Interactive Learning via Physical Human Feedback using Uncertainty-Aware Energy Tanks

Edoardo Fiorini[1], Markus Knauer[1,2], Thomas Eiband[1], Maged Iskandar[1], and João Silvério[1]

***Abstract*—Learning from demonstration (LfD) offers an intuitive alternative to manual coding by leveraging natural human behavior, while Human-Robot Interaction (HRI) provides an intuitive means to refine and adapt learned skills. Among interaction modalities, physical contact is a natural and effective way to convey intent. In order to leverage such modality, robots need to be able to distinguish physical contacts coming from deliberate human interactions (e.g. to correct a learned skill) from those from environmental factors (e.g. task-related). In this paper, we introduce a novel interactive framework for physically modulating learned robot skills. Our method builds on a state-of-the-art energy-tank-based intention detection approach to capture degree-of-freedom(DoF)-specific modulations and, accordingly, incorporate user-defined via-points into the learned skills. In contrast to common approaches, corrections are applied selectively to the relevant DoFs, preserving the original skill behavior in the remaining dimensions. Moreover, we leverage uncertainty in the demonstration data to modulate the tank dynamics, allowing users more or less time to intervene in regions of different uncertainty. We validate our approach on a torque-controlled 7-DoF robot executing a learned task of inserting a bearing ring, where physical human corrections are used to adapt to changes in the environment.**

***Index Terms*—Learning from demonstration, Human-Robot interaction, Variable Impedance Control**

I. Introduction

LEARNING from demonstration (LfD) [1] promises to be a no-code method for robot programming. To address variability in tasks and environments, interactive approaches have been proposed to augment the capabilities of classical LfD, in which expert inputs are provided during task execution to update and refine the learned model [2]. There are various ways for humans to provide this input in Human-Robot Interaction (HRI), such as through graphical user interfaces, voice recognition, vision-based systems, and physical-contact interaction [2], [3]. While the first set of methods are based on digital communication channels, the latter offers a more direct and intuitive mode of collaboration, resembling how humans interact with one another. Collaborative robots are a key enabler of such approaches, as they offer physical interaction capabilities (e.g., compliance [4]) that classical industrial robots typically lack. Another crucial aspect is to equip robots with the ability to learn continuously over time [5]. In tactile contacts, it is essential that input from the human operator, typically captured by the robot's force/torque sensors, is clearly distinguished from task-related environmental interactions. For this to work effectively, the robot must be able to correctly interpret the human's intentions to facilitate smooth and accurate interaction. Once a user interaction is detected, it can be exploited to trigger different robot behaviors, e.g. activating or correcting a skill. Skill corrections improve sample efficiency by eliminating the need to collect new data or retrain a model from scratch whenever the environment changes. Moreover, they allow incremental learning by refining faulty skills, saving time for the user.

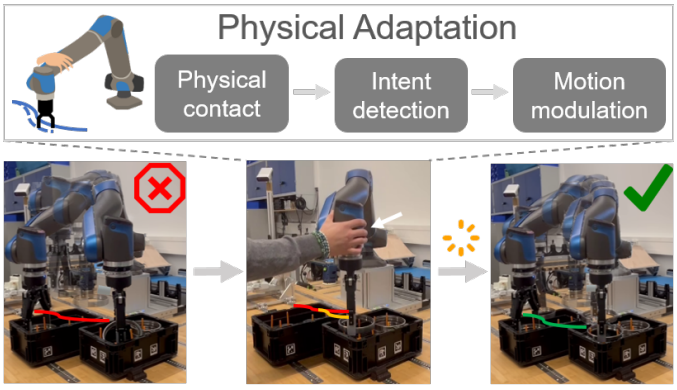


Fig. 1: Overview of the proposed framework. **Left**: the trajectory generated by the learned skill is executed but fails, since the placing phase is not precise enough; **Center**: the operator can intervene during the trajectory and correct the model by touching the manipulator along any part of its surface. This way, the intention is calculated from the physical contact and the motion is modulated; **Right**: the model is updated and the new trajectory containing the corrections is executed ensuring the success of the task.

In this work, we propose an interactive learning approach that combines physical intention detection with learning from demonstration to perform online skill refinement with the human in the loop as illustrated in Figure 1. For this we introduce an energy-tank-based solution, inspired by works like [6], [7], that effectively detects human intention across all degrees of freedom of the Cartesian space independently. We employ a non-parametric, kernel-based LfD algorithm, kernelized Movement Primitives (KMP) [8], that enables incorporation of physical human feedback by adding via-points to the learned model. Unlike state-of-the-art methods, we propose defining via-points in a DoF-specific manner, where the user expresses his feedback by moving the robot

Manuscript received: October, 30, 2025; Accepted February, 14, 2026.

This paper was recommended for publication by Editor Editor name upon evaluation of the Associate Editor and Reviewers' comments. This work was partially funded by the DLR project "Factory of the Future Extended" and the European Union's Horizon Research and Innovation Programme under Grant 101136067 (INVERSE). (*Corresponding author: Edoardo Fiorini.*)

[1]All authors are with the German Aerospace Center (DLR), Robotics and Mechatronics Center (RMC), Münchener Str. 20, 82234 Weßling, Germany. first.last@dlr.de

[2]Markus Knauer is with Technical University of Munich (TUM), School of Computation, Information and Technology (CIT), Arcisstr. 21, 80333 Munich, Germany. markus.knauer@tum.de

Digital Object Identifier (DOI): see top of this page.

in task space during execution, and the corresponding DoF-specific refinement is automatically retrieved. By leveraging an external force estimation [9], integrated with the force-torque and joint-torque sensors of our torque-controlled 7-DoF robot (Fig. 4), we enable interaction along the entire robot body—not just at the end-effector. Moreover, we dynamically modulate the stiffness of the robot's impedance controller using an interaction estimation from the energy, ensuring a smooth and intuitive user's experience. In addition, the uncertainty derived from the learned distribution, which generates the robot reference trajectory, is exploited to regulate the energy tank dynamics. According to the high or low variability in the data, the energy tank empties at different rate, allowing the robot to remain compliant during different task phases, enabling precise movement and effective user interaction, such as for trajectory corrections. In summary we:

1) Introduce a trajectory correction mechanism (Section IV-D) along desired degrees of freedom (Section IV-A), facilitated by an interactive variable impedance controller (Section IV-C).
2) Leverage the covariance of KMPs to incorporate uncertainty into the energy tank model (Section IV-B), thereby connecting interaction and model variance through energy tank dynamics.

We validate our approach on a torque-controlled, 7-DoF, DLR SARA robot [10] performing a bearing ring insertion in a measuring station. We train the skill from demonstrations and use physical corrective feedback to enable adaptation to changes in the task or enviroment. We provide a discussion of the approach and the obtained results in Section VI and concluding remarks and possible extensions in Section VII.

II. Related Work

Several approaches have been proposed to detect desired human input. Machine learning techniques can be employed to classify the origin of input signals [11] as well as force-based threshold methods are engaged to discriminate between environmental noise and intentional forces exerted by a human operator in collaborative tasks. In [12] a force threshold is used to detect when new via-points must be added to the original trajectory, while [13] proposes a human authority parameter which allows the user to hand over task control based on detected external and joint forces. Van der Spaa et al. [14] present a disagreement-aware variable impedance control framework that leverages both human virtual work computed from external forces and a safety threshold to define a human disagreement factor, enabling interactive corrections. In [15], the switch between robot autonomy and human intervention to add new learned skills is ensured by an autonomy level based on measured external wrench. A common issue with all these detection approaches is that they are not always sufficiently robust when dealing with tasks that require interaction with the environment, as they may be triggered even by forces inherent to the task. Instead, our work exploits the injected energy to address this issue during physical contact.

Khoramshahi et al. [6] propose an intention detection method based on an energy tank, combined with admittance control, to enable the robot to transition between an autonomous leader and a passive follower, where it tracks the desired human motion. While this approach exploits an admittance controller, we employ a variable impedance control adapting their proposed formulation in order to connect human intention and robot stiffness. Other works modulate the robot's dynamic response through adjustments to its stiffness matrix. In [13], the stiffness matrix is adapted based on the magnitude of the external wrench, whereas [16] modifies it in response to variations in the input signal. Additionally, [17] presents a framework that dynamically switches between position and impedance control modes according to user requirements, thereby allowing the human operator to take over control of the robot when necessary. While works such as [6] have demonstrated effectiveness in connecting intention detection and shared autonomy, to the best of our knowledge, they have yet to be integrated with learning methods, as we propose in our framework.

There are several movement primitive (MP) approaches compatible with via-point-based skill modulation in LfD. In ProMP [18] a new desired observation is added to the model applying the Bayes theorem, whereas DMP [19] exploits attractor properties to shape the behavior of dynamical system models. Huang et al. [8] introduce KMPs, where new via-points are associated with a conditional probability distribution tailored to the specific task requirements. Compared to the alternatives, KMPs not only permit via-point-based modulations but also capture aleatoric uncertainty (unlike DMPs) and epistemic uncertainty (unlike ProMPs), making them a natural choice for imitation learning. This MP representation is utilized in works such as [20], where an external 3D input device enables trajectory modulation at the desired moment, [12], where immediate corrections are achieved through human-robot physical contact and [21], where the exploration of a reinforcement learning agent is guided by the model uncertainty. Trajectory modulation is a common approach to adaptation in HRI, going beyond learned MPs. Broad et al. [22] introduce motion corrections through natural language commands, while in [23] Losey et al. presented a method for modifying a trajectory in real-time by combining haptic feedback with constrained optimization techniques. However, although these methods have demonstrated strong capabilities in interacting and modulating the trajectory, they never consider DoF-specific corrections, as we do in our work. Indeed, we avoid affecting DoFs that do not require correction, enabling smooth, incremental skill refinement without altering well-performing components.

III. Preliminaries

A. Energy-tank-based intention detection

Energy-tank-based approaches are widely used to design controllers that ensure system passivity and optimize energy efficiency [24]. In [6], Khoramshahi et al. propose an energy-tank-based human-intention detection method. Denoting the external wrench measured by the robot and the estimated human-intended force as $\boldsymbol{F}_e$ and $\boldsymbol{F}_h$ respectively, they are related as

$$\boldsymbol{F}_h = h\boldsymbol{F}_e, \tag{1}$$

where $h \in [0,1]$ describes the interaction ratio. A value of $h = 0$ means that no intention is detected when an external force is applied, while $h = 1$ is reached when the external force is the result of an intentional human contact. The intensity of the interaction can be estimated using the input P_i and output P_o powers of the system

$$P_i = \dot{\boldsymbol{x}}^T \boldsymbol{F}_e, \qquad P_o = \dot{\boldsymbol{x}}^T \boldsymbol{F}_h, \tag{2}$$

where $\dot{\boldsymbol{x}}$ is the end-effector Cartesian velocity. An energy tank, with state E and dynamics

$$\dot{E} = P_i - P_o - (1-h)P_d, \tag{3}$$

is considered, where $P_d > 0$ is the dissipative rate which lowers the tank energy when no input power is applied. The energy in the tank is computed at t by integrating (3) as $E(t) = \int_0^t \dot{E}(t)dt$. The human interaction ratio h is computed according to the energy stored in the tank as

$$h = \begin{cases} 0, & \text{if} \quad E \leq E^* \\ (E - E^*)/(E_{\max} - E^*), & \text{otherwise} \end{cases} \tag{4}$$

where $E_{\max}$ and E^* are the tank size and the energy level that triggers the start of interaction ratio h computation, respectively. As per (1)–(4), externally applied forces increase the energy in the tank through P_i, which is positive when the robot moves along the direction of the applied force. In addition, the dissipative factor P_d helps reject high frequency, low magnitude forces. The combination of these terms and the resulting dynamics into (4) provide a measure of when an external force is the result of a purposeful human interaction.

B. Impedance controller

Given a manipulator with N_q DoF, its rigid-body dynamics is represented by [25] $\boldsymbol{B}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{\tau} + \boldsymbol{\tau}^{ext}$, with the generalized coordinate vector $\boldsymbol{q} \in \mathbb{R}^{N_q}$ and the inertia, Coriolis and centrifugal, and gravity matrices $\boldsymbol{B} \in \mathbb{R}^{N_q \times N_q}$, $\boldsymbol{C} \in \mathbb{R}^{N_q \times N_q}$, and $\boldsymbol{g} \in \mathbb{R}^{N_q}$ respectively; $\boldsymbol{\tau}$, $\boldsymbol{\tau}^{\text{ext}} \in \mathbb{R}^{N_q}$ represent the input command and external forces.

Compliant end-effector robot behavior can be achieved using a Cartesian impedance controller [4]

$$\boldsymbol{\tau} = \boldsymbol{J}(\boldsymbol{q})^T \boldsymbol{F} + \boldsymbol{g}(\boldsymbol{q}), \tag{5}$$

$$\text{with} \quad \boldsymbol{F} = \boldsymbol{K}_P \tilde{\boldsymbol{x}} + \boldsymbol{K}_D \dot{\tilde{\boldsymbol{x}}} + \boldsymbol{F}_d \tag{6}$$

where, given M Cartesian DoFs, $\boldsymbol{F} \in \mathbb{R}^M$ is the task-space wrench needed to obtain a desired spring-damper behavior starting from a virtual spring deflection $\tilde{\boldsymbol{x}} = \hat{\boldsymbol{x}} - \boldsymbol{x}$, with $\hat{\boldsymbol{x}}$ and $\boldsymbol{x} \in \mathbb{R}^M$ the desired and observed operational-space coordinates, respectively; $\boldsymbol{K}_P$ and $\boldsymbol{K}_D \in \mathbb{R}^{M \times M}$ define the stiffness and damping matrices, and $\boldsymbol{F}_d$ is a desired force to be applied on the environment which can be additionally learned.

C. External force estimation

The external interaction wrench $\boldsymbol{F}_e \in \mathbb{R}^M$ is mapped to the joint torques through the Jacobian matrix of the robot $\boldsymbol{J}(\boldsymbol{q}) \in \mathbb{R}^{M \times N_q}$ as

$$\boldsymbol{\tau}^{ext} = \boldsymbol{J}(\boldsymbol{q})^T \boldsymbol{F}_e, \tag{7}$$

where $M = 6$ in case the full Cartesian space is considered. The joint-level external forces can be estimated by monitoring the generalized momentum [9], thus avoiding the need to invert the inertia matrix. This method also decouples the estimated joint-wise external torques and avoids the requirement for joint acceleration. The momentum-based residual vector [26] can be expressed as

$$\boldsymbol{r}(t) = \boldsymbol{K}_o \left(\boldsymbol{p}(t) - \int_0^t (\boldsymbol{\tau} - \boldsymbol{n}(\boldsymbol{q},\dot{\boldsymbol{q}}) + \boldsymbol{r})\, dt - \boldsymbol{p}(0) \right), \tag{8}$$

where $\boldsymbol{p} = \boldsymbol{B}(\boldsymbol{q})\dot{\boldsymbol{q}}$ denotes the generalized momentum. The expression $\boldsymbol{n}(\boldsymbol{q},\dot{\boldsymbol{q}}) = \boldsymbol{g}(\boldsymbol{q}) + \boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} - \dot{\boldsymbol{B}}(\boldsymbol{q})\dot{\boldsymbol{q}}$ is introduced for better readability. Here, $\boldsymbol{K}_o$ denotes a diagonal gain matrix.

Theoretically, the observer gain could be very high $\boldsymbol{K}_o \to \infty$, then the residual convergences quickly to the external torque, i.e., $\boldsymbol{r} \approx \boldsymbol{\tau}^{ext}$. In practice, the observer gain represents a trade-off between the convergence rate of the residual and the noise sensitivity in the measured signal.

In general, the residual is induced by interactions occurring in the task space and projected into the joint space according to (7). Assuming that the Jacobian $\boldsymbol{J}(\boldsymbol{q})$ is known, an estimate $\hat{\boldsymbol{F}}_e \in \mathbb{R}^M$ of the external wrench can be computed as

$$\hat{\boldsymbol{F}}_e = \left(\boldsymbol{J}(\boldsymbol{q})^T\right)^{\#} \boldsymbol{r}, \tag{9}$$

where the symbol $\#$ denotes the generalized pseudoinverse.

D. Kernelized Movement Primitives (KMPs)

KMPs [8] are widely employed in LfD to model and predict the distribution of an output variable $\boldsymbol{\xi} \in \mathbb{R}^O$ conditioned on observed inputs $\boldsymbol{s} \in \mathbb{R}^I$. A KMP is constructed from a *reference trajectory distribution* consisting of N Gaussian components, each characterized by parameters $\{\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n\}_{n=1}^N$. These parameters are derived from a set of U human demonstrations, denoted by $\{\{\boldsymbol{s}_{n,u}, \boldsymbol{\xi}_{n,u}\}_{n=1}^N\}_{u=1}^U$, via Gaussian Mixture Models (GMMs) over the input space $\boldsymbol{s}_{n=1,\ldots,N}$.

Given a novel test input $\boldsymbol{s}^*$, the predictive mean and covariance of the output $\boldsymbol{\xi}(\boldsymbol{s}^*)$ can be computed as follows:

$$\mathbb{E}\left[\boldsymbol{\xi}(\boldsymbol{s}^*)\right] = \boldsymbol{k}^* \left(\boldsymbol{K} + \lambda_1 \boldsymbol{\Sigma}\right)^{-1} \boldsymbol{\mu}, \tag{10}$$

$$\operatorname{cov}\left[\boldsymbol{\xi}(\boldsymbol{s}^*)\right] = \alpha \left(\boldsymbol{k}^{**} - \boldsymbol{k}^* \left(\boldsymbol{K} + \lambda_2 \boldsymbol{\Sigma}\right)^{-1} \boldsymbol{k}^{*\top} \right), \tag{11}$$

where $\boldsymbol{K} = [\hat{\boldsymbol{k}}(\boldsymbol{s}_1)^\top, \ldots, \hat{\boldsymbol{k}}(\boldsymbol{s}_N)^\top]$, $\boldsymbol{k}^* = \hat{\boldsymbol{k}}(\boldsymbol{s}^*)$, with $\hat{\boldsymbol{k}}(\boldsymbol{s}_i) = [\boldsymbol{k}(\boldsymbol{s}_i, \boldsymbol{s}_1), \ldots, \boldsymbol{k}(\boldsymbol{s}_i, \boldsymbol{s}_N)]$, $\boldsymbol{k}^{**} = \boldsymbol{k}(\boldsymbol{s}^*, \boldsymbol{s}^*)$, $\boldsymbol{k}(\boldsymbol{s}_i, \boldsymbol{s}_j) = k(\boldsymbol{s}_i, \boldsymbol{s}_j)\boldsymbol{I}$ and $k(\boldsymbol{s}_i, \boldsymbol{s}_j)$ is a kernel function. Moreover, $\boldsymbol{\mu} = \left[\boldsymbol{\mu}_1^\top \ldots \boldsymbol{\mu}_N^\top\right]^\top$, $\boldsymbol{\Sigma} = \operatorname{blockdiag}\left(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_N\right)$ and λ_1, λ_2, α, are regularization hyperparameters. From (10)—(11), it follows that when the covariance $\boldsymbol{\Sigma}_n$ associated with a particular mean $\boldsymbol{\mu}_n$ is sufficiently small, the predicted expectation at the corresponding input $\boldsymbol{s}_n$ will closely approximate $\boldsymbol{\mu}_n$. This characteristic offers a systematic and effective means for trajectory modulation. Specifically, if one wishes to ensure that the predicted expectation at a new input $\bar{\boldsymbol{s}}$ passes through a desired point $\bar{\boldsymbol{\mu}}$, it is sufficient to augment the reference distribution by adding the pair $\{\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}\}$, where $\bar{\boldsymbol{\Sigma}}$ is chosen to be sufficiently small. This addition guarantees that the mean prediction given by (10) aligns closely with $\bar{\boldsymbol{\mu}}$, while simultaneously reducing the uncertainty described by the covariance in (11) to approximate $\bar{\boldsymbol{\Sigma}}$.

IV. Methodology

The proposed approach combines energy tanks (IV-A) with uncertainty-dependent power dissipation rates (IV-B) into a variable impedance controller that switches control authority between human and robot (IV-C) allowing for DoF-specific via-point definition (IV-D). These elements make up the *online* component of the framework, as they are updated in real time, while the *offline* component involves the training of the trajectory distribution (III-D). A schematic overview of the proposed approach is depicted in Fig. 2.

A. DoF-specific, Energy-tank-based Intention Detection from Contacts Anywhere on the Robot

A key aspect of the solution presented in [6], is that all dimensions of the external force $\boldsymbol{F}_e$ contribute to the computation of h. Specifically, when the power of the system P_i is calculated according to (2), these are combined with the Cartesian velocity resulting in a scalar value, which fills the energy tank. This leads to one interaction ratio for all DoFs in the operational space. As a result, it is challenging to determine which DoF the user intends to correct and it could lead to unnecessary corrections along some of the directions. Note that, compared to [6], in our work we are not using $\boldsymbol{F}_e$ directly, but the framework takes into account external force estimation $\hat{\boldsymbol{F}}_e$ at any point on the robot.

Given $\hat{\boldsymbol{F}}_e \in \mathbb{R}^M$ (computed as described in III-C), we propose to define one energy tank per DoF $m = 1, \ldots, M$. This decouples all the elements, creating different energy tanks, capable of detecting intention in a DoF-specific manner. The input of each energy tank is the force/torque and linear/angular Cartesian velocity depending on whether it is a linear or angular DoF that, for simplicity, we denote as $\hat{\boldsymbol{F}}_e$ and $\dot{\boldsymbol{x}}$ in our notation. For this, we reformulate (1)–(3) as:

$$F_h^m = h_m \hat{F}_e^m \tag{12}$$

$$P_i^m = \dot{x}_m \hat{F}_e^m \qquad P_o^m = \dot{x}_m F_h^m \tag{13}$$

$$\dot{E}_m = P_i^m - P_o^m - (1 - h_m) P_d^m \tag{14}$$

where h_m is a DoF-specific interaction ratio, computed from (4) using E_m. This way, users are not constrained to interact with only the end-effector, when trying to correct a skill, but are able to interact with any point on the robot, making our approach more intuitive and versatile. Note that $\hat{\boldsymbol{F}}_e$ and $\dot{\boldsymbol{x}}$ are mapped to the robot base for simplicity, but they can also be computed in the object or other reference frames.

B. Uncertainty-aware (U-A) energy tank dynamics

A common challenge with energy-tank-based approaches is their parameterization. In our scenario, it is difficult to select parameters that, a priori, are optimal for all types of required corrections since different types of them require different time intervals during which the robot must be compliant. For example, in parts of the task where small but precise corrections are needed, one may want the robot to stay compliant for longer, once an intention is detected. To address this limitation, we propose to let the duration of the intention detection depend on the data. Specifically, making $\boldsymbol{P}_d$ depend on the uncertainty

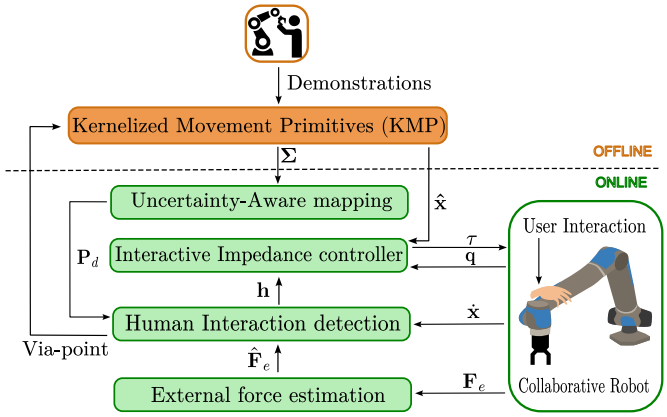


Fig. 2: A schematic representation of the proposed approach. **OFFLINE**: a KMP model is trained after a human demonstrates the task by hand-guiding the robot; **ONLINE**: the user can interact with the robot while it is following the original trajectory, adapting its behavior. $\boldsymbol{F}_e$ is the external interaction wrench, $\hat{\boldsymbol{F}}_e$ is the external wrench estimation along the body of the robot, and h is the intention ratio coming from the detected intention.

of the model that generates $\hat{\boldsymbol{x}}$, which in our case is a KMP. By doing so, the rate at which the energy tank empties is optimized throughout the task execution. Let us denote the KMP covariance prediction in (11) as $\hat{\boldsymbol{\Sigma}}$. At each time step, the m-th diagonal element σ_m^2 can be used in the m-th energy tank dissipative factor $\boldsymbol{P}_d$, leading to:

$$P_d^m = P_{d_{\min}} + \frac{(\sigma_m^2 - \sigma_{\min}^2)(P_{d_{\max}} - P_{d_{\min}})}{\sigma_{\max}^2 - \sigma_{\min}^2} \tag{15}$$

where $P_{d_{\max}}$ and $P_{d_{\min}}$ are the desired maximum and minimum dissipative term, while $\sigma_{\max}^2$ and $\sigma_{\min}^2$ are the maximum and minimum covariance range value. Since the term $\boldsymbol{P}_d$ controls the rate at which the tank dissipates energy when no input power is provided, this will be faster when the variance of the corresponding degree of freedom is higher, thus closer to $\sigma_{\max}^2$. Conversely, the dynamics will be slower in the opposite case. In other words, when there is low uncertainty in the model, the tank will empty slower, ensuring more interaction time for the operator. Conversely, in the opposite case, the interaction time will be shorter. Although the mapping (15) is well-suited to our experimental scenario, it can be adapted to different user or task needs. For instance, inverting the relation can be trivially achieved with one sign change. Regardless of the implementation, to the best of our knowledge, our work is the first proposing to associate model uncertainty with tank dynamics, grounded by a realistic use case.

C. Intention-aware variable impedance controller

The intention captured by the architecture plays a key role in enabling the robot to smoothly become more compliant when a detected external contact is a user intention, to allow the user to modulate the motion. The estimated interaction ratio h_m modulates the impedance control (6) at each loop. The Cartesian wrench used in (5) becomes

$$\boldsymbol{F} = [\boldsymbol{I} - \mathrm{diag}(h_1, \ldots, h_M)]\boldsymbol{K}_P \tilde{\boldsymbol{x}} + \boldsymbol{K}_D \dot{\tilde{\boldsymbol{x}}} + \boldsymbol{F}_d \tag{16}$$

where $h_1, \ldots, h_M$ modulates the position/orientation elements of the stiffness matrix $\boldsymbol{K}_P$. Hence, the robot compliance increases with h_m, reaching full compliance when $h_m = 1$.

D. DoF-specific, intention-aware via-point definition

We employ the DoF-specific intention detection to selectively modulate a demonstrated skill along the relevant DoFs. For this we modify the KMP via-point definition mechanism, described in Section III-D, to add new via-points based on the values of h_m from which a via-point is added. Let us define an interaction index threshold h_{th} as the value of h_m from which via-points should be added. Considering the definition of a via-point $\bar{\boldsymbol{\mu}} \in \mathbb{R}^M$ and the precision matrix $\bar{\Sigma}$ as introduced in Section III-D, and if an intention is detected on the m-th DoF at time t, i.e. $h_{m,t} \geq h_{th}$, a via-point is created with

$$\bar{\boldsymbol{s}} = \boldsymbol{s}_t, \quad \bar{\mu}_m = \bar{x}_{m,t}, \quad \bar{\boldsymbol{\Sigma}} = (\boldsymbol{I}-\boldsymbol{D})\boldsymbol{\Sigma}_t(\boldsymbol{I}-\boldsymbol{D})+\gamma\boldsymbol{D} \quad (17)$$

where $\bar{x}_{m,t}$ is the value of the m-th DoF at t, $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{e}_m)$, $\boldsymbol{e}_m$ is the standard basis vector with 1 at the m-th position and 0 elsewhere, and γ is a small scalar factor. The remaining DoFs $\bar{\mu}_i$, $i \neq m$, are set to the corresponding values of $\boldsymbol{\mu}_n$ in the *reference trajectory distribution* whose input $\boldsymbol{s}_n$ is the closest to $\boldsymbol{s}_t$.

For instance, assuming that $m = 3$ and $h_{3,t} \geq h_{th}$, given

$$\boldsymbol{\mu}_t = \begin{bmatrix}\mu_1 & \mu_2 & \mu_3\end{bmatrix}^T, \quad \boldsymbol{\Sigma}_t = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (18)$$

at input $\bar{\boldsymbol{s}} = \boldsymbol{s}_t$ the final result will be

$$\bar{\boldsymbol{\mu}}_t = \begin{bmatrix}\mu_1 & \mu_2 & \bar{x}_{3,t}\end{bmatrix}^T, \quad \bar{\boldsymbol{\Sigma}}_t = \begin{bmatrix} a & b & 0 \\ d & e & 0 \\ 0 & 0 & \gamma \end{bmatrix}. \quad (19)$$

Indeed, the proposed approach only modifies the learned trajectory distribution along the DoFs that need corrections, leaving the rest of the model unchanged. This is a key differrence with respect to [12] and [20] which assume a diagonal covariance matrix $\bar{\boldsymbol{\Sigma}}_t = \gamma\boldsymbol{I}$ by default, implying that when it is used to define robot stiffness (as in [12]), all DoFs become stiff. Our complete method is summarized in Algorithm 1.

Algorithm 1 Physical adaptation with U-A energy tanks

```
Collect demonstrations {{s_{n,u}, ξ_{n,u}}_{n=1}^N}_{u=1}^U
Set α, λ1, λ2, γ, and define k(.,.)
Set energy tanks E_max, E*
Set tank dynamics params h_th, P_{d_max}, P_{d_min}, σ²_max, σ²_min
Set stiffness K_P
Initialize D_vp = { }
loop for each n = 1, ..., N
    Compute the reference trajectory distribution μ̂ and Σ̂ according to (10)–(11)
    Assign pose from μ to x̂
    Set D = diag(0, 0, 0, 0, 0, 0)
    loop for each m = 1, ..., M
        Compute Ė_m with (14)
        Compute h_m with (4)
        Compute P_d^m according to (15)
        Compute F according to (16)
        if h_m ≥ h_th then
            D += diag(e_m)
        end if
    end loop
    Append D to D_vp
end loop
For each element of D_vp update KMP using (17)
```

V. Evaluation

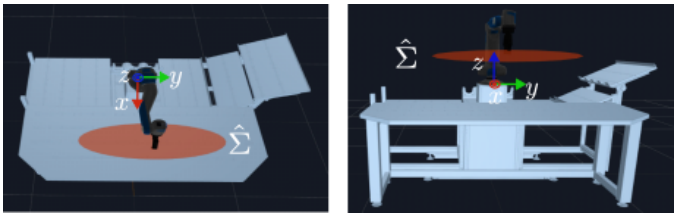
We evaluate our approach on a torque-controlled, 7-DoF, robot in two separate experiments: a proof of concept with simplified conditions, particularly a constant desired pose and constant covariance matrix (we refer to it as *fixed pose* V-A, see Fig. 3) and a learned task of inserting a bearing ring in a measuring station (we refer to it as *bearing ring insertion task with learned trajectory distribution* V-B, see Fig. 4). The first aims to explain the dynamics of our framework, starting from human interaction detection and how this influences the impedance controller. It also accounts for uncertainty awareness, so as to provide a clear overview of the contribution of this work in a relatively simple case. While this first experiment validates the core features of our approach, the second focuses on applying the method to a more realistic scenario with dynamic environment and greater task complexity. This time, we consider a KMP learned from demonstrations with predicted means and covariances as in (10)–(11).

A. Fixed pose

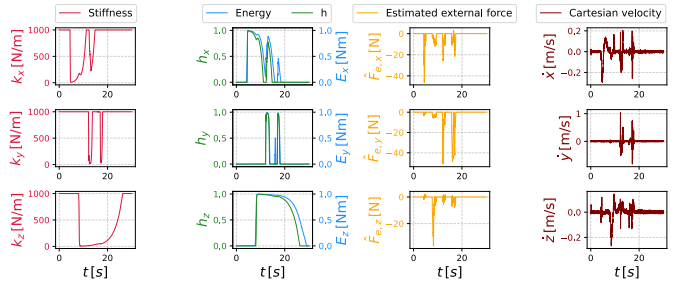
The fixed commanded pose $\hat{\boldsymbol{x}}$ used in the experimental setup is shown in Fig. 4-A. The human operator interacts along the desired DoF filling the correspondent energy tank E_m. As consequence, the interaction ratios h_m are computed and used to update the stiffness of the impedance controller (16). We use $\boldsymbol{K}_P = \mathrm{diag}(1000, 1000, 1000, 100, 100, 100)$, while a constant covariance matrix $\hat{\boldsymbol{\Sigma}} = \mathrm{diag}(0.1,\ 1.0,\ 0.001)$ is set, whose visualization is shown in Fig. 3-a. We expect that the axis with the lower uncertainty (z-axis) stays compliant for longer than the others, since the tank empties more slowly. In detail, $P_{d_{\max}} = 0.5$, $P_{d_{\min}} = 0.01$, $\sigma_{\max} = 1$, and $\sigma_{\min} = 0.01$. Exploiting the mapping presented in (15), a constant Σ results in a constant P_d for each DoF. Regarding energy tanks, we use for all DoFs $E_{\max} = 1.0$ and $E^* = 0.5$. Figure 3-b displays the variation of all these relevant quantities over time, focusing on the Cartesian space position part. It is important to emphasize that during the interaction, the external force estimation is used as the force measurement source, allowing the user to touch the manipulator across its entire surface. Refer to the accompanying video for a detailed visualization of the interaction.

B. Bearing ring insertion task with learned trajectory distribution

In the second experiment, we focus on a ring insertion task with learned trajectory distribution, which is challenging to program by hand. Figure 4 shows the experimental setup, where the goal is to take a ring from the box (Fig. 4-B) and place it in a measuring station (Fig. 4-C). Specifically, the latter demands highly accurate pose control for insertion and interaction with the environment, which significantly increases



(a) Digital twin of the DLR SARA robot and visualization of covariance matrix used within the experiment, $\hat{\mathbf{\Sigma}} = \text{diag}(0.1,\ 1.0,\ 0.001)$. **Left**: Top view; **Right**: Front view.



(b) Robot stiffness, intention rate/energy, estimated force and Cartesian velocity, respectively, during the interaction.

Fig. 3: Uncertainty-aware energy tank dynamics and variable impedance controller for a fixed pose tracked by a torque-controlled robot, focusing only on position.

its complexity compared to traditional pick-and-place tasks that are performed in free space. Changes in the environment (e.g., obstacles) and/or variations in object tolerances that increase precision demands, make the ideal scenario for applying our method. In order to grasp the ring we mounted a Robotiq 140 two-finger gripper at robot end-effector, whose state (open/closed) is also part of the learned distribution. The KMP model was trained based on three demonstrations given by the user kinesthetically (OFFLINE part in Fig. 2), initializing it from GMMs with 12 components. We employed a RBF kernel [27] with length scale $l = 7 \times 10^{-3}$, $N = 200$ inputs, and choosing the following hyperparameters $\alpha = 1.0$, $\lambda_1 = 1 \times 10^{-3}$, and $\lambda_2 = 1 \times 10^{-2}$. The $\boldsymbol{s}$ input is time t, and $\boldsymbol{\xi}$ contains end-effector position $\boldsymbol{x}$ and orientation $\boldsymbol{\phi}$, as axis-angle, as well as gripper state (a scalar between 0 and 1 depending on whether it Is open or closed). In order to better highlight the features of our method, after the user gave their demonstrations we added an obstacle (blue boxes) in the middle of the trajectory and the position of the ring measurement station has been shifted by a few centimeters; therefore, to correctly perform the task in the new setup, the human operator has to modulate the original model's trajectory (Fig. 4-B). The proposed interactive framework (ONLINE part in Fig. 2) is tested by physically interacting with the robot along its body to correct the learned trajectory during task execution. We set an interaction index threshold $h_{th} = 0.8$. When the interaction indices reached $h_m \geq h_{th}$, according to the energy tank equation, via-points were added at the corresponding DoF at that time instant using (17) with $\gamma = 1 \times 10^{-5}$. Note that this behavior is supported by the variable impedance controller (IV-C) which makes the robot compliant. In the meantime, the uncertainty-aware module manages the rate for interaction according to the variance of the model generating a variable $\boldsymbol{P}_d$ term. Figure 5

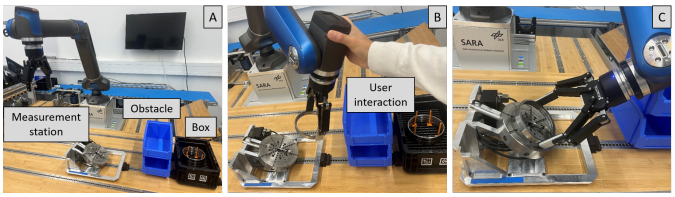


Fig. 4: The experimental setup consists of a torque-controlled 7-DoF, DLR SARA robot mounted on a manufacturing workstation, a ring storage unit, a dedicated station for precise ring measurement, and a set of boxes as obstacle. **A** shows the placement of elements within the scene; **B**: shows the user interaction to correct the trajectory in order to avoid obstacle; **C** highlights the non-trivial orientation that must be maintained for the ring placement.

shows the behavior of all elements presented so far. In our experiments, demonstrations, forces and corrections are all represented in the robot base. For the energy tank parameters we used $E_{\max,xyz} = 0.4$, $E^*_{xyz} = 0.36$, $E_{\max,\phi_x} = 0.35$, $E^*_{\phi_x} = 0.33$, $E_{\max,\phi_{yz}} = 0.6$, and $E^*_{\phi_{yz}} = 0.58$. Regarding uncertainty-aware mapping, we set $P_{d_{\max}} = 0.8$, $P_{d_{\min}} = 0.01$, $\sigma_{\max} = 10^{-2}$, and $\sigma_{\min} = 10^{-4}$. To effectively motivate the necessity of the uncertainty-aware module, we compared the execution of the correction using this module against a fixed value of $\boldsymbol{P}_d$ (Fig. 6).

VI. Discussion

A. Analysis of experimental results

Figure 3 presents the obtained results in the *fixed pose* experiment. On the left and middle-left side, the evolution of the commanded stiffness against the energy and the interaction ratios, displayed in light blue and green, respectively, is shown. While these latter two quantities are directly proportional to one another, following a similar trend despite differing in scale, the commanded stiffness exhibits an inverse relationship with respect to them, as governed by the discussed above control Eq. (16). Indeed, the user begins to interact with the system, filling the corresponded DoF-specific energy tank, leading to a compliance modulation, enabling a smoother interaction. The estimated external force and Cartesian velocity used to compute $\dot{E}$, are shown in the second half of Fig. 3. Overall, when these two quantities are aligned, i.e., they have the same sign or, more intuitively, their vectors point in the same direction, they contribute to fill the DoF-specific tanks. The parameter P_d of each tank is governed by (15), associating high covariance with high P_d. As proof of concept, we employed fixed but significantly different variance across different DoFs. In detail, the P_d value along the y-axis is high, compared to the rest, resulting in an almost immediate return to the original stiffness after the user's interaction or, alternatively, it requires much more force to keep the intention on. Conversely, along the z-axis, which represents the opposite extreme, the energy tank empties slowly, leading to a longer interaction.

In the second experiment, the left side of Fig. 5 shows the full end-effector Cartesian pose: the dark blue trajectory is the path learned via demonstration, and the violet trajectory is the re-generated path after adding the user's desired via-points, marked with red stars. As for *fixed pose*, the center-left part displays the h_m and E_m framework dynamics, for the different DoFs. The two trajectories follow a similar pattern, except in the regions where via-points are introduced. Here, the violet

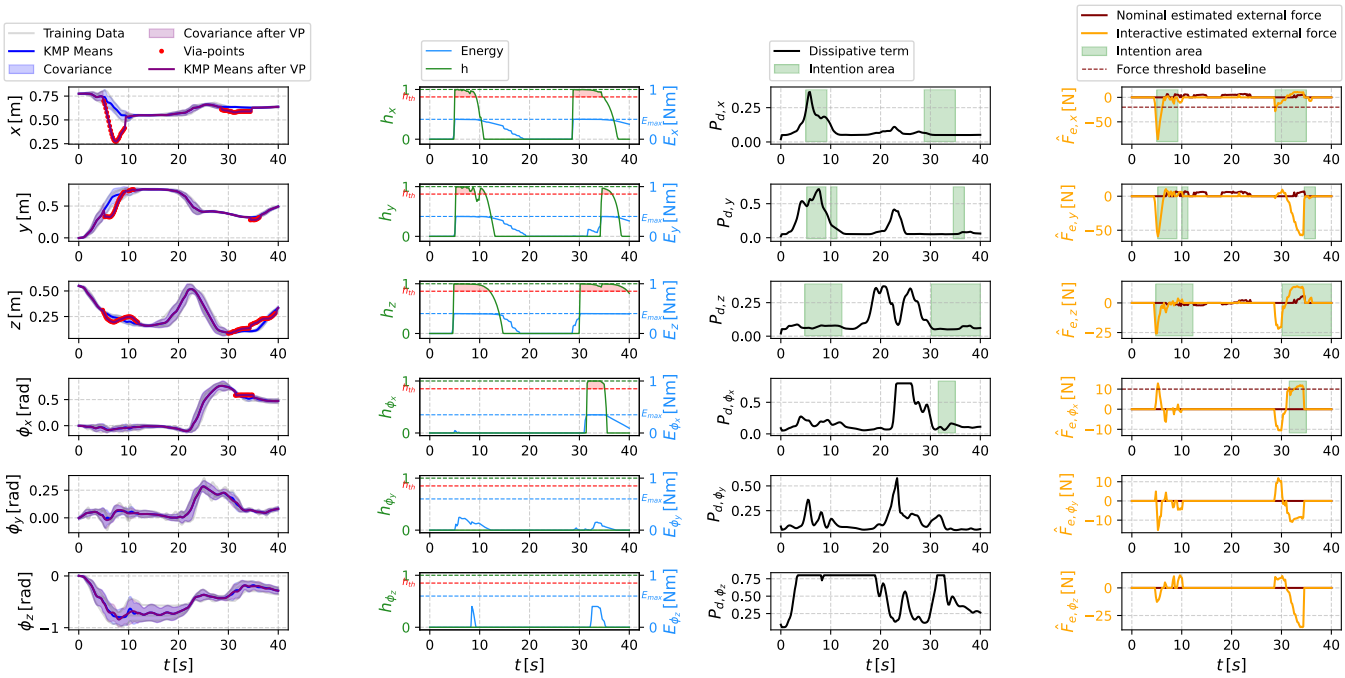


Fig. 5: The plot shows the obtained results during the correction of a manufacturing ring measurment task performed by a torque-controlled robot along all different Cartesian DoF. **Left**: dark blue line is the trajectory predicted by KMP model, purple line is the one generated by KMP model with added via points, red star are via-points, grey the given demonstrations, semi-transparent blue the covaraince before correction, and semi-transparent violet the covariance after via-points; **Center-left**: the interaction index h and the energy E; **Center-right**: the dissipative factor P_d; **Right**: The forces measured by external observer during the human-robot body contact.

line deviates from the original trajectory, passing through the via-points and then rejoining the initial path, effectively performing the intended task correction. Each via-point added to the trajectory corresponds to a h_m value greater than h_{th}, as depicted by the red dashed line, underlining that the new via-point is only added when human intention is detected by the proposed method. Specifically, these are added to avoid the blue box obstacle ($0.1s$–$0.25s$) and to place the ring at the new ring measurement station position, ($0.55s$–$1.0s$). As discussed in (IV-D), via-points are added DoF-specific, so that the covariance matrix is adjusted only along the corresponding axis. First column of Fig. 5 highlights this by showing the trajectory distribution before and after modulation is applied, denoting that the covariance is reduced to zero in the region containing the red points only. Center-right shows the trend of P_d using the same uncertainty-aware mapping employed in the *fixed pose* scenario, but with a real distribution this time. This leads to a lower P_d in critical regions such as ring picking and insertion ($10s$–$20s$ and $30s$-$40s$) that leaves the user with more time for fine, precise corrections, since the tanks empty slower. Indeed, the user successfully gave a correction in the final phase of the task, refer to the supplementary video for a detailed visualization of the corrected trajectory execution.

The right side of Fig. 5 shows the recorded estimated external forces during task execution, with user interaction and nominal task forces indicated in orange and brown, respectively. Several force peaks can be observed, but only those contributing to an increase in the system's energy are detected as intentions. Force increases that do not trigger intention detection are attributed to routine environmental interactions, such as picking up or inserting the ring, and are correctly ignored by our system. As highlighted by the brown dashed line, our method demonstrates to be more robust than threshold-based approaches. For example, in the force component $\hat{F}_{e,\phi_x}$, the magnitude observed during the correction phase ($30s$–$35s$) is lower than the peak that occurred earlier in the task (at $\sim 6s$), which would have been mistakenly identified as an intention by simple threshold methods. Note that in some task phases, interactive forces are comparable to nominal ones, which threshold-based methods may misclassify as intentions.

We compare our method with the same trajectory corrected with the *uncertainty-aware* mode disabled and a small, fixed P_d. A constant P_d suffices to refine the insertion pose, which requires fine, precise corrections. However, the experiment shows that optimizing P_d for one type of correction can compromise others, reinforcing the need for a task-moment-dependent P_d. In Figure 6, the stiffness trends for the two cases are shown on the left: the red line corresponds to the uncertainty-aware enabled case, and the black line to the disabled case. On the right, the corresponding energy tank dynamics are shown. After the first user correction to avoid the obstacle ($5s$–$10s$), the tanks did not empty fast enough for the robot to regain adequate stiffness, preventing it from resuming the task and picking the ring (see video for details).

B. Limitations

Despite the positive results, some limitations of our approach are worth noting. First, because the power (2) depends on both force and velocity, the framework is not well suited to environments containing non-rigid objects. In such cases, external forces exerted on the robot by an object returning to its original shape after contact may be misinterpreted as user intent. Second, the method assumes that demonstrations

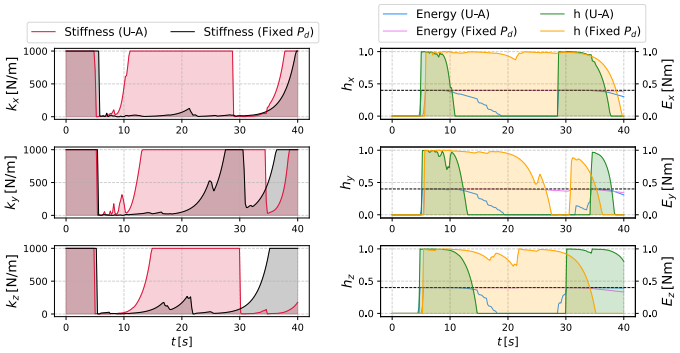


Fig. 6: The plot depicts the different behavior along Cartesian space position using either uncertainty-awareness (U-A) or fixed P_d during the ring measurment task. **Left**: red line is the stiffness modulation using U-A, while black line keeping a fixed P_d; **Right**: energy tank and intention term h during U-A (light blue and violet lines) and fixed P_d (green and orange lines) circumstance are compared.

are sufficiently diverse in the task regions that allow for variation (e.g., approaching the ring box) and, conversely, consistent in regions that do not (e.g., inserting the ring). In practice, this means that users should be encouraged to provide demonstrations that are diverse where flexibility is expected and consistent where precision is required.

VII. Conclusion and Future Work

We proposed a method for detecting human intention through physical contact along the entire robot body, adapting a learned skill in a DoF-specific manner. The core of our approach lies in generating an interaction ratio using an energy tank formulation, which modulates the stiffness of the robot's impedance controller and triggers the definition of corrective via-points. We introduced an uncertainty-aware mechanism that modulates the tank emptying rate by exploiting the variance of the learned distribution, ensuring task-phase-dependent compliance behavior. We demonstrated the effectiveness of our method in a bearing ring insertion task, where human intention dynamically adjusted the robot's trajectory through new via-points, effectively correcting the task execution. In future work, we aim to extend our method to include force corrections and investigate its applicability in human-in-the-loop reinforcement learning.

References

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

[2] C. Celemin, R. Pérez-Dattari, E. Chisari, G. Franzese, L. de Souza Rosa, R. Prakash, Z. Ajanović, M. Ferraz, A. Valada, J. Kober *et al.*, "Interactive imitation learning in robotics: A survey," *Foundations and Trends® in Robotics*, vol. 10, no. 1-2, pp. 1–197, 2022.

[3] M. Iskandar, A. Albu-Schäffer, and A. Dietrich, "Intrinsic sense of touch for intuitive physical human-robot interaction," *Science Robotics*, vol. 9, no. 93, p. eadn4008, 2024.

[4] N. Hogan, "Impedance control: An approach to manipulation," in *1984 American Control Conference*, 1984, pp. 304–313.

[5] T. Lesort, V. Lomonaco, A. Stoian, D. Maltoni, D. Filliat, and N. Díaz-Rodríguez, "Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges," *Information Fusion*, vol. 58, pp. 52–68, 2020.

[6] M. Khoramshahi and A. Billard, "A dynamical system approach for detection and reaction to human guidance in physical human–robot interaction," *Autonomous Robots*, vol. 44, no. 8, pp. 1411–1429, Nov 2020.

[7] A. Kucukyilmaz, T. Sezgin, and C. Basdogan, "Recognition of haptic interaction patterns in dyadic joint object manipulation," *IEEE transactions on haptics*, 12 2014.

[8] Y. Huang, L. Rozo, J. Silvério, and D. Caldwell, "Kernelized movement primitives," *The International Journal of Robotics Research (IJRR*, vol. 38, pp. 833–852, 05 2019.

[9] M. Iskandar, O. Eiberger, A. Albu-Schäffer, A. De Luca, and A. Dietrich, "Collision detection, identification, and localization on the DLR sara robot with sensing redundancy," in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2021, pp. 3111–3117.

[10] M. Iskandar, C. Ott, O. Eiberger, M. Keppler, A. Albu-Schäffer, and A. Dietrich, "Joint-level control of the dlr lightweight robot sara," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 8903–8910.

[11] F. Franzel, T. Eiband, and D. Lee, "Detection of collaboration and collision events during contact task execution," in *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, 2021, pp. 376–383.

[12] M. Knauer, A. Albu-Schäffer, F. Stulp, and J. Silvério, "Interactive incremental learning of generalizable skills with local trajectory modulation," *IEEE Robotics and Automation Letters*, vol. 10, no. 4, pp. 3398–3405, 2025.

[13] S. V. Jadav, J. Heidersberger, C. Ott, and D. Lee, "Shared autonomy via variable impedance control and virtual potential fields for encoding human demonstrations," *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, pp. 15 151–15 157, 2024.

[14] L. van der Spaa, J. Kober, and M. Gienger, "Simultaneously learning intentions and preferences during physical human-robot cooperation," *Autonomous Robots*, vol. 48, no. 4, p. 11, Jun 2024.

[15] X. Chen, T. Ni, K. Karacan, H. Sadeghian, and S. Haddadin, "Online transfer and adaptation of tactile skill: A teleoperation framework," in *8th Annual Conference on Robot Learning*, 2024.

[16] J. Silvério, Y. Huang, L. Rozo, F. Abu-Dakka, and D. G. Caldwell, "Uncertainty-aware imitation learning using kernelized movement primitives," in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, Macau, China, November 2019.

[17] J. Silvério, Y. Huang, L. Rozo, S. Calinon, and D. G. Caldwell, "Probabilistic learning of torque controllers from kinematic and force constraints," in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, Madrid, Spain, October 2018.

[18] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in Neural Information Processing Systems*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds., vol. 26. Curran Associates, Inc., 2013.

[19] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328–373, 2013.

[20] G. Quere, F. Stulp, D. Filliat, and J. Silvério, "A probabilistic approach for learning and adapting shared control skills with the human in the loop," in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2024, pp. 15 728–15 734.

[21] A. Padalkar, F. Stulp, G. Neumann, and J. Silvério, "Towards safe and efficient learning in the wild: Guiding rl with constrained uncertainty-aware movement primitives," *IEEE Robotics and Automation Letters*, vol. 10, no. 7, pp. 6880–6887, 2025.

[22] A. Broad, J. Arkin, N. Ratliff, T. Howard, and B. Argall, "Real-time natural language corrections for assistive robotic manipulators," *The International Journal of Robotics Research*, vol. 36, no. 5-7, pp. 684–698, 2017.

[23] D. P. Losey and M. K. O'Malley, "Trajectory deformations from physical human–robot interaction," *IEEE Transactions on Robotics*, vol. 34, no. 1, pp. 126–138, 2018.

[24] F. Califano, R. Rashad, C. Secchi, and S. Stramigioli, "On the use of energy tanks for robotic systems," in *Human-Friendly Robotics 2022*, P. Borja, C. Della Santina, L. Peternel, and E. Torta, Eds. Cham: Springer International Publishing, 2023, pp. 174–188.

[25] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*. Springer Publishing Company, Incorporated, 2010.

[26] A. De Luca, A. Albu-Schaffer, S. Haddadin, and G. Hirzinger, "Collision detection and safe reaction with the DLR-III lightweight manipulator arm," in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2006, pp. 1623–1630.

[27] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.