Neurocomputing 543 (2023) 126276

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Fuzzy logic nonzero-sum game-based distributed approximated optimal control of modular robot manipulators with human-robot collaboration ☆

Tianjiao An [a], Xinye Zhu [a], Mingchao Zhu [b], Bing Ma [a], Bo Dong [a,*]

[a] Department of Control Science and Engineering, Changchun University of Technology, Changchun 130012, China
[b] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

ARTICLE INFO

Article history:
Received 17 November 2022
Revised 29 January 2023
Accepted 22 April 2023
Available online 27 April 2023
Communicated by Zidong Wang

Keywords:
Adaptive dynamic programming
Nonzero-sum game
Optimal control
Fuzzy systems
Modular robot manipulator

ABSTRACT

In this paper, a fuzzy logic nonzero-sum game-based distributed approximated optimal control scheme is presented for modular robot manipulators (MRMs) with human-robot collaboration (HRC) tasks. The MRM dynamic model is formulated by using joint torque feedback (JTF) technique. Based on the differential game strategy, the optimal control problem of HRC task-oriented MRM systems is transformed into a nonzero-sum game problem of multiple robotic subsystems. By taking advantage of the adaptive dynamic programming (ADP) algorithm, the distributed approximate optimal control policy under HRC tasks is developed by a novel fuzzy logic nonzero-sum game manner for solving the coupled Hamilton–Jacobian (HJ) equation. The trajectory tracking error under HRC task of the closed-loop MRM system is proved to be ultimately uniformly bounded (UUB) using the Lyapunov theory. Finally, experiment results have been presented, which demonstrate the advantage and effectiveness of the developed method.

1. Introduction

The collaborative robots, also known as cobots, are a research hotspot having a significant academic and application value in several application domains such as the assisted industrial operation, cooperative assembly, entertainment, rehabilitation and medical treatment [1,2]. Although the existed robotic system is engaged in physical human-robot interaction (pHRI) tasks, and it has a high precision and safety, traditional cobots still have disadvantages such as large size, difficult assembly and excessive degrees of freedom. Modular robot manipulator (MRM) is a kind of robots with standard modules and interfaces. Its configuration can be recombined and assembled, according to different task requirements. Due to the fact that it is reconfigurable and adaptable, and it has a flexible structure, MRM is especially suitable for human robot collaboration (HRC) tasks in complex environment. Hence, it is necessary to investigate the control method of modular cobot in order to guarantee the robustness and security of the system.

The HRC task is performed by a couple of autonomous individuals. Therefore, it is crucial to deal with the information exchange between robot and human. As a branch of game theory, differential game is used to describe and analyze the situation of interactive behavior, which can formulate the information interaction between participant and robot [3,4]. Several studies investigated the differential game-based pHRI control issue. For instance, Li et al. [5] solve the problem of stable interaction between human and robot based on differential game combining with an observer, and both sides successfully perform the interaction task with the minimum cost. Based on the differential game theory, the interaction behavior between the participant and the robot is analyzed, then the interactive control strategy is obtained by the policy iteration technique [6]. Based on the concept of differential game, each module of the MRM system can be considered as a participant with an individual policy, and operates in group with a general quadratic performance index function [7]. However, in the human-MRM interaction control issue, the control policy corresponding to each module is not only related to its own state, but also referred to the state of other modules and participants. In order to obtain the differential game-based pHRI control strategy, it is necessary to address the system's coupled Nash equilibrium solution. Therefore, the ideal method should obtain the analytical solutions of Nash

☆ This work is supported by the National Natural Science Foundation of China (62173047), the Scientific Technological Development Plan Project in Jilin Province of China (YDZJ202201ZYTS508).
* Corresponding author.

https://doi.org/10.1016/j.neucom.2023.126276
0925-2312/

equilibrium under differential game-based HRC task. Simultaneously, it should be possess with a low power consumption.

Adaptive dynamic programming (ADP), as a branch of reinforcement learning, is a valid method to solve optimal control issues [8]. It interacts with the environment by imitating the learning ability of organisms, and uses interactive data to continuously learn and improve its own strategy until the system performance becomes optimal. The ADP theory recently became a key method to address the approximate optimal control system issues of discrete-time [9–11], continuous-time [12–14], data-driven [15–17], robotic systems with input/output constraints [18–20], uncertainties disturbances [21–23] and multiple sensor failures [24,25]. For MRM system with coupled model uncertainties, the HJB equation of the system can be solved by the Weierstrass high-order approximation theorem [26,27] based on fuzzy logic, and the optimal control strategy can then be obtained. ADP method based on fuzzy logic has gained attention in solving electromechanical systems [28], network systems [29] and nonlinear large-scale systems [30]. In the process of pHRI, human can perceive whether the interacting force is massive, minor or moderate, rather than an exact value. Therefore, using the fuzzy logic, the membership function is introduced into the process of solving the force, and the human subjected force is described by more states, which can increase the security of the process of pHRI. However, few studies fully consider fuzzy logic-based approximate optimal human-MRM interaction control issue.

Motivated by the existing studies and the developing challenges, in this paper, a fuzzy logic nonzero-sum game-based distributed approximated optimal control scheme of MRMs is assessed for HRC tasks. The MRM dynamic model is considered as an integration of joint subsystem models, associated with interconnected dynamic couplings (IDCs) between modules. Based on the nonzero-sum differential game strategy, the optimal control problem of closed-loop robotic systems is transformed into a nonzero-sum game problem of multiple subsystems. By making the full use of ADP, the cost function, which is estimated by the fuzzy logic, is implemented to solve the coupled HJ equation of the physical human-MRM interaction system, thus facilitating the acquisition of the fuzzy logic-based nonzero-sum game scheme. The trajectory tracking error under the HRC task of the closed-loop MRM system is proved to be UUB. The experimental results are then presented to demonstrate the advantage and efficiency of the proposed method.

The main contributions of this paper are summarized as:

1. The original investigation introduces the nonzero-sum game theory to address the coupled Nash equilibrium solution under HRC tasks for the acquisition of the approximate optimal pHRI control strategy of MRM systems. An experimental verification of the efficiency and advantage of the control method is performed, and the experimental results are analyzed.
2. The fuzzy logic-based nonzero-sum game scheme is introduced into ADP to solve the coupled HJ equation of the physical human-MRM interaction system, based on the principle of Bellman optimality. The human subjected force is formulated by the hierarchy, which increases the comfort as well as security of the process of pHRI by fuzzy logic. The effectiveness of the approximated optimal control algorithm is proved by stability analysis.

2. Dynamic analysis and problem formulation

2.1. Dynamic model of MRM

By considering an MRM using the joint torque feedback (JTF) technique [31], the ith subsystem dynamic model is:

$$I_{im}\gamma_i\ddot{q}_i + \frac{\tau_{is}}{\gamma_i} + f_{ir}(q_i,\dot{q}_i) + I_{ic}(q,\dot{q},\ddot{q}) = \tau_i + J_i^T f, \tag{1}$$

where subscript i is the ith joint module subsystem, γ_i represents the gear ratio, q_i denotes the joint position, τ_{is} is the coupled joint torque, $f_{ir}(q_i,\dot{q}_i)$ represents the lumped joint friction, $I_{ic}(q,\dot{q},\ddot{q})$ denotes the IDC effect among subsystems, τ_i is the control torque, J_i denotes the Jacobi matrix, and f represents the human force input that the interaction force exert on the end-effector. The analysis of the properties is summarized as follows:

(1) The lumped joint friction
The joint friction term $f_{ir}(q_i,\dot{q}_i)$ is expressed as:

$$f_{ir}(q_i,\dot{q}_i) = \hat{f}_{ib}\dot{q}_i + \left(\hat{f}_{is}e^{\left(-\hat{f}_{i\tau}\dot{q}_i^2\right)} + \hat{f}_{ic}\right)\mathrm{sgn}(\dot{q}_i) + f_{ip}(q_i,\dot{q}_i) + Y_i(\dot{q}_i)\widetilde{F}_{ir}, \tag{2}$$

in which

$$Y_i(\dot{q}_i) = \left[f_{ib} - \hat{f}_{ib}, f_{ic} - \hat{f}_{ic}, f_{is} - \hat{f}_{is}, f_{i\tau} - \hat{f}_{i\tau}\right]^T, \tag{3}$$

where $f_{ip}(q_i,\dot{q}_i)$ is the position dependency friction term, $f_{ib}, f_{i\tau}$ are viscous and Stribect friction effect, f_{is}, f_{ic} are static and Coulomb friction parameters. Furthermore, $\hat{f}_{ib}, \hat{f}_{ic}, \hat{f}_{is}, \hat{f}_{i\tau}$ are the estimated values of $f_{ib}, f_{ic}, f_{is}, f_{i\tau}$.

Remark 1. $f_{ib}, f_{ic}, f_{is}, f_{i\tau}$ and estimations are bounded, so that $\widetilde{F}_{ir}$ is bounded as $\left|\widetilde{F}_{ir}\right| \leqslant b_{iFrm}(m = 1,2,3,4)$, and b_{iFrm} is a known positive constant. Accordingly, one can obtain $Y_i(\dot{q}_i)\widetilde{F}_{ir}$ which is given as $\left|Y_i(\dot{q}_i)\widetilde{F}_{ir}\right| \leqslant Y_i(\dot{q}_i)b_{iFrm}$. Besides, $|f_{ip}(q_i,\dot{q}_i)| \leqslant b_{iFp}$, in which b_{iFp} is a known positive constant bound.

(2) The interconnected dynamic coupling
The nonlinear IDC term can be expressed as:

$$\begin{aligned} I_{ic}(q,\dot{q},\ddot{q}) &= I_{im}\sum_{j=1}^{i-1} v_{mi}^T v_{lj}\ddot{q}_j + I_{im}\sum_{j=2}^{i-1}\sum_{k=1}^{j-1} v_{mi}^T (v_{lk} \times v_{lj})\dot{q}_k\dot{q}_j \\ &= I_{im}\sum_{j=1}^{i-1} D_j^i\ddot{q}_j + I_{im}\sum_{j=2}^{i-1}\sum_{k=1}^{j-1}\Theta_{kj}^i\dot{q}_k\dot{q}_j \\ &= \sum_{j=1}^{i-1}\left[I_{im}\widehat{D}_j^i, I_{im}\right]\left[\ddot{q}_j, \widetilde{D}_j^i\ddot{q}_j\right]^T \\ &+ \sum_{j=2}^{i-1}\sum_{k=1}^{j-1}\left[I_{im}\hat{\Theta}_{kj}^i, I_{im}\right]\left[\ddot{q}_j, \tilde{\Theta}_{kj}^i\dot{q}_k\dot{q}_j\right]^T, \end{aligned} \tag{4}$$

in which v_{mi}, v_{lj}, v_{lk} represent the unit vectors along with the ith, jth and kth joint rotation axes, respectively. Accordingly, one can define $D_j^i = v_{mi}^T v_{lj}$ and $\Theta_{kj}^i = v_{mi}^T(v_{lk} \times v_{lj})$. Moreover, we also have the relations that $\widehat{D}_j^i = D_j^i - \widetilde{D}_j^i$ and $\hat{\Theta}_{kj}^i = \Theta_{kj}^i - \tilde{\Theta}_{kj}^i$, in which $\widehat{D}_j^i, \hat{\Theta}_{kj}^i$ denote the estimated values of D_j^i, Θ_{kj}^i as well as $\widetilde{D}_j^i, \tilde{\Theta}_{kj}^i$ are alignment errors.

Remark 2. According to the definition of v_{mi}, v_{lk}, v_{lj} in (4), we obtain that the vector products are bounded by $\left|D_j^i\right| = |v_{mi}^T v_{lj}| < 1$ and $\left|\Theta_{kj}^i\right| = |v_{mi}^T(v_{lk} \times v_{lj})| < 1$. Moreover, we also conclude that $I_{ic}(q,\dot{q},\ddot{q})$ is bounded and the up-bound is given as $|I_{ic}(q,\dot{q},\ddot{q})| \leqslant b_{iI}$.

Define state vector $x_i = [x_{i1}, x_{i2}]^T = [q_i, \dot{q}_i]^T$ and the control input $u_i = \tau_i$. One has the ith subsystem state space:

$$\begin{cases} \dot{x}_{i1} = x_{i2} \\ \dot{x}_{i2} = f_i(x) + g_i u_i + \varpi_i(x), \end{cases} \tag{5}$$

where

$$g_i = (I_{im}\gamma_i)^{-1}$$
$$f_i(x) = -g_i \begin{pmatrix} \left(\hat{f}_{is} e^{\left(-\hat{f}_{ir}\dot{x}_{i1}^2\right)} + \hat{f}_{ic}\right)\mathrm{sgn}(x_{i2}) + f_{ip}(x_{i1}, x_{i2}) \\ +\hat{f}_{ib}x_{i2} + Y_i(x_{i2})\widetilde{F}_{ir} + \frac{\tau_{is}}{\gamma_i} + I_{ic}(x, \dot{x}, \ddot{x}) \end{pmatrix} \tag{6}$$
$$\varpi_i(x) = -g_i J_i^T f.$$

2.2. Human limb model and motion intention estimation

In physical human robot interaction, the human force is considered as the only external force exerting on the robot end-effector [32]. The HRC control transfers the human force input into the motion commands of the robot:

$$-C_H\dot{z} + G_H(z_{Hd} - z) = f, \tag{7}$$

where C_H, G_H are the dampers, spring unknown diagonal matrices of the human, z is the robot actual position in Cartesian space which can be calculated as $z(t) = \xi(q), q(t) = [q_1, \cdots, q_i, \cdots, q_n]^T$ is the position vector in the joint space, $\xi(\cdot)$ is a mapping matrix from joint space to Cartesian space, z_{Hd} denotes the trajectory planned in the human which is referred to as the motion intention of the human and the robot.

Remark 3. The up-bound of interaction force f is b_f which guarantees that, once a feasible HRC task is determined, it is possible to guarantee that the MRM will converge to its goal. If the human interaction force is without a finite bound, it is not possible to guarantee the convergence that the trajectory tracking error is UUB.

The human motion intention while interacting with a robot z_{Hd} can be expressed as [33]:

$$z_{Hd} = \Lambda(f, z, \dot{z}), \tag{8}$$

where $\Lambda(\cdot)$ is considered as an unknown nonlinear function.

Furthermore, it is known that z_{Hd} is difficult to obtain since the human may change its limb during the collaboration task. Considering the concept of radial basis function NN, the human motion intention while interacting with a robot and its estimation are given by:

$$z_{Hd} = W_x^T S(f, z, \dot{z}) + \varepsilon, \hat{z}_{Hd} = \widehat{W}_x^T S(f, z, \dot{z}), \tag{9}$$

where ε is the estimation error, $\widehat{W}_x$ is the estimated value of ideal weight W_x and S represents the Gaussian function.

The gradient descent algorithm is used to obtain $\widehat{W}_x$ in (9). In order to make the MRM actively and easily move toward its human's intended position, $\widehat{W}_x$ is adjusted online using the following cost function respect to the interaction force $E = \frac{1}{2}\left\|f^2\right\|$. Therefore:

$$\dot{\widehat{W}}_x = -\alpha\prime \frac{\partial E}{\partial \widehat{W}_x} = -\alpha\prime f G_H S = -\alpha_A f S, \tag{10}$$

where $\alpha\prime$ is a positive scalar, $\alpha_A = \alpha' G_H$.

We can get $\widehat{W}_x$ as:

$$\widehat{W}_x(t) = \widehat{W}_x(0) - \alpha_A \int_0^t (f(\upsilon)S(\upsilon))d\upsilon. \tag{11}$$

Therefore, one can obtain $\hat{z}_{Hd}$ in (9).

The control object consists in optimally ensuring that the considered tracking error of MRM systems under pHRI is UUB. In the next section, a distributed approximate optimal control of MRMs for HRC using fuzzy logic nonzero-sum game, is presented.

3. Fuzzy logic nonzero-sum game-based distributed approximated optimal control

3.1. Problem transformation

In this paper, a fuzzy logic nonzero-sum game-based distributed approximated optimal control scheme is developed to guarantee that the end-effector of MRM can actively move toward human's planned intention under HRC task. In addition, in the nonzero-sum differential game, the object consists in ensuring the minimization of each performance index for all the modules as well as the cost between the human and the robot, as the Nash equilibrium. Therefore, in order to facilitate the design of the controller, n modules can be deemed as n players.

Since the estimation of human intention in Cartesian spa-ce z_{Hd} is obtained in the previous section, the estimated human motion intention $\hat{q}$ can be defined in the joint space as $\hat{q}(t) = \xi^{-1}(z_{Hd})$. By augmenting the subsystem dynamics, the state space can be expressed as:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_D(x) + \sum_{m=1}^{n} G_m u_m + \varpi(x), \end{cases} \tag{12}$$

where $x = \left[x_1^T, x_2^T\right]^T \in R^{2n}$ is the global state of the MRM system, in which the vectors x_1, x_2 are given by $x_1 = [x_{11}, \ldots, x_{i1}, \ldots, x_{n1}]^T \in R^n$ and $x_2 = [x_{12}, \ldots, x_{i2}, \ldots, x_{n2}]^T \in R^n$. Moreover, we have $f_D(x) = [f_1(x), \ldots, f_i(x), \ldots, f_n(x)]^T, G_m = [0, \ldots, 0, g_m, 0, \ldots, 0]^T$, $\varpi(x) = [\varpi_1(x), \ldots, \varpi_i(x), \ldots, \varpi_n(x)]^T$, where $g_m = (I_{mm}\gamma_m)^{-1}$, $m = 1, \ldots, n$.

Define the cost fuction:

$$\begin{aligned} J_i(\dot{e}, u_1, \ldots, u_n, f) &= \int_t^\infty \left(\dot{e}^T Q_i \dot{e} + \sum_{m=1}^{n} u_m^T R_{im} u_m + \varpi^T \varpi\right) d\tau \\ &= \int_t^\infty U_i(\dot{e}, u_1, \ldots, u_n, f) d\tau, \end{aligned} \tag{13}$$

where position error is $e = [e_1, e_2, \cdots, e_n]^T = x_1 - x_d$ and velocity error vector is $\dot{e} = [\dot{e}_1, \dot{e}_2, \cdots, \dot{e}_n]^T = x_2 - \dot{x}_d, x_d = \hat{q}(t)$ denotes the estimated human motion vectors in joint space, Q_i, R_{im} are determined positive definite matrices and $U_i(\dot{e}, u_1, \ldots, u_n, f)$ represents the utility function.

Definition 1. The control policy u_i is denoted as admissible with regard to the cost function (13), if u_i stabilizes the manipulator system (12) on compact set ψ, and $J_i(\dot{e}, u_1, \cdots u_n, f)$ is finite for arbitrary $x \in \psi$.

Using the infinitesimal version of (13), the Hamiltonian function is expressed as:

$$H_i(\dot{e},u_1,\ldots,u_n,f,\nabla J_i)=U_i(\dot{e},u_1,\ldots,u_n,f)+(\nabla J_i)^T\left(f_D(x)+\sum_{m=1}^{n}G_mu_m+\varpi(x)-\ddot{x}_d\right), \tag{14}$$

where $\nabla J_i(\dot{e})=\frac{\partial J_i(\dot{e})}{\partial \dot{e}}$ is the partial derivative of $J_i(\dot{e})$. Moreover, the optimal cost function can be defined as:

$$J_i^*(\dot{e},u_1,\ldots,u_n,f)=\min_{u_i}\int_t^\infty U_i(\dot{e},u_1,\ldots,u_n,f)d\tau. \tag{15}$$

According to the stationary condition $\frac{\partial H_i}{\partial u_i}=0$, the local distributed approximated optimal control policy u_i^* is:

$$u_i^*=-\frac{1}{2}R_{ii}^{-1}G_i^T\nabla J_i^*. \tag{16}$$

Similarly, based on $\frac{\partial H_i}{\partial f}=0$, the local approximated optimal interaction force is given by:

$$\varpi^*=-\frac{1}{2}\nabla J_n^*. \tag{17}$$

Afterwards, by substituting (13), (16) and (17) into the Hamiltonian function (14), the coupled HJ equation can be expressed as:

$$\begin{aligned}&(\nabla J_i^*)^T\left(f_D(x)-\frac{1}{2}\sum_{m=1}^{n}G_mR_{mm}^{-1}G_m^T\nabla J_m^*-\frac{1}{2}\nabla J_n^*-\ddot{x}_d\right)+\dot{e}^TQ_i\dot{e}\\&+\frac{1}{4}\sum_{m=1}^{n}(\nabla J_i^*)^TG_mR_{mm}^{-1}R_{im}R_{mm}^{-1}G_m^T(\nabla J_m^*)+\frac{1}{4}(\nabla J_n^*)^T\nabla J_n^*\\&=0.\end{aligned} \tag{18}$$

Remark 4. The developed distributed approximated optimal control policy u_i^* among MRM systems affords superior expandability as well as handlability than the centralized mechanism, which is not suitable for computation and transmission manner. The applied distributed control mechanism averts MRM systems' majorization of high dimensionality that only utilizes each module subsystem's local and neighborhood information. As a consequence, the cost function of each player can be optimized accompanying fewer cost.

3.2. Derivation of fuzzy logic

Since in the process of pHRI, the participant cannot accurately perceive the exact force that is acting on the robot's end-effector, human can only feel the comfort level in the interaction process, whether it is compliance or strenuous. Therefore, using the concept of fuzzy logic, the following IF-THEN rules are provided:

R_i^k: If x_{i1} is F_{i1}^k, x_{i2} is F_{i2}^k, ..., x_{im} is F_{im}^k.

Then y_i is $G_i^k, k=1,2,\cdots,M$where $x_{i\cdot}$ and y_i respectively denote the fuzzy logic's input and output, $F_{i\cdot}^k$ and G_i^k represent the fuzzy set, $\delta_{F_{i\cdot}^k}$ and $\delta_{G_i^k}$ are respectively the fuzzy membership functions and M is the quantity of the rules.

Based on the fuzzy rules, the fuzzy logic is as follows:

$$y_i=\frac{\sum_{k=1}^{M}\bar{y}_k\prod_{\cdot=1}^{m}\delta_{F_{i\cdot}^k}}{\sum_{k=1}^{M}\left(\prod_{\cdot=1}^{m}\delta_{F_{i\cdot}^k}\right)}, \tag{19}$$

where $\bar{y}_k=\max\delta_{G_i^k}$.

The definition of the fuzzy basis function is then given as follows:

$$\delta_k=\frac{\prod_{\cdot=1}^{m}\delta_{F_{i\cdot}^k}}{\sum_{k=1}^{M}\left(\prod_{\cdot=1}^{m}\delta_{F_{i\cdot}^k}\right)}. \tag{20}$$

Hence, (19) is expressed as:

$$y_i=W_i^T\delta. \tag{21}$$

Lemma 1 [34]. *y_i is considered as a finite continuous function, then:*

$$\sup_{x_i}\left|y_i-W_i^T\delta\right|\leqslant\psi_i, \tag{22}$$

where ψ_i is an arbitrary small constant.

Remark 5. Since the position tracking error generating from human motion while interacting with the robot is time-varying, it is with great significance for controller design to consider the mentioned information. Besides, participants are affected by comfort level during pHRI instead of precise value. Motivating by above comments, the fuzzy logic-based distributed approximated optimal control strategies are adapted according to the position as well as velocity errors to further improve the comfort of interaction and accuracy of calculation.

3.3. Approximate solution of the nonzero-sum game by the fuzzy logic

Dynamic compensation is a significant part in robotic control. In this study, a compensator controller which comprises dynamic information-based model and fuzzy logic no-nzero-sum game-based optimal control is designed.

$$u_i^*=u_{i1}+u_{i2}^*, \tag{23}$$

where u_{i1} is dealt with the dynamic model $f_i(x)$ and u_{i2}^* is the optimal compensation issue of the uncertainties as well as HRC tasks.

According to (6), u_{i1} can be designed as:

$$u_{i1}=-\begin{pmatrix}-\left(\hat{f}_{is}e^{\left(-\hat{f}_{i\tau}x_{i2}^2\right)}+\hat{f}_{ic}\right)\operatorname{sgn}(x_{i2})\\-\hat{f}_{ib}x_{i2}-g_i^{-1}\ddot{x}_{id}-\frac{\tau_{is}}{\gamma_i}\end{pmatrix}. \tag{24}$$

The optimal compensation control problem is then transformed into a nonzero-sum game issue.

The fuzzy logic is developed to approximate the cost function (13) as follows:

$$J_i^*(\dot{e})=W_{ic}^T\phi_{ic}(\dot{e})+\varepsilon_{ic}, \tag{25}$$

where W_{ic} is the fuzzy logic ideal vector, ε_{ic} is the finite approximate error and $\phi_{ic}(\dot{e})$ represents fuzzy basis function.

The gradient of the approximated cost function can be obtained as:

$$\nabla J_i^*(\dot{e})=\nabla\phi_{ic}^T(\dot{e})W_{ic}+\nabla\varepsilon_{ic}, \tag{26}$$

where $\nabla\phi_{ic}(\dot{e})=\partial\phi_{ic}(\dot{e})/\partial\dot{e}$ is the gradient of the fuzzy basis function and $\nabla\varepsilon_{ic}$ is the gradient error.

By substituting (26) into (16), the optimal control policy is given by:

$$u_{i2}^* = -\frac{1}{2}R_{ii}^{-1}G_i^T\left(\nabla\phi_{ic}^T(\dot{e})W_{ic} + \nabla\varepsilon_{ic}\right). \tag{27}$$

Substituting (26) into (27), one yields:

$$\begin{aligned} H_i(\dot{e}, u_1, \ldots, u_n, f, W_c) &= \dot{e}^T Q_i \dot{e} + \sum_{m=1}^{n} u_m^T R_{im} u_m + \varpi^T\varpi \\ &\quad + (\nabla J_i)^T\left(f_D(x) + \sum_{m=1}^{n} G_m u_m + \varpi(x) - \ddot{x}_d\right) \\ &\quad - e_{icH} \\ &= 0, \end{aligned} \tag{28}$$

where e_{icH} denotes the residual error.

The estimated optimal cost function is given by:

$$\widehat{J}_i^*(\dot{e}) = \widehat{W}_{ic}^T\phi_{ic}(\dot{e}). \tag{29}$$

Based on (27) and (29), the approximate optimal control is expressed as:

$$\hat{u}_{i2}^* = -\frac{1}{2}R_{ii}^{-1}G_i^T\nabla\phi_{ic}^T(\dot{e})\widehat{W}_{ic}. \tag{30}$$

According to (28), the approximated Hamiltonian function is given by:

$$\begin{aligned} \widehat{H}_i\left(\dot{e}, \hat{u}_1, \ldots, \hat{u}_n, \hat{f}, \widehat{W}_c\right) &= \dot{e}^T Q_i \dot{e} + \sum_{m=1}^{n} u_m^T R_{im} u_m + \varpi^T\varpi \\ &\quad + \widehat{W}_{ic}^T\nabla\phi_{ic}(\dot{e})\ddot{e} \\ &= 0. \end{aligned} \tag{31}$$

The approximated Hamiltonian error function e_{ic} is defined as follows:

$$e_{ic} = \widehat{H}_i - H_i, \tag{32}$$

where $e_{ic} = \widehat{H}_i$ is obtained using (28), (32).

When defines the vector estimate error $\widetilde{W}_{ic} = W_{ic} - \widehat{W}_{ic}$, it can be deduced that $e_{ic} = e_{icH} - \widetilde{W}_{ic}^T\nabla\phi_{ic}(\dot{e})\ddot{e}$, by combining (28), (31) with (32).

The update law is designed as:

$$\dot{\widehat{W}}_{ic} = -\alpha_i e_{ic}\nabla\phi_{ic}(\dot{e})\ddot{e}, \tag{33}$$

where α_i is the updated rate of the fuzzy logic.

υ_i is denoted as $\nabla\phi_{ic}(\dot{e})\ddot{e}$, and a positive constant υ_{iL} that $\|\upsilon_i\| \leqslant \upsilon_{iL}$ is assumed. Thus, we have:

$$\dot{\widetilde{W}}_{ic} = -\dot{\widehat{W}}_{ic} = \alpha_i e_{ic}\nabla\phi_{ic}(\dot{e})\ddot{e} = \alpha_i\left(e_{icH} - \widetilde{W}_{ic}^T\upsilon_i\right)\upsilon_i. \tag{34}$$

Remark 6. The fuzzy logic is leveraged to decompose the complicated MRM system into quantity of fuzzy subsystems as players in the nonzero-sum differential game, which are corresponding to IF-THEN rules. Then, the unknown cost function (13), which is estimated by the fuzzy logic, is implemented to solve the coupled HJ equation of the physical human-MRM interaction system. Unlike the conventional critic NN to deal with cost function in the nonzero-sum game that tough to determine the center location or quantities of the radial basis function NNs, which increase the computation space and communication protocol. By utilizing the promising fuzzy logic technique, the solution of nonzero-sum differential game can be predigested. Afterwards, one can obtain the proposed fuzzy logic nonzero-sum game-based distributed approximated optimal control scheme.

Theorem 1. *Considered the cost function which is approximated by (25), with an ideal weight W_{ic}, and the estimated cost function given by (29) that is built with approximated weight $\widehat{W}_{ic}$, if the weight of the fuzzy logic is updated by (33), then the weight approximation error is guaranteed to be UUB.*

Proof. Select the Lyapunov function:

$$V_{ic}(t) = \frac{1}{2\alpha_i}\widetilde{W}_{ic}^T\widetilde{W}_{ic}. \tag{35}$$

The time derivative of $V_{ic}(t)$ is obtained as:

$$\begin{aligned} \dot{V}_{ic}(t) &= \tfrac{1}{\alpha_i}\widetilde{W}_{ic}^T\dot{\widetilde{W}}_{ic} = \tfrac{1}{\alpha_i}\widetilde{W}_{ic}^T\alpha_i\left(e_{icH} - \widetilde{W}_{ic}^T\upsilon_i\right)\upsilon_i \\ &= \widetilde{W}_{ic}^T\left(e_{icH} - \widetilde{W}_{ic}^T\upsilon_i\right)\upsilon_i = \widetilde{W}_{ic}^T e_{icH}\upsilon_i - \left\|\widetilde{W}_{ic}^T\upsilon_i\right\|^2 \\ &\leqslant \tfrac{1}{2}\|e_{icH}\|^2 - \tfrac{1}{2}\left\|\widetilde{W}_{ic}^T\upsilon_i\right\|^2, \end{aligned} \tag{36}$$

where $\dot{V}_{ic}(t) \leqslant 0$ when $\dot{e}$ lies outside $\Omega_{i1} = \left\{\widetilde{W}_{ic} : \left\|\widetilde{W}_{ic}\right\| \leqslant \frac{e_{icH}}{\upsilon_{iL}}\right\}$. This completes the proof.

Based on the compensation-based control policy (24) and the optimal control policy (30), $\hat{u}_i^*$ is given by:

$$\begin{aligned} \hat{u}_i^* &= u_{i1} + \hat{u}_{i2}^* \\ &= -\begin{pmatrix} -\left(\hat{f}_{is}e^{\left(-\hat{f}_{i\tau}x_{i2}^2\right)} + \hat{f}_{ic}\right)\mathrm{sgn}(x_{i2}) \\ -\hat{f}_{ib}x_{i2} - g_i^{-1}\ddot{x}_{id} - \frac{\tau_{is}}{\gamma_i} \end{pmatrix} - \tfrac{1}{2}R_{ii}^{-1}G_i^T\nabla\phi_{ic}^T(\dot{e})\widehat{W}_{ic}. \end{aligned} \tag{37}$$

The structural diagram of the proposed fuzzy logic nonze-ro-sum game-based distributed approximated optimal control is illustrated in Fig. 1.

Theorem 2. Given an MRM system with the dynamic model of each joint subsystem represented in (1), and the state space formulated in (12), the trajectory tracking error of the closed-loop robotic system is UUB under the HRC task, by the presented fuzzy logic nonzero-sum game-based distributed approximated optimal control derived in (37).

Proof. Select $J_i^*(\dot{e})$ as Lyapunov function, and derivative it as follows:

$$\dot{V}(t) = (\nabla J_i^*)^T\left(f_D(x) + \sum_{m=1}^{n} G_m u_m + \varpi(x) - \ddot{x}_d\right). \tag{38}$$

By considering the coupled HJ equation formulated in (18), it yields:

$$\begin{aligned} (\nabla J_i^*)^T(f_D(x) - \ddot{x}_d) &= -\dot{e}^T Q_i \dot{e} - \frac{1}{4}(\nabla J_n^*)^T\nabla J_n^* \\ &\quad - \frac{1}{4}\sum_{m=1}^{n}(\nabla J_m^*)^T G_m R_{mm}^{-1} R_{im} R_{mm}^{-1} G_m^T(\nabla J_m^*) \\ &\quad + \frac{1}{2}\sum_{m=1}^{n}(\nabla J_m^*)^T G_m R_{mm}^{-1} G_m^T\nabla J_m^*. \end{aligned} \tag{39}$$

Combining (39) into (38), we obtain:

$$\begin{aligned} \dot{V}(t) &= -\dot{e}^T Q_i \dot{e} - \left(\sum_{m=1}^{n} G_m(u_m^* - \hat{u}_m)\right) - \frac{1}{4}(\nabla J_n^*)^T\nabla J_n^* \\ &\quad - \frac{1}{4}\sum_{m=1}^{n}(\nabla J_m^*)^T G_m R_{mm}^{-1} R_{im} R_{mm}^{-1} G_m^T(\nabla J_m^*) \\ &\quad + \frac{1}{2}\sum_{m=1}^{n}(\nabla J_m^*)^T G_m R_{mm}^{-1} G_m^T\nabla J_m^*. \end{aligned} \tag{40}$$

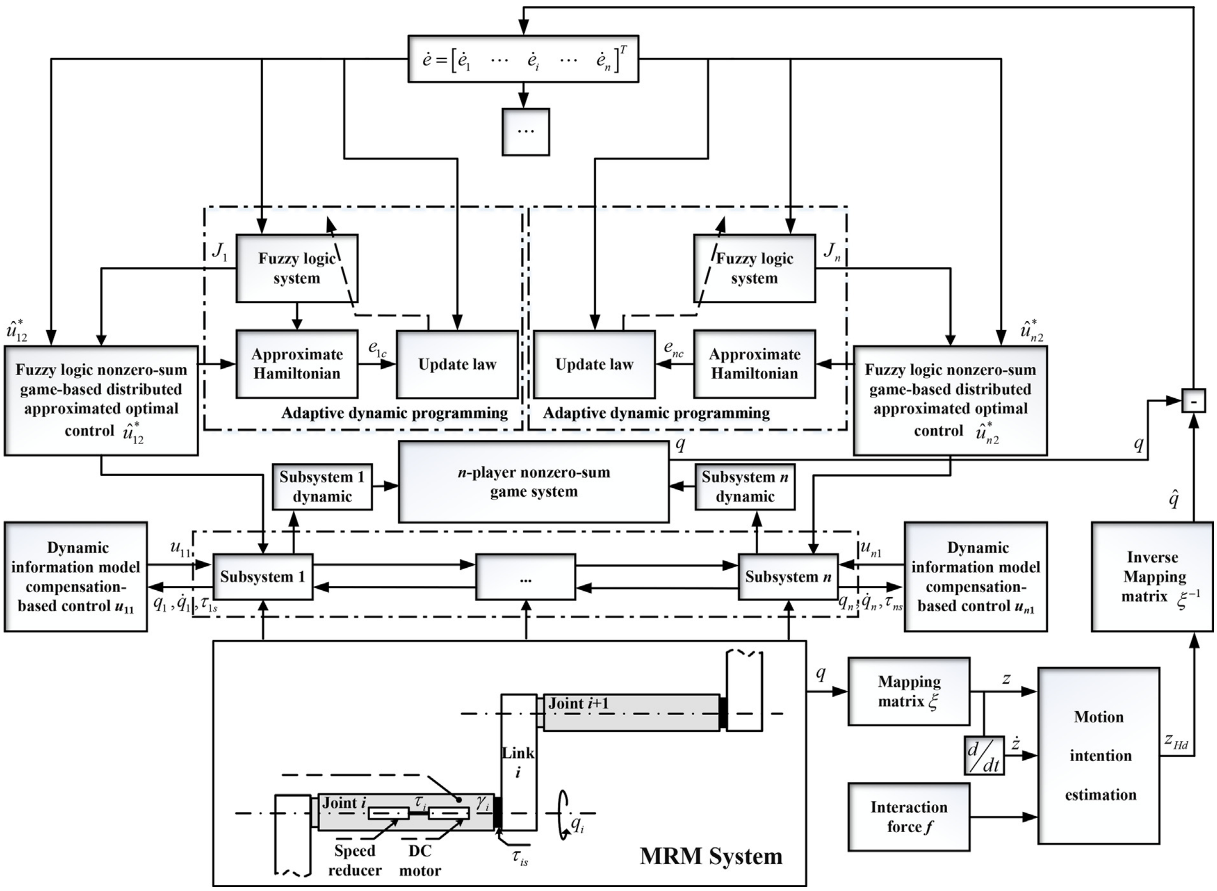

Fig. 1. Structural diagram of the fuzzy logic nonzero-sum game-based distributed approximated optimal control.

Combining (38) and (40), one obtains:

$$\begin{aligned}\dot{V}(t) = & -\dot{e}^T Q_i \dot{e} - \tfrac{1}{4}\left(\nabla J_n^*\right)^T \nabla J_n^* + \tfrac{1}{2}\left(\nabla\phi_{ic}^T(\dot{e}) W_{ic} + \nabla\varepsilon_{ic}\right)^T \\ & \left(\sum_{m=1}^{n} G_m R_{mm}^{-1}\left(G_m^T \nabla\phi_{mc}^T(\dot{e})\widetilde{W}_{mc} + G_m^T \nabla\varepsilon_{mc}\right)\right) \\ & -\tfrac{1}{4}\sum_{m=1}^{n}\left(\nabla J_m^*(\dot{e})\right)^T G_m R_{mm}^{-1} R_{im} R_{mm}^{-1} G_m^T \nabla J_m^*(\dot{e}) \\ = & -\dot{e}^T Q_i \dot{e} + \Pi_J,\end{aligned} \tag{41}$$

in which Π_J has up-bound:

$$\begin{aligned}\Pi_J \leqslant & \left\| \begin{array}{l} -\tfrac{1}{4}\sum_{m=1}^{n}\left(\nabla J_m^*(\dot{e})\right)^T G_m R_{mm}^{-1} R_{im} R_{mm}^{-1} G_m^T \nabla J_m^*(\dot{e}) \\ +\tfrac{1}{2}\left(\nabla\phi_{ic}^T(\dot{e}) W_{ic} + \nabla\varepsilon_{ic}\right)^T \\ \left(\sum_{m=1}^{n} G_m R_{mm}^{-1}\left(G_m^T \nabla\phi_{mc}^T(\dot{e})\widetilde{W}_{mc} + G_m^T \nabla\varepsilon_{mc}\right)\right) \end{array} \right\| \\ \leqslant & \ \pi_J,\end{aligned} \tag{42}$$

where π_J is a computable constant.

Combining (42), $\dot{V}(t)$ has upper bound by:

$$\dot{V}(t) \leqslant -\dot{e}^T Q_i \dot{e} + \pi_J \leqslant -\lambda_{\min}(Q_i)\|\dot{e}\|^2 + \pi_J. \tag{43}$$

If $\dot{e}$ lies outside:

$$\Omega = \left\{ \dot{e} : \|\dot{e}\| \leqslant \sqrt{\frac{\pi_J}{\lambda_{\min}(Q_i)}} \right\}, \tag{44}$$

(38) is negative. Therefore, one obtains $\dot{V}(t) < 0$ for any $\dot{e} \neq 0$ when (44) is satisfied. The trajectory tracking error under HRC task is proved to be UUB under (37). This completes the proof.

4. Experiments

4.1. Experiment setup

The proposed control methods are verified by two degrees of freedom MRM experimental platform. The detailed experimental device can be seen from Fig. 2. The measured joint control torque can be delivered by the joint torque sensor. Furthermore, the joint position information can be acquired by the absolute as well as incremental encoder. The data acquisition board serves as the medium through which the software environment (Simulink) interacts with hardware information (robotic module). The proposed fuzzy logic nonz-ero-sum game-based distributed approximated optimal control algorithm, which is with continuous time condition, and needs to be implemented discretely. The control system is built via Simulink environment, which can realize the discrete implementation automatically. The situation of pHRI that is handshaking task with MRM is considered (cf. Fig. 2c). According to the requirements of human robot collaboration tasks, the first major principle is always to ensure the safety of the participant. An emergency stop is used to ensure the safety in the case of an accident in the interaction process. In trajectory planning, interactive tasks are relatively slow to guarantee the human security.

An experiment is performed to determine if the requirements of position tracking performance and control torque optimization are

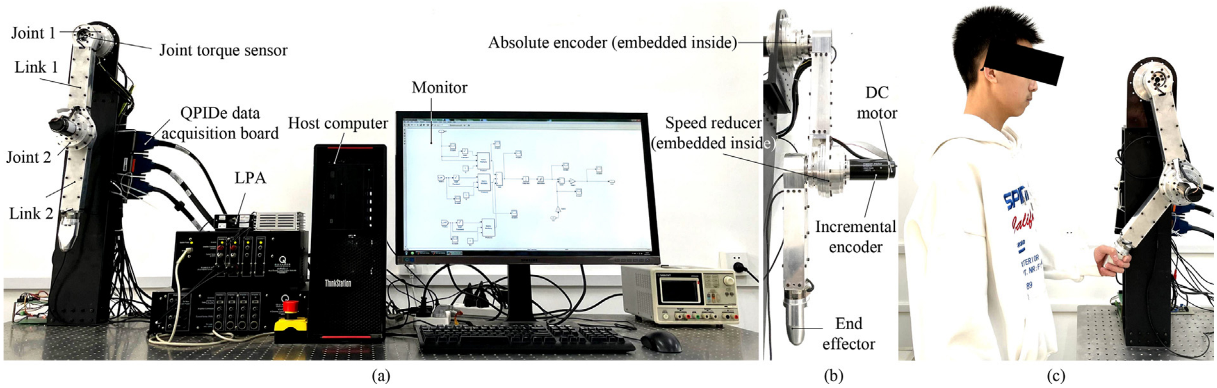

Fig. 2. Experimental situation (a) Experimental platform setup (b) Experimental joint module (c) Experimental with HRC.

Table 1
Fuzzy logic-based controller rule base.

Fuzzy logic input		Velocity tracking error		
		Positive	Zero	Negative
Position tracking error	Positive	Positive big	Positive small	Zero
	Zero	Positive small	Zero	Negative small
	Negative	Zero	Negative small	Negative big

met. Furthermore, humans should feel as comfortable as possible when shaking hands with the robot. The related dynamic model and control design are given by:

$\hat{f}_{ib} = 12mNm/rad, \hat{f}_{ic} = 30mNm, \hat{f}_{is} = 40mNm, \hat{f}_{i\tau} = 20s^2/rad^2$, $I_{im} = 120gcm^2, \gamma_i = \alpha_i = 0.9, Q_i = R_{im} = I, \alpha\prime = 3.72, b_f = 2.6$.

For the fuzzy logic nonzero-sum game-based distributed approximated optimal control in this paper, the position as well as velocity tracking error $e, \dot{e}$ mean the input of the fuzzy logic and the cost function J_i represents the output, thus derivation of the local distributed approximated optimal control policy as well as the optimal interaction force. The defined fuzzy sets $F_{i.}^{k}$ consists of positive, zero as well as negative, and G_i^k includes positive big, positive small, zero, negative small as well as negative big. The membership functions $\delta_{F_{i.}^{k}}$ and $\delta_{G_{i}^{k}}$ are selected as triangle membership function. On the premise of ensuring system performance, the quantity of fuzzy rules should be minimized for simplifying controller design and reducing computation burden. Therefore, 3^m fuzzy rules are adopted in the rule base, which is shown in Table 1.

4.2. Experiment results

The experimental results demonstrate the performance of position tracking, tracking error, control torque and fuzzy logic weight. Two types of control methods are used to verify the validity of the proposed method: the existed learning-based tracking control without fuzzy logic nonzero-sum game method, such as actor-critic-based optimal control, (e.g. [35,36]), and the proposed fuzzy logic nonzero-sum game-based distributed approximate optimal control method. The two control methods are used in handshaking HRC tasks. Note that in the figures, (a) represents joint one, and (b) denotes joint two.

(1) Position tracking performance
Figs. 3 and 4 present the position tracking and tracking error curves in joint space under handshaking tasks, obtained by the existed learning-based tracking control and proposed fuzzy logic nonzero-sum game-based approximate optimal control methods, respectively. As the optimal compensation of the model uncertainties and coupling effects is not implemented in the existed learning-based optimal control method, the error curves are featured with obvious chattering and noise effects. The amplitude of tracking error tends to become larger in some time and can be restored to an acceptable range in a short period via the proposed fuzzy logic nonzero-sum game-based distributed approximate optimal control method due to effective friction compensation. Unlike the conventional stabilization control method, the proposed optimal control method can guarantee track-

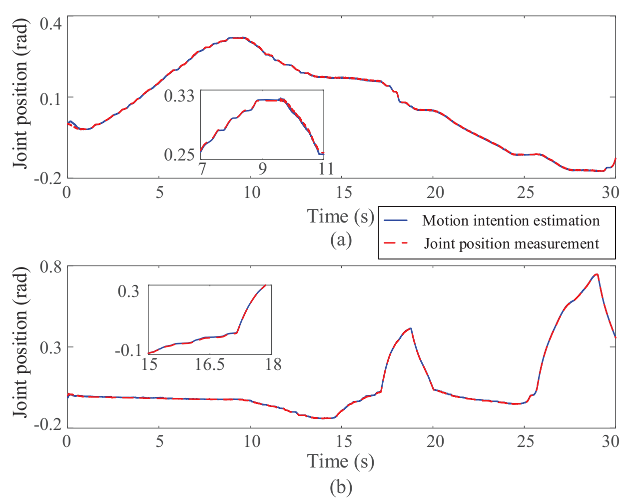

Fig. 3. Motion intention estimation and position tracking curves in joint space under handshaking tasks via the proposed optimal control method.

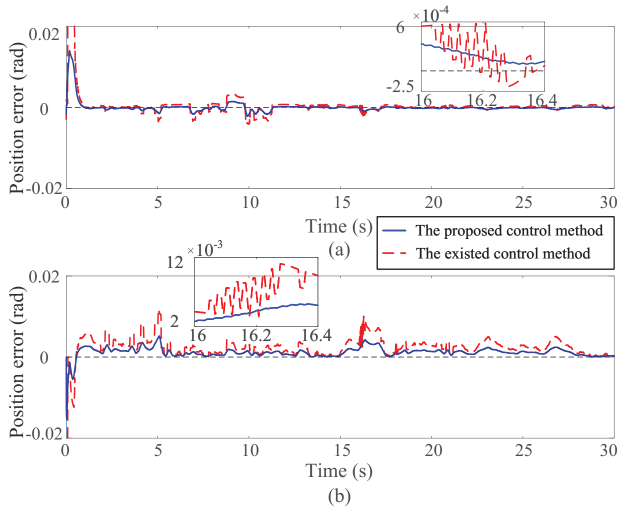

Fig. 4. Position tracking error curves in joint space under handshaking tasks via the existed and the proposed optimal control method.

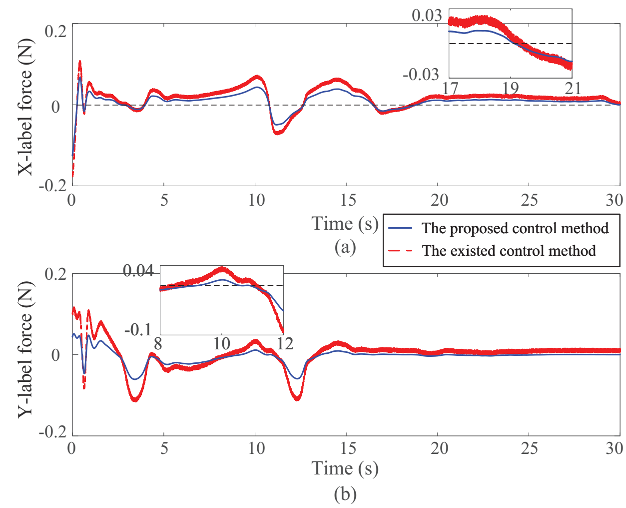

Fig. 6. Interaction force curves under handshaking tasks via the existed and the proposed optimal control method.

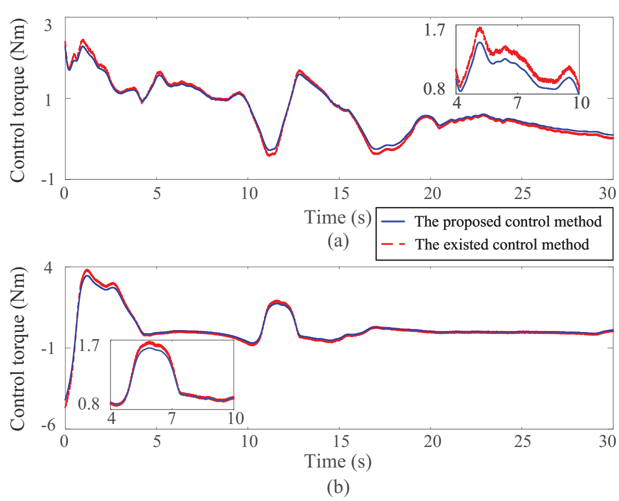

Fig. 5. Control torque curves under handshaking tasks via the existed and the proposed optimal control method.

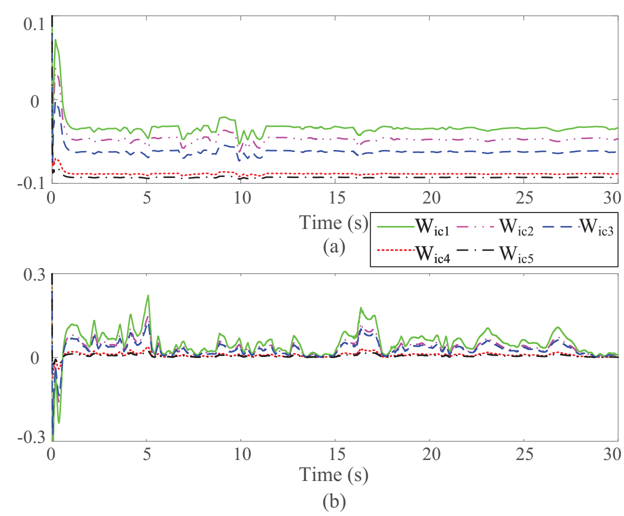

Fig. 7. Fuzzy logic weight vector curves under handshaking tasks via the proposed optimal control method.

ing error restraint in a limited scope, which experimentally demonstrates the ultimately uniformly bounded performance.

(2) Control torque

Fig. 5 shows the control torque curves under handshaking tasks using the existed learning-based tracking control method and the proposed fuzzy logic nonzero-sum game-based distributed approximate optimal control method. It can be observed that, when the trajectory sharply changes, the control torque is instantaneously increased, which may affect the durability of the DC motors. Furthermore, the control torque curves show serious chattering effects under the existed control method, which may reduce the precision of the trajectory tracking. Using the fuzzy logic nonzero-sum game-based distributed approximate optimal control method, the output torques use the output power of motors, and the instant increase of control torque is restrained within safe limits that no more than 4Nm.

(3) Interaction force

Fig. 6 shows the interaction force curves under handshaking tasks by the existed learning-based tracking control method and the proposed fuzzy logic nonzero-sum game-based distributed approximate optimal control method. In the process of pHRI, the interaction force acts on the contact part of the participant with the robot that is the direct factor affecting the degree of comfort. The contact force curves appear in a two-dimensional space, due to the fact that the MRM has 2-DoF and the joint axis is assembled in parallel. It can be seen that the interaction force is gentle and mild less than or equal to 0.1 N, under the fuzzy logic-based nonzero-sum game scheme, and without strong chattering phenomenon. This facilitates the accurate implementation of the fuzzy logic technique, which guarantees the comfort and security of the human.

(4) Fuzzy logic weight
Fig. 7 shows the fuzzy logic weight curves under handshaking tasks obtained by the proposed fuzzy logic nonzero-sum game-based distributed approximate optimal control method. The fuzzy logic nonzero-sum game-based distributed approximate optimal control policies can be obtained using the converged weights (cf. Fig. 5). Thus, the weight can reflect the human intention in real time according to the position tracking error (cf. Fig. 4).

Based on the experimental results, the closed-loop MRM systems under the HRC task have better performance than the existed methods in terms of position tracking, control torque, and interaction force under the proposed fuzzy logic nonzero-sum game-based distributed approximate optimal control method.

5. Conclusion

This paper proposed a fuzzy logic nonzero-sum game-based distributed approximate optimal control scheme for MRMs under HRC tasks. The MRM dynamic model is considered as an integration of joint subsystem models associated with interconnected dynamic coupling effects. Based on the differential game strategy, the optimal control problem of closed-loop robotic systems is transformed into a nonz-ero-sum game issue of multiple subsystems. Using the ADP algorithm, the cost function is developed by fuzzy logic and implemented to solve the coupled HJ equation, which facilitates the acquisition of Nash equilibrium solutions. The trajectory tracking error of the closed-loop MRM system is UUB under the HRC task using the Lyapunov theory. Finally, the performed experiments demonstrate the efficiency of the proposed method.

Security remains a major barrier to the widespread utilize of adaptive dynamic programming algorithms in the real world, especially for MRM system under pHRI tasks. How to guarantee the task performance and safety characteristic of MRM system is a bottleneck problem that demands to be solved urgently.

CRediT authorship contribution statement

Tianjiao An: Writing - original draft. **Xinye Zhu:** Software. **Mingchao Zhu:** Data curation. **Bo Dong:** Conceptualization, Methodology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] F. Ferraguti, L. Talignani, L. Sabattini, M. Bonfe, C. Fantuzzi, C. Secchi, A variable admittance control strategy for stable physical human-robot interaction, Int. J. Robot. Res. 38 (6) (2019) 747–765.
[2] Y. She, S. Song, H. Su, J. Wang, A parametric study of compliant link design for safe physical human-robot interaction, Robotica. 39 (10) (2021) 1739–1759.
[3] J. Neumann, Zur Theorie der Gesellschaftsspiele, Math. Ann. 100 (1928) 295–320.
[4] N. Jarrasse, T. Charalambous, E. Burdet, A framework to describe, analyze and generate interactive motor behaviors, PLoS ONE. 7 (2012) .
[5] Y. Li, G. Carboni, F. Gonzalez, D. Campolo, E. Burdet, Differential game theory for versatile physical human-robot interaction, Nat. Mach. Intell. 1 (1) (2019) 36–43.
[6] Y. Li, K. Tee, R. Yan, W. Chan, Y. Wu, A framework of human-robot coordination based on game theory and policy iteration, IEEE T. Robot. 32 (6) (2016) 1408–1418.
[7] B. Ma, Y. Li, T. An, B. Dong, Compensator-critic structure-based neuro-optimal control of modular robot manipulators with uncertain environmental contacts using non-zero-sum games, Knowl-Based Syst. 224 (2021) .
[8] H. Zhang, J. Zhang, G. Yang, Y. Luo, Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming, IEEE T. Fuzzy Syst. 23 (1) (2015) 152–163.
[9] H. Jiang, H. Zhang, X. Xie, J. Han, Neural-network-based learning algorithms for cooperative games of discrete-time multi-player systems with control constraints via adaptive dynamic programming, Neurocomputing. 344 (2019) 13–19.
[10] C. Li, J. Ding, F. Lewis, Y. Chai, A novel adaptive dynamic programming based on tracking error for nonlinear discrete-time systems, Automatica. 129 (2021) 1.
[11] Q. Wei, R. Song, Z. Liao, B. Li, F. Lewis, Discrete-time impulsive adaptive dynamic programming, IEEE T. Cybernetics. 50 (10) (2020) 4293–4306.
[12] S. Xue, B. Luo, D. Liu, Y. Yang, Constrained event-triggered H_∞ control based on adaptive dynamic programming with concurrent learning, IEEE T. Syst. Man Cy. A. 52 (1) (2022) 357–369.
[13] B. Zhao, D. Liu, Event-triggered decentralized tracking control of modular reconfigurable robots through adaptive dynamic programming, IEEE T. Ind. Electron. 67 (4) (2020) 3054–3064.
[14] L. Liu, W. Zhao, Y. Liu, S. Tong, Y. Wang, Adaptive finite-time neural networks control of nonlinear systems with multiple objective constraints and application to electromechanical system, IEEE T. Neur. Net. Lear. 32 (12) (2021) 5416–5426.
[15] Z. Li, L. Wu, Y. Xu, S. Moazeni, Z. Tang, Multi-stage real-time operation of a multi-energy microgrid with electrical and thermal energy storage assets: a data-driven MPC-ADP approach, IEEE T. Smart Grid. 13 (1) (2022) 213–226.
[16] Q. Wei, Z. Liao, R. Song, P. Zhang, Z. Wang, J. Xiao, Self-learning optimal control for ice-storage air conditioning systems via data-based adaptive dynamic programming, IEEE T. Ind. Electron. 68 (4) (2021) 3599–3608.
[17] L. Cui, S. Wang, J. Zhang, D. Zhang, J. Lai, Y. Zheng, Z. Zhang, Z. Jiang, Learning-based balance control of wheel-legged robots, IEEE Robot Autom. Let. 6 (4) (2021) 7667–7674.
[18] Y. Liu, W. Zhao, L. Liu, D. Li, S. Tong, C.L. Chen, Adaptive neural network control for a class of nonlinear systems with function constraints on states, IEEE T. Neur. Net. Lear. doi: 10.1109/TNNLS.2021.3107600.
[19] X. Wang, D. Ding, H. Dong, X. Zhang, Neural-network-based control for discrete-time nonlinear systems with input saturation under stochastic communication protocol, IEEE-CAA J. Automatic. 8 (4) (2021) 766–778.
[20] S. Zhang, B. Zhao, D. Liu, Y. Zhang, Observer-based event-triggered control for zero-sum games of input constrained multi-player nonlinear systems, Neural Networks. 144 (2021) 101–112.
[21] B. Dong, T. An, X. Zhu, Y. Li, K. Liu, Zero-sum game-based neuro-optimal control of modular robot manipulators with uncertain disturbance using critic only policy iteration, Neurocomputing. 450 (2021) 183–196.
[22] H. Zhang, J. Park, D. Yue, Z. Zhao, Nearly optimal integral sliding-mode consensus control for multiagent systems with disturbances, IEEE T. Syst. Man Cy. A. 51 (8) (2021) 4741–4750.
[23] X. Gao, T. Li, L. Yuan, W. Bai, Robust fuzzy adaptive output feedback optimal tracking control for dynamic positioning of marine vessels with unknown disturbances and uncertain dynamics, Int. J. Fuzzy Syst. 23 (7) (2021) 2283–2296.
[24] H. Su, H. Zhang, H. Jiang, Y. Wen, Decentralized event-triggered adaptive control of discrete-time nonzero-sum games over wireless sensor-actuator networks with input constraints, IEEE T. Neur. Net. Lear. 31 (10) (2020) 4254–4266.
[25] L. Liu, Y. Liu, D. Li, S. Tong, Z. Wang, Barrier Lyapunov function-based adaptive fuzzy FTC for switched systems and its applications to resistance-inductance-capacitance circuit system, IEEE T. Cybernetics. 50 (8) (2020) 3491–3502.
[26] W. Rudin, Principles of mathematical analysis, New York, NY, McGraw-Hill, USA, 1976.
[27] L. Zadeh, Fuzzy sets, Inform. Control 8 (1965) 338–353.
[28] S. Luo, F. Lewis, Y. Song, H. Ouakad, Accelerated adaptive fuzzy optimal control of three coupled fractional-order chaotic electromechanical transducers, IEEE T. Fuzzy Syst. 29 (7) (2021) 1701–1714.
[29] L. An, G. Yang, Optimal transmission power scheduling of networked control systems via fuzzy adaptive dynamic programming, IEEE T. Fuzzy Syst. 29 (6) (2021) 1629–1639.
[30] S. Tong, K. Sun, S. Sui, Observer-based adaptive fuzzy decentralized optimal control design for strict-feedback nonlinear large-scale systems, IEEE T. Fuzzy Syst. 26 (2) (2018) 569–584.
[31] G. Liu, Y. Liu, A. Goldenberg, Design, analysis, and control of a spring-assisted modular and reconfigurable robot, IEEE-ASME T. Mech. 16 (4) (2011) 695–706.
[32] M. Rahman, R. Ikeura, K. Mizutani, Investigation of the impedance characteristic of human arm for development of robots to cooperate with humans, JSME Int. J. Series C. 45 (2) (2002) 510–518.

[33] X. Yu, Y. Li, S. Zhang, C. Xue, Y. Wang, Estimation of human impedance and motion intention for constrained human-robot interaction, Neurocomputing. https://10.1016/j.neucom.2019.07.104.
[34] K. Sun, Y. Li, S. Tong, Fuzzy adaptive output feedback optimal control design for strict-feedback nonlinear systems, IEEE T. Syst. Man Cy. A. 47 (1) (2017) 33–44.
[35] Y. Li, X. Zhu, T. An, B. Dong, Critic-observer-based decentralized force/position approximate optimal control for modular and reconfigurable manipulators with uncertain environmental constraints, Complex Intell. Syst. 2022 (8) (2022) 525–539.
[36] J. Zhao, J. Na, G. Gao, Robust tracking control of uncertain nonlinear systems with adaptive dynamic programming, Neurocomputing. 471 (2022) 21–30.

Tianjiao An received his M.S. degree from Changchun University of Technology, China, in 2020. He is currently working towards a Ph.D. degree in the Department of Control Science and Engineering, Changchun University of Technology, China. His research interests include robot control, and adaptive dynamic programming.

Xinye Zhu received his B.S. and M.S. degrees from the Changchun University of Technology, China, in 2018 and 2021, respectively. His research interests include robot control and adaptive dynamic programming.

Mingchao Zhu received his M.S. and Ph.D degrees from Jilin University, China in 2006 and 2009, respectively. He is currently a researcher in the CAS Key Laboratory of On-orbit Manufacturing and Integration for Space Optics System, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin, China and the Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing, China. His research interests include intelligent mechanical and robot control.

Bing Ma received her B.S. degree and Ph.D degree from Changchun University of Technology, China in 2016 and 2021,respectively. She is currently a lecturer in Changchun University of Technology. Her research interests include robot control, position-force control and adaptive dynamic programming.

Bo Dong received his M.S. and Ph.D degrees from Jilin University, China in 2012 and 2015, respectively. He is currently a Professor in Changchun University of Technology, China. His research interests include intelligent mechanical and robot control