Abstract—In this paper, an admittance adaptation method for robots to interact with unknown environment is developed. The environment to be interact with is modelled as a linear system in the state-space form. In the presence of the unknown environment dynamics, an observer in robot joint space is employed to estimate the interaction torque. Admittance control is adopted to regulate the dynamic behavior at the interaction point when the robot interacts with unknown environment. An adaptive neural controller using radial basis function is employed to guarantee trajectory tracking. A cost function is defined to achieve the interaction performance of torque regulation and trajectory tracking, which is minimised by adaptation of the admittance model. Simulation studies on a robot manipulator are carried out to verify the effectiveness of the proposed method.

Index Terms—optimal adaptive control; robot—environment interaction; observer; neural networks (NNs); admittance control

I. INTRODUCTION

With the development of robot technology, robots have been widely used in various fields such as education, industry and entertainment, etc. In these applications, robots are required to interact with external environment [1]–[3]. Therefore, robots interacting with the environment has received great attention and much effort has been made on this topic. Although it has been investigated for more than decades, there are still many open problems not 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 a compliant behavior, there are three approaches that are widely applied: admittance control, hybrid position/force control and impedance control.

The concept of impedance control introduced by Hogan has been a classical control method in robotics [4]. The aim of impedance control is to develop a relationship between the interaction force and the position of the robot. 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 appears 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 environment force and a desired admittance model, a virtual desired trajectory is obtained and tracked. Then, the compliant behavior is realized by trajectory adaptation. Traditional control method of a robot manipulator is model-based control, which usually has a good control performance [7]. However, this method heavily depends on the accuracy of a robot model which cannot be guaranteed in many cases. Therefore, adaptive control methods have been widely studied and applied to practical systems [8]. These methods can approximate uncertainties of a system by using tools such as neural network (NN), wavelet network, and fuzzy logic system, etc [9]–[11]. Another key element in admittance control system is the force sensor. Force sensors are regarded as a media for communication between a robot and environment. However, force sensors equipped on the manipulators may cause inconvenience and are usually costly. Due to these reasons, sensorless control schemes have been received great attention. There are two main methods for estimating the external force: disturbance observer approach and force observer approach based on knowledge of motor torques. In [12], the disturbance observer approach with knowledge of joint angle has been analyzed. In [13], a force observer for collision detection based on the generalized momentum has been introduced. In [14], an collision detection method is first developed for rigid robot arms and other robots with elastic joints.

Under impedance/admittance control, robots are governed to be compliant to interaction force exerted by the environment. If the environment is passive, a passive impedance/admittance model is imposed to the robot for safety. However, obtaining desired impedance/admittance model is not a easy work 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. It has been generally acknowledged that such an ability of improving performance by repeating a task is an important control strategy and has been widely studied. In [15], an associative search network (ASN) learning scheme is presented for learning control parameters for robot to complete a wall-following task. In [16], neural networks based method is applied to regulate impedance parameters of the end-effector of a robot. However, the disadvantage of the learning method is that it requires a robot to repeat operations to learn the desired impedance parameters which may cause inconvenience in many situations. Therefore, the impedance adaptation method has been widely studied [17]. In [18], strategies of switching
among different values of impedance parameters are developed for dissipating the energy of the system. In [19], the impedance adaptation is investigated for robots to interact with unknown environments.

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. There has been much research effort in literatures. 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 a minimal cost. In [20], the LQR is used to determine desired impedance parameters with the environmental dynamics known. In [21], a target impedance is adjusted by online solutions of the defined LQR problem based on environment stiffness and damping. Better than the fixed impedance parameters obtained from 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 [22], the solution of a Riccati equation could be difficult to find with the unknown dynamics of a environment. Therefore, when the dynamics of a environment is unknown, approaches proposed above may not be used. To copy with this problem, adaptive dynamic programming (ADP) has received much attention and been widely studied [23]–[27]. ADP is a very useful tool in solving optimization and optimal control problems. Based on the idea of ADP, a control action is modified based on the feedback information of a environment. There are many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown interaction performance including the trajectory tracking error is considered to be a damping-stiffness model, which is a linear system in the state-space form. An observer based on the feedback information of a environment 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 many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown interaction performance including the trajectory tracking error is considered to be a damping-stiffness model, which is a linear system in the state-space form. An observer based on the feedback information of a environment 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 many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown interaction performance including the trajectory tracking error is considered to be a damping-stiffness model, which is a linear system in the state-space form. An observer based on the feedback information of a environment 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 many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown interaction performance including the trajectory tracking error is considered to be a damping-stiffness model, which is a linear system in the state-space form. An observer based on the feedback information of a environment 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 many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown interaction performance including the trajectory tracking error is considered to be a damping-stiffness model, which is a linear system in the state-space form. An observer based on the feedback information of a environment 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 many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown interaction performance including the trajectory tracking error is considered to be a damping-stiffness model, which is a linear system in the state-space form. An observer based on the feedback information of a environment 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 many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedance/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.
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 obtain admittance parameters subject to an unknown environment based on the adaptive optimal control method. With the interaction torque estimated from an observer, the virtual desired trajectory \( q_r \) in the joint space is generated. The inner-loop of the system is to guarantee the trajectory tracking with adaptive control scheme.

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

\[ f_{\text{ext}} = f(x_r, x_d) \]  

(6)

where \( x_r \in \mathbb{R}^n \) is the virtual desired trajectory in the Cartesian space and \( f(\cdot) \) is the function of the admittance model. To be specific, a target admittance model in joint space is described as below

\[ M_d \dot{J}(q)(\dot{q}_r - \dot{q}_d) + (M_d \ddot{J}(q) + C_d \dot{J}(q))(\dot{q}_r - \dot{q}_d) + G_d(\kappa(q_r) - \kappa(q_d)) = -J(q)T \tau_{\text{ext}} \]  

(7)

where \( M_d, C_d \) and \( G_d \) are the desired inertia, damping and stiffness matrices, respectively.

Remark 2: Model (7) is a general admittance model which defines the relationship between interaction torque and joint angles. In certain situations, a more simplified stiffness model may be adopted

\[ G_d(q_r - q_d) = -\tau_{\text{ext}} \]  

(8)

Besides, \( M_d, C_d \) and \( G_d \) are constant matrices which implies identical characteristics in all directions. Obviously, in order to have different characteristics in different directions, these three matrices should be defined as positive definite matrices with different diagonal elements. When \( \tau_{\text{ext}} = 0 \), \( q_r \) is the desired trajectory to be tracked by in the absence of interaction. However, when \( \tau_{\text{ext}} \) is not null, the virtual desired trajectory \( q_r \) will be generated.

The adaptive control scheme is to let robot follow the desired trajectory and drive the tracking error \( e_q = q - q_r \) into a small neighborhood of 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 to quantify the interaction performance

\[ V(t) = \int_0^\infty (q - q_d)^T Q(q - q_d) + \dot{\tau}_{\text{ext}}^T R \dot{\tau}_{\text{ext}} \, dt \]  

(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: Different cost functions similar to (9) have been discussed in some related works [21]. In a traditional LQR problem, a cost function includes control input and trajectory tracking errors. The optimal control performance can be achieved by specifying feedback gains. In this paper, both the robot and environment systems are taken into consideration.

III. METHODOLOGY

The objective of the admittance adaptation is to obtain the target admittance model according the changing environment. The inner-loop of the system is to guarantee the tracking performance. 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 method requiring computation of joint accelerations or the inversion of the inertia matrix [33], the generalized momentum approach assumes that only motor torque \( \tau \), joint angle \( q \) and joint velocity \( \dot{q} \) are available. In [33] the generalized momentum is defined as

\[ p = M(q)\dot{q} \]  

(10)

Its differential form with respect to time is

\[ \dot{p} = \dot{M}(q, \dot{q})\dot{q} + \tau - C(q, \dot{q})\ddot{q} - G(q) - \tau_{\text{ext}} \]  

(11)

Substituting (11) into (4), we have

\[ \dot{p} = \ddot{M}(q, \dot{q})\dot{q} + \tau - C(q, \dot{q})\ddot{q} - G(q) - \tau_{\text{ext}} \]  

(12)

Considering that the matrix \( M \) is symmetric and positive definite and Coriolis matrix is expressed using Christoffel
symbols [32], the time derivative of inertia matrix $M$ can be written as
\[ \dot{M} = C + CT \]  
Substituting (13) into (12), we have
\[ \dot{p} = CT(q, \dot{q})\dot{q} + \tau - G(q) - \tau_{ext} \]  
It is obvious that equation (14) based on the generalized momentum does not involve joint angle acceleration $\ddot{q}$. Finally, the external torques can be modelled as
\[ \tau_{ext} = A\tau_{ext} + w_r \]  
where $w_r$ is the uncertainty, $w_r \sim N(0, R_r)$. Usually, the matrix $A_r$ is defined as $A_r = 0_{n \times u}$. However, a negative diagonal matrix can reduce the offset of the estimation of disturbances. Then, equation (14) can be rewritten as
\[ \dot{p} = u - \tau_{ext} \]  
where $u$ is defined as
\[ u = \tau + CT(q, \dot{q})\dot{q} - G(q) \]  
The above equations can be combined and reformulated in the space-state form
\[ \begin{bmatrix} \dot{p} \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & -I_n \\ A_c & 0 \end{bmatrix} \begin{bmatrix} p \\ \tau_{ext} \end{bmatrix} + \begin{bmatrix} I_n & 0_n \\ 0_n & 0_n \end{bmatrix} u + \begin{bmatrix} 0 \\ w_r \end{bmatrix} \]  
where $v$ is the measurement noise $v \sim N(0, R_c)$. It can be easily proved that this system is observable. Since $q$ and $\dot{q}$ are able to be measured, the generalized momentum $p = M(q)\dot{q}$ can be regarded as a measured variable. Then, a state observer is designed
\[ \begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = A_c \dot{x} + B_c u + L(y - \hat{y}) \]  
where $L$ in the system can be calculated by
\[ L = PC_c^T R_c^{-1} \]  
where the matrix $P$ can be calculated by the algebraic Riccati equation (ARE) [34]
\[ A_cP + PA_c^T - PC_c^T R_c^{-1}C_cP + Q_c = 0 \]  
where $Q_c$ is the uncertainty of the state, written as
\[ Q_c = diag([0, R_c]) \]  
A schematic overview of the observer is shown in Fig. 3. As shown in equation (19), the output $y = C_c x(t)$ is compared with $C_c \dot{x}(t)$. 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
\[ \begin{bmatrix} \tau_{ext} \\ \hat{\tau} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} \]  
where $\dot{\hat{x}}$ is defined as
\[ \dot{\hat{x}} = A_c \dot{x} + B_c u + L(y - \hat{y}) \]  
\[ \hat{\tau} = L \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} \]  
\[ \hat{\tau} = L \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} \]  
B. Adaptive Optimal Control
The adaptive optimal control strategy is proposed in [31], which is outlined in the following. Consider a continuous-time linear system:
\[ \ddot{\xi}(t) = A\xi(t) + Bu(t) \]  
where $\xi \in \mathbb{R}^m$ is the system state variable, $u \in \mathbb{R}^l$ is the system input, $A \in \mathbb{R}^{m \times m}$ and $B \in \mathbb{R}^{m \times r}$ are the system matrix and input matrix assumed to be constant and unknown. The following optimal control input
\[ u = -K\xi \]  
which can minimize the cost function as follows
\[ V = \int_0^\infty (\xi^T Q\xi + u^T R u) dt \]  
The solution to this problem is similar to that of the LQR problem. The LQR provides a systematic way to find feedback gains that guarantee the optimal control performance. In optimal control theory [35], when $A$ and $B$ are known, there exists a symmetric positive matrix $P^*$, which is the solution of the ARE
\[ PA + A^T P + Q - PBR^{-1}B^T P = 0 \]  
Then, we can obtain the optimal feedback gain matrix
\[ K^* = -R^{-1}B^T P^* \]  
Therefore, we can obtain the optimal control input from equation (28). Next, we give the on-line learning algorithm to obtain the optimal control input subject to unknown dynamics of the environment.
For convenience, let us introduce the following definitions: $P_k \in \mathbb{R}^{m \times m} \rightarrow P_k \in \mathbb{R}^{m \times (m+1)}$ and $\xi \in \mathbb{R}^m \rightarrow \mathbb{R}^{m \times (m+1)}$ where $P_k$ is a symmetric matrix
\[ P_k = [P_{k11}, P_{k12}, \ldots, P_{k1m}, P_{k22}, P_{k23}, \ldots, P_{km}^m]^T \]  
\[ \bar{\xi} = [\xi_1^2, \xi_1\xi_2, \ldots, \xi_1\xi_m, \xi_2^2, \xi_2\xi_3, \ldots, \xi_m^2]^T \]  
\[ \delta\xi = [\xi(t_1) - \bar{\xi}(t_0), \xi(t_2) - \bar{\xi}(t_1), \ldots, \xi(t_l) - \bar{\xi}(t_{l-1})]^T \]  
\[ I_{\xi\xi} = \int_{t_0}^{t_1} \xi \otimes \xi dt, \int_{t_1}^{t_2} \xi \otimes \xi dt, \ldots, \int_{t_{l-1}}^{t_l} \xi \otimes \xi dt \]  
\[ I_{\xi u} = \int_{t_0}^{t_1} \xi \otimes u dt, \int_{t_1}^{t_2} \xi \otimes u dt, \ldots, \int_{t_{l-1}}^{t_l} \xi \otimes u dt \]
where \( l \) is a positive integer and \( \otimes \) is the Kronecker product. Now, we let \( u = K_0 \xi + \phi \) be the initial input, \( l \in [t_0, t_l] \). \( \phi \) 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

\[
\text{rank}([I_{\xi \xi}, I_{\xi u}]) = \frac{m(m+1)}{2} + nr
\]

After the rank condition is satisfied, we can solve \( P_k \) and \( K_{k+1} \) according to the following equation

\[
\begin{bmatrix}
\hat{P}_k \\
\text{vec}(K_{k+1})
\end{bmatrix} = (\Theta_k^T \Theta_k)^{-1} \Theta_k^T \Xi_k
\]

\[
\Theta_k \text{ and } \Xi_k \text{ are defined as}
\]

\[
\Theta_k = [\delta_{t\xi}, -2I_{t\xi}(I_m \otimes K_f^TR) - 2I_{t\xi}(I_m \otimes R)]
\]

\[
\Xi_k = -I_{t\xi} \text{vec}(Q_k)
\]

\[
Q_k = Q + K_f^TRK_k
\]

where \( I_m \) is the \( m \)-dimensional unit matrix, \( \text{vec}(\cdot) \) is the function to transfer a matrix to a vector. Then, we repeat the calculation until \( \|P_k - P_{k-1}\| < \varepsilon \), where \( \varepsilon \) is a small constant defined by the designer. 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 equation (5). The cost function defined in the equation (9) is to be minimised. Comparing the cost function in this paper with the general linear system (24), we need to make them identical. Define the state variable

\[
\xi = [q^T, q_d^T]^T
\]

Then, equation (9) can be

\[
V = \int_0^\infty \left( [q^T, q_d^T]Q'[q_{\text{ext}}] + \hat{\tau}_{\text{ext}}^T R \hat{\tau}_{\text{ext}} \right) dt
\]

\[
= \int_0^\infty (\xi^T Q' \xi + \hat{\tau}_{\text{ext}}^T R \hat{\tau}_{\text{ext}}) dt
\]

where

\[
Q' = \begin{bmatrix}
Q & -Q \\
-Q & Q
\end{bmatrix}
\]

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

\[
\dot{\xi} = A \xi + B \hat{\tau}_{\text{ext}}
\]

where

\[
A = \begin{bmatrix}
-C_{E}^{-1}G_E & 0 \\
0 & I_n
\end{bmatrix}, B = \begin{bmatrix}
-C_{E}^{-1} \\
0
\end{bmatrix}
\]

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 adaptive optimal control method discussed above to obtain the control input as follows to minimise the cost function

\[
\hat{\tau}_{\text{ext}} = -K_k \xi
\]

### 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 and \( \phi \) is the exploration noise. Compute \( \delta_{t\xi}, I_{t\xi}, I_{tu} \) 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 until \( \|P_k - P_{k-1}\| < \varepsilon \), where \( \varepsilon \) is a small constant.
4. Use \( u = K_k \xi \) as the approximated optimal control input.

where \( K_k \) will be obtained by the on-line learning algorithm discussed in the previous section.

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

\[
P^*_k = \begin{bmatrix}
P_1 & P_2 \\
P_1 & \ast
\end{bmatrix}
\]

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

\[
\hat{\tau}_{\text{ext}} = R^{-1}P_1q - R^{-1}P_2q_d
\]

Comparing equation (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 virtual desired trajectory \( q_d \) in joint space and the inner-loop is to guarantee the trajectory tracking.

### D. RBFNN

Radial basis function neural network (RBFNN) has capabilities of approximating any continuous function [36] \( f(\theta): R^m \rightarrow R \) as follows

\[
f(\theta) = W^T Z(\theta) + \varepsilon(\theta)
\]

where the input vector \( \theta \in \Omega \subset \mathbb{R}^m \), weight vector \( W = [w_1, w_2, ..., w_d] \in \mathbb{R}^d \), \( d \) 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

\[
Z_i(\theta) = \exp\left[\frac{-\|\theta - u_i\|^2}{\eta_i^2}\right], \quad i = 1, ..., d
\]

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

\[
f(\theta) = W^*^T Z(\theta) + \varepsilon^*(\theta)
\]

The ideal weight vector \( W^* \) is defined as the value of \( W \) which minimizes \( \varepsilon(\theta) \) for all \( \theta \in \Omega \subset \mathbb{R}^m \)

\[
W^* = \arg \min_{W \in R^m} \{ \sup_{\theta \in R^m} |f(\theta) - W^T Z(\theta)| \}
\]

In general, the ideal weights \( W^* \) are unknown and need to be estimated by \( W \). The weight estimation errors are defined as

\[
\tilde{W} = W^* - \hat{W}
\]
**E. Controller Design**

The inner-loop of the system is to guarantee tracking performance. An adaptive neural based controller is designed to achieve the objective. Considering the dynamics of robot manipulator (4), we define

\[
\begin{align*}
    s &= \dot{e}_q - \Lambda e_q \\
    v &= \ddot{q}_r + \Lambda e_q \\
    \Lambda &= \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)
\end{align*}
\]

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

\[
M(q)\ddot{s} + C(q, \dot{q})s + G(q) + M(q)\ddot{q} + C(q, \dot{q})v = \tau - \tau_{ext}
\]

Design the control torque input

\[
\tau = \dot{G} + \dot{M}v + \dot{C}v + \tau_{ext} - Ks
\]

where \( \dot{G}(q), \dot{M}(q) \) and \( \dot{C}(q, \dot{q}) \) are the estimates of \( G(q), M(q) \) and \( C(q, \dot{q}) \); \( K = \text{diag}(k_1, k_2, \ldots, k_i) \) with \( k_i > \frac{1}{2} \). The closed-loop dynamics can be written as

\[
M(q)\ddot{s} + C(q, \dot{q})s + Ks = -(M(M)\ddot{q} - (C - \dot{C}))v - (G + \dot{G}) + (\tau_{ext} - \tau_{ext})
\]

Using the approximation method, we have

\[
\begin{align*}
    M(q) &= W_M^T Z_M(q) + \varepsilon_M(q) \\
    C(q, \dot{q}) &= W_C^T Z_C(q, \dot{q}) + \varepsilon_C(q) \\
    G(q) &= W_G^T Z_G(q) + \varepsilon_G(q)
\end{align*}
\]

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

\[
\begin{align*}
    Z_M(q) &= \text{diag}(Z_q, ..., Z_q) \\
    Z_C(q, \dot{q}) &= \text{diag}(Z_q, Z_q, ..., Z_q, Z_q) \\
    Z_G(q) &= \text{diag}(Z_q, ..., Z_q)
\end{align*}
\]

where

\[
\begin{align*}
    Z_q &= [\psi(||q - q_1||), ... , \psi(||q - q_n||)]^T \\
    Z_q &= [\psi(||\dot{q} - \dot{q}_1||), ... , \psi(||\dot{q} - \dot{q}_n||)]^T
\end{align*}
\]

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

\[
\begin{align*}
    \hat{M}(q) &= \hat{W}_M^T Z_M(q) \\
    \hat{C}(q, \dot{q}) &= \hat{W}_C^T Z_C(q, \dot{q}) \\
    \hat{G}(q) &= \hat{W}_G^T Z_G(q)
\end{align*}
\]

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

\[
M(q)\ddot{s} + C(q, \dot{q})s + Ks = -\hat{W}_M^T Z_M \ddot{q} - \hat{W}_C^T Z_C \dot{v} - \hat{W}_G^T Z_G \varepsilon - \varepsilon_r
\]

where \( \hat{W}_M = W_M^* - \hat{W}_M, \hat{W}_C = W_C^* - \hat{W}_C \) and \( \hat{W}_G = W_G^* - \hat{W}_G \) and \( \varepsilon_r = \tau_{ext} - \tau_{ext} \). Considering the Lyapunov function

\[
V = \frac{1}{2}s^T Ms + \frac{1}{2}\text{tr}(\hat{W}_M^T Q_M \hat{W}_M + \hat{W}_C^T Q_C \hat{W}_C + \hat{W}_G^T Q_G \hat{W}_G)
\]

where \( Q_M, Q_C, Q_G \) are positive definite matrices to be set by the designer. The derivative of \( V \) can be written as

\[
\dot{V} = s^T M \dot{s} + \frac{1}{2} \text{tr}(\hat{W}_M^T Q_M \dot{\hat{W}}_M + \hat{W}_C^T Q_C \dot{\hat{W}}_C + \hat{W}_G^T Q_G \dot{\hat{W}}_G)
\]

By using the estimation of the weight matrices \( \hat{W}_M, \hat{W}_C, \hat{W}_G \) to approximate \( W_M^*, W_C^*, W_G^* \), the errors between the actual and the ideal RBFNN can be expressed as

\[
\begin{align*}
W_M^T Z_M(q) - W_M^* Z_M(q) &= \hat{W}_M^T Z_M(q) \\
W_C^T Z_C(q) - W_C^* Z_C(q) &= \hat{W}_C^T Z_C(q) \\
W_G^T Z_G(q) - W_G^* Z_G(q) &= \hat{W}_G^T Z_G(q)
\end{align*}
\]

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

\[
\hat{W}_M = \hat{W}_M \\
\hat{W}_C = \hat{W}_C \\
\hat{W}_G = \hat{W}_G
\]

Considering \( 2C(q, \dot{q}) - \dot{M}(q) \) is a skew-symmetric matrix [32] and (54), we have

\[
\dot{V} = s^T M \dot{s} + s^T C \dot{s}
\]

\[
\begin{align*}
    &+ \text{tr}(\hat{W}_M^T Q_M \dot{\hat{W}}_M + \hat{W}_C^T Q_C \dot{\hat{W}}_C + \hat{W}_G^T Q_G \dot{\hat{W}}_G) \\
    &= -s^T(Ks + \hat{W}_M^T Z_M + \hat{W}_C^T Z_C \dot{v} + \hat{W}_G^T Z_G \varepsilon + \varepsilon_r) + \text{tr}(\hat{W}_M^T Q_M \dot{\hat{W}}_M + \hat{W}_C^T Q_C \dot{\hat{W}}_C + \hat{W}_G^T Q_G \dot{\hat{W}}_G) \\
    &= \text{tr}(\hat{W}_M^T (Z_M \dot{s}^T + Q_M \dot{\hat{W}}_M)) \\
    &\quad - \text{tr}(\hat{W}_C^T (Z_C \dot{s}^T + Q_C \dot{\hat{W}}_C)) \\
    &\quad - \text{tr}(\hat{W}_G^T (Z_G \dot{s}^T + Q_G \dot{\hat{W}}_G))
\end{align*}
\]

The update law is designed as

\[
\begin{align*}
    \dot{\hat{W}}_M &= -Q_M \hat{W}_M + s^T e_r + \text{tr}[\sigma_M \hat{W}_M \hat{W}_M] \\
    \dot{\hat{W}}_C &= -Q_C \hat{W}_C + \text{tr}[\sigma_C \hat{W}_C \hat{W}_C] \\
    \dot{\hat{W}}_G &= -Q_G \hat{W}_G + \text{tr}[\sigma_G \hat{W}_G \hat{W}_G]
\end{align*}
\]

where \( \sigma_M, \sigma_C, \sigma_G \) are constants to be specified by the designer. Substituting (60) into (59), the derivative of \( V \) is

\[
\dot{V} = -s^T Ks - s^T e_r + \text{tr}[\sigma_M \hat{W}_M \hat{W}_M] + \text{tr}[\sigma_C \hat{W}_C \hat{W}_C] + \text{tr}[\sigma_G \hat{W}_G \hat{W}_G]
\]

Using Young’s inequality [37]

\[
\begin{align*}
    \text{tr}[\hat{W}_M^T \hat{W}_M] &\leq \frac{1}{2} ||\hat{W}_M||_F^2 + \frac{1}{2} ||\hat{W}_C||_F^2 \\
    -s^T e_r &\leq \frac{1}{2} s^T s + \frac{1}{2} e_r^2 \\
    \text{tr}[\sigma_M \hat{W}_M \hat{W}_M] &\leq \frac{1}{2} ||\hat{W}_M||_F^2 \\
    -\sigma_C ||\hat{W}_C||_F^2 &\leq \frac{1}{2} ||\hat{W}_C||_F^2 \\
    -\sigma_G ||\hat{W}_G||_F^2 &\leq \frac{1}{2} ||\hat{W}_G||_F^2
\end{align*}
\]

Then (61) can be written as

\[
\dot{V} \leq -s^T Ks + \frac{1}{2} ||s||^2 + \frac{1}{2} ||e_r||^2 + \alpha
\]

\[
-\frac{\sigma_M}{2} ||\hat{W}_M||_F^2 - \frac{\sigma_C}{2} ||\hat{W}_C||_F^2 - \frac{\sigma_G}{2} ||\hat{W}_G||_F^2
\]
where $\alpha = \frac{\sigma_M}{2} ||W_M||^2 + \frac{\sigma_C}{2} ||W_C||^2 + \frac{\sigma_G}{2} ||W_G||^2$. A sufficient condition for $V \leq 0$ is that
\[
\alpha \leq s^T(K - \frac{1}{2}I)s + \frac{1}{2}||e_\tau||^2 + \frac{\sigma_M}{2}||\tilde{W}_M||^2 + \frac{\sigma_C}{2}||W_C||^2 + \frac{\sigma_G}{2}||W_G||^2
\] (64)

where $I$ is the unit matrix. Let $\chi$ denotes the state variable comprised of $e_\tau, s, \tilde{W}_M, W_C, W_G$ defined in the Lyapunov function candidate, and it follows from (64) that
\[
V(\chi) < 0, \forall ||\chi|| > \rho
\] (65)

where $\rho$ is a positive constant. In other words, the time derivative of $V(\chi)$ is negative inside the set $\Omega_s = \{ ||\chi|| \leq \rho \}$, which is defined in Theorem 1 as below, or equivalently, all $\chi(t)$ that start outside $\Omega_s$ will enter the set within a finite time, and will remain inside the set afterwards. Choose $0 < V(\chi) < \epsilon < c$, and suppose that the sets $\Omega_s = \{ V(\chi) \leq \epsilon \}$ and $\Omega_c = \{ V(\chi) \leq c \}$.

Let
\[
\Upsilon = \{ \epsilon \leq V(\chi) \leq \epsilon \} = \Omega_c - \Omega_s
\] (66)

It is known that the time derivative of $V(\chi)$ is negative inside $\Upsilon$, that is
\[
\dot{V}(\chi(t)) < 0, \forall \chi \in \Upsilon, \forall t \geq t_0
\] (67)

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

We can conclude the convergence and stability results in Theorem 1.

**Theorem 1:** Using the Uniformly Ultimately Bounded (U-UB) theorem, the tracking error $s$, estimation error $e_\tau$ and weight errors $\tilde{W}_M, W_C, W_G$ will fall into an set $\Omega_s$, where the bounding set $\Omega_s$ is defined in (68) and shown in Fig. 4.

\[
\Omega_s = \left\{ (||\tilde{W}_M||, ||W_C||, ||W_G||, ||e_\tau||, ||s||) \mid \frac{\sigma_M}{2}||\tilde{W}_M||^2 + \frac{\sigma_C}{2}||W_C||^2 + \frac{\sigma_G}{2}||W_G||^2 + \frac{s^T(K - \frac{1}{2}I)s}{\alpha} + \frac{||e_\tau||^2}{\alpha} \leq 1 \right\}
\] (68)

As shown in Fig. 4, the bounding set $\Omega_s$ is the area in the first quadrant, passing through the points $(\frac{\sigma_M}{2}||\tilde{W}_M||^2 = \alpha, ||e_\tau||^2 = 0, ||s||^2 = 0)$, $(\frac{\sigma_C}{2}||W_C||^2 = 0, ||e_\tau||^2 = ||\tilde{W}_M||^2 = 0, ||s||^2 = 0)$. In Fig. 4, we define
\[
\text{when } \frac{\sigma_M}{2}||\tilde{W}_M||^2 = \alpha, \tilde{W} = \omega
\]
\[
\text{when } \frac{s^T(K - \frac{1}{2}I)s}{\alpha} = \alpha, s = \beta
\] (69)

Since $||\tilde{W}_M||$ is a bounded constant, $||\tilde{W}_M|| = ||W_C|| = ||W_G||$ is bounded. Since $||e_\tau||$ is bounded, $||\tilde{e}_\tau|| = ||e_\tau + \tau_{ext}||$ is bounded. With the bounded signals $q_e$ and $q_i$, according to (46), $||\dot{\tau}||$ is bounded and $||\dot{q}|| = ||\dot{e}_q + q_i||$ is bounded as well. Therefore, the norm of state variable $\chi$ is bounded.

**IV. Simulation Studies**

In this section, we consider a 2-link robot arm in physical interaction with unknown environment. The simulation environment is shown in Fig. 5. The environment torque is applied at the end-effector of the robot arm in the Y direction. 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 virtual desired trajectory $q_r$. The parameters of the robot arm are shown in Table I.

![Environment](image)

**Fig. 5.** An overview of the scenario.

**TABLE I**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1$</td>
<td>Mass of link 1</td>
<td>2.0 kg</td>
</tr>
<tr>
<td>$m_2$</td>
<td>Mass of link 2</td>
<td>2.0 kg</td>
</tr>
<tr>
<td>$l_1$</td>
<td>Length of link 1</td>
<td>0.20 m</td>
</tr>
<tr>
<td>$l_2$</td>
<td>Length of link 2</td>
<td>0.20 m</td>
</tr>
<tr>
<td>$I_1$</td>
<td>Inertia moment of link 1</td>
<td>0.027 kgm²</td>
</tr>
<tr>
<td>$I_1$</td>
<td>Inertia moment of link 2</td>
<td>0.027 kgm²</td>
</tr>
</tbody>
</table>

A. Interaction Performance

It is assumed that the dynamics of the environment can be defined as: $0.01\dot{q} + (q - 0.3) = -\tau_{ext}$. The LQR method is used to verify the effectiveness of the proposed method. If the matrices $A$ and $B$ are known, we can use optimal solutions of the ARE to obtain the optimal parameters of the admittance
model. 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 \(P_0 = 10I_p\), where \(I_p\) represents the \(p\) dimensional unit matrix. Then, the second step is conducted and stops until ||\(P_k|| < 0.02\), where ||·|| denotes the Euclidean norm.

The weights of the cost function are given by \(Q = 1\) and \(R = 1\). The desired admittance model is obtained as \(\hat{\tau}_{ext} = -0.4142q + 0.0702q_t\), based on the known \(A\) and \(B\). Simulation results are shown in Figs. 6-9. In Fig. 6, the desired trajectories of the robot arm in joint space are shown. At the beginning, there is a large error between the LQR and proposed method, due to the initial admittance model: \(\hat{\tau}_{ext} = -q + 0.1(q - 0.3)\) 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. The error of admittance parameters between LQR and proposed method is shown in Fig. 7. The convergence of \(P_k\) and \(K_k\) to optimal values based on LQR are illustrated. The error is defined as ||\(K_k - K^*\)||. 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: \(\hat{\tau}_{ext} = -0.4173q + 0.00904q_t\). The symmetric positive matrix with \(Q = 1\) and \(R = 1\), \(P_k\) from the proposed algorithm and \(P^*\) from LQR are shown below:

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

To further verify the correctness of the proposed method, different weights of the cost function are given by \(Q = 5\) and \(R = 1\). Simulation results are shown in Figs. 10-13. Similarly, the desired admittance model is obtained as \(\hat{\tau}_{ext} = -1.4495q + 0.2033q_t\) based on the known \(A\) and \(B\). After three steps of iteration, the admittance model is obtained as \(\hat{\tau}_{ext} = -1.5374q + 0.2063q_t\) with the proposed method. The desired trajectory \(q_d\) and virtual desired trajectory \(q_v\) with \(Q = 5\) and \(R = 1\) are illustrated in Fig. 10. As shown 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 \(-0.35N.m\) at 7s, 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 below:

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

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 joint angle of the robot arm is $q(0) = [2.6, -2.6]^T$ and the desired trajectory of the robot arm is given by $q_d = [0.3 + 0.2e^{-t}, 0.3e^{-t}]^T$. The weight matrices are initialized as $\hat{W}_M(0) = \mathbf{0}, \hat{W}_C(0) = \mathbf{0}, \hat{W}_G(0) = \mathbf{0}$. The results of tracking performance are shown in Figs. 14-20. In Fig. 14, the desired joint angle trajectory and actual trajectory are shown and the tracking errors are illustrated in Fig. 15. It can be found that the proposed control algorithm can make the tracking errors converge to a small neighborhood of zero. The function approximation performance of the our implemented neural network is depicted in Figs. 16-18. As seen from the figures, we can observe that the output of the implemented neural network could follow the approximated nonlinear function’s dynamics ($M(q), C(q, \dot{q}), G(q)$), which implies that the implemented neural networks have the ability to approximate the nonlinear function with satisfactory tracking performance. As shown in Fig. 19, the weight matrices of the system shows a trend of convergence and the control input is shown in Fig. 20. With the tracking errors and torque regulation, as shown in Fig. 8, the cost function is minimised, as shown in Fig. 9.

**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. A neural 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 minimised. Simulation studies have verified the effectiveness of our proposed method.

REFERENCES


