

# A Framework for Autonomous Impedance Regulation of Robots Based on Imitation Learning and Optimal Control

Yuqiang Wu, Fei Zhao, Tao Tao, and Arash Ajoudani

**Abstract**—In this work, we propose a framework to address the autonomous impedance regulation problem of robots in a class of constrained manipulation tasks. In this framework, a human arm endpoint stiffness model is used to extract the task stiffness geometry along the constrained trajectory, which is then encoded offline and reproduced online by a Gaussian Mixture Model (GMM) and the Gaussian Mixture Regression (GMR), respectively. Furthermore, the full Cartesian impedance of the robot is formulated through an optimal control problem, i.e., the Linear-Quadratic Regulator (LQR), in which the task stiffness geometry (extracted from human demonstrations) is considered as the time-varying weighting matrix  $Q$ . The optimal impedance is eventually realised by the robot through a task geometry consistent Cartesian impedance controller. A tank-based passivity observer is implemented to give evidence on the stability of the system during online impedance variations. To evaluate the performance of the framework, a comparative experiment with three different impedance settings (i.e., the proposed framework, the framework without LQR and the framework without GMM/GMR) for Franka Emika Panda to perform a door opening task was conducted. The results reveal that our framework outperforms the other two, in terms of tracking error and the interaction forces.

## I. INTRODUCTION

In the last decades, Cartesian impedance control techniques have been developed for the new generation of torque controlled light-weight robots to guarantee safe interaction with humans and their surroundings [1]. In such control techniques, the impedance parameters regulate the relationship between the trajectory tracking errors and the interaction forces, playing a crucial role in the interaction performance of the robots for a given task. Nevertheless, for tasks that require continuously varying interaction patterns, the adjustment of the impedance parameters remains a difficult issue, also considering the dynamic uncertainties [2], [3].

Inspired by the human manipulation capability, researchers have been trying to understand how human beings regulate their arm impedance in performing interactive tasks. In this direction, authors in [4] proposed a 2D stiffness model of the human arm by assuming that the arm endpoint stiffness is linearly related to the magnitude of the joint torques. Relying on this assumption, an arm control model including

This work has been carried out in HRI<sup>2</sup> Lab, Istituto Italiano di Tecnologia, Genoa, Italy. This work was supported in part by the ERC starting grant Ergo-Lean (Grant Agreement No.850932), in part by the National Natural Science Foundation of China under Grant 91748208, and in part by the Department of Science and Technology of Shaanxi Province under Grant 2017ZDL-G-3-1. In this work, Y. Wu was also supported by the China Scholarship Council.

Yuqiang Wu, Fei Zhao and Tao Tao are with the State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University, Shaanxi Key Laboratory of Intelligent Robots and School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an Shaanxi 710049, China; Yuqiang Wu is also with the HRI<sup>2</sup> Lab of IIT; Arash Ajoudani is with the HRI<sup>2</sup> Lab of IIT. wuyuqiang@stu.xjtu.edu.cn

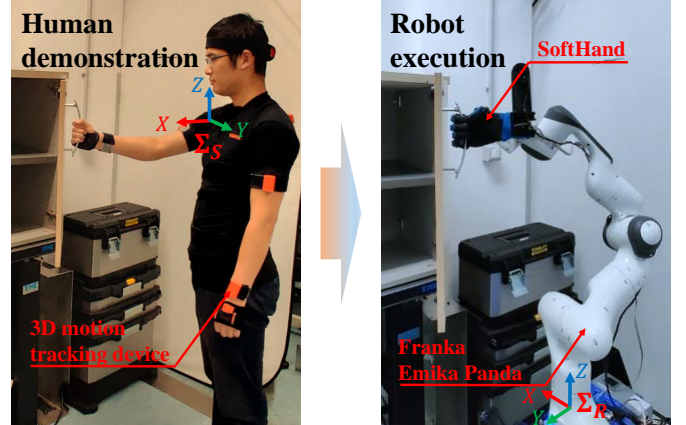


Fig. 1: The task stiffness geometry is extracted from human observations (left), and encoded and reproduced by the robot (right) using the proposed autonomous control framework.

the adaptation of force and impedance was proposed in [5], leading to a similar adaptation behavior observed in human experiments. The online human arm impedance estimation techniques in 3D was first proposed under the concept of teleimpedance control [6], and then extended to the entire arm workspace through a reduced-complexity representation in [7]. An extension of this work to enable transferring of whole-body human motor behaviours to teleoperated robots was presented in [8], where a MOBILE Collaborative robotic Assistant (MOCA) was operated to execute complex loco-manipulation tasks. Although teleimpedance control has shown a high potential in the execution of remote interactive tasks, in most applications, an autonomous approach is preferred.

To achieve autonomous interaction control of robots, the stiffness profile can be encoded offline and reproduced online. An elegant and efficient way to implement this is through imitation learning. In this direction, Dynamic Movement Primitives (DMPs) [9], which consists of a canonical system, a transformation system, and a nonlinear function estimator is widely used to encode complex behaviors in human motor control and robotics. However, due to the dependency of the canonical system on time or phase input in this framework, it can hardly deal with multi-dimensional input data. Another popular movement representation technique is through Gaussian Mixture Regression (GMR) [10]. Firstly, it models the joint probability density of the input and output data in the form of a Gaussian Mixture Model (GMM), which can be estimated by an Expectation-Maximization (EM) procedure. Furthermore, the regression function is derived by computing the conditional probability distribution of GMM. It directly maps the input to output and both can be multi-dimensional. In addition, GMR is a probabilistic model, which can not

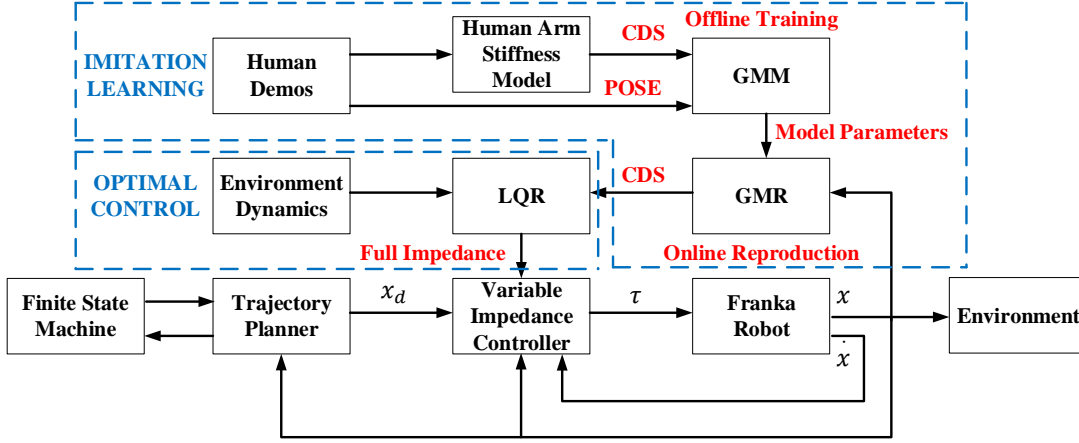


Fig. 2: The architecture of the proposed framework. The full impedance is regulated by an imitation learning algorithm (GMM/GMR) and an optimal control method (LQR), which is then realised by the robot Cartesian impedance controller.

only predict the mean value but also covariance of the output variables. In [11], the robot's movement was encoded and reproduced by GMM/GMR. They consider the predicted covariance as uncertainty in the environment and then use its inverse as the stiffness in Cartesian impedance control of robots. The same idea but with a minimal intervention control framework was established in [12]. However, the inverse of the covariance embedded in multiple movement demonstrations doesn't correspond to the physical meaning of stiffness, and this framework will not work in the constrained motion scenarios, such as door opening and valve turning.

Instead of extracting stiffness information from variability in multiple demonstrations, there are some works about estimating it based on interaction forces and tracking errors, followed by imitation learning algorithms to reproduce the learned stiffness profiles online. The authors in [13] proposed an approach to collaboratively assemble a wooden table by a human subject and a robot. By taking into account the sensed force information, a Weighted Least-Squares (WLS) method is employed to estimate the stiffness. In [14], the authors used a similar framework but adopted a sliding window technique to carry out local stiffness estimation instead of WLS. However, in this way, the position and force data are usually demonstrated via kinesthetic teaching, which may cause discontinuity in a continuous interaction process. It is not very reliable also due to the fact that the first estimation result of stiffness usually does not comply with Symmetric Positive-Definite (SPD) constraints. A more direct way to teach robot compliant manipulation is through human-robot interaction proposed in [3]. However, for a task requiring continuously varying impedance, this way may bring inconvenience to the teacher.

Some works consider impedance regulation as an optimal control problem [15], [16]. In this field, most of the work are based on Linear Quadratic Regulator (LQR) formulation, however, the choices of  $Q$  and  $R$  matrices play an important role which are usually hard to design for individual tasks. In most cases,  $Q$  and  $R$  are fixed which results in a hard-coded behavior. To achieve an online adaptation, adaptive optimal control or Reinforcement Learning (RL) can be employed. An example of this approach is presented in [2], where a Policy

Improvement with Path Integrals (PI<sup>2</sup>) algorithm is proposed to accomplish variable impedance control for a given task. In [17], an approach called OutPut FeedBack Adaptive Dynamic Programming (OPFB ADP) is proposed to solve the problem by only using measured output data, considering that in most robot-environment interaction applications the dynamic processes are partially observable. A detailed discussion of the progress in RL for variable impedance control is beyond the scope of this article.

In this work, we propose a framework that merges the advantages of imitation learning and optimal control to achieve an adaptive and optimal interaction performance with a generalization potential. Our aim is to transfer task stiffness geometry, provided by human demonstrations, to robots in an autonomous and optimal way. The door opening has been chosen as a representative example of tasks with constrained and continuously interactive motion as shown in Fig. 1. Due to a potential instability of the system acting under variable impedance, we developed a tank-based passivity observer to monitor the stability of the system. In addition, to verify the synergistic contribution of the imitation learning and optimal control in this framework, two additional impedance settings were tested: one used GMR to reproduce the varying task stiffness geometry but with a constant volume coefficient instead of LQR, the other one was the opposite, i.e., a constant task stiffness geometry with LQR. The performances of the controllers under different impedance settings were compared in terms of stiffness profiles, tracking error, and interaction forces.

## II. FRAMEWORK

The overall framework is depicted in Fig. 2. Inspired by the imitation learning approach, we attempt to learn the task stiffness geometry from human demonstrations. In contrast with the work in [12]–[14], we propose to utilize physically meaningful stiffness information, which is extracted directly from a human arm endpoint stiffness model. In this model, the arm endpoint stiffness is decomposed into two components, i.e., the Configuration Dependent Stiffness (CDS) and the Common Mode Stiffness (CMS), as proposed in [18], [19].

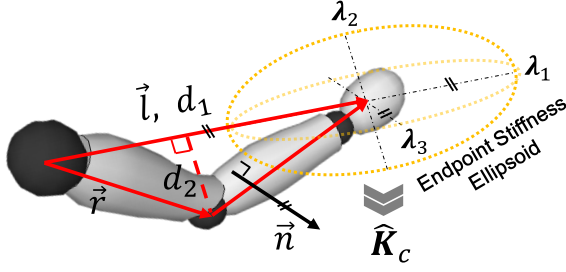


Fig. 3: The geometric relationship between the human arm configuration and the principal axes of the endpoint stiffness ellipsoid.  $\vec{n}$  is perpendicular to the plane formed by  $\vec{l}$  and  $\vec{r}$ .  $\vec{l}$  and  $\vec{n}$  are parallel to the major and minor principal axes of the endpoint stiffness ellipsoid, respectively.

Under the assumption that human beings adapt their arm endpoint stiffness ellipsoid's geometry (represented by CDS) and volume (represented by CMS) to match the interaction requirements of a task, we propose to only encode human demonstrated CDS by GMM offline (pose as input while CDS as output in the model) and reproduce it online by GMR. The CMS component, which regulates the volume of the endpoint stiffness ellipsoid, is not observed from human demonstrations since it can depend on task specifics. For instance, in door opening or valve turning tasks, the CMS would depend on the friction properties or the elastic response of a spring-loaded door or a valve. Instead, to construct the complete impedance (i.e. the full stiffness and the damping term), the LQR is adopted to take CDS as the weighting matrix  $Q$  in an optimal control problem to match different environment dynamics. Eventually, the optimal impedance profile is implemented by a task geometry consistent Cartesian variable impedance controller, whose desired pose is provided by a trajectory planner and a Finite State Machine (FSM). In the following sections, each part of the overall framework will be explained in more detail.

#### A. Learning CDS based on GMM/GMR

Imitation learning has been widely used in modeling of the task kinematics, while there are very few works about learning the dynamics. The barrier is the difficulty to extract physically meaningful dynamic parameters from human demonstrations. To exploit the potential of imitation learning in the execution of physical interaction tasks in this work, a human arm endpoint stiffness model is first introduced. Next, the GMM/GMR is applied to encode and reproduce the task stiffness geometry through the CDS.

1) *Human arm stiffness model and demonstrations:* In [19], we propose a principally simplified, and intuitive online model of the human arm endpoint stiffness. The new model is based on the large dependency of the shape and orientation of the stiffness ellipsoid on the arm configuration. In fact, previous studies [20], [21] indicate that i) the major principal axis of the arm endpoint stiffness ellipse passes close to the hand-shoulder vector, and ii) when the arm is extended and the hand moves further from the shoulder, the ellipse becomes more elongated and conversely, it becomes more isotropic. Inspired by these findings, the new model in 3D is constructed

by the ellipsoid's principal axes and their lengths, namely the eigenvectors and eigenvalues of the stiffness matrix, based on the arm configuration.

As shown in Fig. 3, we use a two-segment human arm skeleton structure in the 3D space. The hand-forearm and the upper arm segments in this model permit us to form an arm triangle at any non-singular configuration. Relying on the dominant contribution of the arm configuration to the endpoint stiffness geometry, we propose to use the vector from the centre of shoulder joint to the position of the hand ( $\vec{l} \in \mathbb{R}^3$ ), to identify the major principal direction of the human arm endpoint stiffness ellipsoid. The minor principal axis direction ( $\vec{n} \in \mathbb{R}^3$ ), instead, is defined to be perpendicular to the arm triangle plane

$$\vec{n} = \vec{r} \times \vec{l}, \quad (1)$$

where  $\vec{r} \in \mathbb{R}^3$  represents the vector from the centre of shoulder to the centre of elbow. The remaining principal axis of the stiffness ellipsoid, which lies on the arm triangle plane, is calculated based on the orthogonality of the three principal axes. Under the assumption that the ratio of the length of median principal axis to the major principal axis of the stiffness ellipsoid is inversely proportional to the distance  $d_1 \in \mathbb{R}$ , from the hand position to the centre of shoulder, while the ratio of the length of the minor principal axis to the major principal axis is proportional to the distance  $d_2 \in \mathbb{R}$ , from the centre of the elbow to the major principal axis, the estimated endpoint stiffness matrix  $\hat{K}_c \in \mathbb{R}^{3 \times 3}$  is formulated by

$$\hat{K}_c = V A_{cc} D_s V^T. \quad (2)$$

Here,  $A_{cc} \in \mathbb{R}$  is the co-contraction activation index of human arm muscles corresponding to CMS, while  $V D_s V^T \in \mathbb{R}^{3 \times 3}$  corresponding to CDS, and  $V \in \mathbb{R}^{3 \times 3}$  and  $D_s \in \mathbb{R}^{3 \times 3}$  are respectively eigenvectors representing the orientation and eigenvalues representing the shape of the stiffness ellipsoid and formulated by:

$$V = \left[ \frac{\vec{l}}{\|\vec{l}\|}, \frac{(\vec{r} \times \vec{l}) \times \vec{l}}{\|(\vec{r} \times \vec{l}) \times \vec{l}\|}, \frac{\vec{r} \times \vec{l}}{\|\vec{r} \times \vec{l}\|} \right], \quad (3)$$

$$D_s = \text{Diag}(1, \alpha_1/d_1, \alpha_2/d_2), \quad (4)$$

where  $\alpha_1 \in \mathbb{R}$  and  $\alpha_2 \in \mathbb{R}$  are subject related parameters in the model. For the detailed explanation and further model parameter identification, please refer to [19].

According to the introduced model, to extract CDS information from human demonstrations for a given task, only positional information of the shoulder, elbow and hand is necessary which is quite easy to get. The CDS information can be extracted followed by a post-processing of the collected 3D position data according to the proposed model.

2) *Encoding CDS profile by GMM/GMR:* In this part, we first give a recap of GMM/GMR algorithm and then introduce how to apply it to encode the CDS profiles. The Gaussian mixture distribution can be written as a linear superposition of Gaussians in the form

$$p(\xi) = \sum_{k=1}^K \pi_k \mathcal{N}(\xi | \mu_k, \Sigma_k), \quad (5)$$

where  $\xi \in \mathbb{R}^d$  and  $p(\xi) \in \mathbb{R}$ , respectively, represent the vector of variables and joint probability distribution.  $\pi_k \in \mathbb{R}$ ,  $\mu_k \in \mathbb{R}^d$  and  $\Sigma_k \in \mathbb{R}^{d \times d}$  represent the prior probability, mean and covariance of the  $k$ -th Gaussian component respectively, while  $K \in \mathbb{R}$  is the total number of Gaussian components.

The goal of the offline training phase is to maximize the log likelihood function Eq. (6) with respect to the model parameters ( $\pi_k$ ,  $\mu_k$  and  $\Sigma_k$ ), which results in an EM procedure (E-step in (7) and M-step in (8) and (9)) to iteratively update the model parameters until convergence

$$\ln p([\xi_1, \xi_2, \dots, \xi_N]) = \sum_{n=1}^N \ln \left\{ \sum_{k=1}^K \pi_k \mathcal{N}(\xi_n | \mu_k, \Sigma_k) \right\}, \quad (6)$$

$$\gamma_{n,k} = \frac{\pi_k \mathcal{N}(\xi_n | \mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\xi_n | \mu_j, \Sigma_j)}, \quad (7)$$

$$\pi_k^{new} = \frac{\sum_{n=1}^N \gamma_{n,k}}{N}, \quad \mu_k^{new} = \frac{\sum_{n=1}^N \gamma_{n,k} \xi_n}{\sum_{n=1}^N \gamma_{n,k}}, \quad (8)$$

$$\Sigma_k^{new} = \frac{\sum_{n=1}^N \gamma_{n,k} (\xi_n - \mu_k^{new})(\xi_n - \mu_k^{new})^T}{\sum_{n=1}^N \gamma_{n,k}}. \quad (9)$$

Here,  $N \in \mathbb{R}$  is the total number of data points in a training dataset and  $\xi_n \in \mathbb{R}^d$  is the  $n$ -th training data point.

After modeling the joint probability distribution of the training data offline, to derive GMR, we use superscripts  $\mathcal{I}$  and  $\mathcal{O}$  to denote the dimensions of the input and output variables. At iteration  $n$ ,  $\eta_n^{\mathcal{I}}$  and  $\eta_n^{\mathcal{O}}$  respectively represent the input and output variables. With this notation, the data point  $\eta_n \in \mathbb{R}^d$ , mean  $\mu_k$  and covariance  $\Sigma_k$  of the  $k$ -th Gaussian component can be decomposed as

$$\eta_n = \begin{bmatrix} \eta_n^{\mathcal{I}} \\ \eta_n^{\mathcal{O}} \end{bmatrix}, \mu_k = \begin{bmatrix} \mu_k^{\mathcal{I}} \\ \mu_k^{\mathcal{O}} \end{bmatrix}, \Sigma_k = \begin{bmatrix} \Sigma_k^{\mathcal{II}} & \Sigma_k^{\mathcal{IO}} \\ \Sigma_k^{\mathcal{OI}} & \Sigma_k^{\mathcal{OO}} \end{bmatrix}. \quad (10)$$

In the online reproduction phase, the best estimation of output  $\hat{\eta}_n^{\mathcal{O}}$  for a given input  $\eta_n^{\mathcal{I}}$  is the mean  $\hat{\mu}_n$  of the conditional probability distribution  $\hat{\eta}_n^{\mathcal{O}} | \eta_n^{\mathcal{I}} \sim \mathcal{N}(\hat{\mu}_n, \hat{\Sigma}_n)$ , with parameters which was reported in [22]:

$$\hat{\mu}_n = \mathbb{E}(\hat{\eta}_n^{\mathcal{O}} | \eta_n^{\mathcal{I}}) = \sum_{k=1}^K h_k(\eta_n^{\mathcal{I}}) \mu_k^{\mathcal{O}}(\eta_n^{\mathcal{I}}), \quad (11)$$

where

$$h_k(\eta_n^{\mathcal{I}}) = \frac{\pi_k \mathcal{N}(\eta_n^{\mathcal{I}} | \mu_k^{\mathcal{I}}, \Sigma_k^{\mathcal{II}})}{\sum_{j=1}^K \pi_j \mathcal{N}(\eta_n^{\mathcal{I}} | \mu_j^{\mathcal{I}}, \Sigma_j^{\mathcal{II}})}, \quad (12)$$

$$\mu_k^{\mathcal{O}}(\eta_n^{\mathcal{I}}) = \mu_k^{\mathcal{O}} + \Sigma_k^{\mathcal{OI}} (\Sigma_k^{\mathcal{II}})^{-1} (\eta_n^{\mathcal{I}} - \mu_k^{\mathcal{I}}). \quad (13)$$

In this framework, only the mean value is used although covariance can be derived as well. human arm endpoint pose and CDS are respectively considered as input and output of GMM. Pose data can be represented in vector form while the CDS profile is a series of SPD matrices, it's not possible to apply GMM to encode it directly considering GMM can only deal with independent variables in vector form. In [14], two different ways to represent stiffness matrix are provided. To make it simple, we use Cholesky decomposition to decompose

CDS into the product of a lower triangular matrix and its transpose, and then we arrange its elements in a vector form. Next, the offline training phase takes the compact pose and the CDS vectors as the vector of variables  $\xi$ . Subsequently, the GMR reproduces the CDS vector online by taking robot end-effector pose as input, followed by a reverse process of Cholesky decomposition to construct CDS matrix whose SPD property is guaranteed.

### B. Constructing full impedance through optimal control

For the control of robots, we need to derive the full stiffness, and eventually the full impedance from CDS information. To take environment dynamics into account and construct the full impedance from an optimal point of view, [12] proposed to formulate it as a LQR problem in which the predicted covariance extracted from multiple movement demonstrations was adopted as the weighting matrix  $Q$ . This part follows the same idea with the difference that instead of using the predicted covariance, we are using the physically meaningful CDS formulated by the human arm endpoint stiffness model. In the following, a brief description of LQR problem is given. Consider the discrete state space equation of a linear environment system as:

$$\zeta_{n+1} = A\zeta_n + Bu_n, \quad (14)$$

where  $\zeta_n = [x_n^T, \dot{x}_n^T]^T \in \mathbb{R}^{2 \times p}$  is the system state,  $x_n \in \mathbb{R}^p$  and  $\dot{x}_n \in \mathbb{R}^p$  respectively represent the pose and velocity.  $u_n \in \mathbb{R}^r$  is the system input, and  $A \in \mathbb{R}^{(2 \times p) \times (2 \times p)}$  and  $B \in \mathbb{R}^{(2 \times p) \times r}$  are the discrete state and input matrices respectively. For the subsequent form of control input:

$$u_n = \hat{K}_n(\hat{x}_n - x_n) - \hat{D}_n \dot{x}_n, \quad (15)$$

where  $\hat{x}_n \in \mathbb{R}^p$  is the desired pose,  $\hat{K}_n \in \mathbb{R}^{r \times p}$  and  $\hat{D}_n \in \mathbb{R}^{r \times p}$  are complete stiffness and damping matrices estimated by LQR with time-varying CDS. For an infinite-horizon, discrete-time LQR, the solution is obtained by minimizing the following cost function:

$$J = \sum_{n=0}^{\infty} (\hat{x}_n - x_n)^T Q_n (\hat{x}_n - x_n) + u_n^T R u_n, \quad (16)$$

where  $Q_n \in \mathbb{R}^{p \times p}$  and  $R \in \mathbb{R}^{r \times r}$  are the weighting matrix and the time-varying  $Q_n$  will be set to CDS in this framework. The feedback gain is updated by

$$[\hat{K}_n, \hat{D}_n] = (R + B^T P B)^{-1} B^T P A, \quad (17)$$

where  $P \in \mathbb{R}^{(2 \times p) \times (2 \times p)}$  is the unique positive definite solution to the following Discrete time Algebraic Riccati Equation (DARE):

$$P = A^T A - A^T P B (R + B^T P B)^{-1} B^T P A + Q_n, \quad (18)$$

where  $Q_n \in \mathbb{R}^{(2 \times p) \times (2 \times p)}$  and  $Q_n = \begin{bmatrix} Q_n & 0_{p \times p} \\ 0_{p \times p} & 0_{p \times p} \end{bmatrix}$ .

In this framework, compared with the heuristic way to construct impedance in [11], LQR is a more systematic method, which takes the environment dynamics into account and holds a good balance between the tracking error and the control

input (which equals to the interaction force in the point view of Cartesian impedance control). The environment dynamics is set to a double integrator system similar to [12]. In the following, we discuss how to implement the full impedance profile on robots by using a task geometry consistent Cartesian impedance controller whose choice of parameters is further discussed in Sec. III.

### C. Robot controller

1) *Task geometry consistent Cartesian impedance control:* At present, the most widely used Cartesian impedance controller for the new generation light-weight robots (such as KUKA LWR and Franka Emika Panda) follows the Jacobian transpose scheme, in which the measurement of external wrench is not necessary by avoiding the inertia shaping. A systematic description of this scheme is reported in [1], where the analytic Jacobian is used due to the minimal representation of the orientation. Such a representation is known to suffer from singularity and the task geometry consistency problem, i.e., the equivalent rotational stiffness between the orientation displacement and the elastic moment, which depends on the desired end-effector orientation. This issue was first discussed in [23], in which an inverse dynamics strategy is employed to implement Cartesian impedance control, resulting in the requirement of measuring contact forces and moments. Although only the translational stiffness is modeled in Sec. II-A and is dominant in most applications, the rotational stiffness still needs to be set when the robot executes a task. To eliminate the effects brought by task geometry consistency, a derivation of task geometry consistent Cartesian impedance control is performed by following the Jacobian transpose scheme.

In contrast with [1], instead of representing the operational space dynamic equation by using analytic Jacobian, here the geometric Jacobian  $J(q) \in \mathbb{R}^{6 \times n}$  is adopted to address the task geometry consistency problem. Without an explicit representation of Cartesian coordinates by using  $v_e = J(q)\dot{q}$  and  $\dot{v}_e = \dot{J}(q)\dot{q} + J(q)\ddot{q}$  and neglecting the gravity term for simplicity, the dynamic model for the manipulator in Cartesian space is

$$\Lambda(q)\dot{v}_e + \mu(q, \dot{q})v_e = F_\tau - F_{ext}, \quad (19)$$

where  $v_e \in \mathbb{R}^6$  is the end-effector velocity,  $q \in \mathbb{R}^n$  is the joint coordinates.  $F_\tau \in \mathbb{R}^6$  and  $F_{ext} \in \mathbb{R}^6$  are respectively the Cartesian control input and external wrench.  $\Lambda(q) \in \mathbb{R}^{6 \times 6}$  and  $\mu(q, \dot{q}) \in \mathbb{R}^{6 \times 6}$  are the Cartesian inertia and Coriolis/centrifugal matrices respectively

$$\Lambda(q) = J(q)^{-T} M(q) J(q)^{-1}, \quad (20)$$

$$\mu(q, \dot{q}) = J(q)^{-T} (C(q, \dot{q}) - M(q) J(q)^{-1} \dot{J}(q)) J(q)^{-1}, \quad (21)$$

where  $M(q)$  and  $C(q, \dot{q})$  are respectively the joint space inertia and Coriolis/centrifugal matrices. The desired Cartesian impedance relationship derived from an energy-based argument when considering task geometry consistency is [23]:

$$\Lambda(q)\dot{\tilde{v}}_{ed} + (\mu(q, \dot{q}) + \hat{D}_n)\tilde{v}_{ed} + \hat{K}_n \begin{bmatrix} p_{ed} \\ o_{ed} \end{bmatrix} = F_{ext}, \quad (22)$$

where  $\tilde{v}_{ed} \in \mathbb{R}^6$  is the velocity error and it equals to  $v_d - v_e$ ,  $v_d \in \mathbb{R}^6$  is the desired end-effector velocity,  $p_{ed} \in \mathbb{R}^3$  is the position error and the orientation error is represented in axis/angle  $o_{ed} \in \mathbb{R}^3$  and  $o_{ed} = r \sin(\theta)$ ,  $R(\theta, r) = R_d R_e^T$ .  $R_d \in \mathbb{R}^{3 \times 3}$  and  $R_e \in \mathbb{R}^{3 \times 3}$  are respectively the rotation matrices from the desired and actual end-effector frames to base frame.  $\hat{D}_n$  and  $\hat{K}_n$  are the impedance parameters derived from Eq. (17), with  $r = p = 6$  in Cartesian space.

By adding Eq. (19) and Eq. (22) to eliminate  $\dot{v}_e$  as in [1], the following control law is derived:

$$F_\tau = \Lambda(q)\dot{v}_d + \mu(q, \dot{q})v_d + \hat{D}_n\tilde{v}_{ed} + \hat{K}_n \begin{bmatrix} p_{ed} \\ o_{ed} \end{bmatrix}. \quad (23)$$

The joint torque command can be computed by following the Jacobian transpose scheme.

$$\begin{aligned} \tau &= J(q)^T F_\tau \\ &= J(q)^T (\Lambda(q)\dot{v}_d + \mu(q, \dot{q})v_d + \hat{D}_n\tilde{v}_{ed} + \hat{K}_n \begin{bmatrix} p_{ed} \\ o_{ed} \end{bmatrix}). \end{aligned} \quad (24)$$

If  $v_d = 0$ , then the simplified task geometry consistent Cartesian impedance controller is

$$\tau = J(q)^T (-\hat{D}_n J(q)\dot{q} + \hat{K}_n \begin{bmatrix} p_{ed} \\ o_{ed} \end{bmatrix}). \quad (25)$$

2) *Tank-based system passivity observer:* Considering that the overall framework assumes a variable impedance control, the passivity of the system and hence the stability may be lost if the impedance values change too fast. In this framework, a tank-based approach proposed in [24] is used to monitor the stability of the system. Formally, the extended dynamics is given by:

$$\begin{aligned} \Lambda(q)\dot{\tilde{v}}_{ed} + (\mu(q, \dot{q}) + \hat{D}_n)\tilde{v}_{ed} + \hat{K}_n^c \begin{bmatrix} p_{ed} \\ o_{ed} \end{bmatrix} - \omega x_t &= F_{ext}, \\ \dot{x}_t &= \frac{\sigma}{x_t} (\tilde{v}_{ed}^\top \hat{D}_n \tilde{v}_{ed}) - \omega^\top \tilde{v}_{ed}. \end{aligned} \quad (26)$$

where  $\hat{K}_n = \hat{K}_n^c + \hat{K}_n^v$ ,  $\hat{K}_n^c \in \mathbb{R}^{6 \times 6}$  is the constant stiffness term and  $\hat{K}_n^v \in \mathbb{R}^{6 \times 6}$  is the time-varying stiffness. The scalar  $x_t \in \mathbb{R}$  is the state associated with the tank and the tank energy  $T \in \mathbb{R}^+$  and  $T = \frac{1}{2}x_t^2$ .  $\sigma \in \mathbb{R}$  and  $\omega \in \mathbb{R}^6$  respectively are

$$\sigma = \begin{cases} 1 & \text{if } T \leq \bar{T}_u \\ 0 & \text{otherwise} \end{cases}, \quad \omega = \begin{cases} -\frac{\hat{K}_n^v [p_{ed}, o_{ed}]^\top}{x_t} & \text{if } T \geq \bar{T}_l \\ 0 & \text{otherwise} \end{cases}. \quad (27)$$

Here,  $\bar{T}_u \in \mathbb{R}^+$  is a suitable, application dependent, upper bound that avoid excessive energy storing in the tank while  $\bar{T}_l \in \mathbb{R}^+$  is a lower bound below, which energy cannot be extracted by the tank for avoiding singularities in Eq. (26) and thus the time-varying stiffness  $\hat{K}_n^v$  will be removed. For a detailed analysis of the system passivity, please refer to [24].

### III. EXPERIMENTS AND RESULTS

The experimental setup is shown in Fig. 1. On the human side of this setup, the MVN Biomech suit (Xsens Tech) was used for collecting the 3D position information required to construct CDS patterns and to train the GMM. On the robot side of this setup, the torque-controlled Franka Emika



Panda arm was used, which was equipped with the Pisa/IIT SoftHand for grasping of the door handle. For the door opening task, the significant motion and interaction happened on X-Y plane. Therefore, in the following, all the stiffness profiles, trajectories and forces data are illustrated on X-Y plane.

As shown in Fig. 4, the subject repeated the door opening task for five times. The black lines represent human arm endpoint trajectories, while the green ellipses represent the demonstrated human arm endpoint stiffness profile. The CDS part was constructed according to the collected 3D position data of human hand, elbow, and shoulder at 100Hz with model parameters ( $\alpha_1 = 0.2450, \alpha_2 = 2.7118$ ), which was reported in [19]. The CMS value was selected experimentally to match the range of trajectories, since it was neither used in the training of the GMM nor for reproduction by GMR. Since the motion was constrained, the demonstrated trajectories didn't show much variability, which means that the method proposed in [12] would fail to provide meaningful stiffness profiles along the trajectory.

After constructing the input (endpoint trajectory) and output (endpoint stiffness) data of the GMM, an EM training process was performed to estimate the model parameters. The reproduction of CDS by GMR was simulated by taking a planned circular trajectory as input. The results are depicted in Fig. 5.

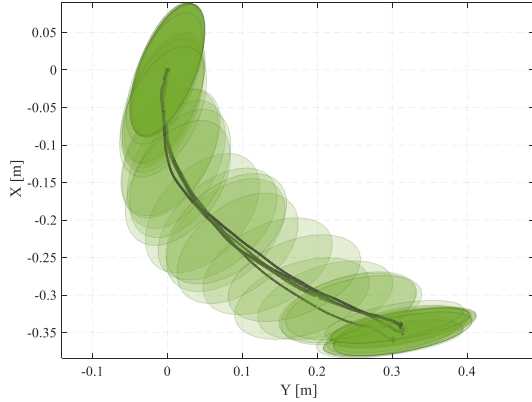


Fig. 4: Human arm endpoint trajectories (black lines) and stiffness profiles (green ellipses) during the demonstration phase. The task stiffness was calculated by the model (resulting by Eq. (2)) and only CDS was extracted from human demonstrations.

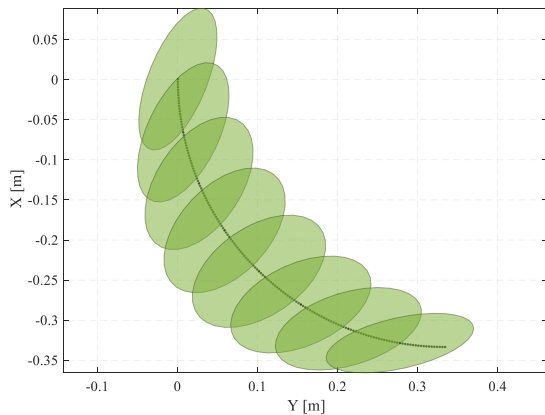


Fig. 5: CDS reproduction (green ellipses) based on a planned circular trajectory (black line) as input.

To reveal the effectiveness and the added value of the two

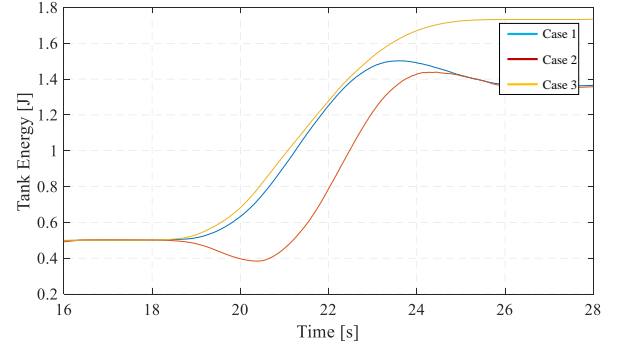


Fig. 6: Tank energy variations in *Case 1* (blue), *Case 2* (red) and *Case 3* (yellow). The lower bound was set  $\bar{T}_l = 0.2J$ .

main components of the proposed framework (i.e., imitation learning and optimal control), a comparative experiment was conducted under three different cases.

*Case 1:* This case implements the proposed framework: The CDS extracted from human demonstrations was encoded by GMM offline and reproduced by GMR online (current end-effector position as the input). For LQR, the weighting matrix  $Q$  was set to CDS and  $R$  was set to  $1 \times 10^{-6} \mathbf{I}$  ( $\mathbf{I}$  is identity matrix), which was decided by limiting the resulted largest stiffness within the stiffness range of Franka Emika Panda robot.

*Case 2:* In this case, the process of CDS information was the same as in *Case 1*. But instead of using LQR, a constant CMS was set to 1040.37 which is decided by keeping average area of stiffness ellipses on X-Y plane along the trajectory is the same in *Case 1* and *Case 2*. The damping term was set to guarantee a damping ratio of  $\frac{1}{\sqrt{2}}$ .

*Case 3:* Here, instead of encoding and reproducing the time-varying CDS by GMM/GMR, a constant CDS was used, which equals to the value of the generated CDS at the starting point in *Case 1*, and then a LQR procedure was followed.

In all above cases, the desired path was generated by the same circular trajectory planner, and the environment dynamics was set to a double integrator system. The tank-based stability observer was used to monitor system passivity. As shown in Fig. 6, the tank energy was above the lower bound ( $\bar{T}_l = 0.2J$ ) in all three cases, which means that the full stiffness including the constant (i.e. the average stiffness along the whole path was considered as the constant stiffness in each case) and varying parts was realised. The resulted trajectories and the stiffness profiles of Franka Emika Panda in three cases are shown in Fig. 7. Compared with *Case 1*, the stiffness ellipses in *Case 2* without the LQR process were more elongated, which may cause large interaction force when the door is in the middle. The stiffness ellipses in *Case 3*, which used a constant CDS with the LQR process demonstrated a more isotropic shape in comparison to those in *Case 2*. However, due to the use of a constant CDS, although the resulted stiffness matched the task requirements in the beginning of the trajectory, however, it didn't yield a desired stiffness ellipse geometry in the end.

Due to the difference in impedance behaviors in the above three cases, the trajectory tracking and the interacting performance of the robot were different. The resulting tracking errors

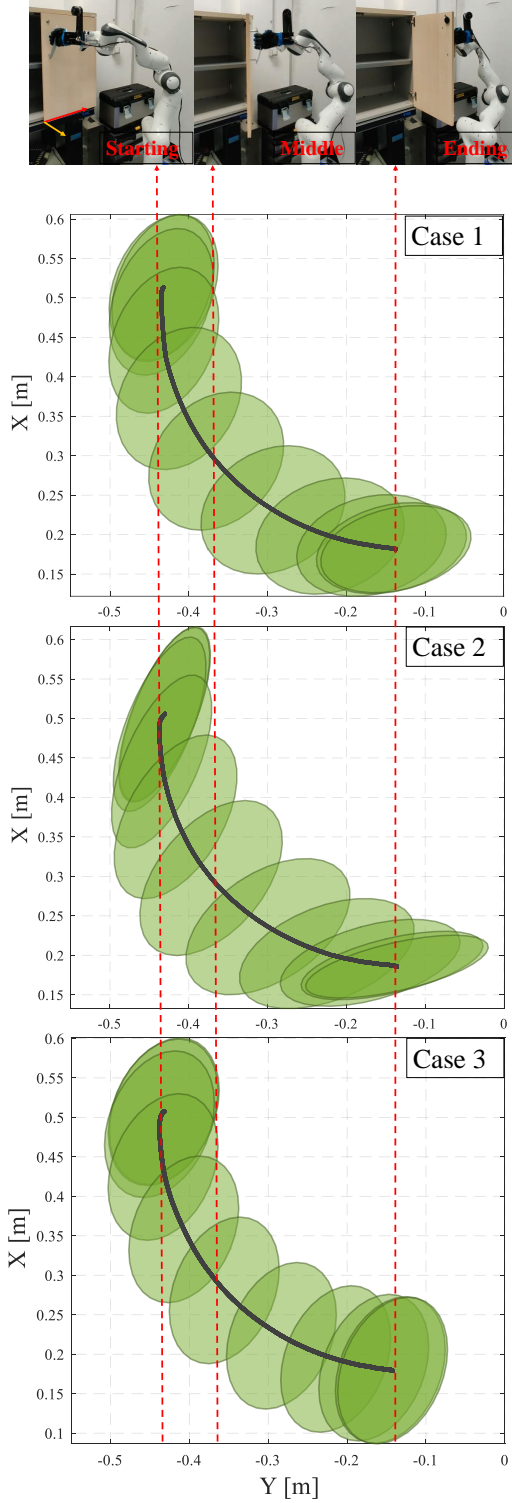


Fig. 7: The executed trajectories (black lines) and stiffness profiles (green ellipses) in three cases. The snapshots shows the starting, middle and ending phases of the door opening experiment, the solid red and yellow arrows attached on the door respectively indicate the constrained-motion and free-motion directions.

and the interaction forces along the constrained-motion and free-motion directions (as shown in Fig. 7, the solid red and yellow arrows attached on the door respectively indicate the constrained-motion and free-motion directions.) are illustrated in Fig. 8 and Fig. 9, respectively. The door opening task lasted 10s, starting and ending at 17s and 27s respectively. The door

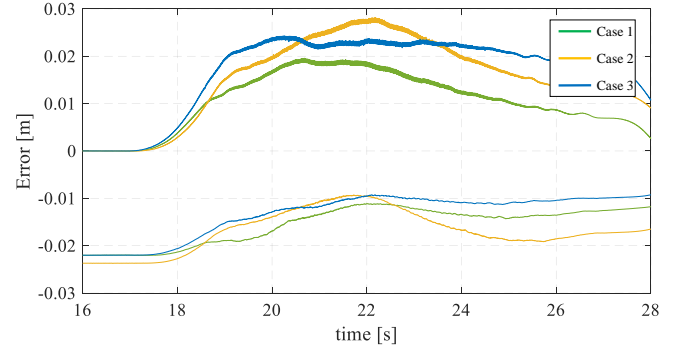


Fig. 8: Tracking error in the constrained (bottom) and free (top) motion directions in *Case 1* (green), *Case 2* (yellow) and *Case 3* (blue).

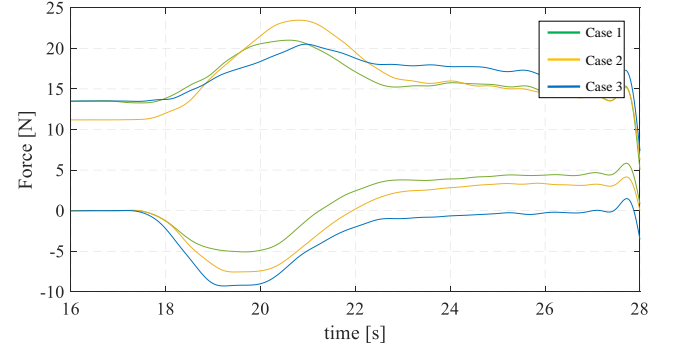


Fig. 9: Interaction forces (control input) in the constrained (top) and free (bottom) motion directions in *Case 1* (green), *Case 2* (yellow) and *Case 3* (blue).

was in the middle at 22s. In the free-motion direction, the tracking errors (top in Fig. 8) in three cases were initialised by the same zero value, thus resulted in the same zero force initial values (bottom in Fig. 9). *Case 1* showed less tracking errors and less control input compared with *Case 2* and *Case 3* along almost the entire path. In the constrained-motion direction in which a compliant behavior was expected, due to the existence of an initial position mismatch and different stiffness settings, *Case 2* showed a different initial tracking error (bottom in Fig. 8) and interaction force (top in Fig. 9) behavior compared with *Case 1* and *Case 3* which were using the same stiffness values at the starting point. During the door moving to the middle, the tracking error in all three cases were decreasing, and it was less in *Case 2* and *Case 3* due to a larger stiffness than that in *Case 1*. However, the interaction force in *Case 2* was much larger than the other two, which was unnecessary. When the robot moved the door from the middle to the end, the tracking error in *Case 1* and *Case 2* was increasing since the stiffness in that direction was decreasing. In *Case 3* the tracking error didn't show obvious changes because of the near constant stiffness settings, however, it resulted in a larger interaction force due to the inconsistency between stiffness and task geometry in the end.

Due to the fact that the varying CDS encoded by GMM/GMR is not directly related to time but to current end-effector position, the online generated impedance parameters were always synchronized with the task geometry even if external interventions (external forces were applied to confront the door's motion at around 11s and 27s) were applied, as

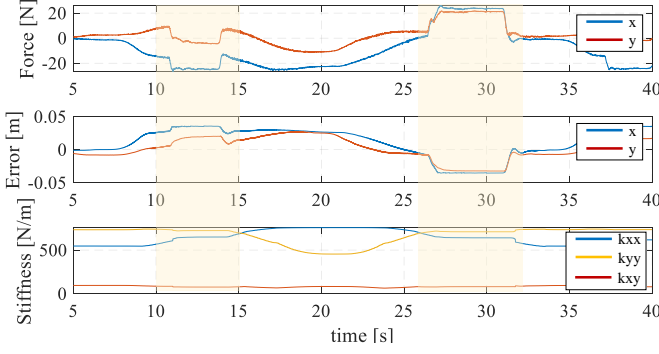


Fig. 10: External intervention happened during the period covered by yellow area, stiffness synchronises with task performing phase. When the tracking error reached 0.05m, the trajectory planner stopped updating to avoid large interaction forces in the free-motion direction.

shown in Fig. 10. In the plots, when an external force was applied and the end-effector stopped moving, GMR stopped updating as well to avoid task stiffness geometry mismatch which would cause large interaction forces in the constrained-motion direction. This highlights the dependency of the GMR on task states and not on task timing, guaranteeing the adaptation of the stiffness profile according to the task requirements.

Although the method proposed in this work outperforms the other two, one may argue that the resulted stiffness profile doesn't perfectly match the task requirements, e.g., in the middle of the task, where the major principle axis of the stiffness ellipse was in the constrained-motion direction. If we consider the problem from the view of a moving task frame attached to the door, a more intuitive and easier solution for the stiffness regulation problem can be a constant task stiffness geometry ellipse in the task frame, whose major and minor principle axes are respectively aligned with the constrained-motion and free-motion directions. However, this approach is rather heuristic based on prior assumptions and knowledge of the task. Instead, the proposed framework has the capacity to encode more complex interaction behaviours from human demonstrations and reproduce them via strong machine learning and optimal control algorithms.

#### IV. CONCLUSIONS

In this work, we proposed a framework for autonomous impedance regulation of robots based on imitation learning and optimal control, which resulted in task- and dynamics-aware impedance planning without an explicit time dependency. The effectiveness of the framework was evaluated by conducting a comparative experiment including three different cases for a door opening task, as an illustrative example of constrained and continuously interactive actions. Results provide evidence on the superior performance of the proposed framework in terms of tracking error and interaction forces, through optimal and adaptive planning of robot interaction parameters.

#### REFERENCES

- [1] C. Ott, *Cartesian impedance control of redundant and flexible-joint robots*. Springer, 2008.
- [2] 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.

- [3] K. Kronander and A. Billard, "Learning compliant manipulation through kinesthetic and tactile human-robot interaction," *IEEE transactions on haptics*, vol. 7, no. 3, pp. 367–380, 2013.
- [4] K. P. Tee, E. Burdet, C.-M. Chew, and T. E. Milner, "A model of force and impedance in human arm movements," *Biological cybernetics*, vol. 90, no. 5, pp. 368–375, 2004.
- [5] E. Burdet, K. P. Tee, I. Mareels, T. E. Milner, C.-M. Chew, D. W. Franklin, R. Osu, and M. Kawato, "Stability and motor adaptation in human arm movements," *Biological cybernetics*, vol. 94, no. 1, pp. 20–32, 2006.
- [6] A. Ajoudani, N. Tsagarakis, and A. Bicchi, "Tele-impedance: Teleoperation with impedance regulation using a body-machine interface," *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1642–1656, 2012.
- [7] A. Ajoudani, C. Fang, N. Tsagarakis, and A. Bicchi, "Reduced-complexity representation of the human arm active endpoint stiffness for supervisory control of remote manipulation," *The International Journal of Robotics Research*, vol. 37, no. 1, pp. 155–167, 2018.
- [8] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, and A. Ajoudani, "A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3593–3600, 2019.
- [9] 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.
- [10] S. Calinon and D. Lee, "Learning control," in *Humanoid Robotics: a Reference*, P. Vadakkepat and A. Goswami, Eds. Springer, 2019, pp. 1261–1312.
- [11] 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.
- [12] S. Calinon, D. Bruno, and D. G. Caldwell, "A task-parameterized probabilistic model with minimal intervention control," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 3339–3344.
- [13] L. D. Roza, S. Calinon, D. Caldwell, P. Jiménez, and C. Torras, "Learning collaborative impedance-based robot behaviors," in *Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.
- [14] F. J. Abu-Dakka, L. Roza, and D. G. Caldwell, "Force-based variable impedance learning for robotic manipulation," *Robotics and Autonomous Systems*, vol. 109, pp. 156–167, 2018.
- [15] M. Matinfar and K. Hashtrudi-Zaad, "Optimization-based robot compliance control: Geometric and linear quadratic approaches," *The International Journal of Robotics Research*, vol. 24, no. 8, pp. 645–656, 2005.
- [16] S. S. Ge, Y. Li, and C. Wang, "Impedance adaptation for optimal robot-environment interaction," *International Journal of Control*, vol. 87, no. 2, pp. 249–263, 2014.
- [17] F. L. Lewis and K. G. Vamvoudakis, "Reinforcement learning for partially observable dynamic processes: Adaptive dynamic programming using measured output data," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 1, pp. 14–25, 2010.
- [18] A. Ajoudani, M. Gabbicini, N. Tsagarakis, A. Albu-Schäffer, and A. Bicchi, "Teleimpedance: Exploring the role of common-mode and configuration-dependant stiffness," in *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*. IEEE, 2012, pp. 363–369.
- [19] Y. Wu, F. Zhao, W. Kim, and A. Ajoudani, "An intuitive formulation of the human arm active endpoint stiffness," *Sensors*, vol. 20, no. 18, p. 5357, 2020.
- [20] F. A. Mussa-Ivaldi, N. Hogan, and E. Bizzi, "Neural, mechanical, and geometric factors subserving arm posture in humans," *Journal of Neuroscience*, vol. 5, no. 10, pp. 2732–2743, 1985.
- [21] V. M. Zatsiorsky and V. M. Zatsiorskij, *Kinetics of human motion*. Human Kinetics, 2002.
- [22] Y. Huang, L. Roza, J. Silvério, and D. G. Caldwell, "Kernelized movement primitives," *The International Journal of Robotics Research*, vol. 38, no. 7, pp. 833–852, 2019.
- [23] F. Caccavale, C. Natale, B. Siciliano, and L. Villani, "Six-dof impedance control based on angle/axis representations," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 2, pp. 289–300, 1999.
- [24] F. Ferraguti, C. Secchi, and C. Fantuzzi, "A tank-based approach to impedance control with variable stiffness," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 4948–4953.