Event-triggered-based cooperative game optimal tracking control for modular robot manipulator with constrained input

Yuanchun Li · Chongyang Wei · Tianjiao An · Bing Ma · Bo Dong

Received: 24 November 2021 / Accepted: 1 June 2022 / Published online: 23 June 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract In this paper, a cooperative game optimal tracking control method based on event-triggered mechanism for constrained input modular robot manipulators (MRMs) system is introduced. According to the joint torque feedback technique, the dynamics model of the constrained input subsystem is established and the global state space equation is derived. The control inputs of \( n \) joints in the MRM system with constrained input are taken as \( n \) participants of cooperative game, and the tracking control problem of the manipulator system is transformed into the optimal control problem based on the cooperative game. Next, a fusion function containing position and velocity errors is defined to construct the performance index function. In order to improve the control performance and robustness of the manipulator system, part of the known model information is used to devise controller, the model uncertainty is dealt by the neural network (NN) observer, and the optimal compensation control strategy is used to deal with internal disturbance such as sensor measurement error and transmission ripple due to power fluctuations, electromagnetic effects, noise and vibration. Based on the adaptive dynamic programming algorithm and event-triggered mechanism, the optimal tracking control strategy is obtained by approximately solving the event-triggered Hamilton-Jacobi-Bellman equation with the critic NN. The Lyapunov theory proves that trajectory tracking error of MRM system with constrained input is uniformly ultimately bounded. Finally, the experimental results demonstrate the effectiveness of the proposed control method.

Keywords Modular robot manipulators · Constrained input · Cooperative game [1] · Adaptive dynamic programming · Event-triggered mechanism

1 Introduction

Modular robot manipulators (MRMs) [2,3] are composed of modules with standard interfaces, which have high flexibility and wide adaptability, and have attracted great attention from researchers in the field of robotics. Compared with the traditional robot which has a single configuration and cannot be combined arbitrarily, the MRM can be added or reduced modules to accomplish different tasks, such as deep-sea exploration [4], disaster rescue [5] and extreme environment operations [6]. Constrained input [7] is a com-
mon phenomenon in the actual engineering system due to the constraint of physical conditions. For example, the output torque and speed of the motor have maximum values, and the valve switch cannot be adjusted arbitrarily large. The occurrence of these situations is often accompanied by the degradation of system performance and even loss of stability. Since each module of MRM system contains a DC motor, the research on constrained input MRM system has profound significance. For a class of constrained input nonlinear systems, an antisaturation adaptive backstepping controller based on wavelet network is designed in [8]. In [9], a time-varying sliding mode control scheme for the control constrain problem of uncertain nonlinear systems is presented. A backstepping controller is designed by using hyperbolic tangent function to deal with saturation function in [10]. The constrained input problem of the controller is regarded as the constraint condition of the objective function to be optimized in [11]. However, the methods mentioned above ignore the comprehensive optimization of control performance and energy consumption. To solve this problem, optimal control is introduced into the controller design.

The optimal control theory includes two important branches of dynamic programming theory [12] and maximum principle [13]. On account of dynamic programming can cause curse of dimensionality [14], Werbos et al. [15] first proposed the idea of using neural network (NN) adaptive approximation to solve Bellman equation, which laid a foundation for the appearance of NN approximators in adaptive dynamic programming (ADP). In the 21st century, the ADP method has been widely recognized and developed in theory. Lewis et al. [16] proved the convergence of Hamilton-Jacobi-Bellman (HJB) equation by using the heuristic dynamic programming (HDP) algorithm based on value iteration. Jiang et al. [17] elaborated a value iteration algorithm for a continuous-time system with completely unknown dynamics and conducted a strict convergence analysis on the algorithm. Liu et al. [18] used the policy iteration ADP algorithm to solve the optimal control of the nonlinear system and analyzed the convergence. Cheng et al. [19] transformed the adaptive cruise control problem into the optimal tracking control problem of complex nonlinear system and then designed the adaptive cruise controller by using ADP and experience playback technology. Wei et al. [20] elaborated a distributed policy iteration algorithm, which only updated the control strategy of one controller at a time, effectively reducing the calculation burden of each iteration. Shi et al. [21] stated a data-based controller design method for the output feedback optimal control problem of continuous-time system with dynamic uncertainty through ADP technology, which improved the robustness of the system. With the continuous exploration of researchers, ADP is not only applied in theory to devise controllers of linear/nonlinear systems with input/output constraints [22,23], external disturbances [24], mismatched interconnections [25], but also devoted to practical systems such as manipulator system [26], marine vessels system [27], power system [28], network control system [29].

For a common MRM system, if the control input of each joint module is considered as a decision-maker, the MRM system can be supposed to an interactive system with multi-player dynamic decision-making. Differential game combining optimal control and game theory is an ideal tool to deal with multi-player decision problem, especially the ADP-oriented differential game has made numerous research achievements [30,31]. According to the cooperative or competitive relationship among participants, game problem can be mainly divided into zero-sum differential game [32], nonzero-sum differential game [33] and fully cooperative differential game [34]. The participants in the zero-sum game are in complete competition relation. For the two-player zero-sum game, the idea is that both players expect to choose appropriate strategies to maximize and minimize the value function, respectively, which is similar to the design of $H_{\infty}$ optimal controller [35]. Jiang et al. [36] used the ADP algorithm to overcome the difficulty of accurately finding the analytic solution of Hamilton-Jacobi-Isaacs (HJI) equation for the multi-player zero-sum differential game problem and adopted the form of single network to reduce the structural complexity. Dong et al. [37] elaborated a zero-sum differential game based on ADP for the optimal tracking problem of MRMs system with uncertain disturbance and successfully solved the optimal control strategy by using the single critic NN. For nonzero-sum differential game, there is both competition and cooperation among players, and each player expects to select strategy to minimize their performance index function. For the nonzero-sum game problem of multi-player unknown nonlinear system, [38] designed an actor-critic structure to approximate optimal solution by using the data-based ADP algorithm. [39] transformed the optimal control of a nonlinear intercon-
nected system into a nonzero-sum differential game problem, and the optimal solution of Hamilton-Jacobi (HJ) equation was obtained by using the proposed distributed ADP algorithm. Zhao et al. [40] proposed a reinforcement learning (RL) method to solve HJ equation for the nonlinear system optimal control problem of nonzero-sum game. Ma et al. [41] transformed the optimal tracking problem into a nonzero-sum game problem and introduced compensator-critic structure to solve the optimal solution of MRM system under uncertain environment. In the cooperative game, there is only a cooperative relationship among players, and each player expects to choose a strategy to minimize the value function for the team as a whole. Zhang et al. [42] exposted the data-driven ADP algorithm to study the cooperative game problem with constrained input. Li et al. [43] expounded a new adaptive Q-learning algorithm to solve cooperative linear quadratic dynamic game. Mu et al. [1] transformed the cooperative differential game problem into the optimal control problem by defining the global performance index function and successfully applied cooperative game to power system.

It is worth noting that most of the above literatures update the designed controller with time-triggered method of fixed sampling period. Due to the advantages of computational efficiency, event-triggered control methods have attracted extensive attention from academic circles in recent years. In consideration of MRM system’s unexpected computational burden and the waste of communication resources caused by sampling periodicity in stable operation, event-triggered mechanism needs to be introduced into the design of MRM system controller. The robust control of nonlinear system was investigated based on event-triggered in [44] which not only has certain robustness but also saves the communication resources between the controller and the control plants. Xue et al. [45] converted the constrained tracking control problem into the optimal adjustment problem through integral reinforcement learning algorithm and then applied the event-triggered mechanism to reduce the data transmission pressure. Mu et al. [46] elaborated a modified event-triggered optimal control scheme by combining global dual heuristic dynamic programming (GDHP) algorithm with event-triggered mechanism to dispose constrained input problems of nonlinear system. Zhu et al. [47] proposed an event-triggered control method for constrained input continuous time nonlinear system and successfully applied it to the overhead crane system. Based on ADP algorithm and event-triggered mechanism, [48] and [49] considered zero-sum and nonzero-sum differential game of continuous-time nonlinear system, respectively, and solve their optimal control problems. However, the optimal control scheme of cooperative game based on event-triggered mechanism is rarely studied.

Inspired by the above literatures, this paper proposes an event-triggered cooperative game optimal tracking control method for MRM system with constrained input. Different from previous optimal control methods [47,50], this paper focuses on cooperative game and fully considers the superiority of game theory in dealing with multi-player dynamic decision-making. Taking each joint module control input of the MRM system with constrained input as a decision-maker, it can be assumed as a multi-player interactive system. Cooperative game theory considers the information exchange between each joint of the manipulator system. The overall control performance and energy consumption of the MRM system are optimized by the cooperation of all joints. And the cooperative game has been successfully applied to real robot system. Then, the event-triggered control method can greatly reduce the updating times of the controller and effectively save the communication burden between the controller and the manipulator. The contributions of this paper are summarized as follows:

1. To the authors’ knowledge, this is the first time that cooperative game is introduced into a real robot system such as MRM. Considering the information exchange among the joints of the MRM system with constrained input, the control input of $n$ joint modules is regarded as $n$ participants in the cooperative game; then, the tracking control problem of the manipulator system can be transformed into the optimal control problem of the cooperative game among $n$ participants.

2. Considering the unexpected computational burden of MRM system and the communication resource waste caused by periodic sampling during stable operation, we introduce event-triggered mechanism based on cooperative game optimal tracking control method. By solving the event-triggered HJB equation, the cooperative game optimal tracking control strategy with aperiodic sampling is
obtained, which effectively reduces the updating times of the manipulator system controller.

The rest of the article is arranged as follows: In Sect. 2, the dynamic model of the constrained input MRM system is established. The design of optimal tracking controller for cooperative game based on event-triggered mechanism is discussed in Sect. 3. The experimental demonstration and conclusion are in Sects. 4 and 5.

2 Problem statement and dynamic model

For an n-degree of freedom (n-DOF) MRM system, it is actually composed of n motors and connecting rods in series. Due to the limitation of physical conditions, the output torque of the motor has a maximum value. This saturation phenomenon will lead to the system control performance degradation and even loss of stability. Therefore, considering a constrained input MRM system with n-DOF and referring to the joint torque feedback (JTF) technique modeling method of the manipulator [51], the dynamic model of the constrained input MRM system can be represented by a constrained input subsystem with interconnected dynamic coupling (IDC). The dynamic model of the ith constrained input subsystem can be shown as:

\[ I_{mi} \dot{q}_i + f_i(q_i, \dot{q}_i) + z_i(q, \dot{q}, \ddot{q}) + \frac{\tau_{si}}{\gamma_i} \dot{q}_i = \tau_i, \quad \| \tau_i \| \leq \beta_i, \tag{1} \]

where the subscript i reflects the ith joint subsystem, \( q_i \), \( \dot{q}_i \) and \( \ddot{q}_i \) are the ith joint position, velocity and acceleration, \( \tau_{si} \) indicates joint coupling torque which can be measured by torque sensors, \( \tau_i \) denotes the control torque which is less than or equal to a known positive constant \( \beta_i \), \( I_{mi} \) refers to the moment of inertia of the rotor relative to the rotation axis, \( \gamma_i \) represents the gear ratio, \( f_i(q_i, \dot{q}_i) \) denotes the joint friction torque, \( z_i(q, \dot{q}, \ddot{q}) \) means dynamic coupling torque among the subsystems and \( d_i(q_i) \) is the internal disturbance torque. The properties of friction term, IDC term and internal disturbance torque term in (1) are analyzed as follows.

(1) Joint friction

For joint friction torque \( f_i(q_i, \dot{q}_i) \), it can be considered as a function of joint position and velocity, which is defined as follows:

\[ f_i(q_i, \dot{q}_i) = f_{bi} \dot{q}_i + \left( f_{si} e^{-f_{si} \dot{q}_i^2} + f_{ci} \right) \text{sgn} (\dot{q}_i) + f_{pi} (q_i, \dot{q}_i), \quad \tag{2} \]

where \( f_{bi} \) is viscous friction coefficient, \( f_{si} \) denotes static friction-related parameter, \( f_{ci} \) is a positive parameter related to Stribeck effect, \( f_{pi}(q_i, \dot{q}_i) \) is the position dependency of friction term, \( \text{sgn} (\dot{q}_i) \) stands for sign function which has the following form:

\[ \text{sgn} (\dot{q}_i) = \begin{cases} 1 & \text{if } \dot{q}_i > 0 \\ 0 & \text{if } \dot{q}_i = 0 \\ -1 & \text{if } \dot{q}_i < 0 \end{cases}. \tag{3} \]

According to the linearization scheme [52], the friction term can be modified as:

\[ f_i(q_i, \dot{q}_i) \approx \hat{f}_{bi} \dot{q}_i + \left( \hat{f}_{si} e^{-\hat{f}_{si} \dot{q}_i^2} + \hat{f}_{ci} \right) \text{sgn} (\dot{q}_i) + \hat{f}_{pi}(q_i, \dot{q}_i) + Y(q_i) \hat{F}_i, \tag{4} \]

where \( \hat{f}_{bi}, \hat{f}_{ci}, \hat{f}_{si} \) and \( \hat{f}_{ti} \) are estimations of friction parameters \( f_{bi}, f_{ci}, f_{si}, \) and \( f_{ti} \), respectively, \( \hat{F}_i = [\hat{f}_{bi} - f_{bi}, \hat{f}_{ci} - f_{ci}, \hat{f}_{si} - f_{si}, \hat{f}_{ti} - f_{ti}]^T \) indicates the parameter uncertainty vector of friction term, \( Y(q_i) \) is defined in the following form:

\[ Y(q_i) = \begin{bmatrix} \dot{q}_i, \text{sgn} (\dot{q}_i), e^{-\hat{f}_{si} \dot{q}_i^2} \text{sgn} (\dot{q}_i), \\ -\hat{f}_{si} \dot{q}_i^2 e^{-\hat{f}_{si} \dot{q}_i^2} \text{sgn} (\dot{q}_i) \end{bmatrix}. \tag{5} \]

Property 1 It can be seen from (2) and (4) that \( f_{bi}, f_{ci}, f_{si} \) and \( f_{ti} \) and their estimated values are bounded, hence each element in the parameter uncertainty vector of friction term \( \hat{F}_i \) is also bounded to satisfy \( |\hat{F}_{il}| \leq \eta_{F_{il}} \), where \( \eta_{F_{il}} \) is a given positive constant , the subscript l represents the lth element of vector \( \hat{F}_i \) and \( l = 1, 2, 3, 4 \). The approximation error of friction \( Y(q_i) \hat{F}_i \) satisfies \( |Y(q_i) \hat{F}_i| \leq |Y(q_i)| \eta_{F_{F11}}, \eta_{F_{F12}}, \eta_{F_{F13}}, \eta_{F_{F14}}^T \). In addition, \( \hat{f}_{pi}(q_i, \dot{q}_i) \) is bounded as \( \| \hat{f}_{pi}(q_i, \dot{q}_i) \| \leq \eta_{f_{pi}} \), where \( \eta_{f_{pi}} \) is a positive constant.

Remark 1 As \( f_{bi}, f_{ci}, f_{si} \) and \( f_{ti} \) are all physical parameters of practical significance in the bound friction torque passively generated by motor rotation, the friction term parameters and their estimated values are bounded. Further, we can learn from (2) and (4) that the frictional parameter uncertainties \( f_{bi} - \hat{f}_{bi}, f_{ci} - \hat{f}_{ci}, f_{si} - \hat{f}_{si} \) and \( f_{ti} - \hat{f}_{ti} \) are also bounded.
Remark 2 The desired trajectories tracked in experiments and projects are given by users in practical applications. Therefore, it is feasible to assume that there is a usable positive constant for the uncertain parameter. The upper limit value of the uncertainty parameter can be easily determined according to the actual characteristics of the model uncertainty.

(2) Interconnected dynamic coupling For the IDC term, it can be considered as a complex nonlinear function as follows:

\[ z_i(q, \dot{q}, \ddot{q}) = I_{mi} \sum_{c=1}^{i-1} z_{mi}^T \dot{z}_{lc} \ddot{q}_c + I_{mi} \sum_{c=1}^{i-1} z_{mi} z_{lk} \dot{q}_k \dot{q}_c \]

\[ + I_{mi} \sum_{c=2}^{i-1} \sum_{k=1}^{c-1} z_{mi}^T (z_{lk} \times z_{lc}) \dot{q}_k \dot{q}_c, \]

where \( z_{mi} \) is the unit vectors along the axis of rotation the \( i \)th rotor, \( z_{lk} \) and \( z_{lc} \) denote the unit vectors along the axis of rotation the \( k \)th and \( c \)th joint.

Redefine (6) to the following form:

\[ z_i(q, \dot{q}, \ddot{q}) = I_{mi} \sum_{c=1}^{i-1} z_{mi} \dot{q}_k \dot{q}_c + I_{mi} \sum_{c=1}^{i-1} z_{mi} \dot{q}_c \]

\[ = \sum_{c=2}^{i-1} \sum_{k=1}^{c-1} I_{mi} \tilde{\Theta}_{kc} \dot{q}_k \dot{q}_c + \sum_{c=1}^{i-1} I_{mi} \tilde{\Theta}_{kc} \dot{q}_c \]

\[ = \sum_{c=2}^{i-1} \sum_{k=1}^{c-1} \left[ I_{mi} \tilde{\Theta}_{kc} \dot{q}_k \dot{q}_c \right]^T + \left( I_{mi} \tilde{\Theta}_{kc} \dot{q}_c \right)^T \]

\[ \triangleq z_i z_i + z_i \theta, \]

where \( z_i z_i = z_{mi}^T z_{lk} \), \( \dot{q}_k \dot{q}_c = z_{mi}^T (z_{lk} \times z_{lc}) \). \( \tilde{\Theta}_{kc} \) and \( \dot{\Theta}_{kc} \) are estimations of \( \dot{\Theta}_{kc} \) and \( \dot{\Theta}_{kc} \) that satisfy \( z_i z_i = z_{mi}^T \tilde{\Theta}_{kc} \) and \( \dot{\Theta}_{kc} = \dot{\Theta}_{kc} - \dot{\Theta}_{kc} \), respectively. \( \tilde{\Theta}_{kc} \) and \( \dot{\Theta}_{kc} \) are alignment errors.

Property 2 Since \( z_{mi} \), \( z_{lk} \) and \( z_{lc} \) are all unit vectors, we have \( |z_{mi}| = |z_{mi}^T z_{lk}| \leq 1 \) and \( |\dot{\Theta}_{kc}| = |z_{mi}^T (z_{lk} \times z_{lc})| \leq 1 \). We also know that if the \( i \)th and \( c \)th joints \( (1 < c, k < i - 1) \) are stable, \( z_i (q, \dot{q}, \ddot{q}) \) is bounded. \( |z_i z_i| \) and \( |z_i \theta| \) satisfy inequalities \( |z_i z_i| \leq \eta_{z_i z_i} \) and \( |z_i \theta| \leq \eta_{z_i \theta} \), where \( \eta_{z_i z_i} \) and \( \eta_{z_i \theta} \) are known positive constants.

(3) Internal disturbance torque \( d_i(q_i) \) stands for internal disturbance torque, which mainly due to power fluctuations, electromagnetic effects, noise and vibration caused by the torque sensor measurement deviation.

(4) State space formulation

According to the dynamic model of the \( i \)th constrained input joint subsystem (1), we can define the state variable \( xi = [x_{i1}, x_{i2}]^T = [q_i, \dot{q_i}]^T \in \mathbb{R}^{2 \times 1} \) and the parameter \( Bi = (Imi \gamma_i)^{-1} \in \mathbb{R}^+ \). Then, the state space equation of the \( i \)th constrained input subsystem can be written as:

\[ \Gamma_i : \begin{cases} \dot{x}_{i1} = x_{i2} \\ \dot{x}_{i2} = \Xi_i(x_i) + \Psi_i(x_i) + h_i(x_i) + B_i u_i. \end{cases} \]

where \( \Xi_i(x_i) = -B_i \left( \hat{f}_{pi} \hat{q}_i + \left( \hat{f}_{si} \hat{q}_i \tilde{\tau}_i \hat{\tau}_i \right) \right) \) means the known and measurable part of the subsystem dynamic model, \( \Psi_i(x_i) = -B_i \left( \hat{f}_{pi} \hat{q}_i + f_{pi} (q_i, \dot{q}_i) + Y_i (q_i, \tilde{F}_i) \right) \) includes model uncertainties such as IDC term and friction error, \( h_i(x_i) = -B_i d_i q_i \) is internal disturbance term, \( u_i = \tau_i \in \mathbb{R}^{1 \times 1} \) refers to control torque and is less than or equal to a known positive constant \( \beta_i \).

Next, a cooperative game optimal tracking control method is proposed by transforming the tracking control problem of manipulator system into cooperative game problem. On the basis of the above control method, the event-triggered mechanism is applied to change the traditional periodic sampling method into non-periodic sampling, so as to reduce the computational burden while ensuring the control performance.

3 Event-triggered-based cooperative game optimal tracking control

3.1 Problem transformation

Since the saturation function is not differentiable, we resort the hyperbolic tangent function \( \tanh (\cdot) \) with well mathematical properties such as continuous differentiability and boundedness to construct the smooth saturation function. Instead of directly employing the hyperbolic tangent function itself, the saturation function is employed by multiplying its independent variable and amplitude by a certain value. The saturation function can satisfy the requirement that the controller’s gain is bounded and the rate of change is adjustable. Referring to [47, 53], the saturation function of this paper is
defined as follows:
\[
y = f(u) = \beta \tanh \left( \frac{u}{\beta} \right),
\]
(9)
where \( \beta > 0 \) is the amplitude coefficient factor of the function, and the critical value of the function is affected by the magnitude of \( \beta \).

For a constrained input MRM system with \( n \)-DOF, the trajectory tracking control problem of the manipulator system can be transformed into an optimal control problem based on cooperative game when the control inputs of \( n \) joint modules’ manipulator system are regarded as \( n \) participants in the cooperative game. Define \( x = [x_1, x_2]^T = [q, \dot{q}]^T \in \mathbb{R}^{2n \times 1} \) as a global state vector of the constrained input MRM system, where \( x_1 = [x_{11}, x_{21}, \ldots, x_{11}, \ldots, x_{n1}]^T \in \mathbb{R}^{n \times 1} \), \( x_2 = [x_{12}, x_{22}, \ldots, x_{12}, \ldots, x_{n2}]^T \in \mathbb{R}^{n \times 1} \). The state space formulation of the augmented subsystem with \( n \) constrained participants is as follows:
\[
\Gamma : \begin{cases} 
\dot{x}_1 = x_2 \\
\dot{x}_2 = \Xi(x) + \Psi(x) + h(x) + \sum_{i=1}^{n} D_{ii} u_i, \quad (10) 
\end{cases}
\]
where \( \Xi(x) = [\Xi_1(x_1), \ldots, \Xi_n(x_n)]^T \in \mathbb{R}^{n \times 1} \), \( \Psi(x) = [\Psi_1(x_1), \ldots, \Psi_2(x_2), \ldots, \Psi_n(x_n)]^T \in \mathbb{R}^{n \times 1}, h(x) = [h_1(x_1), \ldots, h_n(x_n)]^T \in \mathbb{R}^{n \times 1}, \)
\( D_{ii} = \begin{bmatrix} 0, \ldots, B_{ii}, \ldots, 0 \end{bmatrix} \in \mathbb{R}^{n \times 1} \).

In order to achieve the optimal trajectory tracking control tasks, we define a fusion error function \( s(x) = x_2 - x_{2d} + \alpha_s (x_1 - x_{1d}) \) including joint position and joint velocity errors. According to the augmented state space description (10), the continuously differentiable infinite horizon global performance index function is defined as:
\[
J(s, u_1, \ldots, u_n) = \int_0^\infty \left( s^T (x(t)) Q s(x(t)) + \sum_{i=1}^{n} U_i (u_i) + h_m^T h_m \right) dt,
\]
(11)
where \( x_{1d}, x_{2d} \) and \( \dot{x}_{2d} \) refer to the desired position, velocity and acceleration of the joint, respectively, \( Q \) is a symmetric positive definite matrix, such that \( Q^T = Q, \alpha_s = diag(\alpha_1, \alpha_2, \ldots, \alpha_n) \) is a diagonal matrix, \( h_m = [h_{1m}, \ldots, h_{im}, \ldots, h_{nm}]^T \in \mathbb{R}^{n \times 1} \) is a function vector such that the value of \( h_{im} \) is more than or equal to \( h_i \) at any time. Obviously, the internal disturbance term \( h(x) \) and the internal disturbance upper bound function vector \( h_m \) satisfy the relation
\[
h^T (x) h(x) \leq h_m^T h_m.
\]
\( U_i (u_i) \) is a positive definite nonquadratic function that handles constrained input. \( \Upsilon_i (\Omega) \) is a set of admissible control strategies.

**Definition 1** (Admissible control) [54]. For \( \forall i, i = 1, 2, \ldots, n \), a feedback control law \( u_i \) is defined as admissible to performance index function (11) on the compact set \( \Omega \), denoted by \( u_i \in \Upsilon_i (\Omega) \), if \( u_i \) is continuous on \( \Omega \) with \( u_i (0) = 0, u_i \) can stabilize the system (10) on \( \Omega \) and guarantee that the performance index function \( J(s, u_1, \ldots, u_n) \) (11) is finite.

**Remark 3** The participant of cooperative game problem is the control input of MRM system’s each joint, and the optimization target is also the overall performance and energy consumption, which involves all information of the system. Hence, the performance index function based on cooperative game design includes the state of the whole system.

**Remark 4** The upper bound of internal disturbance term \( h_i \) is \( \eta_{him} \), whose existence ensures the stable operation of MRM system when it performs feasible tasks. If the internal disturbance does not possess a finite boundary, the trajectory tracking error of the manipulator system cannot be guaranteed to be uniformly ultimately bounded (UUB).

**Remark 5** Function \( h_{im} \) has an upper bound \( \eta_{him} \), and the system is stable and completely controllable. A fully controllable MRM system does not have uncontrollable or unstable state variables. Therefore, \( J(s, u_1, \ldots, u_n) \) is finite.

The positive definite nonquadratic function \( U_i (u_i) \) is expressed as follows:
\[
U_i (u_i) = \int_0^{u_i} 2\beta_i \tanh^{-1} \left( \frac{u_i}{\beta_i} \right) R_i dw,
\]
(12)
where \( R_i \in \mathbb{R}^{1 \times 1} \) means a positive definite matrix. By integrating \( U_i (u_i) \), we can solve the expression:
\[
U_i (u_i) = 2\beta_i u_i R_i \tanh^{-1} \left( \frac{u_i}{\beta_i} \right) + \beta_i^2 R_i \ln \left( 1 - \frac{u_i^2}{\beta_i^2} \right).
\]
(13)
Remark 6 For the purpose of satisfy the requirements of differentiable saturation function, bounded controller gain and adjustable rate of change, the saturation function (9) is introduced. For the sake of ensuring the differentiable performance index function and the comprehensive optimization of the control performance and energy consumption of the constrained input MRM system, we introduce the positive definite quadratic form of the error function and the positive definite non-quadratic form of the control torque [47, 53] into the integrand function of the performance index.

On the basis of Definition 1 and performance index function (11), we can get:

\[ 0 = s^T \left( x \left( i \right) \right) Q s \left( x \left( i \right) \right) + h_m^T h_m + \sum_{i=1}^{n} U_i \left( u_i^* \right) \]

\[ + \nabla J^T \left( s \right) \left( \Xi \left( s \right) + \Psi \left( s \right) + h \left( s \right) + \vartheta + \sum_{i=1}^{n} D_i u_i \right), \]

where \( \vartheta = [\vartheta_1, \vartheta_2, \cdots, \vartheta_n]^T = -\dot{x}_{2d} + \sigma_s \left( \dot{x}_1 - \dot{x}_{1d} \right) \) is the remainder of \( \dot{s} \left( x \right) \) expansion, \( \nabla J \left( s \right) \) denotes the partial derivative of \( J \left( s \right) \). The Hamiltonian function can be defined based on (14):

\[ H \left( s, u_1, \cdots, u_n, \nabla J \right) = s^T Q s + \sum_{i=1}^{n} U_i \left( u_i \right) \]

\[ + h_m^T h_m + \nabla J^T \left( \Xi + \Psi + h + \vartheta \right) \]

\[ + \sum_{i=1}^{n} D_i u_i \right). \]

For cooperative game problem for constrained input MRM system with \( n \) joints, we need to find a set of solutions \( \left( u_1^*, \cdots, u_n^*, \cdots, u_n^* \right) \) that satisfy the following Pareto equilibrium conditions:

\[ J^* \left( s \right) = J \left( s, u_1^*, \cdots, u_i^*, \cdots, u_n^* \right) \]

\[ \leq \min \{ J \left( s, u_1^*, \cdots, u_i, \cdots, u_n^* \right), \cdots, \]

\[ J \left( s, u_1^*, \cdots, u_i, \cdots, u_w, \cdots, u_n^* \right), \cdots, \]

\[ J \left( s, u_1, \cdots, u_n \right) \}, \]

where \( i, w = 1, 2, \cdots, n, \) but \( i \neq w \). According to the optimal control theory, \( \left( u_1^*, \cdots, u_i^*, \cdots, u_n^* \right) \) satisfy the following stationary conditions:

\[ \frac{\partial H \left( s, u_1, \cdots, u_n, \nabla J^* \left( s \right) \right)}{\partial u_i} = 0. \]

(17)

Based on (17), the optimal control law with constrained input is as follows:

\[ u_i^* = -\beta_i \tanh \left( \frac{1}{2\beta_i} R_i^{-1} D_i^T \nabla J^* \left( s \right) \right). \]

(18)

The minimization of the positive definite quadratic of the joint position and velocity fusion error function ensures the trajectory tracking performance of the manipulator system, and the optimal of the control torque ensures the minimum energy consumption of the system. At this point, changing the control strategy of any participant will not lead to increase overall cost of the manipulator system, which is called Pareto equilibrium. By substituting the constrained input optimal control law (18) into the Hamiltonian function (15), the time-triggered HJB equation is obtained as follows:

\[ 0 = s^T Q s + h_m^T h_m + \sum_{i=1}^{n} U_i \left( u_i^* \right) \]

\[ + \left( \nabla J^* \right)^T \left( \Xi + \Psi + h + \vartheta + \left( \sum_{i=1}^{n} D_i u_i^* \right) \right) \]

\[ = s^T Q s + h_m^T h_m + \left( \nabla J^* \right)^T \left( \Xi + \Psi + h + \vartheta \right) \]

\[ + \sum_{i=1}^{n} \beta_i^2 R_i \ln \left( 1 - \tanh^2 \left( N_i^* \left( s \right) \right) \right) \]

where \( N_i^* \left( s \right) = \frac{1}{2\beta_i} R_i^{-1} D_i^T \nabla J^* \left( s \right). \)

3.2 NN observer design

In order to improve the robustness and control precision of the system, we can apply known model information to devise controller. Redefine \( N_i^* \left( s \right) = -N_{i1} \left( s \right) + N_{i2}^* \left( s \right) \), then the constrained input control law is as follows:

\[ u_i^* \left( s \right) = -\beta_i \tanh \left( -N_{i1} \left( s \right) + N_{i2}^* \left( s \right) \right), \]

(20)

in which \( N_{i1} \left( s \right) = N_{i1\Xi} \left( s \right) + N_{i1\Psi} \left( s \right). \) \( N_{i1\Xi} \left( s \right) \) is devised for the part of known model information, and \( N_{i1\Psi} \left( s \right) \) is applicable to compensate the model uncertainty and \( N_{i2}^* \left( s \right) \) is employed as the optimal compensation for the internal disturbance term \( d_i \left( q_i \right). \)

By substituting (20) into (19), the modified time-triggered HJB equation can be obtained:

\[ s^T Q s - \left( \nabla J^* \right)^T \left( \sum_{i=1}^{n} D_i \beta_i \tanh \left( N_{i2}^* - N_{i1\Xi} - N_{i1\Psi} \right) \right) \]

\[ + h_m^T h_m + \sum_{i=1}^{n} U_i \left( u_i^* \right) \]

\[ + \left( \nabla J^* \right)^T \left( \Xi + \Psi + h + \vartheta \right) = 0 \]
On the basis of (8), \( N_{i1Z} \) can be devised:

\[
N_{i1Z} = \frac{1}{\beta_i} \left( \hat{f}_{bi} \dot{q}_i + \left( \hat{f}_{si} e^{-\hat{f}_{i1}^2} + \hat{f}_{ci} \right) \text{sgn}(\dot{q}_i) + \frac{\tau_{si}}{\gamma_i} - B_{i1}^{-1} \dot{\theta}_i \right) \tag{22}
\]

Next, we utilize NN observer to estimate model uncertainties \( N_{i1\psi} \) such as IDC term and unmeasured friction term. \( N_{i1\psi} \) is reformulated in the following form:

\[
\Psi_i(x) = W_{\psi_i} \delta_{\psi_i}(x) + \varepsilon_{\psi_i}(x), \tag{23}
\]

where \( W_{\psi_i} \in \mathbb{R}^{l_{\psi_i}} \) refers to the ideal weight vector, \( \delta_{\psi_i}(x) \in \mathbb{R}^{l_{\psi_i}} \) means the radial basis function and \( l_{\psi_i} \) is the number of hidden layer neurons. \( \varepsilon_{\psi_i}(x) \) denotes the approximate residual of the NN which is bounded and satisfied \( \| \varepsilon_{\psi_i}(x) \| \leq \varepsilon_{\psi_{im}} \), where \( \varepsilon_{\psi_{im}} \) indicates a positive constant.

Since the NN ideal weight is unknown in (23), \( \hat{\Psi}_i(x) \) can be obtained by replacing \( W_{\psi_i} \) with the estimated value \( \hat{W}_{\psi_i} \):

\[
\hat{\Psi}_i(x) = \hat{W}_{\psi_i} \delta_{\psi_i}(x). \tag{24}
\]

On the grounds of the state space equation of the subsystem with constrained input (8) and the estimated value \( \hat{\Psi}_i \) (24), the NN observer is devised as:

\[
\begin{align*}
\dot{x}_i &= \hat{x}_i + k_{i1} \left( x_i - \hat{x}_i \right) \\
\dot{\hat{x}}_i &= \Xi_i(x_i) + h_i(x_i) + B_i u_i + \hat{\Psi}_i(x) \tag{25}
\end{align*}
\]

where \( \hat{x}_i = [\hat{x}_{i1}, \hat{x}_{i2}]^T \) means the observed value of the system state \( x_i \), \( k_{i1} = \text{diag}(k_{i1o}, k_{i2o}) \) is the diagonal matrix, in which each element is a positive constant. The NN observer error is defined as \( e_{i0} = [e_{i1o}, e_{i2o}]^T = [x_{i1} - \hat{x}_{i1}, x_{i2} - \hat{x}_{i2}]^T \), which can be obtained as follows:

\[
\begin{align*}
\dot{e}_{i1o} &= e_{i2o} - k_{i1o} e_{i1o} \\
\dot{e}_{i2o} &= \hat{W}_{\psi_i} \delta_{\psi_i}(x) + \varepsilon_{\psi_i}(x) - k_{i2o} e_{i2o}, \tag{26}
\end{align*}
\]

where \( \hat{W}_{\psi_i} = W_{\psi_i} - \hat{W}_{\psi_i} \) signifies the weight approximation error and it updates through:

\[
\dot{\hat{W}}_{\psi_i} = \alpha_{\psi_i} e_{i2o} \delta_{\psi_i}(x). \tag{27}
\]

The compensation for the uncertainty part \( N_{i1\psi} \) of the model is as follows:

\[
N_{i1\psi} = \frac{1}{\beta_i} B_{i1}^{-1} \dot{\hat{\Psi}}_i. \tag{28}
\]

Therefore, we can get the expression for \( N_{i1}(s) \):

\[
N_{i1}(s) = N_{i12} + N_{i1\psi} = \frac{1}{\beta_i} \left( \hat{f}_{bi} \dot{q}_i - B_{i1}^{-1} \dot{\theta}_i + \frac{\tau_{si}}{\gamma_i} - B_{i1}^{-1} \dot{\theta}_i \right) + B_{i1}^{-1} \dot{\hat{\Psi}}_i. \tag{29}
\]

By combining (20) and (29), the time-triggered optimal tracking control law for an n-DOF constrained input MRM system is obtained as follows:

\[
u_n^s(s) = -\beta_i \tanh \left( \frac{1}{\beta_i} \left( \hat{f}_{bi} \dot{q}_i + \frac{\tau_{si}}{\gamma_i} - B_{i1}^{-1} \dot{\theta}_i \right) \right) + \left( \hat{f}_{si} e^{-\hat{f}_{i1}^2} + \hat{f}_{ci} \right) \text{sgn}(\dot{q}_i) + B_{i1}^{-1} \dot{\hat{\Psi}}_i + N_{i12}^s(s). \tag{30}
\]

**Theorem 1** Considering the \( i \)th constrained input subsystem of MRM (1) with model uncertainty (7), the overall state space equation is given by (10), and (26) represents the estimation error of the devised NN observer (25). The global observation error of the NN observer is UUB under the given weight update rate (27).

**Proof** Take the following Lyapunov function:

\[
L = \sum_{i=1}^{n} \left( \frac{1}{2} e_{i1o}^T e_{i1o} + \frac{1}{2} \hat{W}_{\psi_i}^T \alpha_{\psi_i}^\dagger \hat{W}_{\psi_i} \right). \tag{31}
\]

The derivative of (31) is as follows:

\[
\dot{L} = \sum_{i=1}^{n} \left( e_{i1o}^T e_{i1o} + \hat{W}_{\psi_i}^T \alpha_{\psi_i}^\dagger \dot{\hat{W}}_{\psi_i} \right) = \sum_{i=1}^{n} \left( e_{i1o}^T e_{i1o} - k_{i1o} e_{i1o}^2 + e_{i2o} \hat{W}_{\psi_i}^T \delta_{\psi_i}(x) + e_{i1o} e_{i2o} \hat{W}_{\psi_i}^T \delta_{\psi_i}(x) - k_{i2o} e_{i2o}^2 - \hat{W}_{\psi_i}^T \alpha_{\psi_i}^\dagger \hat{W}_{\psi_i} \right) \leq \sum_{i=1}^{n} \left( e_{i1o}^2 - k_{i1o} e_{i1o}^2 - k_{i2o} e_{i2o}^2 + e_{i2o} \varepsilon_{\psi_{im}} \right) \leq \sum_{i=1}^{n} \left( \frac{1}{2} - k_{i1o} \right) e_{i1o}^2 + \left( 1 - k_{i2o} \right) e_{i2o}^2 + \frac{1}{2} \varepsilon_{\psi_{im}}^2. \tag{32}
\]

Therefore, the global observation error \( e_{i0} \) of the NN observer is UUB with \( k_{i1o} > \frac{1}{2} \) and \( k_{i2o} > 1 \), if \( e_{i2o} \) is outside the compact set \( \Omega_{\psi_i} = \left\{ e_{i2o} : \| e_{i2o} \| \leq \frac{\varepsilon_{\psi_{im}}}{2(k_{i2o} - 1)} \right\} \). \( \Box \)
3.3 Implementation of the critic NN

The optimal control law (18) is solved in the form of periodic sampling and the HJB Eq. (19) is derived under the time-triggered. Fixed sampling frequency control method will increase the computing burden of MRM system and waste communication resources, which is not conducive to real-time control under the limited network bandwidth. Therefore, the event-triggered mechanism is introduced into the design of controller. The control input is updated only if the system violates the triggering condition and the control strategy is replaced by the value of the previous sampling instant when the triggering condition is met, which can effectively reduce the calculation burden and save communication resources. Suppose \( \{t_j\}_{j=0}^{+\infty} \) is a monotonically increasing sequence composed of triggering instant, where \( t_j \) satisfies \( 0 < t_j < t_{j+1} \) and \( \lim_{j \to \infty} t_j = \infty, \ j \in \{0, 1, \ldots \} \). In event-triggered control, the system state is sampled at non-periodic sampling times \( \{t_0, \ldots, t_j, t_{j+1}, \ldots \} \). To obtain the triggering condition, the difference between the sampled state \( s(x_j) \) and the actual state \( s(x) \) is defined as a gap function:

\[
e_j(t) = s(x_j) - s(x), \quad t \in [t_j, t_{j+1}), \quad j = 0, 1, \ldots, \infty,
\]

where \( s(x_j) = \begin{cases} s(x_j) = s_j, & t \in [t_j, t_{j+1}) \\ s(x), & t = t_{j+1} \end{cases} \). At the triggering instants, \( e_j(t) = 0 \). Based on (20), the optimal control law under event-triggered is obtained as follows:

\[
u^*_i(s_j) = -\beta_i \tanh^{-1}(-N_{i1}(s_j) + N_{i2}^+(s_j)),
\]

(34)

The event-triggered Hamiltonian function is as follows:

\[
H(s_j, u_{i1}, \ldots u_{in}, \nabla J^T(s_j))
\]

\[
= s_j^T Q s_j + \sum_{i=1}^{n} U_i(u_{ij})
\]

\[
+ b_m^T h_m + \nabla J^T(s_j)(\Xi + \Psi + h + \vartheta + \sum_{i=1}^{n} D_{ij}u_{ij})
\]

Substitute (34) into (35) to obtain the event-triggered HJB equation:

\[
0 = s_j^T Q s_j + b_m^T h_m + \nabla J^T(s_j)(\Xi + \Psi + h + \vartheta) + \sum_{i=1}^{n} \beta_i^2 R_i \ln \left(1 - \tanh^2(-N_{i1}(s_j) + N_{i2}^+(s_j))\right).
\]

(36)

**Assumption 1** For any state \((s, s_j)\), \(N_{i1}^+\) exists in (19), \(N_{i1}\) exists in (29) and \(u^*\) exists in (20) are all Lipschitz continuous and there exist positive constants that satisfy the following inequalities:

\[
\begin{align*}
&\|N_{i1}^+(s_j) - N_{i1}^+(s)\| \leq l_{Nij} \|e_j\|, \\
&\|N_{i1}(s_j) - N_{i1}(s)\| \leq l_{Ni1j} \|e_j\|, \\
&\|u^*(s_j) - u^*(s)\| \leq l_{uij} \|e_j\|.
\end{align*}
\]

(37)

**Remark 7** According to (56), it can be concluded that two adjacent internal sampling time intervals \(T_j\) satisfy \(T_j = t_{j+1} - t_j \geq \frac{1}{L(1+P)}, \ j = 0, 1, \ldots, \infty\), where \(L\) is Lipschitz constant. If \(L\) is unbounded, it causes an infinite increase in the number of times the event-triggered. It can be seen from the experimental results that the number of events triggered is limited, and \(L\) must be a bounded positive constant. Therefore, it is reasonable and feasible to assume that the controller is Lipschitz continuous for \(e_j(t)\) in Assumption 1.

Since the HJB Eq. (19) is a partial differential equation, whose analytical solution is difficult to be obtained, the properties of universal approximation of NN are used to approximate the performance index function \(J(s)\) to solve the HJB equation. Under the event-triggered mechanism, we utilize NN to approximate the performance index function \(J(s_j)\) under sampling state, and the optimal control strategy under event-triggered is solved approximately. The \(J(s_j)\) can be reconstructed by NN in the following form:

\[
J(s_j) = W_c^T \delta_{c}(s_j) + \varepsilon_c(s_j),
\]

(38)

where \(W_c \in \mathbb{R}^l\) refers to the ideal weight vector, \(l\) is the number of hidden layer neurons, \(\delta_{c}(s_j)\) means the activation function and \(\varepsilon_c\) denotes the approximate residual. The gradient of (38) is as follows:

\[
\nabla J(s_j) = (\nabla \delta_{c}(s_j))^T W_c + \nabla \varepsilon_c(s_j).
\]

(39)

Based on the ADP framework of the 3.1 “Problem transformation” section, the optimal compensation policy under event-triggered \(N_{i2}^+(s_j)\) can be gained:

\[
N_{i2}^+(s_j) = \frac{1}{2\beta_i} R_i^{-1} D_{i1}^T \left( (\nabla \delta_{c}(s_j))^T W_c + \nabla \varepsilon_c(s_j) \right).
\]

(40)

Combining (34) and (39), the optimal control law of event-triggered based on critic NN is as follows:

\[
u^*_i(s_j) = -\beta_i \tanh^{-1}(-N_{i1}(s_j))
\]

\[+ \frac{1}{2\beta_i} R_i^{-1} D_{i1}^T \left( (\nabla \delta_{c}(s_j))^T W_c + \nabla \varepsilon_c(s_j) \right).
\]

(41)
According to (35) and (40), the Hamiltonian function is transformed into:

\[
H(s_j, u_{1j}^*, \ldots, u_{nj}^*) = s_j^T Q s_j + h_m^T h_m + \sum_{i=1}^{n} U_i (u_{ij}^*)
\]

\[+(\nabla \delta_c (s_j))^T W_c \left( \Xi + \Psi + h + \theta + \sum_{i=1}^{n} D_{ij} u_{ij}^* \right) \]

\[= -\nabla \delta_c (s_j) s_j \triangleq \delta_c H, \quad (42)\]

where \(\delta_c H\) is the approximate residual of the NN which satisfies \(|\delta_c H| \leq c_{HM}\) and \(c_{HM}\) is a positive constant.

Since the critic NN ideal weight \(W_c\) is unknown, the approximate weight \(\hat{W}_c\) is employed to replace \(W_c\), and \(\hat{J}(s_j)\) can be obtained as follows:

\[
\hat{J}(s_j) = \hat{W}_c^T \delta_c (s_j),
\]

\[where \(\hat{W}_c \in \mathbb{R}^l\) means the estimated weight vector.

Similarly, the gradient vector of (43) can be obtained:

\[
\nabla \hat{J}(s_j) = (\nabla \delta_c (s_j))^T \hat{W}_c.
\]

(44)

Similar to (40), the approximate optimal compensation policy under event-triggered \(N_{12}(s_j)\) is as follows:

\[
\hat{N}_{12}(s_j) = \frac{1}{2} \beta_l R_i^{-1} D_{ij}^T (\nabla \delta_c (s_j))^T \hat{W}_c.
\]

(45)

The approximate optimal control law of event-triggered based on critic NN can be obtained:

\[
\hat{u}_i(s_j) = -\beta_l \tanh^{-1} (-N_{11}(s_j)) + \frac{1}{2} \beta_l R_i^{-1} D_{ij}^T (\nabla \delta_c (s_j))^T \hat{W}_c.
\]

(46)

Combining (35) and (45), the approximate Hamiltonian function is as follows:

\[
\hat{H}(s_j, \hat{u}_{1j}, \ldots, \hat{u}_{nj})
\]

\[= (\nabla \delta_c (s_j))^T \hat{W}_c \left( \Xi + \Psi + h + \theta + \sum_{i=1}^{n} D_{ij} \hat{u}_{ij} \right)
\]

\[+ s_j^T Q s_j + h_m^T h_m + \sum_{i=1}^{n} U_i (\hat{u}_{ij}) \triangleq \delta_c H.
\]

(47)

For purpose of obtaining the approximate weight vector \(\hat{W}_c\), the objective function \(E_c = \frac{1}{2} \delta_c H^2\) is selected and the weight is gained by normalized steepest descent method [57, 58]. \(\hat{W}_c\) updates by:

\[
\dot{\hat{W}}_c = -\alpha_c \frac{1}{(1 + \phi^T \phi)^2} \frac{\partial E_c}{\partial \hat{W}_c}
\]

\[= -\alpha_c \frac{\delta_c H}{(1 + \phi^T \phi)^2} \hat{W}_c
\]

\[= -\alpha_c \frac{\phi}{(1 + \phi^T \phi)^2} (\nabla \delta_c (s_j))^T \hat{W}_c.
\]

(48)

where \(\phi = (\nabla \delta_c (s_j))^T \hat{s}_j\), \(\tilde{W}_c = W_c - \hat{W}_c\), \(\hat{W}_c\) refers to the weight approximation error vector.

**Assumption 2** (The persistence of excitation (PE) condition [59]) The signal \(\phi = \frac{\phi}{1 + \phi^T \phi}\) is persistently excited, that is, for \(\forall \tau, t \in [0, \infty),\) there exist positive constants \(\chi_1, \chi_2\) and \(T\), such that the following in equation is satisfied:

\[
\chi_1 I_{n \times n} \leq \int_{t}^{t+T} \phi \phi^T \, dz \leq \chi_2 I_{n \times n}.
\]

**Theorem 2** Considering the dynamic model (10) of the constrained input MRM system, if the weight updating rate given by (48) is satisfied, the weight approximation error vector \(\hat{W}_c\) of the critic NN is UUB.

**Proof** Take the Lyapunov function as:

\[
V_c = \frac{1}{2} \hat{W}_c^T \hat{W}_c.
\]

(49)

Take the derivative of (49) with respect to time,

\[
\dot{V}_c = \hat{W}_c^T \dot{\hat{W}}_c
\]

\[= \alpha_c W_c \frac{\phi}{(1 + \phi^T \phi)^2} (\delta_c H - \phi^T \hat{W}_c)
\]

\[= \alpha_c W_c \phi \delta_c H (1 + \phi^T \phi)^2 - \alpha_c W_c \phi \phi^T \hat{W}_c
\]

\[\leq - \left( \alpha_c - \frac{1}{4} \right) \lambda^2_{\min} \left( \phi \phi^T \right) \hat{W}_c^2
\]

\[+ \frac{\alpha_c^2 \epsilon_{HM}}{(1 + \phi^T \phi)^2},
\]

when \(\alpha_c > \frac{1}{4}\) and \(\hat{W}_c\) is outside the compact set

\[\Omega_c = \left\{ \hat{W}_c : \| \hat{W}_c \| \leq \sqrt{\frac{\alpha_c^2 \epsilon_{HM}}{(\alpha_c - \frac{1}{4}) \lambda^2_{\min} (\phi \phi^T)}} \right\}, \]

the weight approximation error of the critic NN is UUB.
Remark 8 Assumption 2 implies $\chi_1 \leq \lambda_{\min} \left( -T \phi \phi \right) \leq \chi_2$, that is, $\lambda_{\min} \left( -T \phi \phi \right)$ is bounded, which effectively ensures that the weight parameters of the neural network can converge to a bounded range in the learning adjustment stage.

Remark 9 In the critic NN training process, since the optimal control law $u_i^* (s_j)$ usually cannot be obtained, we use the approximate optimal control law $\hat{u}_i (s_j)$ instead [58], and it is also applied to the constrained input MRM system as the control torque given by (46). The structure diagram of event-triggered-based cooperative game optimal tracking control method proposed in this article is given in Fig. 1.

3.4 Stability proof

For the MRM system with constrained input, we devise an approximate optimal control law under event-triggered (46) to accomplish the trajectory tracking task. How to ensure the stable operation of the system is discussed here, and we give the following assumption and theorem to make the closed-loop system stable.

Assumption 3 Let $\nabla \delta_{cm}$, $\nabla \varepsilon_{cm}$ and $\tilde{W}_{cm}$ all be positive constants [58] and satisfied:

1. $\| \nabla \delta_c \| \leq \nabla \delta_{cm}$,
2. $\| \nabla \varepsilon_c \| \leq \nabla \varepsilon_{cm}$,
3. $\| \tilde{W}_c \| \leq \tilde{W}_{cm}$.

Theorem 3 Considering an $n$-DOF MRM system with constrained input whose dynamic model described as (1), the model uncertainty give in (4) and (7), the internal disturbance term existed in (8), and the overall state space equation derive by (10). The trajectory tracking error of the closed-loop system is UUB under the proposed event-triggered optimal control law (46), if the triggering condition

$$\| e_j \| \leq \left( \frac{\lambda_{\min} (Q) - \xi^2}{\lambda_{\max} (Q) + 2\sigma} \right) \| s_j \|^2 + \sum_{i=1}^n U_i (\hat{u}_{ij})$$

$$+ \frac{\lambda_{\max} (Q) + 2\sigma}{\lambda_{\max} (Q) + 2\sigma} \left( \sum_{i=1}^n (\beta_i + 1) \left( \sum_{i=1}^n U_i (\hat{u}_{ij}) \right) \right)$$

$$= \left( e_j (T) \right)$$

is satisfied, where $\sigma = \sum_{i=1}^n \beta_i (\lambda_{\max} (Q) + 2\sigma)$.

Proof The Lyapunov function is chosen as:

$$V (t) = V_1 (t) + V_2 (t) = J^* (s) + J^* (s_j),$$

where $J^* (s)$ refers to optimal performance index function that $J^* (s) > 0$ for any $s \neq 0$. Therefore, $J^* (s)$ is a positive definite function. In addition, we can see that $J^* (s_j) = 0$. By event-triggered mechanism, the stability proof of closed-loop system can be divided into the following two cases.

(a) The event is not triggered, that is $t \in [t_j, t_{j+1})$, take the partial derivative of Lyapunov function (52) is as follows:

$$\dot{V} (t) = \nabla J^* (s) \left( \Xi + \Psi + h + \vartheta + \sum_{i=1}^n D_{ij} \hat{u}_{ij} \right).$$

According to (21) and (30), we can get:

$$(\nabla J^* (s))^T (\Xi + \Psi + h + \vartheta) = -s^T Q s - h_m^T h_m$$

$$- \sum_{i=1}^n \beta_i^2 R_i \ln \left( 1 - \tanh^2 \left( -N_i (s) + N_i^2 (s) \right) \right).$$

(54)

$$\left( \nabla J^* (s) \right)^T D_{ii} = 2\beta_i N_i^2 (s) (s) R_i.$$ (55)

Combining (20), (54) and (55), Lyapunov function (52) can be transformed into:

$$\dot{V} (t) = -s^T Q s - h_m^T h_m - \left( \nabla J^* (s) \right)^T \sum_{i=1}^n D_{ij} u_i (s_j)$$

$$- \sum_{i=1}^n \beta_i^2 R_i \ln \left( 1 - \tanh^2 \left( -N_i (s) + N_i^2 (s) \right) \right)$$

$$= -s^T Q s - h_m^T h_m - \sum_{i=1}^n \int_0^T \ln \left( 1 - \tanh^2 \left( -N_i (s) + N_i^2 (s) \right) \right) d w$$

$$- \left( \nabla J^* (s) \right)^T \sum_{i=1}^n D_{ij} u_i (s_j) + \sum_{i=1}^n \int_0^T \frac{\hat{u}_{ij}}{w} R_i d w$$

$$- \sum_{i=1}^n \beta_i \left( \nabla J^* (s) \right)^T D_{ii} \tanh \left( -N_i (s) + N_i^2 (s) \right).$$

Springer
Fig. 1 Schematic of the proposed control method

\[ -s^T Q s - h_m^T h_m - \sum_{i=1}^{n} \int_{u_i}^{u_i^+} 2\beta_i \tanh^{-1} \left( \frac{w}{\beta_i} \right) R_i d w \]

\[ = -s^T Q s - h_m^T h_m - \sum_{i=1}^{n} U_i (\hat{u}_{ij}) + \sum_{i=1}^{n} \int_{u_i}^{\hat{u}_{ij}} 2\beta_i (N_i^x (s) + \tanh^{-1} \left( \frac{w}{\beta_i} \right)) R_i d w \]

\[ + \sum_{i=1}^{n} \int_{u_i}^{\hat{u}_{ij}} 2\beta_i (N_i^x (s)) R_i d w. \quad (56) \]

\[ -s^T Q s \text{ in (56) can be simplified as follows:} \]

\[ -s^T Q s = -(s_j - e_j)^T Q (s_j - e_j) \]

\[ = -s_j^T Q s_j - e_j^T Q e_j + 2s_j^T Q e_j \]

\[ \leq -s_j^T Q s_j - e_j^T Q e_j + \frac{1}{2} s_j^T Q s_j + 2e_j^T Q e_j \]

\[ = -\frac{1}{2} s_j^T Q s_j + e_j^T Q e_j. \quad (57) \]

Let \( w = \beta_i \tanh (z), \) and then \( \sum_{i=1}^{n} \int_{u_i}^{\hat{u}_{ij}} 2\beta_i (N_i^x (s) + \tanh^{-1} \left( \frac{w}{\beta_i} \right)) R_i d w \) in (56) can be written as:

\[ \sum_{i=1}^{n} \int_{u_i}^{\hat{u}_{ij}} 2\beta_i (N_i^x (s) + \tanh^{-1} \left( \frac{w}{\beta_i} \right)) R_i d w \]

\[ \leq \sum_{i=1}^{n} \beta_i \left( (N_i^x (s))^T R_i N_i^x (s) + (\hat{u}_{ij} - u_i^+)^T R_i (\hat{u}_{ij} - u_i^+) \right) \]
The derivative of the Lyapunov function $\dot{V}(t) < 0$ when (60) satisfies the trigger condition given by (51).

(b) The event is triggered, that is $\forall t = t_{j+1}$. The difference form of (52) is as follows:

$$\Delta V(t) = V(s_{j}(t_{j+1}^{+})) - V(s_{j}(t_{j+1}^{-}))$$

$$= V_{1}(s_{j}(t_{j+1}^{+})) - V_{1}(s_{j}(t_{j+1}^{-})) + V_{2}(s_{j}(t_{j+1}^{+})) - V_{2}(s_{j}(t_{j}))$$

$$= J^{*}(s_{j}(t_{j+1}^{+})) - J^{*}(s_{j}(t_{j+1}^{-})) + J^{*}(s_{j}(t_{j+1}^{+})) - J^{*}(s_{j}(t_{j})),$$

where $s_{j}(t_{j+1}^{+}) = \lim_{x \to 0} (s_{j}(t_{j+1}^{+} + x))$, $s_{j}(t_{j+1}^{-}) = \lim_{x \to 0} (s_{j}(t_{j+1}^{-} - x))$. We know from the proof of case a that $V_{1}(t) \leq 0$, therefore,

$$V_{1}(s_{j}(t_{j+1}^{+})) - V_{1}(s_{j}(t_{j+1}^{-})) \leq 0.$$  (62)

The difference Eq. (61) can be reduced to:

$$V(t) \leq V_{2}(s_{j}(t_{j+1}^{+})) - V_{2}(s_{j}(t_{j})) = J^{*}(s_{j}(t_{j+1}^{+})) - J^{*}(s_{j}(t_{j}))$$

$$\leq -\kappa(\|e_{j}(t_{j+1})\|) \leq 0,$$

where $\kappa(\cdot)$ is a class-$\kappa$ function [58] and $e_{j}(t_{j+1}) = s_{j}(t_{j+1}) - s_{j}(t_{j}).$

Combining case (a) with case (b), we can conclude that the MRM system with constrained input is stable in real time. □

3.5 Zeno behavior analysis

For an event-triggered MRM system with constrained input, the minimum triggering interval $\Delta t_{\min} = \min_{j \in \mathbb{N}} \{ (t_{j+1} - t_{j}) \}$ may be zero, which results in the system being driven an infinite number of times, leading to an infinite number of controller updates, these behaviors are called Zeno Behaviors [60]. This phenomenon is not in line with the original intention of applying event-triggered to save resources, and it is difficult to achieve in the actual system, hence it should be avoided. Next, we will prove that the minimum triggering time interval $\Delta t_{\min}$ under the event-triggered control has a nonzero lower bound, which can effectively avoid Zeno behavior.
Assumption 4 Assuming the state space Eq. (10) of the MRM system with constrained input is Lipschitz continuous, then $\|\dot{x}\|$ has an upper bound and therefore $\|\dot{x}\| \leq \|\dot{x}_2 + v\| \leq \|\ddot{x}_2\| + \|v\| \leq \|\ddot{x}_2\| + v_m$ also has an upper bound, where $v_m$ denotes a positive constant. The upper bound of $\|\dot{x}\|$ can be defined as:

$$\|\dot{x}\| \leq k \|x\| + k\varphi,$$  \hspace{1cm} (64)

where $\varphi, k$ are positive constants, and $\varphi$ is tiny.

Theorem 4 For an $n$-DOF MRM system with constrained input (1) whose state space described as (10), the minimum triggering interval $\Delta_{t_{\text{min}}}$ has a positive lower bound

$$\Delta_{t_{\text{min}}} \geq \frac{1}{k} \ln \left( 1 + \frac{e_j (t_j + 1)}{\|s_j\| + \varphi} \right) > 0,$$ \hspace{1cm} (65)

if the triggering condition given by (51) is satisfied, where $t_{j+1} = \lim_{\delta \to 0} (t_{j+1} - \delta)$.

Proof The derivative of gap function (33) is as follows:

$$\frac{de_j (t)}{dt} = \dot{e}_j (t) = -\dot{s}_j, \forall t \in [t_j, t_{j+1}) , j \in \mathbb{N}.$$ \hspace{1cm} (66)

Combining (64) and (66), we have

$$\|\dot{e}_j\| \leq k \|s\| + k\varphi = k \|s_j - e_j (t)\| + k\varphi \leq k \|e_j (t)\| + k (\|s_j\| + \varphi), \forall t \in [t_j, t_{j+1}).$$ \hspace{1cm} (67)

Depending on the property of the event-triggered mechanism, at the triggering instant $e_j (t_j) = 0$, the solution of (67) is as follows:

$$\|e_j\| \leq \int_{t_j}^{t} e^{(k(t-\tau))} k (\|s_j\| + \varphi)d\tau$$
$$= (\|s_j\| + \varphi) \left( e^{k(t-t_j))} - 1 \right).$$ \hspace{1cm} (68)

Through (68), we can know that the triggering condition (51) of jth sampling is satisfied

$$\Delta_{t_{\text{min}}} = \min \{ t_{j+1} - t_j \}$$
$$\geq \frac{1}{k} \ln \left( 1 + \frac{e_j (t_{j+1})}{\|s_j\| + \varphi} \right) > 0, \forall t \in \mathbb{N}.$$ \hspace{1cm} (69)

4 Experiments

In this part, the feasibility of the proposed method is verified by using a 2-DOF MRM experimental platform.

4.1 Experimental setup

Figure 2 shows a 2-DOF MRM experimental platform to verify the proposed control algorithm. Each joint module includes: a DC motor, a reducer, a torque sensor, an absolute encoder and an incremental encoder. Two metal rods connect two identical joint modules in series. The DC motor product number is 388987. It has a rated voltage of 48 V and a rated torque of 0.74 Nm. The reducer is embedded in the DC motor, and its gear ratio is 100:1. The two DC motors are driven by a linear power amplifier (LPA). The 500-line incremental encoder and the 19-bit absolute encoder measure position information of the connecting rod side and the motor side, respectively. QPIDe data acquisition card is used for real-time collection of experimental data and then through the MATLAB/SIMULINK software installed on the host computer for data processing. Based on the QUARC module communication between the host and QPIDe equipment, the controller built on the host’s MATLAB/SI-MULINK software can realize the real-time control of the 2-DOF MRM.

In this paper, experiments with tracking tasks are established for 2-DOF constrained input MRM to verify the effectiveness of the proposed algorithm. The desired trajectory of the two joints is as follows:

$$q_{1d} = -\frac{\pi}{12} \sin \left( \frac{\pi}{40}t \right) - \frac{\pi}{40} \sin \left( \frac{\pi}{40} \frac{\pi}{30}t \right) - \frac{3}{200},$$

$$q_{2d} = -\frac{\pi}{10} \sin \left( \frac{\pi}{40} \frac{\pi}{30}t \right) - \frac{\pi}{60} \sin \left( \frac{\pi}{60} \frac{\pi}{60}t \right) + \frac{1}{20}.$$}

For the NN observer and the critic NN, we choose the radial basis function NN to approximate the model uncertainty and performance index function, respectively. The NN observer chooses 2-3-1 structure, that is 2 input neurons, 3 hidden layer neurons and 1 output neuron. The initial weight vector of the NN observer is defined as $\hat{W}_{\psi/0} = \left[ \hat{\hat{x}}_{\psi/1}, \hat{\hat{x}}_{\psi/2}, \hat{\hat{x}}_{\psi/3} \right]^{T}$, and its initial value is $\hat{W}_{\psi/0} = \hat{W}_{\psi/0} = [0.1, 0.2, 0.3]^{T}$. The activation function of the NN observer is selected as $\delta (s_m) = \exp (-\|s_m - c_m\|^{2}/2b_{m}^{2})$, where $c_1 = \ldots$
Fig. 2 Experimental platform. a Experimental platform setup, b Joint module structure

$[-1, 0, 1], c_2 = [-2, 0, 2], m = 1, 2, b_l = 1.5, l = 1, 2, 3$. The critic NN chooses 4-3-1 structure whose initial weight vector is defined as $\hat{W}_{c0} = [\hat{W}_{c1}, \hat{W}_{c2}, \hat{W}_{c3}]^T$, and the initial value of weight vector is selected as $\hat{W}_{c0} = [1, 3, 5]^T$. The activation function is $\delta_c(s_m) = \exp\left(-\|s_m - c_m\|^2/2b_l^2\right)$, where $c_1 = c_3 = [-1, 0, 1], c_2 = c_4 = [-2, 0, 2], m = 1, 2, 3, 4, b_l = 1.5, l = 1, 2, 3$. Other parameters are given in Table 1.

4.2 Experimental results and analysis

Figures 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 show the experimental comparison between the ADP-based event-triggered optimal control method [26, 50] and the proposed event-triggered-based cooperative game optimal tracking control method for MRM system with constrained input. In the event-triggered control, we take the real-time state of MRM system with constrained input measured by the sensor under time-triggered as the system state under event-triggered.

(1) Trajectory tracking curves analysis

Figures 3 and 4, respectively, demonstrate the trajectory tracking curves of two joints with the existing ADP-based event-triggered control method and the proposed event-triggered-based cooperative game optimal tracking control method. From the figures, one observes that the actual trajectory can keep up with the desired trajectory commendably under both control methods. In addition, from the local magnification view, we know that the existing control method can track desired trajectory at about 0.4 s, and the proposed event-triggered-based cooperative game optimal tracking control method ensures that the two joints can quickly catch up at 0.2 s. Therefore, the proposed control method has preferable dynamic tracking performance than the existing control method.

Figures 5 and 6, respectively, exhibit the trajectory tracking error curves of two joints with the existing control method and the proposed control method. From these experimental results, one concludes that the manipulator system can quickly learn and track the desired trajectory at the beginning under the two control methods and ensure the trajectory tracking error within an ideal range. After the desired trajectory is quickly tracked by the actual trajectory, the trajectory tracking error of the proposed control method is within $\pm 4 \times 10^{-3}$ rad, which is less than the trajectory tracking error of the existing control method. The reason is the event-triggered-based cooperative game optimal tracking control method proposed in this paper not only compensates the model uncertainties such as friction term, IDC term and internal disturbance term, but also ensures the optimal overall performance by considering the cooperation among the various joints. Therefore, the steady-state tracking performance of the cooperative game optimal tracking control method based on event triggering mechanism is more excellent than the existing control method.

(2) Control torque curves analysis
Table 1  Parameters setting

| Parameter type | Name | Value          | Name | Value          | Name | Value          |
|----------------|------|----------------|------|----------------|------|----------------|
| Model          | \( \hat{f}_{bi} \) | 30 mN ms/rad    | \( \hat{f}_{ci} \) | 60 mN m       | \( \hat{f}_{si} \) | 80 mN m       |
|                | \( \gamma_i \)  | 100            | \( l_{mi} \)    | 1350 g cm²    | \( \hat{f}_{ci} \) | 50 s²/rad²   |
| Uncertainty    | \( \eta_{Fi1} \) | 12 mN ms/rad    | \( \eta_{Fi2} \) | 0.02 Nm        | \( \eta_{Fi3} \) | 0.03 Nm        |
|                | \( \eta_{Fi4} \) | 20s²/rad²       | \( \eta_{\theta_1} \) | 0.15 Nm      | \( \eta_{\theta_2} \) | 0.15 Nm      |
|                | \( e_{\psi_{im}} \) | 0.02          | \( \nabla \delta_{cm} \) | 5           | \( \nabla \theta_m \) | 0.001         |
|                | \( \bar{W}_{cm} \) | 0.01           | \( \eta_{\theta_{im}} \) | 0.1       | \( \nu_m \) | 0.5           |
| Control        | \( Q \)    | \( I_{2 \times 2} \) | \( R_{1, R_2} \) | \( I_{1 \times 1} \) | \( \xi \) | 0.7           |
| parameters     | \( l_{N/1} \) | 1.35           | \( l_{aij} \)   | 1.35         | \( l_{N/1} \) | 1.35         |
|                | \( \alpha \) | 0.6           | \( \alpha_{\psi_{i}} \) | 0.75      | \( \alpha_{s1} \) | 1.7           |
|                | \( \alpha_{2} \) | 1.2           | \( k_{io1} \)   | 1.5         | \( k_{io2} \) | 1.1           |

![Fig. 3](image1.png)  
**Fig. 3** Trajectory tracking curves of the existing control method.  
(a) Joint 1, (b) Joint 2

![Fig. 4](image2.png)  
**Fig. 4** Trajectory tracking curves of the proposed control method.  
(a) Joint 1, (b) Joint 2

Figures 7 and 8 display the control torque curves of the existing control method and the proposed control method, respectively. From the experimental results, one observes that the upper and lower bounds of the control torque for the constrained input MRM system are ±0.6 Nm, and the exceeding parts are kept at ±0.6 Nm. In addition, under the event-triggered control method, the control torque has an obvious holding stage. Within this triggering time interval, the system state and control torque both adopt the value of the last triggering point and will not be updated until the next triggering instant. This non-periodic sampling method reduces the updating times of the control strategy and the communication burden between the controller and the controlled plant. In addition, Fig. 7 illustrates that the control torque is with obvious chattering phenomenon during the operation of the system, which will not only affect the control performance of the system, but also reduces the durability of the DC motor inside the manipulator. A smooth control torque curve is shown in Fig. 8. The vibration amplitude and frequency of the control torque devised under the event-triggered-based cooperative game optimal tracking control method are distinctly smaller, which enhances the control precision of the manipulator system and prolongates the operational lifespan of the motor.

(3) Event-triggered curves analysis

© Springer
Figure 9 illustrates the overall trigger threshold and trigger condition curves. Due to the offset of the desired trajectory at the initial time, the value of the trigger condition and the trigger threshold are both large, but the trigger condition is still within the range of the trigger threshold, which indicates the event-triggered control method can ensure the normal operation of the system at the initial instant. In the local magnification, the trigger condition and the trigger threshold have an aperiodic intersection instant. The intersection of two points indicates that the system violates the trigger condition and the event is triggered. At this point, the system state is sampled and the control law is updated, which
remains unchanged until the next intersection of the two curves.

Figure 10 describes the comparison curves of controller update times under the time-triggered and the event-triggered. As we can see from the figure, the update times of the event-triggered controller are about half of the time-triggered controller, which not only decrease computational cost of the system, but also effectively cut down the communication burden between the controller and the controlled model. In addition, the time-triggered controller update times curve is a straight line with constant slope, on account of the time-triggered controller employed periodic sampling communication. The event-triggered control method adopts non-periodic communication mode, as a result the slope of the update times curve of the event-triggered controller changes constantly. Where the slope is larger, the system triggering more times to ensure the control performance and vice versa.

(4) Weight curves analysis

Figure 11 is the weight curves of the NN observer used to compensate model uncertainty. Note that the NN observer proposed in this paper is designed under the time-triggered because it requires real-time position and velocity information of each joint, but the NN observer based compensation control law is under the event-triggered. From the experimental result curves in the figure, one concludes that the weights of the NN observer converge to a certain range, which verifies Theorem 1 in an experimental way and fully proves the effectiveness of the NN observer approximation model uncertainty.

Figure 12 exhibits the weight curves of the critic NN with the proposed optimal control method. Similar to the weight curves of the NN observer, the critic NN weight curves also converge to a bounded range and Theorem 2 can be verified experimentally. With the event-triggered mechanism, the weight of the critic NN will be updated only when the trigger condition is violated, and the other times remain unchanged. The weight curves and control torque curves have the same update frequency. Not only that, the proposed cooperative game optimal control scheme based on event-triggered mechanism adopts single critic NN to decrease the complexity of the system structure, and the estimation value of critic NN weight makes it feasible to solve the optimal control strategy proposed in this paper.
Experimental results illustrate that the event-triggered-based cooperative game optimal tracking control method proposed in this paper is superior to the traditional control method considering the comprehensive influence factors such as control precision, energy consumption and communication burden of MRM system with constrained input. Both theoretical derivation and experimental results can be verified repeatedly.

5 Conclusion

In this paper, we solve the problem of optimal tracking control in cooperative game of MRM system with constrained input based on event-triggered mechanism. The dynamic model of constrained input subsystem is established by using JTF technology, and the entirety state space description is obtained by defining the global state variables. The cooperative game theory is introduced into the derivation of controller, and the tracking control problem of MRM system with constrained input is transformed into the optimal control problem based on cooperative game. Next, the global performance index function is constructed, and the optimal tracking controller based on cooperative game is proposed by using partial known model information and ADP framework. Then, the event-triggered mechanism is applied to the above controller, and the system state is sampled aperiodic according to the designed trigger conditions. The approximate optimal control law is gained by solving the event-triggered HJB equation through the critic NN. Lyapunov theory is used to prove the tracking error is UUB of closed-loop constrained input MRM system. From the experimental results, one concludes that the event-triggered-based cooperative game optimal tracking control method proposed in this paper can not only satisfy the control accuracy, but also reduce the update times of the controller and alleviate the calculation burden, which fully proves the validity of the proposed control method.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Grant No. 61773075, 61703055 and 62173047), the Scientific Technological Development Plan Project in Jilin Province of China (Grant Nos. 20200801056GH) and the Science and Technology project of Jilin Provincial Education Department of China during the 13th Five-Year Plan Period (JJKH202 00672KJ, JJKH20200673KJ and JJKH20200674KJ).

Funding The authors have not disclosed any funding.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declared no potential conflicts of interest with respect to the research, authorship and publication of this article.

References

1. Mu, C., Wang, K., Ni, Z., Sun, C.: Cooperative differential game-based optimal control and its application to power systems. IEEE Trans. Industr. Inform. 16(8), 5169–5179 (2020)
2. Yun, A., Moon, D., Ha, J., Kang, S., Lee, W.: Modman: an advanced reconfigurable manipulator system with genderless connector and automatic kinematic modeling algorithm. IEEE Robot. Autom. Lett. 5(3), 4225–4232 (2020)
3. Fei, Y., Gao, H.: Nonlinear dynamic modeling on multi-spherical modular soft robots. Nonlinear Dyn. 78(2), 831–838 (2014)
4. Brantner, G., Khatib, O.: Controlling ocean one: human-robot collaboration for deep-sea manipulation. J. Field Robot. 38(1), 28–51 (2020)
5. Aggravi, M., Elsherif, A., Giordano, P.R., Pacchierotti, C.: Haptic-enabled decentralized control of a heterogeneous human-robot team for search and rescue in partially-known environments. IEEE Robot. Autom. Lett. 6(3), 4843–4850 (2021)
6. Ginting, M.F., Otsu, K., Edlund, J., Gao, J., Aghamohammadi, A.: Chord: distributed data-sharing via hybrid ROS 1 and 2 for multi-robot exploration of large-scale complex environments. IEEE Robot. Autom. Lett. 6(3), 5064–5071 (2021)
7. Zhao, S., Liang, H., Du, P., Qi, S.: Adaptive NN finite-time tracking control of output constrained nonlinear system with input saturation. Nonlinear Dyn. 92(4), 1845–1856 (2018)
8. Lars, Q., Sonneveldt, P., Chu, J.: Nonlinear flight control design using constrained adaptive backstepping. J. Guid. Control Dyn. 30(2), 322–336 (2007)
9. Jin, Y.Q., Liu, X.D., Wei, Q., Hou, C.Z.: Time-varying sliding mode control for a class of uncertain MIMO nonlinear system subject to control input constraint. Sci. China Inf. Sci. 53(1), 88–100 (2010)
10. Li, Y., Tong, S., Li, T.: Direct adaptive fuzzy backstepping control of uncertain nonlinear systems in the presence of input saturation. Neural Comput. Appl. 23(5), 1207–1216 (2013)
11. Kurtz, M.J., Henson, M.A.: Input-output linearizing control of constrained nonlinear processes. J. Process Contr. 7(1), 3–17 (1997)
12. Bellman, R.: Dynamic programming. Science 153(3731), 34–37 (1966)
13. Ciarlet, P.G., Raviart, P.A.: Maximum principle and uniform convergence for the finite element method. Comput. Methods Appl. Mech. Eng. 2(1), 17–31 (1976)
14. Sniedovich, M., Lew, A.: Dynamic programming: an overview. Control Cybern. 35(3), 513–533 (2006)
15. Werbos, P.J.: Backpropagation through time: what it does and how to do it. Proc. IEEE 78(10), 1550–1560 (1990)
16. Al-Tamimi, A., Lewis, F.L.: Discrete-time nonlinear HJB solution using approximate dynamic programming: convergence proof. IEEE Trans. Syst. Man Cybern. Syst. 38(4), 943–949 (2008)
17. Tao, B., Jiang, Z.P.: Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design. Automatica 71, 348–360 (2016)
18. Liu, D., Wei, Q.: Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. IEEE Trans. Neural Netw. Learn. Syst. 25(3), 621–634 (2014)
19. Wang, B., Zhao, D., Cheng, J.: Adaptive cruise control via adaptive dynamic programming with experience replay. Soft Comput. 23(12), 4131–4144 (2019)
20. Wei, Q., Li, H., Yang, X., He, H.: Continuous-time distributed policy iteration for multi controller nonlinear systems. IEEE Trans. Cybern. 51(5), 2372–2383 (2021)
21. Shi, Z., Wang, Z.: Adaptive output-feedback optimal control for continuous-time linear systems based on adaptive dynamic programming approach. Neurocomputing 438, 334–344 (2021)
22. Zhao, B., Liu, D., Luo, C.: Reinforcement learning-based optimal stabilization for unknown nonlinear systems subject to inputs with uncertain constraints. IEEE Trans. Neural Netw. Learn. Syst. 31(10), 4330–4340 (2020)
23. Shao, X., Yi, S.: Neural network based constrained output feedback control for mems gyroscopes considering scarce transmission bandwidth. IEEE Trans. Cybern. (2022). https://doi.org/10.1109/TCYB.2021.3070137
24. Wang, D., Liu, D., Li, H., Luo, B., Ma, H.: An approximate optimal control approach for robust stabilization of a class of discrete-time nonlinear systems with uncertainties. IEEE Trans. Syst. Man Cybern. Syst. 46(5), 713–717 (2016)
25. Zhao, B., Luo, F., Lin, H., Liu, D.: Particle swarm optimized neural networks based local tracking control scheme of unknown nonlinear interconnected systems. Neural Netw. 134, 54–63 (2021)
26. Dong, B., An, T., Zhou, F., Liu, K., Li, Y.: Decentralized robust zero-sum neuro-optimal control for modular robot manipulators in contact with uncertain environments: theory and experimental verification. Nonlinear Dyn. 97(13), 503–524 (2019)
27. Gao, X., Bai, W.W., Li, T.S., Yuan, L.E., Long, Y.: Broad learning system-based adaptive optimal control design for dynamic positioning of marine vessels. Nonlinear Dyn. 105(2), 1593–1609 (2021)
28. An, L., Yang, G.H.: Optimal transmission power scheduling of networked control systems via fuzzy adaptive dynamic programming. IEEE Trans. Fuzzy Syst. 29(6), 1629–1639 (2020)
29. Yang, H., Li, Y., Liu, Z., Yuan, H.: Adaptive dynamic programming for security of networked control systems with actuator saturation. Inf. Sci. 460, 51–64 (2018)
30. Zhang, H., Zhang, J., Yang, G., Luo, Y.: Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming. IEEE Trans. Fuzzy Syst. 23(1), 152–163 (2015)
31. Ge, Y.Y., Liu, X.K., Li, Y.: Pareto optimal control of the mean-field stochastic systems by adaptive dynamic programming algorithm. ISA Trans. 102, 81–90 (2020)
32. Zhang, H., Wei, Q., Liu, D.: An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. Automatica 47, 207–214 (2011)
33. Zhang, H., Cui, L., Luo, Y.: Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network ADP. IEEE Trans. Syst. Man Cybern. Syst. 43(1), 206–216 (2012)
34. Miao, X.N., Gang, X.: Cooperative differential game model based on trade-off between energy and delay for wireless sensor networks. Ann. Oper. Res. 206(1), 297–310 (2013)
35. Yang, X., Gao, Z., Zhang, J.: Event-driven H∞ control with critic learning for nonlinear systems. Neural Netw. 132, 30–42 (2020)
36. Jiang, H., Zhang, H., Han, J., Zhang, K.: Iterative adaptive dynamic programming methods with neural network implementation for multi-player zero-sum games. Neurocomputing 307(13), 54–60 (2018)
37. Dong, B., An, T., Zhu, X., Li, Y., Liu, K.: Zero-sum game-based neuro-optimal control of modular robot manipulators with uncertain disturbance using critic only policy iteration. Neurocomputing 450(2), 183–196 (2021)
38. Jiang, H., Zhang, H., Xia, G., Cui, X.: Data-based approximate optimal control for nonzero-sum games of multi-player systems using adaptive dynamic programming. Neurocomputing 275(31), 192–199 (2018)
39. Narayan, V., Sahoo, A., Jagannathan, S., George, K.: Approximate optimal distributed control of nonlinear interconnected systems using event-triggered nonzero-sum games. IEEE Trans. Neural Netw. Learn. Syst. 30(5), 1512–1522 (2019)
40. Zhao, J.: Neural networks-based optimal tracking control for nonzero-sum games of multi-player continuous-time nonlinear systems via reinforcement learning. Neurocomputing 412(1), 167–176 (2020)
41. Ma, B., Li, Y., An, T., Dong, B.: Compensator-critic structure-based neuro-optimal control of modular robot manipulators with uncertain environmental contacts using non-zero-sum games. Knowl. Based Syst. 224, 107100–107114 (2021)
42. Zhang, Q., Zhao, D., Zhu, Y.: Data-driven adaptive dynamic programming for continuous-time fully cooperative games with partially constrained inputs. Neurocomputing 238, 377–386 (2017)
43. Li, X., Peng, Z., Jiao, L., Cai, J.: Online adaptive Q-learning method for fully cooperative linear quadratic games. Sci. China Inf. Sci. 62(12), 2201–2215 (2019)
44. Zhang, Q., Zhao, D., Wang, D.: Event-based robust control for uncertain nonlinear systems using adaptive dynamic programming. IEEE Trans. Neural Netw. Learn Syst. 29(1), 37–50 (2018)
45. Xue, S., Luo, B., Liu, D., Li, Y.: Adaptive dynamic programming-based event-triggered control for unknown continuous-time nonlinear systems with input constraints. Neurocomputing 396, 191–200 (2019)
46. Mu, C., Liao, K., Wang, K.: Event-triggered design for discrete-time nonlinear systems with control constraints. Nonlinear Dyn. 103(3), 2645–2657 (2021)
47. Zhu, Y., Zhao, D., He, H., Ji, J.: Event-triggered optimal control for partially unknown constrained-input systems via adaptive dynamic programming. IEEE Trans. Ind. Electron. 64(5), 4101–4109 (2017)
48. Shan, X., Luo, B., Liu, D.: Event-triggered adaptive dynamic programming for zero-sum game of partially unknown continuous-time nonlinear systems. IEEE Trans. Syst. Man Cybern. Syst. 50(9), 3189–3199 (2018)
49. Zhang, H., Su, H., Zhang, K., Luo, Y.: Event-triggered adaptive dynamic programming algorithm for non-zero-sum games of unknown nonlinear systems via generalized fuzzy hyperbolic models. IEEE Trans. Fuzzy Syst. 27(11), 2202–2214 (2019)
50. Ma, B., Li, Y.: Compensator-critic structure-based event-triggered decentralized tracking control of modular robot manipulators: theory and experimental verification. Complex Intell. Syst. (2021). https://doi.org/10.1007/s40747-021-00359-0
51. Liu, G., Abdul, S., Goldenberg, A.A.: Distributed control of modular and reconfigurable robot with torque sensing. Robotica 26(1), 75–84 (2008)
52. Liu, G.: Decomposition-based friction compensation of mechanical systems. Mechatronics 12(5), 755–769 (2002)
53. Modares, H., Lewis, F.L., Naghibi, M.: Integral reinforcement learning and experience replay for adaptive optimal control of partially unknown constrained input continuous time systems. Automatica 50(1), 193–202 (2014)
54. Lv, Y., Ren, X.: Approximate Nash solutions for multiplayer mixed-zero-sum game with reinforcement learning. IEEE Trans. Syst. Man Cybern. Syst. 49(12), 2739–3750 (2018)
55. Vamvoudakis, K.G.: Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems. Automatica 44(3), 282–293 (2014)
56. Dong, L., Yuan, X., Sun, C.: Event-triggered receding horizon control via actor-critic design. Sci. China Inf. Sci. 63(5), 1–15 (2020)
57. Zhong, X., He, H.: An event-triggered ADP control approach for continuous-time system with unknown internal states. IEEE Trans. Cybern. 47(3), 683–694 (2017)
58. Wang, D., Mu, C., Liu, D.: On mixed data and event driven design for adaptive-critic-based nonlinear $H_\infty$ control. IEEE Trans. Neural Netw. Learn. Syst. 29(4), 993–1005 (2017)
59. Vamvoudakis, K.G., Lewis, F.L.: Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. Automatica 46(5), 878–888 (2010)
60. Lamperski, A., Ames, D.: Lyapunov theory for Zero stabilility. IEEE Trans. Autom. Control 58(1), 100–112 (2013)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.