

# A Passivity-Based Variable Impedance Controller for Incremental Learning of Periodic Interactive Tasks

Matteo Dalle Vedove<sup>1,2</sup>, Edoardo Lamon<sup>3,4</sup>, Daniele Fontanelli<sup>1</sup>, Luigi Palopoli<sup>3</sup>, and Matteo Saveriano<sup>1</sup>

**Abstract**—In intelligent manufacturing, robots are asked to dynamically adapt their behaviours without reducing productivity. Human teaching, where an operator physically interacts with the robot to demonstrate a new task, is a promising strategy to quickly and intuitively reconfigure the production line. However, physical guidance during task execution poses challenges in terms of both operator safety and system usability. In this paper, we solve this issue by designing a variable impedance control strategy that regulates the interaction with the environment and the physical demonstrations, explicitly preventing at the same time passivity violations. We derive constraints to limit not only the exchanged energy with the environment but also the exchanged power, resulting in smoother interactions. By monitoring the energy flow between the robot and the environment, we are able to distinguish between disturbances (to be rejected) and physical guidance (to be accomplished), enabling smooth and controlled transitions from teaching to execution and vice versa. The effectiveness of the proposed approach is validated in wiping tasks with a real robotic manipulator.

## I. INTRODUCTION

The introduction of collaborative robotic arms in manufacturing processes brought several benefits in terms of safety, flexibility, and ease of use. The limitations in actuation torque contribute to intrinsic safety by constraining the robot physical capabilities. Moreover, their control architecture can incorporate compliant behaviours by dynamically regulating the robot impedance of the robot based on the task requirements and safety protocols, enabling safe interaction with unstructured environments. Second, it is possible to quickly teach new skills from human demonstrations to the robot intuitively. The versatility in the execution of a variety of tasks, combined with the need of adapting the production process to new products, makes collaborative robot a flexible tool for reconfigurable manufacturing systems.

In this context, Programming by Demonstration (PbD) represent one of the noteworthy methods to transfer physical skills from human demonstrations, making robot programming faster, easier and more accessible to people without prior robotics expertise, such as factory workers. To learn from multiple demonstrated trajectories, PbD requires a

strategy to parameterise and encode the taught demonstration. Examples of the most used are (task-parameterized) Gaussian Mixture Model (GMM) [1], Dynamical Movement Primitives (DMPs) [2], [3], Probabilistic Movement Primitives (ProMP) [4], and Kernelized Movement Primitives (KMP) [5]. The demonstrated trajectories could be obtained from visual observations or from kinesthetic teaching [6]. In the context of collaborative robots, whose design envisions physical interactions with humans, the latter has proven to be more effective, since it does not require any retargeting from the teacher to the robot. While at the kinematic level human demonstrations have proven a high level of reliability and precision in repetitions, learning contact forces through kinesthetic teaching is still complex due to the high variability of the human ability to accurately exert the same amount of force [7]. For this reason, in the case of physically interactive tasks, PbD needs to be coupled with adaptive control schemes that are able to adjust the interaction forces with the environment and the teacher. This regulation is particularly complex in the case of incremental or continual learning applications, where the learning phase is interleaved on-the-fly with the execution phase.

In this direction, variable impedance strategies allow the robot to interact safely with the environment, avoiding the application of high forces that could harm the human or damage the robot or the surroundings, especially in case of unexpected perturbation of the system, while at the same time ensuring accurate trajectory tracking according to the task requirements [3]. This ability, in addition, if coupled with a system capable of distinguishing the desired physical interactions from the unexpected ones, enables the robot to behave differently in teaching (being more compliant) and execution phases (being stiffer).

Several examples of PbD with variable interactive controllers are available in literature [8]–[10]. For instance, in [10] authors exploit two DMPs to learn the Cartesian motion and the interaction forces, and replicate it with a hybrid force-position controller. The teaching phase is achieved through gravity compensation obtained by reducing the controller stiffness according to the Cartesian position error as well as the external forces exerted at the end-effector. However, high forces are required to perturb the system when autonomous, as the stiffness is high in such a case. Moreover, these approaches do not tackle energy considerations, which are paramount to ensure the system's stability. The stability of such schemes is often addressed by introducing energy tanks, that ensure the system passivity with time varying damping [11] or stiffness [12]. Recent approaches [13], [14]

We acknowledge the support of the MUR PNRR project FAIR - Future AI Research (PE00000013), European Union projects INVERSE (grant agreement no. 101136067) and MAGICIAN (grant agreement no. 101120731)

<sup>1</sup>Department of Industrial Engineering, Università di Trento, Trento, Italy. [matteo.dallevedove@unitn.it](mailto:matteo.dallevedove@unitn.it)

<sup>2</sup>DRIM, Ph.D. of national interest in Robotics and Intelligent Machines.

<sup>3</sup>Department of Information Engineering and Computer Science, Università di Trento, Trento, Italy.

<sup>4</sup>Human-Robot Interfaces and Interaction, Istituto Italiano di Tecnologia, Genoa, Italy.

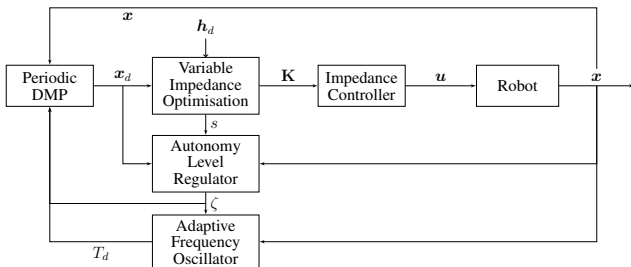


Fig. 1. High level overview of the proposed framework.

show that the dynamics of the energy tank can be designed to further ensure stability in case of free motion.

In this work, we develop an incremental learning approach to an industrial human-robot collaborative application in which robot is required to execute a periodic task, and the worker can physically interact with the robot by updating its desired trajectory through kinesthetic teaching. While the taught Cartesian trajectory is encoded with DMPs, the desired force is constant and specified by the task requirements. The demonstrated task's period is estimated through a custom adaptive frequency oscillator [15]. The proposed strategy is able to interleave learning and execution phases without the need of two different controller strategies, by regulating the level of compliance, formulated as a quadratic program (QP) optimisation, according to the measured interaction, the tracking error, and the energy exchanged by the system. The energy tank not only ensures the system's passivity, but is also used in the optimisation problem to regulate the input/output energy flow. By monitoring such flow, the system can distinguish desired demonstrations from other types of physical interactions, such as disturbances. We evaluate the proposed approach with a proof-of-concept polishing/cleaning/wiping task on arbitrary-shaped surfaces, where three tasks, with different positions and orientations, are successfully taught with few demonstrations and executed on-the-fly by the robot.

## II. METHODOLOGY

To this aim, we integrate three modules that work interactively to learn, execute, and adapt the required periodic motion. In particular, low-level control is relegated to a variable impedance controller that modulates the stiffness to generate a target force at the contact point. Interaction safety with the human demonstrator is ensured by energy tanks embedded in an optimisation problem that modulates the stiffness and avoids the instantaneous generation of excessive forces. The end-effector trajectory is incrementally learned from a human demonstration and encoded into a periodic DMP that considers both position and orientation. To let the system work without any prior information regarding the demonstration period, we add to the framework an adaptive frequency oscillator which online estimates the fundamental frequency directly from the demonstration data. The different modules are coordinated by an autonomy level regulator, a unit that's responsible for identifying the autonomy level  $\gamma \in [0, 1]$  of the system that, in turn, tunes the learning factor  $\zeta \in [0, 1]$  that's used in both the dynamic movement

primitive and the adaptive frequency oscillator. The overall framework is depicted in Fig. 1.

### A. Variable Impedance Control

To have a singularity-free description of the robot end-effector state, we consider the pose  $\mathbf{x} = (\mathbf{p}^\top, \mathbf{q}^\top)^\top \in \mathbb{R}^7$  as encoded by its position  $\mathbf{p} \in \mathbb{R}^3$  and a unit quaternion  $\mathbf{q} \in \mathbb{S}^3 \subset \mathbb{R}^4$  to represent orientation.

Let  $\mathbf{h}^{ext} = (\mathbf{f}^{ext\top}, \boldsymbol{\tau}^{ext\top})^\top \in \mathbb{R}^6$  be the generalised force applied on the end-effector, with  $\mathbf{f}^{ext}, \boldsymbol{\tau}^{ext} \in \mathbb{R}^3$  respectively a force and a torque, then the objective of a Cartesian variable impedance controller is to render the robot closed-loop dynamics into a mass-spring-damper model in the form

$$\mathbf{h}^{ext} = \boldsymbol{\Lambda} \ddot{\tilde{\mathbf{x}}} + \mathbf{D} \dot{\tilde{\mathbf{x}}} + \mathbf{K} \tilde{\mathbf{x}}. \quad (1)$$

In (1), symmetric positive-definite matrices  $\boldsymbol{\Lambda}, \mathbf{D}, \mathbf{K} \in \mathbb{R}^{6 \times 6}$  are respectively the desired Cartesian inertia, damping and stiffness; to have a geometrically consistent definition of rotation difference in the quaternion domain, the Cartesian error is defined as

$$\tilde{\mathbf{x}} = \begin{pmatrix} \tilde{\mathbf{p}} \\ \tilde{\mathbf{q}} \end{pmatrix} = \begin{pmatrix} \mathbf{p}^{des} - \mathbf{p} \\ \text{Log}(\mathbf{q}^{des} * \bar{\mathbf{q}}) \end{pmatrix} \in \mathbb{R}^6, \quad (2)$$

with  $\mathbf{x}_d = (\mathbf{p}^{des\top}, \mathbf{q}^{des\top})^\top$  the desired pose,  $\text{Log}(\cdot) : \mathbb{S}^3 \mapsto \mathbb{R}^3$  the logarithmic map of the quaternion manifold,  $*$  the quaternion multiplication operator, and  $\bar{\mathbf{q}}$  the conjugate quaternion of  $\mathbf{q}$ . To facilitate the analysis, (1) is expressed in the compliance reference frame with respect to which the system is principal, i.e.  $\boldsymbol{\Lambda}, \mathbf{D}, \mathbf{K}$  are all diagonal. In this paper, we chose as compliance frame the tool centre point.

Starting from (1), we plan to modulate the desired stiffness  $\mathbf{K}$  (and, accordingly, the damping  $\mathbf{D}$ ) to generate some desired interaction behaviour depending both on the autonomy level and the task to accomplish. In this scenario, it is known that a variable stiffness controller may generate non passive control actions [12]. Therefore, we enforce the passivity of the power port  $\dot{\tilde{\mathbf{x}}}^\top \mathbf{h}^{ext}$  using energy tanks [12]. To this end, we consider the tank dynamics

$$\dot{s} = \frac{\sigma}{s} \dot{\tilde{\mathbf{x}}}^\top \mathbf{D} \dot{\tilde{\mathbf{x}}} - \frac{\mathbf{k}^\top}{s} \dot{\tilde{\mathbf{x}}}, \quad (3)$$

where  $s \in \mathbb{R}$  is the state of the tank,  $\sigma \in \{0, 1\}$  is a coefficient modulating the energy storage, and  $\mathbf{k}(t)$  is an extra input defined as:

$$\mathbf{k}(t) = \begin{cases} -\mathbf{K}^\delta \tilde{\mathbf{x}} & \text{if } T(s) > T_{min} \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

with  $T(s) = \frac{1}{2}s^2$  the tank energy, which activates the extra input whenever is contains a minimum energy quantity  $T_{min} \in \mathbb{R}^+$ . In addition, let  $\mathbf{K}^{min} \in \mathbb{R}^{6 \times 6}$  be the minimum allowed stiffness for the impedance controller, then  $\mathbf{K}^\delta \in \mathbb{R}^{6 \times 6}$  is the time-varying component of the stiffness, thus satisfying  $\mathbf{K}(t) = \mathbf{K}^{min} + \mathbf{K}^\delta(t)$ .

With (3) and (4), to guarantee the system passivity, it is sufficient to choose a stiffness  $\mathbf{K}$  that does not drive the energy tank below its minimum value  $T_{min}$ .

To modulate the stiffness of the impedance model (1), we cast the optimisation problem to a QP, thus ensuring that a solution can be found within each control period. In our framework, stiffness is modulated to accommodate two tasks: one aiming at producing a target wrench  $\mathbf{h}^{des} \in \mathbb{R}^6$  at the end-effector, and a second regulating a desired stiffness  $\mathbf{K}^{des} \in \mathbb{R}^{6 \times 6}$  [16]. The optimisation problem must take into account the energy tank constraint  $T(s) \geq T_{min}$ ; this condition alone, however, does not prevent the energy tank from an instantaneous drain by selecting a high-enough stiffness, leading to a complete loss of performance [17]. To overcome this limitation, we add a constraint to the power flow of the energy tank, limiting the energy extraction rate to a lower bound  $\rho \in \mathbb{R}^-$ , i.e.  $\dot{T}(s) \geq \rho$ . The resulting QP is formulated as

$$\begin{aligned} & \arg \min_{\mathbf{K} \in \mathbb{R}^{6 \times 6}} \frac{1}{2} \left( \|\mathbf{h}^{ext} - \mathbf{h}^{des}\|_{\mathbf{Q}}^2 + \|\mathbf{K} - \mathbf{K}^{des}\|_{\mathbf{R}}^2 \right) \\ & \text{subject to} \\ & \mathbf{K}^{min} \preceq \mathbf{K}^d \preceq \mathbf{K}^{max} \\ & \mathbf{h}^{min} \preceq \mathbf{h}^{ext} \preceq \mathbf{h}^{max} \\ & -\tilde{\mathbf{x}}^\top \mathbf{K} \dot{\mathbf{x}} \leq \sigma \tilde{\mathbf{x}}^\top \mathbf{D} \dot{\mathbf{x}} - \tilde{\mathbf{x}}^\top \mathbf{K}^{min} \dot{\mathbf{x}} + \frac{T_{t-1} - T_{min}}{\Delta t} \\ & -\tilde{\mathbf{p}}^\top \mathbf{K}_p \dot{\mathbf{p}} \leq \sigma \tilde{\mathbf{p}}^\top \mathbf{D}_p \dot{\mathbf{p}} - \tilde{\mathbf{p}}^\top \mathbf{K}_p^{min} \dot{\mathbf{p}} - \rho_p \\ & -\tilde{\mathbf{q}}^\top \mathbf{K}_q \dot{\mathbf{q}} \leq \sigma \tilde{\mathbf{q}}^\top \mathbf{D}_q \dot{\mathbf{q}} - \tilde{\mathbf{q}}^\top \mathbf{K}_q^{min} \dot{\mathbf{q}} - \rho_q, \end{aligned} \quad (5)$$

where  $\mathbf{Q}, \mathbf{R} \in \mathbb{R}^{6 \times 6}$  are diagonal positive definite weighting matrices,  $\mathbf{K}^{min}, \mathbf{K}^{max} \in \mathbb{R}^{6 \times 6}$  are diagonal matrices representing the minimum and maximum allowed stiffness,  $\mathbf{h}^{ext} \in \mathbb{R}^6$  is the wrench of the impedance interaction model (1), and  $\mathbf{h}^{max}, \mathbf{h}^{min} \in \mathbb{R}^6$  respectively the maximum and minimum wrench that the robot is allowed to exert. The symbol  $\preceq$  represents the matrix inequality. The third to last constraint avoids the complete drainage of the tank given the energy  $T_{t-1}$  at the previous control cycle that takes  $\Delta t$  to execute. The last equations instead constraint energy extraction rate; to take into account different scales of both position and orientation error, the power constraint has been divided in 2 individual pieces; in light of this,  $\mathbf{K}_p, \mathbf{K}_q \in \mathbb{R}^{3 \times 3}$  are the stiffness sub-blocks of  $\mathbf{K}$  associated respectively to the position and orientation error, and  $\rho_p, \rho_q \in \mathbb{R}^-$  are the power limits for both position and orientation.

### B. Incremental Learning of Periodic Tasks

To incrementally learn the trajectory of the end-effector, we exploit the Dynamic Movement Primitive (DMP) framework [3]. Periodic DMP encodes motion trajectories in a spring-damper like system

$$\begin{cases} \dot{z} = \Omega(\alpha(\beta(g - y) - z) + f(\phi)) \\ \dot{y} = \Omega z \\ \dot{\phi} = \Omega, \end{cases} \quad (6)$$

with  $\Omega$  the inverse of the task period,  $g$  the motion goal, and  $f(\phi)$  is a non-linear excitation term that must be learned from demonstration. Conventionally,  $f(\phi)$  is chosen as the

linear combination of  $N$  Gaussian kernels, i.e.

$$f(\phi) = \frac{\sum_{i=1}^N \Psi_i(\phi) w_i}{\sum_{i=1}^N \Psi_i(\phi)}, \quad (7)$$

with

$$\Psi_i(\phi) = \exp\left(h_i(\cos(\phi - c_i) - 1)\right). \quad (8)$$

Given that a demonstration  $y_d, \dot{y}_d, \ddot{y}_d$  provides a desired shape  $f_d$  for the forcing term as

$$f_d(t_j) = \frac{\ddot{y}_d(t_j)}{\Omega^2} - \alpha \left( \beta(g - y_d(t_j)) - \frac{\dot{y}_d(t_j)}{\Omega} \right), \quad (9)$$

it is possible to incrementally learn the weights  $w_i$  using a recursive least-squares method with forgetting factor  $\lambda$  as follows:

$$\begin{aligned} w_i(t_{j+1}) &= w_i(t_j) + \zeta \Psi_i(\phi_i) P_i(t_{j+1}) e(t_j) \\ e(t_j) &= f_d(t_j) - w_i(t_j) \\ P_i(t_{j+1}) &= \frac{1}{\lambda} \left( P_i(t_j) - \frac{P_i^2(t_j)}{\frac{\lambda}{\Psi_i} + P_i(t_j)} \right), \end{aligned} \quad (10)$$

where the system is initialised with  $w_i(0) = 0, P_i(0) = 1$  for  $i = 1, \dots, N$ . Note that in (10), the update is weighted also by the learning factor  $\zeta$  that is specific to the framework we propose.

While (6) is suitable to encode position variables, non-singular descriptions of orientation inherently come with constraints, which makes their learning non-trivial. To overcome this issue, we use the formulation proposed in [18] that exploits the Riemannian property of the quaternion manifold to encode orientation in a DMP:

$$\begin{cases} \dot{\boldsymbol{\eta}} = \boldsymbol{\Omega}(\alpha_z(\beta_z 2 \text{Log}(\mathbf{q} * \bar{\mathbf{q}}_d) - \boldsymbol{\eta}) + \mathbf{f}(\phi)) \\ \dot{\mathbf{q}} = \boldsymbol{\Omega}^{\frac{1}{2}} \boldsymbol{\eta} * \mathbf{q}, \end{cases} \quad (11)$$

with  $\boldsymbol{\Omega} = \boldsymbol{\Omega}_{\mathbb{I}_{3 \times 3}}$ . Differently from [18], quaternion periodic DMP learning is performed incrementally using (10) for each one of the 3 coefficients of  $\mathbf{f}(\phi) \in \mathbb{R}^3$ , thus no coupling in the forcing terms is considered.

### C. Online Period Estimation

A crucial aspect in learning periodic motions is the estimation of the period of the demonstrated trajectory. To address this issue, we augment the adaptive frequency oscillator (AFO) proposed in [15] with the learning rate  $\zeta$ . AOFs are dynamical systems whose goal is to learn the Fourier series, truncated at the first order, of the input signal, which is, in this case, one component of the desired trajectory. Let  $\hat{l} = \alpha\varphi\omega + \beta \cos \varphi \in \mathbb{R}$  be the first-order expansion of the signal  $l \in \mathbb{R}$ , and calling  $e = l - \hat{l} \in \mathbb{R}$  the corresponding error, then parameters  $\alpha, \beta, \omega \in \mathbb{R}$  are estimated integrating the dynamics:

$$\begin{cases} \dot{\varphi} = \omega - \zeta a e \sin \varphi \\ \dot{\omega} = -\zeta a e \sin \varphi \\ \dot{\alpha} = \zeta \eta e \cos \varphi \\ \dot{\beta} = \zeta \eta e \sin \varphi \end{cases}. \quad (12)$$

Here,  $\omega, \varphi \in \mathbb{R}$  are respectively the basic frequency and the phase of the signal, while  $\alpha, \beta$  are parameters modulating

the amplitude of the reconstructed signal;  $a \in \mathbb{R}^+$  is a coupling constant that drives the convergence of the frequency estimates, while  $\eta \in \mathbb{R}^+$  is the learning coefficient of the modulating terms. Once the dynamics of the AFO converges to the fundamental angular frequency  $\omega$  of the demonstration, the corresponding period  $T \in \mathbb{R}$  can be simply obtained as

$$T = \frac{2\pi}{\omega}. \quad (13)$$

In our case, multiple signal sources can be used to estimate the period of the motion, i.e. 3 for the Cartesian position and 4 for the quaternion trajectory of the end-effector. Even though each DMP is allowed to evolve at its own time constant  $\tau$ , we stick to a common period to avoid synchronisation issues between different tasks. To properly select the shared demonstration period  $T_d$ , we let each independent  $i$ -th source estimate its period  $T_i$  according to (12) and (13); then, adapting from [10], we select  $T_d$  as

$$T_d = \arg \max_i \left\{ T_i \frac{\bar{l}_i}{\max_j \{\bar{l}_j\}} \right\}, \quad (14)$$

where  $\bar{l}_i = S(\text{Var}\{l_i\}) \in (0, 1)$  is the logistic sigmoid function computed on the variance  $\text{Var}\{l_i\}$  of the  $i$ -th signal. Tuning the steepness  $a_l \in \mathbb{R}^+$  and the threshold  $\sigma_0^2 \in \mathbb{R}^+$  of

$$S(x) = \frac{1}{1 + e^{-a_l(x - \sigma_0^2)}}, \quad (15)$$

it is possible to effectively penalise period estimates associated to signals with low variance. For instance, while demonstrating a motion on the  $x, y$  plane, the AFO associated with the  $z$  coordinate, learning from noisy data, could converge to an unreasonable period. The scaling term  $\bar{l}_i / \max\{\bar{l}_j\}$  in (14) handles such conditions by artificially reducing the periods associated with less interesting signals, effectively discouraging their selection for the overall learning framework.

#### D. Autonomy Level Regulator

We introduced an autonomy factor  $\gamma \in [0, 1]$  which is a variable describing the ability of the system to autonomously execute the task ( $\gamma = 1$ ) or to stand still waiting for a new demonstration ( $\gamma = 0$ ). To this extent, we propose a modification of the automation dynamic proposed in [10] that embeds the energy tank state in order to relax requirements on the tracking error. In particular, we select the autonomy factor  $\gamma(t)$  dynamics as described by

$$\dot{\gamma} = \begin{cases} \max\{\gamma_r, 0\} & \gamma = 0 \\ \gamma_r & 0 < \gamma < 1, \\ \min\{\gamma_r, 0\} & \gamma = 1 \end{cases}, \quad (16)$$

with

$$\gamma_r = S(\|\dot{\mathbf{x}}\|_{\mathbf{G}}^2)(\gamma_{r,min} + k_r \gamma) \left( 1 - \nu_1 \|\dot{\mathbf{x}}\|_{\mathbf{L}}^2 - \nu_2 \max\{0, \xi - \xi_{min}\} \right). \quad (17)$$

Here,  $\gamma_{r,min}, k_r, \nu_1, \nu_2, \xi_{min} \in \mathbb{R}^+$  are all parameters that must be tuned, while  $S(\cdot)$  is a sigmoid function, as in (15), that is used to avoid premature learning while standing still.

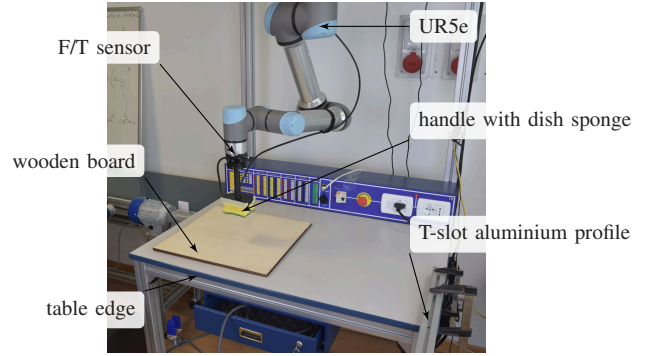


Fig. 2. Experimental setup consisting of a ceil-mounted UR5e manipulator, the SenONE BOTA System force-torque sensor, the 3D printed handle and the dish sponge. For the experiment, there is a floating wooden board and a table-fixed T-slot aluminium profile.

In (17), we consider an additional state  $\xi$  whose dynamic is associated to the power-flow of the energy tank according to the linear system

$$\dot{\xi} = \frac{1}{\tau_\xi} (\dot{T}(s) - \xi). \quad (18)$$

b With such choice, we effectively create a low-pass filter for the tank power signal, thus rejecting high-frequency disturbances that might arise during contact with rough surfaces. This, in addition with the threshold in (17), allows to quickly reset the autonomy factor whenever lots of energy is pumped in the tank, i.e., when the human is acting on the robot.

### III. EXPERIMENTS

#### A. Setup

The framework is evaluated in a proof-of-concept robotic setup, depicted in Fig. 2. The system is made of a ceil-mounted 6 DoF manipulator, the Universal Robot UR5e, with force-torque sensor, the SensONE BOTA System, mounted between the wrist and the end-effector. For the demonstration, the tool is made by a 3D-printed handle, at which end is attached a dish sponge used for the polishing operation.

The UR5e robot is position controlled, thus direct compliance control is not possible. To overcome this technological limitation, we employ the Forward Dynamics Compliance Control (FDCC) algorithm [19], a method that enables impedance, admittance, and force control in a single strategy for position-controlled robots.

To evaluate the proposed framework, we test the system while performing a wiping task on surfaces of different materials. To achieve the goal, we set as desired wrench for the QP problem (5) the one with only a positive action oriented along the positive  $z$  axis of the end-effector:

$$\mathbf{h}^{des} = (0, 0, 5\text{N}, 0, 0, 0)^\top. \quad (19)$$

To have a robot which is fully compliant while learning ( $\gamma = 0$ ), and stiff when autonomous ( $\gamma = 1$ ) to reduce tracking errors, the desired stiffness is modulated as

$$\mathbf{K}^{des} = \mathbf{K}^{min} + \gamma^2 (\mathbf{K}^{am} - \mathbf{K}^{min}), \quad (20)$$

with  $\mathbf{K}^{min} = \text{blkdiag}\{\mathbf{I}_{3 \times 3} \text{ N/m}, 0.001 \mathbf{I}_{3 \times 3} \text{ Nm/rad}\}$  and  $\mathbf{K}^{am} = \text{diag}\{600, 600, 600, 60, 60, 70\}$  the stiffness in autonomous mode; for the experiment, maximum allowed stiffness has been set to  $\mathbf{K}^{max} = \text{blkdiag}\{1000 \mathbf{I}_{3 \times 3}, 100 \mathbf{I}_{3 \times 3}\}$ .

To have a proper trade-off between the two complementary tasks in (5), tuning the weighting matrices  $\mathbf{Q}, \mathbf{R}$  is of key importance. In this case, we set the former to  $\mathbf{Q} = \text{blkdiag}\{3200 \mathbf{I}_{3 \times 3}, 32 \mathbf{I}_{3 \times 3}\}$ , while the latter is modulated, similarly to (20), as  $\mathbf{R} = \mathbf{R}^{min} + \gamma^2(\mathbf{R}^{am} - \mathbf{R}^{min})$ , with  $\mathbf{R}^{min} = \text{blkdiag}\{0.1 \mathbf{I}_{3 \times 3}, 0.01 \mathbf{I}_{3 \times 3}\}$  and  $\mathbf{R}^{am} = \text{diag}\{10, 10, 1.0, 10, 10, 10\}$ . These choices are made to guarantee maximum compliance for  $\gamma = 0$ , while in autonomous mode the controller can freely modulate the stiffness to track the force on the  $z$  component, while translation in the  $x, y$  plane and orientation is regulated through bigger stiffness gains. To complete the QP problem definition, we set  $\mathbf{h}^{max} = (40, 40, 40, 10, 10, 10)$ ,  $\mathbf{h}^{min} = -\mathbf{h}^{max}$ ,  $T_{min} = 0.1$ ,  $\rho_p = -1.1$ , and  $\rho_q = -1, 4$ .

Concerning the DMPs, we set  $\alpha = 48$ ,  $\beta = 12$ ,  $N = 25$  basis functions and forgetting factor  $\lambda = 0.999$  for both position and orientation DMP, while their goals are initialised to the starting configuration of the robot (to avoid unexpected movements when the robot starts with no excitation term learned). Gains of the adaptive frequency oscillator are set to  $a = 50$ ,  $\eta = 1$ , and to avoid divergence of the dynamical system, we constrained  $\omega \in [0.7, \pi]$  and  $|\alpha| = |\beta| \leq 0.3$ .

Finally, the parameters for the autonomy level dynamics are  $\gamma_{r,min} = 0.1$ ,  $k_r = 0.25$ ,  $\nu_1 = 450$ ,  $\nu_2 = 15$ ,  $\xi_{min} = 0.8$ ,  $\tau_\xi = 0.5$ ,  $\mathbf{L} = \text{diag}\{1, 1, 0.2, 0.1, 0.1, 0.1\}$  and  $\mathbf{G} = \text{blkdiag}\{\mathbf{I}_{3 \times 3}, 0.1 \mathbf{I}_{3 \times 3}\}$ . In the following demonstration, the learning rate  $\zeta$  is computed as  $(1 - \gamma)^2$ .

## B. Results and Discussion

To show the effectiveness of the proposed framework, we run experiments involving the teaching of different wiping tasks within the same robot execution, proving how the system can easily adapt to different requirements. In particular, Fig. 3 reports a demonstration in which the robot: 1) starts from a standstill configuration, then 2) the operator demonstrates how to clean the edge of the table by moving along it and after a few autonomous executions, 3) the system is taught a circular motion on a wooden surface. At this stage, 4) we show the safety of the proposed optimised variable impedance controller by moving the plane on which the robot is executing his wiping task. Finally, 5) the human operator demonstrates how to polish a T-slot aluminium profile with a motion that requires the end-effector re-orientation to properly carry out the task.

At the beginning of the demonstration, when the human takes over the robot, energy is injected in the tank and the autonomy factor drops to 0. One can observe a time delay between the proper energy injection and the factor drop, but this is due to the low-pass filter (18). Additionally, while learning from a new demonstration, the automation sometimes dips down: this is related to the adaptive frequency oscillator that, at the same moment, converges to the actual demonstration period. This sudden change in the estimated

period negatively affects the learning performance of the DMP, but convergence at the desired motion is restored quickly. To properly converge at the desired motion with no prior knowledge of the task, the system takes about 20s for planar tasks, i.e. 4 to 5 full demonstrations. In case of a demonstration involving the active learning of an orientation trajectory, the convergence rate is slower due to two primary reasons: firstly, the convergence rate of the orientation DMP is slower, and, secondly, the virtualisation mechanism that enables compliance control of a position-controlled robot generates an apparent rotational stiffness, increasing the difficulty of humans to demonstrate repeatable motion patterns. However, the system is able to converge even slowly in this scenario. Additionally, these issues can be mitigated using torque-controlled robots for which compliance control is properly defined, without the need for virtualisation mechanisms.

During autonomous execution, the optimisation problem modulates the stiffness in the  $x, y$  plane towards the  $\mathbf{K}^{am}$  value, leading to a tracking error within the  $2mm$  set for the application; to reduce tracking error, it is sufficient to increase  $\mathbf{K}^{am}$ . Concerning the force-tracking requirements along the  $z$  axis, actual applied forces stay mainly within the 5N error range. If one seeks better force regulation,  $\mathbf{R}^{am}$  can be accordingly adjusted. Still, during execution, we preferred a trade-off toward a higher stiffness to guarantee more contact in situations at which the DMP provides targets just slightly below the surface. Finally, the choice (20) of the desired stiffness is a valuable feedback for the demonstrator, since as the system gets autonomous, the stiffness is perceived back by the human.

## IV. CONCLUSION

In this paper, we present our preliminary work on a incremental learning framework that combines a safe variable impedance control, adaptive frequency oscillator, and dynamic movement primitives to perform wiping tasks. Differently from previous works in the field, the DMP also learns the orientation, thus, we extended the problem formulation of variable stiffness optimisation to the orientation. In doing so, we enabled the execution of the wiping task on any smooth surface, not necessarily flat ones. In addition, we propose a novel formulation for autonomy-level dynamics that embeds the energy-tank knowledge to relax tracking error satisfaction to drive the learning phase of the module, while, in the meantime, guaranteeing a good disturbance rejection. In future works, we will address some limitations of the proposed architecture. First, we will join the dynamic movement primitive module with the adaptive frequency oscillator to enable a faster convergence of the DMP towards the demonstrated task, even in conditions where end-effector re-orientation is required. Second, we will integrate continual learning techniques to avoid the inefficient loss of the acquired demonstration knowledge that, as of now, is discarded by the system.

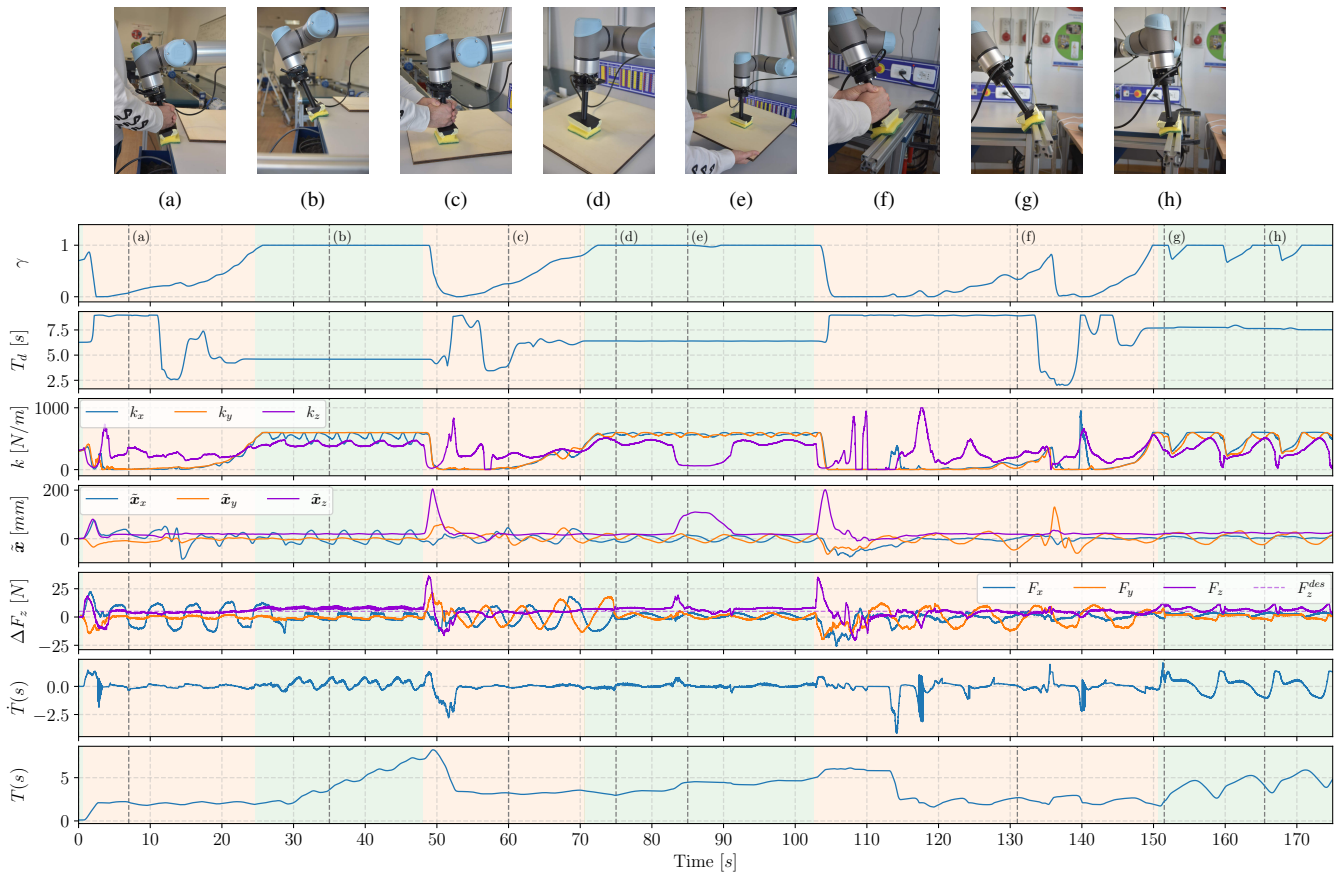


Fig. 3. Snapshots taken from the demonstration (available at <https://youtu.be/2fCSnqNbqz0>) and collected data: (a) the human demonstrates how to clean the edge of a table, and (b) the robot executes it. Then (c), the human demonstrates a circular motion on a wooden table, and (d) the robot executes it (e), even when the object is perturbed. Finally, (f) the human shows a motion that requires the tilting of the robot end-effector, as for the poses shown in (g) and (h). Green areas represent time windows where the robot is autonomous, while orange ones represent time windows when the human is interacting.

## REFERENCES

- [1] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," *Intelligent service robotics*, vol. 9, pp. 1–29, 2016.
- [2] 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.
- [3] M. Saveriano, F. J. Abu-Dakka, A. Kramberger, and L. Peternel, "Dynamic movement primitives in robotics: A tutorial survey," *The International Journal of Robotics Research*, vol. 42, no. 13, pp. 1133–1184, 2023.
- [4] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," *Advances in neural information processing systems*, vol. 26, 2013.
- [5] Y. Huang, L. Rozo, J. Silvério, and D. G. Caldwell, "Kernelized movement primitives," *The International Journal of Robotics Research*, vol. 38, no. 7, pp. 833–852, 2019.
- [6] A. G. Billard, S. Calinon, and R. Dillmann, "Learning from humans," *Springer handbook of robotics*, pp. 1995–2014, 2016.
- [7] L. Rozo, S. Calinon, D. G. Caldwell, P. Jimenez, and C. Torras, "Learning physical collaborative robot behaviors from human demonstrations," *IEEE Transactions on Robotics*, vol. 32, no. 3, pp. 513–527, 2016.
- [8] A. Gams, T. Petrič, M. Do, B. Nemeč, J. Morimoto, T. Asfour, and A. Ude, "Adaptation and coaching of periodic motion primitives through physical and visual interaction," *Robotics and Autonomous Systems*, vol. 75, pp. 340–351, 2016.
- [9] T. Kastritsi, F. Dimeas, and Z. Doulgeri, "Progressive automation with DMP synchronization and variable stiffness control," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3789–3796, 2018.
- [10] F. Dimeas and Z. Doulgeri, "Progressive automation of periodic tasks on planar surfaces of unknown pose with hybrid force/position control," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020, pp. 5246–5252.
- [11] K. Kronander and A. Billard, "Passive interaction control with dynamical systems," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 106–113, 2015.
- [12] F. Ferraguti, N. Preda, A. Manurung, M. Bonfe, O. Lambercy, R. Gassert, R. Muradore, P. Fiorini, and C. Secchi, "An energy tank-based interactive control architecture for autonomous and teleoperated robotic surgery," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1073–1088, 2015.
- [13] M. Saveriano, "An energy-based approach to ensure the stability of learned dynamical systems," in *IEEE International conference on robotics and automation*, 2020, pp. 4407–4413.
- [14] Y. Michel, M. Saveriano, and D. Lee, "A passivity-based approach for variable stiffness control with dynamical systems," *IEEE Transactions on Automation Science and Engineering*, 2023.
- [15] T. Petrič, A. Gams, A. J. Ijspeert, and L. Žlajpah, "On-line frequency adaptation and movement imitation for rhythmic robotic tasks," *The International Journal of Robotics Research*, vol. 30, no. 14, pp. 1775–1788, 2011.
- [16] J. Zhao, A. Giammarino, E. Lamon, J. Gandarias, E. Momi, and A. Ajoudani, "A Hybrid Learning and Optimization Framework to Achieve Physically Interactive Tasks With Mobile Manipulators," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, 2022.
- [17] L. Beber, E. Lamon, D. Nardi, D. Fontanelli, M. Saveriano, and L. Palopoli, "A passive variable impedance control strategy with viscoelastic parameters estimation of soft tissues for safe ultrasonography," in *IEEE International Conference on Robotics and Automation*, 2024.
- [18] F. J. Abu-Dakka, M. Saveriano, and L. Peternel, "Periodic dmp formulation for quaternion trajectories," in *IEEE International Conference on Advanced Robotics*, 2021, pp. 658–663.
- [19] S. Scherzinger, A. Roennau, and R. Dillmann, "Forward Dynamics Compliance Control (FDCC): A new approach to cartesian compliance for robotic manipulators," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2017-September, pp. 4568–4575, 12 2017.