Journal of Simulation and Analysis of Novel Technologies in Mechanical Engineering
14 (1) (2022) 0027~0046
HTTP://JSME.IAUKHSH.AC.IR

Self-tuning consensus on directed graph in the case of time-varying nonhomogeneous input gains

Mozhgan Bagherboum*

Department of Mathematics, Khomeinishahr Branch, Islamic Azad University, Isfahan, Iran

*m.bagherboum@iaukhsh.ac.ir

(Manuscript Received --- 06 May 2022; Revised --- 07 June 2022; Accepted --- 21 June 2022)

Abstract

In this paper, the problem of self-tuning of coupling parameters in multi-agent systems is considered. Agent dynamics are described by a discrete-time double integrator with time-varying nonhomogeneous input gain. The coupling parameters defining the strength of agents interactions are locally self-tuning by each node based on the velocities of its neighbors. The cost function is equal to the square of the local error between the agent velocity and the weighted average of the velocities of interacting neighbors. So, the proposed algorithm is the normalized gradient algorithm which is minimized the square of the local error between the agent velocity and the one step delayed average of the velocities of its neighbors. Provided that the underlying graph is strongly connected, it is shown that the sequence of the inter-agent coupling parameters generated by the proposed algorithm is convergent. Also, assuming the suitable initial condition on coupling parameters, it is proved that the network achieves average consensus. In other words, the agent velocities converge toward the average of the initial velocities values. Furthermore, the distance among agents converges to a finite limit. Simulation results illustrate effectiveness of the proposed method.

Keywords: Multi-agent system, Unmanned aerial vehicle, Self-tuning consensus, Double integrator agent dynamic, Directed graph.

1- Introduction

Multi-agent system (MAS) or distributed artificial intelligence is an approach in which groups of agents work together to accomplish a mission or perform a task. The importance of such cooperating systems arises from their several applications in mobile robots, vehicles, unmanned aerial vehicles (UAVs), autonomous underwater vehicles (AUVs), routers, sensors, processors, ships, missiles, aircraft, and satellites [1], which are control objects with certain coordination ability. Usually, the kinds of duties assigned to MASs are those that cannot be accomplished by a single agent, or they are such that the cost and the difficulty of employing a single complex agent for performing them are greater than using groups of simpler agents [2]. In order to perform the assigned missions accurately and successfully, the agents have to communicate, cooperate and coordinate with one another, and engage in bilateral negotiations.

Some of the interesting missions and tasks can be performed by using multiple mobile robots, such as exploration, surveillance, cooperative search, mapping of different environments, distributed manipulation, and moving large objects [3]. In such missions, a group of robots may have several advantages over a single robot, including improving stability, flexibility, and reliability of the system [4], saving time, and low cost. Indeed, by using a MAS instead of a single-agent one, the following advantages are gained: (1) complicated and dangerous missions can be undertaken, (2) cost of performing a task can be reduced, (3) parallel processing can be implemented, (4) stability can be achieved and error reduced, (5) a MAS is scalable, and most importantly, (6) since no single element is responsible for thinking and decision-making in a MAS (in other words, decision-making is distributed), the system can continue functioning even if a part of it is disabled.

In the cooperative control of MASs, agents interact locally with each other to achieve the desired macroscopic objective of the swarm system or MAS. Up to now, there are many different research branches in the cooperative control of MASs which include consensus control, formation control, containment control, formation-containment control, consensus tracking control, flocking control, pursuit-evasion control, distributed filtering [5].

Consensus control problem is one of the most fundamental and important problems in cooperative control of MASs. Consensus means the interaction between groups of agents in a team, via sensors or a communication network, to reach an agreement on a common value or state. In mobile robots, this common state can be a position where the robots must meet in an autonomous manner [6]. The consensus control of continuous-time MASs with general noises and delays is studied in [7]. To achieve consensus, agents in a swarm system often interact with each other locally. For each agent, the local interaction is realized by constructing distributed controller or protocol using neighboring relative information. Information exchange is an integral part of MASs. During an exchange of information, each agent updates the state of its current information based on the information it receives from its neighbors. Before introducing the consensus algorithms and protocols, it is necessary to model each agent of a MAS by means of dynamic structures and to present a specific consensus algorithm for that agent. For this purpose, based on the type of application and the required precision considered for a consensus problem, different models such as the single-integrator model, double-integrator model and the high-order model can be used [8].

In this article, we concern on the motion equations of some vehicles defined with double-integrator dynamic models. Contrary to the consensus problems in the single-integrator dynamic model, in which all the information states converge to constant and identical values, sometimes we need to get some information states converge to one constant value and the remaining information states to another constant value; in this case, we have to use double-integrator dynamic models. The double-integrator model is one of the simplest dynamic models for omnidirectional mobile robots. Ever since the advent of control theory, the double-integrator control system has attracted a great deal of attention and is now a symbol of the controllers that achieve minimum execution time and minimum fuel consumption.

During the last two decades there has been an avalanche of papers covering various aspects of consensus problems with applications to decentralized and cooperative control, flocking and distributed formation, distributed optimization and estimation, communication and sensor networks, synchronization of coupled oscillators. Consideration has been given to a variety of topics, including convergence of various consensus protocols, consensus in the case of time varying network topologies, nonlinear system dynamics, quantization of information and noisy channels, packet drops and communication delays, randomized consensus algorithms and asynchronous algorithms.

Recent advances in MASs coordination were first reported in 1995 by Vicsek et al. [9]. In this motivational article, they proposed a minimal model that described the phase transition of self-driven particles by a novel type of dynamics. Jadbabaie et al. [10] applied graph and matrix theories to present a theoretical description of consensus in the Vicsek model. The properties of the Laplacian matrix were used to analyze the problem of consensus in first-order MASs in [11]. Also, the relationship between algebraic graph connection, convergence rate, and maximum threshold of time-delay tolerance in consensus problems was investigated. Ren and Beard [12] explored the subject of consensus in second-order MASs and highlighted the importance of the topology of interaction between agents, including the directed spanning tree, for attaining asymptotic agreement. By introducing the Laplacian matrix, consensus problems have entered the theoretical analysis phase [13] and since then, the graph theory has become an important tool for the theoretical analysis of consensus problems.

The theory of adaptive consensus deals with two distinct issues one of them is the problem of leaderless consensus where the aim is not to follow predetermined leader, but rather for all agents to achieve agreement on a common, unknown in advance value of their state. In [14], the case of known identical linear agent dynamics and undirected graph topology is analyzed, whereas in [15] identical linear and Lipschitz nonlinear dynamics on directed graphs are studied. In both articles, it is assumed that the high frequency gain (parameter multiplying agent input signal) or its sign (control direction) is known, and it is the same for all agents. They proved the interesting result that each agent state converges toward the average of its neighbors' states.

A continuous time consensus problem of second order systems governed by a directed graph is considered in [16]. The authors show that the error between any two agent positions converges to zero. They also show that in case of absolute velocity damping all velocities converge to zero, while in the case of relative velocity damping the difference between agent velocities converges to zero. In recent work by Chen et al. [17] continuous time adaptive consensus with unknown identical control directions is considered. The authors analyze an undirected graph and show that the difference between agent states tends to zero. In [18] the consensus problem of networked mechanical systems with time-varying delay and jointly connected topologies is considered. It is assumed that the high-frequency gain is known, and the inter-agent coupling term in the consensus protocol is nonadaptive with a fixed gain whose value is the same for all agents.

In 2015, Radenkovic and Tadi [19], present consensus protocol over a network of MASs with discrete time integrator dynamics. The coupling coefficients defining interaction among agents are adaptively adjusted in time. In fact, they investigate the underlying graph where the input gain of any agent was different and an unknown scalar parameter. They are shown that the agent states converge toward the average of the initial state values and all agent states asymptotically reach consensus equal to the average of initial state values. In [20], the authors consider a network of heterogeneous agents whose dynamics are described by a double integrator discrete time model with the constant input gain (independent of time) of unknown magnitude. The proposed algorithm considers for each agent to locally tune the inter-agent coupling parameter. They proved all agent velocities converge to the same value, and the distance between any two agents converges to a finite limit. The proposed algorithm was a normalized gradient recursion based on minimizing the square of the error between an individual agent’s state (velocity) and the one step delayed average of its own state and the states of its neighbors.

At present, the research on MAS is mostly based on homogeneous agents, and the research on heterogeneous agents is even less. In [21], a novel distributed adaptive controllers for leaderless synchronization in networks of identical discrete-time dynamical systems is proposed. It is proved that all agent outputs converge toward an emerging, unknown in advance, synchronization trajectory. This trajectory is not available for use by the agent’s controllers, and its pattern is determined by the initial conditions and the internal model built-in in the distributed adaptation mechanism. Compared to [17], where a synchronization trajectory is a constant, they allow arbitrary form for a synchronization trajectory. Shi et al. [22] have addressed the consensus problem of heterogeneous second-order MASs under linear and non-linear conditions. In this paper, a valid control law is given for each agent based on the communication with their neighbors, and sufficient condition is obtained to determine some parameters in controller. Zhao et al. [23] analyzed the consensus of heterogeneous MASs. They studied the average consensus of heterogeneous multi-agents, including continuous-time consensus protocol, discrete-time consensus protocol, consensus with time delay, and consensus of switching topology.

In this paper, we investigate consensus protocol where the coupling parameters defining the strength of agent's interactions are locally self-tuning. Each agent locally tunes the strength of interaction with neighboring agents by using a normalized gradient algorithm (NGA). The tuning algorithm minimizes the square of the error between an individual agent’s velocity and the one step delayed average of its own velocity and the velocities of its neighbors. Agent dynamics are described by a discrete-time double integrator with time-varying nonhomogeneous input gain. Assuming that the underlying graph is directed and strongly connected and it satisfies some additional constraints, it is proved that the sequence of coupling parameters is convergent. Also, it is shown all velocities converge toward the same constant value, where it is the average of the initial velocities. In addition, it is proved that the distance between any two agents converges to a finite limit. Finally, the validity of the theoretical method has been illustrated by numerical simulation.

Based on our previous discussion, the main contributions of this paper are three-fold, which is shown as follows.

1) The network topology is assumed to be a directed graph, whereas in [17] an undirected graph is considered.
2) The consensus problem is investigated for multi-agent networks with time-varying nonhomogeneous input gain, which is more general than the constant input gain in [17].
3) In [17] it is shown the agent velocities converge toward the same constant value, while we prove that they converge toward the average of the initial velocities values.

The remainder of the paper is arranged as follows. Problem statement is given in Section 2 and the relation of self-tuning consensus to Kuramoto synchronization is discussed. Section 3 presents the proposed algorithm and also global stability and convergence of self-tuning consensus are given in this section. A simulation example is presented in Section 4.

2- Problem formulation

It is natural to model information exchange among vehicles by directed or undirected graphs. Suppose that a team consists of N vehicles. A directed graph is a pair $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_N\}$ is a finite nonempty node set and $E \subseteq V \times V$ is an edge set of ordered pairs of nodes, called edges. The edge (v_i, v_j) in the edge set of a directed graph denotes that vehicle v_j can obtain information from vehicle v_i, but not necessarily vice versa. If an edge $(v_i, v_j) \in E$, then node v_i is a neighbor of node v_j. The set of neighbors of node v_i is denoted as $\mathcal{N}_i$.

In sequence, consider a cooperative group of N agents where the dynamics of the ith agent are described by the following discrete time system

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{2-1}$$

$$v_i(t+1) = v_i(t) + \beta_i(t)\, u_i(t) \tag{2-2}$$

where time $t \geq 0$ take on nonnegative integer values, $x_i(t) \epsilon \mathbb{R}$ and $v_i(t) \epsilon \mathbb{R}$ are the position and velocity respectively. $u_i(t) \epsilon \mathbb{R}$ is the control signal or consensus protocol of the agent, which is a function of the ith agent velocity and the velocities of its neighbors. Moreover, $\beta_i(t) \epsilon \mathbb{R}$ is an unknown input gain. The model defined by Eq. (2-2) can be viewed as a discrete time version of a kinematic model,

$$\frac{d}{d\,\tau} v_i(\tau) = \frac{1}{m_i} u_i(\tau), \qquad \tau \geq 0,$$

for $i = 1, \dots, N,$ where $v_i(\tau)$ is velocity and $u_i(\tau)$ is driving force of the ith agent respectively, where m_i is its mass. Such model can be of interest in analyzing flock behavior or considering the problem of achieving common velocity in a formation of unmanned aerial vehicles. In this case due to the fuel consumption during flight, mass m_i can be a slowly time-varying quantity. The parameter $\beta_i(t)$ in Eq. (2-2) can be interpreted as the inverse of mass m_i. The above described scenario can be encountered in problems of velocity coordination in robot formation control. As is described in [17], for robots produced in the same batch it is reasonable to assume that all units have the same dynamics and the same control directions.

It is plausible to assume that in natural phenomena such as flocks, each agent adjusts its velocity so that it is as close as possible to the average of the velocities of its neighbors, i.e. $u_i(t)$ is proportional to the local velocity mismatch defined by $\varphi_i(t) = \bar{v}_i(t) - v_i(t)$, where $\bar{v}_i(t) = \frac{1}{1+N_i}\sum_{j\in\mathcal{N}_i'} v_j(t)$, where N_i is the cardinality of $\mathcal{N}_i$, and $\mathcal{N}_i' = \mathcal{N}_i \cup \{i\}$. Thus $u_i(t) = \theta_i(t)\,\varphi_i(t)$, where $\theta_{ij}(t) \in \mathbb{R}$ are integrant coupling parameters to be determined so that all agent velocities achieve average consensus. The network theory literature often refers to $u_i(t)$ as the consensus protocol and we say that the network achieves average consensus if $v_c = \frac{1}{N}\sum_{i=1}^{N} v_i(0)$.

The aim of the average consensus protocol is to converge the values of all the nodes in a network to the mean value of data measured by the initial nodes. Average consensus has various applications. The calculation of distributed mean consensus in wireless sensor networks has been explored in [24]; and for this purpose, a completely distributed algorithm has been presented, which is able to average the data measured in the network itself.
In this paper, we consider the consensus protocol as the weighted average of the velocities of the agents as follow

$$u_i(t) = \sum_{j\epsilon\mathcal{N}_i}\theta_{ij}(t)(v_j(t) - v_i(t)) \qquad (2\text{-}3)$$

where the coupling coefficients $\theta_{ij}(t)$ defining interaction among agents are adaptively adjusted in time. Note that the control signal given in Eq. (2-3) can be written in the form

$$u_i(t) = \boldsymbol{\theta_i}(t)^{\mathrm{T}}\boldsymbol{\varphi_i}(t) \qquad (2\text{-}4)$$

where

$$\boldsymbol{\theta_i}(t)^T = [\theta_{i1}(t)I_{i1}, \dots, \theta_{iN}(t)I_{iN}] \qquad (2\text{-}5)$$

and

$$\boldsymbol{\varphi_i}(t)^T = [\varepsilon_{i1}(t), \dots, \varepsilon_{iN}(t)] \qquad (2\text{-}6)$$

where $\varepsilon_{ij}(t) = \big(v_j(t) - v_i(t)\big)I_{ij}$ and I_{ij} is an indicator function given by

$$I_{ij} = \begin{cases} 1, & j \in \mathcal{N}_i \\ 0, & j \notin \mathcal{N}_i. \end{cases}$$

After substituting Eq. (2-4) in Eq. (2-2) we arrive at the following evolution of $v_i(t)$,

$$v_i(t+1) = v_i(t) + \beta_i(t)\boldsymbol{\theta_i}(t)^T\boldsymbol{\varphi_i}(t),$$

$$i \in V. \qquad (2\text{-}7)$$

Now that we have derived model (2-7) we can relate multi-agent velocity consensus to frequency synchronization problems in linearly phase coupled oscillators described by

$$\dot{x}_i(\tau) = \Omega_i + K_i\frac{1}{1+N_i}\sum_{j\,\epsilon\,\mathcal{N}_i'}\big(x_j(\tau) - x_i(\tau)\big),$$

$$i \in V \qquad (2\text{-}8)$$

where $x_i(\tau)$ is the phase of the ith oscillator, Ω_i is its natural frequency and K_i is the coupling gain. Let frequency $\dot{x}_i(\tau)$ at time $\tau = \mathrm{t}T_s, t = 0, 1, 2, \dots,$ be approximated by

$$\frac{d}{d\tau}x_i(\tau)\bigg|_{\tau=tT_s} = \frac{x_i\big((t+1)T_s\big) - x_i(tT_s)}{T_s}$$

where T_s is the sampling interval. Then from Eq. (2-8) we can write

$$\frac{x_i\big((t+1)T_s\big) - x_i(tT_s)}{T_s} = \Omega_i +$$

$$K_i\frac{1}{1+N_i}\sum_{j\,\epsilon\,\mathcal{N}_i'}\big(x_j(tT_s) - x_i(tT_s)\big).$$

By assuming $\psi_i(t) = \frac{1}{1+N_i}\sum_{j\,\epsilon\,\mathcal{N}_i'}\big(x_j(t) - x_i(t)\big)$ and $\theta_i = K_iT_s$, this relationship will be as follows

$$v_i((t+1)T_s) = \Omega_iT_s + \theta_i\psi_i(tT_s) \qquad (2\text{-}9)$$

$v_i\big((t+1)T_s\big) = x_i\big((t+1)T_s\big) - x_i(t\,T_s),$

$v_i(0) = \Omega_i T_s.$

The initial condition $v_i(0)$ is determined so that $v_i(t) = 0$ and $x_i(t) = 0$ for all $t < 0$. Note that for the sake of simpler notation the constant T_s has been omitted in signal arguments, i.e. $x_i(tT_s) = x_i(t), t \geq 0$. Obviously we can think of $v_i(t)$ as the ''normalized frequency'' at a discrete time t. Then for any agent $i \in V$, Eq. (2-9) can be written as follow

$$v_i(t+1) = v_i(t) + \theta_i(\psi_i(t) - \psi_i(t-1))$$

$$= v_i(t) + \frac{\theta_i}{1+N_i}\sum_{j\in\mathcal{N}_i'}\big(v_j(t) - v_i(t)\big)$$

$$= v_i(t) + \boldsymbol{\theta_i}(t)^{\mathrm{T}}\boldsymbol{\varphi_i}(t), \quad v_i(0) = \Omega_i T_s$$

which is the same as consensus model (2-7) with $\boldsymbol{\theta_i}(t)^T = \left[\frac{\theta_i}{1+N_i}, \dots, \frac{\theta_i}{1+N_i}\right]$.

3- Convergence of self-tuning consensus

In this section, without using any global information, the coupling coefficients are generated by minimizing its local cost function. The cost function is equal to the square of the local error between the agent velocity and the weighted average of the velocities of interacting neighbors. Therefor the proposed algorithm is a normalized gradient recursion based on minimizing the square of the error between an agent velocity and the one step delayed average of its own velocity and the velocities of its neighbors. Assuming that the underlying graph is connected, it is shown that the sequence of the inter-agent coupling parameters generated by the local cost function is convergent and all agent velocities achieve strict-sense consensus; i.e. $\lim_{t\to\infty} v_j(t) - v_i(t) = 0,\ \forall i,j \in V$ and $\lim_{t\to\infty} v_i(t) = v_c, \forall i \in V$ for some finite v_c. In other words, each agent adjusts its velocity so that it is as close as possible to the average of the initial velocities of its neighbors; i.e. $v_c = \frac{1}{N}\sum_{i=1}^{N} v_i(0)$. Furthermore, the distance between any two agents converges to a finite limit.

As previously stated, we consider a double integrator multi-agent as follow

$$x_i(t+1) = x_i(t) + v_i(t+1)$$

$$v_i(t+1) = v_i(t) + \beta_i(t)\,\boldsymbol{\theta_i}(t)^{\mathrm{T}}\boldsymbol{\varphi_i}(t). \quad (3\text{-}1)$$

It can be written in the compact form as

$$\mathbf{v}(\mathrm{t}+1) = \mathbf{W}(\mathrm{t})\mathbf{v}(\mathrm{t}) \quad (3\text{-}2)$$

where $\mathbf{v}(\mathrm{t})$ is given by $\mathbf{v}(\mathrm{t}) = [v_1(\mathrm{t}), \dots, v_{\mathrm{N}}(\mathrm{t})]$ and $\mathbf{W}(\mathrm{t})$ is a $N \times N$ matrix defined as follows

$$\mathbf{W}(\mathrm{t}) = \big[w_{\mathrm{ij}}(\mathrm{t})\big], \quad (3\text{-}3)$$

$$w_{ij}(t) = \begin{cases} \beta_i(t)\theta_{ij}(t), & j \in \mathcal{N}_i \\ 1 - \beta_i(t)\sum_{j\in\mathcal{N}_i}\theta_{ij}(t), & i = j \\ 0, & otherwise. \end{cases}$$

It is well-known that the condition for the iteration in Eq. (3-2) to converge to the average of initial velocities is $\boldsymbol{l}^T = [1, \dots, 1]$ to be left and right eigenvector of $\mathbf{W}(\mathrm{t})$ corresponding to the eigenvalue $\lambda_1 = 1$, i.e. $\boldsymbol{l}^T\mathbf{W}(\mathrm{t}) = \boldsymbol{l}^T, \mathbf{W}(\mathrm{t})\boldsymbol{l} = \boldsymbol{l}$. Then, the sum of the velocities is time invariant, i.e. $\boldsymbol{l}^{\boldsymbol{T}}\mathbf{V}(\mathrm{t}+1) = \boldsymbol{l}^{\boldsymbol{T}}\mathbf{V}(\mathrm{t}) = \cdots = \boldsymbol{l}^{\boldsymbol{T}}\mathbf{V}(0)$, and $\boldsymbol{l}$ (or scalar multiple of it) is a fixed point of the recursion defined by Eq. (3-2).

Agent $i\epsilon\, V$ tunes coupling parameter $\boldsymbol{\theta_i}(t)$ so that the following local cost function is minimized.

$$\boldsymbol{J_i}(\boldsymbol{\theta_i}) = \frac{1}{2}(v_i(t+1) - \bar{\bar{v}}_i(t+1))^2,$$

$$i \in V \quad (3\text{-}4)$$

where $\bar{\bar{v}}_i(t+1)$ represents the one step delayed weighted average of the ith agents neighbors' velocities, including its own velocity, i.e.

$$\bar{\bar{v}}_i(t+1) = \sum_{j=1}^{N} m_{ij} v_i(t) \qquad (3\text{-}5)$$

where

$$m_{ij} = \begin{cases} \dfrac{1-\alpha_i}{1+N_i}, & 0 \le \alpha_i < 1, j \in \mathcal{N}_i \\ \alpha_i + \dfrac{1-\alpha_i}{1+N_i}, & j = i \\ 0, & otherwise. \end{cases}$$

It is obvious from Eq. (3-5) that when $\alpha_i = 0$, the average $\bar{\bar{v}}_i(t+1)$ becomes equal to the uniformly weighted average. In addition when calculating $\bar{\bar{v}}_i(t+1)$ agent i assigns more weight to its value than to any individual neighbor. By virtue of the fact that from Eqs. (3-1) (3-4) and (3-5) the gradient of $\boldsymbol{J}_i(\boldsymbol{\theta}_i)$ with respect to $\boldsymbol{\theta}_i(t)$ is

$$\frac{\partial \boldsymbol{J}_i(\boldsymbol{\theta}_i)}{\partial \boldsymbol{\theta}_i} = \left(v_i(t+1) - \bar{\bar{v}}_i(t+1)\right) \beta_i(t)\, \boldsymbol{\varphi}_i(t).$$

Since from Eq. (2-7), $\frac{\partial v_i(t+1)}{\partial \boldsymbol{\theta}_i(t)} = \beta_i(t)\boldsymbol{\varphi}_i(t)$ gradient based minimization of the cost function given in Eq. (3-4) suggests the following updating rule for $\boldsymbol{\theta}_i(t)$ and $\boldsymbol{\theta}_i(t)$ can be tuned by the following recursive procedure

$$\boldsymbol{\theta}_i(t+1) = \boldsymbol{\theta}_i(t) - \beta_i(t)\, \boldsymbol{\varphi}_i(t)\, (v_i(t+1) - \bar{\bar{v}}_i(t+1)$$

However, since $\beta_i(t)$ is unknown, instead of the previous equation, agent i can use the following normalized gradient algorithm

$$\boldsymbol{\theta}_i(t+1) = \boldsymbol{\theta}_i(t) - \frac{\mu_i}{r_i(t)}\, sgn\left(\beta_i(t)\right) \boldsymbol{\varphi}_i(t)\, e_i(t+1) \qquad (3\text{-}6)$$

where it is assumed that $sgn(\beta_i(t))$, the sign of $\beta_i(t)$ is known, $\mu_i > 0$ is the algorithm step size, $\boldsymbol{\varphi}_i(t)$ is the local velocity mismatch defined by (2-6), $e_i(t+1)$ is the cost function error given by

$$e_i(\mathrm{t}+1) = v_i(t+1) - \bar{\bar{v}}_i(t+1) \qquad (3\text{-}7)$$

with $\bar{\bar{v}}_i(t+1)$ defined by (3-5), and $r_i(t)$ being the gradient normalizer given by

$$r_i(t) = 1 + \|\boldsymbol{\varphi}_i(t)\|^2, \qquad i\epsilon V.$$

Recursion (3-6) starts with some finite initial $\boldsymbol{\theta}_i(0)$.
In sequel, we consider the following vectors

$$\boldsymbol{e}(t)^T = [e_1(t), \dots, e_N(t)]$$

$$\boldsymbol{\Phi}_i(\mathrm{t}) = \mathbf{v}(\mathrm{t}) - v_i(t)\boldsymbol{l} \qquad (3\text{-}8)$$

and we show that sequences $\{e_i(t)\}_{\mathrm{t}\ge 0}$ and $\{\boldsymbol{\varphi}_i(t)\}_{\mathrm{t}\ge 0}$ generated by the algorithm (3-6) and (3-7) along with system dynamics (2-7) have finite total energies for all finite initial conditions $x_i(0)$, $\boldsymbol{\theta}_i(0)$, $v_i(0), i \in V$.

Lemma 1. Let the underlying directed graph G is strongly connected. Then for all $i = 1, \dots, N$, $\forall\, n \ge 0$,

$$\sum_{t=0}^{n} \|\boldsymbol{\Phi}_i(t+1)\|^2 \le c \sum_{t=0}^{n} \|\boldsymbol{e}(t+1)\|^2,$$

where c is positive constant.

Proof. From (3-5) and (3-7) we can write

$$\mathbf{v}(t+1) = \mathbf{M}\, \mathbf{v}(t) + \boldsymbol{e}(t+1)$$

where $\mathbf{M}$ is the $N \times N$ matrix given by $\mathbf{M} = [m_{ij}]$ with elements m_{ij} defined by (3-5). By using the fact that $\sum_{j=1}^{N} m_{ij} = 1$ for all $i\epsilon V$, we conclude that $\lambda_1 = 1$ is an eigenvalue of $\mathbf{M}$ with the corresponding right eigenvector $\boldsymbol{l}^T = [1, \dots, 1]$, i.e. $\mathbf{M}\boldsymbol{l} = \boldsymbol{l}$. Since $\mathbf{M}$ is a stochastic matrix, $\lambda_1 = 1$ is its maximal eigenvalue [25, p. 527]. By virtue of the fact that G is a connected graph, the nonnegative matrix $\mathbf{M}$ is irreducible implying that $\lambda_1 = 1$ is an algebraically simple eigenvalue of $\mathbf{M}$ [25, Theorems 6.2.24, p. 362 and 8.4.4, p. 508]. Furthermore, if the nonnegative matrix is irreducible and any main diagonal element is positive, such matrix must be primitive [25, Theorems 8.5.2, p. 516 and 8.5.10, p. 520]. Consequently, except $\lambda_1 = 1$, the rest of the eigenvalues λ_i of the matrix $\mathbf{M}$ satisfy the condition $|\lambda_i| < 1, i = 2, \dots, N$. Let $\boldsymbol{y}_1$ be the left eigenvector of $\mathbf{M}$ associated to $\lambda_1 = 1$, and normalized so that $\boldsymbol{y}_1^T \boldsymbol{l} = 1$.

Based on this discussion we can decompose the matrix $\mathbf{M}$ as follows

$$\mathbf{M} = \mathbf{M}_1 + \boldsymbol{l}\boldsymbol{y}_1^T, \quad \boldsymbol{l}^T \mathbf{y}_1 = 1 \tag{3-9}$$

where all eigenvalues of $\mathbf{M}_1$ are inside the unit circle, i.e. the spectral radius of $\mathbf{M}_1$ satisfies $\rho(\mathbf{M}_1) < 1$. Hence from Eq. (3-9), one can obtain

$$\mathbf{v}(t+1) = \mathbf{M}_1 \mathbf{v}(t) + \boldsymbol{l}\boldsymbol{y}_1^T \mathbf{v}(t) + \boldsymbol{e}(t+1).$$

Define

$$\mathbf{H}(q^{-1}) = (I - q^{-1}\mathbf{M}_1)^{-1}$$

where q^{-1} is the unit delay operator. The fact $\rho(\mathbf{M}_1) < 1$ implies that $\mathbf{H}(q^{-1})$ is a stable operator. Since $\mathbf{M}_1 \boldsymbol{l} = 0$ we have $\mathbf{H}(q^{-1})\boldsymbol{l} = \boldsymbol{l}$. After multiplying both sides of the last equation with $\mathbf{H}(q^{-1})$ we can write

$$\mathbf{H}(q^{-1})\mathbf{v}(t+1) = \mathbf{H}(q^{-1})\mathbf{M}_1 \mathbf{v}(t) + \boldsymbol{l}\boldsymbol{y}_1^T \mathbf{v}(t) + \mathbf{H}(q^{-1})\boldsymbol{e}(t+1).$$

But

$$\mathbf{H}(q^{-1})\mathbf{v}(t+1) = \mathbf{v}(t+1) + \sum_{k=1}^{\infty} q^{-k+1}\mathbf{M}_1^k \mathbf{v}(t) = \mathbf{v}(t+1) + \mathbf{H}(q^{-1})\mathbf{M}_1 \mathbf{v}(t)$$

then the previous equation gives

$$\mathbf{v}(t+1) = \boldsymbol{l}\boldsymbol{y}_1^T \mathbf{v}(t) + \mathbf{H}(q^{-1})\boldsymbol{e}(t+1). \tag{3-10}$$

Now using the fact that $\boldsymbol{y}_1^T \boldsymbol{l} = 1$ we can derive from Eq. (3-8)

$$\boldsymbol{\Phi}_i(t+1) = \mathbf{v}(t+1) - v_i(t+1)\boldsymbol{l} = \\ = \boldsymbol{l}\boldsymbol{y}_1^T[\mathbf{v}(t) - v_i(t)\boldsymbol{l}] \\ + \boldsymbol{l}[v_i(t) - v_i(t+1)] + \mathbf{H}(q^{-1})\boldsymbol{e}(t+1).$$

Note that given by Eq. (3-5) can be written in the form

$$\bar{\bar{v}}_i(t+1) = \sum_{j=1}^{N} m_{ij} v_j(t) \\ = v_i(t) + \boldsymbol{a}_i^T \boldsymbol{\Phi}_i(t) \tag{3-11}$$

where $\boldsymbol{a}_i^T$ is defined as follows

$$\boldsymbol{a}_i^T = [a_{i1} \dots . a_{iN}],$$

$$a_{ij} = \begin{cases} \frac{1-\alpha_i}{1+N_i}, & j \in \mathcal{N}_i, 0 \le \alpha_i \le 1 \\ 0, & \text{otherwise.} \end{cases} \tag{3-12}$$

Then the error $e_i(t+1)$ given by Eq. (3-7) can be expressed as

$$e_i(t+1) = v_i(t+1) - \bar{\bar{v}}_i(t+1) \\ = v_i(t+1) - v_i(t) - \boldsymbol{a}_i^T \boldsymbol{\Phi}_i(t).$$

From Eqs. (3-8) and (3-10) we can derive

$$\boldsymbol{\Phi}_i(t+1) = \boldsymbol{l}\boldsymbol{y}_1^T \boldsymbol{\Phi}_i(t) + \boldsymbol{l}\left(-\boldsymbol{a}_i^T \boldsymbol{\Phi}_i(t) - e_i(t+1)\right) + \mathbf{H}(q^{-1})\boldsymbol{e}(t+1) \\ = \boldsymbol{l}(\boldsymbol{y}_1^T - \boldsymbol{a}_i^T)\boldsymbol{\Phi}_i(t) - e_i(t+1)\boldsymbol{l} + \mathbf{H}(q^{-1})\boldsymbol{e}(t+1).$$

Define the following matrix

$$\mathbf{Q}_i = \boldsymbol{l}(\boldsymbol{y}_1 - \boldsymbol{a}_i)^T, \ i \in \mathrm{V}.$$

Then one can obtain

$$\boldsymbol{\Phi}_i(t+1) = \mathbf{Q}_i \boldsymbol{\Phi}_i(t) - e_i(t+1)\boldsymbol{l} \\ + \mathbf{H}(q^{-1})\boldsymbol{e}(t+1).$$

With assuming $\mathbf{L}_i(q^{-1}) = (I - q^{-1}\mathbf{Q}_i)^{-1}$, we have

$$\mathbf{L}_i(q^{-1})\mathbf{\Phi}_i(t+1) = \mathbf{L}_i(q^{-1})\mathbf{Q}_i\mathbf{\Phi}_i(t) - \mathbf{L}_i(q^{-1})e_i(t+1)\boldsymbol{l} + \mathbf{L}_i(q^{-1})\boldsymbol{H}(q^{-1})\boldsymbol{e}(t+1).$$

Again, q^{-1} is the unit delay operator, according to Eq. (3-10), we have

$$\mathbf{L}_i(q^{-1})\mathbf{\Phi}_i(t+1) = \mathbf{\Phi}_i(t+1) + \sum_{k=1}^{\infty} q^{-k+1}\mathbf{Q}_i^k\mathbf{\Phi}_i(t) = \mathbf{\Phi}_i(t+1) + \mathbf{L}_i(q^{-1})\mathbf{Q}_i\mathbf{\Phi}_i(t).$$

Hence

$$\mathbf{\Phi}_i(t+1) = -\mathbf{L}_i(q^{-1})e_i(t+1)\boldsymbol{l} + \mathbf{L}_i(q^{-1})\mathbf{H}(q^{-1})\boldsymbol{e}(t+1) = \mathbf{L}_i(q^{-1})[\mathbf{H}(q^{-1})\boldsymbol{e}(t+1) - e_i(t+1)\boldsymbol{l}]. \quad (3\text{-}13)$$

Note that the matrix $\mathbf{Q}_i$ is of rank 1 and its only nonzero eigenvalue is $\rho_1 = \boldsymbol{l}^T(\boldsymbol{y}_1 - \boldsymbol{a}_i)$. Since $\boldsymbol{l}^T\mathbf{y}_1 = 1$ and from (3-12) $\frac{1+\alpha_i N_i}{1+N_i} < 1$, we have

$$\rho_1 = \boldsymbol{l}^T(\boldsymbol{y}_1 - \boldsymbol{a}_i) = 1 - \boldsymbol{l}^T\boldsymbol{a}_i = 1 - \sum_{j=1}^{N} a_{ij} = 1 - \frac{(1-\alpha_i)N_i}{1+N_i} = \frac{1+\alpha_i N_i}{1+N_i} < 1,$$

where we have used the fact that $\alpha_i < 1$. Hence, all eigenvalues of $\mathbf{Q}_i$ are inside the unit circle and consequently $\mathbf{L}_i(q^{-1})$ is a stable operator. By virtue of the fact that $\boldsymbol{H}(q^{-1})$ is a stable operator, $\|\mathbf{H}(q^{-1})\| \le 1$ and $\|\mathbf{L}_i(q^{-1})\| \le 1$. Then from Eq. (3-13) we have

$$\|\mathbf{\Phi}_i(t+1)\| \le \|\mathbf{L}_i(q^{-1})\|[\|\mathbf{H}(q^{-1})\|\|\boldsymbol{e}(t+1)\| + \|e_i(t+1)\boldsymbol{l}\|] \le \left[\|\mathbf{H}(q^{-1})\|\|\boldsymbol{e}(t+1)\| + \sqrt{N}\|\boldsymbol{e}(t+1)\|\right] \le \left(1+\sqrt{N}\right)\|\boldsymbol{e}(t+1)\| = c_1\|\boldsymbol{e}(t+1)\|$$

where $e_i(t+1)$ is absorbed by $\boldsymbol{e}(t+1)$. Summing up both sides of above inequality from $t = 0$ to $t = n$ yields

$$\sum_{t=0}^{n}\|\mathbf{\Phi}_i(t+1)\|^2 \le c_1^2\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2.$$

Thus the lemma is proved.

In the following, we prove that the sequence of coupling parameters is convergent.

Lemma 2. Let the underlying directed graph G is strongly connected. Also, let the sign of β_i in Eq. (2-2) and the upper bound β_{imax} of $|\beta_i(t)|$ are known to agent *i* and the lower bound β_{imin} of $|\beta_i(t)|$ is nonzero. The step size μ_i in Eq. (3-6) satisfies $\mu_i < \frac{2}{\beta_{imax}}$, for all $1 \le i \le N$. Then there exist some positive constants c_2, c_3 and c_4 so that for all initial conditions $x_i(0)$, $v_i(0)$ and $\theta_{ij}(0)$ $(1 \le i, j \le N)$ the MAS given by Eq. (2-7) and the parameter estimation algorithm defined by (3-6) provide

(1) $\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2 \le c_2 < \infty, \quad \forall n \ge 0$

(2) $\sum_{t=0}^{n}\|\mathbf{\Phi}_i(t)\|^2 \le c_3 < \infty, \ \forall n \ge 0,$

$\forall\, i \in \mathrm{V}$

(3) $\sum_{t=0}^{n}(v_i(t) - v_j(t))^2 \le c_4 < \infty,$

$\forall n \ge 0, \quad \forall i, j \in \mathrm{V}$

(4) $\lim_{t\to\infty}\boldsymbol{\theta}_i(t) = \bar{\boldsymbol{\theta}}_i, \quad i \in \mathrm{V},$

for some finite $\bar{\boldsymbol{\theta}}_i$.

Proof. In view of Eq. (2-6), Eq. (3-11) can be written in the form

$$\bar{\bar{v}}_i(t+1) = v_i(t) + \boldsymbol{a}_i^T\boldsymbol{\varphi}_i(t).$$

Note that according to Eqs. (2-7) and (3-7) we have

$$e_i(t+1) = v_i(t+1) - \bar{\bar{v}}_i(t+1)$$
$$= v_i(t) + \beta_i(t)\boldsymbol{\theta}_i^T(t)\boldsymbol{\varphi}_i(t) - \left(v_i(t) + \boldsymbol{a}_i^T\boldsymbol{\varphi}_i(t)\right)$$
$$= (\beta_i(t)\boldsymbol{\theta}_i(t) - \boldsymbol{a}_i)^T\boldsymbol{\varphi}_i(t).$$

Now, by defining

$$\widetilde{\boldsymbol{\theta}}_i(t) = \beta_i(t)\boldsymbol{\theta}_i(t) - \boldsymbol{a}_i \qquad (3\text{-}14)$$

the parameter error $e_i(t+1)$ can be expressed in terms of $\widetilde{\boldsymbol{\theta}}_i(t)$ as follows

$$e_i(t+1) = \widetilde{\boldsymbol{\theta}}_i^T(t)\boldsymbol{\varphi}_i(t) \qquad (3\text{-}15)$$

where the parameter errors $\widetilde{\boldsymbol{\theta}}_i$ are in effect governed by the recursion given by Eq. (3-6). According to the definitions given by Eqs. (2-5) and (3-14) and from Eq. (3-6) the following recursion can be obtained for the components $\widetilde{\boldsymbol{\theta}}_i$

$$\tilde{\theta}_{ij}(t+1) =$$
$$= \tilde{\theta}_{ij}(t) - \frac{\mu_i}{r_i(t)}|\beta_i(t)|\varepsilon_{ij}(t)e_i(t+1) \qquad (3\text{-}16)$$

Now, we define $L(t) = \sum_{i=1}^{N}\left\|\widetilde{\boldsymbol{\theta}}_i(t)\right\|^2$, where $\left\|\widetilde{\boldsymbol{\theta}}_i(t)\right\|^2 = \sum_{j=1}^{N}\tilde{\theta}_{ij}^2(t)I_{ij}$. Clearly $L(n) \geq 0$, for all $n \in \mathbb{N}$. Also, in view of the recursion relation (3-16), the sequence $\{L(n)\}$ is non-increasing. So $L(n) \leq L(0)$, which shows $\{L(n)\}$ is a bounded sequence. Moreover, Eq. (3-16) implies that

$$L(t+1) \leq L(t) - 2\sum_{i=1}^{N}\frac{\mu_i}{r_i(t)}|\beta_i(t)|e_i(t+1)\sum_{j=1}^{N}\tilde{\theta}_{ij}(t)\, I_{ij}\, \varepsilon_{ij}(t) + \sum_{i=1}^{N}\frac{\mu_i^2}{r_i^2(t)}\beta_i^2(t)e_i^2(t+1)\sum_{j=1}^{N}\varepsilon_{ij}^2(t)I_{ij}.$$

Also, relations Eqs. (2-5), (2-6) and (3-15) imply that

$$e_i(t+1) = \widetilde{\boldsymbol{\theta}}_i(t)^T\boldsymbol{\varphi}_i(t)$$
$$= \textstyle\sum_{j=1}^{N}\tilde{\theta}_{ij}(t)\varepsilon_{ij}(t)I_{ij} \qquad (3\text{-}17)$$

By using $\sum_{j=1}^{N}\varepsilon_{ij}^2(t) = \|\boldsymbol{\varphi}_i(t)\|^2$ it follows that

$$L(t+1) \leq L(t) - 2\sum_{i=1}^{N}\frac{\mu_i}{r_i(t)}|\beta_i(t)|e_i^2(t+1) + \sum_{i=1}^{N}\frac{\mu_i^2}{r_i(t)}\beta_i^2(t)\frac{\|\boldsymbol{\varphi}_i(t)\|^2}{r_i(t)}e_i^2(t+1)$$

Define $r_i(t) = 1 + \|\boldsymbol{\varphi}_i(t)\|^2$, we can obtain

$$L(t+1) \leq L(t) - 2\sum_{i=1}^{N}\frac{\mu_i}{r_i(t)}|\beta_i(t)|e_i^2(t+1) + \sum_{i=1}^{N}\frac{\mu_i^2}{r_i(t)}\beta_i^2(t)e_i^2(t+1)$$

or

$$L(t+1) \leq L(t) - 2\sum_{i=1}^{N}\mu_i\,|\beta_i(t)|\left(1 - \frac{\mu_i\,|\beta_i(t)|}{2}\right)\frac{e_i^2(t+1)}{r_i(t)}.$$

Summing up both sides of the previous equation from $t = 0$ to $t = n$ gives

$$L(n+1) \leq L(0) - 2\sum_{t=0}^{n}\sum_{i=1}^{N}\mu_i|\beta_i(t)|(1 - \frac{\mu_i\,|\beta_i(t)|}{2})\frac{e_i^2(t+1)}{r_i(t)}.$$

Since by assumption $\mu_i < \frac{2}{\beta_{imax}}$ and the step size μ_i satisfies $1 - \frac{\mu_i\,|\beta_i(t)|}{2} > 0$, we have

$$2\min_{1\leq i\leq N}\left\{\mu_i\,\beta_{imin}\left(1 - \frac{\mu_i\beta_{imax}}{2}\right)\right\}\sum_{t=0}^{n}\sum_{i=1}^{N}\frac{e_i^2(t+1)}{r_i(t)}$$
$$\leq 2\sum_{t=0}^{n}\sum_{i=1}^{N}\mu_i|\beta_i(t)|\left(1 - \frac{|\beta_i(t)|\mu_i}{2}\right)\frac{e_i^2(t+1)}{r_i(t)}$$

$$\leq L(0) - L(n+1).$$

Hence

$$\sum_{t=0}^{n}\sum_{i=1}^{N}\frac{e_i^2(t+1)}{r_i(t)}$$

$$\leq \frac{L(0) - L(n+1)}{2\min\limits_{1\leq i\leq N}\left\{\mu_i\,\beta_{imin}(1-\frac{\mu_i\beta_{imax}}{2})\right\}} = k_1 < \infty$$

for some positive constant k_1. Define $\bar{r}(t) = 1 + \sum_{i=1}^{N}\sum_{k=0}^{t}\|\boldsymbol{\varphi}_i(k)\|^2$. Since $\bar{r}(t) \geq r_i(t)$ for $1 \leq i \leq N$ and $\|\boldsymbol{e}(t+1)\|^2 = \sum_{i=1}^{N} e_i^2(t+1)$, the previous relation implies

$$\sum_{t=0}^{n}\frac{\|\boldsymbol{e}(t+1)\|^2}{\bar{r}(t)} \leq \sum_{t=0}^{n}\sum_{i=1}^{N}\frac{e_i^2(t+1)}{r_i(t)}$$

$$\leq k_1 < \infty. \quad (3\text{-}18)$$

Then $\sum_{t=0}^{n}\frac{\|\boldsymbol{e}(t+1)\|^2}{\bar{r}(t)}$ is convergent.

Since from Eqs. (2-6) and (3-8), $\|\boldsymbol{\varphi}_i(t)\| \leq \|\boldsymbol{\Phi}_i(t)\|, \forall t \geq 0$, we have $r_i(t) \leq 1 + \|\boldsymbol{\Phi}_i(t)\|^2$. Then Lemma 3-1 implies that for some positive constant c,

$$\bar{r}(n) \leq 1 + \sum_{i=1}^{N}\sum_{t=0}^{n}\|\boldsymbol{\Phi}_i(t)\|^2$$

$$\leq 1 + cN\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2. \quad (3\text{-}19)$$

If $\lim\limits_{n\to\infty}\bar{r}(n) = \infty$, by Kronecker's Lemma [26, p. 503], one can derive $\lim\limits_{n\to\infty}\frac{1}{\bar{r}(n)}\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2 = 0$. Hence there exists N_1 such that $\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2 \leq \frac{\bar{r}(n)}{2cN}$, for $n \geq N_1$. Then Eq. (3-19) follow $\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2 \leq \frac{1}{2cN} + \frac{1}{2}\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2$ or $\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2 \leq \frac{1}{cN}$, for $n \geq N_1$, which gives $\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2 \leq k_2$, for all n.

On the other hand, $\{\bar{r}(n)\}$ is non-increasing. Thus if $\lim\limits_{n\to\infty}\bar{r}(n) = \bar{r}$, then $\bar{r}(n) \leq \bar{r}$ and from Eq. (3-18) we can derive $\frac{\sum_{t=0}^{n}\|\boldsymbol{e}(t+1)\|^2}{\bar{r}} \leq \sum_{t=0}^{n}\frac{\|\boldsymbol{e}(t+1)\|^2}{\bar{r}(t)} \leq k_1 < \infty$, and this proves Statement (1). Statement (2) is a consequence of Statement (1) and Lemma 1. Observe that Lemma 1 and Statement (1) yield $\sum_{t=0}^{n}\|\varphi_i(t)\|^2 < \infty,\ \forall\, n \geq 0,\ i \in V$. This relation together with (3-8) implies Statement (3). Note that Statement (3) is stronger than wide-sense consensus and it implies

$$\lim_{t\to\infty} v_j(t) - v_i(t) = 0, \forall i,j \in V.$$

We now prove that the parameter sequence $\{\theta_{ij}(t)\}_{t\geq 0}, t \geq 0, 1 \leq i,j \leq N,$ has a limit. From Eq. (3-16), we have

$$\tilde{\theta}_{ij}(k+1)I_{ij} = \tilde{\theta}_{ij}(k)I_{ij}$$

$$-\frac{\mu_i}{r_i(k)}|\beta_i(k)|\,\varepsilon_{ij}(k)e_i(k+1)$$

After summing both sides of the previous equation from $k = 0$ to $k = n$, it follows that

$$\tilde{\theta}_{ij}(t+1)I_{ij} = \tilde{\theta}_{ij}(0)\,I_{ij} \quad (3\text{-}20)$$

$$-\sum_{k=0}^{t}\frac{\mu_i}{r_i(k)}|\beta_i(k)|\,\varepsilon_{ij}(k)e_i(k+1).$$

Consider now an infinite series R_{ij} defined by

$$R_{ij}(t) = \sum_{k=0}^{t}\frac{\beta_i(k)\,\varepsilon_{ij}(k)\,e_i(k+1)}{r_i(k)}.$$

Since by Eqs. (2-6) and (3-8), $|\varepsilon_{ij}(k)| \leq \|\boldsymbol{\varphi}_i(k)\| \leq \|\boldsymbol{\Phi}_i(k)\|$, we have

$$\sum_{k=0}^{t}\left|\frac{\beta_i(k)\,\varepsilon_{ij}(k)\,e_i(k+1)}{r_i(k)}\right|$$

$$\leq \beta_{imax}\sum_{k=0}^{t}\|\boldsymbol{\Phi}_i(k)\||e_i(k+1)|$$

where we used the fact that $r_i(k) \geq 1$. Then from the recent equation and Cauchy-Schwartz's inequality, it follows that

$$\sum_{k=0}^{t}\left|\frac{\beta_i(k)\,\varepsilon_{ij}(k)\,e_i(k+1)}{r_i(k)}\right|$$

$$\leq \beta_{imax}\left(\sum_{k=0}^{t}\|\boldsymbol{\Phi}_i(k)\|^2\right)^{\frac{1}{2}}\left(\sum_{k=0}^{t}|e_i(k+1)|^2\right)^{\frac{1}{2}} \leq k_3 < \infty$$

for $1 \leq i, j \leq N$. Thus the infinite series R_{ij} is absolutely convergent. Hence, Eq. (3-20) implies that

$$\lim_{t\to\infty}\tilde{\theta}_{ij}(t+1)I_{ij} = \tilde{\theta}_{ij}(0)I_{ij} - \mu_i \lim_{t\to\infty} R_{ij}(t).$$

Hence $\lim_{t\to\infty}\tilde{\theta}_{ij}(t+1)$ or $\lim_{t\to\infty}\tilde{\theta}_{ij}(t)$ exists, where it follows that $\lim_{t\to\infty}\boldsymbol{\theta}_i(t)$ exists. It proves Statement (4) of the theorem.

Remark. The stability analysis in case of the unknown control directions can be a fairly complex task algebraically. For the sake of presenting clear and easy to follow proofs, we investigated our analysis by considering the case of known sign of the parameter $\beta_i(t)$ in (2.2). Practically, if multiple agents such as robots are produced in batches, then it is reasonable to assume that all agents have similar models and the same control directions. Therefore, it is reasonable that the control directions of all agents (i.e., signs of) are the same. It should be mentioned that the gradient-based protocol presented in (3-6) cannot handle the case of the unknown sign of parameter $\beta_i(t)$.

We now turn to Eqs. (3-2) and (3-3). Observe that from Eq. (3-3), Lemma 3-2 and boundedness $|\beta_i(t)|$, it follows $\limsup_{t\to\infty}\mathbf{W}(t) = \bar{\mathbf{W}}$, where

$$\bar{\mathbf{W}} = [\bar{w}_{ij}],$$

$$\bar{w}_{ij} = \begin{cases} \bar{\beta}_i\,\bar{\theta}_{ij}, & j \in \mathcal{N}_i \\ 1 - \sum_{j\in\mathcal{N}_i}\bar{\beta}_i\,\bar{\theta}_{ij}, & j = i \\ 0, & \text{otherwise.} \end{cases}$$

Hence from Eq. (3-2) we can write $\mathbf{v}(t+1) = \bar{\mathbf{W}}\,\mathbf{v}(t) + \widetilde{\mathbf{W}}(t)\mathbf{v}(t)$, with $\widetilde{\mathbf{W}}(t) = \mathbf{W}(t) - \bar{\mathbf{W}}$, and thus $\limsup_{t\to\infty}\widetilde{\mathbf{W}}(t) = 0$. In sequence, we prove $\mathbf{W}(t)$ and $\bar{\mathbf{W}}$ are nonnegative matrices provided that the following constraint is fulfilled on the initial conditions $\boldsymbol{\theta}_i(0)$, $i \in \mathrm{V}$.

Assumption 3-3. In (3-6), $\theta_{ij}(0)$ is selected so that $0.5 < \alpha_i < 1$ and

$$0 < \theta_{ij}(0)\,sgn(\beta_i(0)) < \frac{(2N_i+3)(1-\alpha_i)}{2(1+N_i)^2 \max\limits_{1\leq i\leq N}\{|\beta_i(0)|\}}, \qquad j \in \mathcal{N}_i,$$

$$0 < \theta_{ij}(0)\,sgn(\beta_i(0)) < \frac{(1-\alpha_i)}{2(1+N_i)\sqrt{N-N_i}\max\limits_{1\leq i\leq N}\{|\beta_i(0)|\}}, \qquad j \notin \mathcal{N}_i.$$

Recall that α_i is a parameter defining the weight of $v_i(t)$ in the average $\bar{\bar{v}}_i(t)$ given by Eq. (3-5). After squaring up both sides of Eq. (3-16) and summing up from $j = 1$ to $j = N$, it follows that

$$\sum_{j=1}^{N}\tilde{\theta}_{ij}^2(t+1) = \sum_{j=1}^{N}\tilde{\theta}_{ij}^2(t) - \frac{2\mu_i}{r_i(t)}|\beta_i(t)|\,e_i(t+1)\sum_{j=1}^{N}\tilde{\theta}_{ij}(t)\varepsilon_{ij}(t) + \frac{\mu_i^2}{r_i^2(t)}\beta_i^2(t)\,e_i^2(t+1)\sum_{j=1}^{N}\varepsilon_{ij}^2(t).$$

Then from Eqs. (2-5), (2-6) and (3-17) we derive

$$\|\tilde{\boldsymbol{\theta}}_i(t+1)\|^2 = \|\tilde{\boldsymbol{\theta}}_i(t)\|^2 - \frac{2\mu_i}{r_i(t)}|\beta_i(t)|e_i^2(t+1) + \frac{\mu_i^2}{r_i^2(t)}\beta_i^2(t)\,\|\boldsymbol{\varphi}_i(t)\|^2 e_i^2(t+1)$$

$$\leq \left\|\tilde{\boldsymbol{\theta}}_i(t)\right\|^2 - \frac{2\mu_i}{r_i(t)}|\beta_i(t)|e_i^2(t+1)$$

$$+\frac{{\mu_i}^2}{r_i(t)}\beta_i^2(t)\, e_i^2(t+1)$$

$$= \left\|\tilde{\boldsymbol{\theta}}_i(t)\right\|^2 - \frac{2\mu_i}{r_i(t)}|\beta_i(t)|\left(1-\frac{\mu_i|\beta_i(t)|}{2}\right)e_i^2(t+1)$$

Hence, in view of $\mu_i < \frac{2}{\beta_{imax}}$, we obtain $\left\|\tilde{\boldsymbol{\theta}}_i(t+1)\right\|^2 \leq \left\|\tilde{\boldsymbol{\theta}}_i(t)\right\|^2$. Therefore $\left\|\tilde{\boldsymbol{\theta}}_i(t)\right\|^2 \leq \left\|\tilde{\boldsymbol{\theta}}_i(0)\right\|^2$. On the other hands, according to Assumption 3-3,

$$0 < \theta_{ij}(0)\, sgn\big(\beta_i(0)\big) < \frac{(2N_i+3)(1-\alpha_i)}{2(1+N_i)^2 \max\limits_{1\leq i\leq N}\{|\beta_i(0)|\}},\ j \in \mathcal{N}_i.$$

Then $0 < \theta_{ij}(0)\beta_i(0) < \frac{(2N_i+3)(1-\alpha_i)}{2(1+N_i)^2}$ or $\tilde{\theta}_{ij}(0) = \beta_i(0)\theta_{ij}(0) - \frac{1-\alpha_i}{1+N_i} < \frac{(1-\alpha_i)}{2(1+N_i)^2}$,

$j \in \mathcal{N}_i$.

Also, $0 < \theta_{ij}(0)\, sgn\big(\beta_i(0)\big) <$

$$\frac{(1-\alpha_i)}{2(1+N_i)\sqrt{N-N_i}\max\limits_{1\leq i\leq N}\{|\beta_i(0)|\}},\ j \notin \mathcal{N}_i.$$

Then

$0 < \beta_i(0)\theta_{ij}(0) - 0 < \frac{(1-\alpha_i)}{2(1+N_i)\sqrt{N-N_i}}$ or

$0 < \tilde{\theta}_{ij}(0) < \frac{(1-\alpha_i)}{2(1+N_i)\sqrt{N-N_i}},\ j \notin \mathcal{N}_i$.

It follows that

$$\tilde{\theta}_{ij}^2(t) \leq \left\|\tilde{\boldsymbol{\theta}}_i(t)\right\|^2 \leq \left\|\tilde{\boldsymbol{\theta}}_i(0)\right\|^2 = \sum_{j=1}^{N}\tilde{\theta}_{ij}^2(0)$$

$$= \sum_{j\in\mathcal{N}_i}\left(\tilde{\theta}_{ij}(0)\right)^2 + \sum_{j\notin\mathcal{N}_i}\left(\tilde{\theta}_{ij}(0)\right)^2$$

$$\leq \sum_{j\in\mathcal{N}_i}\frac{(1-\alpha_i)^2}{4(1+N_i)^4} + \sum_{j\notin\mathcal{N}_i}\frac{(1-\alpha_i)^2}{4(1+N_i)^2(N-N_i)}$$

$$\leq \frac{(1-\alpha_i)^2}{4(1+N_i)^2} + \frac{(1-\alpha_i)^2}{4(1+N_i)^2} < \frac{(1-\alpha_i)^2}{(1+N_i)^2}.$$

for all $i \in V$, $j \in \mathcal{N}_i$. Hence

$\tilde{\theta}_{ij}^2(t) < \frac{(1-\alpha_i)^2}{(1+N_i)^2}, j \in \mathcal{N}_i$, or

$\left(\beta_i(t)\theta_{ij}(t) - \frac{1-\alpha_i}{1+N_i}\right)^2 < \frac{(1-\alpha_i)^2}{(1+N_i)^2}$.

Therefore $\beta_i^2(t)\theta_{ij}^2(t) - \frac{2(1-\alpha_i)}{1+N_i}\beta_i(t)\theta_{ij}(t) + \frac{(1-\alpha_i)^2}{(1+N_i)^2} < \frac{(1-\alpha_i)^2}{(1+N_i)^2}$ or $0 < \beta_i^2(t)\theta_{ij}^2(t) < \frac{2(1-\alpha_i)}{1+N_i}\beta_i(t)\,\theta_{ij}(t)$, for all $j \in \mathcal{N}_i$. It implies that $\beta_i(t)\theta_{ij}(t) > 0$ and $\beta_i(t)\theta_{ij}(t) < \frac{2(1-\alpha_i)}{1+N_i}$.

On the other hands,

$$1 - \sum_{j\in\mathcal{N}_i}\beta_i(t)\theta_{ij}(t) > 1 - \sum_{j\in\mathcal{N}_i}\frac{2(1-\alpha_i)}{1+N_i}$$

$$= 1 - \frac{2N_i(1-\alpha_i)}{1+N_i} > 2\alpha_i - 1 > 0.$$

Thus, all matrix components of $\mathbf{W}$(t) are nonnegative. i.e. $w_{ij}(t) \geq 0$. Also $\bar{\beta}_i\bar{\theta}_{ij} \geq 0$ and $\bar{\beta}_i\bar{\theta}_{ij} \leq \frac{2(1-\alpha_i)}{1+N_i},\ 1 \leq i \leq N$. Thus $\overline{\mathbf{W}}$ is a nonnegative matrix. Since by construction it is a row stochastic matrix, $\lambda_1 = 1$ is its maximal eigenvalue and $\boldsymbol{l}^T = [1,\ldots,1]$ is its right eigenvector [25, p. 527]. By the fact that the corresponding graph is strongly connected, $\overline{\mathbf{W}}$ is an irreducible matrix [25, Theorem 6.2.24, p. 362]. Then by the Perron–Frobenius theorem for nonnegative matrices, $\lambda_1 = 1$ is an algebraically simple eigenvalue [25, Theorem 8.4.4, p. 508]. Since $\bar{w}_{ii} = 1 - \sum_{j\epsilon\mathcal{N}_i}\bar{\beta}_i\bar{\theta}_{ij} > 0$ it follows that $\overline{\mathbf{W}}$ is a primitive matrix, i.e. it has only one eigenvalue of maximum modulus [25, Theorems 8.5.2, p. 516 and 8.5.10, p. 520].

It is obvious that $\boldsymbol{l}^T = [1,\ldots,1]$ is the right eigenvector of $\overline{\mathbf{W}}$ corresponding to $\lambda_1 = 1$. Let $\boldsymbol{y}_2$ be the left eigenvector associated to the eigenvalue $\lambda_1 = 1$ and normalized so that $\boldsymbol{l}^T\boldsymbol{y}_2 = 1$. Based on the above discussion matrix $\bar{\mathbf{W}}$ can be decomposed into $\bar{\mathbf{W}} = \mathbf{W}_1 + \boldsymbol{l}^T \boldsymbol{y}_2$, where

$$\mathbf{W}_1 \boldsymbol{l} = 0, \quad \mathbf{W}_1^T \boldsymbol{y}_2 = 0, \quad \text{and} \quad \rho(\mathbf{W}_1) < 1 \tag{3-21}$$

with $\rho(\mathbf{W}_1)$ being the spectral radius of $\mathbf{W}_1$. In the following, we show that the MAS achieves consensus and all agent velocities converge to the average of initial velocity values. In addition, the distance between any two members of the group converges toward a finite limit.

Theorem 3-4. Let the underlying directed graph G is strongly connected. Also, let the sign of β_i in Eq. (2-2) and the upper bound β_{imax} of $|\beta_i(t)|$ are known to agent i and the lower bound β_{imin} of $|\beta_i(t)|$ is nonzero. The step size μ_i in Eq. (3-6) satisfies $\mu_i < \frac{2}{\beta_{imax}}$, for all $1 \le i \le n$. Also, let Assumptions (3-3) hold. Then

(1) $\lim_{t\to\infty} v_i(t) = \frac{\sum_{i=1}^{N} v_i(0)}{N}, \quad 1 \le i \le N$

(2) $\lim_{t\to\infty}(x_j(t) - x_i(t)) = \bar{x}_{ij}, \ \left|\bar{x}_{ij}\right| < \infty$

for all $1 \le i, j \le N$.

Proof. Let

$$\boldsymbol{z}(t+1) = \mathbf{v}(t+1) - \boldsymbol{l}\boldsymbol{y}_2^T \mathbf{v}(t) \tag{3-22}$$

where $\boldsymbol{y}_2$ is the left eigenvector of $\bar{\mathbf{W}}$ corresponding to $\lambda_1 = 1$. We first show that $\|\boldsymbol{z}(t+1)\| \le c\rho^t$, for some $0 < \rho < 1$ and $0 < c < \infty$. To prove this claim, after substituting $\bar{\mathbf{W}} = \mathbf{W}_1 + \boldsymbol{l}^T \boldsymbol{y}_2$ in $\mathbf{v}(t+1) = \bar{\mathbf{W}}\,\mathbf{v}(t) + \widetilde{\mathbf{W}}(t)\mathbf{v}(t)$ we obtain $\mathbf{v}(t+1) = \mathbf{W}_1\mathbf{v}(t) + \boldsymbol{l}\boldsymbol{y}_2^T\mathbf{v}(t) + \widetilde{\mathbf{W}}(t)\mathbf{v}(t)$. Since

$\mathbf{W}_1\boldsymbol{l} = 0$ and $\widetilde{\mathbf{W}}(t)\boldsymbol{l} = (\mathbf{W}(t) - \bar{\mathbf{W}})\boldsymbol{l} = 0$, it follows that

$$\begin{aligned}\boldsymbol{z}(t+1) &= \mathbf{W}_1\mathbf{v}(t) + \widetilde{\mathbf{W}}(t)\mathbf{v}(t)\\ &= \mathbf{W}_1\left(\boldsymbol{z}(t) + \boldsymbol{l}\boldsymbol{y}_2^T\mathbf{v}(t-1)\right)\\ &\quad + \widetilde{\mathbf{W}}(t)\left(\boldsymbol{z}(t) + \boldsymbol{l}\boldsymbol{y}_2^T\mathbf{v}(t-1)\right)\\ &= \left(\mathbf{W}_1 + \widetilde{\mathbf{W}}(t)\right)\boldsymbol{z}(t).\end{aligned}$$

Then $\boldsymbol{z}(t+1) = \prod_{k=1}^{t}(\mathbf{W}_1 + \widetilde{\mathbf{W}}\,(k))\boldsymbol{z}(1)$. By assumption $\boldsymbol{T}(t+1, n) = \prod_{k=n}^{t}(\mathbf{W}_1 + \widetilde{\mathbf{W}}(k))$, the previous equation can be written $\boldsymbol{z}(t+1) = \boldsymbol{T}(t+1,1)\boldsymbol{z}(1)$. By using [27, Lemma A.2.13, p. 310], there exits $0 < \rho < 1$ such that $\|\boldsymbol{T}(t+1, n)\| \le c\,\rho^{t+1-n}$, for $t \ge n$. Hence

$$\|\boldsymbol{z}(t+1)\| = \|\boldsymbol{T}(t+1,1)\|\|\boldsymbol{z}(1)\|$$

$$\le c\rho^t\,\|\boldsymbol{z}(1)\| = c'\rho^t \tag{3-23}$$

for $t > 0$. From (3-22)

$$\mathbf{v}(t+1) = \boldsymbol{z}(t+1) + \boldsymbol{l}\boldsymbol{y}_2^T\mathbf{v}(t)$$

$$= \boldsymbol{z}(t+1) + \boldsymbol{l}\boldsymbol{y}_2^T\big(\boldsymbol{z}(t) + \boldsymbol{l}\boldsymbol{y}_2^T\mathbf{v}(t-1)\big)$$

$$= (\boldsymbol{l}\boldsymbol{y}_2^T)^{\mathbf{2}}\,\mathbf{v}(t-1) + \sum_{k=t-1}^{t}(\boldsymbol{l}\boldsymbol{y}_2^T)^{\boldsymbol{t-k}}\,\boldsymbol{z}(k+1)$$

$$= \cdots = (\boldsymbol{l}\boldsymbol{y}_2^T)^{\boldsymbol{t+1}}\,\mathbf{v}(0) + \sum_{k=0}^{t}(\boldsymbol{l}\boldsymbol{y}_2^T)^{\boldsymbol{t-k}}\,\boldsymbol{z}(k+1)$$

where $\boldsymbol{P} = \boldsymbol{l}\boldsymbol{y}_2^T$ is an idempotent matrix, i.e. $\boldsymbol{P}^k = \boldsymbol{P}$, $k \ge 1$. Then, the above relation can be written as follows

$$\mathbf{v}(t+1) = \boldsymbol{P}\,\mathbf{v}(0) + \boldsymbol{P}\sum_{k=0}^{t-1}\boldsymbol{z}(k+1) + \boldsymbol{z}(t+1).$$

Since by Eq. (3-23) $\sum_{k=0}^{\infty}\boldsymbol{z}(k+1)$ is an absolutely convergent series, the recent equation implies that $\lim_{t\to\infty}\mathbf{v}(t)$ exists. Furthermore, Eq. (3-2) implies that $\boldsymbol{l}^T\mathbf{v}(t+1) = \boldsymbol{l}^T\mathbf{v}(t) = \cdots = \boldsymbol{l}^T\mathbf{v}(0)$.

On the other hands, from Eq. (3-8) together with Statement (2) of Lemma 3-2, we can derive $\lim_{t\to\infty}\mathbf{v}(t) - v_i(t)\boldsymbol{l} = 0$ or

$\lim_{t\to\infty} \boldsymbol{l}^T \mathbf{v}(t) - v_i(t)\, \boldsymbol{l}^T \boldsymbol{l} = 0$. Since $\boldsymbol{l}^T\, \boldsymbol{l} = N$ it follows that

$$\lim_{t\to\infty} v_i(t) = \frac{\boldsymbol{l}^T \mathbf{v}(0)}{N} = \frac{\sum_{i=1}^{N} v_i(0)}{N},$$

$i = 1, 2, \ldots, N$. This proves Statement (1).

Next we show the validity of Statement (2). From Eq. (2-1) it follows that

$$x_i(t+1) - x_j(t+1) = x_i(0) - x_j(0) + \sum_{k=1}^{t+1} \left(v_i(k) - v_j(k)\right), \quad \forall i, j \in V.$$

Obviously we need to demonstrate that the partial sum on the right hand side of the previous equation is convergent. From Eq. (3-22) we can write $z_i(t+1) = v_i(t+1) - \boldsymbol{y}_2^T v(t), i \in V$, where $z_i(t)$ is the ith component of the vector $z(t)$. Hence $v_i(t+1) - v_j(t+1) = z_i(t+1) - z_j(t+1)$, for all $i, j \in V$. Then from Eq. (3-23) one obtains

$$\left|v_i(t+1) - v_j(t+1)\right|$$
$$\leq |z_i(t+1)| + \left|z_j(t+1)\right|$$
$$\leq \|\boldsymbol{z}(t+1)\| + \|\boldsymbol{z}(t+1)\| \leq 2c'\rho^t.$$

Then $\sum_{k=1}^{\infty}(v_i(k) - v_j(k))$ is an absolutely convergent series and it implies that $\lim_{t\longrightarrow\infty} x_i(t) - x_j(t)$ exists. Thus the theorem is proved.

4- Simulation experiment

Consider a network of seven agents characterized by a directed graph whose topology is defined by the following adjacency matrix

$$A_d = \begin{bmatrix} 0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0 \\ 0&0&0&1&0&0&1 \\ 0&0&0&0&0&1&0 \\ 1&0&0&0&0&0&1 \\ 0&1&0&0&0&0&0 \\ 1&1&0&1&0&1&0 \end{bmatrix}$$

$A_d(i,j) = 1$ signifies that agent i directly receives information from agent j. $A_d(i,j) = 0$ means that agent i cannot receive any information from agent j. Let $\beta(t) = [\beta_1(t) \ldots \beta_N(t)]$ be a vector with $\beta_i(t)$ being parameters from Eq. (3-1) and they are defined as follow

$$\beta_1(t) = 1 + 0.6\ \sin\left(\frac{\pi}{100}t\right),$$

$$\beta_2(t) = -1 - 0.4\ \cos\left(\frac{\pi}{100}t\right),$$

$$\beta_3(t) = 1 - 0.5\ \cos\left(\frac{\pi}{100}t\right),$$

$$\beta_4(t) = -0.45 + 0.3 \sin\left(\frac{\pi}{100}t\right),$$

$$\beta_5(t) = 0.8 + 0.5\ \cos\left(\frac{\pi}{100}t\right),$$

$$\beta_6(t) = 0.70 + 0.6\ \cos\left(\frac{\pi}{100}t\right),$$

$$\beta_7(t) = 0.2 + 0.1 \sin\left(\frac{\pi}{100}t\right).$$

Initial states of the model (3-1) and (3-2) are selected as $x_i(0) = 0.5 \times (-2)^{i+1}$ and $v_i(0) = i(-1)^i, i = 1, 2, \ldots, 7$. In Eq. (3-6) the algorithm step size is set to $\mu_i = 0.95, i = 1, 2, \ldots, 7$. Fig. 1 shows that all velocities $v_i(t), i = 1, 2, \ldots, 7$, converge to the same value (-0.5714), where is equal to the average of the initial velocities.

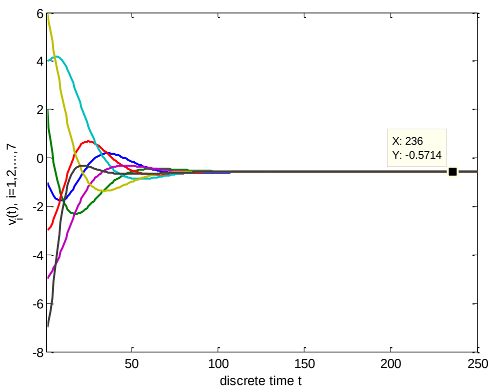

Fig. 1 Convergence of agent velocities $v_i(t)$.

Fig. 2 illustrates that the distance between the third and the sixth agent converges to a constant and all agents move in its own direction with the same distance.

Fig. 3 shows that the coupling parameters are convergent, where they are locally self-tuned by using normalized gradient algorithm.

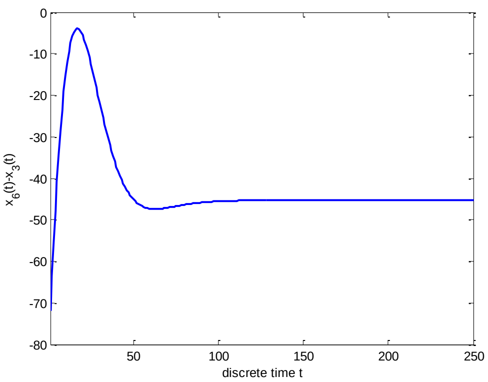

Fig. 2 Evolution of distance between the 3rd and 6th agent.

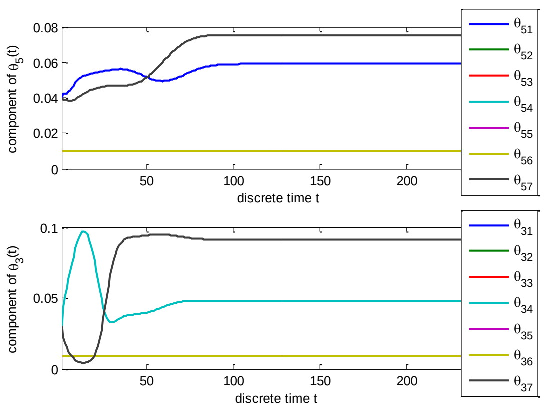

Fig. 3 Evolution of the parameter vector $\theta_3(t)$ and $\theta_5(t)$.

Fig. 4 depicts the consensus protocol that it causes the convergence of velocity and state and it converges to zero.

5- Conclusion

Emergence of a synchronized collective behavior in MASs is a topic of significant interest in various fields of science and engineering. Focal point of study in multi-agent coordination is understanding the consensus phenomenon where a number of autonomous agents reach a state of agreement. This intriguing phenomenon of collective behavior is observed in natural and manmade systems in biology, chemistry, physics and engineering, as well as in the art and social contexts.

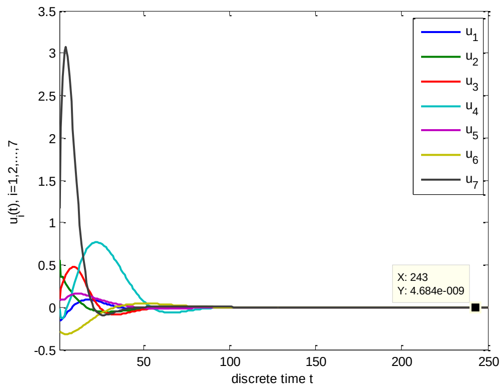

Fig. 4 The convergence of the control signal of the agents.

In this paper we considered a multi agent networked system where the coupling parameters are locally tuned. Each node locally tunes its coupling parameters by using normalized gradient algorithm (NGA) recursion. Provided that the network graph is strongly connected, it is shown that the coupling parameter sequence converges. Also, under additional constraints specified by Assumption 3-3, it is proved that all agent velocities converge toward the same constant value, and this value, in contrast to [17], is the average of initial velocities; i.e. the network achieves average consensus. In addition, it is shown that the distance between any two members of the group converges. Finally, theoretically derived conclusions are confirmed by simulation results and figures.

In spite of the extensive research works on consensus, still there are many unresolved challenges and issues. On the one hand, most of the research efforts in this field are focused on the theoretical studies and the results of the developments and innovations in this area are mainly verified through simulation and less attention is paid to hardware implementation. While in the actual implementation of a MAS, different factors such as presence of noise and delay in transmission channels and information exchange between agents, communication interferences, and the external disturbances will adversely affect the system and disrupt its operation. On the other hand, because of the high complexity of some subjects, either they have been totally ignored by the researchers or seldom explored even in the theoretical studies. There seems to be a deep void in the studies related to the mentioned research themes and especially the high-order nonlinear nonhomogeneous networks with the switching topology and in the presence of delay and noise. Hopefully, more researchers will attempt to study the stated subjects in their future works, and we will witness the actual implementation of consensus in the real-world applications.

References

[1] Lewis, M. A., & Tan, K. H. (1997). High precision formation control of mobile robots using virtual structures, *Autonomous Robots*, 4(4), 387-403.

[2] Zhu, J., Lu, J., & Yu, X. (2012). Flocking of multi-agent non-holonomic systems with proximity graphs. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(1),199-210.

[3] Das, A. K., Fierro, R., Kumar, V., Ostrowski, J. P., Spletzer, J., & Taylor, C. J. (2002). A vision-based formation control framework. *IEEE Transactions on Robotics and Automation*, 18(5), 813-825.

[4] Yamchi, M. H., & Esfanjani, R. M. (2017). Formation control of networked mobile robots with guaranteed obstacle and collision avoidance. *Robotica*, 35(6), 1365-1377.

[5] Xiwang, D. (2016). Formation Control of Swarm Systems. *Formation and Containment Control for High-order Linear Swarm Systems*. Springer, Berlin, Heidelberg, 53-103.

[6] Qiao, W., & Sipahi, R. (2016). Consensus control under communication delay in a three robot system: design and experiments. *IEEE transactions on control systems technology*, 24, 687-694.

[7] Zong, X., Zhang, J. F., & Yin, G. (2021). Stochastic Consensus Control of Multi-agent Systems under General Noises and Delays. In *Developments in Advanced Control and Intelligent Automation for Complex Systems*, Springer, Cham, 225-254.

[8] Hu, G. (2012). Robust consensus tracking of a class of second-order multi-agent dynamic systems. *Systems & Control Letters*, 61(1), 134-142.

[9] Vicsek, T., Czirok, A., Jacob, E. B., Cohen, I., & Schochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75, 1226-1229.

[10] Jadbabaie, A., Lin, J., & Morse, S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, l48, 988-1001.

[11] Olfati-Saber, R., & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9), 1520-1533.

[12] Ren, W., & Beard, R. W. (2005). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE transactions on automatic control*, 50(5), 655-661.

[13] Li, Y., & Tan, C. (2019). A survey of the consensus for multi-agent systems. *Systems Science & Control Engineering*, 7(1), 468-482.

[14] Li, Z., Ren, W., Liu, X., & Fu, M. (2013). Consensus of multi-agent systems with general linear and Lipschitz nonlinear dynamics using distributed adaptive protocols. *IEEE Transactions on Automatic Control*, 58(7), 1786-1791.

[15] Li, Z., Ren, W., Liu, X., & Xie, L. (2013). Distributed consensus of linear multi-agent systems with adaptive dynamic protocols. *Automatica*, 49, 1986-1995.

[16] Mei, J., Ren, W., & Chen, J. (2014). Consensus of second order heterogeneous multi-agent system under a directed graph. *Proceeding of the 2014 American Control Conference*, Portland, Oregon, 802-807.

[17] Chen, W., Li, X., Ren, W., & Wen, C. (2014). Adaptive consensus of multi-agent systems with unknown identical control directions based on a novel Nussbaum-type function. *IEEE Transactions on Automatic Control*, 59 (7), 1887-1892.

[18] Liu, Y., Min, H., Liu, Z. & Liao, S. (2014). Distributed adaptive consensus for multiple mechanical systems with switching topologies and time-varying delay. *Systems & Control Letters*, 64, 119-126.

[19] Radenkovic, M. S., & Tadi, M. (2015). Self-tuning average consensus in complex networks. *Journal of the Franklin Institute*, 352, 1152-1168.

[20] Radenkovic, M. S., & Golkowski, M. (2015). Distributed self-tuning consensus and synchronization. *System & Control Letters*, 76, 66-73.

[21] Radenkovic M. S., & Krstic, M. (2018). Distributed adaptive consensus and synchronization in complex networks of dynamical systems. *Automatica*, 91, 233-243.

[22] Shi, Y., Yin, Y., Liu, C., Liu, C., Liu, F. (2018). Consensus for heterogeneous multi-agent systems with second-order linear and nonlinear dynamics. *2018 Chinese Control And Decision Conference (CCDC). IEEE*, 6068-6071.

[23] Zhao, J., Dai, F., & Song, Y. (2022). Consensus of Heterogeneous Mixed-Order Multi-agent Systems Including UGV and UAV. *Proceedings of 2021 Chinese Intelligent Systems Conference*, Springer, Singapore, 202-210.

[24] Cardei, M., & Wu, J. (2004). Coverage in wireless sensor networks. *Handbook of Sensor Networks*, M. Ilyas, Ed. CRC Press, West Palm Beach, FL.

[25] Horn, R., & Johnson, C. (1985). *Matrix analysis*. Cambridge University Press, Cambridge, UK.

[26] Goodwin, G., & Sin, K. S. (1985). *Adaptive filtering, prediction and control*. Prentice Hall, New Jersey.

[27] Mareels I., & Polderman, J. W. (1996). *Adaptive systems: An introduction*. Birkhäuser.