

# A Hybrid Learning and Optimization Framework to Achieve Physically Interactive Tasks with Mobile Manipulators

Jianzhuang Zhao<sup>1,2,+</sup>, Alberto Giammarino<sup>1,+</sup>, Edoardo Lamon<sup>1</sup>, Juan M. Gandarias<sup>1</sup>, Elena De Momi<sup>2</sup>, and Arash Ajoudani<sup>1</sup>

**Abstract**—This paper proposes a hybrid learning and optimization framework for mobile manipulators for complex and physically interactive tasks. The framework exploits an admittance-type physical interface to obtain intuitive and simplified human demonstrations and Gaussian Mixture Model (GMM)/Gaussian Mixture Regression (GMR) to encode and generate the learned task requirements in terms of position, velocity, and force profiles. Next, using the desired trajectories and force profiles generated by GMM/GMR, the impedance parameters of a Cartesian impedance controller are optimized online through a Quadratic Program augmented with an energy tank to ensure the passivity of the controlled system. Two experiments are conducted to validate the framework, comparing our method with two approaches with constant stiffness (high and low). The results showed that the proposed method outperforms the other two cases in terms of trajectory tracking and generated interaction forces, even in the presence of disturbances such as unexpected end-effector collisions.

**Index Terms**—Compliance and Impedance Control, Mobile Manipulation, Imitation Learning.

## I. INTRODUCTION

**D**UE to the manufacturing shift from mass to customized production processes and the trend of an aging population in recent years, more flexibility and safety are required for robots in industrial and logistic scenarios. Collaborative Mobile Manipulators (CMM), which integrate a collaborative manipulator on a mobile platform, are designed with these objectives, combining the precision and manipulation capabilities of a robotic arm and the locomotion potential of the mobile platforms [1]. Because of these advantages, CMMs have demonstrated high suitability to different tasks, not only in industrial manufacturing and warehouse automation but also

Manuscript received: February, 24, 2022; Revised May, 15, 2022; Accepted June, 12, 2022. This paper was recommended for publication by Editor Clement Gosselin upon evaluation of the Associate Editor and Reviewers' comments. This work was supported in part by the European Research Council's (ERC) starting grant Ergo-Lean (GA 850932), and in part by European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 871237 (SOPHIA). <sup>+</sup> Contributed equally to this work.

<sup>1</sup> Jianzhuang Zhao, Alberto Giammarino, Edoardo Lamon, Juan M. Gandarias, and Arash Ajoudani are with the Human-Robot Interfaces and physical Interaction, Istituto Italiano di Tecnologia, Genoa, Italy jianzhuang.zhao@iit.it; alberto.giammarino@iit.it; edoardo.lamon@iit.it; juan.gandarias@iit.it; arash.ajoudani@iit.it.

<sup>2</sup> Jianzhuang Zhao and Elena De Momi are with the Dept. of Electronics, Information and Bioengineering, Politecnico di Milano, Italy elena.demomi@polimi.it.

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

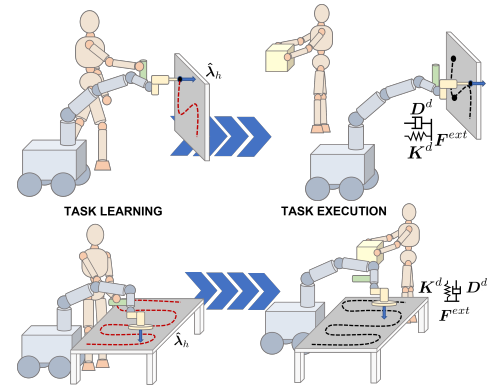


Fig. 1. Left side: teaching a complex physically interactive task. Right side: Execution of the learned task adapting the Cartesian impedance.

in service and domestic applications [1], [2]. Nevertheless, integrating different components makes the control and trajectory planning harder for CMM. Aiming at overcoming the synchronization issues of decoupled strategies, whole-body controllers have been presented. In particular, whole-body torque-based strategies could regulate not only the interaction forces with the environment but also the motion distribution at the joint level, generating different motion patterns [3], [4].

In addition, when facing unstructured environments, where interactions may change unexpectedly, collaborative robots (cobots) are expected to adapt their performance online to unforeseen situations. While dealing with physically interactive tasks (see Fig. 1), hybrid position-force approaches are favorable w.r.t. pure force control for safety reasons. Nevertheless, such schemes presented an unstable behavior during the contact phase [5]; hence it appears convenient to render virtual forces through an indirect method such as impedance control.

To model the system response to interactions with the environment, impedance controllers require the tuning of a high number of design parameters (generally stiffness, damping and inertia). Usually, high impedance is required in case of precise motions, whereas lower impedance is favorable in case of interactions with the environment. However, how to optimally tune impedance parameters, namely variable impedance control (VIC), is still an open issue. For instance, in [6] an impedance model based on a mass-damper system is used to express the desired dynamics of the error between the guessed

position of the environment and the actual desired position, which is the one allowing force tracking. In particular, based on the force error the damping is tuned online and the desired position obtained from the resulting dynamic model is sent to a position-controlled robot. Although this method can achieve force tracking on different surfaces, the desired trajectory computed is altered based on the force error (e.g., when the surface contact is lost), which may result in an unsafe behavior of the robot. Differently, authors in [5], [7] propose different strategies to self-tune the impedance parameters according to the position error. While a high tracking error increases the stiffness, the interaction force is exploited to detect unplanned contacts and recover a compliant behavior. In [5], moreover, the authors try to further minimize the interaction with the environment. Nevertheless, sometimes high interaction forces are a task requirement and hence becoming compliant is not a suitable strategy. Additionally, human-in-the-loop approaches allow for online tuning of the robot behavior (varying damping and inertia) transferring the human skills for specific collaborative tasks as a trade-off between precision and speed of execution [8]. Other methods exploit EMG sensors to estimate joint stiffness and generate reference trajectories at the same time [9], [10]. Although these human impedance-transfer approaches for fixed base robotic arms can benefit from real-time adaptation to the changing environment, especially in teleoperation, EMG sensors and motion capture systems are required, making them unsuitable for industrial and domestic applications.

Furthermore, to generate autonomous tasks for CMM, optimal impedance profiles are not sufficient, also reference trajectories should be defined. Imitation Learning (IL) paradigm has proven to be an efficient way to obtain desired motion given a predefined set of trajectories demonstrated by a teacher [11]. The IL paradigm includes different algorithms to encode the taught demonstrations, such as Dynamical Movement Primitives (DMPs) [12] and Gaussian Mixture Model (GMM) / Gaussian Mixture Regression (GMR) [11], among the most used. In the context of cobots, one IL methodology, i.e. kinesi-*thetic* teaching [11], has gained increasing popularity since it does not require any retargeting from teacher to the robot. Its implementation usually requires gravity compensation or admittance control through force sensing (force-torque (F/T) sensors at the end-effector or joint torque sensors).

To the best of the authors knowledge, currently there exist few intuitive interfaces that fully exploit the potential of CMMs in kinesi-*thetic* teaching. In the context of fixed-base manipulators, several works faced the problem of simultaneous learning of trajectories and force profiles through impedance models [13] using GMM/GMR. In [14], to generate safe robot behavior, the variable stiffness is tuned according to the inverse of the position covariance of the GMM. In [15], two different semi-positive definite (SPD) stiffness representations are investigated, while stiffness profiles are obtained with a least-squares estimator using a linear interaction model, based on position/force demonstrations. A similar approach based on regularized regression and force sensing was used by the authors in [16], where the GMM maps the relationship between external force (input) and stiffness (output). Then,

the stiffness is generated online by GMR based on current external force during teleoperation, where high external force will result in high stiffness. Nonetheless, the high stiffness of robots may hurt humans if the large force is generated by human disturbances. A different methodology, based on reinforcement learning (RL), models the robot policy by DMP that includes trajectory and impedance parameters. The policy parameters are then offline optimized using a variant of policy improvement with path integrals [17]. However, this RL method and the iteration optimization method in [5] requires several simulated executions of the desired task and they cannot deal with unforeseen disturbances online. In addition, most of the methods in the literature do not directly ensure the stability of the VIC, whose property is fundamental for a safe Human-Robot Collaboration (HRC) [13].

To address these issues, we propose a method to learn loco-manipulation tasks exploiting the paradigm of VIC (see Fig. 1). Our goal is to teach our Mobile Collaborative robotic Assistant (MOCA) from human physical demonstrations through IL and generate robust and safe autonomous behaviors by regulating the interaction forces between the robot and the unstructured environment in quasi-static conditions. Thanks to an admittance-type interface [18], complex interactive tasks are intuitively demonstrated and then encoded by a GMM that learns the task trajectory and interaction force maps between the robot and the environment simultaneously. The desired trajectory and force profile are generated using GMR. Based on the learned force, the desired impedance parameters are computed online as the result of a quadratic program (QP), which ensures the lowest stiffness required to achieve the task, subject to force and stiffness limits imposed by safety requirements. Moreover, an energy tank-based passivity constraint is added to ensure the controller's stability. The desired trajectory and generated stiffness are finally sent to the MOCA weighted whole-body impedance controller (see Fig. 2).

The method is evaluated with a table cleaning task, where two different experiments are conducted. First of all, the task is taught in nominal conditions. Then, MOCA replicates the learned task in the same conditions and in the presence of external disturbances coming from the environment and from unexpected physical interactions with the human. The results show that the impedance parameters' tuning yields similar (good) performances in all conditions, allowing at the same time trajectory and force tracking and robust and compliant interactions compared to constant stiffness solutions.

## II. PRELIMINARIES

### A. MOCA Platform and Whole-Body Impedance Control

MOCA is a research robotic platform designed for HRC, with loco-manipulation capabilities that make it potentially suitable for logistics [3] and flexible manufacturing [19]. It is composed of the lightweight torque-controlled 7-DoFs Franka Emika Panda robotic arm, mounted on top of the velocity-controlled 3-DoFs Robotnik SUMMIT-XL STEEL mobile platform.



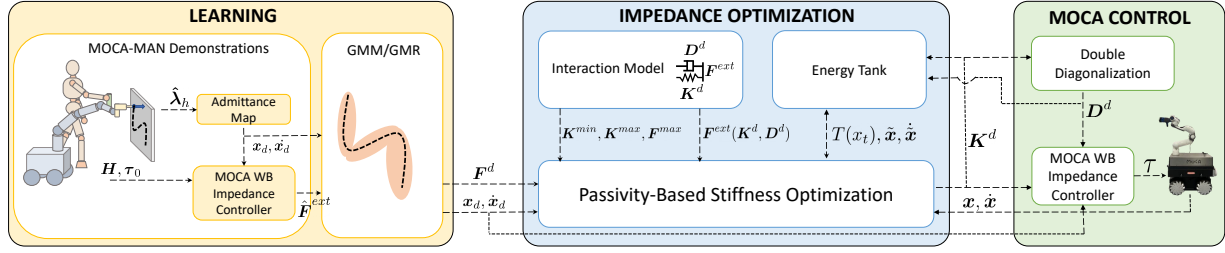


Fig. 2. The proposed framework scheme. Humans can teach MOCA the interactive tasks directly through an admittance-type physical interface, and the desired trajectories and interaction force are replicated with GMM/GMR. Then, the desired force is sent to a QP-based algorithm to optimize online stiffness. A tank energy constraint ensures the passivity of the system. Finally, the desired trajectory and stiffness are sent to MOCA's whole-body impedance controller.

The whole-body dynamic model [4], [20] can be formulated as:

$$\begin{pmatrix} M_v & \mathbf{0} \\ \mathbf{0} & M_a(q_a) \end{pmatrix} \begin{pmatrix} \ddot{q}_m \\ \ddot{q}_a \end{pmatrix} + \begin{pmatrix} D_v & \mathbf{0} \\ \mathbf{0} & C_a(q_a, \dot{q}_a) \end{pmatrix} \begin{pmatrix} \dot{q}_m \\ \dot{q}_a \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ g_a(q_a) \end{pmatrix} = \begin{pmatrix} \tau_m^{vir} \\ \tau_a \end{pmatrix} + \begin{pmatrix} \tau_m^{ext} \\ \tau_a^{ext} \end{pmatrix}, \quad (1)$$

where  $M_v \in \mathbb{R}^{n_b \times n_b}$  and  $D_v \in \mathbb{R}^{n_b \times n_b}$  are the virtual inertia and virtual damping of the mobile platform,  $\dot{q}_m \in \mathbb{R}^{n_b}$  is its input velocity,  $\tau_m^{ext} \in \mathbb{R}^{n_b}$  and  $\tau_m^{vir} \in \mathbb{R}^{n_b}$  are the related external and the virtual torque. With respect to the arm,  $q_a, \dot{q}_a$ , and  $\ddot{q}_a \in \mathbb{R}^{n_a}$  are the joint angles, velocities and accelerations vectors,  $M_a \in \mathbb{R}^{n_a \times n_a}$  is the symmetric and positive definite inertia matrix of the arm,  $C_a \in \mathbb{R}^{n_a}$  is the Coriolis and centrifugal force vector,  $g_a \in \mathbb{R}^{n_a}$  is the gravity,  $\tau_a \in \mathbb{R}^{n_a}$ , and  $\tau_a^{ext} \in \mathbb{R}^{n_a}$  are the commanded torque vector and external torque vector, respectively. This model can be summarized by

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) = \tau + \tau^{ext}. \quad (2)$$

where  $M(q) \in \mathbb{R}^{n \times n}$  ( $n = n_a + n_b$ ) is the symmetric positive definite joint-space inertia matrix,  $C(q, \dot{q}) \in \mathbb{R}^{n \times n}$  is the joint-space Coriolis/centrifugal matrix, and  $g(q) \in \mathbb{R}^n$  the joint-space gravity. Finally,  $\tau \in \mathbb{R}^n$  and  $\tau^{ext} \in \mathbb{R}^n$  represent joint-space input and external torque.

The MOCA Cartesian impedance controller is formulated as a prioritized weighted inverse dynamics algorithm and can be obtained by solving the problem of finding the torque vector  $\tau \in \mathbb{R}^m$  ( $m \leq 6$ ), according to the norm induced by the positive definite weighting matrix  $W \in \mathbb{R}^{n \times n}$ ,

$$\min_{\tau \in \mathbb{R}^n} \frac{1}{2} \|\tau - \tau_0\|_W^2 \quad \text{s.t. } F = \bar{J}^T \tau \quad (3)$$

where  $\bar{J} = M^{-1}J^T\Lambda$  is the dynamically consistent pseudo-inverse of the Jacobian matrix  $J(q)$ , and the constraint  $F = \bar{J}^T \tau = \Lambda JM^{-1}\tau$  is the general relationship between the generalized joint torques and the operational forces, and  $\Lambda = (JM^{-1}J^T)^{-1} \in \mathbb{R}^{m \times m}$  is the Cartesian inertia. The closed-form solution results in:

$$\tau = W^{-1}M^{-1}J^T\Lambda_W\Lambda^{-1}F + (I - W^{-1}M^{-1}J^T\Lambda_W JM^{-1})\tau_0, \quad (4)$$

where  $\Lambda_W = (JM^{-1}W^{-1}M^{-1}J^T)^{-1}$  is the weighted Cartesian inertia, analogous to the Cartesian inertia  $\Lambda$ . The weighting matrix  $W$  is generally defined as  $W(q) = H^T M^{-1}(q)H$ , where  $H \in \mathbb{R}^{n \times n}$  is the tunable positive definite weight matrix of the controller, that is used to generate different motion modes [3].

Finally,  $F$  is computed to generate the desired closed-loop behaviour, according to the Cartesian impedance law:

$$F = -D^d \dot{\tilde{x}} - K^d \tilde{x}, \quad (5)$$

where  $\tilde{x} = x_d - x \in \mathbb{R}^6$  is the Cartesian pose error computed with respect to the desired Cartesian equilibrium pose  $x_d$ , and  $D^d, K^d \in \mathbb{R}^{m \times m}$  are the desired Cartesian damping and stiffness matrices, respectively. Moreover,  $\tau_0$  could contain different contributions, such as joint impedance, collision and self-collision avoidance, joint limits avoidance, etc.

### B. GMM & GMR

The desired trajectories and force profiles are generated using GMM and GMR. This choice is due to the low number of hyperparameters that need to be tuned (i.e. only  $K$ , the number of Gaussians). The parameters of GMM can be estimated by Expectation-Maximization (EM) algorithm [21] with an offline training process that makes use of the demonstrations. To make the representation easier to understand, we give the following definition:  $\eta^{\mathcal{I}}$  and  $\eta^{\mathcal{O}}$  respectively denote the input and output variables on which the training is carried out, where the superscripts  $\mathcal{I}$  and  $\mathcal{O}$  stand for their dimensions, respectively. Given an input variable  $\eta^{\mathcal{I}}$ , the best estimation of output  $\hat{\eta}^{\mathcal{O}}$  is the mean  $\hat{\mu}$  of the conditional probability distribution  $\hat{\eta}^{\mathcal{O}} | \eta^{\mathcal{I}} \sim \mathcal{N}(\hat{\mu}, \hat{\Sigma})$ , which is computed by GMR [11].

## III. METHODOLOGY

### A. Desired Trajectory and Interaction Force Hybrid Learning Exploiting an admittance-type physical interface.

Aiming at demonstrating desired end-effector trajectories and interaction wrenches to MOCA, we integrated kinesthetic teaching with an admittance-type physical interface. The interface is depicted in Fig. 3. It consists of i) an Arduino Nano microcontroller connected to a 4-buttons panel that allows the user to configure different functionalities and communicate with the robot through the Robot Operating System (ROS)

middleware suite, ii) a F/T sensor to measure the user interaction wrenches with a physical part that the human can easily grasp and iii) an end-effector tool. A detailed explanation of the hardware and design of an admittance-type physical interface can be found in [18].

The measured human wrenches  $\hat{\lambda}_h \in \mathbb{R}^m$ , are used to change the input Cartesian equilibrium pose in (5) as input of an admittance controller which implements the following desired dynamics:

$$M^{adm}\ddot{x}_d + D^{adm}\dot{x}_d = \hat{\lambda}_h, \quad (6)$$

where  $M^{adm}, D^{adm} \in \mathbb{R}^{m \times m}$  are respectively the admittance mass and damping. This way, MOCA can move in space in the same direction of the applied wrench. During the demonstration, the aforementioned buttons enable the human teacher to control some functionalities of the admittance mapping in (6) and of the whole-body controller presented in section II-A. In particular, the human has control over the loco-manipulation behavior of the robot (through selection of  $\tau_0$  and  $H$ ) and over the desired admittance ( $M^{adm}$  and  $D^{adm}$ ). In addition, the human can also enable and disable rotations and translations of the end-effector.

The additional features introduced by an admittance-type physical interface allow to exploit the full potential of a mobile base robot and ease the demonstration process, considering the human preference in different parts of the demonstrated behavior. Thanks to selecting the loco-manipulation mode, the human teacher can move the robot in a theoretically infinite workspace during the demonstration and keep the mobile base still when an interaction of the end-effector with the environment is planned. Indeed, since the mobile base movements, controlled at lower frequency w.r.t. the arm, perturb the end-effector motion, demonstrating a desired interaction with the environment can result challenging. Then, the human teacher can tune the system's responsiveness to the interaction by selecting the admittance according to his/her experience, preference, and specific requirements of the task: high admittance for long unconstrained motion and low admittance for short and precise interactions with the environment. Finally, selecting only a subset of the task space axes could help execute the task and speed up the learning procedure.

Desired Cartesian equilibrium pose  $x_d$  and twist  $\dot{x}_d$  are obtained as the solution of (6). Since the robotic arm used is torque controlled, the interaction wrenches of the robot end-effector with the environment  $\hat{F}^{ext}$  can be estimated through the joint torque measurements of the arm. Hence, during the demonstration, the human teacher can guide the robot through the admittance-type physical interface while the robot arm estimates its interaction with the environment. Then, a GMM is trained from the data obtained from the demonstrations, which input variable is time  $\eta^T = t$  and output variables are  $\eta^O = \begin{bmatrix} x_d^T & \dot{x}_d^T & \hat{F}^{ext^T} \end{bmatrix}^T$ . Finally, the desired trajectory and force can be generated by GMR.

### B. QP-Based Stiffness Online Optimization

The outputs of the GMR are fed to a QP that allows to online modulate the stiffness of the Cartesian whole-body impedance

controller presented in section II-A. The QP is formulated as follows:

$$\begin{aligned} \min_{\substack{K_i^d \in \mathbb{R}^{m \times m} \\ i \in \{1, \dots, N\}}} & \frac{1}{2} \sum_{i=1}^N \left( \|F_i^{ext} - F_i^d\|_Q^2 + \|K_i^d - K^{min}\|_R^2 \right) \\ \text{s.t.} & \quad K^{min} \leq K_i^d \leq K^{max} \quad i \in \{1, \dots, N\} \\ & \quad -F^{max} \leq F_i^{ext} \leq F^{max} \quad i \in \{1, \dots, N\} \end{aligned} \quad (7)$$

where  $i$  is the time step,  $N$  is the length of the time window,  $Q$  and  $R \in \mathbb{R}^{m \times m}$  are diagonal positive definite weighting matrices,  $K_i^d \in \mathbb{R}^{m \times m}$  is the desired stiffness of the Cartesian whole body impedance controller at time step  $i$ ,  $K^{min}$  and  $K^{max} \in \mathbb{R}^{m \times m}$  are respectively minimum and maximum allowed stiffness,  $F_i^{ext} \in \mathbb{R}^m$  is the wrench of the impedance interaction model at time step  $i$ , that can be modeled with impedance-like laws,  $F_i^d \in \mathbb{R}^m$  is the learned desired interaction wrench at time step  $i$  (i.e., output of the GMR) and  $F^{max} \in \mathbb{R}^m$  is the maximum wrench that the robot can exert. The constraint inequality between vectors is element-wise. The optimization problem formulated above trades off the tracking of the desired wrench with the requirement of keeping a small stiffness.

The desired impedance interaction model can be expressed in different possible ways. To render a desired mass-spring-damper system,  $F^{ext}$  is expressed as:

$$F^{ext} = \Lambda^d \ddot{\tilde{x}} + D^d \dot{\tilde{x}} + K^d \tilde{x}, \quad (8)$$

where  $\Lambda^d \in \mathbb{R}^{m \times m}$  is the desired inertia matrix. However, the interaction wrench must be measured precisely to render the desired inertia matrix, which is referred to as inertia shaping (e.g., using an F/T sensor). Unfortunately, a precise measure is often not available, and the following interaction model is rendered [22]:

$$F^{ext} = \Lambda(x) \ddot{\tilde{x}} + (\mu(x, \dot{\tilde{x}}) + D^d) \dot{\tilde{x}} + K^d \tilde{x}, \quad (9)$$

where no inertia shaping is performed, since the actual Cartesian inertia of the manipulator  $\Lambda(x)$  is used. Note that, in order to keep the physical coherence of the interaction model, the Coriolis and centrifugal terms  $\mu(x, \dot{\tilde{x}})$  must be added to the desired damping since those arise from a configuration dependent inertia [8], [22]. In practice, since also the acceleration signal is often noisy, the following simplified model is used:

$$F^{ext} = (\mu(x, \dot{\tilde{x}}) + D^d) \dot{\tilde{x}} + K^d \tilde{x}. \quad (10)$$

### C. Tank Energy Based Passivity Constraint

The QP formulated in section III-B computes a Cartesian stiffness at each time step that is sent to the Cartesian impedance whole-body controller presented in section II-A. It is well known that VIC might violate the passivity of the system [23]; thus, the stability of the controlled system is not guaranteed. In order to ensure system stability, the QP formulation previously presented can be augmented through a passivity constraint for the power port  $\dot{x}F^{ext}$ . The interaction model of the variable Cartesian impedance can be written as a port-Hamiltonian system and augmented through an energy

tank [23]. The scalar differential equation that describes the tank dynamics is:

$$\dot{x}_t = \frac{\sigma}{x_t} \dot{\mathbf{x}}^T \mathbf{D}^d \dot{\mathbf{x}} - \frac{\mathbf{w}^T}{x_t} \dot{\mathbf{x}}, \quad (11)$$

where  $x_t \in \mathbb{R}$  is the state of the tank that stores energy  $T(x_t) = \frac{1}{2}x_t^2$ ,  $\sigma \in \{0, 1\}$  is used to enable and disable the dissipated energy storage in case a maximum limit is reached, and  $\mathbf{w}$  is the extra input of the port-Hamiltonian dynamics which can be written as:

$$\mathbf{w}(t) = \begin{cases} -\mathbf{K}^v(t)\tilde{\mathbf{x}} & \text{if } T(x_t) > \varepsilon \\ 0 & \text{otherwise,} \end{cases} \quad (12)$$

where  $\mathbf{K}^v(t)$  is the variable part of the stiffness so that  $\mathbf{K}^d(t) = \mathbf{K}^{min} + \mathbf{K}^v(t)$  and  $\varepsilon > 0$  is the minimum energy that the tank is allowed to store. The tank energy is initialized so that  $T(x_t(0)) > \varepsilon$ . The constraint used to augment the QP is derived by integrating the energy tank over time and enforcing it to be higher than its minimum  $\varepsilon$ :

$$T(x_t(t)) = T(x_t(t-1)) + \dot{T}(x_t(t))\Delta t > \varepsilon, \quad (13)$$

where  $\Delta t$  is the time step and

$$\dot{T}(x_t) = x_t \dot{x}_t = \sigma \dot{\mathbf{x}}^T \mathbf{D}^d \dot{\mathbf{x}} - \mathbf{w}^T \dot{\mathbf{x}} \quad (14)$$

$$= \begin{cases} \sigma \dot{\mathbf{x}}^T \mathbf{D}^d \dot{\mathbf{x}} + \tilde{\mathbf{x}}^T \mathbf{K}^v \dot{\mathbf{x}} & \text{if } T(x_t) > \varepsilon \\ \sigma \dot{\mathbf{x}}^T \mathbf{D}^d \dot{\mathbf{x}} & \text{otherwise.} \end{cases} \quad (15)$$

Note from (15) that the constraint can be expressed as a linear function of the QP optimization variable only when  $T > \varepsilon$ . Instead, when  $T \leq \varepsilon$  the stiffness is constrained to take on its minimum value ( $\mathbf{K}^d = \mathbf{K}^{min}$ ).

#### D. Technical Details

While in the previous subsection, the method has been explained with generality, some simplifying hypotheses have been assumed in the implemented method. First of all, in the QP, the interaction model in (10) is used, where  $\mathbf{K}^d$ ,  $\mathbf{D}^d$  are assumed diagonal and  $\dot{\mathbf{x}}_d$  is assumed equal to zero to comply with the whole-body controller formulation. Moreover, only the translational part of the interaction model is considered in the optimization, while the rotational part is kept constant ( $m = 3$ ). The value of  $\mathbf{D}^d$  in the optimization at the next control loop is computed through double diagonalization at the previous optimization step  $\mathbf{K}^d$ , with critical damping factor [22], where the diagonal elements are:

$$\mathbf{d}^d(t) = 2 \cdot 0.707 \cdot \sqrt{\mathbf{k}^d(t-1)}, \quad (16)$$

where  $\mathbf{d}^d(t) \in \mathbb{R}^3$  is the vector of the diagonal components of the desired Cartesian damping matrix at time  $t$  and  $\mathbf{k}^d(t-1) \in \mathbb{R}^3$  is the vector of the diagonal components of the desired Cartesian stiffness matrix at time  $t-1$ . The damping is computed using the desired stiffness at the previous time step to preserve the cost function's quadratic nature in the QP formulation. Finally, only a one-time step is considered in each optimization ( $N = 1$ ).

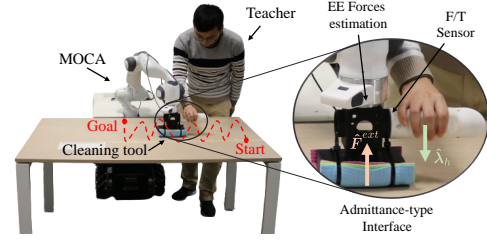


Fig. 3. Experimental setup. The teacher demonstrates the table cleaning task to MOCA from Start to Goal. A cleaning tool is attached to the end-effector of MOCA. The end-effector forces  $\mathbf{F}^{ext}$  are estimated thanks to the torque sensors integrated in the robot joints. An additional F/T sensor is included in the admittance-type interface to measure human interaction wrenches  $\hat{\lambda}_h$  and decouple them from  $\mathbf{F}^{ext}$ .

The problem in (7) can be rewritten as:

$$\begin{aligned} \min_{\mathbf{k}^d \in \mathbb{R}^3} & \frac{1}{2} \left( \|\mathbf{F}^{ext} - \mathbf{F}^d\|_Q^2 + \|\text{diag}\{\mathbf{k}^d\} - \mathbf{K}^{min}\|_R^2 \right) \\ \text{s.t.} & \quad \mathbf{k}^{min} \leq \mathbf{k}^d \leq \mathbf{k}^{max} \\ & \quad -\mathbf{F}^{max} \leq \mathbf{F}^{ext} \leq \mathbf{F}^{max} \\ & \quad T(x_t) \geq \varepsilon \end{aligned} \quad (17)$$

where  $\text{diag}\{\cdot\}$  is the diagonal operator and the vector inequalities are elementwise. Note that both  $\mathbf{F}^{ext}$  and  $T(x_t)$  are linear functions of the optimization variable  $\mathbf{k}^d$ , thus the last two inequality constraints can be easily expressed in the generic form  $\mathbf{C}\mathbf{k}^d \leq \mathbf{d}$ . Moreover, if  $T(x_t) < \varepsilon$ , then  $\mathbf{k}^d = \mathbf{k}^{min}$ . Note also that, since only the translational part of the impedance model is considered, the GMM is trained to encode only the linear part of velocities and forces.

## IV. EXPERIMENTS AND RESULTS

### A. Experimental Setup

The experiments carried out in this work consist of human demonstrations to train the GMM model and autonomous repetition of a table cleaning task in two different conditions, i.e. without and with external disturbances. For each condition, three stiffness settings are tested: i) low constant stiffness (LS), where  $\mathbf{k}^d = \mathbf{k}^{min}$ , ii) high constant stiffness (HS), where  $\mathbf{k}^d = \mathbf{k}^{max}$  and iii) optimized stiffness (OS), where  $\mathbf{k}^d$  is found online through our QP formulation in (17). The QP solution was computed in C++ using ALGLIB QP-BLEIC solver<sup>1</sup> on Ubuntu 18.04. All experiments were run on a computer with an Intel Core i7-4790S 3.2 GHz  $\times$  8-cores CPU and 16 GB RAM. The same desired end-effector trajectory, generated by GMR, is used for all the stiffness settings, while the desired interaction force is employed only by OS.

The experimental setup for the human demonstrations is shown in Fig. 3 along with the path followed by the human. The human teacher grasps an admittance-type physical interface and can use the four buttons as described in section III-A. The demonstrated trajectory starts from the *start*, which coincides with the top-left part of the table from the human viewpoint. Then, the human guides the robot end-effector to clean the table, moving parallel to the shortest side of the

<sup>1</sup>[www.alglib.net/optimization/quadraticprogramming.php](http://www.alglib.net/optimization/quadraticprogramming.php)

table, keeping the sponge in contact with the table's surface. When the bottom of the table is reached, the human guides the robot in free motion to the top of the table, shifted to the right w.r.t. the previous trajectory of an amount equal to the cleaning tool width. This procedure is repeated six times until the *goal* is reached. Three demonstrations are performed, all from the same human teacher, where desired position and interaction force of the end-effector with the environment are recorded and are later used to train the GMM.

A constant Cartesian impedance for the whole-body controller of MOCA is used during the demonstrations, with a stiffness of  $\mathbf{K}^d = \text{diag}\{500, 500, 500, 50, 50, 50\}$  and a damping computed through the double diagonalization formula. Also, joint stiffness is added as secondary task to prevent joint limits reaching and low manipulability configurations,  $\boldsymbol{\tau}_0 = -\mathbf{D}_0 \dot{\mathbf{q}} - \mathbf{K}_0(\mathbf{q} - \mathbf{q}_0)$ , where  $\mathbf{K}_0, \mathbf{D}_0 \in \mathbb{R}^{n \times n}$  are joint stiffness and damping, while  $\mathbf{q}_0 \in \mathbb{R}^n$  is the desired joint configuration. The teacher can change the admittance level to three discrete values corresponding to low ( $\mathbf{M}^{adm} = \text{diag}\{6, 6, 6\}$ ,  $\mathbf{D}^{adm} = \text{diag}\{40, 40, 40\}$ ), medium ( $\mathbf{M}^{adm} = \text{diag}\{4, 4, 4\}$ ,  $\mathbf{D}^{adm} = \text{diag}\{30, 30, 30\}$ ) and high ( $\mathbf{M}^{adm} = \text{diag}\{2, 2, 2\}$ ,  $\mathbf{D}^{adm} = \text{diag}\{20, 20, 20\}$ ) admittance using a button of an admittance-type physical interface. However, a constant high admittance is always chosen, in order to make the human forces negligible w.r.t. the ones coming from the environment. Moreover, the teacher can also switch between locomotion (large mobile base movements) and manipulation (mobile base fixed) mode, since a button of an admittance-type physical interface allows to change  $\mathbf{H}$ ,  $\mathbf{K}_0$  and  $\mathbf{q}_0$  simultaneously. Particularly,  $\mathbf{H}$  and  $\mathbf{K}_0$  can take on two discrete values ( $\mathbf{H} = \text{diag}\{10 \cdot \mathbf{1}_{n_b}, 2 \cdot \mathbf{1}_{n_a}\}$ ,  $\mathbf{K}_0 = \text{diag}\{2 \cdot \mathbf{1}_n\}$  for manipulation and  $\mathbf{H} = \text{diag}\{2 \cdot \mathbf{1}_{n_b}, 10 \cdot \mathbf{1}_{n_a}\}$ ,  $\mathbf{K}_0 = \text{diag}\{50 \cdot \mathbf{1}_n\}$  for locomotion), while  $\mathbf{q}_0$  is set to the current arm configuration when a switch between the two modes occurs.

The same experimental setup of the demonstrations is used for the autonomous repetition of the task. Two conditions are tested for each stiffness setting (LS, HS, and OS). In the first one, i.e., the autonomous task repetition without external disturbances, the robot repeats the demonstration in the same environment displayed during the demonstration. While in the second one, i.e., the autonomous task repetition with external disturbances, the environment is perturbed during the execution of the task. As previously mentioned, the cleaning task consists of 6 cleaning movements that go from the top to the bottom of the table, alternated with five free motions to return to the top of the table before starting the next cleaning movement. The following perturbations are applied during the experiment (see also Fig. 4 (bottom)):

- 1) Single slow table lifting and lowering during the 1<sup>st</sup> cleaning movement.
- 2) Single fast table lifting and lowering during the 2<sup>nd</sup> cleaning movement.
- 3) Repeated low frequency and high amplitude table lifting and lowering during the 3<sup>rd</sup> cleaning movement.
- 4) Repeated high frequency and low amplitude table lifting and lowering during the 4<sup>th</sup> cleaning movement.
- 5) Collision with a human during the 5<sup>th</sup> cleaning motion.

- 6) Collision with a human during the last free motion.

The values of the parameters used during the autonomous repetition experiments are  $\mathbf{k}^{min} = [200, 200, 200]$ ,  $\mathbf{k}^{max} = [1000, 1000, 1000]$ ,  $\mathbf{F}^{max} = [60, 60, 60]$ ,  $\varepsilon = 0.4$ ,  $\mathbf{x}_t(0) = \mathbf{1}$ ,  $\mathbf{Q} = \text{diag}\{3200, 3200, 3200\}$  and  $\mathbf{R} = \text{diag}\{1, 1, 1\}$ . The minimum stiffness is chosen in order to have high compliance but ensuring an acceptable position tracking performance, while the maximum stiffness is selected in order to track the desired motion with high accuracy while ensuring a stable behavior. The maximum force corresponds to the maximum payload of the robot, and the initial energy of the tank is slightly higher than the energy threshold ( $\varepsilon$ ). Finally,  $\mathbf{Q}$  and  $\mathbf{R}$  are chosen experimentally to give higher priority to the force tracking.

## B. Experimental Results

The results of the autonomous task repetition without external disturbances for LS, HS, and OS are shown in Fig. 5. For all the experiments, force tracking and position tracking performance are reported, while the stiffness values and energy in the tank are shown only for OS. Although LS allows a compliant behavior, it is not able to track the desired force along  $z$ , which is necessary for the cleaning performance (Fig. 5(b)). Conversely, for HS (Fig. 5(c)) a too large force is applied along  $z$ . This results in the stick and slip phenomenon, as can be viewed from the measured force and position trend along the  $x$  axis. On the other hand, OS (Fig. 5(a)) generates the right amount of compliance needed to track the desired force during the interaction with the table while keeping low stiffness in free motion when there are no interactions. The two sources of energy variations in the tank are the dissipated energy and the exchanged energy due to the variable stiffness (see (15)). The energy in the tank has an increasing trend, and it never falls below the threshold  $\varepsilon$ , meaning that the passivity of the system is preserved, although sometimes energy drops due to stiffness variations.

The results of the autonomous task repetition with external disturbances for LS, HS, and OS are shown in Fig. 6. In Fig. 6(b), the results of the autonomous task repetition with external disturbances for LS are shown. Although the robot is always compliant to environment uncertainties introduced by external disturbances, it cannot perform the task properly since insufficient force is exerted on the table. On the contrary, for HS (Fig. 6(c1)), as soon as the table is lifted, the interaction force exceeds the robot payload limit. The same situation occurs when a human collides along  $x$  as it can be viewed in Fig. 6(c2) on the right. Finally, the results for OS are reported in Fig. 6(a). For all the types of disturbances applied along  $z$ , the robot can optimize the stiffness to track the desired force. The high-frequency disturbance makes the force tracking more challenging, but the performance is still acceptable. Then, the stiffness is kept low along  $x$ , resulting in a compliant behavior when the human collides with the robot along that direction while interacting with the table and during free motion. In addition, the energy in the tank never falls below the threshold, and it has an overall increasing trend. At the same time, it is evident how the external disturbances affected the additional energy injected (e.g., during table lifting) and extracted (e.g., during table lowering) from the tank.



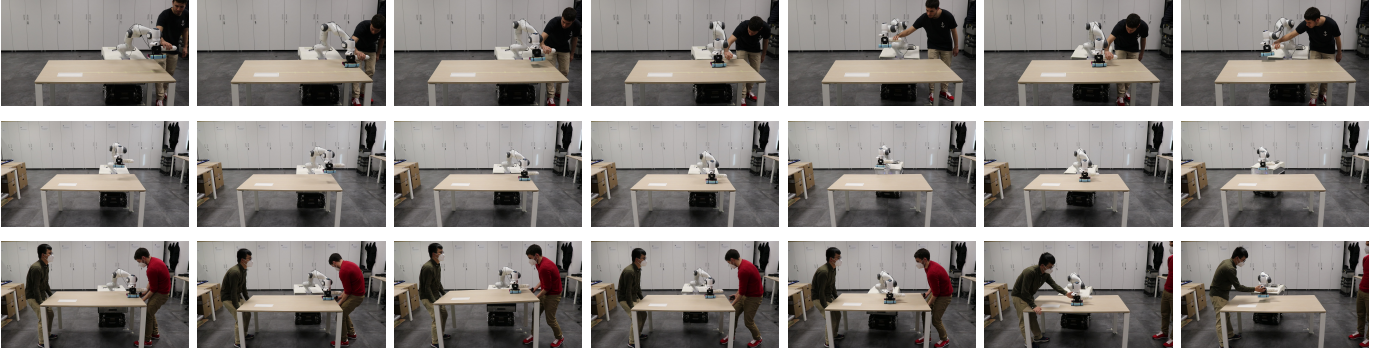


Fig. 4. Snapshots of the experiments. (top) Kinesthetic teaching of MOCA using an admittance-type physical interface. (middle) Experiment 1: autonomous task execution in nominal conditions. (bottom) Experiment 2: autonomous task execution with random disturbances. The desk’s height was changed in different frequencies and scales in the first five photos. The human stopped MOCA’s end-effector during the cleaning and free motion phase in the last two photos, respectively. A video of the experiments is available in the multimedia extension and at the link: <https://youtu.be/Ini4twYy4Rk>.

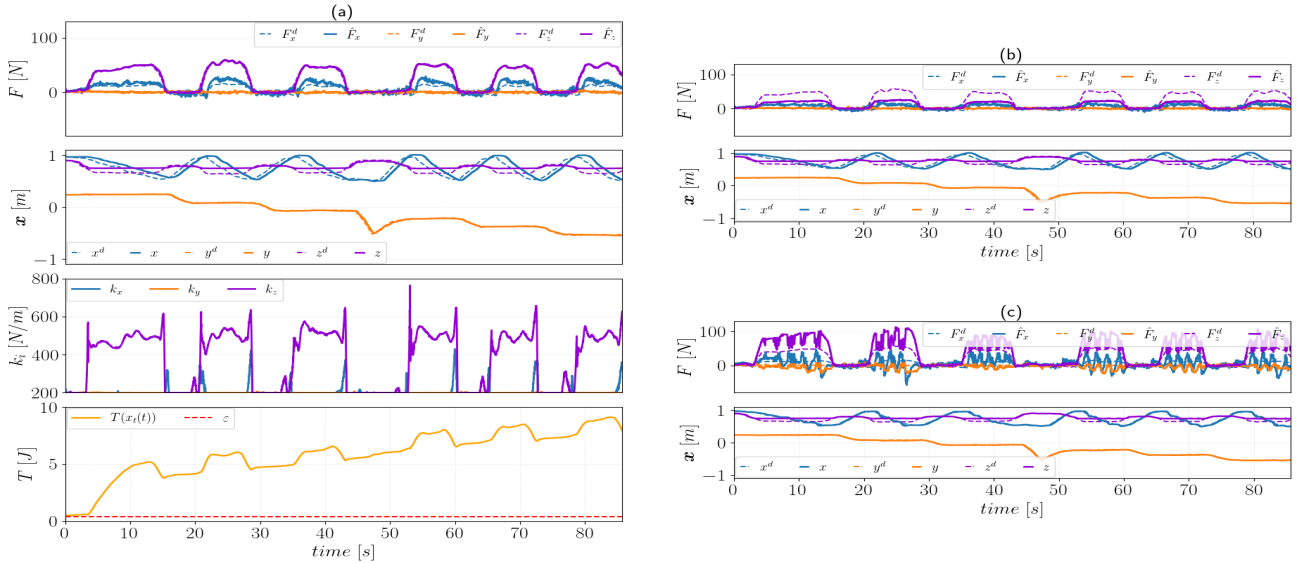


Fig. 5. Results of autonomous task execution experiments in nominal conditions. (a). OS settings, from top to bottom are desired interactive force  $\mathbf{F}^d$  and estimated force  $\hat{\mathbf{F}}^{ext}$  during experiments, desired position  $\mathbf{X}^d$  and real position  $\mathbf{X}$ , generated stiffness  $\mathbf{K}$ , and tank energy  $T(x_t)$ . (b). LS settings, from top to bottom are desired interactive force  $\mathbf{F}^d$  and estimated force  $\hat{\mathbf{F}}^{ext}$  during experiments, desired position  $\mathbf{X}^d$  and real position  $\mathbf{X}$ . (c). HS settings, same contents with (b).

## V. DISCUSSION AND CONCLUSION

The results described in section IV-B show that overall OS outperforms LS and HS. Although LS keeps high compliance and robustness to external perturbations and uncertainties, it cannot achieve satisfactory cleaning performance. Instead, HS failed in performance, safety, and robustness since it renders a rigid behavior at the end-effector, resulting in high interaction forces with the environment since it rejects the external disturbances during motion. On the other hand, OS is capable of exerting the right amount of force for the cleaning task and, at the same time keeping as much compliance as possible when no interaction is required.

The framework developed in this paper allowed to easily teach a mobile manipulator to perform an interactive task. The target tasks consist of complex interaction with the environment and unconstrained motions in a larger workspace w.r.t. the reachable workspace of a fixed-base robotic arm. Moreover, the robot can robustly apply the learned interaction force while maintaining free-motion compliance. If, instead,

the kinesthetic demonstrations were used to learn only desired end-effector desired position and velocity trajectories and a constant stiffness was used as it is done for HS and LS, a trade-off should have been found between ensuring compliance and high task-related performance. Even if achieving an acceptable trade-off was feasible, it would not have been possible to obtain the robustness to external perturbations that our framework accomplishes. In addition, our approach ensures the passivity of the system through an energy tank-based constraint and allows to include payload limits and stiffness bounds as constraints in the optimization problem.

Future work will include the learning procedure of all the relevant parameters of the Cartesian impedance whole-body controller, such as the motion modes and secondary tasks through an admittance-type physical interface. In this work, the task considered allowed to choose a priori some suitable controller parameters that would have ensured the successful loco-manipulation behavior of the robotic platform during the whole task execution, but in more complex environments,



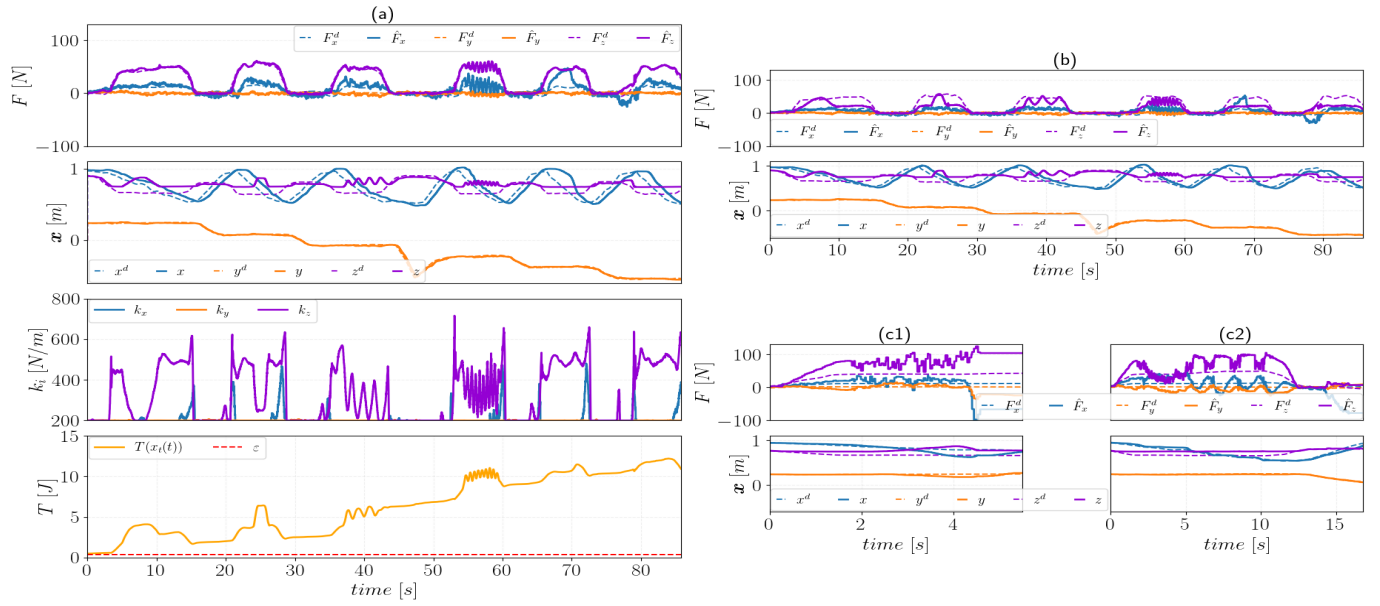


Fig. 6. Results of the autonomous task repetition with external disturbances. (a). OS (b). LS (c) HS. The plotted variables are the same of Fig. 5. Plot (c1) shows only disturbance type 1) since the robot stopped due to violation of safety limits. Plot (c2) shows disturbances type 5) and 6), where the robot again stops due to violation of safety limits. The characterizing parameters of the first four perturbations for OS and LS respectively are (A: Amplitude; V: velocity; f: frequency): 1)  $A=(0.1246, 0.1230)\text{m}$ ,  $V=(0.0417, 0.0464)\text{m/s}$ ; 2)  $A=(0.1163, 0.1443)\text{m}$ ,  $V=(0.1256, 0.1163)\text{m/s}$ ; 3)  $A=(0.0758, 0.0986)\text{m}$ ,  $f=(0.8645, 0.6755)\text{Hz}$ ; and 4)  $A=(0.0435, 0.0428)\text{m}$ ,  $f=(2.0024, 2.2422)\text{Hz}$ .

this is not always the case. Furthermore, the self-tuning of the weighting matrices of the QP formulation ( $Q$  and  $R$ ) represents another appealing future development.

## REFERENCES

- [1] M. A. Roa, D. Berenson, and W. Huang, "Mobile manipulation: toward smart manufacturing [tc spotlight]," *IEEE Robotics & Automation Magazine*, vol. 22, no. 4, pp. 14–15, 2015.
- [2] S. Bogh, C. Schou, T. Rühr, Y. Kogan, A. Dömel, M. Brucker, C. Eberst, R. Tornese, C. Sprunk, G. D. Tipaldi *et al.*, "Integration and assessment of multiple mobile manipulators in a real-world industrial production facility," in *ISR/Robotik 2014; 41st International Symposium on Robotics*. VDE, 2014, pp. 1–8.
- [3] E. Lamon, M. Leonori, W. Kim, and A. Ajoudani, "Towards an intelligent collaborative robotic system for mixed case palletizing," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 9128–9134.
- [4] Y. Wu, E. Lamon, F. Zhao, W. Kim, and A. Ajoudani, "Unified approach for hybrid motion control of moca based on weighted whole-body cartesian impedance formulation," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3505–3512, 2021.
- [5] G. Averta and N. Hogan, "Enhancing robot-environment physical interaction via optimal impedance profiles," in *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechanics (BioRob)*, 2020, pp. 973–980.
- [6] J. Duan, Y. Gan, M. Chen, and X. Dai, "Adaptive variable impedance control for dynamic contact force tracking in uncertain environment," *Robotics and Autonomous Systems*, vol. 102, pp. 54–65, 2018.
- [7] P. Balatti, D. Kanoulas, N. Tsagarakis, and A. Ajoudani, "A method for autonomous robotic manipulation through exploratory interactions with uncertain environments," *Autonomous Robots*, vol. 44, no. 8, pp. 1395–1410, 2020.
- [8] F. Ficuciello, L. Villani, and B. Siciliano, "Variable impedance control of redundant manipulators for intuitive human-robot physical interaction," *IEEE Transactions on Robotics*, vol. 31, no. 4, 2015.
- [9] L. Peternel, T. Petrič, E. Oztop, and J. Babič, "Teaching robots to cooperate with humans in dynamic manipulation tasks based on multimodal human-in-the-loop approach," *Autonomous robots*, vol. 36, no. 1, pp. 123–136, 2014.
- [10] Y. Wu, F. Zhao, T. Tao, and A. Ajoudani, "A Framework for Autonomous Impedance Regulation of Robots Based on Imitation Learning and Optimal Control," *IEEE Robotics and Automation Letters*, vol. 6, no. 1, pp. 127–134, 2020.
- [11] S. Calinon and D. Lee, "Learning control," in *Humanoid Robotics: a Reference*. Springer, 2019, pp. 1261–1312.
- [12] 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.
- [13] F. J. Abu-Dakka and M. Saveriano, "Variable impedance control and learning—a review," *Frontiers in Robotics and AI*, p. 177, 2020.
- [14] S. Calinon, I. Sardellitti, and D. G. Caldwell, "Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, pp. 249–254.
- [15] F. J. Abu-Dakka, L. Rozo, and D. G. Caldwell, "Force-based variable impedance learning for robotic manipulation," *Robotics and Autonomous Systems*, vol. 109, pp. 156–167, 2018.
- [16] Y. Michel, R. Rahal, C. Pacchierotti, P. R. Giordano, and D. Lee, "Bilateral teleoperation with adaptive impedance control for contact tasks," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5429–5436, 2021.
- [17] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 820–833, 2011.
- [18] J. M. Gandarias, P. Balatti, E. Lamon, M. Lorenzini, and A. Ajoudani, "Enhancing flexibility and adaptability in conjoined human-robot industrial tasks with a minimalist physical interface," in *2022 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, p. in press.
- [19] F. Fusaro, E. Lamon, E. D. Momi, and A. Ajoudani, "A human-aware method to plan complex cooperative and autonomous tasks using behavior trees," in *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*, 2021, pp. 522–529.
- [20] A. Dietrich, K. Bussmann, F. Petit, P. Kotyczka, C. Ott, B. Lohmann, and A. Albu-Schäffer, "Whole-body impedance control of wheeled mobile manipulators," *Autonomous Robots*, vol. 40, no. 3, pp. 505–517, 2016.
- [21] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.
- [22] C. Ott, *Cartesian impedance control of redundant and flexible-joint robots*. Springer, 2008.
- [23] 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.