Prescribed-Time Formation Control for a Class of Multi-agent Systems via Fuzzy Reinforcement Learning
Yan Zhang, Mohammed Chadli, Zhengrong Xiang

To cite this version:
Yan Zhang, Mohammed Chadli, Zhengrong Xiang. Prescribed-Time Formation Control for a Class of Multi-agent Systems via Fuzzy Reinforcement Learning. IEEE Transactions on Fuzzy Systems, 2023, 31 (12), pp.4195–4204. 10.1109/TFUZZ.2023.3277480. hal-04117871

HAL Id: hal-04117871
https://hal.science/hal-04117871
Submitted on 13 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Prescribed-Time Formation Control for a Class of Multiagent Systems via Fuzzy Reinforcement Learning

Yan Zhang, Mohammed Chadli, Senior Member, IEEE, and Zhengrong Xiang, Member, IEEE

Abstract—This article concerns optimal prescribed-time formation control for a class of nonlinear multiagent systems (MASs). Optimal control depends on the solution of the Hamilton–Jacobi–Bellman equation, which is hard to be calculated directly due to its inherent nonlinearity. To overcome this difficulty, the reinforcement learning strategy with fuzzy logic systems is proposed, in which identifier, actor, and critic are used to estimate unknown nonlinear dynamics, implement control behavior, and evaluate system performance, respectively. Different from the existing optimal control algorithms, a new performance index function considering formation error cost and control input energy cost is constructed to achieve optimal formation control of MASs within a prescribed time. The presented control strategy can ensure that the formation error converges to the desired accuracy within a prescribed time. Finally, the validity of the presented strategy is verified via a simulation example.

Index Terms—Formation control, fuzzy logic systems, Hamilton–Jacobi–Bellman, nonlinear multiagent systems (MASs), prescribed-time control, reinforcement learning (RL).

I. INTRODUCTION

COOPERATIVE control of multiagent systems (MASs) has drawn widespread concern from scholars because of its extensive application. Cooperative control includes consensus control, containment control, formation control, etc. Formation control refers to constructing control protocols to stabilize agents’ positions with respect to each other and to achieve specified geometric shapes. The formation control of MASs has been applied in the fields of mobile robots [1], underactuated marine surface vessels [2], and unmanned aerial vehicles [3]. Zhang et al. [4] presented a formation-tracking control scheme for linear MASs. Then, an event-triggered formation containment control strategy was developed in [5] for linear MASs. Furthermore, the formation control issues for nonlinear MASs were further investigated in [6], [7], [8], and [9]. The results obtained in [4], [5], [6], [7], [8], and [9] only realized the expected formation; however, the optimal performance was not considered.

Optimal control is a research topic with important academic value and broad application potential [10], [11]. To make the controlled system complete the tasks with the least resources, Bellman in [12] and Pontryagin in [13] put forward the optimal control. Usually, it is expected that the optimal function can be got by solving the Hamilton–Jacobi–Bellman (HJB) equation, and then, the optimal controller can be constructed [14]. As a partial differential equation with multiple nonlinear terms, it is hard to find an analytical solution to the HJB equation. As a machine learning method, reinforcement learning (RL) was inspired by animal behavior and gradually became a powerful tool. Under the actor–critic architecture, the optimal consensus and formation control of MASs concerned in [15], [16], and [17], [18], [19], respectively. It should be pointed out that most RL algorithms [15], [16], [17], [18], [19] need prior knowledge of system dynamics while the nonlinear terms of most actual dynamic systems may be unknown. By applying RL under the actor–critic–identifier (ACI) architecture, Wen et al. [20], [21] developed the optimal control strategies for MASs. However, the optimal control results [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21] can only ensure the expected performance when the time variable goes to infinity.

Convergence rate, as a crucial performance criterion, is one of the hot topics in the control field [22], [23]. To improve the convergence rate, some finite-time formation control results of MASs were reported in [24], [25], [26], and [27]. The settling time in [24], [25], [26], and [27] is dependent on initial conditions, which limits the application of finite-time control. To address this issue, some fixed-time formation control issues were investigated in [28], [29], [30], and [31]. It should be mentioned that the settling time in [28], [29], [30], and [31] depends on the design parameters, which makes it difficult to adjust the design parameters according to actual control requirements. For this reason, some prescribed-time consensus control issues for MASs were investigated in [32], [33], [34], [35], [36], [37], and [38]. However, the schemes developed in [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], and [38] did not take into account optimal performances. To the author’s knowledge, for a class of uncertain nonlinear MASs, the...
prescribed-time formation control via RL is still an interesting and challenging topic, which is rarely reported.

Motivated by the above statements, a prescribed-time formation control strategy is presented for a class of MASs, which can guarantee that the formation error converges to the desired accuracy within a prescribed time. The main innovations are as follows.

1) Different from the existing prescribed-time control [32], [33], [34], [35], [36], [37], [38], the optimal control problem is considered. By constructing an auxiliary function and an error transformation function, the original performance index function is rewritten, which lays the foundation for realizing the prescribed-time formation control of MASs with the least resources.

2) The proposed strategy can eliminate the assumption of persistence excitation, which is required by most optimal control algorithms (such as [19] and [20]). In addition, new actor updating laws are designed to implement training sufficiently.

3) Compared with the existing optimal control results for MASs (such as [12], [13], [14], [15], [16], [17], [18], [19], [20], [21]), the settling time is not influenced by initial conditions and design parameters and can be prescribed.

The rest is organized as follows. Section II gives the problem description and some preliminaries. The prescribed-time control scheme is developed in Section III. Thereafter, the validity of the strategy is verified in Section IV via a simulation example. Finally Section V concludes the article.

II. PROBLEM DESCRIPTION AND SOME PRELIMINARIES

A. Graph Theory

The undirected connected graph \( G = (\mathcal{V}, \chi, A) \) is modeled to describe the communication network among agents, where \( \mathcal{V} = \{v_1, v_2, \ldots, v_n\} \), \( \chi \subseteq \mathcal{V} \times \mathcal{V} \), and \( A = [a_{ij}] \) represent vertex set, edge set, and adjacency matrix, respectively. If there is an information flow from \( v_j \) to \( v_i \), then \( v_j \in \chi \) and \( a_{ji} = 1 \), and \( v_j \) is a neighbor of \( v_i \); otherwise, \( v_j \notin \chi \) and \( a_{ij} = 0 \). The neighbor node set is denoted by \( N_i = \{v_j \mid (v_i, v_j) \in \chi\} \). If and only if \( a_{ij} = a_{ji} \), \( G \) is an undirected graph. If any pair of distinct vertices has a path, the undirected graph is connected, where a path from \( v_i \) to \( v_j \) is an edge sequence in the form of \( (v_i, v_{i1}, v_{i2}, \ldots, v_{it}, v_j) \).

\[ L = D - A \in \mathbb{R}^{n \times n} \quad \text{and} \quad B = \text{diag}\{b_1, b_2, \ldots, b_n\} \]

denote the Laplacian matrix and the communication matrix between agent and leader, respectively, where \( D = \text{diag}\{\sum_{j=1}^{n} a_{1j}, \ldots, \sum_{j=1}^{n} a_{nj}\} \). Assume that at least one agent is connected to the leader, which means \( \sum_{i=1}^{n} b_i > 0 \).

B. Fuzzy Logic Systems

IF–THEN rules:

\[ \text{Rule} \; l: \; \text{IF} \; x_1 \; \text{is} \; A_{11}^l, \; x_2 \; \text{is} \; A_{22}^l, \ldots, \; x_k \; \text{is} \; A_{kk}^l \; \text{THEN:} \; y_B = B^l. \]

where \( x = [x_1, \ldots, x_k]^T \in \mathbb{R}^k \) and \( y_B \) denote the input and output of \( l \) represents the number of fuzzy rules. \( \mu_{A_i^l}(x) \) and \( \mu_{B_i^l}(y_B) \) denote the membership functions of fuzzy sets \( A_i^l \) and \( B^l \), respectively.

Combining the singleton fuzzifier, product inference, and center average defuzzifier, \( y_B \) can be expressed as

\[ y_B(x) = \frac{\sum_{i=1}^{L} \phi_i^{k} \prod_{j=1}^{k} \mu_{A_j^l}(x_i) \prod_{j=1}^{k} \mu_{A_j^l}(x_i)}{\sum_{i=1}^{L} \prod_{j=1}^{k} \mu_{A_j^l}(x_i)} \]

where \( L \) denotes the total number of fuzzy rules, \( \phi_i^k \) satisfies \( \mu_{B_i^l}(\phi_i^k) = \max\{\mu_{B_i^l}(y_B)\} \forall y_B \in \mathbb{R} \). Let

\[ \phi_i^k = \frac{\prod_{j=1}^{k} \mu_{A_j^l}(x_i) \prod_{j=1}^{k} \mu_{A_j^l}(x_i)}{\sum_{i=1}^{L} \prod_{j=1}^{k} \mu_{A_j^l}(x_i)} \]

and \( \phi = [\phi_1^k, \ldots, \phi_L^k]^T \) and \( \Phi = [\Phi_1^k, \ldots, \Phi_L^k]^T \).

Then, \( y_B(x) \) is represented as

\[ y_B(x) = \Phi^T \phi(x). \]

Lemma 1 (see [17]): For \( \forall \varepsilon > 0 \), there exists FLS \( \Phi^T \phi(x) \) such that \( \sup_{x \in \Xi} |f(x) - \Phi^T \phi(x)| < \varepsilon \), where \( f(x) \) denotes a continuous function defined on a compact set \( \Xi \).

C. Problem Description

Consider the following multiagent system (MAS):

\[ \dot{x}_i = u_i + f_i(x_i), \; i = 1, \ldots, n \]

where \( x_i = [x_{i1}, \ldots, x_{im}]^T \in \mathbb{R}^m \), \( u_i = [u_{i1}, \ldots, u_{im}]^T \in \mathbb{R}^m \) and \( f_i(\cdot) \in \mathbb{R}^m \) represent the system state, control input, and unknown continuous nonlinear function, respectively.

Assumption 1: The desired formation trajectory \( x_d \in \mathbb{R}^m \) and its derivative \( \dot{x}_d \in \mathbb{R}^m \) are bounded.

Define the coordinate transformations as

\[ z_i = x_i - x_d - \epsilon_i \]

where \( z_i = [z_{i1}, \ldots, z_{im}]^T \in \mathbb{R}^m \) denotes the tracking error; \( \epsilon_i = [\epsilon_{i1}, \ldots, \epsilon_{im}]^T \in \mathbb{R}^m \) denotes the relative position between the leader and the agent \( i \), which describes the formation mode.

Remark 1: \( x_d \) and \( \dot{x}_d \) represent the desired formation position and speed, respectively. In actual formation control, the given formation position and speed must be bounded. Otherwise, the formation would not be feasible. Assumption 1 is usually used in formation control, such as [20] and [21].

The formation error is defined as

\[ e_i = \sum_{j \in N_i} a_{ij}(x_i - \epsilon_i - x_j + \epsilon_j) + b_i(x_i - x_d - \epsilon_i) \]

\[ = \sum_{j \in N_i} a_{ij}(z_i - z_j) + b_i z_i \]

where \( \epsilon_i = [\epsilon_{i1}, \ldots, \epsilon_{im}]^T \in \mathbb{R}^m \).

Definition 1: The practical prescribed-time formation control of MAS (1) can be realized if there exists protocol \( u_i \) such that \( \|e_i\| < \|\varepsilon_i\| \) for \( \forall t > T \), where \( T > 0 \) is a settling time and \( \varepsilon_i = [\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{im}]^T \in \mathbb{R}^m \) denotes the desired accuracy.
Define the performance index function as follows:

$$J_i(e_i) = \int_t^\infty e^{-\alpha_i(t-T)}(e_i^T Q_i e_i + u_i^T R_i u_i)d\tau$$  \hspace{1cm} (4)$$

where \(\alpha_i > 0\) represents the discount factor, and \(Q_i\) and \(R_i\) denote the symmetric positive-definite matrices.

**Definition 2** (see [40]): For MAS (1), if 1) \(u_i\) is continuous; 2) \(u_i(0) = 0\); 3) \(u_i\) stabilizes (1) on the set \(\Omega\); 4) \(J_i(e_i)\) is finite, then the control protocol \(u_i\) is admissible on \(\Omega\), which is represented as \(u_i \in \mathcal{W}(\Omega)\).

**Remark 2:** To obtain the optimal formation control protocol \(u^*\) so that MAS (1) can achieve the desired formation, the performance index function includes the formation error cost and the control input cost is considered in (4).

The tasks of this article are to present a prescribed-time control algorithm for MAS (1) such that 1) the formation error (3) converges to the desired accuracy within a prescribed time; 2) the performance index function (4) can be minimized.

### D. Necessary Lemmas

**Lemma 2** (see [41]): The Laplacian matrix of the undirected connected graph is irreducible.

**Lemma 3** (see [21]): If the Laplacian matrix is irreducible, then \(\mathcal{L} + \mathcal{B}\) is a positive-definite matrix.

**Lemma 4** (see [39]): For \(i, j \in R, \) one has

$$i_j \leq \frac{m^p}{p} \|i\|^p + \frac{1}{qm^q} \|j\|^q$$  \hspace{1cm} (5)$$

where \(m > 0\), and \(p, q > 1\) with \((p-1)(q-1) = 1\).

### III. PRESCRIBED-TIME FORMATION CONTROL

#### A. Error Transformation Function

To realize the prescribed-time control performance of MAS (1), an auxiliary function is given as follows:

$$\zeta = \begin{cases} \frac{(1 - \tau) (T - \tau)}{T}^2 + \tau, & 0 < \tau < T \\ \tau, & \tau \geq T \end{cases}$$  \hspace{1cm} (6)$$

where \(0 < \tau < 1\) represents the design parameter, and \(T > 0\) denotes the prescribed time. The following properties are obtained from (6): 1) \(\zeta\) is strictly decreasing over \((0, T)\), with \(\zeta(0) = 1\), and \(\zeta = \tau\) for \(\forall t \geq T\); 2) \(\zeta\) is smooth, and \(\dot{\zeta}\) and \(\ddot{\zeta}\) are bounded for all \(t \geq 0\).

Define the error transformation function as \(\xi_i = [\xi_{i1}, \ldots, \xi_{im}]^T \in R^{m}\). \(\xi_i\) is constructed as follows:

$$\xi_i = \tan\left(\frac{\pi \xi_i \ell}{\psi \ell}\right), \ell = 1, \ldots, m$$  \hspace{1cm} (7)$$

where \(\psi > 0\) is a constant, and \(\psi(0) = 0\).

**Remark 3:** For \(\forall t \geq 0\), if \(\xi_{i\ell}\) is bounded, it can be observed that \(\psi \xi_{i\ell} < e_{i\ell} < \psi \xi_{i\ell}\) from (7).

### B. Optimal Formation Control

From (7), we can obtain that

$$e_{i\ell} = \frac{2\psi \xi_{i\ell}}{\pi} \arctan(\xi_{i\ell}), \ell = 1, \ldots, n, \ell = 1, \ldots, m.$$  \hspace{1cm} (8)$$

Then, \(e_i\) can be further expressed as \(e_i = \left[\frac{2\psi \xi}{\pi} \arctan(\xi_1), \ldots, \frac{2\psi \xi}{\pi} \arctan(\xi_m)\right]^T\).

Let \(\xi_i = [\psi_{i1}, \ldots, \psi_{im}]^T\). The performance index function (4) can be represented as

$$J_i = \int_t^\infty e^{-\alpha_i(t-T)}\left(\left(\frac{2\psi \xi}{\pi} \psi_{i1}\right)^T Q_i \left(\frac{2\psi \xi}{\pi} \psi_{i1}\right) + u_i^T R_i u_i\right)d\tau.$$  \hspace{1cm} (9)$$

**Remark 4:** Through the transformation function (7), the error \(e_{i\ell}\) is rewritten as (8) and embedded into the performance index function (4). This operation can ensure that the performance index function includes formation error cost and control input energy cost while achieving the prescribed-time formation control performance of MAS (1).

Taking the optimal formation control \(u^* = [u_{i1}^*, \ldots, u_{im}^*]^T \in R^{m}\) into (9), we can get the optimal performance index function as follows:

$$J_i^* = \min_{u_i \in \mathcal{W}(\Omega)} \left\{ \int_t^\infty e^{-\alpha_i(t-T)}\left(\left(\frac{2\psi \xi}{\pi} \psi_{i1}\right)^T Q_i \left(\frac{2\psi \xi}{\pi} \psi_{i1}\right) + u_i^T R_i u_i\right)d\tau \right\}$$

$$= \int_t^\infty e^{-\alpha_i(t-T)}\left(\left(\frac{2\psi \xi}{\pi} \psi_{i1}\right)^T Q_i \left(\frac{2\psi \xi}{\pi} \psi_{i1}\right) + u_i^T R_i u_i\right)d\tau.$$  \hspace{1cm} (10)$$

From (2) and (3), taking the time derivative of \(e_i\) yields

$$\dot{e}_i = c_i f_i + c_i u_i - b_i \dot{x}_i - \sum_{j \in N_i} a_{ij} \dot{x}_j$$  \hspace{1cm} (11)$$

where \(c_i = \sum_{j \in N_i} a_{ij} + b_i\).

From (11) and (12), we have

$$\dot{\xi}_{i\ell} = \frac{\dot{\xi}_{i\ell}}{1 + \xi_{i\ell}^2} = \frac{1}{1 + \xi_{i\ell}^2} K_i(\dot{e}_{i\ell} - \gamma_{i\ell})$$  \hspace{1cm} (12)$$

where \(K_i = \frac{\pi}{2\psi \xi_{i\ell}} \cos^2\left(\frac{\pi \xi_{i\ell} \ell}{2\psi \xi_{i\ell}}\right)\) and \(\gamma_{i\ell} = \frac{\psi_{i\ell} \dot{\xi}_{i\ell}}{\xi_{i\ell}}\).

From (13), we have

$$\dot{v}_i = K_i(\dot{e}_i - \gamma_i) = K_i\left(c_i f_i + c_i u_i - b_i \dot{x}_i - \sum_{j \in N_i} a_{ij} \dot{x}_j - \gamma_i\right)$$  \hspace{1cm} (13)$$

where \(K_i = \text{diag}\{K_{i1}, \ldots, K_{im}\}\), and \(\gamma_i = [\gamma_{i1}, \ldots, \gamma_{im}]^T\).
By calculating the time derivative of (10), we can get HJB equation as follows:

\[ H_i \left( v_i, ζ_i, u_i^*; \frac{∂J_i^*}{∂v_i}, \frac{∂J_i^*}{∂ζ_i} \right) \]

\[ = \left( \frac{2νC}{π} v_i \right)^T Q_i \left( \frac{2νC}{π} v_i \right) + u_i^T R_i u_i^* - α_i J_i^* + \frac{∂J_i^*}{∂v_i} \] 
\[ \times \left[ \begin{array}{c} \tilde{K}_i \left( c_i f_i + c_i u_i^* - b_i \dot{x}_d - \sum_{j \in N_i} a_{ij} \dot{x}_j - γ_i \right) \end{array} \right] \]
\[ + \frac{∂J_i^*}{∂ζ_i} \frac{dζ_i}{dt} \]
\[ = \left( \frac{2νC}{π} v_i \right)^T Q_i \left( \frac{2νC}{π} v_i \right) + u_i^T R_i u_i^* - α_i J_i^* + \frac{∂J_i^*}{∂v_i} \left[ \begin{array}{c} \tilde{K}_i \left( c_i f_i + c_i u_i^* - b_i \dot{x}_d - \sum_{j \in N_i} a_{ij} \dot{x}_j - γ_i \right) \end{array} \right] \]
\[ + \frac{∂J_i^*}{∂ζ_i} \frac{dζ_i}{dt} \]
\[ = 0 \tag{14} \]

where \( K_i = \text{diag} \{ K_{i1}, \ldots, K_{im} \} \). The right-hand side of (10) is existent and unique, it can be known that \( u_i^* \) is the unique solution of (14). By solving \( \frac{∂H_i}{∂v_i} \left( v_i, ζ_i, u_i^*; \frac{∂J_i^*}{∂v_i}, \frac{∂J_i^*}{∂ζ_i} \right) = 0 \), one can get the optimal control input as follows:

\[ u_i^* = -c_i \frac{2}{R_i^2} K_i^T \frac{∂J_i^*}{∂ζ_i}. \tag{15} \]

Substituting (15) into (14) yields

\[ \left( \frac{2νC}{π} v_i \right)^T Q_i \left( \frac{2νC}{π} v_i \right) - α_i J_i^* + \frac{∂J_i^*}{∂ζ_i} \frac{dζ_i}{dt} \]
\[ + \frac{∂J_i^*}{∂ζ_i} \left[ \begin{array}{c} \tilde{K}_i \left( c_i f_i + b_i \dot{x}_d - a_{ij} \dot{x}_j - γ_i \right) \end{array} \right] \]
\[ - c_i^T \frac{∂J_i^*}{∂ζ_i} K_i R_i^{-1} K_i^T \frac{∂J_i^*}{∂ζ_i} = 0. \tag{16} \]

\textbf{Remark 5:} It is expected to get the term \( \frac{∂J_i^*}{∂ζ_i} \) by solving (16). Due to that the HJB equation (16) contains completely unknown system dynamics, it is difficult or even impossible to calculate its solution.

RL methods, as a powerful tool, will be used to address the optimal formation issue.

\textbf{C. ACI Design}

To achieve the optimal formation control, \( \frac{∂J_i^*}{∂ζ_i} \) is segmented as

\[ \frac{∂J_i^*}{∂ζ_i} = \frac{2h_i}{c_i} K_i^{-2} ξ_i + \frac{2}{c_i} K_i^{-2} \tilde{f}_i(X_i) + \frac{1}{c_i} K_i^{-2} J_i^0(\mathcal{X}_i) \tag{17} \]

where \( h_i > \frac{3}{4} \) is a design parameter, \( J_i^0(\mathcal{X}_i) = -2h_i ξ_i - 2 \tilde{f}_i + c_i^2 K_i^2 d_i \frac{∂J_i}{∂ζ_i} \), and \( \tilde{f}_i(X_i) = R_i K_i (f_i(x_i) - ̄x_d - μ^{-1} γ_i) \) with \( μ = \lambda_{\text{max}}(\mathcal{L}), \mathcal{L} = \dot{\mathcal{L}} + B, \mathcal{X}_i = [x_i, ξ_i, x_d, ̄x_d, ζ_i, ̄ζ_i]^T, X_i = [x_i, x_d, ̄x_d, ζ_i, ̄ζ_i]^T \).

By substituting (17) into (15), we have

\[ u_i^* = -h_i R_i^{-1} K_i^{-1} ξ_i - R_i^{-1} K_i^{-1} \tilde{f}_i(X_i) \]
\[ - \frac{1}{2} h_i K_i^{-1} ζ_i^T J_i^0(\mathcal{X}_i). \tag{18} \]

The unknown terms \( \tilde{f}_i(X_i) \) and \( J_i^0(\mathcal{X}_i) \) are continuous, by Lemma 1, for \( \forall e_{f_i} \in R^m \) and \( \forall \epsilon_i \in R^m \), there exist FLSs \( \Psi_i^T \varphi f_i(X_i) \) and \( \Psi_i^T \varphi \epsilon_i(\mathcal{X}_i) \) such that

\[ \tilde{f}_i(X_i) = \Psi_i^T \varphi f_i(X_i) + \epsilon f_i(X_i) \tag{19} \]
\[ J_i^0(\mathcal{X}_i) = \Psi_i^T \varphi \epsilon_i(\mathcal{X}_i) + \epsilon \epsilon_i(\mathcal{X}_i) \tag{20} \]

where \( \Psi_i^T \in R_i^{p_i \times m} \) and \( \Psi_i^T \in R_i^{p_i \times m} \) denote the optimal parameter matrices; \( p_i \) and \( p_i \) denote the fuzzy rule numbers; \( \varphi f_i \in R_i^p \) and \( \varphi \epsilon_i \in R_i^p \) denote the fuzzy basis function vectors; the approximation errors \( e_{f_i} \in R^m \) and \( \epsilon_i \in R^m \) satisfy \( \|e_{f_i}\| \leq \delta_{f_i} \) and \( \|\epsilon_i\| \leq \delta_i \); \( \delta_{f_i} \) and \( \delta_i \) are constants.

Substituting (19) and (20) into (17) and (18), we get

\[ \frac{∂J_i^*}{∂ζ_i} = \frac{2h_i}{c_i} K_i^{-2} ξ_i + \frac{2}{c_i} K_i^{-2} \Psi_i^T \varphi f_i(X_i) \]
\[ + \frac{2}{c_i} K_i^{-2} e_{f_i}(X_i) + \frac{1}{c_i} K_i^{-2} \Psi_i^T \varphi \epsilon_i(\mathcal{X}_i) \]
\[ + \frac{1}{c_i} K_i^{-2} \epsilon_i(\mathcal{X}_i) \tag{21} \]

\[ u_i^* = -h_i R_i^{-1} K_i^{-1} ξ_i - R_i^{-1} K_i^{-1} \Psi_i^T \varphi f_i(X_i) \]
\[ - R_i^{-1} K_i^{-1} \epsilon_{f_i}(X_i) - \frac{1}{2} R_i^{-1} K_i^{-1} \Psi_i^T \varphi \epsilon_i(\mathcal{X}_i) \]
\[ - \frac{1}{2} R_i^{-1} K_i^{-1} \epsilon_i(\mathcal{X}_i). \tag{22} \]

Because \( \Psi_i^T \) and \( \Psi_i^T \) are unknown, (21) and (22) are not available. To this end, we design the following identifier, critic, and actor.

First, to identify the unknown dynamic function, we design the identifier as follows:

\[ \hat{f}_i(X_i) = \hat{Ψ}_i \varphi f_i(X_i) \tag{23} \]

where \( \hat{f}_i(X_i) \) and \( \hat{Ψ}_i \in R_i^{p_i \times m} \) denote the output of FLS and identifier parameter matrices, respectively. In addition, we design the identifier updating law as follows:

\[ \dot{\hat{Ψ}}_f = P_i (\mu \varphi f_i(X_i) \xi_i^T R_i^{-1} - \omega_i \hat{Ψ}_f) \tag{24} \]

where \( \omega_i > 0 \) denotes a design parameter, and \( P_i \in R_i^{p_i \times p_i} \) denotes a positive-definite matrix. To evaluate the control performance, we design the following critic:

\[ \frac{∂J_i^*}{∂ζ_i} = \frac{2h_i}{c_i} K_i^{-2} ξ_i + \frac{2}{c_i} K_i^{-2} \hat{Ψ}_i \varphi f_i(X_i) \]
\[ + \frac{1}{c_i} K_i^{-2} \hat{Ψ}_i \varphi \epsilon_i(\mathcal{X}_i) \tag{25} \]
where $\partial J_i$ and $\tilde{\Psi}_{ci}$ $\in R^{p_2 \times m}$ denotes the critic parameter matrix. Then, we design the following updating law for (25):
\[
\dot{\tilde{\Psi}}_{ci} = -\eta_{ci}(\varphi_i(\mathcal{X}_i)\varphi_i^T(\mathcal{X}_i) + jI_{p_2})\tilde{\Psi}_{ci}
\] (26)
where $\eta_{ci} > 0$ and $j > 0$ represent design parameters. To implement the control behavior, the actor is designed as follows:
\[
u_i = -h_i R_i^{-1} K_i^{-1} \xi_i - R_i^{-1} K_i^{-1} \tilde{\Psi}_{fi} \varphi_i(X_i)
- \frac{1}{2} R_i^{-1} K_i^{-1} \tilde{\Psi}_{ai} \varphi_i(X_i)
\] (27)
where $h_i > \frac{1}{2}$, $\nu_{ci} \in R^{p_2 \times m}$ represents the actor parameter matrix. We construct the following actor updating law:
\[
\dot{\nu}_{ci} = - (\varphi_i(\mathcal{X}_i)\varphi_i^T(\mathcal{X}_i) + jI_{p_2})(\eta_{ai}(\tilde{\Psi}_{ai} - \tilde{\Psi}_{ci}) + \eta_{ci}\tilde{\Psi}_{ci})
\] (28)
where $\eta_{ai} > 0$ represents the actor design parameter.

**Remark 6:** The actor updating laws in [21] was constructed as $\dot{\tilde{\Psi}}_{ai} = -\varphi_i(\mathcal{X}_i)\varphi_i^T(\mathcal{X}_i)(\eta_{ai}(\tilde{\Psi}_{ai} - \tilde{\Psi}_{ci}) + \eta_{ci}\tilde{\Psi}_{ci})$. If the term $\tilde{\Psi}_{ai} - \tilde{\Psi}_{ci}$ falls on the value of zero eigenvectors of $\varphi_i(\mathcal{X}_i)\varphi_i^T(\mathcal{X}_i)$, the training may be terminated. To this end, $jI_{p_2}$ is introduced in (28) for sufficient training.

**Remark 7:** The actor $u_i$ is expected to satisfy $H_i(v_i, \zeta, u_i, \epsilon_i, \gamma_i, \alpha_i) \rightarrow 0$. Inserting (25) and (27) into (14), we can get the approximated HJB equation as follows:
\[
H_i(v_i, \zeta, u_i, \epsilon_i, \gamma_i, \alpha_i, \varphi_i, \tilde{\Psi}_{ai}, \tilde{\Psi}_{ci})
\] (29)
where $\eta_{ai} > 0$ and $\eta_{ci} > 0$. The inequality (29) means that $\frac{\partial H_i(v_i, \zeta, u_i, \epsilon_i, \gamma_i, \alpha_i)}{\partial \tilde{\Psi}_{ai}} = 0$ holds by the designed RL updating laws (26) and (28). Furthermore, we get that $H_i(v_i, \zeta, u_i, \epsilon_i, \gamma_i, \alpha_i) \rightarrow 0$.

**D. Main Results**

**Theorem 1:** Consider MAS (1) under Assumption 1, if the identifier in (23), the critic in (25), and the actor in (27) with the updating laws (24), (26), and (28) are employed, then the formation error converges to the desired accuracy within a prescribed time, where the design parameters $h_i$, $\eta_{ai}$, and $\eta_{ci}$ satisfy $h_i > \frac{1}{2}$ and $\eta_{ai} > \eta_{ci} > \frac{2}{h_i^2} > \frac{\mu R_i^2}{4}$.

**Proof:** Construct the Lyapunov function as follows:
\[
V = \frac{1}{2} T ^\xi T + \frac{1}{2} \sum_{i=1}^{n} (\tilde{\Psi}_{fi} P_i^{-1} \tilde{\Psi}_{fi})
+ \frac{1}{2} \sum_{i=1}^{n} (\tilde{\Psi}_{ai} P_i^{-1} \tilde{\Psi}_{ai})
\] (30)
where $\xi = [\xi_1, \cdots, \xi_n]^T \in R^{m \times n}$, $\tilde{\Psi}_{fi} = \tilde{\Psi}_{fi} - \Psi_{fi}^*$, $\tilde{\Psi}_{ai} = \tilde{\Psi}_{ai} - \Psi_{ai}^*$, $\tilde{\Psi}_{ci} = \tilde{\Psi}_{ci} - \Psi_{ci}^*$.

From (3), (7), (24), (26), and (28), we have
\[
\dot{\tilde{\Psi}} = \xi^T[K(\tilde{\mathcal{L}} \otimes I_m)\tilde{z} - K\gamma]
+ \sum_{i=1}^{n} T \tilde{\Psi}_{fi}(\mu R_i^{-1} \varphi_i(X_i)\xi^T - \omega_i \tilde{\Psi}_{fi})
\] (31)
\[
- \sum_{i=1}^{n} T \tilde{\Psi}_{ai}(\varphi_i(\mathcal{X}_i)\varphi_i^T(\mathcal{X}_i) + jI_{p_2})\tilde{\Psi}_{ci}
- \sum_{i=1}^{n} T \tilde{\Psi}_{ai}(\varphi_i(\mathcal{X}_i)\varphi_i^T(\mathcal{X}_i) + jI_{p_2})
\] (32)
\[
\times (\tilde{\Psi}_{ai} - \tilde{\Psi}_{ci}) + \eta_{ai}\tilde{\Psi}_{ci})
\leq \mu \sum_{i=1}^{n} \xi_i^T K_i (f_i + u_i - \dot{x}_i) - \sum_{i=1}^{n} \xi_i^T K_i \gamma_i
+ \sum_{i=1}^{n} T \tilde{\Psi}_{fi}(\mu R_i^{-1} \varphi_i(X_i)\xi^T - \omega_i \tilde{\Psi}_{fi})
\] (33)
where \( K = \text{diag}(K_1, \ldots, K_n) \in \mathbb{R}^{n \times n} \), \( z_i^T \in \mathbb{R}^{n} \), \( \gamma = [\gamma_1^T, \ldots, \gamma_n^T] \in \mathbb{R}^{n \times n} \), and \( \mu = \lambda_{\max}(L) \).

Substituting (19) and (27) into (31), we have

\[
\dot{V} \leq \sum_{i=1}^{n} \xi_i^2 \left( -\mu R_i^{-1} \xi_i - \mu R_i^{-1} \frac{d\eta_i}{dt} \right) + \sum_{i=1}^{n} \text{Tr} \left\{ \eta_i \frac{d}{dt} \tilde{\Psi}_{ai} \tilde{\Psi}_{ai}^T \left( \mathcal{A}_i \right) \right\} + \sum_{i=1}^{n} \text{Tr} \left\{ \eta_i \tilde{\Psi}_{ai}^T \left( \mathcal{A}_i \right) \tilde{\Psi}_{ai} \left( \mathcal{A}_i \right) \right\} + \sum_{i=1}^{n} \text{Tr} \left\{ \eta_i \tilde{\Psi}_{ai}^T \left( \mathcal{A}_i \right) \tilde{\Psi}_{ai} \left( \mathcal{B}_i \right) \right\} + \sum_{i=1}^{n} \text{Tr} \left\{ \eta_i \tilde{\Psi}_{ai}^T \left( \mathcal{B}_i \right) \tilde{\Psi}_{ai} \left( \mathcal{B}_i \right) \right\}
\]

Substituting (33)–(37) into (32), we can get

\[
\dot{V} \leq -\sum_{i=1}^{n} \frac{n R_i^{-1}}{4} \left( h_i - \frac{3}{4} \right) \| \xi_i \|^2 - \sum_{i=1}^{n} \frac{n R_i^{-1}}{4} \text{Tr} \left\{ \tilde{\Psi}_{ci}^T \tilde{\Psi}_{ci} \right\}
\]

(38)
There exists a constant $\bar{\rho}$ such that $\rho \leq \bar{\rho}$. Then, we can get that

$$
V \leq -\frac{\nu_1}{2} \sum_{i=1}^{n} \|\xi_i\|^2 - \frac{\nu_2}{2} \sum_{i=1}^{n} Tr\{\hat{\Psi}_c^T \hat{P}_i^{-1} \hat{\Psi}_f\} - \frac{\nu_3}{2} \sum_{i=1}^{n} Tr\{\hat{\Psi}_e^T \hat{\Psi}_e\} + \bar{\rho}
$$

where $\nu_1 = \min_{i=1,\ldots,n} \left\{2 \left(\mu R_i^{-1} h_i - 3 \mu R_i^{-1} \right) \right\}$, $\nu_2 = \min_{i=1,\ldots,n} \left\{\lambda_{\max}(P_{i1}) \right\}$, $\nu_3 = \min_{i=1,\ldots,n} \left\{\eta_\phi \lambda_{\min}(\varphi_i) \right\}$, $\nu = \min\{\nu_1, \nu_2, \nu_3\}$, $\lambda_{\max}(\varphi_i)$ represents the minimal eigenvalue of matrix $\varphi_i(X_i)$.

From (40), we can get that $V \in L_\infty$. Then, $\xi_i \in L_\infty$, $\hat{\Psi}_f \in L_\infty$, and $\hat{\Psi}_e \in L_\infty$ can be obtained. As $\xi_i$ is bounded, it follows from (7) that $-\rho_\xi \in (-\bar{\rho}, \bar{\rho})$. Due to $\xi = \tau$ for $t \geq T$, we get

$$
-\hat{\rho} \xi_i (t - \tau) + b_i z_i(t) \leq \bar{\rho} \xi_i(t)
$$

for $t \geq T$, with $\xi = \left[\rho_\xi, \ldots, \rho_\xi \right]^T \in \mathbb{R}^n$, and then, the expected formation can be realized. It can be further obtained that the formation error can converge to the specified accuracy within a prescribed time.

**Remark 8:** The performance index functions specified in [15], [16], [17], [18], [19], [20], and [21] did not involve performance weights while the results presented in this article can make a tradeoff between the formation error cost and control input energy cost by choosing appropriate performance weights. By choosing $R_i$, which is smaller than $Q_i$, the formation error can converge to the desired accuracy with a faster convergence speed. Compared with the results in [15], [16], [17], [18], [19], [20], and [21], the presented control scheme is more conducive to engineering application.

**Remark 9:** It should be mentioned that the following knowledge on MAS (1) is required: the number of agents; the general connection between agents; the formation information loaded in each agent.

**Remark 10:** It is obvious that the controller parameter $h_i$ should be selected under the condition that $h_i > \frac{2}{\alpha_2}$. In practical engineering, the Laplacian matrix $L$, the communication matrix $B$, and the weight matrix $R_i$ can be determined in advance, that is, the value of $\frac{2R_i^{-1}}{\alpha_2}$ can be obtained in advance. Then, the selection of $h_i$ should satisfy $\eta_\alpha > \frac{\mu R_i^{-1}}{2}$. Finally, the selection of $\eta_\alpha$ should satisfy $\eta_\alpha > \eta_\phi > \frac{\eta_\alpha}{\alpha_2}$.

The block diagram of MAS is given in Fig. 1. Moreover, the proposed design procedure is given in Algorithm 1.

**IV. SIMULATION EXAMPLE**

In this section, to verify the feasibility of the scheme, the following MAS is considered:

$$
\dot{x}_i = -\tau_i x_i - \left[\begin{array}{c}
0.3x_{i1} \cos^2(\kappa_i x_{i1}) \\
x_{i2} - 0.5 \sin^2(\kappa_i x_{i2})
\end{array}\right] + u_i
$$

where $\kappa_i = \{1, 2, 3, 4\}$

$$
\begin{align*}
\eta_\alpha &= \left[\begin{array}{c}
-0.3, 0.1, -0.3, 0.1
\end{array}\right]^T, \\
\kappa_1, \kappa_2, \kappa_3, \kappa_4^T &= [0.5, 0.4, -5.5, -1.5]^T.
\end{align*}
$$

The adjacency matrix is selected as

$$
A = \left[\begin{array}{cccc}
0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0
\end{array}\right]
$$

**Algorithm 1: Prescribed-Time Formation Control via Fuzzy RL**

**Input:** $x_i(0), \epsilon_i, A, D, B, L, T, \epsilon_i - \epsilon_0 < 1, R_i > 0, \mu = \lambda_{\max}(L)$

**Initialization:** $x_1(0), x_2(0), \Phi_f(0), \Phi_e(0), \Phi_a(0)$

**Parameters:**

$$
\rho_\phi > 0, P_i > 0, \omega_i > 0, \eta_\alpha > 0, \eta_\phi > 0, i > 0, h_i > \frac{3}{2}. \quad \eta_\alpha > n_i^2 > \frac{\mu R_i^{-1}}{4}
$$

**for** follower agent $i$ 

- Compute $e_i$ with (3)
- Compute $\xi$ with (6)
- Compute $\xi_i$ with (7)
- Define fuzzy IF–THEN rules, input
  - $X_i = [x_i, x_{i2}, x_{i3}, x_{i4}, \xi_i, \xi_i, \xi_i]^T$, determine $X_i^0$, determine fuzzy membership functions $\mu_{\hat{\Psi}}(X_i)$, and then determine fuzzy basis function vector $\varphi_f(X_i)$
- Define fuzzy IF–THEN rules, input
  - $\mathcal{X}_i = [x_i, x_{i2}, x_{i3}, x_{i4}, \xi_i, \xi_i, \xi_i]^T$, determine fuzzy membership functions $\mu_{\hat{\Psi}}^0(\mathcal{X}_i)$, and then determine fuzzy basis function vector $\varphi_f(\mathcal{X}_i)$
- Compute $u_i$ with (27)
- Update $\Psi_f$ with (24)
- Update $\Psi_e$ with (26)
- Update $\Psi_a$ with (28)

**end for**
then \( D = \text{diag}\{1, 2, 1\} \), and the Laplacian matrix is set as
\[
L = \begin{bmatrix}
1 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 1 \\
\end{bmatrix}
\]
and the communication matrix is set as \( B = \text{diag}\{1, 0, 0, 0, 0\} \).

The fuzzy sets of FLS \( \Psi^T_{f_i} \varphi_i(X_i) \) are defined over the interval \([-23/2, 23/2]\). Define the vector \( X_i \) as \( X_i = [x_i, x_d, \dot{x}_d, \dot{\zeta}_i] \) and for \( l = 1, 2, \ldots, 24 \), define
\[
X_i^0 = \begin{bmatrix}
\begin{bmatrix}
-25/2 + l, -25/2 + l \\
\vdots \\
-25/2 + l, -25/2 + l \\
\end{bmatrix}
\end{bmatrix}^T.
\]

Thus, the fuzzy membership functions of FLS \( \Psi^T_{f_i} \varphi_i(X_i) \) are given as
\[
\mu^i_{f_i}(X_i) = \exp\left(-\frac{(X_i - X_i^0)^T(X_i - X_i^0)}{2}\right).
\]

Then, one can obtain the fuzzy basis function vector as \( \varphi_i(X_i) = [\varphi^1_i(X_i), \varphi^2_i(X_i), \ldots, \varphi^{24}_i(X_i)] \), where
\[
\varphi^1_i(X_i) = \sum_{l=1}^{24} \mu^l_{f_i}. 
\]

The fuzzy sets of FLS \( \Psi^T_{i} \varphi_i(\mathcal{X}_i) \) are defined over the interval \([-11/2, 11/2]\). Define the vector \( \mathcal{X}_i \) as \( \mathcal{X}_i = [x_i, \xi_i, x_d, \dot{x}_d, \dot{\zeta}_i] \) and for \( \ell = 1, 2, \ldots, 12 \), define
\[
\mathcal{X}_i^0 = \begin{bmatrix}
\begin{bmatrix}
-13/2 + \ell, -13/2 + \ell \\
\vdots \\
-13/2 + \ell, -13/2 + \ell \\
\end{bmatrix}
\end{bmatrix}^T.
\]

Thus, the fuzzy membership functions of FLS \( \Psi^T_{i} \varphi_i(\mathcal{X}_i) \) are given as
\[
\mu^i_{f_i}(\mathcal{X}_i) = \exp\left(-\frac{(\mathcal{X}_i - \mathcal{X}_i^0)^T(\mathcal{X}_i - \mathcal{X}_i^0)}{2}\right).
\]

Then, one can get the fuzzy basis function vector as \( \varphi_i(\mathcal{X}_i) = [\varphi^1_i(\mathcal{X}_i), \varphi^2_i(\mathcal{X}_i), \ldots, \varphi^{12}_i(\mathcal{X}_i)] \), where \( \varphi^i(\mathcal{X}_i) = \sum_{i=1}^{12} \mu^i_{f_i} \).

For identifier updating laws (24), critic updating laws (26), and actor updating laws (28), the initial values are set as \( \Psi_{f_1}(0) = \Psi_{f_2}(0) = \Psi_{f_3}(0) = \Psi_{f_4}(0) = [0.1]_{24 \times 2} \), \( \Psi_{a_1}(0) = [0.92]_{12 \times 2} \), \( \Psi_{a_2}(0) = [0.94]_{12 \times 2} \), \( \Psi_{a_3}(0) = [0.95]_{12 \times 2} \), \( \Psi_{a_4}(0) = [0.96]_{12 \times 2} \), and \( \Psi_{a_1}(0) = [0.90]_{12 \times 2} \), \( \Psi_{a_2}(0) = [0.91]_{12 \times 2} \), \( \Psi_{a_3}(0) = [0.90]_{12 \times 2} \), \( \Psi_{a_4}(0) = [0.91]_{12 \times 2} \). The design parameters are selected as \( \omega = 2, \omega = 1.5, \omega = 3, \omega = 1.5, \eta = 1, j = 1, h = h = h = h = 45; \eta = 2 = \eta = \eta = \eta = 2 \). \( R = \text{diag}\{0.8, 0.8\}; \quad P_1 = \text{diag}\{0.5, \ldots, 0.5\}, P_2 = \text{diag}\{0.7, \ldots, 0.7\}, P_3 = \text{diag}\{0.3, \ldots, 0.3\}, P_4 = \text{diag}\{0.6, \ldots, 0.6\}. \) It can be seen from Fig. 2 that the formation can be achieved by the presented strategy. Fig. 3 shows the curves of errors. As observed from Fig. 3 that the errors \( z_i \) converge to the expected accuracy within 3(s). Figs. 4–6 show the norm of the identifier, critic, and actor parameter matrices. From Figs. 2 to 6, we can see that the presented method can realize the control goal. The comparison results between this article and the scheme in [20]
are displayed in Figs. 7–9. Fig. 7(a) and (b) shows the formation of agents under the proposed scheme and formation control [20], respectively. Fig. 8(a) and (b) shows formation errors $z_{i1}$ under the proposed scheme and formation control [20], respectively. Fig. 9(a) and (b) shows formation errors $z_{i2}$ under the proposed scheme and formation control [20], respectively. The proposed scheme shows higher convergence accuracy than that in [20].

V. CONCLUSION

In this article, the prescribed-time formation control problem was investigated for a class of MASs. First, a new auxiliary function was given, and then, a new performance index function considering formation error cost and control input energy cost was constructed. Then, by developing the RL algorithm with ACI structure, an optimal prescribed-time control approach was presented, which can minimize the constructed performance index function. In addition, the formation can be achieved, and the settling time can be prespecified as required. The validity of the presented control strategy was verified by an example.
Future work will focus on the prescribed-time formation control for MASs subjected to DoS attacks via RL.

References


