Neural Networks Enhanced Adaptive Admittance Control of Optimized Robot–Environment Interaction

Chenguang Yang, *Senior Member, IEEE*, Guangzhu Peng, Yanan Li, *Member, IEEE*, Rongxin Cui, *Member, IEEE*, Long Cheng, *Senior Member, IEEE*, and Zhijun Li, *Senior Member, IEEE*

***Abstract*—In this paper, an admittance adaptation method has been developed for robots to interact with unknown environments. The environment to be interacted with is modeled as a linear system. In the presence of the unknown dynamics of environments, an observer in robot joint space is employed to estimate the interaction torque, and admittance control is adopted to regulate the robot behavior at interaction points. An adaptive neural controller using the radial basis function is employed to guarantee trajectory tracking. A cost function that defines the interaction performance of torque regulation and trajectory tracking is minimized by admittance adaptation. To verify the proposed method, simulation studies on a robot manipulator are conducted.**

***Index Terms*—Admittance control, neural networks (NNs), observer, optimal adaptive control, robot–environment interaction.**

I. Introduction

With the rapid development of robot technology, robots have been widely studied in various areas such as education, industry, entertainment, etc. In these areas, robots inevitably interact with external environments [1]–[3]. Therefore, robots interacting with the environment have received great attention and much effort has been made in this paper. Although it has been investigated for more than several decades, there are still many open problems to be solved, due to the high expectation of robots in more general scenarios and the complex environment in which robots are working. In order to achieve compliant behavior, there are three approaches that are widely applied: 1) admittance control; 2) hybrid position/force control; and 3) impedance control.

The concept of impedance control introduced by Hogan [4] has been a very popular control method in robotics. Impedance control aims to regulate robot behavior at physical interaction points. The core idea of impedance control is that the controller should modulate the mechanical impedance, which is a mapping from generalized velocities to generalized force. This control approach has proven to be feasible and robust [5]. Another approach is admittance control, which was introduced by Mason [6]. In a generalized admittance control system, with the measurements of environmental force and a desired admittance model, a reference trajectory is obtained and tracked. Then, the compliant behavior is realized by trajectory adaptation. Traditional control methods of a robot manipulator are usually based on system models, which could have good control performance [7]. However, these methods require accurate robot models, which cannot be obtained in many cases. Therefore, adaptive control methods have attracted lots of attention and have been applied in practical systems [8]. Intelligent tools were used to approximate uncertainties in the adaptive control [9]–[13]. Another key element in the admittance control system is the force sensor. Force sensors are regarded as a medium for communication between a robot and the environment. However, force sensors equipped on the manipulators may cause inconvenience and are usually costly. Due to these reasons, sensorless control schemes have received great attention. There are two main methods for estimating the external force: 1) disturbance observer approach and 2) force observer approach based on the knowledge of motor torques. In [14], the disturbance observer approach with knowledge of a joint angle has been analyzed. In [15], a force observer based on the generalized momentum for collision detection has been introduced. In [16], the collision detection problem of rigid robot arms and other robots with elastic joints has been studied.

Under impedance/admittance control, an impedance/admittance model is of great importance. In practice,

Manuscript received November 28, 2017; revised February 15, 2018; accepted April 4, 2018. This work was supported in part by the National Nature Science Foundation under Grant 61473120, Grant 61633016, and Grant 61472325, in part by the Science and Technology Planning Project of Guangzhou under Grant 201607010006, in part by the Fundamental Research Funds for the Central Universities under Grant 2017ZD057, in part by the Research Fund for Young Top-Notch Talent of National Ten Thousand Talent Program, and in part by the Beijing Municipal Natural Science Foundation under Grant 4162066. This paper was recommended by Associate Editor J. Q. Gan.

C. Yang, G. Peng, and Z. Li are with the Key Laboratory of Autonomous Systems and Networked Control, College of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China (e-mail: cyang@ieee.org; gz.peng@qq.com; zjli@ieee.org).

Y. Li is with the Department of Engineering and Design, University of Sussex, Brighton BN1 9RH, U.K. (e-mail: yl557@sussex.ac.uk).

R. Cui is with the School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: r.cui@nwpu.edu.cn).

L. Cheng is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: long.cheng@ia.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2018.2828654

obtaining the desired impedance/admittance model is not easy due to the complexity of the environmental dynamics. Moreover, a fixed impedance/admittance model could not be applied to all situations. To solve this problem, iterative learning has been widely studied for robots to adapt to unknown environments. This approach aims to introduce human learning skills to robots by repeating a task, and has been widely studied. In [17], an associative search network-learning scheme was presented for learning control parameters for robot to complete a wall-following task. In [18], the neural-network (NN)-based method was applied to regulate impedance parameters of the end-effector of a robot. However, the disadvantage of the learning method is that repeated operations for robots are needed to learn impedance parameters, which may bring a lot inconvenience in practical situations. Therefore, the impedance adaptation method has been studied in [19]. In [20], strategies of switching among different values of impedance parameters were developed for dissipating the energy of the system. In [21], an impedance adaptation method for robots to interact with unknown environment was proposed.

The control objective of interaction control is to achieve force regulation and trajectory tracking. Thus, optimization should be taken into consideration, since it is the compromise of these two objectives. The well-known linear quadratic regulator (LQR) is widely acknowledged as an important solution of optimal control, which is concerned with operating a dynamic system at minimal cost. In [22], desired impedance parameters were found by the LQR method with the environmental dynamics known. In [23], a target impedance was adjusted by online solutions of the defined LQR problem based on environment stiffness and damping. Better than the fixed impedance parameters obtained from the LQR technique, the algorithm shows greater adaptability for a wide range of environments. However, the dynamics of the environment is also assumed to be known in this paper above. As presented in [24], the solution of a Riccati equation could be difficult to find with the unknown dynamics of an environment. Therefore, when the dynamics of a environment is unknown, the approaches proposed above may not be used. To cope with this problem, adaptive dynamic programming (ADP) has received much attention and has been widely studied [25]–[29], where the ADP was used to solve the optimization problems. Based on the idea of ADP, a control action is modified by the feedback information of an environment. There are many ADP approaches such as heuristic dynamic programming and Q-learning. The advantage of ADP is that only partial information of the system under control needs to be known. In [30], a least-square-based algorithm was used to update impedance parameters. In [31], the reinforcement learning algorithm was adopted to accomplish variable impedance control. However, in many situations, the learning process is still needed to obtain parameters of an impedance/admittance model [32]. The optimal control method with an unknown environment proposed in [33] was applied and developed. The environmental model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, an admittance adaptation method is proposed to adjust admittance parameters and achieve the optimal interaction behavior with unknown environmental dynamics. An environment with unknown dynamics is taken into consideration and modeled as a linear system in the state-space form. The control objective is to minimize the defined cost function to have good interaction performance. An observer based on the generalized momentum approach is used to estimate the interaction torque, while the tracking performance is guaranteed by an NN-based controller.

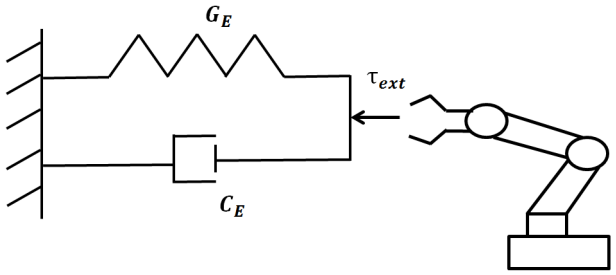


Fig. 1. Model of damping-stiffness environment.

The rest of this paper is organized as follows. The dynamics of a robot and control objective as well as the unknown environmental dynamics are introduced in Section II. Section III introduces the methodology including estimation of the interaction torque, admittance adaptation, and controller design. In Section IV, simulations studies are given. The conclusion is drawn in Section V.

II. Preliminaries

A. System Description

In this system, we consider that a robot arm is interacting with an environment. The forward kinematics can be expressed by

$$x = \kappa(q) \tag{1}$$

where $\kappa(\cdot)$, x, $q \in \mathbb{R}^n$, and n denote forward kinematics function, pose of the end-effector in the Cartesian space, joint vectors, and the DOF number, respectively. The derivative of (1) along time is

$$\dot{x} = J(q)\dot{q} \tag{2}$$

where matrix $J(q) \in \mathbb{R}^{n\times n}$ is the manipulator Jacobian. Further differentiating (2) along time, yields

$$\ddot{x} = \dot{J}(q)\dot{q} + J(q)\ddot{q}. \tag{3}$$

The dynamics of a robot arm can be described by

$$M_q(q)\ddot{q} + C_q(q,\dot{q})\dot{q} + G_q(q) + \tau_{\text{ext}} = \tau \tag{4}$$

where $M_q(q) \in \mathbb{R}^{n\times n}$ is the inertia matrix in joint space. $C_q(q,\dot{q}) \in \mathbb{R}^{n\times n}$ is the Coriolis and centripetal coupling matrix. $G_q(q) \in \mathbb{R}^{n\times 1}$ is the gravity loading. $\dot{q}$ and $\ddot{q}$ are joint velocities and accelerations, respectively. $\tau \in \mathbb{R}^n$ is the control input. $\tau_{\text{ext}} \in \mathbb{R}^n$ denotes the vector of joint torque exerted by the environment.

Property 1: The matrix $M_q(q)$ is symmetric and positive definite [34].

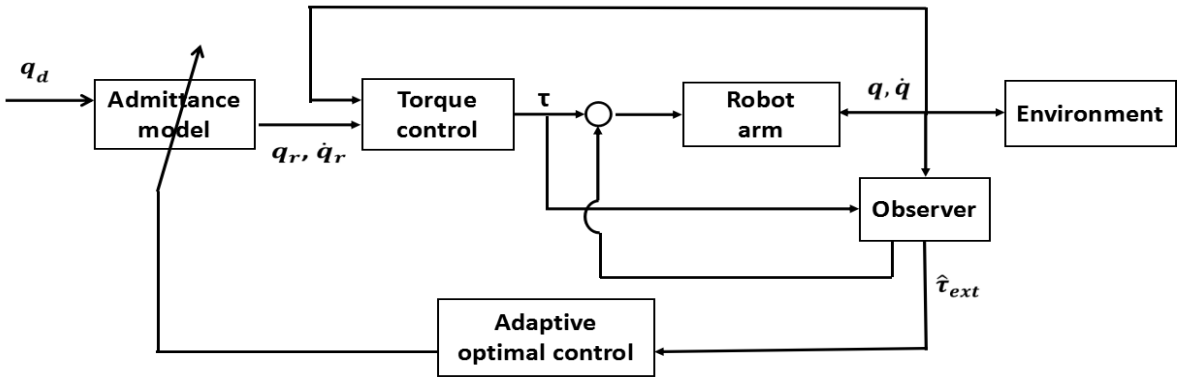


Fig. 2. Control diagram.

Property 2: The matrix $2C_q(q,\dot{q}) - \dot{M}_q(q)$ is skew-symmetric [34].

Now, let us consider the modeling of the environment. In this paper, a damping-stiffness environment model is considered, as shown in Fig. 1

$$C_E\dot{q} + G_E q = -\tau_{\text{ext}} \tag{5}$$

where C_E and G_E denote damping and stiffness matrices with unknown constants, respectively.

Remark 1: Without loss of generality, the spring-damper system is usually used to describe the model of an environment. Compared to the model (5), a simple second-order equation is used as the model for an environment. In the model, mass, damping and stiffness are considered: $M_E\ddot{q} + C_E\dot{q} + G_E q = -\tau_{\text{ext}}$. In general, a large range of environments can be represented by these two models. For analysis convenience, the spring-damper system is considered as a time-invariant system, i.e., the three coefficient matrices are constant matrices.

B. Control Strategy

The system diagram is given in Fig. 2. The outer-loop of the system is to determine admittance parameters subject to an unknown environment based on the adaptive optimal control method. With the interaction torque estimated from an observer, a reference trajectory can be generated from the admittance model. The inner-loop of the system is to guarantee the manipulator could track the reference trajectory q_r effectively.

In general, the desired admittance model in the Cartesian space is

$$f_{\text{ext}} = f(x_r, x_d) \tag{6}$$

where $x_r \in \mathbb{R}^n$ is the reference trajectory in the Cartesian space and $f(\cdot)$ is the function of the admittance model. To define the relationship between interaction torque and joint angles, a general admittance model is described as

$$\begin{aligned} &M_d J(q)(\ddot{q}_r - \ddot{q}_d) + \big(M_d\dot{J}(q) + C_d J(q)\big)(\dot{q}_r - \dot{q}_d) \\ &\quad + G_d(\kappa(q_r) - \kappa(q_d)) = -J^{-T}(q)\tau_{\text{ext}} \end{aligned} \tag{7}$$

where M_d, C_d, and G_d denote desired parameter matrices of the admittance mode.

Remark 2: In certain situations, a more simplified stiffness model may be adopted

$$G_d(q_r - q_d) = -\tau_{\text{ext}}. \tag{8}$$

Besides, M_d, C_d, and G_d are constant matrices with different diagonal elements, respectively. When $\tau_{\text{ext}} = 0$, q_d is the desired trajectory to be tracked in the absence of interaction. However, when τ_{ext} is not null, the reference trajectory q_r will be generated [35].

The adaptive control scheme is to let robot follow the reference trajectory and drive the tracking error $e_q = q - q_r$ near zero. The design of adaptive control scheme will be discussed in the next section.

C. Control Objective

The control objective is to achieve an optimal interaction performance and the following cost function is defined:

$$V(t) = \int_0^\infty \big((q - q_d)^T Q(q - q_d) + \hat{\tau}_{\text{ext}}^T R\hat{\tau}_{\text{ext}}\big)dt \tag{9}$$

where $Q = Q^T \in \mathbb{R}^{n\times n}$ is positive definite, describing the weight of tracking errors, and $R \in \mathbb{R}^{n\times n}$ is the weight of the interaction torque. By minimizing $V(t)$, a desired interaction performance can be achieved.

Remark 3: Cost function (9) is a tradeoff between trajectory tracking and torque regulation and a similar form has been studied in [23]. In a classical LQR problem, trajectory tracking and control inputs are included in the quadratic function, which is solved by specified feedback gains to achieve the optimal control performance. In this paper, both robot and environmental systems are taken into consideration.

III. Methodology

The objective of the admittance adaptation is to obtain the target admittance model according to the changing environment. The inner-loop of the system is to ensure the robot arm track the reference trajectory effectively. The interaction torque from the environment is estimated by the force observer.

A. Torque Estimation

In this section, an observer based on the generalized momentum approach is used to estimate the external torque

in joint space. Compared with traditional methods that require computation of joint accelerations or the inversion of the inertia matrix [36], the generalized momentum approach assumes that only motor torque τ, joint angle q and joint velocity $\dot{q}$ are available. In [36], the generalized momentum is defined as

$$p = M_q(q)\dot{q}. \tag{10}$$

Its differential form along time is

$$\dot{p} = \dot{M}_q\dot{q} + M_q\ddot{q}. \tag{11}$$

Substituting (11) into (4), one have

$$\dot{p} = \dot{M}_q(q)\dot{q} + \tau - C_q(q,\dot{q})\dot{q} - G_q(q) - \tau_{\text{ext}}. \tag{12}$$

Consider Property 1 and Coriolis matrix with Christoffel symbols [34], the time derivative of inertia matrix M is

$$\dot{M} = C + C^T. \tag{13}$$

Substituting (13) into (12), we have

$$\dot{p} = C_q^T(q,\dot{q})\dot{q} + \tau - G_q(q) - \tau_{\text{ext}}. \tag{14}$$

It is obvious that (14) based on the generalized momentum does not involve joint angle acceleration $\ddot{q}$. Finally, the external torques can be modeled as

$$\dot{\tau}_{\text{ext}} = A_\tau\tau_{\text{ext}} + w_\tau \tag{15}$$

where w_τ is the uncertainty, $w_\tau \sim N(0, Q_\tau)$. Usually, the matrix A_τ is defined as $A_\tau = 0_{n\times n}$. However, a negative diagonal matrix can reduce the offset of the estimation of disturbances. Then, (14) can be rewritten as

$$\dot{p} = u - \tau_{\text{ext}} \tag{16}$$

where u is defined as

$$u = \tau + C_q^T(q,\dot{q})\dot{q} - G_q(q). \tag{17}$$

The above equations can be combined and reformulated in the state-space form [37]

$$\underbrace{\begin{bmatrix}\dot{p}\\ \dot{\tau}_{\text{ext}}\end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix}0_n & -I_n\\ 0_n & A_\tau\end{bmatrix}}_{A_c}\underbrace{\begin{bmatrix}p\\ \tau_{\text{ext}}\end{bmatrix}}_{x} + \underbrace{\begin{bmatrix}I_n\\ 0_n\end{bmatrix}}_{B_c}u + \underbrace{\begin{bmatrix}0\\ w_\tau\end{bmatrix}}_{w}$$
$$y = \underbrace{\begin{bmatrix}I_n & 0_n\end{bmatrix}}_{C_c}\begin{bmatrix}p\\ \tau_{\text{ext}}\end{bmatrix} + v \tag{18}$$

where v is the measurement noise $v \sim N(0, R_c)$. It can be proved that this system is observable. Since q and $\dot{q}$ are able to be measured, the generalized momentum $p = M_q(q)\dot{q}$ can be regarded as a measured variable. Then, a state observer is designed

$$\begin{cases}\dot{\hat{x}} = A_c\hat{x} + B_cu + L\big(y - \hat{y}\big)\\ \hat{y} = C_c\hat{x}.\end{cases} \tag{19}$$

The gain matrix L in the system can be calculated by

$$L = PC_c{}^TR_c^{-1} \tag{20}$$

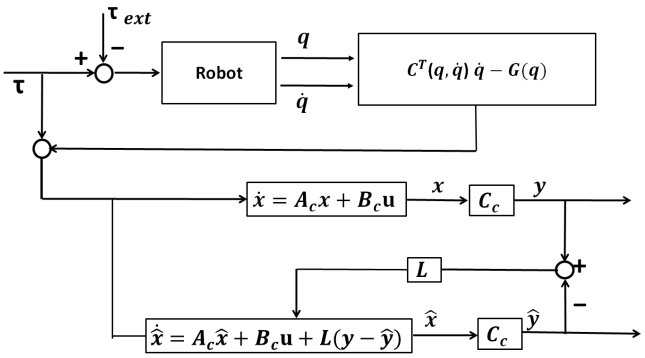


Fig. 3. Diagram of the observer based on generalized momentum approach.

where the matrix P can be solved by the following algebraic Riccati equation (ARE) [38]:

$$A_cP + PA_c{}^T - PC_c{}^TR_c^{-1}C_cP + Q_c = 0 \tag{21}$$

where Q_c is the uncertainty of the state, written as

$$Q_c = \text{diag}([0, Q_\tau]). \tag{22}$$

A schematic overview of the observer is shown in Fig. 3. As shown in (19), the output $y = C_cx(t)$ is compared with $C_c\hat{x}(t)$. If the gain matrix L is properly designed, the difference, passing through the gain matrix, will drive the estimated state to actual state. From the above analysis, we can see that the estimation of states can be obtained from estimated state $\hat{x}$, which can be written as

$$\hat{\tau}_{\text{ext}} = \begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\hat{x}$$
$$\hat{p} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\hat{x}. \tag{23}$$

B. Computational Adaptive Optimal Control

In [33], a computational adaptive optimal control strategy is proposed, which is outlined in the following. Consider a continuous-time linear system

$$\dot{\xi}(t) = A\xi(t) + Bu(t) \tag{24}$$

where $\xi \in \mathbb{R}^m$ is the system state variable. $u \in \mathbb{R}^r$ is the system input. $A \in \mathbb{R}^{m\times m}$ is the system matrix and $B \in \mathbb{R}^{m\times r}$ is the input matrix with being unknown. The following optimal control input:

$$u = -K\xi \tag{25}$$

which can minimize the cost function as follows:

$$V = \int_0^\infty \big(\xi^TQ\xi + u^TR\,u\big)dt. \tag{26}$$

The solution of (26) is similar to that of the LQR problem. In optimal control theory [39], based on known matrices A and B, we can solve ARE to obtain matrix P^* as

$$PA + A^TP + Q - PBR^{-1}B^TP = 0. \tag{27}$$

Then, we can obtain the optimal feedback gain matrix

$$K^* = -R^{-1}B^TP^*. \tag{28}$$

Algorithm 1 Admittance Adaptation Algorithm

1: Choose $u = K_0\xi + \phi$ as the initial admittance model, where K_0 is the initial feedback gain with the exploration noise ϕ. Compute $\delta_{\xi\xi}, I_{\xi\xi}, I_{\xi u}$ until the rank condition in equation (30) is satisfied.
2: Solve P_k and K_{k+1} in equation (32).
3: Let $K+1 \rightarrow K$ and repeat step 2 above until $||P_k - P_{k-1}|| < \varepsilon$ with ε being a small constant.
4: Use $u = K_k\xi$ as the approximated optimal control input.

Therefore, we can obtain the optimal control input from (28). Next, we use the online learning algorithm to obtain the optimal control input subject to unknown dynamics of the environment.

For convenience, let us introduce the following definitions: $\delta_{\xi\xi} \in \mathbb{R}^{l\times(1/2)m(m+1)}$, $I_{\xi\xi} \in \mathbb{R}^{l\times m^2}$, $I_{\xi u} \in \mathbb{R}^{l\times mr}$, $P_k \in \mathbb{R}^{m\times m} \rightarrow \hat{P}_k \in \mathbb{R}^{(1/2)m(m+1)}$, and $\xi \in \mathbb{R}^m \rightarrow \mathbb{R}^{(1/2)m(m+1)}$, where l and P_k denote positive integer and symmetric matrix, respectively

$$\begin{aligned}
\hat{P}_k &= [P_{11},\ 2P_{12},\ldots,\ 2P_{1m},\ P_{22},\ 2P_{23},\ldots,\ P_{mm}]^T\\
\bar{\xi} &= \left[\xi_1^2,\ \xi_1\xi_2,\ldots,\ \xi_1\xi_m,\ \xi_2^2,\ \xi_2\xi_3,\ldots,\ \xi_m^2\right]^T\\
\delta_{\xi\xi} &= \left[\bar{\xi}(t_1)-\bar{\xi}(t_0),\ \bar{\xi}(t_2)-\bar{\xi}(t_1),\ \ldots,\bar{\xi}(t_l)-\bar{\xi}(t_{l-1})\right]^T\\
I_{\xi\xi} &= \left[\int_{t_0}^{t_1}\xi\otimes\xi dt,\ \int_{t_1}^{t_2}\xi\otimes\xi dt,\ldots,\ \int_{t_{l-1}}^{t_l}\xi\otimes\xi dt\right]^T\\
I_{\xi u} &= \left[\int_{t_0}^{t_1}\xi\otimes u dt,\ \int_{t_1}^{t_2}\xi\otimes u dt,\ldots,\ \int_{t_{l-1}}^{t_l}\xi\otimes u dt\right]^T
\end{aligned} \tag{29}$$

where $\otimes$ is the Kronecker product. Now, we let $u = K_0\xi + \phi$ be the initial input, $t \in [t_0,\ t_l]$. ϕ is the exploration noise and K_0 is initial feedback gain, which can stabilize the system. Then, compute $I_{\xi\xi}$ and $I_{\xi u}$ until the following rank condition is satisfied:

$$\text{rank}\left(\left[I_{\xi\xi},\ I_{\xi u}\right]\right) = [(m(m+1))/2] + mr. \tag{30}$$

After the rank condition is satisfied, we can solve P_k and K_{k+1} as

$$\begin{bmatrix}\hat{P}_k\\ \text{vec}(K_{k+1})\end{bmatrix} = \left(\Theta_k^T\ \Theta_k\right)^{-1}\Theta_k^T\Xi_k. \tag{31}$$

Θ_k and Ξ_k are defined as

$$\begin{aligned}
\Theta_k &= \left[\delta_{\xi\xi}, -2I_{\xi\xi}\left(I_m\otimes K_k^TR\right) - 2I_{\xi u}(I_m\otimes R)\right]\\
\Xi_k &= -I_{\xi\xi}vec(Q_k)\\
Q_k &= Q + K_k^TRK_k
\end{aligned} \tag{32}$$

where I_m and $vec(\cdot)$ denote an m-dimensional unit matrix and a transformation from a matrix to vector, respectively. Then, we repeat the calculation until $||P_k - P_{k-1}|| < \varepsilon$ with ε being a small constant. Finally, the optimal feedback gain K_k is obtained. This learning algorithm is summarized in Algorithm 1.

C. Admittance Adaptation

This section is to obtain a target admittance model according to an unknown environment. The environment model is assumed to be damping-stiffness as described in (5). The cost function defined in (9) is to be minimized. Compared with the general linear system (24), we need to make them identical [21]. Define the state variable

$$\xi = \left[q^T, q_d^T\right]^T. \tag{33}$$

Then, (9) can be

$$\begin{aligned}
V &= \int_0^\infty\left([q^Tq_d^T]Q'\begin{bmatrix}q\\ q_d\end{bmatrix} + \hat{\tau}_{\text{ext}}^T\ R\ \hat{\tau}_{\text{ext}}\right)dt\\
&= \int_0^\infty\left(\xi^TQ'\xi + \hat{\tau}_{\text{ext}}^T\ R\ \hat{\tau}_{\text{ext}}\right)dt
\end{aligned} \tag{34}$$

where

$$Q' = \begin{bmatrix}Q & -Q\\ -Q & Q\end{bmatrix}. \tag{35}$$

Combined with the defined state variable, we can rewrite the environment model into state-space form

$$\dot{\xi} = A\xi + B\hat{\tau}_{\text{ext}} \tag{36}$$

where

$$A = \begin{bmatrix}-C_E^{-1}G_E & 0\\ 0 & I_n\end{bmatrix},\quad B = \begin{bmatrix}-C_E^{-1}\\ 0\end{bmatrix}. \tag{37}$$

It is obvious that matrices A and B contain the unknown dynamics of the environment. If $\hat{\tau}_{\text{ext}}$ is taken as the input of the system (36), we can use the method discussed above to obtain the control input as follows to minimize the cost function:

$$\hat{\tau}_{\text{ext}} = -K_k\xi \tag{38}$$

where K_k will be obtained by the online learning algorithm discussed in the previous section.

Understanding (38) in the sense of LQR, the optimal control input is obtained in (28). According to the solution of ARE, we can obtain the optimal matrix

$$P^* = \begin{bmatrix}P_1 & P_2\\ * & *\end{bmatrix} \tag{39}$$

where $P_1 \in \mathbb{R}^{n\times n}$ and $P_2 \in \mathbb{R}^{n\times n}$, and $*$ denotes the useless matrix. Therefore, we have

$$\hat{\tau}_{\text{ext}} = R^{-1}P_1q - R^{-1}P_2q_d. \tag{40}$$

Comparing (40) with the desired admittance model (6), the expected admittance model is obtained to ensure the optimal interaction performance. With the desired trajectory q_d and $\hat{\tau}_{\text{ext}}$ estimated by the observer approach, we can obtain the reference trajectory q_r in joint space.

D. Radial Basis Function Neural Network

Radial basis function NN (RBFNN) has capabilities of approximating any continuous function [40] $f(\theta)$: $R^m \rightarrow R$ as

$$f(\theta) = W^TZ(\theta) + \varepsilon(\theta) \tag{41}$$

where the input vector $\theta \in \Omega_\theta \subset R^m$, weight vector $W = [w_1, w_2, \ldots, w_d] \in R^d$, d denotes the NNs node number $d > 1$, $Z(\theta) = [Z_1(\theta), Z_2(\theta), ..., Z_T(\theta)]^T$, with $Z_i(\theta)$ the basis function usually chosen as Gaussian function as

$$Z_i(\theta) = exp\left[\frac{-\left(\theta - u_i^T\right)(\theta - u_i)}{\eta_i^2}\right],\quad i = 1,\ldots,d \tag{42}$$

where $u_i = [u_{i1}, u_{i2}, \ldots, u_{im}]^T \in R^m$ is the center field and η_i the standard deviation. A continuous function can be approximated by

$$f(\theta) = W^{*T} Z(\theta) + \varepsilon^*(\theta). \tag{43}$$

The W^* is defined as the ideal value of W which minimizes $\varepsilon(\theta)$ for all $\theta \in \Omega_\theta \subset R^m$

$$W^* = \arg \min_{W \subset R^l} \{\sup |f(\theta) - W^T Z(\theta)|\}. \tag{44}$$

In practice, W^* is hard to obtain and should be estimated by $\hat{W}$. The error between the actual weight and its estimation is

$$\tilde{W} = W^* - \hat{W}. \tag{45}$$

E. Controller Design

The inner-loop of the system is to ensure the trajectory of a robot arm can track the reference trajectory. An adaptive neural-based controller is designed to achieve the objective. Considering the dynamics of robot manipulator (4), we define

$$\begin{aligned} s &= \dot{e}_q - \Lambda e_q \\ v &= \dot{q}_r + \Lambda e_q \\ \Lambda &= \text{diag}(\lambda_1, \lambda_2, .., \lambda_n) \end{aligned} \tag{46}$$

where $e_q = q - q_r$ and $\lambda_i > 0$. Substituting (46) into (4), yields

$$M_q(q)\dot{s} + C_q(q, \dot{q})s + G_q(q) + M_q(q)\dot{v} + C_q(q, \dot{q})v = \tau - \tau_{\text{ext}}. \tag{47}$$

The input torque of the controller is

$$\tau = \hat{G} + \hat{M}\dot{v} + \hat{C}v + \hat{\tau}_{\text{ext}} - Ks \tag{48}$$

where $\hat{G}_q(q)$, $\hat{M}_q(q)$, and $\hat{C}_q(q, \dot{q})$ denote the estimation of $G_q(q)$, $M_q(q)$, and $C_q(q, \dot{q})$; $K = \text{diag}(k_1, k_2, \ldots, k_i)$ with $k_i > (1/2)$. Substituting (48) into (47), (47) can be rewritten as

$$\begin{aligned} M_q(q)\dot{s} + C_q(q, \dot{q})s + Ks = &-\left(M - \hat{M}\right)\dot{v} - \left(C - \hat{C}\right)v \\ &- \left(G - \hat{G}\right) + \left(\hat{\tau}_{\text{ext}} - \tau_{\text{ext}}\right). \end{aligned} \tag{49}$$

Using the approximation method, yields

$$\begin{aligned} M_q(q) &= W_M^T Z_M(q) + \varepsilon_M(q) \\ C_q(q, \dot{q}) &= W_C^T Z_C(q, \dot{q}) + \varepsilon_C(q) \\ G_q(q) &= W_G^T Z_G(q) + \varepsilon_G(q) \end{aligned} \tag{50}$$

where W_M, W_C, and W_G are the ideal weight matrices, and $Z_M(q)$, $Z_C(q, \dot{q})$, and $Z_G(q)$ are the basis function matrices. The basis function matrices can be defined as

$$\begin{aligned} Z_M(q) &= \text{diag}\left(Z_q, \ldots, Z_q\right) \\ Z_C(q, \dot{q}) &= \text{diag}\left(\left[Z_q, Z_{\dot{q}}\right]^T, \ldots, \left[Z_q, Z_{\dot{q}}\right]^T\right) \\ Z_G(q) &= \text{diag}\left(Z_q^T, \ldots, Z_q^T\right) \end{aligned} \tag{51}$$

with

$$\begin{aligned} Z_q &= \left[\psi(||q - q_1||), \ldots, \psi(||q - q_n||)\right]^T \\ Z_{\dot{q}} &= \left[\psi(||\dot{q} - \dot{q}_1||), \ldots, \psi(||\dot{q} - \dot{q}_n||)\right]^T. \end{aligned} \tag{52}$$

$\psi(\cdot)$ is defined as the Gaussian function. The estimates of $M(q)$, $C(q, \dot{q})$, and $G(q)$ can be written as

$$\begin{aligned} \hat{M}_q(q) &= \hat{W}_M^T Z_M(q) \\ \hat{C}_q(q, \dot{q}) &= \hat{W}_C^T Z_C(q, \dot{q}) \\ \hat{G}_q(q) &= \hat{W}_G^T Z_G(q). \end{aligned} \tag{53}$$

Substituting (50) and (53) into (49), we have

$$\begin{aligned} M_q(q)\dot{s} + C_q(q, \dot{q})s + Ks = &-\tilde{W}_M^T Z_M \dot{v} - \tilde{W}_C^T Z_C v \\ &- \tilde{W}_G^T Z_G - e_\tau. \end{aligned} \tag{54}$$

$\tilde{W}_{(\cdot)} = W_{(\cdot)}^* - \hat{W}_{(\cdot)}$ and $e_\tau = \tau_{\text{ext}} - \hat{\tau}_{\text{ext}}$. Considering the Lyapunov function

$$V = \frac{1}{2} s^T M s + \frac{1}{2} \text{tr}\left(\tilde{W}_M^T Q_M \tilde{W}_M + \tilde{W}_C^T Q_C \tilde{W}_C + \tilde{W}_G^T Q_G \tilde{W}_G\right) \tag{55}$$

where Q_M, Q_C, and Q_G are positive definite matrices to be set by the designer. Taking derivative along the time, one have

$$\begin{aligned} \dot{V} = &s^T M \dot{s} + \frac{1}{2} s^T \dot{M} s \\ &+ \text{tr}\left(\tilde{W}_M^T Q_M \dot{\tilde{W}}_M + \tilde{W}_C^T Q_C \dot{\tilde{W}}_C + \tilde{W}_G^T Q_G \dot{\tilde{W}}_G\right). \end{aligned} \tag{56}$$

By using the estimation of the weight matrices $\hat{W}_M$, $\hat{W}_C$, and $\hat{W}_G$ to approximate W_M^*, W_C^*, and W_G^*, the errors are expressed as

$$\begin{aligned} W_M^{*T} Z_M(q) - \hat{W}_M^T Z_M(q) &= \tilde{W}_M^T Z_M(q) \\ W_C^{*T} Z_C(q) - \hat{W}_C^T Z_C(q) &= \tilde{W}_C^T Z_C(q) \\ W_G^{*T} Z_G(q) - \hat{W}_G^T Z_G(q) &= \tilde{W}_G^T Z_G(q). \end{aligned} \tag{57}$$

As the ideal weight matrix W^* is a constant vector, we know that

$$\begin{aligned} \dot{\tilde{W}}_M &= \dot{\hat{W}}_M \\ \dot{\tilde{W}}_C &= \dot{\hat{W}}_C \\ \dot{\tilde{W}}_G &= \dot{\hat{W}}_G. \end{aligned} \tag{58}$$

Considering Property 2 and (54), we have

$$\begin{aligned} \dot{V} &= s^T M \dot{s} + s^T C \dot{s} \\ &\quad + \text{tr}\left(\tilde{W}_M^T Q_M \dot{\tilde{W}}_M + \tilde{W}_C^T Q_C \dot{\tilde{W}}_C + \tilde{W}_G^T Q_G \dot{\tilde{W}}_G\right) \\ &= -s^T \left(Ks + \tilde{W}_M^T Z_M \dot{v} + \tilde{W}_C^T Z_C v + \tilde{W}_G^T Z_G + e_\tau\right) \\ &\quad + \text{tr}\left(\tilde{W}_M^T Q_M \dot{\tilde{W}}_M + \tilde{W}_C^T Q_C \dot{\tilde{W}}_C + \tilde{W}_G^T Q_G \dot{\tilde{W}}_G\right) \\ &= -s^T K s - s^T e_\tau \\ &\quad - \text{tr}\left[\tilde{W}_M^T \left(Z_M \dot{v} s^T + Q_M \dot{\hat{W}}_M\right)\right] \\ &\quad - \text{tr}\left[\tilde{W}_C^T \left(Z_C v s^T + Q_C \dot{\hat{W}}_C\right)\right] \\ &\quad - \text{tr}\left[\tilde{W}_C^T \left(Z_G s^T + Q_G \dot{\hat{W}}_G\right)\right]. \end{aligned} \tag{59}$$

The update law is designed as

$$\begin{aligned} \dot{\hat{W}}_M &= -Q_M^{-1}\left(Z_M \dot{v} s^T + \sigma_M \hat{W}_M\right) \\ \dot{\hat{W}}_C &= -Q_C^{-1}\left(Z_C v s^T + \sigma_C \hat{W}_C\right) \\ \dot{\hat{W}}_G &= -Q_G^{-1}\left(Z_G s^T + \sigma_G \hat{W}_G\right) \end{aligned} \tag{60}$$

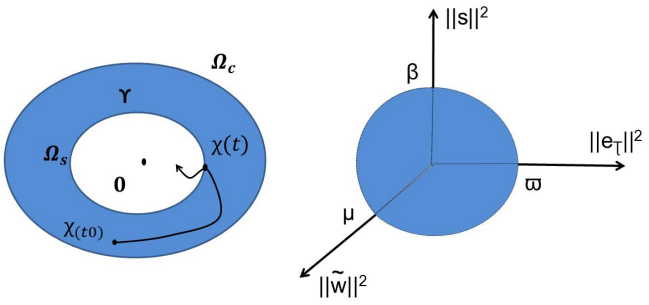


Fig. 4. Representation of UUB and the set Ω_s defined in (68).

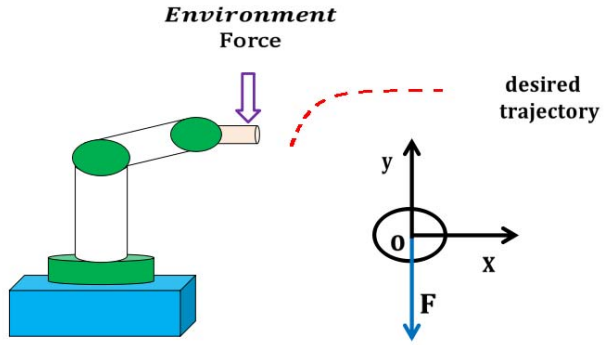


Fig. 5. Overview of the scenario.

where σ_M, σ_G, and σ_C are constants specified by the designer. Substituting (60) into (59), the derivative of V is

$$\dot{V} = -s^T K s - s^T e_\tau + \mathrm{tr}\Big[\sigma_M \tilde{W}_M^T \hat{W}_M\Big] + \mathrm{tr}\Big[\sigma_C \tilde{W}_C^T \hat{W}_C\Big] + \mathrm{tr}\Big[\sigma_G \tilde{W}_G^T \hat{W}_G\Big]. \tag{61}$$

Using Young's inequality [41]

$$\mathrm{tr}\Big[\tilde{W}_{(\cdot)}^T \hat{W}_{(\cdot)}\Big] \le -\frac{1}{2}||\tilde{W}_{(\cdot)}||_F^2 + \frac{1}{2}||W_{(\cdot)}^*||_F^2$$
$$-s^T e_\tau \le \frac{1}{2}s^T s + \frac{1}{2}e_\tau^T e_\tau. \tag{62}$$

Then (61) can be written as

$$\dot{V} \le -s^T K s + \frac{1}{2}||s||^2 + \frac{1}{2}||e_\tau||^2 + \alpha - \frac{\sigma_M}{2}||\tilde{W}_M||^2 - \frac{\sigma_C}{2}||\tilde{W}_C||^2 - \frac{\sigma_G}{2}||\tilde{W}_G||^2 \tag{63}$$

where $\alpha = (\sigma_M/2)||W_M^*||^2 + (\sigma_C/2)||W_C^*||^2 + (\sigma_G/2)||W_G^*||^2$. A sufficient condition for $\dot{V} \le 0$ is that

$$\alpha \le s^T\left(K - \frac{1}{2}I\right)s + \frac{1}{2}||e_\tau||^2 + \frac{\sigma_M}{2}||\tilde{W}_M||^2 + \frac{\sigma_C}{2}||\tilde{W}_C||^2 + \frac{\sigma_G}{2}||\tilde{W}_G||^2 \tag{64}$$

where I is an unit matrix. Let χ denote the state variable comprised of $e_\tau, s, \tilde{W}_M, \tilde{W}_C$, and $\tilde{W}_G$ defined in the Lyapunov function candidate, and it follows from (64) that:

$$\dot{V}(\chi) < 0, \quad \forall ||\chi|| > \varrho \tag{65}$$

where ϱ is a positive constant. In other words, the time derivative of $V(\chi)$ is negative outside the set $\Omega_s = \{||\chi|| \le \varrho\}$, which is defined of Theorem 1 as below, or equivalently, all $\chi(t)$ that start outside Ω_s will enter the set within a limited period of time, and not escape afterwards. Choose $0 < V(\chi) < \epsilon < c$, where ϵ and c are small constants. Denote that the sets $\Omega_s = \{V(\chi) \le \epsilon\}$ and $\Omega_c = \{V(\chi) \le c\}$. Let

$$\Upsilon = \{\epsilon \le V(\chi) \le c\} = \Omega_c - \Omega_s. \tag{66}$$

Note that derivative of $V(\chi)$ along time is negative inside Υ, that is

$$\dot{V}(\chi(t)) < 0, \quad \forall \chi \in \Upsilon, \quad \forall t \ge t_0. \tag{67}$$

Since $\dot{V}$ is negative in $\Upsilon = \{\epsilon \le V(\chi) \le c\}$, which means that in this set $V(\chi(t))$ will decrease monotonically in time until the solution enters the set $\{V(\chi) \le \epsilon\}$. From that time on, $\chi(t)$ cannot leave the set because $\dot{V}$ is negative on its boundary $V(\chi) = \epsilon$. A sketch of the sets is shown in Fig. 4.

The conclusion of convergence analysis can be concluded in Theorem 1.

Theorem 1: Using the uniformly ultimately bounded (UUB), the tracking error s, estimation error e_τ and weight errors $\tilde{W}_{(\cdot)}$ will fall into an set Ω_s, where the bounding set Ω_s is defined in (68) and shown in Fig. 4

$$\Omega_s = \Bigg\{(||\tilde{W}_M||, ||\tilde{W}_C||, ||\tilde{W}_G||, ||e_\tau||, ||s||), \frac{\sigma_M||\tilde{W}_M||^2}{2\alpha} + \frac{\sigma_C||\tilde{W}_C||^2}{2\alpha} + \frac{\sigma_G||\tilde{W}_G||^2}{2\alpha} + \frac{s^T\left(K - \frac{1}{2}I\right)s}{\alpha} + \frac{||e_\tau||^2}{\alpha} \le 1\Bigg\}. \tag{68}$$

As shown in Fig. 4, the bounding set Ω_s is the area in the first quadrant, passing through the points $(||(\sigma_{(\cdot)}/2)\tilde{W}_{(\cdot)}||^2 = \alpha, ||e_\tau||^2 = 0, ||s||^2 = 0)$, $(s^T(K - (1/2)I)s = \alpha, ||\tilde{W}_{(\cdot)}||^2 = 0, ||e_\tau||^2 = 0)$ and $((1/2)||e_\tau||^2 = \alpha, ||\tilde{W}_{(\cdot)}||^2 = 0, ||s||^2 = 0)$. In Fig. 4, we define

$$\text{when } ||\frac{\sigma_{(\cdot)}}{2}\tilde{W}_{(\cdot)}||^2 = \alpha, \ \tilde{W} = \varpi$$
$$\text{when } s^T\left(K - \frac{1}{2}I\right)s = \alpha, \ s = \beta$$
$$\text{when } \frac{1}{2}||e_\tau||^2 = \alpha, \ e_\tau = \mu. \tag{69}$$

Since $||W_{(\cdot)}^*||$ is a bounded constant, $||\hat{W}_{(\cdot)}|| = ||W_{(\cdot)}^* - \tilde{W}_{(\cdot)}||$ is also bounded. Since $||e_\tau||$ is bounded, $||\hat{\tau}_{\text{ext}}|| = ||e_\tau + \tau_{\text{ext}}||$ is bounded. With the bounded signals q_r and $\dot{q}_r$, according to (46), $||v||$ is bounded and $||q|| = ||e_q + q_r||$ is bounded as well. Therefore, the norm of state variable χ is bounded.

IV. Simulation Studies

In this section, a 2-link robot arm is in physical interaction with the unknown environment. Fig. 5 shows the simulation environment. The environment torque is exerted at the end-effector only in the Y-direction. Table I presents the parameters of the robot arm. The desired trajectory will be modified to adjust the interaction torque exerted by the environment so that a compliant behavior is achieved. The controller is to guarantee the robot arm to track a reference trajectory q_r.

TABLE I
PARAMETERS OF THE ROBOT ARM

Parameters	Description	Value
m_1	Mass of link 1	2.0 kg
m_2	Mass of link 2	2.0 kg
l_1	Length of link 1	0.20 m
l_2	Length of link 2	0.20 m
I_1	Inertia moment of link 1	0.027 kgm^2
I_1	Inertia moment of link 2	0.027 kgm^2

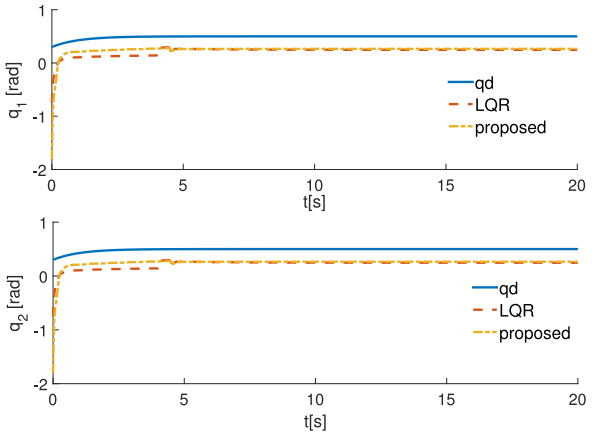


Fig. 6. Desired trajectory q_d and reference trajectory q_r with $Q = 1$ and $R = 1$.

A. Interaction Performance

It is assumed that the dynamics of the environment can be defined as: $0.01\dot{q} + (q - 0.3) = -\tau_{\text{ext}}$. The LQR method is employed to verify the proposed method. Solutions of the ARE are used to gain optimized parameters of the admittance model with known matrices A and B. The proposed method is adopted when the environment model is not available.

In the first step, the initial values of variables of the system should be properly set. The selection of exploration noise is not a trivial task which is used to make estimated parameters converge to the real values. In this paper, the exploration noise is selected as

$$\phi = \sum_{w=1}^{8} \frac{0.04}{w} \sin(wt). \tag{70}$$

The initial feedback gain K_0 should ensure the stability of the system, which is set as $K_0 = [-1, 0.1]$. The initial P_k is set as $P_0 = 10I_m$. Then, the second step is conducted and stops until $||P_k|| < 0.02$, where $||\cdot||$ denotes the Euclidean norm.

Simulation results are presented in Figs. 6–9. The weights of the cost function are given by $Q = 1$ and $R = 1$. Based on LQR method, the desired admittance model is obtained as $\hat{\tau}_{\text{ext}} = -0.4142\dot{q} + 0.0702q_r$ with the known A and B. In Fig. 6, the desired trajectories and reference trajectories of the robot arm based on the LQR and proposed method in joint space are shown. At the beginning, a large error exists in trajectory between the LQR and the proposed method, due to the initial admittance model: $\hat{\tau}_{\text{ext}} = -\dot{q} + 0.1(q - 0.3)$ and the exploration noise. After that, the error is becoming smaller and trajectory of the robot arm with proposed method is coming closely to the trajectory of the robot arm with LQR method. Fig. 7 illustrates the error of admittance parameters between LQR and proposed method, where the error is defined as $||K_k - K^*||$. The convergence of P_k and K_k to optimal values based on LQR are illustrated. It is obvious that the error decreases to around 0.01 after 12 iterations and the Euclidean norm of P_k decreases to around 0.02. After three steps, the admittance model with the proposed method is obtained as: $\hat{\tau}_{\text{ext}} = -0.4173\dot{q} + 0.00904q_r$. The symmetric positive matrix with $Q = 1$ and $R = 1$, P_k from the proposed algorithm and

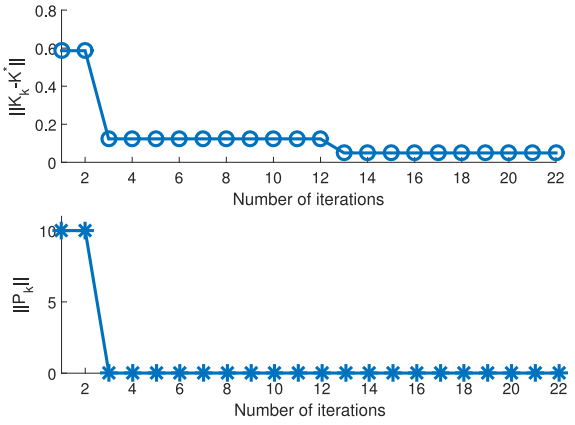


Fig. 7. Convergence of P_k and K_k to their optimal values with $Q = 1$ and $R = 1$.

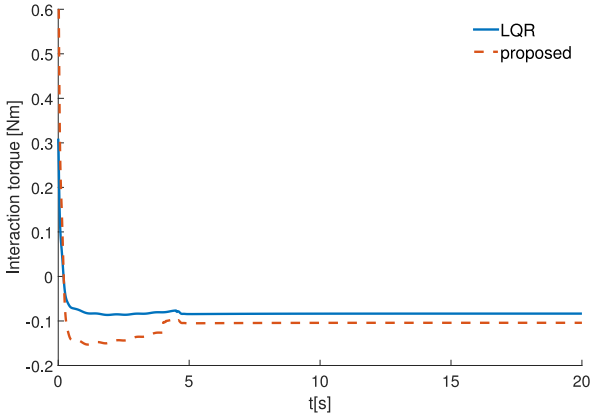


Fig. 8. Interaction torque with $Q = 1$ and $R = 1$.

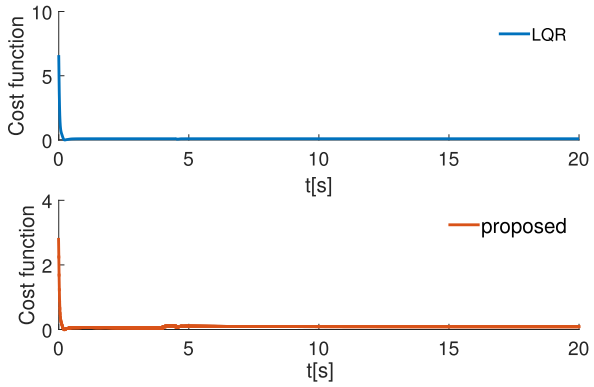


Fig. 9. Value of cost function with $Q = 1$ and $R = 1$.

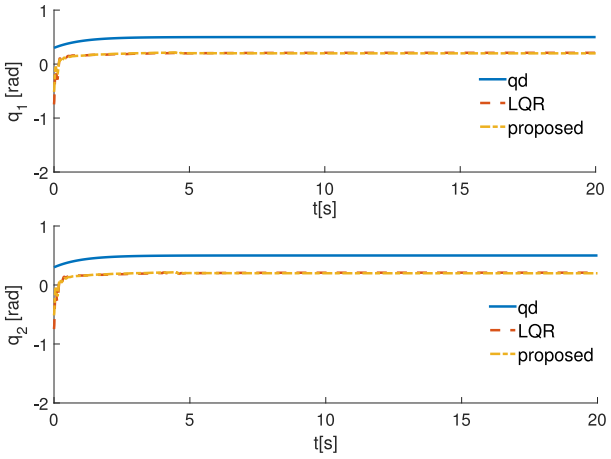


Fig. 10. Desired trajectory q_d and reference trajectory q_r with $Q = 5$ and $R = 1$.

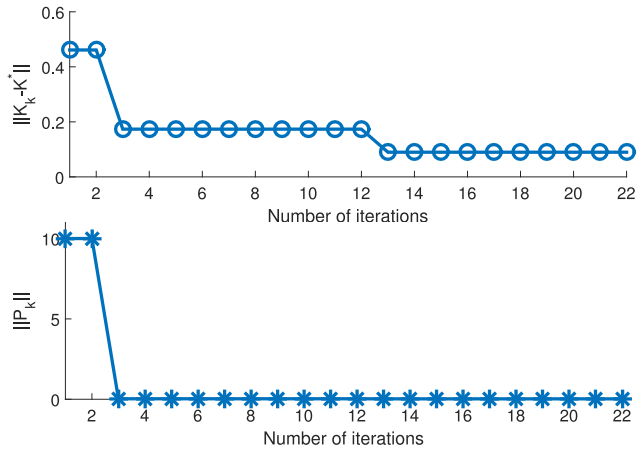


Fig. 11. Convergence of P_k and K_k to their optimal values with $Q = 5$ and $R = 1$.

P^* from LQR are shown as follows:

$$P_k = \begin{bmatrix} 0.0033 & 0.0027 \\ 0.0027 & 0.0077 \end{bmatrix} P^* = \begin{bmatrix} 0.0041 & -0.0007 \\ -0.0007 & 0.0025 \end{bmatrix}. \tag{71}$$

To further verify the correctness of the proposed method, different weights of the cost function are given by $Q = 5$ and $R = 1$, and results are presented in Figs. 10–13. Similarly, the desired admittance model is: $\hat{\tau}_{\text{ext}} = -1.4495\dot{q} + 0.2033q_r$ based on the known A and B. After three steps of iteration, the admittance model is obtained as $\hat{\tau}_{\text{ext}} = -1.5374\dot{q} + 0.2063q_r$ with the proposed method. The desired trajectory q_d and reference trajectory q_r with $Q = 5$ and $R = 1$ are illustrated in Fig. 10. As depicted in Fig. 11, the error $||K_k - K^*||$ between the LQR and proposed method converges to 0.08 after 12 iterations. The interaction torque converges to around -0.35 Nm at 7 s, shown in Fig. 12. The values of cost function are 0.561 and 0.602 with LQR and the proposed method, respectively. Similarly, the symmetric positive matrix with $Q = 5$ and $R = 1$, P_k from the proposed algorithm and P^* from LQR are shown as follows:

$$P_k = \begin{bmatrix} 0.0106 & 0.0068 \\ 0.0068 & 0.0107 \end{bmatrix} \quad P^* = \begin{bmatrix} 0.0145 & -0.0020 \\ -0.0020 & 0.0043 \end{bmatrix}. \tag{72}$$

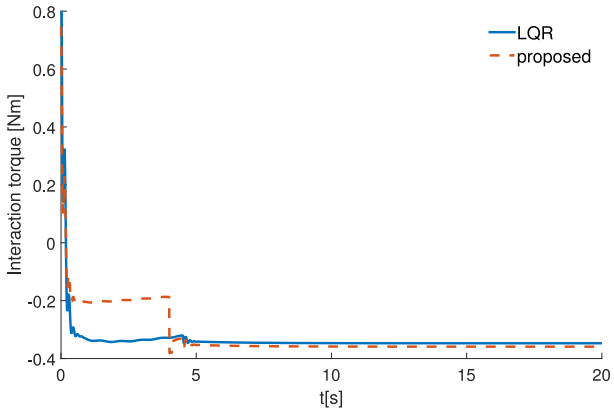


Fig. 12. Interaction torque with $Q = 5$ and $R = 1$.

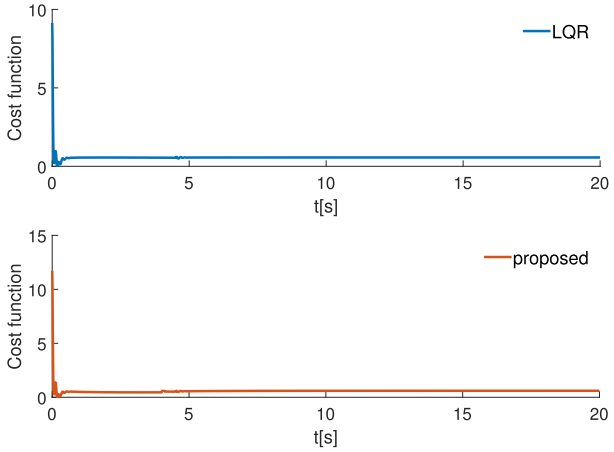


Fig. 13. Value of cost function with $Q = 5$ and $R = 1$.

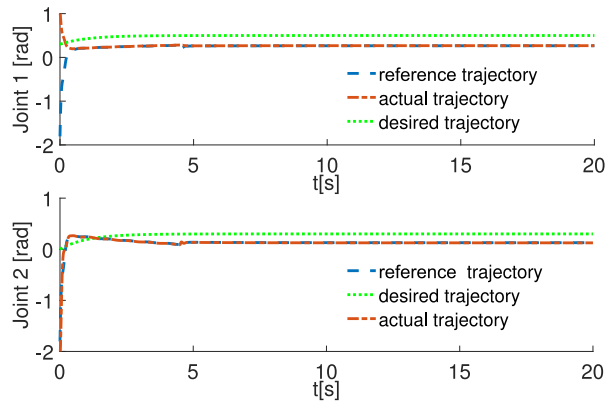


Fig. 14. Tracking performance of each joint.

Comparing the admittance model obtained by the proposed method and that with LQR method, the difference is small and acceptable. The overall results are satisfactory.

B. Trajectory Tracking

The initial configurations of two joints are: $q(0) = [2.6, -2.6]^T$ and the predesigned trajectory q_d is: $q_d = [0.3 + 0.2e^{-t}, 0.3e^{-t}]^T$. The weight matrices are initialized as $\hat{W}_M(0) = \mathbf{0}$, $\hat{W}_C(0) = \mathbf{0}$, and $\hat{W}_G(0) = \mathbf{0}$. The results of tracking performance are presented in Figs. 14–20. In Fig. 14,

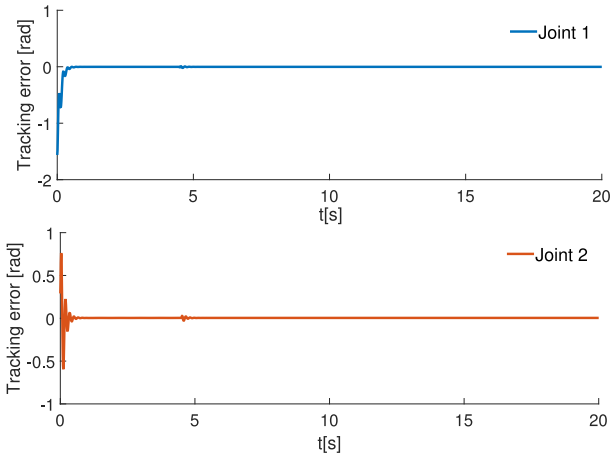


Fig. 15. Tracking error of each joint.

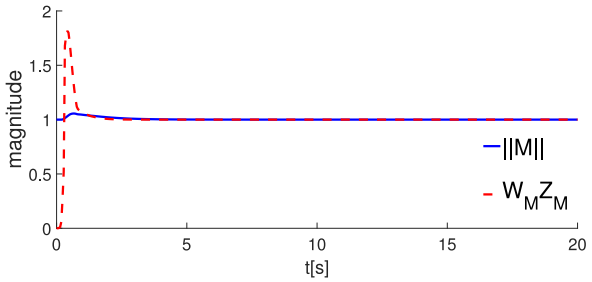


Fig. 16. Function approximation performance of the $M(q)$ matrix by RBFNN.

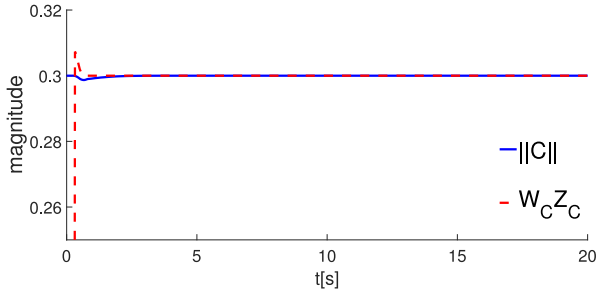


Fig. 17. Function approximation performance of the $C(q, \dot{q})$ matrix by RBFNN.

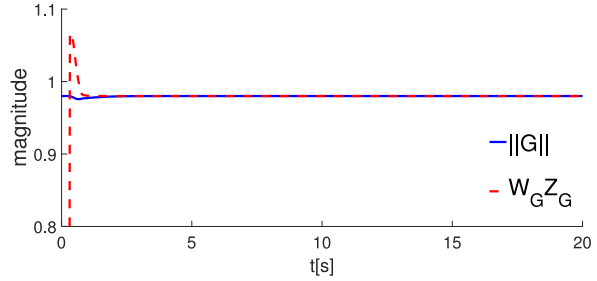


Fig. 18. Function approximation performance of the $G(q)$ matrix by RBFNN.

the actual trajectory can track the reference trajectory, and the tracking errors can converge near to zero, as shown in Fig. 15. The function approximation performance of the our implemented NN is depicted in Figs. 16–18, where $W_{(\cdot)}$ and $Z_{(\cdot)}$ denote the weight matrix and basis function matrix of RBFNN, respectively. As seen from the figures, we can see that the

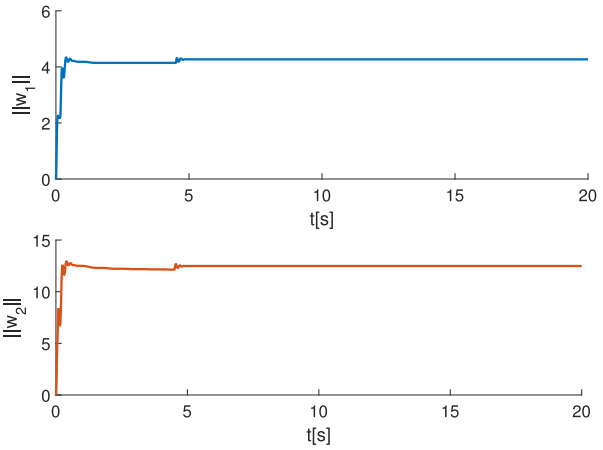


Fig. 19. Norm of weights of the NN for each joint.

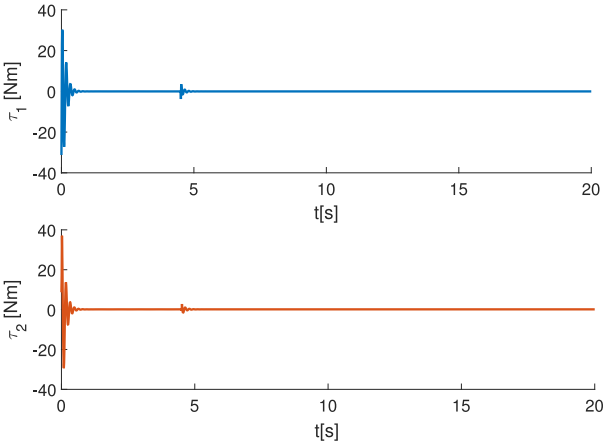


Fig. 20. Control inputs τ_1 and τ_2.

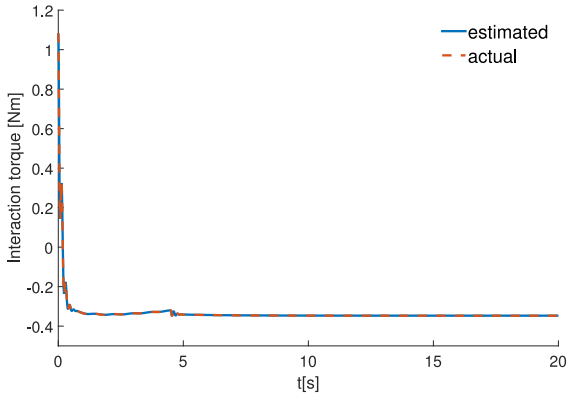


Fig. 21. Estimated interaction torque obtained from the observer based on generalized momentum approach and actual interaction torque.

output of the implemented NN could follow the approximated nonlinear function's dynamics $(M(q), C(q, \dot{q}), G(q))$, which implies that the implemented NNs can successfully approximate the nonlinear function with satisfactory tracking performance. As shown in Fig. 19, the weight matrices of the system shows a trend of convergence. The control input is shown in Fig. 20. With the tracking errors and toque regulation, the cost function is minimized, as shown in Fig. 9.

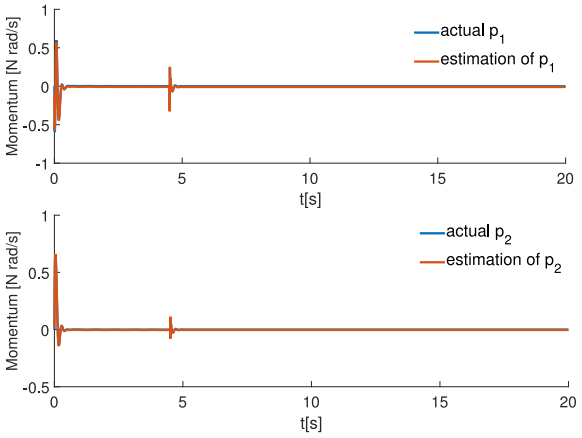


Fig. 22. Estimated generalized momentum of the joints obtained from the torque observer and actual generalized momentum.

C. Torque Observer

The interaction torque is assumed to be applied at the end-effector of the robot arm. In Fig. 21, the actual interaction torque and its estimation are shown, and the estimated results of the generalized momentum of each joint are shown in Fig. 22. As seen from the figures, the output of the implemented torque observer could follow the real torque which implies that the implemented observer have a satisfactory estimation performance.

V. Conclusion

In this paper, a method of admittance adaptation is proposed for robot–environment interaction. The desired admittance model is obtained by optimal adaptive control approach and admittance control is used to regulate the interaction behavior. An NN-based controller is developed to guarantee trajectory tracking and interaction torque is estimated by an observer approach. A cost function that includes tracking errors and interaction torque is minimized. Simulation studies have verified the effectiveness of our proposed method.

References

[1] S. Katsura and K. Ohnishi, "Human cooperative wheelchair for haptic interaction based on dual compliance control," *IEEE Trans. Ind. Electron.*, vol. 51, no. 1, pp. 221–228, Feb. 2004.

[2] H. Wang *et al.*, "Visual servoing of soft robot manipulator in constrained environments with an adaptive controller," *IEEE/ASME Trans. Mechatron.*, vol. 22, no. 1, pp. 41–50, Feb. 2017.

[3] J. Huang, W. Huo, W. Xu, S. Mohammed, and Y. Amirat, "Control of upper-limb power-assist exoskeleton using a human-robot interface based on motion intention recognition," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1257–1270, Oct. 2015.

[4] N. Hogan, "Impedance control: An approach to manipulation: Part II—Implementation," *J. Dyn. Syst. Meas. Control*, vol. 107, no. 1, pp. 8–16, 1985.

[5] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 3, pp. 334–344, Mar. 2016.

[6] M. T. Mason, "Compliance and force control for computer controlled manipulators," *IEEE Trans. Syst., Man, Cybern.*, vol. 11, no. 6, pp. 418–432, Jun. 1981.

[7] A. M. Smith *et al.*, "Novel hybrid adaptive controller for manipulation in complex perturbation environments," *PLoS ONE*, vol. 10, no. 6, 2015, Art. no. e0129281.

[8] H. Wang *et al.*, "Eye-in-hand tracking control of a free-floating space manipulator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 4, pp. 1855–1865, Aug. 2017.

[9] W. He and S. Zhang, "Control design for nonlinear flexible wings of a robotic aircraft," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 1, pp. 351–357, Jan. 2017.

[10] W. He, Z. Yan, C. Sun, and Y. Chen, "Adaptive neural network control of a flapping wing micro aerial vehicle with disturbance observer," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3452–3465, Oct. 2017.

[11] J. Huang, P. Di, T. Fukuda, and T. Matsuno, "Robust model-based online fault detection for mating process of electric connectors in robotic wiring harness assembly systems," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 5, pp. 1207–1215, Sep. 2010.

[12] C. Yang, X. Wang, Z. Li, Y. Li, and C.-Y. Su, "Teleoperation control based on combination of wave variable and neural networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 8, pp. 2125–2136, Aug. 2017.

[13] C. Yang, K. Huang, H. Cheng, Y. Li, and C.-Y. Su, "Haptic identification by ELM-controlled uncertain manipulator," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 8, pp. 2398–2409, Aug. 2017.

[14] A. Alcocera, A. Robertssona, A. Valerac, and R. Johanssona, "Force estimation and control in robot manipulators," in *Proc. Vol. 7th IFAC Symp. Robot Control*, Wrocław, Poland, 2004, p. 55.

[15] A. De Luca and R. Mattone, "Sensorless robot collision detection and hybrid force/motion control," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2005, pp. 999–1004.

[16] A. De Luca, A. Albu-Schaffer, S. Haddadin, and G. Hirzinger, "Collision detection and safe reaction with the DLR-III lightweight manipulator arm," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2006, pp. 1623–1630.

[17] M. Cohen and T. Flash, "Learning impedance parameters for robot control using an associative search network," *IEEE Trans. Robot. Autom.*, vol. 7, no. 3, pp. 382–390, Jun. 1991.

[18] T. Tsuji, K. Ito, and P. G. Morasso, "Neural network learning of robot arm impedance in operational space," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 26, no. 2, pp. 290–298, Apr. 1996.

[19] C. Yang *et al.*, "Interface design of a physical human-robot interaction system for human impedance adaptive skill transfer," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 1, pp. 329–340, Jan. 2018.

[20] R. Z. Stanisic and Á. V. Fernández, "Adjusting the parameters of the mechanical impedance for velocity, impact and force control," *Robotica*, vol. 30, no. 4, pp. 583–597, 2012.

[21] S. S. Ge, Y. Li, and C. Wang, "Impedance adaptation for optimal robot–environment interaction," *Int. J. Control*, vol. 87, no. 2, pp. 249–263, 2014.

[22] R. Johansson and M. W. Spong, "Quadratic optimization of impedance control," in *Proc. IEEE Int. Conf. Robot. Autom.*, San Diego, CA, USA, 1994, pp. 616–621.

[23] M. Matinfar and K. Hashtrudi-Zaad, "Optimization-based robot compliance control: Geometric and linear quadratic approaches," *Int. J. Robot. Res.*, vol. 24, no. 8, pp. 645–656, 2005.

[24] D. Kirk, *Optimal Control Theory: An Introduction*. Mineola, NY, USA: Dover, 2004.

[25] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Belmont, MA, USA: Athena Sci., 1995.

[26] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits Syst. Mag.*, vol. 9, no. 3, pp. 32–50, Sep. 2009.

[27] F.-Y. Wang, H. Zhang, and D. Liu, "Adaptive dynamic programming: An introduction," *IEEE Comput. Intell. Mag.*, vol. 4, no. 2, pp. 39–47, May 2009.

[28] P. J. Werbos, *A Menu of Designs for Reinforcement Learning Over Time*. Cambridge, MA, USA: MIT Press, 1991.

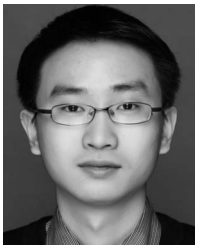
[29] D. A. White and D. A. Sofge, *Handbook of Intelligent Control: Neural, Fuzzy and Adaptive Approaches*. New York, NY, USA: IEEE Press, 1994.

[30] B. Kim, J. Park, S. Park, and S. Kang, "Impedance learning for robotic contact tasks using natural actor-critic algorithm," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 2, pp. 433–443, Apr. 2010.

[31] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *Int. J. Robot. Res.*, vol. 30, no. 7, pp. 820–833, 2011.

[32] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of dynamic systems by learning: A new control theory for servomechanism or mechatronics systems," in *Proc. 23rd IEEE Conf. Decis. Control*, Las Vegas, NV, USA, 1984, pp. 1064–1069.

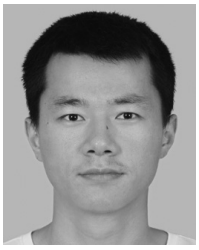
[33] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.
[34] B. Siciliano and O. Khatib, *Handbook of Robotics*. Berlin, Germany: Springer-Verlag, 2016.
[35] C. Yang, Z. Li, and J. Li, "Trajectory planning and optimized adaptive control for a class of wheeled inverted pendulum vehicle models," *IEEE Trans. Cybern.*, vol. 43, no. 1, pp. 24–36, Feb. 2013.
[36] A. Wahrburg, E. Morara, G. Cesari, B. Matthias, and H. Ding, "Cartesian contact force estimation for robotic manipulators using Kalman filters and the generalized momentum," in *Proc. IEEE Int. Conf. Autom. Sci. Eng.*, 2015, pp. 1230–1235.
[37] M. Capurso, M. M. G. Ardakani, R. Johansson, A. Robertsson, and P. Rocco, "Sensorless kinesthetic teaching of robotic manipulators assisted by observer-based force control," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2017, pp. 945–950.
[38] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, vol. 1. New York, NY, USA: Wiley, 1972,
[39] F. L. Lewis and V. L. Syrmos, *Optimal Control*. New York, NY, USA: Wiley, 1995.
[40] T. H. Lee and C. J. Harris, *Adaptive Neural Network Control of Robotic Manipulators*, vol. 19. Singapore: World Sci., 1998.
[41] W. H. Young, "On classes of summable functions and their Fourier series," *Proc. Royal Soc. A Math. Phys. Eng. Sci.*, vol. 87, no. 594, pp. 225–229, 1912.



Chenguang Yang (M'10–SM'16) received the B.Eng. degree in measurement and control from Northwestern Polytechnical University, Xi'an, China, in 2005, and the Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 2010.

He was a Postdoctoral Research Fellow with Imperial College London, London, U.K. His current research interests include robotics and automation.

Dr. Yang was a recipient of the Best Paper Award from the IEEE TRANSACTIONS ON ROBOTICS and a number of international conferences.

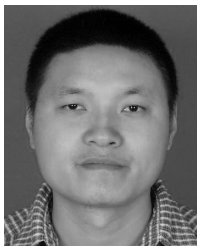


Guangzhu Peng received the B.Eng. degree in automation from Yangtze University, Jingzhou, China, in 2014. He is currently pursuing the M.Eng. degree with the College of Automation Science and Engineering, South China University of Technology, Guangzhou, China.

His current research interests include robotics and intelligent control, etc.

Yanan Li (S'10–M'14) received the B.Eng. and M.Eng. degrees in automation from the Harbin Institute of Technology, Harbin, China, in 2006 and 2008, respectively, and the Ph.D. degree in control engineering from the National University of Singapore, Singapore, in 2013.

He is a Lecturer of Control Engineering with the Department of Engineering and Design, University of Sussex, Brighton, U.K. From 2015 to 2017, he was a Research Associate with the Department of Bioengineering, Imperial College London, London, U.K. From 2013 to 2015, he was a Research Scientist with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. His current research interests include human–robot interaction, assistive robotics, human motor control, and control theory and applications.



Rongxin Cui (M'09) received the B.Eng. degree in automatic control and the Ph.D. degree in control science and engineering from Northwestern Polytechnical University, Xi'an, China, in 2003 and 2008, respectively.

From 2008 to 2010, he was a Research Fellow with the Centre for Offshore Research and Engineering, National University of Singapore, Singapore. He is currently a Professor with the School of Marine Science and Technology, Northwestern Polytechnical University. His current research interests include control of nonlinear systems, cooperative path planning and control for multiple robots, control and navigation for underwater vehicles, and system development.

Dr. Cui servers as an Editor for the *Journal of Intelligent and Robotic Systems*.

Long Cheng (SM'14) received the B.S. degree (Hons.) in control engineering from Nankai University, Tianjin, China, in 2004, and the Ph.D. degree (Hons.) in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2009.

In 2010, he was a Postdoctoral Research Fellow with the Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK, Canada, for eight months. From 2010 to 2011, he was a Postdoctoral Research Fellow with the Mechanical and Industrial Engineering Department, Northeastern University, Boston, MA, USA. From 2013 to 2014, he was a Visiting Scholar with the Electrical and Computer Engineering Department, University of California at Riverside, Riverside, CA, USA. He is currently a Professor with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. He has published over 50 technical papers in peer-refereed journals and prestigious conference proceedings. His current research interests include intelligent control of smart materials, coordination of multiagent systems, and neural networks and their applications to robotics.

Dr. Cheng is an Editorial Board Member of *Neurocomputing*, *Neural Processing Letters*, and the *International Journal of Systems Science*.

Zhijun Li (M'07–SM'09) received the Ph.D. degree in mechatronics from Shanghai Jiao Tong University, Shanghai, China, in 2002.

From 2003 to 2005, he was a Postdoctoral Fellow with the Department of Mechanical Engineering and Intelligent Systems, University of Electro-Communications, Tokyo, Japan. From 2005 to 2006, he was a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, and Nanyang Technological University, Singapore. Since 2012, he has been a Professor with the College of Automation Science and Engineering, South China University of Technology, Guangzhou, China. His current research interests include service robotics, teleoperation systems, nonlinear control, and neural-network optimization.

Dr. Li is serving as an Editor-at-Large for the *Journal of Intelligent and Robotic Systems*, and an Associate Editor for several IEEE TRANSACTIONS. Since 2016, he has been the Co-Chair of the Technical Committee on Biomechatronics and Biorobotics Systems; IEEE Systems, Man and Cybernetics Society; and the Technical Committee on Neuro-Robotics Systems, IEEE Robotics and Automation Society. He was the General Chair and the Program Chair of 2016 and 2017 IEEE Conference on Advanced Robotics and Mechatronics.