IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 16, 2020, accepted October 1, 2020, date of publication October 6, 2020, date of current version October 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3029074

Adaptive Dynamic Programming-Based Fault-Tolerant Position-Force Control of Constrained Reconfigurable Manipulators

BING MA, BO DONG, FAN ZHOU, AND YUANCHUN LI

Department of Control Science and Engineering, Changchun University of Technology, Changchun 130021, China

Corresponding author: Yuanchun Li (liyc@ccut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61773075 and Grant 61703055, in part by the Scientific Technological Development Plan Project in Jilin Province of China under Grant 20200801056GH and Grant 20190103004JH, and in part by the Science and Technology Project of Jilin Provincial Education Department of China During the 13th Five-Year Plan Period under Grant JJKH20200672KJ, Grant JJKH20200673KJ, and Grant JJKH20200674KJ.

ABSTRACT This article presents a novel fault-tolerant position-force optimal control method for constrained reconfigurable manipulators with uncertain actuator failures. On the basis of the radial basis function neural network (RBFNN)-estimated manipulators dynamics, the proposed force-position error fusion function and the estimated actuator failure are utilized to construct an improved optimal performance index function, which reflects the faults and optimizes system comprehensive performance as well as the energy consumption simultaneously. Based on the policy iteration (PI) scheme and the adaptive dynamic programming (ADP) algorithm, the Hamiltonian-Jacobi-Bellman (HJB) equation is solved by constructing the critic neural network (NN), and then the approximated fault-tolerant position-force optimal control policy can be derived correspondingly. The closed-loop manipulator system is proved to be asymptotically stable by using the Lyapunov theory. Finally, simulations are provided to demonstrate the effectiveness of the proposed method.

INDEX TERMS Reconfigurable manipulators, adaptive dynamic programming, fault-tolerant position-force control, optimal control, neural network.

I. INTRODUCTION

The replacement of manual production by manipulators is a significant development trend of digitalization, industrial automation, and intelligence. To meet the requirements of high exploitation and portability in the modern manufacturing industry, the reconfigurable manipulators [1] equipped with standardized modules are capable of adapting to the severe working conditions through the changes in the configurations and by increasing/reducing the degrees of freedom. Nowadays, the reconfigurable manipulators have the potential of wide applications in numerous extreme and restricted environments such as human-robot cooperations, medical rescue operations, aerospace explorations, and such others, needing an effective control strategy to ensure the security and precision of the manipulator systems even under circumstances of uncertain failures.

Fault tolerant control is an advanced method of regulation to ensure the safe operation of the system in the event of the failure of certain components or parameters, which includes passive fault-tolerant control [2], [3], robust fault-tolerant control [4], [5] and active fault-tolerant control [6], [7]. Fault tolerant control possesses the wide application value and the necessary research significance in various fields, especially in intelligent manufacturing industry. Zhu and Li [8] designed the robust $H\infty$ observer to handle the road condition variations for the integrated motor-transmission (IMT) system. To cater to the ever growing demands of the scientific researchers and the practical applications, the manipulators are being widely utilized in complex environments to complete certain special tasks such as grasping, polishing, rehabilitating, etc. In condition of contacting with environments involving even human contacts, among the various possible failures, the actuator failure is considered to be one of the most crucial challenges. The sudden, unexpected and unknown movement of the actuator makes the whole

The associate editor coordinating the review of this manuscript and approving it for publication was Shen Yin.

system go awry and out of control, and at times leads to irreparably serious consequences. Recently, several fault-tolerant position-force control methods have been developed to overcome the aforementioned problems through the application of various approaches and theories. Position-force control, is one such efficient method to deal with the complex robot-environment contacted tasks, drawing extensive attention in the robotics community. Over the decades, multiple kinds of the position force control methods classified as impedance control [9]–[11], hybrid position-force control [12], [13] and parallel position-force control [14], have been actualized for various types of robot manipulators. Doulgeri and Arimoto [15] focused on investigating the force commanded impedance control algorithm of a soft-tipped robotic finger with uncertain kinematics, overcoming the physical interaction between a rigid object and the robotic finger with unknown nonlinearities of the reproducing forces. Javier and Marco [16] presented an adaptive position-force control method alleviating the problem of the uncertainty between the robot and the constraint surface. Moreover, to address the fault-tolerance position-force control problems of the manipulators, Yousef *et al.* [17] devised a force-position active fault accommodation strategy for the legged robots subject under the actuator failures including actuation bias and effective gain degradation. Different from the traditional manipulators, reconfigurable manipulators possess the structural characteristics of variable configurations, that is, the dynamics of the reconfigurable manipulators are uncertain which change with the requirements of the tasks generally. With the development of intelligent algorithm, the powerful learning ability of NN is often applied to estimate arbitrary parameter or function [18], [19]. Taking advantage of the above superiority, Xu *et al.* [19] proposed a neural-approximation-based robust adaptive control for a constrained flexible air-breathing hypersonic vehicle (FAHV) subject, which utilized two RBFNNs are applied to approximate the lumped unknown nonlinearities of the velocity subsystem and the altitude subsystem. For the unknown dynamics of reconfigurable manipulators, Zhou *et al.* [20] proposed a force/position fault-tolerant control method of constrained reconfigurable manipulators, which consisted of a modified sliding mode controller to ensure the force/position tracking performance and a RBFNN-based compensation controller to increase the robustness of the manipulator system. Nonetheless, all these methods mentioned above ignored the problem of the comprehensive optimization of the control performance and power consumption in the event of actuator failures. Evidently, the reconfigurable manipulators are mostly utilized in extreme environments and required to work under an 'optimal' state, that is, to attain a multi-part balancing point thereby ensuring the energy consumption, performance, and stability working at its best balance. To the best of authors' knowledge, there has been very little discussion on the fault-tolerant optimal control methods that directly place the observed actuator failures as a control indicator into the performance index function for robotic manipulators, especially, in case of the reconfigurable manipulators under contact in the complex working environments.

Optimal control has been receiving widespread attentions from the researchers and the manufacturers since the mid-1950s, when it was formed and developed under the promotion of space technology. As an appropriate application for solving the optimal control problems of nonlinear systems, ADP algorithm, which was first proposed by Werbos [21], was considered as an effective approach in avoiding the difficulties of the 'curse of dimensionality'. Today, the ADP-based methods are being utilized in the designing of continue-time [22], [23], discrete time [24]–[26], data driven-based [27], [28] intelligence systems, and the solution of nonlinear optimal control with input/output constraints [29], [30], external disturbances [31]–[33] and actuator failures [34], [35]. Several investigations studied, the optimal control problems of the robot manipulator systems based on the ADP approach. Li *et al.* [36] propounded an adaptive neural network tracking approach based on the reinforcement learning to resolve the problem of the trajectory changing with time for the wheeled mobile robots. Dong *et al.* [37] concentrated on studying the decentralized optimal control method for the reconfigurable manipulators, designing a model-based compensation controller and an ADP-based optimal controller to deal with the influence of the unknown internal dynamics and the interconnected dynamic coupling, respectively. Nevertheless, these ADP-based optimal control methods focused on solving the position tracking problem of the complex nonlinear systems, while there were only a few investigations addressing the problems of actuator failures with reinforcement learning theory of the constrained manipulator systems. Zhao *et al.* [38] suggested an online fault compensation control scheme based on the Policy Iteration (PI) algorithm for a class of affine non-linear systems with actuator failures reconstructed adaptively to achieve online fault compensation. In [39], a Neural-network-based robust hybrid position/force controller was put forth, which included a main controller to track the motion/force trajectory objectives, along with an adaptive neural network controller to compensate for the deficiencies of the manipulators model. However, the method merely considered the optimal compensation in case of the uncertainties between the manipulators and the environment, neglecting the global optimization and the structural characteristics of the reconfigurable manipulators. Indeed, reconfigurable manipulators have been always utilized in constrained and extreme environments such as military battlefields, space explorations and rescue operations, which are supposed to possess the abilities of fault tolerance and optimized energy consumption, taking into account the working efficiency and maintaining the lowest consumption state to prolong their service life. Moreover, the optimal control approaches mentioned above were all limited in solving the optimal compensation control problem of a specific class of manipulator systems and overlooked the implementing optimal fault-tolerant position-force control directly in the case of modeless manipulators. Unfortunately, there have been very few researches that concentrated on investigating the fault-tolerant position-force optimal control approach, considering both the position-force control and the ADP-based optimal control for reconfigurable manipulator systems with uncertain actuator failures.

This article, inspired by the above investigations, proposes a novel fault-tolerant position-force control approach based on ADP algorithm for a class of constrained reconfigurable manipulator systems. First, an adaptive fault observer is proposed, according to the RBFNN-estimated dynamics of the constrained reconfigurable manipulators, to detect the uncertain actuator failures in real-time observation of the whole system. Second, the fault tolerant position-force deviation fusion utility function would be defined containing the information of the estimated actuator failures, the joint tracking errors and the torque deviations. Besides, the PI algorithm would be applied to solve the HJB equation, while the cost function would be approximated by the constructed critic neural network. Subsequently, the ADP-based approximated optimal control strategy shall be obtained directly. The closed-loop manipulator system is proved to be asymptotic stability by adopting the Lyapunov theory. Finally, simulations shall be provided to verify the advantages and effectiveness of the developed method.

The main contributions of this article are summarized as follows:

- Unlike the traditional fault-tolerant control approaches relying mostly on the compensate controllers, directly adding the failures obtained by the fault observer to the actuator with higher specifications in the form of a compensator, this article endeavors to present a fault-tolerant optimal control for the constrained reconfigurable manipulators regarding estimated actuator failure by adaptive fault observer as the one of control indicators into the optimal performance index function. The proposed method not only realizes the observation and compensation of the actuator failures, rather, also considers the service life, the performance and the universality of the actuator, concurrently.
- In variance the existing control methods of manipulators leaved the comprehensive optimization of fault-tolerant control, position-force control and power consumption out of consideration, in this article, a novel fault-tolerant position-force optimal control method for the constrained reconfigurable manipulators is improved. The proposed controller mitigates the problem of model uncertainty, improves the system robustness and comprehensive balances control accuracy, energy consumption and stable operation, simultaneously.

The remainder of this article is arranged as follows: Section II sketches the dynamics model formulation of manipulators system with environmental constraints. The fault-tolerant position-force optimal control scheme of modeless reconfigurable manipulators is proposed in Section III. In Section IV, simulations are provided to verify the advantages and effectiveness of the developed method. Finally, Section V summarizes the result.

II. PROBLEM STATEMENT

In order to acclimatize environments to accomplish a variety of tasks [40], the end-effector contacts with the external environments in general. Considering a constrained reconfigurable manipulator with actuator failure, the dynamic model of the n-DOF manipulators is represented as:

$$M(q)\ddot{q} + C_a(q,\dot{q})\dot{q} + G(q) = u + J_\Phi^T(q) f_c - F_a, \quad (1)$$

where $q, \dot{q}, \ddot{q} \in \mathbb{R}^n$ denote the vector of joint movements, joint velocity, and joint acceleration, correspondingly. $M(q) \in \mathbb{R}^{n\times n}$ is the inertia matrix, $C_a(q,\dot{q}) \in \mathbb{R}^n$ is the centripetal-coriolis matrix, and $G(q) \in \mathbb{R}^n$ is the gravity vector. $u \in \mathbb{R}^n$ is the control torque, which represents the input torque of the manipulator systems generally. $J_\Phi(q) = \nabla_q \Phi(q)$ is the Jacobian matrix, $\Phi(q)$ is the constrained function caused by constrained task space, and f_c defined as the vector of the exerted force to the environment by the manipulators. $F_a \in \mathbb{R}^m$ is an unknown additive actuator failure with bounded as $F_a \le \varepsilon_1 < \infty$, where ε_1 is a positive constant.

The dynamics of constrained reconfigurable manipulator (1) has the following properties [41]:

Property 1 The inertia matrix $M(q)$ is symmetric, positive scalar, and satisfies the following inequalities:
$0 < n_1\|\lambda\|^2 \le \lambda^T M(q)\lambda \le n_2\|\lambda\|^2, \forall\lambda \in R^n$.

Property 2 The inertia and centripetal-coriolis matrices satisfy the following skew-symmetric relationship:

$$\zeta^T\left(\frac{1}{2}\dot{M}(q) - C_a(q,\dot{q})\right)\zeta = 0, \quad \forall\zeta \in R^n,$$

where $\dot{M}(q)$ denotes the time derivatives of the inertia matrix.

By deforming the dynamic model formula (1), the constrained reconfigurable manipulator system state equation is proposed as:

$$I: \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x) + g(x)(u+\tau_c) - g(x)F_a \\ y = x, \end{cases} \quad (2)$$

where $x = [x_1\ x_2]^T = [q\ \dot{q}]^T$, $x_1, x_2 \in \mathbb{R}^n$ are the state vector of system I and y is the exportation of the manipulator system, τ_c is the constrained torque caused by contacting with task environments, $\tau_c = J_\Phi^T(q) f_c$ and $f(x) = M^{-1}(q)(-C_a(q,\dot{q})\dot{q} - G(q))$, $g(x) = M^{-1}(q)$.

Remark 1: In this article, we considered that the task environment of reconfigurable manipulators is known. f_c is the vector of the exerted force to the environment by the manipulators, namely, the interaction force f_c represents the force magnitude which is measurable in task space. In the practical applications of manipulators, f_c is usually bounded to ensure the normal and stable operation of the system, with $|f_c| \le f_M, f_M > 0$.

III. FAULT-TOLERANT POSITION-FORCE OPTIMAL CONTROL BASED ON ADP

A. OPTIMAL CONTROL

For the sake of finding an optimal control policy $u^*(t)$, to ensure the performance and the stability of reconfigurable manipulator systems, enlightened by [42], an improved infinite horizon performance index function is designed as:

$$\underset{u\in\Psi(\Omega)}{Z(m_a)} = \int_0^\infty \left(\rho\hat{F}_a^T(t)\hat{F}_a(t) + N(m_a, u(m_a))\right)dt, \quad (3)$$

where $\hat{F}_a(t) \in \mathbb{R}^m$ is the estimation of the uncertain actuator failure $F_a(t)$, and $\rho > 0$ is a positive constant. $m_a(t)$ is the position-force deviation fusion function defined as $m_a(t) = k_{deq}\dot{e}_q + k_{eq}e_q + k_\tau \int_0^{t_k} e_\tau dt$, $m_{a0}(t) = m_a(0)$, in which $e_q = q - q_d$ is position tracking error, $\dot{e}_q = \dot{q} - \dot{q}_d$ is velocity tracking error, and $e_\tau = \tau_c - \tau_d$ is the contact torque deviation of end-effector, with the desired joint trajectory q_d, the desired velocity $\dot{q}_d$ and the desired contact torque τ_d. k_{deq}, k_{eq}, k_τ are the function coefficients. $N(m_a(t), u(m_a(t))) = {m_a}^T Q m_a + u^T R u$ is the utility function, where $N(0,0) = 0$ and $N(m_a, u(m_a)) \geq 0$. $Q \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{p\times p}$ are the positive definite matrices. $\Psi(\Omega)$ is a series of feasible control policies.

Remark 2: The designed performance index function contains the actuator failures $\hat{F}_a(t)$, position-force errors fusion function $m_a(t)$ and control torque $u(t)$, considering both task implementation and consumption. The proposed method not only observes and compensates the actuator failures, but also attains a multi-part balance point thereby ensuring the whole system stability, performance and energy consumption working at its best.

Assumption 1: The desired joint trajectory q_d, the desired joint velocity $\dot{q}_d$ and the desired joint acceleration $\ddot{q}_d$ are all bounded and known. The desired contact torque τ_d is continuous and known, with the first order integral $\int_0^{t_k} \tau_d dt$ bounded.

Assumption 2: The constrained reconfigurable manipulators keep away from singularities to ensure the full rank of Jacobian matrix.

Definition 1: For the constrained reconfigurable manipulator systems (2), the fault-tolerant position-force control policy $U(m_a)$ is admissible called for the function (3) on a compact set Ω, $Z(m_a)$ is finite, $\forall m_a \in \Omega$. $U(m)$ is continuous on Ω, $U(m_a) \in \Psi(\Omega)$, $U(0) = 0$, $U(m_a) = u(m_a)$ stabilizes $m_a(t)$ on Ω.

Then the infinitesimal version of improved performance index function (3) that is the nonlinear Lyapunov equation is described as:

$$0 = \rho\hat{F}_a^T\hat{F}_a + N(m_a, u(m_a)) + (\nabla Z(m_a))^T \dot{m}_a, \quad (4)$$

$$0 = \rho\hat{F}_a^T\hat{F}_a + N(m_a, u(m_a)) + (\nabla Z(m_a))^T \left(k_{deq}(f(x) + g(x)(u+\tau_c) - g(x)F_a) + v\right), \quad (5)$$

where $v = -k_{deq}\ddot{q}_d + k_{eq}\dot{e}_q + k_\tau e_\tau$, and the term $\nabla Z(m_a)$ shows the partial derivative of $Z(m_a)$ with respect to m_a, $Z(0) = 0$, i.e. $\nabla Z(m_a) = \frac{\partial Z(m_a)}{\partial m_a}$.

Define the HJB function of the problem and the optimal performance index function as:

$$\begin{aligned} &H(m_a, u(m_a), \nabla Z(m_a)) \\ &\quad = \rho\hat{F}_a^T\hat{F}_a + N(m_a, u(m_a)) + (\nabla Z(m_a))^T \dot{m}_a \\ &\quad = \rho\hat{F}_a^T\hat{F}_a + m_a^T Q m_a + u^T R u + (\nabla Z(m_a))^T \\ &\qquad \left(k_{deq}(f(x) + g(x)(u+\tau_c) - g(x)F_a) + v\right), \end{aligned} \quad (6)$$

and

$$Z^*(m_a) = \min_{u(m_a)} \int_0^\infty \left(\rho\hat{F}_a^T(\tau)\hat{F}_a(\tau) + N(m_a u(m_a))\right)d\tau. \quad (7)$$

To attain the optimal control strategy, the solution of HJB equation is

$$0 = \min_{u(m_a)} H\left(m_a, u(m_a), \nabla Z^*(m_a)\right), \quad (8)$$

where $\nabla Z^*(m_a) = \frac{\partial Z^*(m_a)}{\partial m_a}$, if $Z^*(m_a)$ is continuously differentiable, the optimal control by PI algorithm which will be covered in next section can be formulated as

$$u^*(m_a) = -\frac{1}{2}k_{deq}R^{-1}g^T(x)\nabla Z^*(m_a). \quad (9)$$

Combining (5) and (8), by simple transformation, we have

$$\begin{aligned} &(\nabla Z^*(m_a))^T \cdot \left(k_{deq}(f(x) + g(x)((u+\tau_c) - F_a)) + v\right) \\ &\quad = -m_a^T Q m_a - u^T R u - \rho\hat{F}_a^T\hat{F}_a. \end{aligned} \quad (10)$$

B. ADAPTIVE FAULT OBSERVER DESIGN

For the constrained reconfigurable manipulator with actuator failures (2), one can design an adaptive fault observer as

$$\dot{\hat{x}}_o = \begin{cases} \dot{\hat{x}}_{o1} = \hat{x}_{o2} + \alpha_1(x_1 - \hat{x}_{o1}) \\ \dot{\hat{x}}_{o2} = f(\hat{x}_o) + g(\hat{x}_o)(u+\tau_c) - g(\hat{x}_o)\hat{F}_a \\ +\alpha_2(x_2 - \hat{x}_{o2}). \end{cases} \quad (11)$$

where $\hat{x}_o$ is the observation of the system state x, α_1 and α_2 are the positive definite observer gain and $\hat{F}_a$ is the estimation of the actuator failure which can be updated by the following adaptive law as

$$\dot{\hat{F}}_a = -\alpha_F g^T(\hat{x}_o) E_a, \quad (12)$$

where $E_a = x - \hat{x}_o$ is the state observer error and $\alpha_F = \left[\alpha_1\ \alpha_2\right]^T$. Combining (2) and (11), one obtains

$$\begin{aligned} \dot{E}_a &= f(x) + g(x)(u+\tau_c) - g(x)F_a - f(\hat{x}_o) \\ &\quad - g(\hat{x}_o)(u+\tau_c) + g(\hat{x}_o)\hat{F}_a - \alpha_F(x - \hat{x}_o) \\ &= f_e + g_e((u+\tau_c) - F_a) - g(\hat{x}_o)e_F - \alpha_F E_a, \end{aligned} \quad (13)$$

where $f_e = f(x) - f(\hat{x}_o)$ and $g_e = g(x) - g(\hat{x}_o)$ are the observer error of $f(x)$ and $g(x)$, respectively, and define $e_F = F_a - \hat{F}_a$. And then define overall error of observer

$\xi_e = f_e + g_e\left((u+\tau_c) - F_a\right)$, and assume ξ_e is norm-bounded as $\|\xi_e\| \le \varepsilon_2$, ε_2 is a positive constant.

Remark 3: As shown in Property 1-2, the terms of $f(x)$ and $g(x)$ are bounded and the functions of controllable nonlinear systems are bounded, so it is reasonable to assume that the random unknown actuator failure F_a and ξ_e are bounded.

Theorem 1: Consider the dynamic model of the constrained reconfigurable manipulator systems with actuator failures in (2), the observation error of the developed fault observer in (13) is uniformly ultimately bounded (UUB) under the defined adaptive law in (12).

Proof: Selecting the Lyapunov function candidate as:

$$V_1(t) = \frac{1}{2}E_a^T E_a + \frac{1}{2}e_F^T \alpha_F^{-1} e_F. \tag{14}$$

Substituting (13) and adaptive law (12) into the derivative of (14), we have

$$\begin{aligned}\dot{V}_1(t) &= E_a^T \dot{E}_a + \dot{\hat{F}}_a^T \alpha_F^{-1} e_F \\ &= E_a^T\left(f_e + g_e\left((u+\tau_c) - F_a\right) - g\left(\hat{x}\right)e_F\right) \\ &\quad - E_a^T \lambda_{\min}(\alpha_F) E_a - \dot{\hat{F}}_a^T \alpha_F^{-1} e_F \\ &\le -\left(\left(\lambda_{\min}(\alpha_F)\|E_a\|\right) - \varepsilon_2\right)\|E_a\|,\end{aligned} \tag{15}$$

where $\lambda_{\min}(\alpha_F)$ presents the minimum eigenvalue of the matrix. Therefore, according to Lyapunov's direct method, the observation error E_a with the compact set $\Omega_1 = \{E_a : \|E_a\| \le \|\varepsilon_2/\lambda_{\min}(\alpha_F)\|\}$ can be guaranteed to be UUB. This completes the proof of theorem 1.

C. RBFNN-BASED MODEL IDENTIFER OF CONSTRAINED RECONFIGURABLE MANIPULATOR

As we all know, the shape and the degree of freedom (DOF) of reconfigurable manipulators are vary with different and arduous tasks. In this part, the unknown dynamics gain matrixes are obtain by constructing RBFNN identifiers based actual input-output data.

In this part, the RBFNN, with strong capabilities of nonlinear function learning and approximation, is applied to approximate the unknown reconfigurable manipulators dynamics terms:

$$f(x) = W_f^T \delta_f(x) + e_f, \quad \left\|e_f\right\| \le e_{fM}, \tag{16}$$
$$g(x) = W_g^T \delta_g(x) + e_g, \quad \left\|e_g\right\| \le e_{gM}, \tag{17}$$

where W_f and W_g are the ideal NN weights, $\delta_f(x)$ and $\delta_g(x)$ are the NN radial basis function, assuming that there exist constants δ_{fM} and δ_{gM}, such that $\left\|\delta_f(x)\right\| \le \delta_{fM}$ and $\left\|\delta_g(x)\right\| \le \delta_{gM}$. e_f and e_g are the NN approximation errors and bounded, $\left\|e_f\right\| \le e_{fM}$ and $\left\|e_g\right\| \le e_{gM}$, where e_{fM} and e_{gM} are the known constants.

Define $\hat{W}_f$ and $\hat{W}_g$ as the estimations of W_f and W_g, respectively. $\hat{f}(x, \hat{W}_f)$ and $\hat{g}(x, \hat{W}_g)$ are the estimation value of $f(x)$ and $g(x)$, respectively. $\hat{f}(x, \hat{W}_f)$ and $\hat{g}(x, \hat{W}_g)$ are expressed as:

$$\hat{f}(x, \hat{W}_f) = \hat{W}_f^T \delta_f(x), \tag{18}$$
$$\hat{g}(x, \hat{W}_g) = \hat{W}_g^T \delta_g(x). \tag{19}$$

The approximation errors can be defined as $\tilde{W}_f = W_f - \hat{W}_f$ and $\tilde{W}_g = W_g - \hat{W}_g$, thus

$$f(x) - \hat{f}(x, \hat{W}_f) = \tilde{W}_f^T \delta_f(x) + e_f = e_{fH}, \tag{20}$$
$$g(x) - \hat{g}(x, \hat{W}_g) = \tilde{W}_g^T \delta_g(x) + e_g = e_{gH}, \tag{21}$$

where the e_{fH} and e_{gH} are the actual neural network approximation errors and bounded. $\hat{f}(x, \hat{W}_f)$ and $\hat{g}(x, \hat{W}_g)$ are utilized to substitute for the unknown dynamics of the reconfigurable manipulator, and these estimates can be updated as

$$\dot{\hat{W}}_f = \alpha_f k_{deq} {m_a}^T \delta_f(x), \tag{22}$$
$$\dot{\hat{W}}_g = \alpha_g k_{deq} {m_a}^T \delta_g(x)\left((u+\tau_c) - \hat{F}_a\right), \tag{23}$$

where α_f and α_g are positive constants.

Remark 4: As shown in Property 1-2, the unknown terms of $f(x)$ and $g(x)$ are locally Lipschitz and continuous in compact set $\Theta \in \mathbb{R}^n$, there exists the positive constants Υ_f and Υ_g, $\|f(x)\| \le \Upsilon_f \|m_a\|$ and $\|g(x)\| \le \varpi_g$, so it is reasonable to approach the uncertainty of reconfigurable manipulator model by RBFNN approximation ability.

Hence, according to (18) and (19), the state equation of dynamic model space (2) can be updated as

$$\hat{I}: \begin{cases} \dot{\hat{x}}_1 = \hat{x}_2 \\ \dot{\hat{x}}_2 = \hat{f}(x, \hat{W}_f) + \hat{g}(x, \hat{W}_g)\left((u+\tau_c) - F_a\right) \\ \hat{y} = \hat{x}, \end{cases} \tag{24}$$

Combining (18) with (19), the observation of $\dot{m}_a(t)$ as

$$\begin{aligned}\dot{M}_a(t) = k_{deq}&\left(\hat{f}(x, \hat{W}_f) + \hat{g}(x, \hat{W}_g)\left(u(m_a) + \tau_c\right)\right) \\ &- k_{deq}\hat{g}(x, \hat{W}_g)F_a + v,\end{aligned} \tag{25}$$

According to (4) and (25), we have

$$\begin{aligned}0 &= \left(\nabla Z(m_a)\right)^T \cdot \dot{M}_a + N\left(m_a, u(m_a)\right) + \rho \hat{F}_a^T \hat{F}_a \\ &= \left(\nabla Z(m_a)\right)^T \cdot k_{deq}\left(\hat{f}(x, \hat{W}_f) + \hat{g}(x, \hat{W}_g)(u+\tau_c)\right) \\ &\quad + \left(\nabla Z(m_a)\right)^T \cdot \left(-k_{deq}\hat{g}(x, \hat{W}_g)F_a + v\right) \\ &\quad + N\left(m_a, u(m_a)\right) + \rho \hat{F}_a^T \hat{F}_a.\end{aligned} \tag{26}$$

The fault-tolerant position-force optimal control policy for modeless reconfigurable manipulators (9) can updated as

$$u_o^*(m_a) = -\frac{1}{2}k_{deq}R^{-1}\hat{g}^T(x, \hat{W}_g)\nabla Z(m_a). \tag{27}$$

D. ONLINE POLICY ITERATION ALGORITHM

This part introduces the online PI algorithm [43], [44] to solve the HJB equation. The online PI algorithm is composed of the policy evaluation (26) and the policy improvement (27).

Step1: Let $i = 0$, start with an arbitrary initial admissible control law $U_o^{(0)}(m_a)$, and a positive constant σ_P.

Step2: Let $i > 0$, solve $Z^{(i)}(m_a)$ based the control policy $U_o^{(i)}(m_a)$, from

$$0 = \rho \hat{F}_a^T \hat{F}_a + N\left(m_a, U_o^{(i)}(m_a)\right) + \left(\nabla Z^{(i+1)}(m_a)\right)^T \left(k_{deq}\left(\hat{f}(x, \hat{W}_f) + \hat{g}(x, \hat{W}_g)(U_o^{(i)}(m_a) + \tau_c - F_a)\right) + v\right),$$

with $Z^{(i+1)}(0) = 0$.

Step3: Update the control policy $U_o^{(i)}(m_a)$ through

$$U_o^{(i+1)}(m_a) = -\frac{1}{2} k_{deq} R^{-1} \hat{g}^T(x, \hat{W}_g) \nabla Z^{(i+1)}(m_a).$$

Step4: If $\left\|Z^{(i+1)}(m_a) - Z^{(i)}(m_a)\right\| \leq \sigma_P$, stop and get the optimal control strategy; else, let $i = i + 1$ and return to step 2.

This algorithm will converge to the optimal control strategy and the optimal performance index function, i.e. $Z^{(i)}(m_a) \to Z^*(m_a)$ and $U_o^{(i)}(m_a) \to u_o^*(m_a)$ as $i \to \infty$.

E. CRITIC NN IMPLEMENTATION AND STABILITY ANALYSIS

As we all know, NN has the ability to learn arbitrary functions. Aiming at the highly non-linear optimal control problems which cannot be solved by traditional control methods, the hidden layer neuron of multi-layer NN adopts activation function, which has the function of non-linear mapping. This mapping can approximate arbitrary linear/nonlinear function, and provide an effective way to solve the nonlinear control problems. For the aforementioned performance index function $Z(m_a)$, a critic NN $K_N(m_a)$ can be utilized to approximate:

$$K_N(m_a) = W_K^T \delta_K(m_a) + \varepsilon_K, \tag{28}$$

where $W_K \in \mathbb{R}^N$ is the desired weight vector and N is the number of neurons in the hidden-layer, $\delta_K(m_a)$ is the activation function, and ε_K is the NN approximation error. Thus, the partial derivative of $K_N(m_a)$ is

$$\nabla K_N(m_a) = (\nabla \delta_K(m_a))^T W_K + \nabla \varepsilon_K, \tag{29}$$

where $\nabla \delta_K(m_a) = \frac{\partial \delta_K(m_a)}{\partial m_a} \in \mathbb{R}^{N \times n}$ is the partial derivative of $\delta_K(m_a)$ and $\nabla \varepsilon_K$ is also the corresponding partial derivative.

For the constrained reconfigurable manipulator systems (24), combining (26) and (28), we get

$$0 = \rho \hat{F}_a^T \hat{F}_a + N(m_a, u(m_a)) + \left((\nabla \delta_K(m_a))^T W_K + \nabla \varepsilon_K\right)^T \dot{M}_a. \tag{30}$$

Therefore, the Hamiltonian function can be expressed

$$\begin{aligned} H(m_a, u(m_a), W_K) &= \rho \hat{F}_a^T \hat{F}_a + N(m_a, u(m_a)) \\ &\quad + \left((\nabla \delta_K(m_a))^T W_K\right)^T \dot{M}_a \\ &= -\nabla {\varepsilon_K}^T \dot{M}_a = e_{Kh}, \end{aligned} \tag{31}$$

where e_{Kh} is the NN approximation remnants error. The approximated critic NN can be calculated by

$$\hat{K}_N(m_a) = \hat{W}_K^T \delta_K(m_a). \tag{32}$$

Then the partial of derivative $\hat{K}_N(m_a)$ is expressed as

$$\nabla \hat{K}_N(m_a) = (\nabla \delta_K(m_a))^T \hat{W}_K. \tag{33}$$

Therefore, we can get the approximate HJB function as

$$\begin{aligned} H\left(m_a, u(m_a), \hat{W}_K\right) &= \rho \hat{F}_a^T \hat{F}_a + N(m_a, u(m_a)) \\ &\quad + \left((\nabla \delta_K(m_a))^T \hat{W}_K\right)^T \dot{M}_a \\ &= e_K. \end{aligned} \tag{34}$$

We minimized the objective function $E_K = \frac{1}{2} e_K^T e_K$ which is minimized by the gradient decent algorithm, to obtain the appropriate critic NN weight vector $\hat{W}_K$ which is updated by

$$\dot{\hat{W}}_K = -\alpha_K e_K \nabla \delta_K(m_a) \dot{M}_a, \tag{35}$$

where α_K is the updated rate of critic NN. Denote $h_\delta = \nabla \delta_K(m_a) \dot{M}_a$, and assume that there exist a positive constant $h_{\delta L}$ and $\|h_\delta\| \leq h_{\delta L}$.

Define the weight estimation error as

$$\tilde{W}_K = W_K - \hat{W}_K. \tag{36}$$

Through (31), (33) and (34), we obtain

$$e_K = e_{Kh} - \tilde{W}_K^T h_\delta. \tag{37}$$

The weight estimation error can be updated by

$$\dot{\tilde{W}}_K = -\dot{\hat{W}}_K = \alpha_K e_K h_\delta = \alpha_K \left(e_{Kh} - \tilde{W}_K^T h_\delta\right) h_\delta. \tag{38}$$

According to (27) and (29), the desired optimal control policy can be described as

$$u_o^*(m_a) = -\frac{1}{2} k_{deq} R^{-1} \hat{g}^T(x, \hat{W}_g) \left((\nabla \delta_K(m_a))^T W_K + \nabla \varepsilon_K\right). \tag{39}$$

Therefore, the approximated fault-tolerant position-force optimal control law $\hat{u}_o^*$ as the control input torque of the developed optimal control method, is given as

$$\hat{u}_o^*(m_a) = -\frac{1}{2} k_{deq} R^{-1} \hat{g}^T(x, \hat{W}_g) (\nabla \delta_K(m_a))^T \hat{W}_K, \tag{40}$$

and the structural diagram of the proposed fault-tolerant position-force optimal control strategy is illustrated in Fig(1).

Remark 5: Unlike the existing researches which presented the fault-tolerant position-force controllers for manipulators without considering the optimal performance and the energy consumption. In this study, the improved optimal performance index function contains the position-force errors fusion equation and uncertain actuator failures which obtained by adaptive fault observer. Moreover, according to the ADP approach, one presents a novel optimal fault-tolerant position-force control method which optimizes the system performance and enhances the system stability.

Theorem 2: Consider the constrained reconfigurable manipulator system with uncertain actuator failures, if the dynamic model formulation is presented in (24), the weight of the critic NN is updated by (35), then the weight approximation error is UUB.

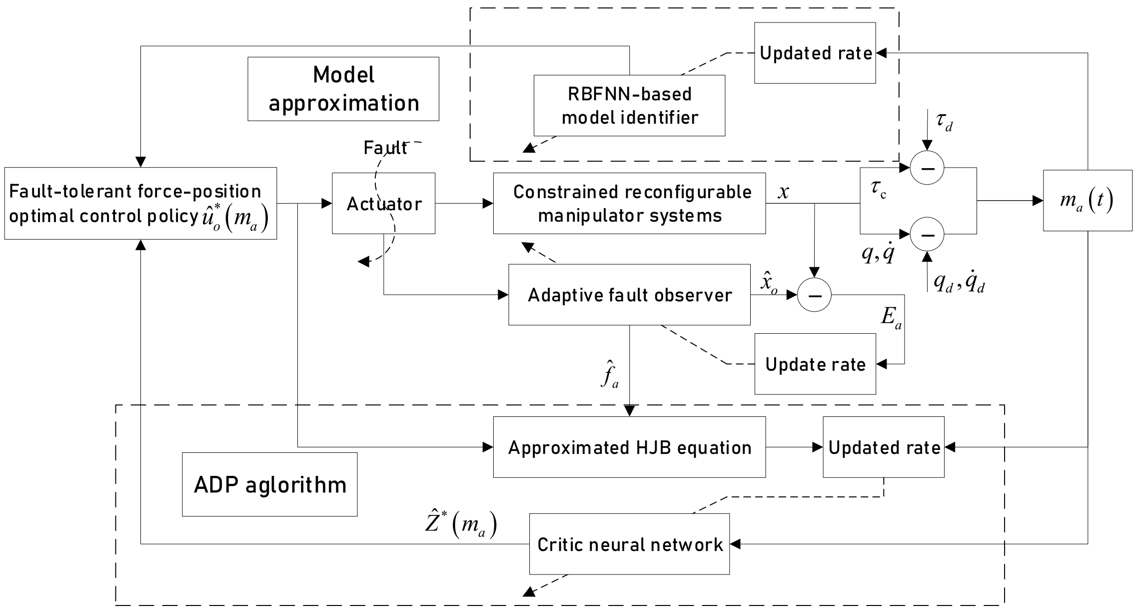


FIGURE 1. **Structural diagram of the proposed optimal control method.**

Proof: Choose the Lyapunov candidate as:

$$V_2(t) = \frac{1}{2\alpha_K}\tilde{W}_K^T \tilde{W}_K. \tag{41}$$

The time derivative of $V_2(t)$ is taken as:

$$\begin{aligned} \dot{V}_2(t) &= \frac{1}{\alpha_K}\tilde{W}_K^T \dot{\tilde{W}}_K \\ &= \tilde{W}_K^T \left(e_{Kh} - \tilde{W}_K^T h_\delta\right) h_\delta \\ &= \tilde{W}_K^T e_{Kh}\upsilon - \left\|\tilde{W}_K^T h_\delta\right\|^2 \\ &\le -\frac{1}{2}\left(\left\|\tilde{W}_K^T h_\delta\right\|^2 - \|e_{Kh}\|^2\right). \end{aligned} \tag{42}$$

Therefore, one can observe that $\dot{V}_2(t) \le 0$ with the compact set $\Omega_2 = \left\{\tilde{W}_K \middle| \left\|\tilde{W}_K\right\| \le \left\|\frac{e_{Kh}}{h_{\delta L}}\right\|\right\}$, and the weight approximation error is verified as UUB.

F. STABILITY ANALYSIS

In this section, we discuss the stability issue of the closed-loop reconfigurable manipulator systems with the constrained environment contacts under the proposed fault-tolerant position-force optimal control policy (40), and the theorem is given as:

Theorem 3: Consider an n-DOF constrained reconfigurable manipulator with the NN-estimated dynamics formulated in (24), the closed-loop manipulator system is asymptotically stable via the fault-tolerant position-force optimal control law proposed (40). Even if actuator failure occurs, the reconfigurable manipulator system is possessed of certain fault tolerance and robustness, that is, the position tracking error and the torque deviation of constrained reconfigurable manipulator can converge to zero asymptotically.

Proof: Select the Lyapunov function candidate as

$$V_3(t) = \frac{1}{2}{m_a}^T m_a + Z^*(m_a) + \frac{1}{2}\tilde{W}_f^T \alpha_f^{-1}\tilde{W}_f + \frac{1}{2}\tilde{W}_g^T \alpha_g^{-1}\tilde{W}_g. \tag{43}$$

Its time derivative as

$$\begin{aligned} \dot{V}_3(t) &= {m_a}^T \cdot \left(k_{deq}\left(\hat{f}\left(x, \hat{W}_f\right) + \hat{g}\left(x, \hat{W}_g\right)(u + \tau_c)\right)\right) \\ &\quad - {m_a}^T \cdot \left(k_{deq}\hat{g}\left(x, \hat{W}_g\right) F_a + v\right) - {m_a}^T Q m_a \\ &\quad - u^T R u - \rho \hat{F}_a^T \hat{F}_a + \tilde{W}_f^T\left(k_{deq}{m_a}^T \delta_f(x)\right) \\ &\quad + \tilde{W}_g^T\left(k_{deq}{m_a}^T \delta_g(x)\left((u + \tau_c) - \hat{F}_a\right)\right). \end{aligned} \tag{44}$$

According the model-free NN weight approximation updated equations (18) (19), assuming $\ddot{q}_d \le \sigma$, $\dot{e}_q \le \xi$, $e_\tau \le \eta$, by Young's inequality, (44) is updated as

$$\begin{aligned} \dot{V}_3(t) &\le k_{deq}\Upsilon_f \|m_a\|^2 + k_{deq}\varpi_g \|m_a\| \|u\| + k_{deq}\varpi_g \tau_c \|m_a\| \\ &\quad + v\|m_a\| - k_{deq}\varpi_g \|m_a\| \|F_a\| - {m_a}^T Q m_a \\ &\quad - u^T R u - \rho \hat{F}_a^T \hat{F}_a \\ &\le k_{deq}\Upsilon_f \|m_a\|^2 + \frac{1}{2}k_{deq}^2 \|m_a\|^2 + \frac{1}{2}\varpi_g^2 \|u\|^2 \\ &\quad + k_{deq}\varpi_g \tau_c \|m_a\| + v\|m_a\| + \frac{1}{2}k_{deq}^2 \|m_a\|^2 \\ &\quad + \frac{1}{2}\varpi_g^2 \|F_a\|^2 - \lambda_{\min}(Q)\|m_a\|^2 - \lambda_{\min}(R)\|u\|^2 \\ &\quad - \rho \hat{F}_a^T \hat{F}_a \\ &\le -\left(-k_{deq}\varpi_g \tau_c - v\right)\|m_a\| - \left(\lambda_{\min}(R) \right. \\ &\quad \left. - \frac{1}{2}\varpi_g^2\right)\|u\|^2 \\ &\quad - \left(\lambda_{\min}(Q) - k_{deq}\Upsilon_f - k_{deq}^2\right)\|m_a\|^2 \end{aligned}$$

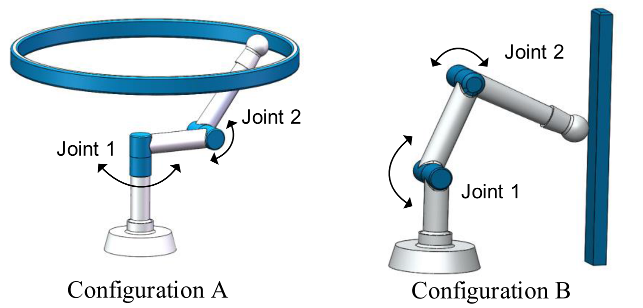


FIGURE 2. 2-DOF reconfigurable manipulators with different configurations for simulation.

$$-\frac{1}{2}\varpi_g^2\left(2\|F_a\|-\left\|F_a-\hat{F}_a\right\|\right)\left(\left\|F_a-\hat{F}_a\right\|\right)$$
$$-\left(\rho-\frac{1}{2}\varpi_g^2\right)\hat{F}_a^T\hat{F}_a. \quad (45)$$

Define $\Phi_1=\lambda_{\min}(Q)-k_{deq}\Upsilon_f-k_{deq}^2$, $\Phi_2=\lambda_{\min}(R)-\frac{1}{2}\varpi_g^2$, and $|\tau_c|\le|\tau_M|$ on account of $|f_c|\le f_M$,define $e_{fM}=\varpi_g\tau_M$. Then

$$\dot{V}_3(t)\le-\left(\Phi_1\|m_a\|-k_{deq}e_{fM}-k_{deq}\sigma\right)\|m_a\|$$
$$-(-k_{eq}\xi-k_\tau\eta)\|m_a\|-\left(\rho-\frac{1}{2}\varpi_g^2\right)\hat{F}_a^T\hat{F}_a$$
$$-\Phi_2\|u\|^2-\frac{1}{2}\varpi_g^2(2\varepsilon_1-\|e_F\|)\|e_F\|. \quad (46)$$

Therefore, we can obtain that $\dot{V}_3(t)\le 0$ when m_a lies outside the compact set

$$\Omega_3=\left\{m_a:\|m_a\|\ge\frac{k_{deq}e_{fM}+k_{deq}\sigma+k_{eq}\xi+k_\tau\eta}{\Phi_1}\right\}$$

and $\|e_F\|\le 2\varepsilon_1$, if the following conditions hold:

$$\begin{cases}\lambda_{\min}(Q)\ge k_{deq}\Upsilon_f+k_{deq}^2\\ \lambda_{\min}(R)\ge\frac{1}{2}\varpi_g^2\\ \rho\ge\frac{1}{2}\varpi_g^2.\end{cases} \quad (47)$$

IV. SIMULATION

In this article, two 2-DOF constrained reconfigurable manipulators with different configurations (see Fig.2) are used for verifying the effectiveness of the fault-tolerant position-force optimal control strategy based on ADP algorithm. The dynamics of constrained reconfigurable manipulators with uncertain actuator failures is selected by referencing our previous work (24), which can be transformed into a form of analytic charts, that is shown in Fig.3 under the sake of analysis configurations aforementioned. The constrained reconfigurable manipulator system parameters, which include the uncertainty up-bound parameters and the control parameters, are represented in Table 1.

Considering the constrained reconfigurable manipulators with two different configurations (see Fig.2), which are assumed that operating in some special constrained environments such as wheels or cylinders (Configuration A). The constrained equation and dynamic model for Configuration A of reconfigurable manipulators can be defined as: $\Phi_A(q)=l_1+l_2\cos(q_2)-1.5=0$ where $l_1=1$ and $l_2=1$ are the lengths of the two manipulator links.

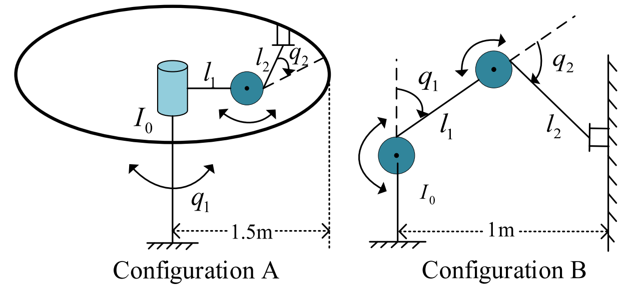


FIGURE 3. The analytic charts of 2-DOF reconfigurable manipulators with different configurations for simulation.

TABLE 1. Parameters setting.

Parameter type	Name	Value	Name	Value
Uncertainty up-bound parameters	ε_1	$8Nm$	ε_2	$0.5rad/s$
	f_M	$15N$	τ_M	$10Nm$
	Υ_f	$2.5s^2/rad^2$	ϖ_g	$3s/rad$
	$h_{\delta L}$	0.5	σ	$5rad^2/s^2$
	ξ	$0.5rad/s$	η	$2Nm$
Control parameters	ρ	0.02	α_1	18
	α_2	15	α_K	0.027
	α_f	0.48	α_g	2.3
	Q	$0.19I_{2\times1}$	R	$0.34I_{2\times1}$
	α_F	$0.05I_{2\times1}$		

The desired trajectory and anticipated contact force of Configuration A are written as follows: $q_{1d}=\sin(2t)+0.2\cos(t)$, $q_{2d}=\frac{\pi}{3}$, $f_d=5N$. The presupposed actuator failure is $F_a=\left[F_{a1}\ F_{a2}\right]^T$ for 2-DOF constrained reconfigurable manipulator of Configuration A, among $F_{a1}=0$ and $F_{a2}=\begin{cases}0 & ,t<30s\\ 3\sin(0.2t)+\cos(t) & ,t\ge30s\end{cases}$. In this article, we utilized the RBFNN to estimate the modeless dynamics, and the NN to approximate the improved optimal performance index function. For the model dynamics, we choose the 1-3-1 RBFNN structure, in which 1 input, 3 hidden layers, and 1 output, for each joint model function. The RBFNN weights are defined as $\hat{W}_f=\left[\hat{W}_{f11},\hat{W}_{f12},\hat{W}_{f13},\hat{W}_{f21},\hat{W}_{f22},\hat{W}_{f23}\right]^T$ and $\hat{W}_g=\left[\hat{W}_{g11},\hat{W}_{g12},\hat{W}_{g13};\hat{W}_{g21},\hat{W}_{g22},\hat{W}_{g23}\right]^T$ in this simulation part, with the initial value $\hat{W}_{f0}=\hat{W}_{g0}=[1,1,1,1,1,1]^T$ for every weights. The radial basis function is chosen as the activation function $\delta_f(x)$ and $\delta_g(x)$ in (16) (17), represented as $\delta_f(x)=\delta_g(x)=\exp\left(-\frac{\|x-c_j\|^2}{2b_j^2}\right)$, with $b_j=1.5$, $j=1,2,3$. For the critic NN approximator, the weight vector of critic NN can be defined as $\hat{W}_K=[\hat{W}_{K1},\hat{W}_{K2},\hat{W}_{K3}]^T$

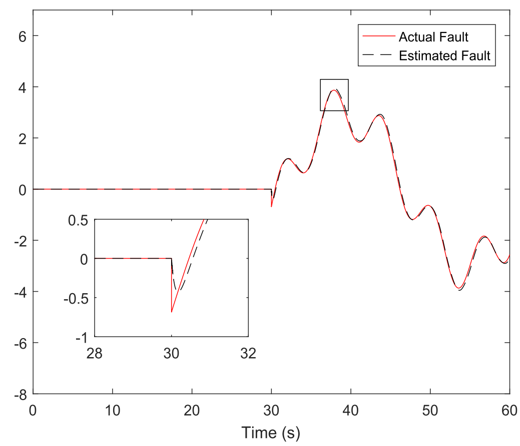


FIGURE 4. **The estimation of the uncertain actuator failure based on the proposed controller for Configuration A.**

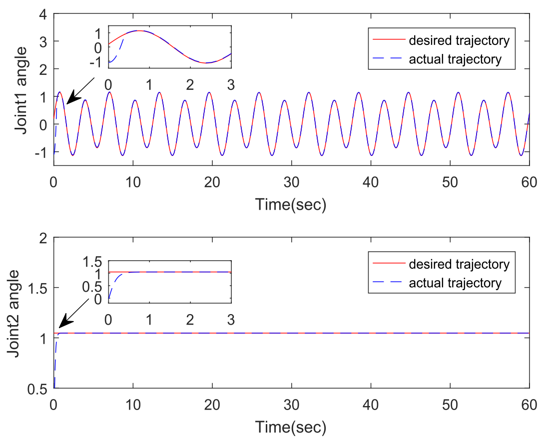


FIGURE 5. **Trajectory tracking curves of the proposed controller for Configuration A.**

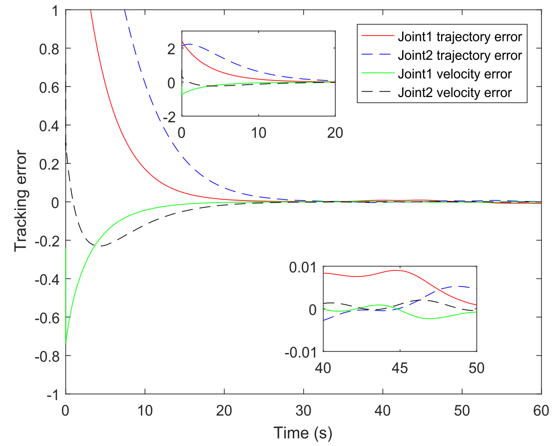


FIGURE 6. **Trajectory tracking errors and velocity tracking errors curves of the exist controller for Configuration A.**

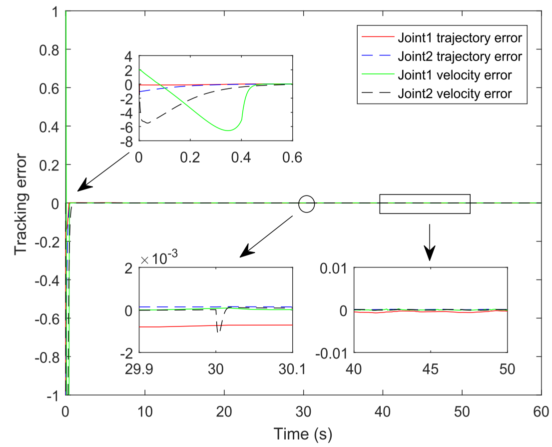


FIGURE 7. **Trajectory tracking errors and velocity tracking errors curves of the proposed controller for Configuration A.**

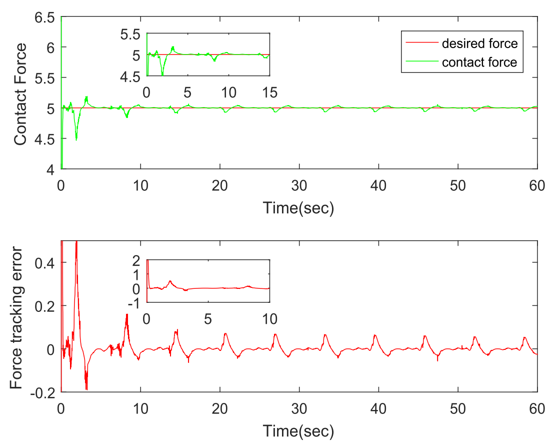


FIGURE 8. **Force tracking curves and force tracking error of the exist controller for Configuration A.**

with their initial value and the weights update rate are given as $\hat{W}_{K0} = [20\ 25\ 30]^T$ and the activation functions for critic NN are selected with three neurons as $\sigma = [m_{a1}^2, m_{a1}m_{a2}, m_{a2}^2]$, where m_{a1} and m_{a2} are position-force deviation fusion function of joint 1 and joint 2, and select $k_{deq} = 1.1$, $k_{eq} = 6.4$, $k_\tau = 4.1$ for Configuration A.

The simulation results include estimated actuator failure, joint position tracking, force tracking, and force tracking deviation of reconfigurable manipulators with different configurations in constrained environment. Two different position-force control schemes are employed in the simulation that contain the traditional robust position-force control method, e.g. [19], and the proposed fault-tolerant position-force optimal control method with adaptive fault observer based on PI scheme and ADP algorithm.

Fig.4-Fig.12 are the simulation results of Configuration A. The estimated value of actuator failure $\hat{F}_{a2}$ is shown as Fig.4, and the red and black dashed lines represent the actual failure and the estimated failure, respectively. One can observe that after the failure occurs at t = 30s, owing to the knowledge of uncertain fault function obtained from proposed adaptive fault observer, then, the estimated failure keeps up with actual failure smoothly in less than one second.

Fig.5 is trajectory tracking curves for Configuration A under the proposed fault-tolerant position-force optimal control policy, and the red and blue dashed lines represent

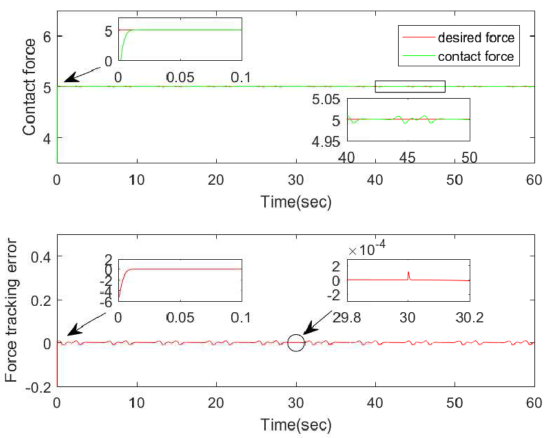


FIGURE 9. Force tracking curves and force tracking error of the proposed controller for Configuration A.

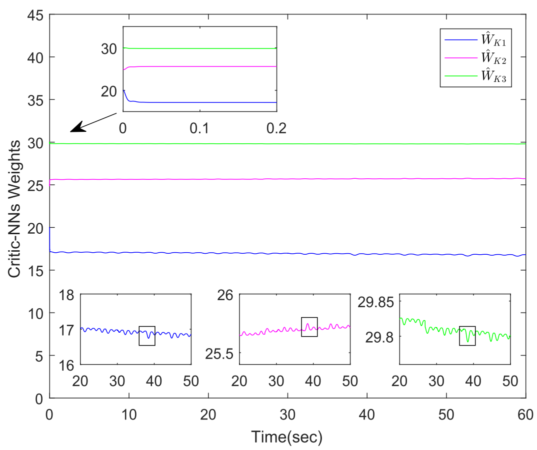


FIGURE 10. Critic NN weights adjustment curves of the proposed controller for Configuration A.

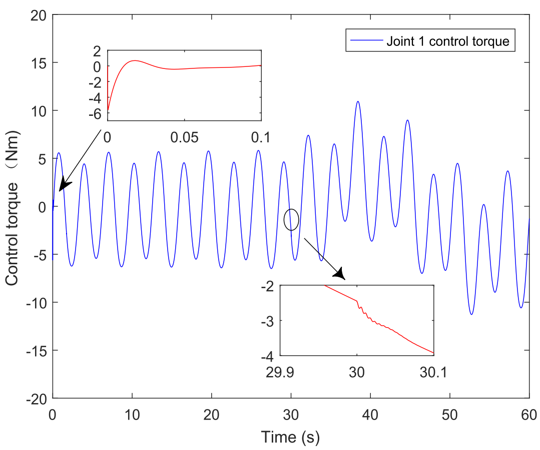


FIGURE 11. Control input torque of fault-tolerant position-force optimal controller for Configuration A-Joint 1.

practical trajectory and desired trajectory respectively. In this picture, one observes that actual joint position can tracking the desired trajectories around the 1s smoothly and stably. Fig.6 and Fig.7 show the tracking errors curves including the trajectory tracking errors and velocity tracking errors

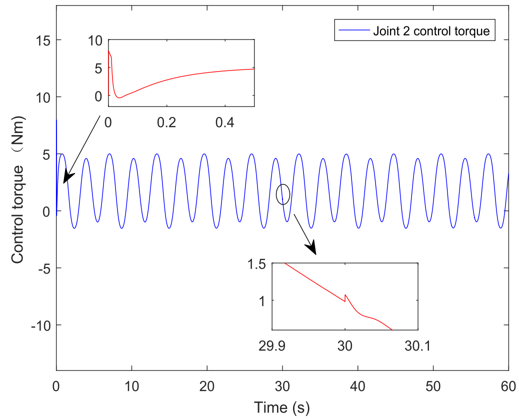


FIGURE 12. Control input torque of fault-tolerant position-force optimal controller for Configuration A-Joint 2.

of Configuration A with the proposed control method and the previous exist one [18], [34], respectively. In Fig.6, one observes that the tracking errors tend to stable at 30s and the stable errors values around 0.01. Comparing with the exist control method, the errors adjustment time of the proposed method is about 0.5s and the errors are stable at less than 2e-3 even after the actuator failure in Fig.7. Shown in Fig.8, the force of the end-effector based on the exist control method is continuous jitter around within ± 0.2 N and with almost 20s of adjustment time to the stable state. However, the precision of force control with the exist controller for manipulators is unacceptable under some special tasks, such as the assistant surgery manipulators, the search and rescue robots, etc. Comparing with Fig.8, one observes that the force tracking curves is gradually stable at 1s and the force tracking precision is less than $\pm 2e - 3$N in Fig.9. And the contact force deviation is in an extremely small range $\pm e - 4$N when the actuator failure occurred t = 30s, since the uncertain actuator failure considered as one of performance indictor is compensated. Given as Fig.10, the critic NN weights converge to $\hat{W}_K = [17.1, 25.7, 29.8]^T$ with the fluctuated value not over 0.1, and stable around 0.1 seconds smoothly, even under the uncertain actuator failure after 30 seconds.Comparing with Fig.4 and Fig.10, the NN weights are constantly changing in the small range after actuator failure occurs, thereby ensures the stable operation of the system under the approximated optimal control torque.

Fig.11 and Fig.12 indicate the control torque curves of the developed control method for Configuration A in this article. From these figures, one draws that the presented control torque curves can be quickly stabilized at about ± 5N within 0.2s. After the fault occurred on the joint 1, the control torque compensated the influence of uncertain failures smoothly and the instantaneous impact on the actuator is less than 0.1N at 30s. Through the analysis of the simulation results of Configuration A, the presented fault-tolerant position-force optimal control method could not only improve the control accuracy and the ability of active fault tolerance, but also effectively stabilize and reduce the output of actuator.

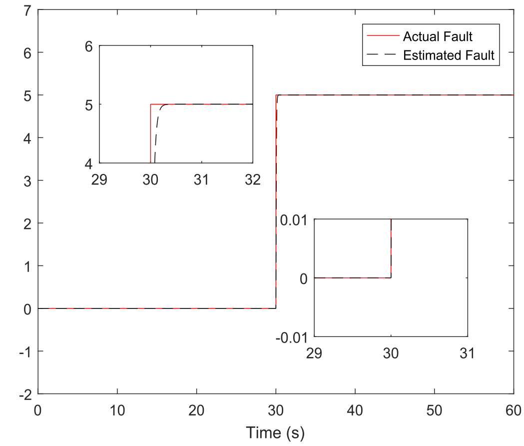


FIGURE 13. The estimation of the uncertain actuator failure based on the proposed controller for Configuration B.

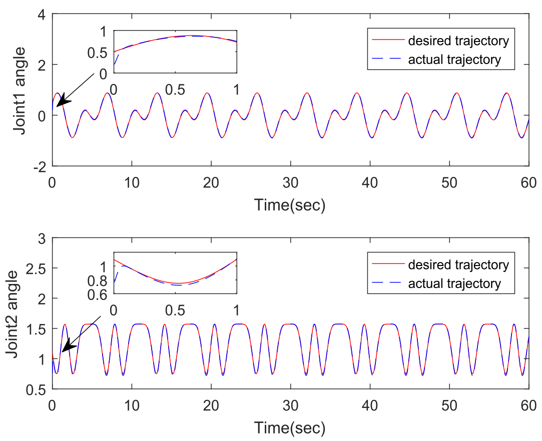


FIGURE 14. Trajectory tracking curves of proposed fault-tolerant position-force optimal controller for Configuration B.

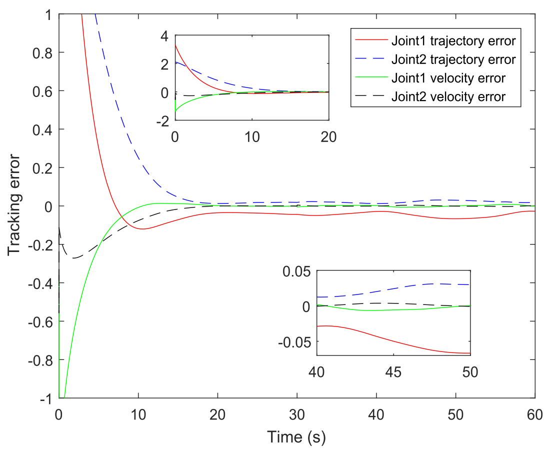


FIGURE 15. Trajectory tracking errors and velocity tracking errors curves of the exist controller for Configuration B.

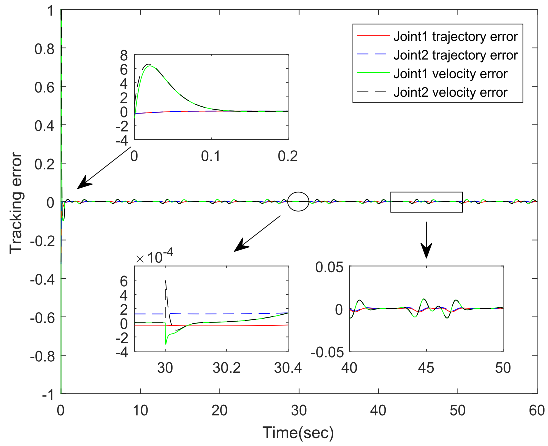


FIGURE 16. Trajectory tracking errors and velocity tracking errors curves of the proposed controller for Configuration B.

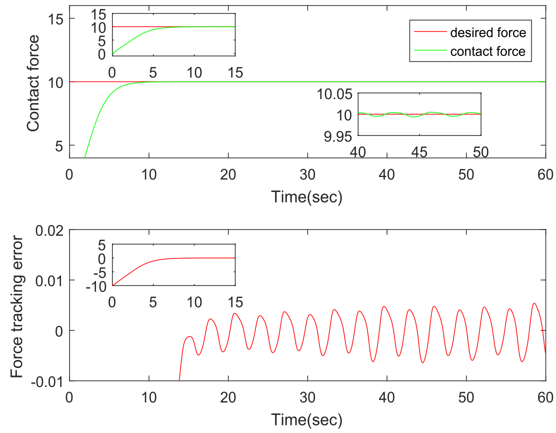


FIGURE 17. Force tracking curves and force tracking error of the exist controller for Configuration B.

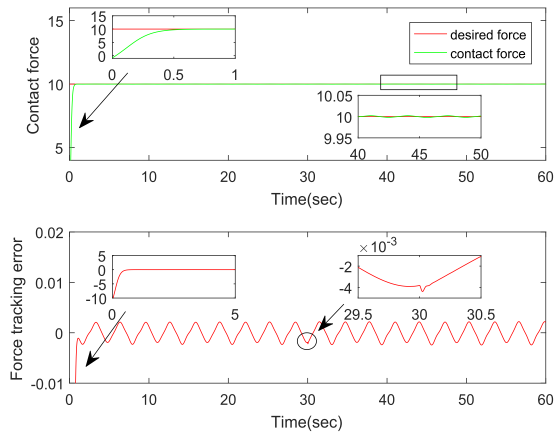


FIGURE 18. Force tracking curves and force tracking error of the proposed controller for Configuration B.

Considering the reconfigurable manipulators with Configurations B (see Fig.2), which working in some special constrained environments such some operation tasks such as polishing or wiping. The constrained equation and dynamic model for Configuration A of reconfigurable manipulators can be defined as: $\Phi_B(q) = l_1 \cos(q_1) + l_2 \cos(q_2) - 1 = 0$ where $l_1 = 1$ and $l_2 = 1$ are the lengths of the two manipulator links.

The desired trajectory and anticipated contact force of Configuration A are written as follows: $q_{1d} = \cos(t) + 0.5\sin(2t)$, $q_{2d} = \arccos\left(\frac{1 - l_1\cos(\cos(t) + 0.5\sin(2t))}{l_2}\right)$,

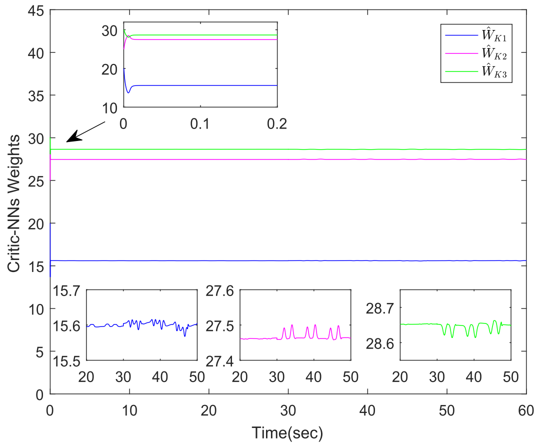


FIGURE 19. Critic NN weights adjustment curves of the proposed controller for Configuration B.

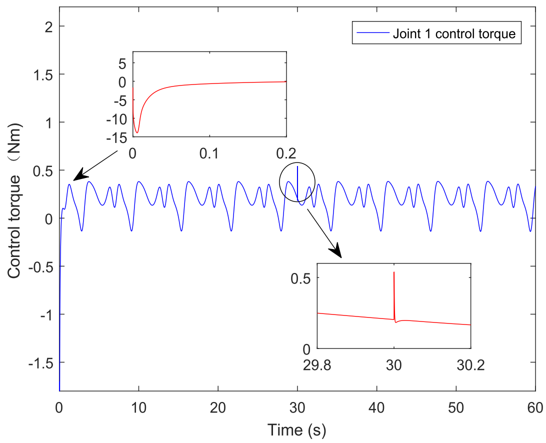


FIGURE 20. Control input torque of fault-tolerant position-force optimal controller for Configuration B-Joint 1.

$f_d = 10N$. The presupposed actuator failure is $F_a = [F_{a1}\ F_{a2}]^T$ for 2-DOF constrained reconfigurable manipulator of Configuration B, among $F_{a1} = \begin{cases} 0, t < 30s \\ 5, t \geq 30s \end{cases}$, $F_{a2} = 0$. And we select $\hat{W}_{K0} = [20\ 25\ 30]^T$, $k_{deq} = 0.95$, $k_{eq} = 3.34$, $k_\tau = 1.45$ for Configuration B.

Fig.13-Fig.21 show the estimated uncertain actuator failure, the joint position tracking, the tracking errors, the force tracking, the force tracking deviation, NN weights adjustment, and the control input torque of reconfigurable manipulators with Configuration B, and the critic NN weights converge to $\hat{W}_K = [15.6, 27.47, 28.65]^T$ with the small range fluctuation. From these pictures, one draws the same clear conclusion as the situation of Configuration A, that the proposed fault-tolerant position-force optimal control method is applicable to different configurations of reconfigurable manipulators under uncertain actuator failures.

By observing the simulation results of Configuration A and Configuration B, we can summarize that the proposed fault-tolerant position-force optimal control policy is effective and

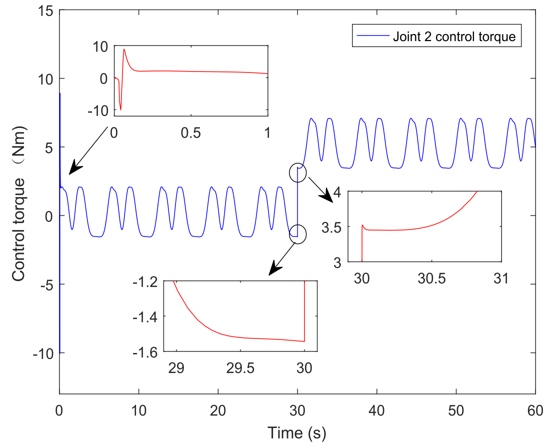


FIGURE 21. Control input torque of fault-tolerant position-force optimal controller for Configuration B-Joint 2.

stable for reconfigurable manipulators to satisfy the requirements of various tasks in constraint environment.

V. CONCLUSION

This article addresses a novel fault-tolerant position-force optimal control problem of the constrained reconfigurable manipulators under the uncertain actuator failures. With the help of the RBFNN-based dynamic model of constrained reconfigurable manipulators, a novel performance index function is constructed including position error, torque deviation and estimated uncertain actuator failures. Then the PI algorithm is utilized to attain the optimal control strategy, a critic NN is constructed to solve the improved HJB equation, and the approximated fault-tolerant position-force optimal control torque can be derived directly. The Lyapunov theory is utilized to prove the asymptotic stability of the closed-loop robotic systems. By comparing the traditional position-force robust controller with simulation study, verified the effectiveness and accuracy of the proposed fault-tolerant position-force optimal control method.

REFERENCES

[1] M. Fujita, H. Kitano, and K. Kageyama, "A reconfigurable robot platform," *Robot. Auto. Syst.*, vol. 29, nos. 2–3, pp. 119–132, Nov. 1999.

[2] M. Benosman and K.-Y. Lum, "Passive Actuators' fault-tolerant control for affine nonlinear systems," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 1, pp. 152–163, Jan. 2010.

[3] L. Su, X. Zhu, and J. Qiu, "Robust passive fault-tolerant control for uncertain non-linear stochastic systems with distributed delays," in *Proc. 29th Chin. Control Conf.*, Beijing, China, 2010, pp. 2024–2028.

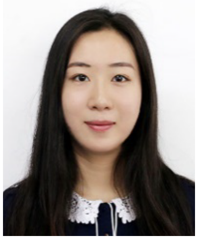
[4] K. Zhao and J. Chen, "Robust adaptive fault-tolerant control of nonlinear systems with assignable tracking performance under time-varying control coefficients," *IEEE Access*, vol. 6, pp. 40757–40767, Jul. 2018.

[5] X. Zhu, H. Zhang, B. Yang, and G. Zhang, "Cloud-based shaft torque estimation for electric vehicle equipped with integrated motor-transmission system," *Mech. Syst. Signal Process.*, vol. 99, no. 15, pp. 647–660, Jan. 2018.

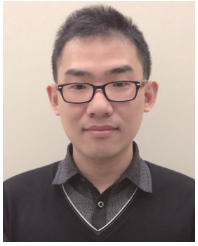
[6] Y. Li, F. Zhou, and B. Zhao, "Dynamic output feedback based active decentralized fault-tolerant control for reconfigurable manipulator with concurrent failures," *Math. Problems Eng.*, vol. 2015, pp. 1–14, Mar. 2015.

[7] W. Guo, A. Qiu, and C. Wen, "Active actuator fault tolerant control based on generalized internal model control and performance compensation," *IEEE Access*, vol. 7, pp. 175514–175521, Dec. 2019.

[8] X. Zhu and W. Li, "Takagi–Sugeno fuzzy model based shaft torque estimation for integrated motor–transmission system," *ISA Trans.*, vol. 93, pp. 14–22, Oct. 2019.
[9] 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, Jul. 2011.
[10] 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, Sep. 2013.
[11] F. Wang, Z. Luo, H. Liu, and L. Wang, "Impedance model based fuzzy force control for robot manipulator contacting with a constrained surface with uncertain errors," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Tianjin, China, Dec. 2010, pp. 1555–1558.
[12] M. Raibert and J. Craig, "Hybrid position/force control of manipulators," *J. Dyn. Syst., Meas. Control*, vol. 103, no. 2, pp. 126–133, Dec. 1980.
[13] N. Kumar, V. Panwar, and N. Sukavanam, "Neural network based hybrid force/position control for robot manipulators," *Int. J. Precis. Eng. Manuf.*, vol. 12, no. 3, pp. 419–426, Jun. 2011.
[14] S. Chiaverini and L. Sciavicco, "The parallel approach to force/position control of robotic manipulators," *IEEE Trans. Robot. Autom.*, vol. 9, no. 4, pp. 361–373, Aug. 1993.
[15] Z. Doulgeri and Y. Karayiannidis, "Force position control for a robot finger with a soft tip and kinematic uncertainties," *Robot. Auto. Syst.*, vol. 55, no. 4, pp. 328–336, Apr. 2007.
[16] J. Pliego-Jimenez and M. A. Arteaga-Perez, "Adaptive position/force control for robot manipulators in contact with a rigid surface with unknown parameters," in *Proc. Eur. Control Conf. (ECC)*, Linz, Austria, Jul. 2015, pp. 3603–3608.
[17] Y. Farid, V. J. Majd, and A. Ehsani-Seresht, "Fractional-order active fault-tolerant force-position controller design for the legged robots using saturated actuator with unknown bias and gain degradation," *Mech. Syst. Signal Process.*, vol. 104, no. 1, pp. 465–486, May 2018.
[18] X. Bu, X. Wu, F. Zhu, J. Huang, Z. Ma, and R. Zhang, "Novel prescribed performance neural control of a flexible air-breathing hypersonic vehicle with unknown initial errors," *ISA Trans.*, vol. 59, pp. 149–159, Nov. 2015.
[19] X. Bu, X. Wu, D. Wei, and J. Huang, "Neural-approximation-based robust adaptive control of flexible air-breathing hypersonic vehicles with parametric uncertainties and control input constraints," *Inf. Sci.*, vols. 346–347, pp. 29–43, Jun. 2016.
[20] F. Zhou, Y. Li, and G. Liu, "Robust decentralized force/position fault-tolerant control for constrained reconfigurable manipulators without torque sensing," *Nonlinear Dyn.*, vol. 89, no. 2, pp. 955–969, Apr. 2017.
[21] P. Werbos, "Approximate dynamic programming for real-time control and neural modeling," in *Handbook of Intelligent Control Neural Fuzzy and Adaptive Approaches*. New York, NY, USA: Van Nostrand, 1992, ch. 13, pp. 493–525.
[22] D. Wang, D. Liu, and Q. Wei, "Finite-horizon neuro-optimal tracking control for a class of discrete-time nonlinear systems using adaptive dynamic programming approach," *Neurocomputing*, vol. 78, no. 1, pp. 14–22, Feb. 2012.
[23] Y. Huang and D. Liu, "Finite-horizon neuro-optimal tracking control for a class of discrete-time nonlinear systems using adaptive dynamic programming approach," *Neurocomputing*, vol. 125, no. 11, pp. 46–56, Feb. 2014.
[24] Q. Wei, D. Liu, and H. Lin, "Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 840–853, Mar. 2016.
[25] X. Yang, D. Liu, Q. Wei, and D. Wang, "Guaranteed cost neural tracking control for a class of uncertain nonlinear systems using adaptive dynamic programming," *Neurocomputing*, vol. 198, pp. 80–90, Jul. 2016.
[26] L. Dong, X. Zhong, C. Sun, and H. He, "Adaptive event-triggered control based on heuristic dynamic programming for nonlinear discrete-time systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 7, pp. 1594–1605, Jul. 2017.
[27] Q.-M. Sun, C. Zhang, N.-Y. Jiang, J.-J. Yu, and L. Xu, "Data-driven nonlinear near-optimal regulation based on multi-dimensional Taylor network dynamic programming," *IEEE Access*, vol. 8, pp. 36476–36484, Feb. 2020.
[28] B. Luo, H.-N. Wu, and H.-X. Li, "Data-based suboptimal neuro-control design with reinforcement learning for dissipative spatially distributed processes," *Ind. Eng. Chem. Res.*, vol. 53, no. 19, pp. 8106–8119, May 2014.
[29] Y. Jiang and Z.-P. Jiang, "Robust adaptive dynamic programming for large-scale systems with an application to multimachine power systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 10, pp. 693–697, Oct. 2012.
[30] B. Zhao, D. Liu, and C. Luo, "Reinforcement learning-based optimal stabilization for unknown nonlinear systems subject to inputs with uncertain constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 27, 2019, doi: 10.1109/TNNLS.2019.2954983.
[31] D. Wang, D. Liu, and H. Li, "An approximate optimal control approach for robust stabilization of a class of discrete-time nonlinear systems with uncertainties," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 5, pp. 713–717, May 2016.
[32] X. Yang, D. Liu, and D. Wang, "Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints," *Int. J. Control*, vol. 87, no. 3, pp. 553–566, Mar. 2014.
[33] Y. Luo and G. Xiao, "ADP-based optimal control for a class of nonlinear discrete-time systems with inequality constraints," in *Proc. IEEE Symp. Adapt. Dyn. Program. Reinforcement Learn. (ADPRL)*, Orlando, FL, USA, Dec. 2014, pp. 1–5.
[34] H. Xia, B. Zhao, F. Zhou, B. Dong, G. Liu, and Y. Li, "Fault tolerant control for reconfigurable manipulators based on adaptive dynamic programming with an improved performance index function," in *Proc. 12th World Congr. Intell. Control Autom. (WCICA)*, Guilin, China, Dec. 2014, pp. 2273–2278.
[35] X. Xiao and X.-J. Li, "Adaptive dynamic programming method-based synchronisation control of a class of complex dynamical networks with unknown dynamics and actuator faults," *IET Control Theory Appl.*, vol. 12, no. 2, pp. 291–298, Jan. 2018.
[36] S. Li, L. Ding, H. Gao, C. Chen, Z. Liu, and Z. Deng, "Adaptive neural network tracking control-based reinforcement learning for wheeled mobile robots with skidding and slipping," *Neurocomputing*, vol. 283, no. 29, pp. 20–30, Mar. 2018.
[37] B. Dong, T. An, F. Zhou, K. Liu, and Y. Li, "Decentralized robust zero-sum neuro-optimal control for modular robot manipulators in contact with uncertain environments: Theory and experimental verification," *Nonlinear Dyn.*, vol. 97, no. 1, pp. 503–524, May 2019.
[38] B. Zhao, D. Liu, and Y. Li, "Online fault compensation control based on policy iteration algorithm for a class of affine non-linear systems with actuator failures," *IET Control Theory Appl.*, vol. 10, no. 15, pp. 1816–1823, Oct. 2016.
[39] M.-H. Ghajar, M. Keshmiri, and J. Bahrami, "Neural-network-based robust hybrid force/position controller for a constrained robot manipulator with uncertainties," *Trans. Inst. Meas. Control*, vol. 40, no. 5, pp. 1625–1636, Feb. 2017.
[40] J. Duan, Y. Gan, M. Chen, and X. Dai, "Adaptive variable impedance control for dynamic contact force tracking in uncertain environment," *Robot. Auto. Syst.*, vol. 102, pp. 54–65, Apr. 2018.
[41] H. P. Singh and N. Sukavanam, "Stability analysis of robust adaptive hybrid position/force controller for robot manipulators using neural network with uncertainties," *Neural Comput. Appl.*, vol. 22, nos. 7–8, pp. 1745–1755, Jun. 2012.
[42] H. Xia, B. Zhao, and Y. Li, "Policy iteration algorithm based fault tolerant tracking control and its application to reconfigurable manipulators," *J. Electr. Eng. Technol.*, vol. 13, no. 4, pp. 1739–1750, Jul. 2018.
[43] Z. Huang, C. Zhang, Y. Zhang, and G. Zhang, "A data-driven online ADP control method for nonlinear system based on policy iteration and nonlinear MIMO decoupling ADRC," *Neurocomputing*, vol. 310, no. 8, pp. 28–37, Oct. 2018.
[44] Q. Wei and D. Liu, "A novel policy iteration based deterministic Q-learning for discrete-time nonlinear systems," *Self-Learn. Optim. Control Nonlinear Syst.*, vol. 103, pp. 85–109, Jun. 2018.



BING MA received the B.S. degree from the Changchun University of Technology, China, in 2016, where she is currently pursuing the Ph.D. degree with the Department of Mechanical Engineering. Her research interests include robot control and adaptive dynamic programming.



BO DONG received the M.S. and Ph.D. degrees from Jilin University, China, in 2012 and 2015, respectively. He is currently an Associate Professor with the Changchun University of Technology, China. He is also a Postdoctoral Fellow with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include intelligent mechanical and robot control.

FAN ZHOU received the M.S. and Ph.D. degrees from the Changchun University of Technology, China, in 2015 and 2018, respectively. She is currently a Lecturer with the Changchun University of Technology. Her research interests include intelligent mechanical, robot control, and robust control.

YUANCHUN LI received the M.S. and Ph.D. degrees from the Harbin Institute of Technology, China, in 1987 and 1990, respectively. He is currently a Professor with the Changchun University of Technology, China. His research interests include complex system modeling, intelligent mechanical, and robot control.