ISA Transactions 138 (2023) 63–73

Contents lists available at ScienceDirect

ISA Transactions

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

Research article

Adaptive formation control of leader–follower mobile robots using reinforcement learning and the Fourier series expansion

Gholamreza Khodamipour, Saeed Khorashadizadeh *, Mohsen Farshad

Faculty of Electrical and Computer Engineering, University of Birjand, 97175/615, Birjand, Iran

ARTICLE INFO

Article history:
Received 15 September 2022
Received in revised form 9 March 2023
Accepted 9 March 2023
Available online 23 March 2023

Keywords:
Formation control
Reinforcement learning
Actor–critic strategy
The Fourier series expansion
Leader–follower mobile robots

ABSTRACT

In this paper, a formation controller for leader–follower mobile robots is presented based on reinforcement learning and the Fourier series expansion. The controller is designed based on the dynamical model in which permanent magnet direct-current (DC) motors are included as actuator. Thus, motor voltages are the control signals and are designed based on the actor–critic strategy which is a well-known approach in the field of reinforcement learning. Stability analysis of formation control of leader–follower mobile robots using the proposed controller verifies that the closed-loop system is globally asymptotically stable. Due to the existence of sinusoidal terms in the model of mobile robots, the Fourier series expansion has been selected to construct the actor and critic, while previous related works utilized neural networks in actor and critic. In comparison with neural networks, the Fourier series expansion is simpler and involves the designer in fewer tuning parameters. In simulation studies, it has been assumed that some follower robots can play the role of leader for the other follower robots behind it. Simulation results show that there is no need to use large number of the sinusoidal terms in the Fourier series expansion and just the first three terms can overcome uncertainties. In addition, the proposed controller reduced the performance index of tracking errors considerably in comparison with radial basis function neural networks (RBFNN).

1. Introduction

Using mobile robots instead of the human is one of the most important achievements in these last decades in unstructured environments which require autonomy. Due to the fast decision making and high accuracy of the robotic systems, mobile robots are widely used in many different applications such as patrolling surveillance, emergency rescue operations planetary exploration, industrial automation, personal services, entertainment, reconnaissance, medical care, and so on [1]. In some applications, several mobile robots must cooperate in a fixed form to increase safety and efficiency. Thus, formation control is of great importance. Formation control is not confined to mobile robots. For example, in order to take the advantages of aerodynamics facts, aircraft flying in a V-shaped formation is usual, since it will minimize fuel loss. This formation is inspired by flocks of birds flying in the sky that use this shape to increase their efficiency [2]. Moreover, in automated formation control, several aircrafts can be flown by one pilot simultaneously that results in reducing fatigue in long-distance flight. Moreover, only one piloted aircraft can control a large group of semi-autonomous pilotless aircraft. Other applications of formation control can be easily observed in spacecrafts. Several small spacecrafts can accurately play the role of sensor arrays or place sensors in fixed geometric shape. Also, box pushing using mobile robots is a well-known problem in cooperative robotics.

Various controllers for formation control have been presented in recent years, such as the behavior-based method [3], the leader-following strategy [3,4], and the consensus-based approach [5]. In this paper, the formation control using leader–follower strategy is studied in which the separation-bearing parameters will be tuned in which the followers stay at a specified separation and bearing from its designated leader. In [6], a velocity control signal is designed by back-stepping kinematic equations of separation and bearing into dynamics. This paper is based on the strategy presented in [6].

Most of presented controllers for mobile robots and leader–follower formation control are based on the torque control strategy in which the control law calculates the applied torques on the mobile robot [7–9]. However, in practical applications, this torque cannot be produced directly by the controller located on the processors. In fact, this torque is generated by actuators such as permanent magnet DC motors. Thus, this question arises that how the motor must be excited to generate that torque calculated by the torque-based controller. To solve this problem, voltage

* Corresponding author.
E-mail address: s.khorashadizadeh@birjand.ac.ir (S. Khorashadizadeh).

https://doi.org/10.1016/j.isatra.2023.03.009
0019-0578/

control strategy has been introduced in which the control input is the applied voltages on the motors' terminals [10]. This paper utilizes this strategy and consequently a voltage-based controller will be proposed.

Generally, the adaptive controllers presented for nonlinear systems can be classified into 2 main categories: Model-based and model-free controllers. Model-based controllers are faced by some fundamental challenges such as uncertainties and sensing requirements [11]. Therefore, model-free controllers are more preferable. Model-free approaches can be classified into 2 main groups: intelligent controllers and function approximation-based controllers. Among intelligent control approaches, neuro-fuzzy controllers have been extensively studied [12,13]. The universal approximation property is the main common feature of these researches. Neural networks such as radial-basis function neural networks (RBFNN) are also intelligent tools that have been used in formation control [14,15].

Reinforcement learning (RL) controllers are also considered as intelligent controllers in which the agents (controllers) learn their strategy and calculate the control action by trial-and-error exploration mixed with reward signals originated by the environment (the system under control) [16]. Recently, deep reinforcement learning has been presented and tested on formation control of mobile robots in which a convolutional neural network is utilized for image processing [17]. Actor–critic is one of important strategies in RL. Usually, neural networks or adaptive fuzzy systems are used to approximate the actors and critics [18–20]. However, neural networks and fuzzy systems are involved in a large number of tuning parameters [21]. Therefore, function approximation-based approaches using the Taylor series [22], Bernstein polynomials [23], Szász–Mirakyan operator [24], the Fourier series expansion and Legendre polynomials [25] have been presented in which the universal approximation property is satisfied and the number of tuning parameters is fewer in comparison with neural networks and fuzzy systems. This paper uses these techniques and proposes a novel actor–critic RL controller.

The proposed controller in this paper can be considered as a combination of intelligent controllers and function approximation-based techniques. It is intelligent due to utilizing reinforcement learning (actor–critic strategy). Also, it is a function approximation-based controller due to using the Fourier series expansion in the actor and critic. As mentioned above, previous related works [18–20] utilize neural networks and fuzzy systems in the actor and critic part and consequently involve the designer in a large number of tuning parameters. However, there are much fewer parameters in the Fourier series expansion. The proposed controller in this paper uses just the first three terms of the Fourier series expansion in the actor and critic.

The innovations of this paper can be summarized as follow. ***Firstly***, a model-free voltage-based controller has been proposed. Although the controller presented in [10] is a voltage-based controller for mobile robots, it needs current sensors and some nominal parameters of the motors, while the proposed controller does not need current sensors. ***Secondly***, an actor–critic RL controller for formation control of leader–follower mobile robots with guaranteed asymptotic stability has been presented. Some RL-based controllers for formation control have been presented in the literature [17] without stability analysis. Although the RL-based formation controller presented in [18] is accompanied by a rigorous stability analysis, it has been tested on some simple multi-agents, while the proposed controller is applied on leader–follower mobile robots that are more complicated due the separation and bearing dynamics between leaders and followers. ***Thirdly***, it is a combination of two categories of model-free controllers (intelligent controllers and function-approximation-based controllers) due to using the Fourier series expansion in an actor–critic-based structure.

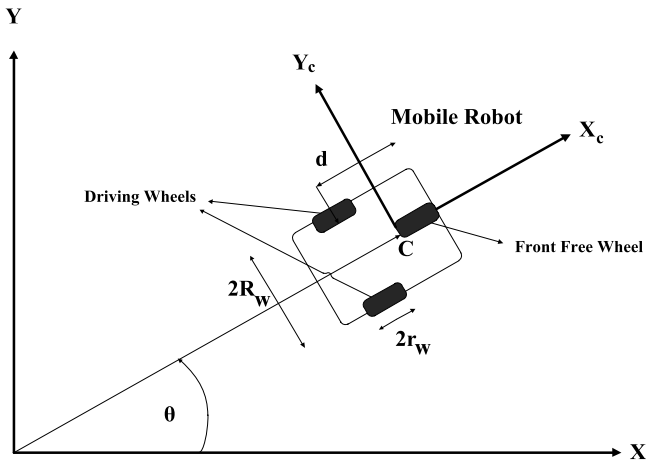

Fig. 1. A non-holonomic mobile robot.

The rest of this paper is organized as follow. In Section 2, the modeling of a mobile robot with its DC motors are explained. In Section 3, the Fourier series expansion is introduced. In Section 4, the proposed RL-based controller is designed and the adaptive laws in actor and critic parts are derived. Stability analysis is discussed in Section 5. In Section 6, simulations and comparisons are presented. In order to highlight the superiority of the designed controller, a formation scenario for control of mobile robots is simulated and the performance of the Fourier series expansion and neural networks are compared. Finally, some concluding remarks are presented in Section 7.

2. Modeling

Let us consider the dynamics of a mobile robot as [26]

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}}+\mathbf{V_m}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}}+\mathbf{F}(\mathbf{q},\dot{\mathbf{q}})+\mathbf{G}(\mathbf{q})+\boldsymbol{\tau_d}(t)=\mathbf{B}\boldsymbol{\tau}(t)-\mathbf{A}(\mathbf{q})^T\boldsymbol{\lambda}(t) \quad (1)$$

where $\mathbf{M}\in\Re^{n\times n}$ is a positive definite symmetric inertia matrix, $\mathbf{V_m}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}}\in\Re^{n\times n}$ denotes the Coriolis and centrifugal torques, $\mathbf{F}\in\Re^{n\times 1}$ is the friction torques and $\mathbf{G}\in\Re^{n\times 1}$ is the vector of gravitational torques. Also, $\boldsymbol{\tau_d}$ is the bounded unknown disturbances and un-modeled dynamics. The matrix $\mathbf{B}\in\Re^{n\times r}$ is the input coefficients and $\boldsymbol{\tau}\in\Re^{n\times 1}$ is a control input. Also, $\mathbf{A}\in\Re^{m\times n}$ is the weight matrix of constraints and $\boldsymbol{\lambda}\in\Re^{m\times 1}$ is a constant force vector and m is the number of constraints. It considered that all constraints equalities are time independence, so

$$\mathbf{A}\dot{\mathbf{q}}=0 \quad (2)$$

Let $\mathbf{S}(q)$ be a full rank matrix with the size of $(n-m)$ formed by a set of smooth and linearly independent vector fields spanning the null space of $\mathbf{A}$, i.e.,

$$\mathbf{S}^T\mathbf{A}^T=0 \quad (3)$$

According to [26], it is possible to use (2) and (3) to find a vector $\mathbf{v}(t)\in\Re^{n-m}$, that

$$\dot{\mathbf{q}}=\mathbf{S}\mathbf{v}(t) \quad (4)$$

Fig. 1 shows a non-holonomic mobile robot which has two driving wheels in the back and a free wheel in the front of robot. Two DC motors as its actuators produce the torques of rear wheels. The vector $\mathbf{q}=[x_c,y_c,\theta]^T$ shows the robot position and orientation in the Cartesian frame, which denotes the actual Cartesian position of vehicle mass center and orientation of the physical robot.

The mobile robot, due to the nonholonomic constraints, can only move in directions normal to the driving wheel axis [27]. Therefore,

$$\dot{y}_c \cos\theta - \dot{x}_c \sin\theta - d\dot{\theta} = 0 \tag{5}$$

The kinematic equation for the front of the mobile robot can be obtained as:

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -d\sin\theta \\ \sin\theta & d\cos\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} = \mathbf{S}\mathbf{v} \quad , \mathbf{v} = \begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{6}$$

In the rest of this paper, the notations v_1 and v_2 are used to show the linear (v) and angular velocity (ω) of the mobile robot, respectively. It is be assumed that

$$\begin{aligned} |v_1| &\le V_{\max} \\ |v_2| &\le W_{\max} \end{aligned} \tag{7}$$

The relationship between motor angular speeds and Cartesian speeds is given by:

$$\dot{\mathbf{q}} = \mathbf{J}(\mathbf{q})\dot{\boldsymbol{\theta}}_{\mathbf{w}} \tag{8}$$

where $\dot{\boldsymbol{\theta}}_{\mathbf{w}} = \begin{bmatrix} \dot{\theta}_{wr} & \dot{\theta}_{wl} \end{bmatrix}^T$ is the angular speeds of right and left motors and the Jacobian matrix $\mathbf{J}(\mathbf{q})$ is in the form of

$$\mathbf{J}(\mathbf{q}) = \frac{r_w}{2} \begin{bmatrix} \cos\theta & \cos\theta \\ \sin\theta & \sin\theta \\ 1/R_w & -1/R_w \end{bmatrix} \tag{9}$$

In this paper, due to the horizontal frame and mobile trajectory, the gravitational vector is zeros, i.e., $\mathbf{G} = 0$, and the matrices in dynamical model of mobile robot in (1) are as follows

$$\begin{aligned} \mathbf{M} &= \begin{bmatrix} m & 0 & md\sin\theta \\ 0 & m & -md\cos\theta \\ md\sin\theta & -md\cos\theta & I \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} md\dot{\theta}^2\cos\theta \\ md\dot{\theta}^2\sin\theta \\ 0 \end{bmatrix} \\ \mathbf{B} &= \frac{1}{r_w} \begin{bmatrix} \cos\theta & \cos\theta \\ \sin\theta & \sin\theta \\ R_w & -R_w \end{bmatrix}, \quad \boldsymbol{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}, \quad \mathbf{A}^T = \begin{bmatrix} -\sin\theta \\ \cos\theta \\ -d \end{bmatrix}, \\ \lambda &= -m(\dot{x}_c\cos\theta + \dot{y}_c\sin\theta)\dot{\theta} \end{aligned} \tag{10}$$

In order to find a proper velocity controller, the dynamic Eq. (1) is rewritten as

$$\mathbf{S}^T\mathbf{M}\mathbf{S}\dot{\mathbf{v}} + \mathbf{S}^T(\mathbf{M}\dot{\mathbf{S}} + \mathbf{V}_m\mathbf{S})\mathbf{v} + \bar{\mathbf{F}} + \bar{\tau}_d = \mathbf{S}^T\mathbf{B}\boldsymbol{\tau} \tag{11}$$

in which the transformation (4) has been used. According to (11),

$$\bar{\mathbf{M}}\dot{\mathbf{v}} + \bar{\mathbf{V}}_m\mathbf{v} + \bar{\mathbf{F}} + \bar{\tau}_d = \bar{\mathbf{B}}\boldsymbol{\tau} \tag{12}$$

in which

$$\bar{\mathbf{M}} = \mathbf{S}^T\mathbf{M}\mathbf{S}, \quad \bar{\mathbf{V}}_m = \mathbf{S}^T(\mathbf{M}\dot{\mathbf{S}} + \mathbf{V}_m\mathbf{S}) \tag{13}$$

$$\bar{\tau} = \bar{\mathbf{B}}\boldsymbol{\tau}, \quad \bar{\mathbf{B}} = \mathbf{S}^T\mathbf{B} \tag{14}$$

Now, to have a complete model of mobile robot including the DC motors as its actuators, the dynamic equations of DC motors are introduced. For the torques in permanent magnet DC motors, we have

$$\mathbf{J}_m\ddot{\boldsymbol{\theta}}_m + \mathbf{B}_m\dot{\boldsymbol{\theta}}_m + \mathbf{r}\,\boldsymbol{\tau} = \mathbf{k}_m\mathbf{I}_a \tag{15}$$

where $\mathbf{J}_m$ is the motor shaft inertia diagonal matrix, $\boldsymbol{\theta}_m \in \Re^n$ is the motors angular position vector ($\boldsymbol{\theta}_m = \mathbf{r}^{-1}\boldsymbol{\theta}_w$), $\mathbf{B}_m$ is the damping, $\mathbf{r}$ is the reduction gear, $\boldsymbol{\tau}$ is the electromagnetics torque, $\mathbf{k}_m$ denotes the diagonal matrix of torque constant and $\mathbf{I}_a$ is the motor current vector. Also, by (8) it can be shown that

$$\dot{\boldsymbol{\theta}}_{\mathbf{w}} = \mathbf{J}(\mathbf{q})^\dagger\dot{\mathbf{q}} \tag{16}$$

where

$$\mathbf{J}(\mathbf{q})^\dagger = [\mathbf{J}(\mathbf{q})^T\mathbf{J}(\mathbf{q})]^{-1}\mathbf{J}(\mathbf{q})^T = \frac{1}{r_w} \begin{bmatrix} \cos\theta & \sin\theta & R_w \\ \cos\theta & \sin\theta & -R_w \end{bmatrix} \tag{17}$$

The derivative of (16) can be computed as

$$\ddot{\boldsymbol{\theta}}_{\mathbf{w}} = \dot{\mathbf{J}}(\mathbf{q})^\dagger\dot{\mathbf{q}} + \mathbf{J}(\mathbf{q})^\dagger\ddot{\mathbf{q}} \tag{18}$$

where

$$\dot{\mathbf{J}}(\mathbf{q})^\dagger = \frac{\dot{\theta}}{r_w} \begin{bmatrix} -\sin\theta & \cos\theta & 0 \\ -\sin\theta & \cos\theta & 0 \end{bmatrix} \tag{19}$$

The electrical equation of permanent magnet DC motors is given by

$$\mathbf{R}\mathbf{I}_a + \mathbf{L}_a\dot{\mathbf{I}}_a + \mathbf{k}_b\dot{\boldsymbol{\theta}}_m = \mathbf{u}\,(t) \tag{20}$$

in which $\mathbf{u} \in \Re^n$ denotes the vector of motor voltages, $\boldsymbol{\theta}_m \in \Re^n$ is angular positions and we have $\boldsymbol{\theta}_m = \mathbf{r}^{-1}\mathbf{q}$, and $\mathbf{R}$, $\mathbf{L}_a$, $\mathbf{k}_b$ are diagonal matrices of resistance, inductance and back electromotive force (EMF), respectively. Substituting $\mathbf{I}_a$ from (15) into (20) and using (16) and (18) result in

$$\begin{aligned} \mathbf{u}(t) = {} & \mathbf{R}\mathbf{k}_m^{-1}(\mathbf{J}_m\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\mathbf{S} + \mathbf{r}\ \bar{\mathbf{B}}^{-1}\bar{\mathbf{M}})\dot{\mathbf{v}} \\ & + \mathbf{R}\mathbf{k}_m^{-1}(\mathbf{J}_m\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\dot{\mathbf{S}} + \mathbf{J}_m\mathbf{r}^{-1}\dot{\mathbf{J}}(\mathbf{q})^\dagger\mathbf{S} \\ & + \mathbf{B}_m\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\mathbf{S} + \mathbf{r}\bar{\mathbf{B}}^{-1}\bar{\mathbf{V}}_{\mathbf{m}})\mathbf{v} \\ & + \mathbf{R}\mathbf{k}_m^{-1}\mathbf{r}\,\bar{\mathbf{B}}^{-1}(\bar{\mathbf{F}} + \bar{\tau}_{\mathbf{d}}) + \mathbf{L}_a\dot{\mathbf{I}}_a + \mathbf{k}_b\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\mathbf{S}\mathbf{v} \end{aligned} \tag{21}$$

Eq. (21) can be simplified to

$$\mathbf{D}\,(\mathbf{q})\,\dot{\mathbf{v}} + \mathbf{N}\,(\mathbf{q}, \dot{\mathbf{q}})\,\mathbf{v} + \mathbf{H}\,(\mathbf{q}) + \mathbf{v}_L = \mathbf{u} \tag{22}$$

where

$$\mathbf{D}\,(\mathbf{q}) = \mathbf{R}\mathbf{k}_m^{-1}(\mathbf{J}_m\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\mathbf{S} + \mathbf{r}\ \bar{\mathbf{B}}^{-1}\bar{\mathbf{M}}) \tag{23}$$

$$\begin{aligned} \mathbf{N}\,(\mathbf{q}, \dot{\mathbf{q}}) = {} & \mathbf{R}\mathbf{k}_m^{-1}(\mathbf{J}_m\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\dot{\mathbf{S}} + \mathbf{J}_m\mathbf{r}^{-1}\dot{\mathbf{J}}(\mathbf{q})^\dagger\mathbf{S} \\ & + \mathbf{B}_m\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\mathbf{S} + \mathbf{r}\bar{\mathbf{B}}^{-1}\bar{\mathbf{V}}_{\mathbf{m}}) + \mathbf{k}_b\mathbf{r}^{-1}\mathbf{J}(\mathbf{q})^\dagger\mathbf{S} \end{aligned} \tag{24}$$

$$\mathbf{H}\,(\mathbf{q}) = \mathbf{R}\mathbf{k}_m^{-1}\mathbf{r}\,\bar{\mathbf{B}}^{-1}(\bar{\mathbf{F}} + \bar{\tau}_{\mathbf{d}}) \tag{25}$$

in which $\mathbf{v}_L = \mathbf{L}_a\dot{\mathbf{I}}_a$ is considered as unmodeled dynamic.

3. The fourier series expansion

According to [28], the non-periodic function $f\,(t)$ defined on $[t_1, t_2]$ can be expressed as

$$f\,(t) = a_0 + \sum_{n=1}^{\infty} a_n\cos\left(\frac{2n\pi t}{T}\right) + b_n\ \sin\left(\frac{2n\pi t}{T}\right) \tag{26}$$

where $T = t_2 - t_1$. Considering the first $m+1$ harmonics, (26) can be rewritten as

$$\begin{aligned} f\,(t) &= a_0 + \sum_{n=1}^{m} a_n\cos\left(\frac{2n\pi t}{T}\right) + b_n\ \sin\left(\frac{2n\pi t}{T}\right) + \varepsilon_m \\ &= {\varphi_f}^T\omega_f + \varepsilon_m \end{aligned} \tag{27}$$

in which

$$\omega_f = [a_0 \quad a_1 \quad b_1 \quad \dots \quad a_m \quad b_m]^T \tag{28}$$

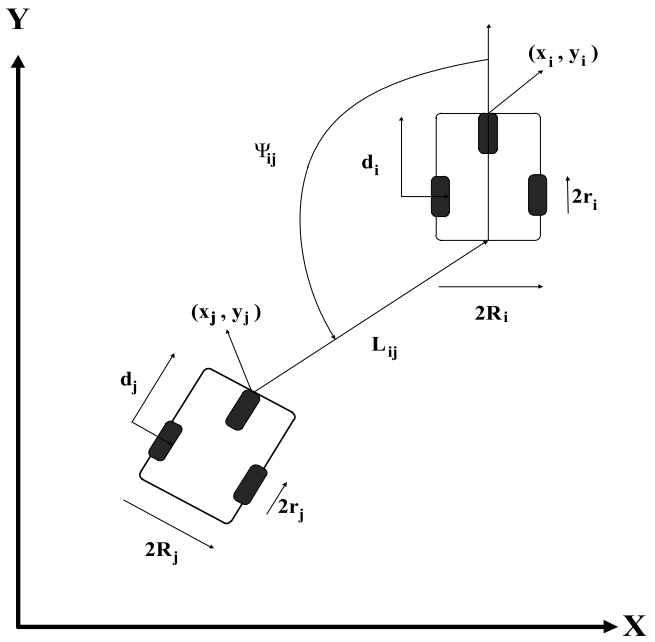

Fig. 2. Separation bearing formation control problem.

$$\varphi_f(t) = \left[1 \quad \cos\left(\frac{2\pi t}{T}\right) \quad \sin\left(\frac{2\pi t}{T}\right) \quad \dots \quad \cos\left(\frac{2m\pi t}{T}\right) \quad \sin\left(\frac{2m\pi t}{T}\right)\right]^T \tag{29}$$

where ε_m denotes approximation error, ω_f is unknown weight vector, and φ_f is the regressor vector.

4. Designing the controller

4.1. Back stepping control design

In this section an RL controller is proposed to solve the leader–follower formation control of mobile robots. In the rest of this paper, indexes "i" denotes the leader robot and "j" denotes the follower one. The introduced controller, due to the voltage control strategy [29], produced voltage of DC motors to achieve

$$\begin{aligned} &\lim_{t\to\infty}(L_{ijd} - L_{ij}) = 0 \\ &\lim_{t\to\infty}(\psi_{ijd} - \psi_{ij}) = 0 \end{aligned} \tag{30}$$

where L_{ij} and L_{ijd} are the measured and desired separation of the follower "j" with respect to leader "i", ψ_{ij} and ψ_{ijd} are the measured and desired bearing of the follower "j" with respect to leader "i". Eq. (30) shows that the separation and bearing errors of the formation control will be zero as $t \to \infty$. Fig. 2 shows an example of leader–follower mobile robot in formation control problem.

Define the tracking controller error for the follower "j" as:

$$e_j = T_{ej}(q_{jr} - q_j) = \begin{bmatrix} e_{j1} \\ e_{j2} \\ e_{j3} \end{bmatrix} = \begin{bmatrix} \cos\theta_j & \sin\theta_j & 0 \\ -\sin\theta_j & \cos\theta_j & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{jr} - x_j \\ y_{jr} - y_j \\ \theta_{jr} - \theta_j \end{bmatrix} \tag{31}$$

$\dot{\mathbf{q}}_{jr} = \begin{bmatrix} \dot{x}_{jr} & \dot{y}_{jr} & \dot{\theta}_{jr} \end{bmatrix}^T$, $\dot{y}_{jr} = v_{1jr}\sin\theta_{jr}$, $\dot{x}_{jr} = v_{1jr}\cos\theta_{jr}$, $\dot{\theta}_{jr} = v_{2jr}$

Now, the reference position of the follower "j" can be computed due to the desired separation L_{ijd} and bearing ψ_{ijd} as [6]

$$x_{jr} = x_i - d_i\cos\theta_i + L_{ijd}\cos(\psi_{ijd} + \theta_i) \tag{32}$$

$$y_{jr} = y_i - d_i\sin\theta_i + L_{ijd}\sin(\psi_{ijd} + \theta_i) \tag{33}$$

The actual measured position of the follower "j" is defined as

$$x_j = x_i - d_i\cos\theta_i + L_{ij}\cos(\psi_{ij} + \theta_i) \tag{34}$$

$$y_j = y_i - d_i\sin\theta_i + L_{ij}\sin(\psi_{ij} + \theta_i) \tag{35}$$

Substitution of (32)–(35) into (31), and applying the trigonometric identities, the kinematic error of formation control problem is achieved as

$$e_j = \begin{bmatrix} e_{j1} \\ e_{j2} \\ e_{j3} \end{bmatrix} = \begin{bmatrix} L_{ijd}\cos(\psi_{ijd} + \theta_{ij}) - L_{ij}\cos(\psi_{ij} + \theta_{ij}) \\ L_{ijd}\sin(\psi_{ijd} + \theta_{ij}) - L_{ij}\sin(\psi_{ij} + \theta_{ij}) \\ \theta_{jr} - \theta_j \end{bmatrix} \tag{36}$$

where $\theta_{ij} = \theta_i - \theta_j$. Due to the constraint of nonholonomic mobile robots [26], the orientation of each robot in the formation problem will not be equal so it is not $\theta_{jr} = \theta_i$ for each robot. However, the reference of the follower "j" is chosen such that

$$\dot{\theta}_{jr} = \frac{1}{d_j}(v_{2i}L_{ijd}\cos(\psi_{ijd} + \theta_{ij}) + v_{1i}\sin\theta_{ij} + k_{j2}e_{j2}) \tag{37}$$

where v_{1j}, v_{2j}, v_{1i} and v_{2i} are the linear and angular velocities of follower and leader robot, respectively. Using (6) it can be computed that [30]

$$\begin{aligned} \dot{L}_{ij} &= v_{1j}\cos(\psi_{ij} + \theta_{ij}) - v_{1i}\cos\psi_{ij} + d_j v_{2j}\sin(\psi_{ij} + \theta_{ij}) \\ \dot{\psi}_{ij} &= \frac{1}{L_{ij}}(v_{1i}\sin\psi_{ij} - v_{1j}\sin(\psi_{ijd} + \theta_{ij}) \\ &\quad + d_j v_{2j}\cos(\psi_{ij} + \theta_{ij}) - L_{ij}v_{2i}) \end{aligned} \tag{38}$$

It is assumed that L_{ijd} and ψ_{ijd} are held constant. By the derivation of (36) and using (38), the error dynamics can be written as

$$\dot{e}_j = \begin{bmatrix} \dot{e}_{j1} \\ \dot{e}_{j2} \\ \dot{e}_{j3} \end{bmatrix} = \begin{bmatrix} -v_{1j} + v_{1i}\cos\theta_{ij} + v_{2j}e_{j2} - v_{2i}L_{ijd}\sin(\psi_{ijd} + \theta_{ij}) \\ -v_{2j}e_{j1} + v_{1i}\sin\theta_{ij} - d_j v_{2j} + v_{2i}L_{ijd}\cos(\psi_{ijd} + \theta_{ij}) \\ (v_{2i}L_{ijd}\cos(\psi_{ijd} + \theta_{ij}) + v_{1i}\sin\theta_{ij} + k_{j2}e_{j2})/d_j - v_{2j} \end{bmatrix} \tag{39}$$

Following velocity control inputs is used to stabilize the kinematic of follower to achieve the desired separation and bearing [6].

$$v_{jc} = \begin{bmatrix} v_{jc1} \\ v_{jc2} \end{bmatrix} = \begin{bmatrix} v_{1i}\cos\theta_{ij} + k_{j1}e_{j1} - v_{2i}L_{ijd}\sin(\psi_{ijd} + \theta_{ij}) \\ \frac{1}{d_j}(v_{2i}L_{ijd}\cos(\psi_{ijd} + \theta_{ij}) + v_{1i}\sin\theta_{ij} + k_{j2}e_{j2} + k_{j3}e_{j3}) \end{bmatrix} \tag{40}$$

where k_{j1}, k_{j2} and k_{j3} are controller positive design constant which will be chosen by Lyapunov methods in stability section. Now, the error of velocity tracking can be defined as

$$e_{jc} = \begin{bmatrix} e_{j4} \\ e_{j5} \end{bmatrix} = v_{jc} - v_j = \begin{bmatrix} v_{jc1} \\ v_{jc2} \end{bmatrix} - \begin{bmatrix} v_{1j} \\ v_{2j} \end{bmatrix} \tag{41}$$

where $v_j = v_{jc} - e_{jc}$. Therefore, Eq. (39) can be rewritten as

$$\dot{e}_j = \begin{bmatrix} \dot{e}_{j1} \\ \dot{e}_{j2} \\ \dot{e}_{j3} \end{bmatrix} = \begin{bmatrix} -k_{j1}e_{j1} + v_{2j}e_{j2} + e_{j4} \\ -k_{j2}e_{j2} - k_{j3}e_{j3} - v_{2j}e_{j1} + d_j e_{j5} \\ -k_{j3}e_{j3}/d_j + e_{j5} \end{bmatrix} \tag{42}$$

Considering the tracking error for leader and its derivative, the velocity control inputs for leaders are found as below:

$$e_i = \begin{bmatrix} e_{i1} \\ e_{i2} \\ e_{i3} \end{bmatrix} = \begin{bmatrix} \cos\theta_i & \sin\theta_i & 0 \\ -\sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r - x_i \\ y_r - y_i \\ \theta_r - \theta_i \end{bmatrix} \tag{43}$$

Therefore, we have

$$\dot{e}_i = \begin{bmatrix} \dot{e}_{i1} \\ \dot{e}_{i2} \\ \dot{e}_{i3} \end{bmatrix} = \begin{bmatrix} -v_{1i} + v_{ir}\cos e_{i3} + v_{2i}e_{i2} \\ -v_{2i}e_{i1} + v_{1ir}\sin e_{i3} \\ v_{2ir} - v_{2i} \end{bmatrix} \tag{44}$$

The control velocity for the leader "i" can be defined as

$$v_{ic} = \begin{bmatrix} v_{ic1} \\ v_{ic2} \end{bmatrix} = \begin{bmatrix} v_{1ir}\cos e_{i3} + k_{i1}e_{i1} \\ v_{2ir} + k_{i2}v_{1ir}e_{i2} + k_{i3}k_{i2}\sin e_{i3} \end{bmatrix} \tag{45}$$

Similar to the follower and Eq. (41), the error of velocity tracking for the leader robot can be defined as

$$e_{ic} = \begin{bmatrix} e_{i4} \\ e_{i5} \end{bmatrix} = v_{ic} - \overline{v}_i = \begin{bmatrix} v_{ic1} \\ v_{ic2} \end{bmatrix} - \begin{bmatrix} v_{1i} \\ v_{2i} \end{bmatrix} \tag{46}$$

where $\overline{v}_i = v_{ic} - e_{ic}$. Therefore, Eq. (44) can be rewritten as

$$\dot{e}_i = \begin{bmatrix} \dot{e}_{i1} \\ \dot{e}_{i2} \\ \dot{e}_{i3} \end{bmatrix} = \begin{bmatrix} -k_{i1}e_{i1} + v_{2i}e_{i2} + e_{i4} \\ -v_{2i}e_{i1} + v_{1ir}\sin e_{i3} \\ -k_{i2}v_{1ir}e_{i2} - k_{i3}k_{i2}\sin e_{i3} + e_{i5} \end{bmatrix} \tag{47}$$

4.2. RL controller

Eq. (22) explains the complete dynamics of mobile robot. The actor–critic RL controller structure is designed similar to the one introduced in [11]. By adding and subtracting $\mathbf{D}_j\dot{\mathbf{v}}_{jc}$ and $\mathbf{N}_j\mathbf{v}_{jc}$ in (22) for jth follower, the Eq. (22) can be rewritten as

$$\mathbf{D}_j\dot{\mathbf{v}}_j - \mathbf{D}_j\dot{\mathbf{v}}_{jc} + \mathbf{D}_j\dot{\mathbf{v}}_{jc} + \mathbf{N}_j\mathbf{v}_j - \mathbf{N}_j\mathbf{v}_{jc} + \mathbf{N}_j\mathbf{v}_{jc} + \mathbf{H}_j + \mathbf{v}_{Lj} = \mathbf{u}_j \tag{48}$$

Now, (48) can be simplified as

$$-\mathbf{D}_j\dot{\mathbf{e}}_{jc} + \mathbf{D}_j\dot{\mathbf{v}}_{jc} - \mathbf{N}_j\mathbf{e}_{jc} + \mathbf{N}_j\mathbf{v}_{jc} + \mathbf{H}_j + \mathbf{v}_{Lj} = \mathbf{u}_j \tag{49}$$

And (49) can be written as

$$\mathbf{D}_j\dot{\mathbf{e}}_{jc} = \mathbf{D}_j\dot{\mathbf{v}}_{jc} - \mathbf{N}_j\mathbf{e}_{jc} + \mathbf{N}_j\mathbf{v}_{jc} + \mathbf{H}_j + \mathbf{v}_{Lj} - \mathbf{u}_j \tag{50}$$

If $\mathbf{f}_j = \mathbf{D}_j\dot{\mathbf{v}}_{jc} + \mathbf{N}_j\mathbf{v}_{jc}$, then (50) changed to

$$\mathbf{D}_j\dot{\mathbf{e}}_{jc} = -\mathbf{N}_j\mathbf{e}_{jc} + \mathbf{f}_j + \mathbf{H}_j + \mathbf{v}_{Lj} - \mathbf{u}_j \tag{51}$$

where $\mathbf{u}_j$ denotes the vector of motor voltages as the input of jth follower mobile robot.

Assumption 1. lets the Fourier series approximation property hold for the function $\mathbf{g}_j = -\mathbf{N}_j\mathbf{e}_{jc} + \mathbf{f}_j + \mathbf{H}_j + \mathbf{v}_{Lj}$ with accuracy ε_{aj} for all followers j, in the compact set S_j.

Assumption 2. The Fourier series reconstruction error for all followers j is bounded by positive values such that $\|\varepsilon_{aj}\| < \delta_j$. (the index a shows the actor part of controller)

Using Fourier series expansion, there exist $\mathbf{g}_j = -\mathbf{N}_j\mathbf{e}_{jc} + \mathbf{f}_j + \mathbf{H}_j + \mathbf{v}_{Lj} = \varphi_j\omega^*{}_{aj} + \varepsilon_{aj}$ where $\omega^*{}_{aj}$ are actual weights and ε_{aj} is the approximation error. This approximation is done in actor part of RL controller to find a proper voltage control. The function g_j and its estimation are given by

$$\mathbf{g}_j = \varphi_j\omega^*_{aj} + \varepsilon_{aj} \tag{52}$$

$$\hat{\mathbf{g}}_j = \varphi_j\hat{\omega}_{aj} + u_{raj} \tag{53}$$

where $\hat{g}_j$ and $\hat{\omega}_{aj}$ are the Fourier series approximation of the g_j and $\omega^*{}_{aj}$, respectively.

The control voltage for the follower robot is designed as:

$$\mathbf{u}_j = K_{j4}\mathbf{e}_{jc} + \hat{\mathbf{g}}_j = K_{j4}\mathbf{e}_{jc} + \varphi_j\hat{\omega}_{aj} + u_{raj} \tag{54}$$

where u_{raj} is the robust term and index aj denotes the actor part of RL controller of jth follower. Substituting the proposed control voltage (54) into dynamic (51), the close loop dynamic of robot could be written as

$$\mathbf{D}_j\dot{\mathbf{e}}_{jc} = -K_{j4}\mathbf{e}_{jc} + \widetilde{\mathbf{g}}_j = -K_{j4}\mathbf{e}_{jc} + \varphi_j\widetilde{\omega}_{aj} + \varepsilon_{aj} - u_{raj} \tag{55}$$

where $\widetilde{g}_j = g_j - \hat{g}_j$ is the Fourier series approximation error and similarly $\widetilde{\omega}_j = \omega_j - \hat{\omega}_j$. In critic part of RL controller, the reinforcement signal (long-term cost function) is produced. This critic network signal is defined as [31]

$$J_j = e_{jc} + \|e_{jc}\|\,\hat{\omega}_{cj}{}^T\varphi_j{}^T \tag{56}$$

in which $\hat{\omega}_{cj}$ is estimation of $\omega^*_{cj} \in \Re^{2m+1}$. The long-term strategic utility function is usually defined as $Q(k) = \min\left\{\alpha^N p(k+1) + \alpha^{N-1}p(k+2) + \cdots + \alpha^{k+1}p(N) + \cdots\right\}$ where $0 < \alpha < 1$ is the discount factor and N is the horizon index and $p(k)$ is the utility function [32]. It should be noted that $Q(k)$ is not available and should be approximated [32]. The term $\varphi_j\hat{\omega}_{cj}$ in (34) represents this estimation. For the ith leader mobile robot, similar to (51), (54), (56), dynamic equation, control law and critic network signal can be written similar to (51), (54), (56) in the form of

$$\mathbf{D}_i\dot{\mathbf{e}}_{ic} = -\mathbf{N}_i\mathbf{e}_{ic} + \mathbf{f}_i + \mathbf{H}_i + \mathbf{v}_{Li} - \mathbf{u}_i \tag{57}$$

$$\mathbf{u}_i = K_{i4}\mathbf{e}_{ic} + \hat{\mathbf{g}}_i = K_{i4}\mathbf{e}_{ic} + \varphi_i\hat{\omega}_{ai} + u_{rai} \tag{58}$$

$$J_i = e_{ic} + \|e_{ic}\|\,\hat{\omega}_{ci}{}^T\varphi_i{}^T \tag{59}$$

5. Stability analysis

Two theorems are needed to complete the stability analysis. The first theorem relates to the follower robot and the second theorem discusses the leader robot and convergence analysis of the whole closed-loop system of leader–follower formation control.

Theorem 1 (*Stability Analysis for Followers*)**.** *The signals in the closed-loop system consisted of the system (51) and controller (54) are bounded and the tracking error e_{jc} asymptotically converges to zero if the following adaptation laws in actor and critic parts are hold*

$$\dot{\hat{\omega}}_{aj}{}^T = K_{cj}J_j^T\varphi_j - \eta_j K_{cj}\|e_{jc}\|\,\hat{\omega}_{aj}{}^T \tag{60}$$

$$\dot{\hat{\omega}}_{cj}{}^T = -K_{\omega j}\|e_{jc}\|\,\hat{\omega}_{aj}{}^T\varphi_j{}^T\varphi_j - \eta_j K_{\omega j}\|e_{jc}\|\,\hat{\omega}_{cj}{}^T \tag{61}$$

in which K_{cj}, $K_{\omega j}$ and η_j are designed positive constant scalars. In fact, it is assumed that ω^*_{aj} and ω^*_{cj} are unknown. Consider the following robust term.

$$u_{raj} = \frac{(\delta_j + \beta_j)(\delta_j + \beta_j)e_{jc}}{\|e_{jc}{}^T(\delta_j + \beta_j)\|} \tag{62}$$

The follower robots will be stable if the controller gains are selected as $k_{j1} > \frac{1}{2}$, $k_{j2} > \frac{d_j}{2}$ and $k_{j3} > \frac{d_j}{2}$. The proof of Theorem 1 is presented in the Appendix.

Assumption 3. Assume that a_j and c_j are upper bounds of the weight matrix ω^*_{aj} and ω^*_{cj}, respectively, i.e., $\|\omega^*_{aj}\| \le a_j$, $\|\omega^*_{cj}\| \le c_j$.

Theorem 2 (*Stability Analysis for Leaders*)**.** *The signals in the closed-loop system consisted of the system (57) and controller (58) are bounded and the tracking error e_{ic} asymptotically converges to zero*

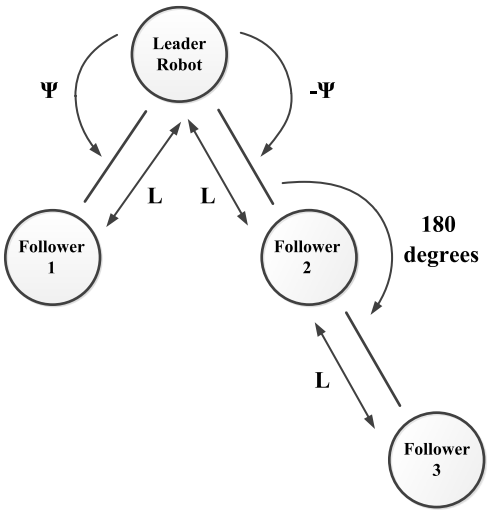

Fig. 3. Desired formation structure.

if the following adaptation laws are hold

$$\dot{\hat{\omega}}_{ai}^{T} = K_{ci} J_i^T \varphi_i - \eta_i K_{ci} \|e_{ic}\| \hat{\omega}_{ai}^{T} \tag{63}$$

$$\dot{\hat{\omega}}_{ci}^{T} = -K_{\omega i} \|e_{ic}\| \hat{\omega}_{ai}^{T} \varphi_i^{T} \varphi_i - \eta_i K_{\omega i} \|e_{ic}\| \hat{\omega}_{ci}^{T} \tag{64}$$

in which K_{ci}, $K_{\omega i}$ and η_i are positive constant scalars.

Assumption 4. Assume that a_i and c_i are upper bounds of the weight matrix ω_{ai}^* and ω_{ci}^*, respectively, i.e., $\|\omega_{ai}^*\| \leq a_i$, $\|\omega_{ci}^*\| \leq c_i$. Consider the following robust term.

$$u_{rai} = \frac{(\delta_i + \beta_i)(\delta_i + \beta_i) e_{ic}}{\left\| e_{ic}^{T} (\delta_i + \beta_i) \right\|} \tag{65}$$

The leader robots will be stable by following conditions, if the controller gains are selected as $k_{i1} > \frac{1}{2}$, $k_{i3} > \frac{1}{2k_{i2}}$. Due to the boundedness of $\|e_{ij}\|$ and $\|\dot{e}_{ij}\|$, it can be shown that all the errors will be zero as $t \to \infty$. The proof of Theorem 2 is given in Appendix.

6. Simulation results

In this section, the performance of the proposed controller is considered. Formation control problem of leader–follower mobile robot using the RL controller with actor critic approach based on the Fourier series expansion is simulated as the introduced method. In addition, this method is based on the voltage control strategy. Therefore, DC motors as the actuators of mobile robot are considered. The desired formation structure of the mobile robots in the simulation is chosen as Fig. 3. This structure is organized by four robots including one leader and three followers. The leader robot must track the desired trajectory given in Fig. 4 and other robots must track the leader with the desired separation and bearing to have the desired formation of Fig. 3. Due to the sinusoidal terms in the dynamics of mobile robots, the Fourier series is utilized in the controller design to decrease the computational complexity. This simulation is done in two approaches, at the first step, the proposed controller is simulated and then to have a proper comparison, the Neural Network controller is simulated [6]. The mobile robot parameters and motor parameters are shown in Tables 1 and 2. Also, the initial position of each robot is shown in Table 3.

RL controller:

In this subsection, the RL controller by the voltage control strategy approach will be simulated. So, the controller produces the voltage of DC motors to achieve the desired formation. Therefore, the dynamics of DC motors is considered. The proposed controller is designed using actor–critic approach based on the Fourier series. In this simulation, four robots by the formation structure of Fig. 3 and first non-zero error of tracking are considered.

Fig. 4 shows the tracking performance and mobile robots in *xy* direction. Fig. 5 shows the error of formation control for all three follower robots. As it is obvious, the formation error will be bounded and converges to zero after a short time. The motor voltages as the inputs of the mobile robots are shown in Fig. 6. These voltages are bounded and smooth. The reason of high voltages in initial times of simulation is large initial errors in the tracking errors presented in Fig. 6. In order to reduce the tracking errors as fast as possible, the control gains have been selected large that results in large initial control signals. In the rest of this section, to show the superiority of the proposed method, another controller is simulated. In the second simulation, the parameters of mobile robots and their motors will be the same.

NN controller:

In this subsection, an NN controller is simulated to control the formation of the leader–follower mobile robots. To have a fare comparison with the proposed controller, the initial positions of robots are chosen the same. The formation errors of follower robots are shown in Fig. 7.

Due to the results of the above simulations, the proposed controller is superior. Table 4 compares these methods. Table 4 shows that the formation error in proposed method is much less than the NN one. Also, the computational complexity due to the use of Fourier series is very low.

7. Conclusion

In this paper, a complete model for mobile robots using DC motors is introduced. Also, an RL controller for a formation control problem is designed. Due to the sinusoidal terms in the dynamics of mobile robots, the Fourier series is utilized to design the proposed controller. The stability of the proposed controller is

Table 1
Robot parameters.

d (m)	r (m)	R (m)	I (kg^2)	m (kg)
0.1	0.02	0.1	0.1	1

Table 2
Motor parameters.

kb	km	rg	Ra	La	Bm	Jm
0.5	0.5	0.5	1	0.00021	0.001	0.002

Table 3
Initial position of mobile robots.

θ	Y	X	Positions
0	0	0	Position of leader
0	−0.86	−0.5	Position of first follower
0	−1.72	−1	Position of second follower
0	0.86	−0.5	Position of third follower

Table 4
NN performance compared with the desired controller.

Integral absolute error (IAE)	Descriptions	Formation controller
10.4116	The proposed method: RL based on Fourier series	FF-RL
15.2798	Neural networks	NN

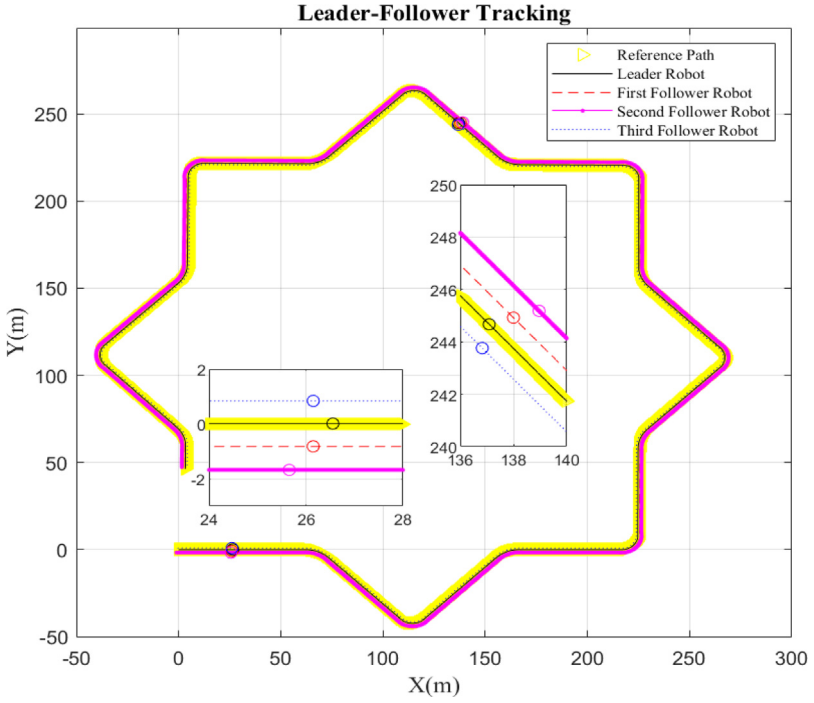

Fig. 4. The performance of the proposed controller.

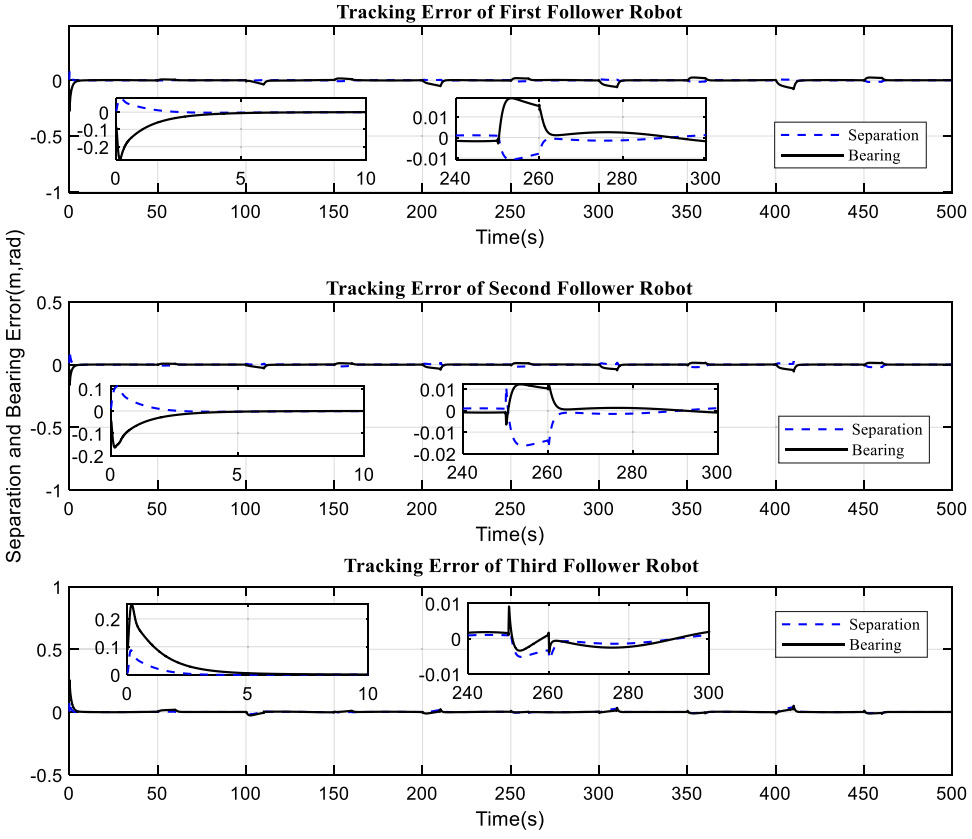

Fig. 5. Tracking error of the proposed controller.

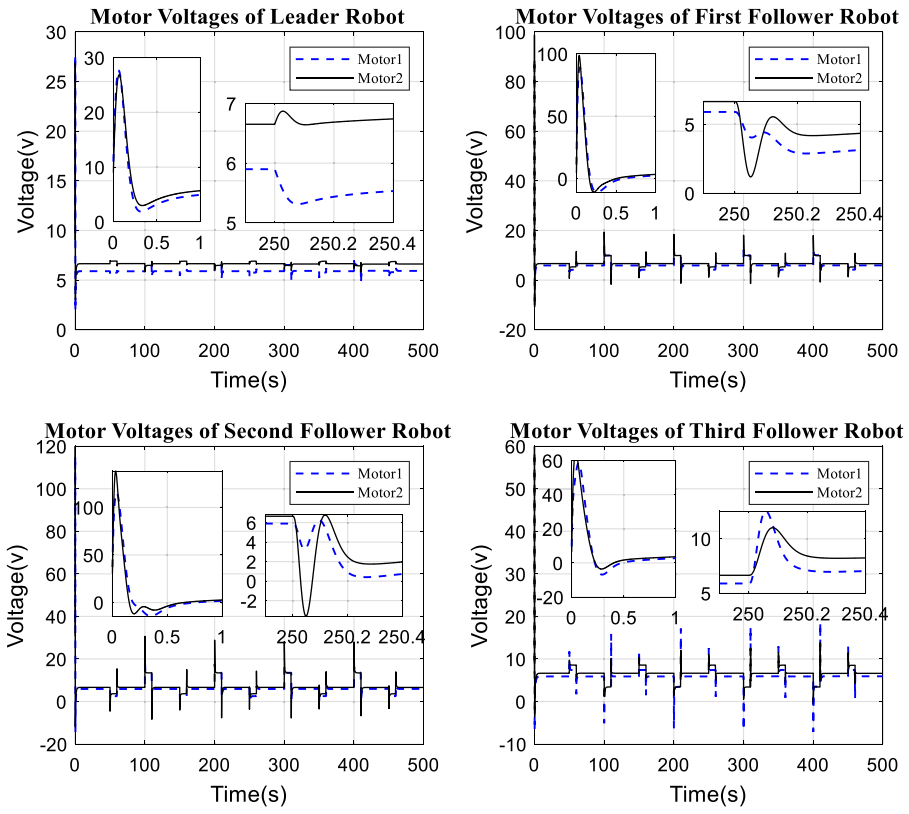

Fig. 6. Motor voltages in the proposed controller.

examined by Lyapunov theorem and it is shown that introduced controller is asymptotically stable. Based on the simulation results, it can be shown that the proposed controller has a smaller formation error in comparison with the NN controller.

Declaration of competing interest

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

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Compliance with ethical standards

The authors declare that this paper has never been submitted to other journals for simultaneous review processes and it has not been published before (partially or completely). Moreover, this paper has not been divided into different sections in order to have more submissions. The data has not been fabricated or changed in favor of the conclusions. In addition, no theory, data, or text belonging to other authors and publications was included as if it was our own and proper acknowledgment to other works has been provided.

Appendix

Proof of Theorem 1. Consider the Lyapunov function

$$V_j = k_{j4}V_{j1} + (k_{j2} + d_j(k_{j2} + k_{j3}))V_{j2} \tag{66}$$

where

$$V_{j1} = \frac{k_{j2}}{2}(e_{j1}^2 + e_{j2}^2) + \frac{d_j k_{j3}}{2}e_{j3}^2 \tag{67}$$

and

$$V_{j2} = \frac{1}{2}{e_{jc}}^T D_j e_{jc} + \frac{1}{2K_{cj}}\tilde{\omega}_{aj}^T\tilde{\omega}_{aj} + \frac{1}{2K_{\omega j}}\tilde{\omega}_{cj}^T\tilde{\omega}_{cj} \tag{68}$$

The time derivation of V_{j1} is given by

$$\dot{V}_{j1} = k_{j2}e_{j1}\dot{e}_{j1} + k_{j2}e_{j2}\dot{e}_{j2} + d_j k_{j3}e_{j3}\dot{e}_{j3} \tag{69}$$

Using (42), (69) can be written as

$$\begin{aligned}\dot{V}_{j1} = &-k_{j2}k_{j1}e_{j1}^2 - k_{j2}^2e_{j2}^2 - k_{j3}^2e_{j3}^2 - k_{j2}k_{j3}e_{j2}e_{j3} + k_{j2}e_{j1}e_{j4}\\ &+ d_j(k_{j2}e_{j2} + k_{j3}e_{j3})e_{j5}\end{aligned} \tag{70}$$

The time derivation of V_{j2} is given by

$$\dot{V}_{j2} = {e_{jc}}^T D_j\dot{e}_{jc} - \frac{1}{K_{cj}}\dot{\hat{\omega}}_{aj}^T\tilde{\omega}_{aj} - \frac{1}{K_{\omega j}}\dot{\hat{\omega}}_{cj}^T\tilde{\omega}_{cj} \tag{71}$$

Substituting (55) into (71) results in

$$\dot{V}_{j2} = -{e_{jc}}^T K_{j4}e_{jc} + {e_{jc}}^T(\varphi_j\tilde{\omega}_{aj} + \varepsilon_{aj} - u_{raj}) - \frac{1}{K_{cj}}\dot{\hat{\omega}}_{aj}^T\tilde{\omega}_{aj} - \frac{1}{K_{\omega j}}\dot{\hat{\omega}}_{cj}^T\tilde{\omega}_{cj} \tag{72}$$

Substituting (59), (60)–(61) into (72) yields in

$$\begin{aligned}\dot{V}_{j2} = &-{e_{jc}}^T K_{j4}e_{jc} + {e_{jc}}^T\varphi_j\tilde{\omega}_{aj} + {e_{jc}}^T(\varepsilon_{aj} - u_{raj}) - {e_{jc}}^T\varphi_j\tilde{\omega}_{aj}\\ &- \|e_{jc}\|\,\hat{\omega}_{cj}^T\varphi_j^T\varphi_j\tilde{\omega}_{aj}\\ &+\eta_j\|e_{jc}\|\,\hat{\omega}_{aj}^T\tilde{\omega}_{aj} + \|e_{jc}\|\,\hat{\omega}_{aj}^T\varphi_j^T\varphi_j\tilde{\omega}_{cj} + \eta_j\|e_{jc}\|\,\hat{\omega}_{cj}^T\tilde{\omega}_{cj}\end{aligned} \tag{73}$$

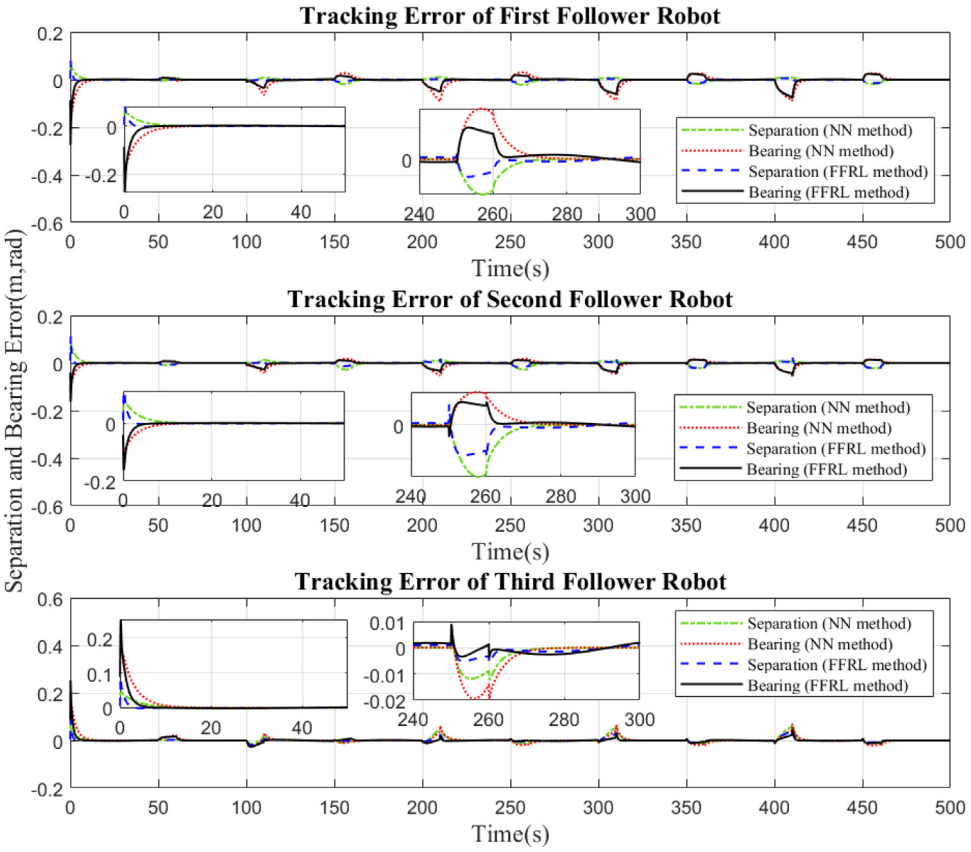

Fig. 7. Separation and bearing error of NN compared with the desired controller.

Therefore, we have

$$\begin{aligned}\dot{V}_{j2} \;=\; & -{e_{jc}}^T K_{j4} e_{jc} + {e_{jc}}^T(\varepsilon_{aj} - u_{raj}) - \left\|e_{jc}\right\| {\hat{\omega}_{cj}}^T {\varphi_j}^T \varphi_j \tilde{\omega}_{aj} \\ & + \eta_j \left\|e_{jc}\right\| {\hat{\omega}_{aj}}^T \tilde{\omega}_{aj} \\ & + \left\|e_{jc}\right\| {\hat{\omega}_{aj}}^T {\varphi_j}^T \varphi_j \tilde{\omega}_{cj} + \eta_j \left\|e_{jc}\right\| {\hat{\omega}_{cj}}^T \tilde{\omega}_{cj}\end{aligned} \tag{74}$$

In other words,

$$\begin{aligned}\dot{V}_{j2} \;\le\; & -{e_{jc}}^T K_{j4} e_{jc} + {e_{jc}}^T(\varepsilon_{aj} - u_{raj}) \\ & + \left\|e_{jc}\right\| \underbrace{\left[-{\hat{\omega}_{cj}}^T {\varphi_j}^T \varphi_j \tilde{\omega}_{aj} + \eta_j {\hat{\omega}_{aj}}^T \tilde{\omega}_{aj} + {\hat{\omega}_{aj}}^T {\varphi_j}^T \varphi_j \tilde{\omega}_{cj} + \eta_j {\hat{\omega}_{cj}}^T \tilde{\omega}_{cj}\right]}_{\alpha_j}\end{aligned} \tag{75}$$

The term α_j is a function of time. Using the definitions $\tilde{\omega}_{aj} = \omega^*_{aj} - \hat{\omega}_{aj}$ and $\tilde{\omega}_{cj} = \omega_{cj} - \hat{\omega}_{cj}$, $\left\|e_{jc}\right\| \alpha_j$ can be rewritten as

$$\begin{aligned}\left\|e_{jc}\right\| \alpha_j \;=\; & -\left\|e_{jc}\right\| {\hat{\omega}_{cj}}^T {\varphi_j}^T \varphi_j \omega^*_{aj} + \left\|e_{jc}\right\| {\hat{\omega}_{cj}}^T {\varphi_j}^T \varphi_j \hat{\omega}_{aj} \\ & + \eta_j \left\|e_{jc}\right\| {\hat{\omega}_{aj}}^T \omega^*_{aj} - \eta_j \left\|e_{jc}\right\| {\hat{\omega}_{aj}}^T \hat{\omega}_{aj} \\ & + \left\|e_{jc}\right\| {\hat{\omega}_{aj}}^T {\varphi_j}^T \varphi_j \omega^*_{cj} - \left\|e_{jc}\right\| {\hat{\omega}_{aj}}^T {\varphi_j}^T \varphi_j \hat{\omega}_{cj} \\ & + \eta_j \left\|e_{jc}\right\| {\hat{\omega}_{cj}}^T \omega^*_{cj} - \eta_j \left\|e_{jc}\right\| {\hat{\omega}_{cj}}^T \hat{\omega}_{cj}\end{aligned} \tag{76}$$

It is obvious that $-\left\|e_{jc}\right\| \eta_j {\hat{\omega}_{aj}}^T \hat{\omega}_{aj}$ and $-\left\|e_{jc}\right\| \eta_j {\hat{\omega}_{cj}}^T \hat{\omega}_{cj}$ are non-positive. Now, consider the definition:

$$\begin{aligned}\alpha'_j \;=\; & -{\hat{\omega}_{cj}}^T {\varphi_j}^T \varphi_j \omega^*_{aj} + {\hat{\omega}_{cj}}^T {\varphi_j}^T \varphi_j \hat{\omega}_{aj} + \eta_j {\hat{\omega}_{aj}}^T \omega^*_{aj} + {\hat{\omega}_{aj}}^T {\varphi_j}^T \varphi_j \omega^*_{cj} \\ & - {\hat{\omega}_{aj}}^T {\varphi_j}^T \varphi_j \hat{\omega}_{cj} + \eta_j {\hat{\omega}_{cj}}^T \omega^*_{cj}\end{aligned} \tag{77}$$

According to Assumption 3, it can be concluded

$$\begin{aligned}\left|\alpha'_j\right| \le\; & + \left\|\hat{\omega}_{cj}\right\| \left\|{\varphi_j}^T \varphi_j\right\| a_j + \left\|\hat{\omega}_{cj}\right\| \left\|{\varphi_j}^T \varphi_j\right\| \left\|\hat{\omega}_{aj}\right\| \\ & + \eta_j \left\|\hat{\omega}_{aj}\right\| a_j + \left\|\hat{\omega}_{aj}\right\| \left\|{\varphi_j}^T \varphi_j\right\| c_j \\ & \underbrace{+ \left\|\hat{\omega}_{aj}\right\| \left\|{\varphi_j}^T \varphi_j\right\| \left\|\hat{\omega}_{cj}\right\| + \eta_j \left\|\hat{\omega}_{cj}\right\| c_j}_{\beta_j}\end{aligned} \tag{78}$$

The time-varying term β_j can be simplified to

$$\begin{aligned}\beta_j = & \left\|{\varphi_j}^T \varphi_j\right\| \left(a_j \left\|\hat{\omega}_{cj}\right\| + 2 \left\|\hat{\omega}_{cj}\right\| \left\|\hat{\omega}_{aj}\right\| + c_j \left\|\hat{\omega}_{aj}\right\|\right) \\ & + \eta_j \left(a_j \left\|\hat{\omega}_{aj}\right\| + c_j \left\|\hat{\omega}_{cj}\right\|\right)\end{aligned} \tag{79}$$

By substituting (78) into (75), (75) can written as

$$\begin{aligned}\dot{V}_{j2} \le\; & -{e_{jc}}^T K_{j4} e_{jc} + e_{jc}(\varepsilon_{aj} - u_{raj}) + \left\|e_{jc}\right\| \beta_j - \left\|e_{jc}\right\| \eta_j {\hat{\omega}_{aj}}^T \hat{\omega}_{aj} \\ & - \left\|e_{jc}\right\| \eta_j {\hat{\omega}_{cj}}^T \hat{\omega}_{cj}\end{aligned} \tag{80}$$

Due to the negative-ness of $-{e_{jc}}^T K_{j4} e_{jc}$, $-\left|Y_j\right| \eta_j {\hat{\omega}_{aj}}^T \hat{\omega}_{aj}$, and $-\left|Y_j\right| \eta_j {\hat{\omega}_{cj}}^T \hat{\omega}_{cj}$, it should be verified that ${e_{jc}}^T(\varepsilon_{aj} - u_{raj}) + \left\|e_{jc}\right\| \beta_j \le 0$. In order to achieve this goal, the robust term u_{raj} is chosen as follows [24]

$$u_{raj} = \frac{(\delta_j + \beta_j)(\delta_j + \beta_j) e_{jc}}{\left\|{e_{jc}}^T(\delta_j + \beta_j)\right\|} \tag{81}$$

Based on $\left\|\varepsilon_{aj}\right\| \le \delta_j$, (86) is satisfied and we have:

$$\dot{V}_{j2} \le -{e_{jc}}^T K_{j4} e_{jc} \tag{82}$$

Due to (66), (70) and (82), and complete squares, it can be shown that:

$$\begin{aligned}\dot{V}_j \le\ & -k_{j4}k_{j2}(k_{j1}-\frac{1}{2})e_{j1}^2 - k_{j4}k_{j2}(k_{j2}-\frac{d_j}{2})e_{j2}^2 - k_{j4}k_{j3}(k_{j3}-\frac{d_j}{2})e_{j3}^2 \\ & -k_{j4}(\frac{k_{j2}}{2}+d_j(k_{j2}+k_{j3}))e_{j4}^2 - k_{j4}(k_{j2}+\frac{d_j}{2}(k_{j2}+k_{j3}))e_{j5}^2 \\ & -\frac{k_{j4}k_{j2}}{2}(e_{j1}-e_{j4})^2 - \frac{d_jk_{j4}k_{j2}}{2}(e_{j2}-e_{j5})^2 \\ & -\frac{d_jk_{j4}k_{j3}}{2}(e_{j3}-e_{j5})^2\end{aligned} \tag{83}$$

Examining (83), it can be concluded that $\dot{V}_j \le 0$ if the controller gains are selected as: $k_{j1} > \frac{1}{2}$, $k_{j2} > \frac{d_j}{2}$ and $k_{j3} > \frac{d_j}{2}$. Therefore, the position tracking errors e_j and the velocity tracking errors e_{jc} converge to zero as well as the weight gain matrix is bounded. Due to the boundedness of $\|e_j\|$, $\|e_{jc}\|$, $\|\dot{e}_j\|$, and $\|\dot{e}_{jc}\|$ for follower j, it can be shown that $\|\ddot{V}_j\| \le \infty$, therefore, using Barbalat's Lemma, $\dot{V}_j \to 0$. Thus, $e_j \to 0$ and $e_{jc} \to 0$ as $t \to \infty$. As a result, the tracking errors of position, orientation and velocity are asymptotically stable.

This completes the proof of Theorem 1.

Proof of Theorem 2. Consider the Lyapunov function

$$V'_{ij} = \sum_{1}^{N} V'_j + V_i \tag{84}$$

where V_j is defined by (66) and is as V_i below:

$$V_i = k_{i4}V_{i1} + (1 + 1/k_{i2})V_{i2} \tag{85}$$

where

$$V_{i1} = \frac{1}{2}(e_{i1}^2 + e_{i2}^2) + \frac{1}{k_{i2}}(1-\cos e_{i3}) \tag{86}$$

and

$$V_{i2} = \frac{1}{2}{e_{ic}}^T D_i e_{ic} + \frac{1}{2K_{ci}}\tilde{\omega}_{ai}{}^T\tilde{\omega}_{ai} + \frac{1}{2K_{\omega i}}\tilde{\omega}_{ci}{}^T\tilde{\omega}_{ci} \tag{87}$$

The time derivation of V_{i1} is given by

$$\dot{V}_{i1} = e_{i1}\dot{e}_{i1} + e_{i2}\dot{e}_{i2} + \frac{1}{k_{i2}}\sin e_{i3}\dot{e}_{i3} \tag{88}$$

According to (47)

$$\begin{aligned}\dot{V}_{i1} =\ & -k_{i1}e_{i1}^2 + e_{i1}e_{i4} + v_{2i}e_{i1}e_{i2} - v_{2i}e_{i1}e_{i2} + e_{i2}v_{1ir}\sin e_{i3} \\ & -k_{i3}\sin^2 e_{i3} + \frac{1}{k_{i2}}e_{i5}\sin e_{i3} - e_{i2}v_{1ir}\sin e_{i3}\end{aligned} \tag{89}$$

then

$$\dot{V}_{i1} = -k_{i1}e_{i1}^2 - k_{i3}\sin^2 e_{i3} + e_{i1}e_{i4} + \frac{1}{k_{i2}}e_{i5}\sin e_{i3} \tag{90}$$

The time derivation of V_{i2} is given by

$$\dot{V}_{i2} = {e_{ic}}^T D_i\dot{e}_{ic} - \frac{1}{K_{ci}}\dot{\tilde{\omega}}_{ai}{}^T\tilde{\omega}_{ai} - \frac{1}{K_{\omega i}}\dot{\tilde{\omega}}_{ci}{}^T\tilde{\omega}_{ci} \tag{91}$$

Applying similar steps to Eq. (71), Based on Assumption 4, $|\varepsilon_i| \le \delta_i$. Now, consider the following control term:

$$u_{rai} = \frac{(\delta_i+\beta_i)(\delta_i+\beta_i)e_{ic}}{\left\|{e_{ic}}^T(\delta_i+\beta_i)\right\|} \tag{92}$$

Thus, according to (92), we have:

$$\dot{V}_{i2} \le -{e_{ic}}^T K_{i4}e_{ic} \tag{93}$$

Due to (85), (90) and (93), and complete squares, it can be shown that:

$$\begin{aligned}\dot{V}_i \le\ & k_{i4}[-(k_{i1}-\frac{1}{2})e_{i1}^2 - (k_{i3}-\frac{1}{k_{i2}})\sin^2 e_{i3} - (\frac{1}{2}+\frac{1}{k_{i2}})e_{i4}^2 \\ & -(1+\frac{1}{2k_{i2}})e_{i5}^2 + \frac{1}{2k_{i2}}e_{i5}^2 - (\frac{1}{\sqrt{2}}e_{i1}+\frac{1}{\sqrt{2}}e_{i4})^2 \\ & -(\frac{1}{\sqrt{2k_{i2}}}e_{i5}+\frac{1}{\sqrt{2k_{i2}}}\sin e_{i3})^2]\end{aligned} \tag{94}$$

Since $-(\frac{1}{\sqrt{2}}e_{i1}+\frac{1}{\sqrt{2}}e_{i4})^2$ and $-(\frac{1}{\sqrt{2k_{i2}}}e_{i5}+\frac{1}{\sqrt{2k_{i2}}}\sin e_{i3})^2$ are non-positive, Eq. (94) can be rewritten as:

$$\begin{aligned}\dot{V}_i \le\ & k_{i4}[-(k_{i1}-\frac{1}{2})e_{i1}^2 - (k_{i3}-\frac{1}{k_{i2}})\sin^2 e_{i3} - (\frac{1}{2}+\frac{1}{k_{i2}})e_{i4}^2 \\ & -(1+\frac{1}{2k_{i2}})e_{i5}^2 + \frac{1}{2k_{i2}}e_{i5}^2]\end{aligned} \tag{95}$$

Examining (95), it can be concluded that $\dot{V}_i \le 0$ if the controller gains are selected as: $k_{i1} > \frac{1}{2}$, $k_{i3} > \frac{1}{2k_{i2}}$. Therefore, the position tracking errors e_i and the velocity tracking errors e_{ic} converge to zero as well as the weight gain matrix is bounded. Also, it can be concluded that $\dot{V}_{ij} \le 0$, so the position, orientation and velocity tracking errors for formation control problem are bounded. Due to the boundedness of $\|e_{ij}\|$ and $\|\dot{e}_{ij}\|$ for the leader i and its followers, it can be shown that $\|\dot{V}_{ij}\| \le \infty$ and $\dot{V}_j$ is uniformly continuous, therefore, using Barbalat's Lemma, $\dot{V}_{ij} \to 0$. Thus, $e_j \to 0$, $e_{jc} \to 0$, $e_{i1} \to 0$, $e_{i3} \to 0$ and $e_{ic} \to 0$ as $t \to \infty$. Using (47), it can verify that $e_{i2} \to 0$ as $t \to \infty$. Therefore, asymptotical stability is verified. This completes the proof of Theorem 2.

Remark 1. In the case when the follower j is a leader for follower $j+1$, the global asymptotic stability of this formation control problem is similar to Theorem 1. In this case, the radially unbounded Lyapunov candidate set as:

$$V''_j = \sum_{j}^{j+1} V_j \tag{96}$$

where V_j is defined by (66). Now, follower j makes a reference for follower $j+1$ in formation control problem. Due to using the same dynamics for jth follower robot (dynamics of follower j is incorporates the dynamics of leader i), so the proof of stability will be the same when the follower j is being considered as leader for follower $j+1$.

References

[1] Rubio F, Valero F, Llopis-Albert C. A review of mobile robots: Concepts, methods, theoretical framework, and applications. Int J Adv Robot Syst 2019;16(2):1729881419839596.

[2] Lewis MA, Tan KH. High precision formation control of mobile robots using virtual structures. Auton Robots 1997;4(4):387–403.

[3] Xi J, Yang X, Yu Z, Liu G. Leader–follower guaranteed-cost consensualization for high-order linear swarm systems with switching topologies. J Franklin Inst 2015;352(4):1343–63.

[4] Wang H, Guo D, Liang X, Chen W, Hu G, Leang KK. Adaptive vision-based leader–follower formation control of mobile robots. IEEE Trans Ind Electron 2017;64(4):2893–902.

[5] Peng Z, Wen G, Rahmani A, Yu Y. Distributed consensus-based formation control for multiple nonholonomic mobile robots with a specified reference trajectory. Internat J Systems Sci 2015;46(8):1447–57.

[6] Dierks T, Jagannathan S. Asymptotic adaptive neural network tracking control of nonholonomic mobile robot formations. J Intell Robot Syst 2009;56(1):153–76.

[7] Lin S, Jia R, Yue M, Xu Y. On composite leader–follower formation control for wheeled mobile robots with adaptive disturbance rejection. Appl Artif Intell 2019;33(14):1306–26.

[8] Li Y, Zhu L, Guo Y. Observer-based multivariable fixed-time formation control of mobile robots. J Syst Eng Electron 2020;31(2):403–14.

[9] Hou R, Cui L, Bu X, Yang J. Distributed formation control for multiple non-holonomic wheeled mobile robots with velocity constraint by using improved data-driven iterative learning. Appl Math Comput 2021;395:125829.
[10] Fateh MM, Arab A. Robust control of a wheeled mobile robot by voltage control strategy. Nonlinear Dynam 2015;79(1):335–48.
[11] Khodamipour G, Khorashadizadeh S, Farshad M. Observer-based adaptive control of robot manipulators using reinforcement learning and the Fourier series expansion. Trans Inst Meas Control 2021;43(10):2307–20.
[12] Cui Y, Liu X, Deng X, Wang Q. Observer-based adaptive fuzzy formation control of nonlinear multi-agent systems with nonstrict-feedback form. Int J Fuzzy Syst 2021;23(3):680–91.
[13] Li Y, Zhang J, Tong S. Fuzzy adaptive optimized leader-following formation control for second-order stochastic multi-agent systems. IEEE Trans Ind Inf 2022;18(9):6026–37.
[14] Yang S, Cao Y, Peng Z, Wen G, Guo K. Distributed formation control of nonholonomic autonomous vehicle via RBF neural network. Mech Syst Signal Process 2017;87:81–95.
[15] Zhang M, Yu X, Ding P, Ou L, Zhang W. Distributed adaptive three-dimension formation control based on improved RBF neural network for non-linear multi-agent time-delay systems. IET Control Theory Appl 2019;13(17):2758–65.
[16] Falco P, Attawia A, Saveriano M, et al. On policy learning robust to irreversible events: An application to robotic in-hand manipulation. IEEE Robot Autom Lett 2018;3(3):1482–9.
[17] Zhou Y, Lu F, Pu G, Ma X, Sun R, Chen HY, Li X. Adaptive leader-follower formation control and obstacle avoidance via deep reinforcement learning. In: 2019 IEEE/RSJ international conference on intelligent robots and systems. IROS, IEEE; 2019, p. 4273–80.
[18] Wen G, Chen CP, Feng J, Zhou N. Optimized multi-agent formation control based on an identifier–actor–critic reinforcement learning algorithm. IEEE Trans Fuzzy Syst 2017;26(5):2719–31.
[19] Tang L, Liu Y-J, Tong S. Adaptive neural control using reinforcement learning for a class of robot manipulator. Neural Comput Appl 2014;25(1):135–41.
[20] Wen G, Chen CP, Li B. Optimized formation control using simplified reinforcement learning for a class of multiagent systems with unknown dynamics. IEEE Trans Ind Electron 2019;67(9):7879–88.
[21] Izadbakhsh A. Robust control design for rigid-link flexible-joint electrically driven robot subjected to constraint: theory and experimental verification. Nonlinear Dynam 2016;85(2):751–65.
[22] Ahmadi SM, Fateh MM. Composite direct adaptive Taylor series–fuzzy controller for the robust asymptotic tracking control of flexible-joint robots. Trans Inst Meas Control 2019;41(14):4023–34.
[23] Izadbakhsh A, Khorashadizadeh S. Robust adaptive control of robot manipulators using Bernstein polynomials as universal approximator. Internat J Robust Nonlinear Control 2020;30(7):2719–35.
[24] Izadbakhsh A, Khorashadizadeh S, Ghandali S. Robust adaptive impedance control of robot manipulators using Szász–Mirakyan operator as universal approximator. ISA Trans 2020;106:1–11.
[25] Zirkohi MM. Direct adaptive function approximation techniques-based control of robot manipulators. J Dyn Syst Meas Control 2018;140(1):011006.
[26] Fierro R, Lewis FL. Control of a nonholonomic mobile robot: backstepping kinematics into dynamics. J Robot Syst 1997;13:149–63.
[27] Lewis FL, Liu K, Yesildirek A. Neural net robot controller with guaranteed tracking performance. IEEE Trans Neural Netw 1995;6(3):703–15.
[28] Khorashadizadeh S, Fateh MM. Uncertainty estimation in robust tracking control of robot manipulators using the Fourier series expansion. Robotica 2017;35(2):310–36.
[29] Fateh MM. On the voltage-based control of robot manipulators. Int J Control Autom Syst 2008;6(5):702–12.
[30] Martinelli A. The odometry error of a mobile robot with a synchronous drive system. IEEE Trans Robot Autom 2002;18(3):399–405.
[31] Tang L, Liu YJ, Tong S. Adaptive neural control using reinforcement learning for a class of robot manipulator. Neural Comput Appl 2014;25(1):135–41.
[32] Zhang X, Zhang H, Liu D, Kim Y. Neural-network-based reinforcement learning controller for nonlinear systems with non-symmetric dead-zone inputs. In: 2009 IEEE symposium on adaptive dynamic programming and reinforcement learning. IEEE; 2009, p. 124–9.