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

Jianzhuang Zhao

***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 Programming augmented with an energy tank to ensure the passivity of the controlled system. 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, in the presence of disturbances such as unexpected end-effector collisions.**

I. INTRODUCTION

Collaborative Mobile Manipulators (CMM), which integrate a collaborative manipulator on a mobile platform, combine the precise 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 in service and domestic applications [1], [2]. Nevertheless, integrating different components makes the control and 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

In addition, when facing unstructured environments, where interactions may change unexpectedly, collaborative robots (cobots) are expected to adapt their performance online to unforeseen situations. To model the system response to interactions with the environment, impedance controllers require the tuning of a high number of design parameters. However, how to optimally tune impedance parameters, namely variable impedance control (VIC), is still an open issue. For instance, authors in [3], [4] 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 [3], 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. Other methods exploit EMG sensors to estimate joint stiffness and generate reference trajectories at the same time [5], [6]. 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 [7]. The IL paradigm includes different algorithms to encode the taught demonstrations, such as Dynamical Movement Primitives (DMPs) [8] and Gaussian Mixture Model (GMM) / Gaussian Mixture Regression (GMR) [7], among the most used. In the context of cobots, one IL methodology, i.e. kinesthetic teaching [7], 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. To the best of the authors knowledge, currently there exist few intuitive interfaces that fully exploit the potential of CMMs in kinesthetic teaching. In the context of fixed-base manipulators, several works faced the problem of simultaneous learning of trajectories and force profiles through impedance models [9] using GMM/GMR. In [10], to generate safe robot behavior, the variable stiffness is tuned according the inverse of the position covariance of the GMM. In [11], two different semi-positive definite stiffness representations are investigated, while stiffness profiles are obtained with a least-squares estimator using a linear interaction model, based on position/force demonstrations. 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 [9].

To address these issues, we propose a method to learn loco-manipulation tasks exploiting the paradigm of VIC. 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 [12], complex interactive

Research Center for Intelligent Robotics, School of Astronautics, Northwestern Polytechnical University, Xi'an, China.

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. 1). The method is evaluated with a table cleaning task. The results show that the impedance parameters' tuning yields good performances in all conditions, allowing at the same time trajectory and force tracking and robust and compliant interactions compared to constant stiffness.

II. 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 (MOCA-MAN). Note that a F/T sensor is integrated to measure the user interaction wrenches. A detailed explanation of the admittance-type physical interface can be found in [12]. The measured human wrenches $\hat{\boldsymbol{\lambda}}_h \in \mathbb{R}^m$, are used to change the input Cartesian equilibrium pose $\boldsymbol{x}_d$, as input of an admittance controller which implements the following desired dynamics:

$$\boldsymbol{M}^{adm}\ddot{\boldsymbol{x}}_d + \boldsymbol{D}^{adm}\dot{\boldsymbol{x}}_d = \hat{\boldsymbol{\lambda}}_h, \tag{1}$$

where $\boldsymbol{M}^{adm}, \boldsymbol{D}^{adm} \in \mathbb{R}^{m\times m}$ 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 (1) and of the whole-body controller.

Desired Cartesian equilibrium pose $\boldsymbol{x}_d$ and twist $\dot{\boldsymbol{x}}_d$ are obtained as the solution of (1). Since the robotic arm used is torque controlled, the interaction wrenches of the robot end-effector with the environment $\hat{\boldsymbol{F}}^{ext}$ 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 $\boldsymbol{\eta}^{\mathcal{I}} = t$ and output variables are $\boldsymbol{\eta}^{\mathcal{O}} = \begin{bmatrix}\boldsymbol{x}_d^T & \dot{\boldsymbol{x}}_d^T & \hat{\boldsymbol{F}}^{ext^T}\end{bmatrix}^T$. 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. The QP is formulated as follows:

$$\begin{aligned}\min_{\substack{\boldsymbol{K}_i^d\in\mathbb{R}^{m\times m}\\ i\in\{1,\dots,N\}}} \quad & \frac{1}{2}\sum_{i=1}^{N}\left(\|\boldsymbol{F}_i^{ext}-\boldsymbol{F}_i^d\|_{\boldsymbol{Q}}^2 + \|\boldsymbol{K}_i^d-\boldsymbol{K}^{min}\|_{\boldsymbol{R}}^2\right)\\ \text{s.t.} \quad & \boldsymbol{K}^{min} \le \boldsymbol{K}_i^d \le \boldsymbol{K}^{max} \quad i\in\{1,\dots,N\}\\ & -\boldsymbol{F}^{max} \le \boldsymbol{F}_i^{ext} \le \boldsymbol{F}^{max} \quad i\in\{1,\dots,N\}\end{aligned} \tag{2}$$

where i is the time step, N is the length of the time window, $\boldsymbol{Q}$ and $\boldsymbol{R} \in \mathbb{R}^{m\times m}$ are diagonal positive definite weighting matrices, $\boldsymbol{K}_i^d \in \mathbb{R}^{m\times m}$ is the desired stiffness of the Cartesian whole body impedance controller at time step i, $\boldsymbol{K}^{min}$ and $\boldsymbol{K}^{max} \in \mathbb{R}^{m\times m}$ are respectively minimum and maximum allowed stiffness, $\boldsymbol{F}_i^{ext} \in \mathbb{R}^m$ is the wrench of the impedance interaction model at time step i, that can be modeled with impedance-like laws, $\boldsymbol{F}_i^d \in \mathbb{R}^m$ is the learned desired interaction wrench at time step i (i.e., output of the GMR) and $\boldsymbol{F}^{max} \in \mathbb{R}^m$ 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, $\boldsymbol{F}^{ext}$ is expressed as:

$$\boldsymbol{F}^{ext} = \left(\boldsymbol{\mu}(\boldsymbol{x},\dot{\boldsymbol{x}}) + \boldsymbol{D}^d\right)\dot{\tilde{\boldsymbol{x}}} + \boldsymbol{K}^d\tilde{\boldsymbol{x}}. \tag{3}$$

where $\tilde{\boldsymbol{x}} = \boldsymbol{x}_d - \boldsymbol{x} \in \mathbb{R}^6$ is the Cartesian pose error, and $\boldsymbol{D}^d$, $\boldsymbol{K}^d \in \mathbb{R}^{m\times m}$ are the desired Cartesian damping and stiffness matrices, respectively. $\boldsymbol{\mu}(\boldsymbol{x},\dot{\boldsymbol{x}})$ is the Coriolis and centrifugal terms.

C. Tank Energy Based Passivity Constraint

The QP formulated in section II-B computes a Cartesian stiffness at each time step that is sent to the Cartesian impedance whole-body controller. It is well known that VIC might violate the passivity of the system [13]; 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 $\dot{\boldsymbol{x}}\boldsymbol{F}^{ext}$. The interaction model of the variable Cartesian impedance can be written as a port-Hamiltonian system and augmented through an energy tank. The scalar differential equation that describes the tank dynamics is:

$$\dot{\mathrm{x}}_t = \frac{\sigma}{\mathrm{x}_t}\dot{\tilde{\boldsymbol{x}}}^T\boldsymbol{D}^d\dot{\tilde{\boldsymbol{x}}} - \frac{\boldsymbol{w}^T}{\mathrm{x}_t}\dot{\tilde{\boldsymbol{x}}}, \tag{4}$$

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

$$\boldsymbol{w}(t) = \begin{cases} -\boldsymbol{K}^v(t)\tilde{\boldsymbol{x}} & \text{if } T(\mathrm{x}_t) > \varepsilon \\ 0 & \text{otherwise,}\end{cases} \tag{5}$$

where $\boldsymbol{K}^v(t)$ is the variable part of the stiffness so that $\boldsymbol{K}^d(t) = \boldsymbol{K}^{min} + \boldsymbol{K}^v(t)$ and $\varepsilon > 0$ is the minimum energy

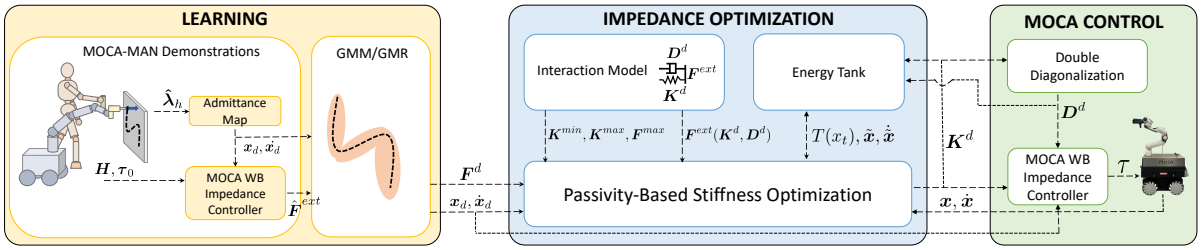

Fig. 1. The proposed framework scheme. Humans can teach MOCA the interactive tasks directly through MOCA-MAN interface, and the desired trajectories and interaction force are replicated with GMM/GMR. Then, the desired force is sent to a QP-based algorithm to optimize online stiffness. A tank energy constraint ensures the passivity of the system. Finally, the desired trajectory and stiffness are sent to MOCA's whole-body impedance controller.

that the tank is allowed to store. The tank energy is initialized so that $T(\mathrm{x}_t(0)) > \varepsilon$. 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 ε:

$$T(\mathrm{x}_t(t)) = T(\mathrm{x}_t(t-1)) + \dot{T}(\mathrm{x}_t(t))\Delta t > \varepsilon, \tag{6}$$

where Δt is the time step and

$$\dot{T}(\mathrm{x}_t) = \mathrm{x}_t\dot{\mathrm{x}}_t = \sigma\dot{\tilde{\boldsymbol{x}}}^T\boldsymbol{D}^d\dot{\tilde{\boldsymbol{x}}} - \boldsymbol{w}^T\dot{\tilde{\boldsymbol{x}}} \tag{7}$$

$$= \begin{cases} \sigma\dot{\tilde{\boldsymbol{x}}}^T\boldsymbol{D}^d\dot{\tilde{\boldsymbol{x}}} + \tilde{\boldsymbol{x}}^T\boldsymbol{K}^v\dot{\tilde{\boldsymbol{x}}} & \text{if } T(\mathrm{x}_t) > \varepsilon \\ \sigma\dot{\tilde{\boldsymbol{x}}}^T\boldsymbol{D}^d\dot{\tilde{\boldsymbol{x}}} & \text{otherwise.} \end{cases} \tag{8}$$

Note from (8) 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 ($\boldsymbol{K}^d = \boldsymbol{K}^{min}$).

III. 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 with disturbances. Three stiffness settings are tested: i) low constant stiffness (LS, $\boldsymbol{k}^d = \boldsymbol{k}^{min}$), ii) high constant stiffness (HS, where $\boldsymbol{k}^d = \boldsymbol{k}^{max}$) and iii) optimized stiffness (OS), where $\boldsymbol{k}^d$ is found online through our QP formulation. 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 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 autonomous repetition experiment (see also Figure 2 (bottom)): 1), Single slow table lifting and lowering during the 1st cleaning movement; 2), Single fast table lifting and lowering during the 2nd cleaning movement; 3), Repeated low frequency and high amplitude table lifting and lowering during the 3rd cleaning movement; 4), Repeated high frequency and low amplitude table lifting and lowering during the 4th cleaning movement. 5), Collision with a human during the 5th cleaning motion; 6), Collision with a human during the last free motion. The values of the parameters used during the autonomous repetition experiments are $\boldsymbol{k}^{min}$=$[200, 200, 200]$, $\boldsymbol{k}^{max}$=$[1000, 1000, 1000]$, $\boldsymbol{F}^{max}$=$[60, 60, 60]$, ε=0.4, $x_t(0)$=1, $\boldsymbol{Q}$=$diag\{3200, 3200, 3200\}$ and $\boldsymbol{R}$=$diag\{1, 1, 1\}$.

B. Experimental Results

The results of the autonomous task repetition with external disturbances for LS, HS, and OS are shown in Fig. 3. In Fig. 3(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. 3(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. 3(c2) on the right. Finally, the results for OS are reported in Fig. 3(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.

IV. DISCUSSION AND CONCLUSION

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,

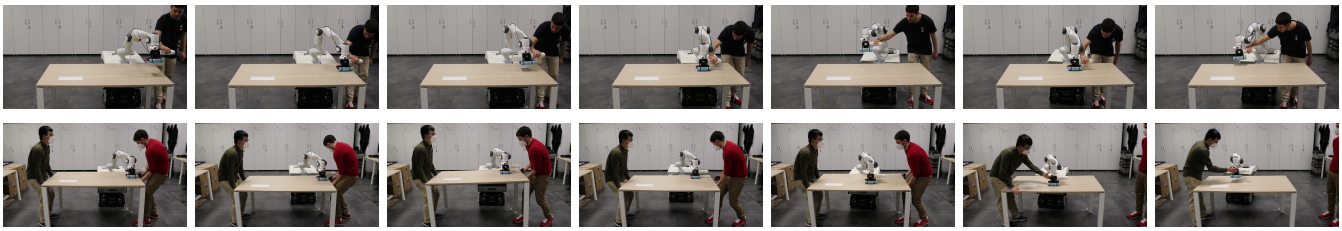

Fig. 2. Snapshots of the experiments. (top) Kinesthetic teaching of MOCA using the MOCA-MAN interface. (bottom) Autonomous task execution with random disturbances. The desk's height was changed in different frequencies and scales in the first five photos. The human stopped MOCA's end-effector during 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/wotzfcpcAU0`.

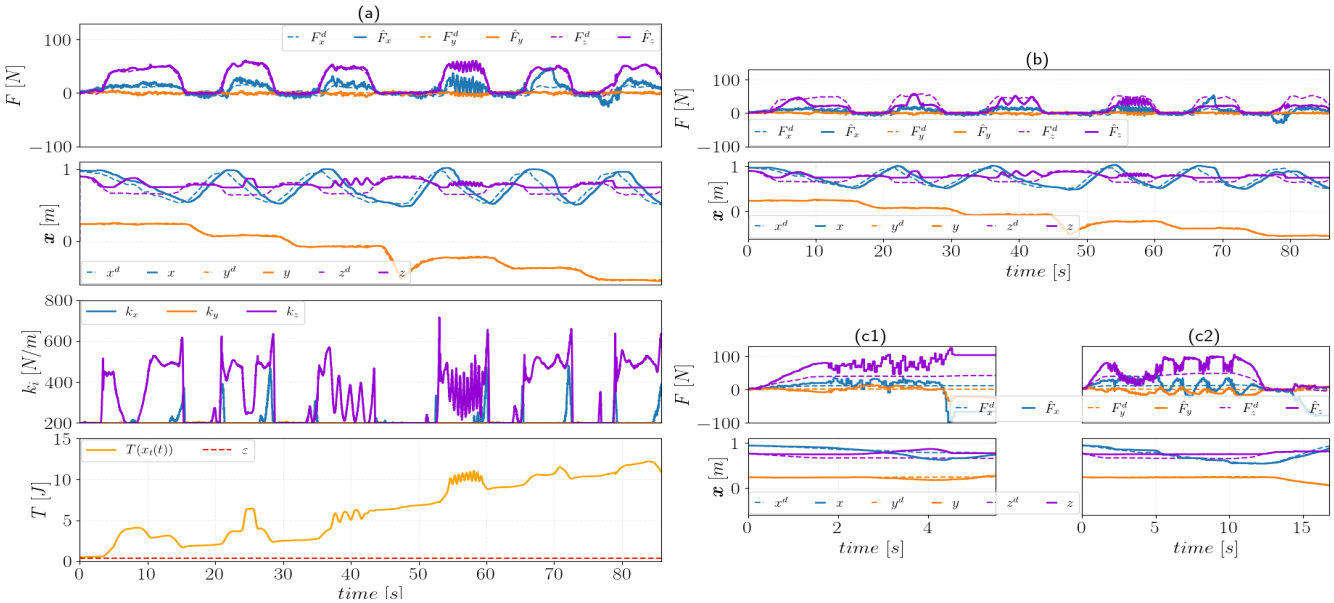

Fig. 3. Results of the autonomous task execution experiments with unplanned disturbances. (a). OS settings, from top to bottom are desired interactive force $\boldsymbol{F}^d$ and estimated force $\hat{\boldsymbol{F}}^{ext}$ during experiments, desired position $\boldsymbol{X}^d$ and real position $\boldsymbol{X}$, generated stiffness $\boldsymbol{K}$, and tank energy $T(\mathrm{x}_t)$. (b). LS settings,from top to bottom are desired interactive force $\boldsymbol{F}^d$ and estimated force $\hat{\boldsymbol{F}}^{ext}$ during experiments, desired position $\boldsymbol{X}^d$ and real position $\boldsymbol{X}$. (c). HS settings, the figure that shows disturbances 1) (in (c1)) and 5), 6) (in (c2)).

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 the self-tuning of the weighting matrices of the QP formulation ($\boldsymbol{Q}$ and $\boldsymbol{R}$).

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] 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 Biomechatronics (BioRob)*, 2020, pp. 973–980.

[4] 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.

[5] L. Peternel, T. Petrič, E. Oztop, and J. Babič, "Teaching robots to cooperate with humans in dynamic manipulation tasks based on multi-modal human-in-the-loop approach," *Autonomous robots*, vol. 36, no. 1, pp. 123–136, 2014.

[6] 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.

[7] S. Calinon and D. Lee, "Learning control," in *Humanoid Robotics: a Reference*. Springer, 2019, pp. 1261–1312.

[8] 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.

[9] F. J. Abu-Dakka and M. Saveriano, "Variable impedance control and learning—a review," *Frontiers in Robotics and AI*, p. 177, 2020.

[10] 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.

[11] 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.

[12] 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.

[13] 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.