Q}_i \left(\begin{bmatrix} \mathbf{x}_s \\ \mathbf{q}_s \end{bmatrix} \right). \quad (8)$$

The control variables (the velocity of the end-effector $\dot{\mathbf{x}}_t^*$, the derivative of the orientation $\dot{\mathbf{q}}_t^*$ and the desired exerted force \mathbf{f}_t^*) are determined online as their expectations over the current state distribution $h_{i,t}$ given the current position and orientation

$$\dot{\mathbf{x}}_t^* = \sum_{i=1}^N h_{i,t} [\boldsymbol{\mu}_i^{\dot{x}} + \Sigma_i^{\dot{x}\dot{x}} (\Sigma_i^{xx})^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_i^x)], \quad (9)$$

$$\dot{\mathbf{q}}_t^* = \sum_{i=1}^N h_{i,t} [\boldsymbol{\mu}_i^{\dot{q}} + \Sigma_i^{\dot{q}\dot{q}} (\Sigma_i^{qq})^{-1} (\mathbf{q}_t - \boldsymbol{\mu}_i^q)], \quad (10)$$

$$\begin{aligned} \mathbf{f}_t^* = \sum_{i=1}^N h_{i,t} [\boldsymbol{\mu}_i^f + \Sigma_i^{f\dot{q}} (\Sigma_i^{qq})^{-1} (\mathbf{q}_t - \boldsymbol{\mu}_i^q) \\ + \Sigma_i^{fx} (\Sigma_i^{xx})^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_i^x)]. \end{aligned} \quad (11)$$

The desired velocity of the end-effector is computed in (9), based on the current position, the trained parameters of the HSMM's states and assuming that the random variables are Gaussian. In the same way, the desired rotational velocity is computed in (10). Finally, the desired exerted force \mathbf{f}_t^* is computed in (11), assuming that both the position and the orientation of the tool are given. The desired position and orientation can then be computed by integration as

$$\mathbf{x}_{t+1}^* = \mathbf{x}_t + \tau \dot{\mathbf{x}}_t^*, \quad (12)$$

$$\mathbf{q}_{t+1}^* = \mathbf{q}_\omega \otimes \mathbf{q}_t^*, \quad (13)$$

with τ being the length of the single timestep and $\mathbf{q}_\omega \equiv e^{\tau \dot{\mathbf{q}}_t^*}$ using $\dot{\mathbf{q}}_t^*$ from (10).

B. Stiffness selection and impedance controller

We utilise a Cartesian impedance controller with varying stiffness \mathbf{k}_c along different Cartesian dimensions for controlling the arm. Our approach can be applied to any Cartesian impedance controller with a control law in the form [18]

$$\begin{aligned} \boldsymbol{\tau}_{cmd} = J^T (\text{diag}(\mathbf{k}_c)(\mathbf{x}_{cmd} - \mathbf{x}_{msr}) + D(\mathbf{d}_c) + \mathbf{F}_{cmd}) \\ + \boldsymbol{\tau}_{dyn}(q, \dot{q}, \ddot{q}), \end{aligned} \quad (14)$$

where $\text{diag}(\mathbf{k}_c)$ denotes the diagonal matrix with the elements of \mathbf{k}_c on the main diagonal. The controller emulates a spring-damper system driven by the desired \mathbf{x}_{cmd} and the measured Cartesian pose \mathbf{x}_{msr} , through programmable Cartesian stiffness \mathbf{k}_c and damping parameters \mathbf{d}_c . The robot can provide a force/torque feedforward term \mathbf{F}_{cmd} . The model of the arm $\boldsymbol{\tau}_{dyn}(q, \dot{q}, \ddot{q})$ compensates for dynamical forces such as gravity and Coriolis forces.

We compute the desired pose \mathbf{x}_{cmd} at each timestep from the HSMM using (12) and (13) and \mathbf{F}_{cmd} using (11). The stiffness coefficient \mathbf{k}_c acts as a weight between the pose term in (14) and the feedforward force term \mathbf{F}_{cmd} . When negligible stiffness is used along a dimension, \mathbf{F}_{cmd} dominates the pose term $\text{diag}(\mathbf{k}_c)(\mathbf{x}_{\text{cmd}} - \mathbf{x}_{\text{msr}})$. Increasing \mathbf{k}_c instead favours the reaching of the commanded pose \mathbf{x}_{cmd} .

We propose a mechanism for selecting, at each timestep of the reproduction, the stiffness parameter \mathbf{k}_c based on the distribution of forces in demonstrations. The mechanism models the force readings as a mixture of two components: one component explaining the readings during the free-space motion phases of the task and another explaining the readings when there is a contact between the robot arm and the environment. We model the force readings in *not in-contact* (denoted as NC) and the *in-contact* (denoted as C) phases, respectively, with a negative exponential distribution and a Gaussian distribution. Fitness to the collected demonstrations drives our modeling choice. We use the Gaussian distribution to be as general as possible since different tasks can have different force distributions. We model the absolute values of force readings since stiffness coefficients \mathbf{k}_c are always positive regardless of the force sign.

The absolute values of force readings F_j along the Cartesian dimension j from the not in-contact phases (NC) of the task are modeled as

$$P(F_j|\text{NC}) = \lambda_j^F e^{-\lambda_j^F F} H(F_j) = \text{Exp}(F_j; \lambda_j^F), \quad (15)$$

where $H(\cdot)$ is the Heaviside step function. The absolute force readings F_j collected during the in-contact phases (C) of the task are instead modeled as

$$P(F_j|\text{C}) = \frac{1}{\sigma_j^F \sqrt{2\pi}} e^{-\frac{(F_j - \mu_j^F)^2}{2(\sigma_j^F)^2}} = \mathcal{N}(F_j; \mu_j^F, \sigma_j^F). \quad (16)$$

Then the mixture of the two distributions in (15) and (16) is

$$F_j \sim P(F_j|\text{C})\phi_j + P(F_j|\text{NC})(1-\phi_j), \quad j = 1, \dots, 3 \quad (17)$$

where F_j is the absolute value of force reading from the set of demonstrations and $\phi_j \equiv P(\text{C})$ is a prior weighting the distributions in the mixture. We optimize the priors ϕ_j and the distribution parameters of the mixture with EM algorithm [19] using data from human demonstrations.

We compute the probability of contact given the absolute exerted force F_j for each Cartesian direction j using Bayes rule as

$$P(\text{C}|F_j) = \frac{P(F_j|\text{C})\phi_j}{P(F_j|\text{C})\phi_j + P(F_j|\text{NC})(1-\phi_j)}. \quad (18)$$

Based on this, for each state $i = 1, \dots, N$ of the HSMM we compute the probability of contact η_i using the absolute value of the trained mean force μ_i^f

$$\eta_i = P(\text{C}||\mu_i^f|). \quad (19)$$

During the reproduction of the task, at each timestep t , we adapt the stiffness parameters \mathbf{k}_c of (14) according to the

expected stiffness over the state and contact probabilities as

$$\mathbf{k}_c(t) = \sum_{i=1}^N (h_{i,t} \text{diag}(\eta_i)(\mathbf{k}_{\max} - \mathbf{k}_{\min}) + \mathbf{k}_{\min}), \quad (20)$$

where $\text{diag}(\eta_i)$ denotes the diagonal matrix with the elements of η_i on the main diagonal. \mathbf{k}_{\max} and \mathbf{k}_{\min} define respectively the maximum and the minimum stiffness allowed for each Cartesian dimension. The choice of \mathbf{k}_{\max} is a tradeoff between the desired compliance of the robot (and thus the safety of the interaction) and the desired accuracy of the pose reproduction. If the minimum allowed stiffness \mathbf{k}_{\min} is set to zero, the Cartesian impedance controller can act as a pure feedforward force controller.

To summarize, at each timestep, we compute the desired stiffness coefficient using the normalized forward variable h_t , that is, the probability of being in each HSMM state, and the state's probability of contact η . The proposed mechanism is able to smoothly vary the stiffness coefficients along all Cartesian directions individually. We can, for example, have the x -axis with stiffness close to \mathbf{k}_{\max} (following the commanded pose \mathbf{x}_{cmd}) and the z -axis with stiffness close to \mathbf{k}_{\min} (exerting the commanded force term \mathbf{F}_{cmd}), all learned automatically from human demonstration.

IV. EXPERIMENTS AND RESULTS

The proposed method was evaluated experimentally to study how the demonstrated tasks are reproduced, by analysing the pose and the exerted forces. The pulling of a door handle and the pushing of a button were chosen as evaluation tasks. In both cases, the robot has to exert force in order to accomplish the task's objective while remaining compliant for safety reasons.

The experiment setup consisted of a 7-DOF KUKA LWR 4+ robotic arm with a six-axis ATI mini 45 F/T sensor rigidly mounted between the robot flange and tool as shown in Fig. 2. The selected tools were a BH8-series BarrettHand in a fixed finger configuration for the handle experiment (Sect. IV-A) and a pen with a softened tip for the button experiment (Sect. IV-B). The arm has integrated torque sensors at each joint that make it capable of programmable active compliance, torque control and gravity compensation. The robot was interfaced to an external computer running the proposed PbD framework through the KUKA *Fast Research Interface* protocol at 100 Hz [18].

During the teaching phase, a human instructor grabbed the flange of the robot and moved the robot in order to teach the pulling down of the handle (see Fig. 1) and the pushing of the button. The demonstrations were provided in gravity compensation mode to ease the demonstrator's effort in moving the system. The tool pose was recorded using joint encoders and contact forces were recorded using the F/T sensor. Experiments consisted of several reproductions but no significant variance was observed between the different trials.

RMSE	Position [m]	Orientation	Force [N]
Handle experiment	0.023	2.874°	6.547

TABLE I: RMS error for the handle experiment of the measured trajectory with respect to the collected demonstrations.

A. Handle pulling experiment

This experiment's purpose was to evaluate the capability of the proposed model to execute a complex task involving changes in the tool's orientation and a non trivial distribution of forces in time and space.

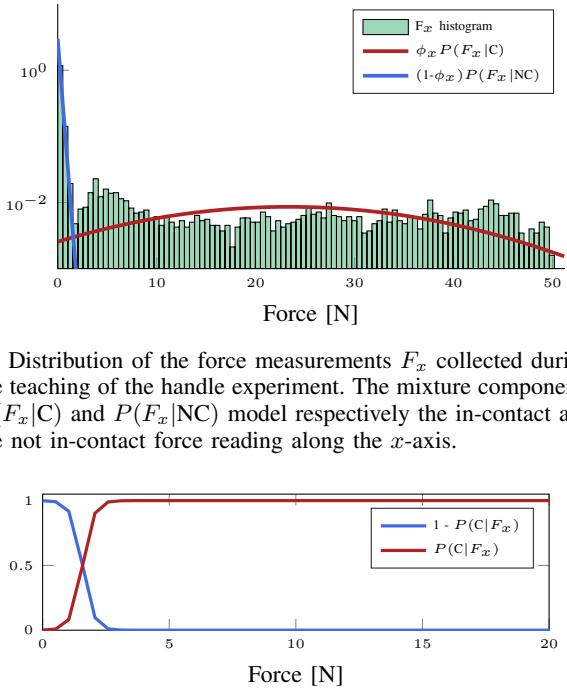
1) *Setting*: For this experiment, a common door handle mechanism was mounted on a rigid structure, at reachable distance from the robot arm. A set of $D = 10$ demonstrations were collected and used to train the HSMM with $N = 15$. During this experiment and the experiment presented in Sect. IV-B, the number of states N of the HSMM was manually chosen. Model selection criteria such as Bayesian Information Criterion (BIC) could be used to find a compromise between accuracy and complexity of the model [20].

During the reproduction, at each timestep t , the commanded pose $[x_t^* \ q_t^*]^T$ from (12) and (13), and the feedforward force f_t^* from (11) were sent to the Cartesian impedance controller (14) at 100 Hz. The damping values d_c were set to $0.7 \frac{\text{Ns}}{\text{m}}$ and $0.7 \frac{\text{Nm s}}{\text{rad}}$ for the linear and the angular motions respectively. The stiffness range parameters k_{\max} and k_{\min} were set respectively to $2000 \frac{\text{N}}{\text{m}}$ and $0 \frac{\text{N}}{\text{m}}$ for the linear motions and to $200 \frac{\text{Nm}}{\text{rad}}$ and $0 \frac{\text{Nm}}{\text{rad}}$ for the angular motions.

2) *Results and analysis*: Figures 4 and 5 show pose and force profiles during a typical reproduction. The robot was able to successfully complete the learned task, reproducing not only the demonstrated pose but also the recorded force profiles while retaining a low stiffness interface with the environment. The combination of HSMM and GMR was able to extract the relevant features from the provided set of demonstrations. Although the demonstrations differ slightly from each other, no alignment of the signals was performed and readings were affected by noise (especially the force readings coming from the F/T sensor).

The stiffness selection mechanism was able to model the two modes (in-contact and not in-contact) from the force readings and compute the probability of contact $P(C|F_j)$ for each Cartesian direction. Fig. 3a shows how the force readings along the x -axis were modeled by two distributions. The mixture was then used to compute the probability of contact $P(C|F_x)$ as shown in Fig. 3b.

During the reproduction, the stiffness selection mechanism adjusted the k_c of the Cartesian Impedance controller based on the state belief $h_{i,t}$ of the HSMM. Fig. 6 shows for the x -axis how the mechanism correctly recognized the HSMM states where there was a contact with the handle (states 8 to 11) and consequently adapted the stiffness to values closer to the selected k_{\min} . While it may first seem counterintuitive to have a low stiffness during contact phases, the low stiffness cancels the spring effect of the impedance controller that



(a) Distribution of the force measurements F_x collected during the teaching of the handle experiment. The mixture components $P(F_x|C)$ and $P(F_x|NC)$ model respectively the in-contact and the not in-contact force reading along the x -axis.

(b) Probability of contact $P(C|F_x)$ given the force measurement along the x -axis for the handle experiment.

Fig. 3: Computation of the probability of contact $P(C|F_x)$ by analyzing the recorded force profiles. The axis limit of the plots are different for a better visualization of $P(C|F_x)$.

opposes the force feedforward. Thus with low stiffness, the commanded force term F_{cmd} influences the control law more, resulting in a better reproduction of the force profile. At the same time, the robot remains safe as the contact forces are limited to those measured during demonstrations.

The RMS errors shown in Table I confirm the accuracy of the reproduction. The RMSE values are computed between the reproduction and the *non-aligned* demonstrations. The small delay between the commanded profiles and the measured profiles shown in Fig. 4 and 5 can be explained by the spring-damper dynamics of the Cartesian impedance controller.

B. Button pushing experiment

The purpose of this experiment was to justify the need for the force to be modeled in the case of in-contact tasks and compare the accuracy in the force profile reproduction of the proposed stiffness selection mechanism against constant stiffness reproduction. To ease the analysis, we chose a button pushing task where relevant forces are exerted only in one Cartesian direction. To further test the capabilities of our model, we explored also the case of small perturbations in the environment by displacing the button 2 cm lower and 2 cm higher.

1) *Setting*: Following demonstration of the task and training of the model, four experimental conditions regarding the reproduction were compared:

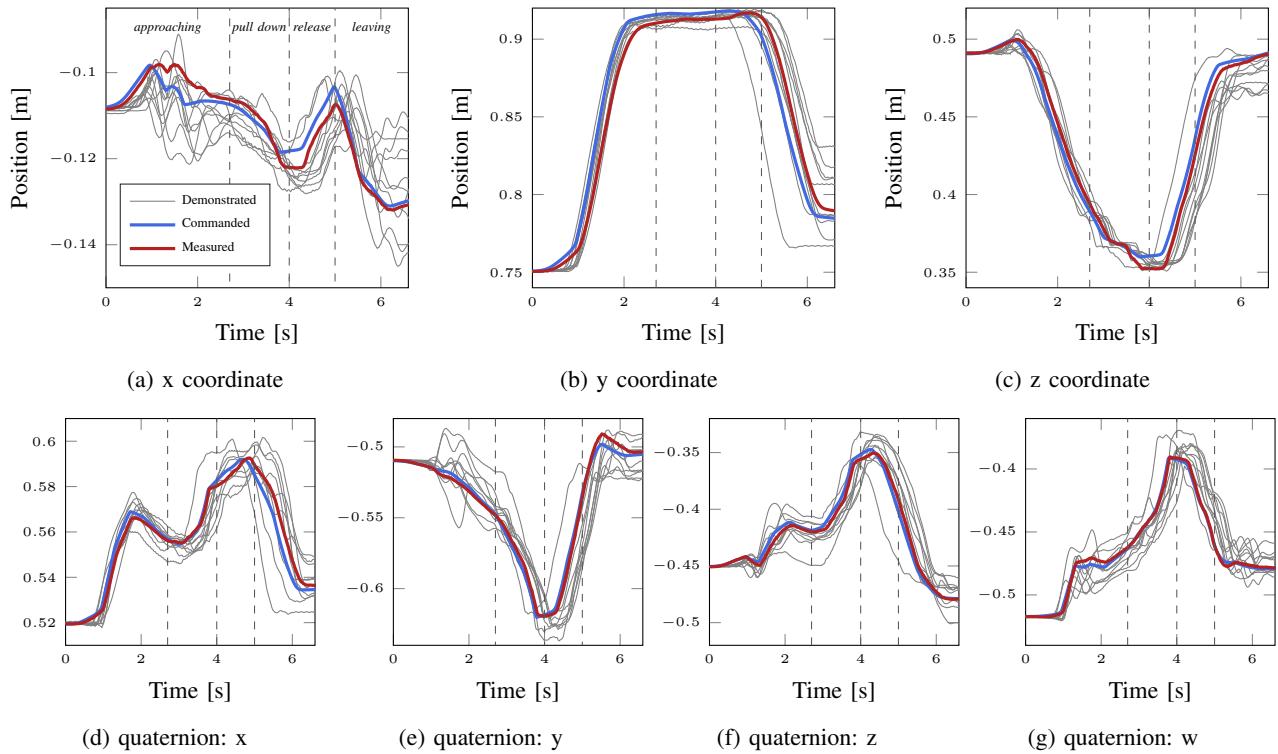


Fig. 4: Pose of the end-effector during a reproduction of the taught task (handle pulling) compared to the demonstrations and the commanded trajectory. The commanded trajectory computed with (12) and (13) generalizes over the provided demonstrations.

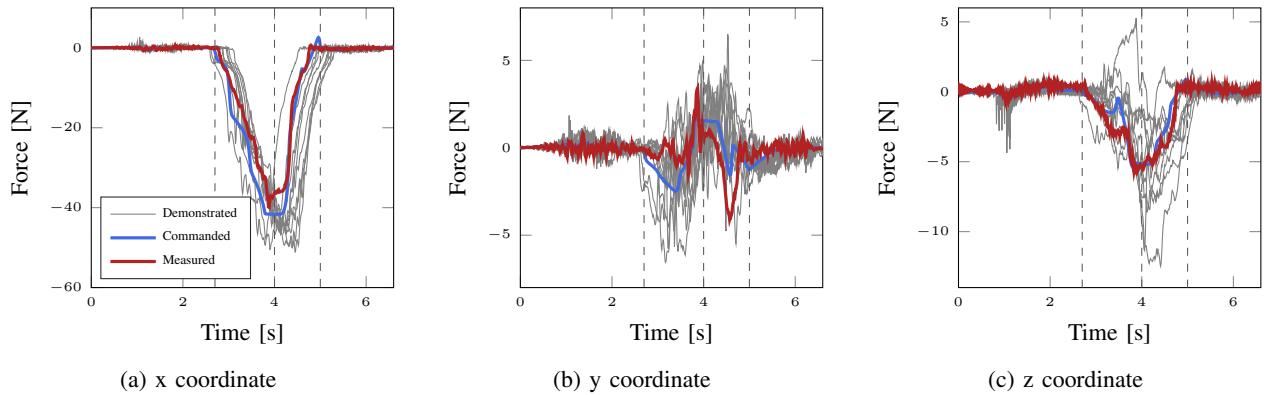


Fig. 5: Force profiles along the Cartesian axes recorded during a reproduction of the taught task (handle pulling) compared to the demonstrated and commanded profiles. The model is able to generalize over the provided set of demonstrations, despite the significant noise coming from the F/T sensor.

- 1) *forceless reproduction*, without training of the force profiles and, consequently, no force control (Condition 1);
- 2) *reproduction with constant stiffness k_c* , with training and feedforward control of the exerted forces (Condition 2);
- 3) *reproduction with the stiffness selection mechanism*, with the training and feedforward control of exerted forces and dynamically set stiffness in the Cartesian impedance controller (Condition 3);
- 4) *reproduction in case of perturbations in the environment*,

same as Condition 3 but the button was placed 2 cm lower or 2 cm higher than in the demonstration (Conditions 4a and 4b, respectively).

The training set consisted of $D = 4$ demonstrations, used to build the HSMM with $N = 10$ states. The damping values d_c were set to $0.7 \frac{Ns}{m}$ and $0.7 \frac{Nm s}{rad}$ for the linear and the angular motions respectively. In Conditions 1 and 2, the stiffness of the Cartesian impedance controller was set to a constant safe value of $k_c = 2000 \frac{N}{m}$ for the linear components and to $200 \frac{Nm}{rad}$ for the angular components. In Condition 1, the

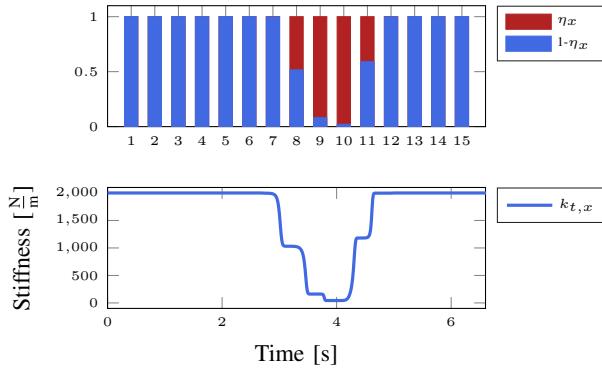


Fig. 6: Working principle of the stiffness selection mechanism. The mechanism defines the probability of contact along the x -axis η_x for each state of the HSMM. Online, the stiffness along the x -axis is adjusted based on the state belief $h_{i,t}$.

commanded pose $\begin{bmatrix} \mathbf{x}_t^* \\ \mathbf{q}_t^* \end{bmatrix}$ from (12) and (13) was sent to the Cartesian impedance controller. Additionally, in Conditions 2, 3 and 4, the force feedforward \mathbf{f}_t^* from (11) was sent to the Cartesian impedance controller. The parameters k_{max} and k_{min} of the stiffness selection mechanism for Conditions 3 and 4, have equal values to the handle experiment in Sect. IV-A.

2) *Results and analysis:* In Conditions 1, 2 and 3, the robot was able to follow the learned trajectory with comparable accuracy of approximately 0.015 m RMSE over the end-effector position and lower than 1° for the orientation (see Table II). In contrast, the force reproductions differed significantly, particularly in tool z direction (direction of the pushing), as illustrated in Fig. 8.

Fig. 8a shows that in Condition 1, i.e. without the force feedforward term \mathbf{F}_{cmd} , the position feedback of the impedance controller did not cause sufficient force for the task to be achieved. In Condition 2, the force reproduction improved but the force component still had an offset of approximately 20 N with respect to the demonstrations during the pushing of the button (4s - 6s) due to the compliance of the Cartesian impedance controller (see Fig. 8b). As a result, the pushing of the button failed also in this case.

Fig. 8c shows the force profile achieved using the proposed stiffness selection mechanism, illustrating good force reproduction with respect to the demonstrations. After a transient, caused by the contact between the button and the tool at $t = 4s$ and the acceleration achieved by the robot arm due to the feedforward term \mathbf{F}_{cmd} , the measured force profile accurately follows the demonstrations with a RMSE of 4.7 N against the 8.3 N and the 10.7 N of Condition 2 and 1 respectively.

To evaluate the capability of the model to adapt to minor changes in the environment, the button was placed 2 cm higher or 2 cm lower with respect to the demonstration set (Conditions 4a and 4b). Fig. 7 shows the position and force profiles along the z -axis in these two Conditions.

RMSE	Position [m]	Orientation	Force [N]
Condition 1	0.015	0.543°	10,769
Condition 2	0.015	0.555°	8,301
Condition 3	0.015	0.613°	4,737
Condition 4a	0.021	0.534°	5,316
Condition 4b	0.016	0.547°	4,483

TABLE II: RMS error for the button experiment of the measured trajectory with respect to the collected demonstrations.

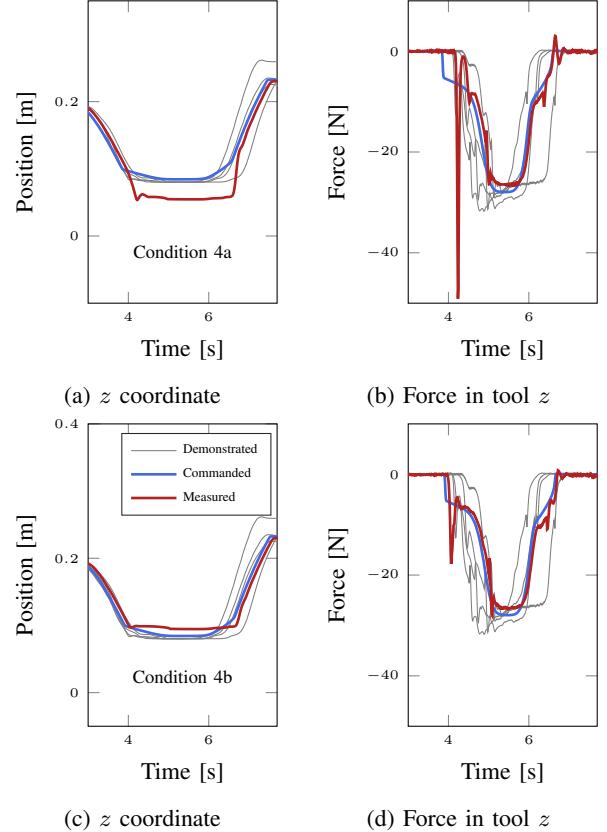


Fig. 7: Experimental conditions 4a (top plots) and 4b (bottom plots): z coordinate of the end effector and force profile along the z tool axis.

Despite the fact that displacements were not present in the training set, the model can cope with them with an acceptable accuracy (see Table II) thanks to the stiffness selection mechanism that, lowering the stiffness during the pushing of the button, allows the reaching of other positions besides the commanded \mathbf{x}_{cmd} . Similar to Condition 3, the force transients are present in both cases of displacement, with an higher peak for Condition 4a. The greater contact forces during the transient in this case are due to the feedforward force term \mathbf{F}_{cmd} acting longer than expected, letting the robot gain more inertia before contact.

V. CONCLUSION

In this paper, we proposed a programming-by-demonstration approach for in-contact tasks that uses a hidden semi-Markov model for encoding the dynamics of the demonstrated trajectories including the force profiles.

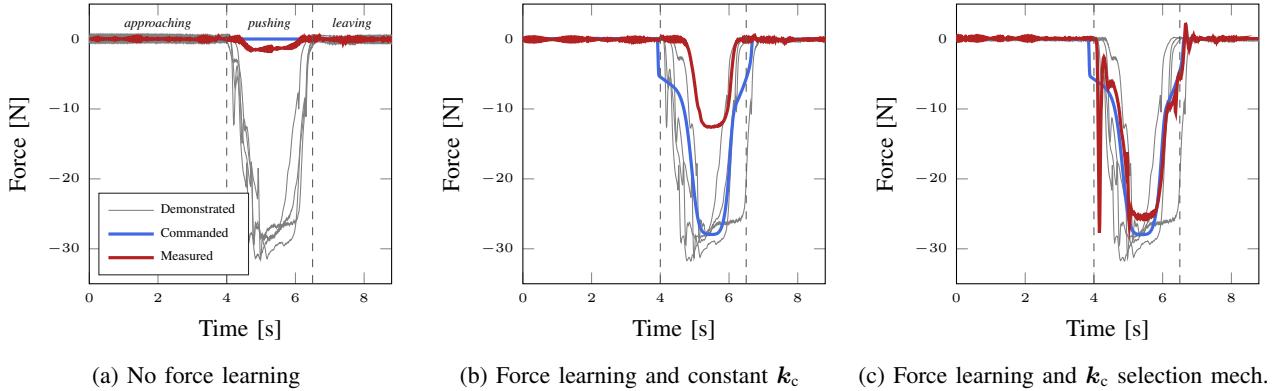


Fig. 8: Force profiles in the tool z direction. The force profile generated by the proposed stiffness selection mechanism in sub-figure (c) was closest to the demonstrated force profiles.

We proposed a mechanism for selecting the stiffness parameters of a Cartesian impedance controller based on the current belief of the HSMM states and the analysis of the recorded force profiles. The state representation allows state-specific control strategies. During the reproduction, the proposed approach achieved safe and compliant operation of an impedance controlled robot while retaining force reproduction capabilities when needed. The approach can select the maximum allowed stiffness, retaining safety of the interaction. A F/T sensor was used during the demonstration to simultaneously teach position and force tasks. The sensor is however not needed during the reproduction due to the feedforward nature of the force control.

In the case of tasks with complex force distributions, the proposed stiffness selection mechanism might fail in modelling the in-contact and not in-contact components of the motion. In that case, more complex distribution models such as mixtures of distributions could be used. However, in tasks such as kneading a mass of dough, the division between force and position requirements is not clear. Nevertheless, the proposed approach of modelling how to combine several control strategies with mixtures of distributions could be used as a starting point towards those complex tasks. However, better understanding of dynamics of such tasks would be essential for future developments.

REFERENCES

- [1] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer handbook of robotics*, pp. 1371–1394, Springer, 2008.
- [2] 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, 2009.
- [3] A. Billard and D. Grollman, "Robot learning by demonstration," *Scholarpedia*, vol. 8, no. 12, p. 3824, 2013.
- [4] A. Montebelli, F. Steinmetz, and V. Kyrki, "On handing down our tools to robots: Single-phase kinesthetic teaching for dynamic in-contact tasks," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp. 5628–5634, IEEE, 2015.
- [5] F. Steinmetz, A. Montebelli, and V. Kyrki, "Simultaneous kinesthetic teaching of positional and force requirements for sequential in-contact tasks," in *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pp. 202–209, IEEE, 2015.
- [6] S. Calinon, F. D'halluin, E. L. Sauser, D. G. Caldwell, and A. G. Billard, "Learning and reproduction of gestures by imitation," *Robotics & Automation Magazine, IEEE*, vol. 17, no. 2, pp. 44–54, 2010.
- [7] S. Calinon, A. Pistillo, and D. G. Caldwell, "Encoding the time and space constraints of a task in explicit-duration hidden markov model," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pp. 3413–3418, IEEE, 2011.
- [8] L. Rozo Castañeda, S. Calinon, D. Caldwell, P. Jimenez Schlegl, and C. Torras, "Learning collaborative impedance-based robot behaviors," in *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, pp. 1422–1428, 2013.
- [9] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, no. 5, pp. 581–603, 2011.
- [10] A. X. Lee, H. Lu, A. Gupta, S. Levine, and P. Abbeel, "Learning force-based manipulation of deformable objects from multiple demonstrations," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 177–184, IEEE, 2015.
- [11] A. Paraschos, E. Rueckert, J. Peters, and G. Neumann, "Model-free Probabilistic Movement Primitives for physical interaction," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2860–2866, IEEE, 2015.
- [12] S. Schaal, P. Mohajerian, and A. Ijspeert, "Dynamics systems vs. optimal control: a unifying view," *Progress in brain research*, vol. 165, pp. 425–445, 2007.
- [13] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in Neural Information Processing Systems (NIPS)*, pp. 2616–2624, 2013.
- [14] S. M. Khansari-Zadeh and A. Billard, "Imitation learning of globally stable non-linear point-to-point robot motions using nonlinear programming," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 2676–2683, IEEE, 2010.
- [15] S. M. Khansari-Zadeh and A. Billard, "Learning stable nonlinear dynamical systems with gaussian mixture models," *Robotics, IEEE Transactions on*, vol. 27, no. 5, pp. 943–957, 2011.
- [16] S.-Z. Yu and H. Kobayashi, "Practical implementation of an efficient forward-backward algorithm for an explicit-duration hidden markov model," *Signal Processing, IEEE Transactions on*, vol. 54, no. 5, pp. 1947–1951, 2006.
- [17] H. Sung, *Gaussian mixture regression and classification*. PhD thesis, Rice University, Houston, Texas.
- [18] G. Schreiber, A. Stemmer, and R. Bischoff, "The fast research interface for the KUKA lightweight robot," in *Proc. of the IEEE Workshop on Innovative Robot Control Architectures for Demanding (Research) Applications – How to Modify and Enhance Commercial Controllers (ICRA 2010)*, IEEE, 2010.
- [19] M. A. Figueiredo, J. M. Leitão, and A. K. Jain, "On fitting mixture models," in *Energy minimization methods in computer vision and pattern recognition*, pp. 54–69, Springer, 1999.
- [20] M. Mühlig, M. Gienger, S. Hellbach, J. J. Steil, and C. Goerick, "Task-level imitation learning using variance-based movement optimization," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pp. 1177–1184, IEEE, 2009.