J Optim Theory Appl (2018) 177:222–253
https://doi.org/10.1007/s10957-018-1268-7

Game-Theory-Based Consensus Learning of Double-Integrator Agents in the Presence of Worst-Case Adversaries

Kyriakos G. Vamvoudakis[1] · **João P. Hespanha**[2]

Received: 3 May 2017 / Accepted: 21 March 2018 / Published online: 29 March 2018

Abstract This work proposes a game-theory-based technique for guaranteeing consensus in unreliable networks by satisfying local objectives. This multi-agent problem is addressed under a distributed framework, in which every agent has to find the best controller against a worst-case adversary so that agreement is reached among the agents in the networked team. The construction of such controllers requires the solution of a system of coupled partial differential equations, which is typically not feasible. The algorithm proposed uses instead three approximators for each agent: one to approximate the value function, one to approximate the control law, and a third one to approximate a worst-case adversary. The tuning laws for every controller and adversary are driven by their neighboring controllers and adversaries, respectively, and neither the controller nor the adversary knows each other's policies. A Lyapunov stability proof ensures that all the signals remain bounded and consensus is asymptotically reached. Simulation results are provided to demonstrate the efficacy of the proposed approach.

Keywords Game theory · Consensus · Hamilton–Jacobi equations · Optimization · Security

✉ Kyriakos G. Vamvoudakis
kyriakos@vt.edu

João P. Hespanha
hespanha@ece.ucsb.edu

[1] Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA, USA

[2] Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA, USA

1 Introduction

In recent years, complex dynamic systems consisting of interacting subsystems have received a significant attention. Such systems arise, for example, in vehicle formation and maneuvering, power systems [1,2], altitude alignment [3], rendezvous problem [4], flocking [5], and consensus seeking and agreement [6]. Large-scale distributed systems are vulnerable to adversaries since traditional consensus algorithms [7–9] typically exhibit fragilities with respect to attacks. There is thus a need for architectures that are resilient against attacks.

Machine learning is an attractive approach to achieving optimal behavior when classical optimization techniques are infeasible, and its importance in real applications has been demonstrated, e.g., in [10] that proposes applied apprenticeship algorithms for learning control policies for helicopters flying in a very wide range of highly aerobatics with a performance close to that of a human expert pilot. However, the application of machine learning techniques in a distributed multi-agent setting is complicated by the fact that all agents will likely be learning and adapting simultaneously, which may prevent the learning process from converging. Game theory provides the appropriate framework to study learning and autonomy in a distributed setting, as recognized in [11–13].

Reinforcement learning [14] is useful when agents take a sequence of actions, based on information or rewards received from the environment. This technique is inspired by dynamic programming and lays the foundation for developing algorithms to update expected utilities and to use them to explore the system's state space. Actor–critic frameworks [15,16] are based on reinforcement learning methods and use an *actor* component that generates actions and a *critic* component that assesses the costs of these actions. Based on this assessment, the actor policy is updated at each learning step [14,17]. Current reinforcement learning proofs of convergence do not hold for most of the multi-agent system problems.

1.1 Related Work

A survey of existing threats in multi-agent systems and models of realistic and rational adversary models is presented in [18,19]. Consensus in the presence of persistent adversaries has been focused on detecting, identifying, and isolating the failing nodes [20], which can be computationally expensive and often requires the use of global information and specific graph connectivity. The adversaries can easily drive the system unstable and make the system operate with an undesired behavior. The authors in [21–23] show the advantages of using game theory in network security. The work of [24] presents a framework to show how models from game theory can be used to defend an electric power system against antagonistic attacks.

Robust consensus algorithm from the information theory side has been proposed in [25,26]. Specifically, distributed averaging in the presence of channel uncertainties is presented in [25] where the network is used twice per cycle without any optimality guarantees. The work of [26] proposes consensus-based distributed decoding algorithms (e.g., Viterbi algorithm) that are robust to random link failures and additive noise present in the inter-sensor links. Performance bounds are provided to determine the number of iterations needed to guarantee a given level of average decoding errors at any given sensor.

Most of the consensus algorithms proposed in the literature to provide resilience against adversaries either are not optimizing a specific performance criteria [27,28] or require the offline solution of Riccati equations [29,30]. The work of [31] formulates the problem of measurement corruption and jamming attacks as centralized finite horizon problems where the adversary wants to maximize the Euclidean distance between the nodes' state and the consensus line. The computations have to be performed offline in order to compute the optimal strategy. The authors in [32] propose to evaluate the cost of every agent by considering constant states for the other agents, and the authors in [33] present a suboptimal phase synchronization adaptive control scheme for networked nonlinear Euler–Lagrange systems with a centralized performance without incorporating the adversarial inputs in it. A distributed clock synchronization protocol based on a consensus algorithm for non-identical double integrators is presented in [34]. The authors provide conditions on the protocol parameters in terms of rate convergence and in terms of steady-state error in the presence of an additive noise by solving offline a centralized optimization algorithm. In [35], the authors propose a controller that suppresses the effect of constant and time-varying disturbances by using information of agent's and neighbors' states. Vamvoudakis et al. [36] presented reinforcement learning algorithms that guarantee optimal performance in networked systems with leaders, but did not consider the effect of adversaries that corrupt the measured data. In [26], the authors propose a distributed resilient formation control algorithm that consists of a formation control block that adapts online to minimize a local formation error, as well as a learning block that collects information in real time to update the estimates of adversaries. The book of [21] models the malware traffic as an H-infinity problem (zero-sum game) where the optimal filtering strategies can also be used in spam filtering, distributed denial of service attacks, etc. This H-infinity framework allows for dynamically changing filtering rules in order to ensure a certain performance level. Moreover, the objectives of the defense and adversaries are diametrically opposed, so the zero-sum assumption is accurate with a performance guarantee in the form of a minimum security level. The work of [37] provides methods for reaching consensus in the presence of malicious agents, but the algorithms proposed are combinatorial in nature and thus computationally expensive.

A multi-agent synchronization problem of continuous-time Markov jump linear systems is presented in [38] where one subsystem aims to mislead the other subsystems to an unfavorable system state. The authors provide a set of coupled Riccati differential equations derived from centralized performances to characterize the feedback Nash equilibrium solution, but require offline computations and without any stability guarantees. The work of [39] achieves tracking for a case of double-integrator agents, when sufficient communication for path distribution is permitted, but without any optimality and robustness guarantees. In [40] the authors consider a network consisting of identical agents with the only measurement given to each agent is a linear combination of the output of the agent relative to that of its own neighbors. Their aim was to design protocols to attenuate the impact of disturbances in the sense of the H-infinity norm of the corresponding transfer function, without enforcing any optimization criteria.

Coupled Hamilton–Jacobi equations that arise in such problems are nonlinear partial differential equations, and it is well known that in general such equations do not admit global classical solutions and if they do, they may not be smooth. But they may have the so-called viscosity solutions [41,42]. Under certain local reachability and observability assumptions, they have local smooth solutions [22,43]. Various other assumptions guarantee the existence of smooth solutions, such as that the dynamics not be bilinear and the value not contain cross-terms in the state and control input.

Our algorithm on the other hand will propose plug-and-play policies while also satisfying user-defined distributed performance criteria formulated as a graphical game. In this graphical game, only neighbors in the graph can interact and the payoffs depend on the actions of the neighbors which give a representation that is exponential in the size of the largest neighborhood. This is in contrast to games in normal form where every agent interacts with any other agent, the payoffs depend on actions of all agents, and the representation is exponential in the number of players. The result of this work is an adaptive control system that learns based on the interplay of agents in a game, to deliver true online gaming behavior.

1.2 Contributions

This paper proposes a game-theory-based consensus learning algorithm for networked systems with N double-integrator dynamics that are being attacked by persistent adversaries. The problem is solved by combining adaptive dynamic programming and game theory. We consider double-integrator dynamics, which are often used to represent single-degree-of-freedom rotational or translational motion [44].

Furthermore, this work provides a relationship between the optimal consensus problem for multi-agent systems with adversarial inputs and graphical Nash equilibrium. The derived coupled Hamilton–Jacobi equations for multi-agent systems are established by Bellman's dynamic programming, and then a formal stability analysis is developed for the learning scheme. The game has $2N$ players, since the input to each double integrator is the sum of two signals controlled by players with opposite goals: a *controller* that wants to achieve consensus and an *adversary* that wants to prevent this goal. We shall see that this problem formulation results in a distributed consensus algorithm that is significantly more robust than the usual linear consensus.

The criteria to be minimized by each agent depend only on local information, and the problem is formulated as a graphical game. Namely, the criteria depend on every agent's own state and controller/adversaries inputs, as well as the controller/adversarial inputs for its neighbors. These optimization criteria will let us design the decision policies based on the graph structure, local control, and adversarial inputs and hence enable us the development of an optimal and distributed consensus algorithm. It is worth noting that even though we are using distributed performance criteria, this does not allow one to find a closed-form solution to the coupled Bellman equation even for the quadratic case due to the state coupling in the neighborhood.

The computation of an exact equilibrium to the game formulated requires the solution of a system of coupled partial differential equations, one for each pair of agents, which does not appear to be feasible. The algorithm proposed uses instead three approximators for every agent that use distributed information. Every agent uses a critic to approximate each value function, an actor to approximate the optimal controller, and a second actor to approximate the worst-case adversary. The tuning laws for each controller and each adversary are driven by the values of their distributed criteria and of their control signals, as well as the criteria and control signals of their neighbors. In particular, the controllers do not require explicit measurements of the adversarial signals and vice versa. The convergence of the game-theoretical actor–critic algorithm and the stability of the closed-loop system are analyzed using Lyapunov-based adaptive control methods. These results require an appropriate notion of persistence of excitation (PE) to guarantee exponential convergence to a bounded region.

1.3 Organization

In Sect. 2 we introduce the problem formulation and an existence result for the Nash equilibrium. A game-theory-based consensus architecture that uses only local information is presented in Sect. 3, where we also provide an online learning algorithm. The effectiveness of the proposed approach is illustrated in Sect. 4 through simulations. Finally, in Sect. 5 we conclude and discuss future work.

2 Problem Formulation

We consider a system consisting of N two-input agents, each modeled by a two-input double integrator:

$$\begin{cases} \dot{q}_i = p_i & q_i \in \mathbb{R}^m \\ \dot{p}_i = u_i + v_i, & p_i \in \mathbb{R}^m \end{cases} \qquad \forall i \in \mathcal{N} := \{1, \dots N\}, \tag{1}$$

where $u_i, v_i \in \mathbb{R}^m$ are the two inputs to agent i, which are controlled by players with opposite goals. We thus have a total of $2N$ players: N of them—which we call *controllers*—select values for $u_i(t)$, $t \geqslant 0$, $i \in \mathcal{N}$, within appropriate sets $\mathcal{U}_i$ and the other N—which we call *adversaries*—select values for the $v_i(t)$, $t \geqslant 0$, $i \in \mathcal{N}$, within appropriate sets $\mathcal{V}_i$.

2.1 Distributed Performance Criteria

For each $i \in \mathcal{N}$, the criteria of the players associated with the inputs u_i and v_i are symmetric and depend on their state $[q_i' \ p_i']$, their inputs $[u_i' \ v_i']$, and also the inputs of a nonempty subset $\mathcal{N}_i \subset \mathcal{N}$ of the other agents. Specifically, u_i and v_i want to minimize and maximize, respectively, the following criteria

$$\begin{aligned} &J_i(p_i(0), q_i(0), p_{\mathcal{N}_i}(0), q_{\mathcal{N}_i}(0); u_i, u_{\mathcal{N}_i}, v_i, v_{\mathcal{N}_i}) \\ &\quad = \frac{1}{2}\int_0^\infty \Big(\|s_i\|^2 + \|u_i\|^2 - \gamma_{\mathrm{ii}}^2\|v_i\|^2 + \sum_{j\in\mathcal{N}_i}\|u_j\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\mathrm{ij}}^2\|v_j\|^2\Big)\mathrm{d}t, \end{aligned} \tag{2}$$

where

$$s_i(t) = \sum_{j \in \mathcal{N}_i} \left(\left[\begin{smallmatrix} q_i \\ p_i \end{smallmatrix} \right] - \left[\begin{smallmatrix} q_j \\ p_j \end{smallmatrix} \right] \right) \in \mathbb{R}^{2m}, \quad \forall t \geqslant 0, \; i \in \mathcal{N}, \tag{3}$$

and $\gamma_{\mathrm{ii}}, \; \gamma_{\mathrm{ij}}, \; \forall i, j \in \mathcal{N}$, are positive constants. In (2) and in the sequel, we use the subscript $\cdot_{\mathcal{N}_i}$ as a shorthand notation for all the subscripts $\cdot_j$, $j \in \mathcal{N}_i$, associated with the neighbors of i. The game just defined is called a *graphical game* [45], because the coupling between the criteria of the different players can be naturally associated with a graph $\mathcal{G}$ with N nodes (one for each pair of opposing agents) and edges defined by the neighboring relations expressed by the sets $\mathcal{N}_i$.

Remark 2.1 Note that, with the performance index (2), we consider attack scenarios where the adversary's goal is to drive the agents to an non-consensus state while remaining stealthy. These kinds of attacks leave more autonomy to the adversary and consider attacks that may be theoretically discernible, but are still stealthy since they do not cause any alarm by the anomaly detector. We shall see in the subsequent theorem (cf. Theorem 2.1) that in order for the systems to be stabilized one needs to pick $\gamma_{\mathrm{ii}} > 1$ and $\gamma_{\mathrm{ij}} < \gamma_{\mathrm{jj}}^2$, $\forall i \in \mathcal{N}$, $j \in \mathcal{N}_i$. These conditions are similar to the ones in [21] (see chapter 7), [22,43] where one needs to pick a $\gamma > \gamma^* \geqslant 0$, where γ^* is the smallest γ such that the system is stabilized.

The variables $s_i(t)$ should be viewed as *consensus tracking errors* that express the (weighted) sum of the errors between the state of agent i and the states of its neighbors in $\mathcal{N}_i$. When all the $s_i(t)$ converge to zero, we say that *consensus is asymptotically reached* in the sense that $q_i = q_j$, $p_i = p_j$, $\forall i, j \in \mathcal{N}$ [6].

We are interested in finding *Nash equilibria policies* u_i^* and v_i^*, $\forall i \in \mathcal{N}$, for the $2N$-player game, in the sense that

$$\begin{aligned} J_i(\cdot; u_i^*, u_{\mathcal{N}_i}^*, v_i, v_{\mathcal{N}_i}^*) &\leqslant J_i^*(\cdot; u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*) \\ &\leqslant J_i(\cdot; u_i, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*), \; \forall v_i, \; u_i, \; i \in \mathcal{N}, \end{aligned} \tag{4}$$

where, for simplicity, we omitted the dependence of criteria (2) on the initial conditions. Essentially, each pair of players u_i and v_i are engaged in a zero-sum game, but the outcomes of all these zero-sum games are coupled through the neighboring relations expressed by the sets $\mathcal{N}_i$. The positive terms on the $\|u_i\|$ and the negative terms on the $\|v_i\|$ in (2) express the fact that the players that select the u_i and the v_i want to achieve their objectives, while keeping these signals small. Recall that J_i is a cost for the player that selects u_i, but a reward for the player that selects v_i. It is clear that criteria (2) enable us to define a general neighborhood optimization framework, while the different signs are due to the zero-sum game formulation. One can also remove the neighboring terms (see the fourth and fifth terms of (2)) since (3) has the information from the neighborhood.

To solve this problem, it is convenient to rewrite the agent dynamics (1) in terms of the consensus tracking errors (3):

$$\begin{aligned}\dot{s}_i = \begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} s_i + d_i \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_i + v_i)\\ - \sum_{j\in\mathcal{N}_i} \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_j + v_j),\ i \in \mathcal{N},\end{aligned} \tag{5}$$

where $d_i := |\mathcal{N}_i| > 0$, and $\mathbf{0}$ and $\mathbf{I}$ denote the $m \times m$ zero and identity matrices, respectively. This shows that $J_i(\cdot)$ in (2) only depends on the initial conditions $p_i(0), q_i(0), p_{\mathcal{N}_i}(0), q_{\mathcal{N}_i}(0)$ through $s_i(0)$. In the sequel, we thus simply write $J_i(s_i(0); u_i, u_{\mathcal{N}_i}, v_i, v_{\mathcal{N}_i})$ for the left-hand side of (2).

2.2 Existence and Stability of Equilibria

The saddle-point conditions (4) can be expressed by $2N$ coupled optimizations:

$$\begin{aligned}&J_i(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*)\\ &\quad = \max_{\mathrm{v_i}} J_i(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i, v_{\mathcal{N}_i}^*), \qquad i \in \mathcal{N}, \quad (6a)\\ &J_i(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*)\\ &\quad = \min_{u_i} J_i(s_i(0); u_i, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*), \qquad i \in \mathcal{N}, \quad (6b)\end{aligned}$$

which share dynamics (5) and the Hamiltonians

$$\begin{aligned}H_i(s_i, \rho_i, u_i, u_{\mathcal{N}_i}, v_i, v_{\mathcal{N}_i}) := \rho_i^T\Bigg(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} s_i\\ + d_i \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_i + v_i) - \sum_{j\in\mathcal{N}_i} \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_j + v_j)\Bigg) + \frac{1}{2}\Bigg(\|s_i\|^2 + \|u_i\|^2\\ - \gamma_{\mathrm{ii}}^2\|v_i\|^2 + \sum_{j\in\mathcal{N}_i}\|u_j\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\mathrm{ij}}^2\|v_j\|^2\Bigg),\ i \in \mathcal{N},\end{aligned}$$

where $\rho_i \in \mathbb{R}^{2m}$ is the co-state (adjoint) variable. For each $i \in \mathcal{N}$, the left-hand side of the two optimizations in (6) can be viewed as a (common) value function to both optimizations

$$\begin{aligned}V_i^*(s_i(0)) &:= \max_{\mathrm{v_i}} J_i(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i, v_{\mathcal{N}_i}^*)\\ &= \min_{u_i} J_i(s_i(0); u_i, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*), \quad i \in \mathcal{N},\end{aligned} \tag{7}$$

that should satisfy the following coupled Bellman equations:

$$H_i\Big(s_i, \frac{\partial V_i^*}{\partial s_i}, u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*\Big) = 0, \quad i \in \mathcal{N}, \tag{8}$$

with boundary conditions $V_i^*(0) = 0$ and controller/adversarial policies given by

$$\begin{aligned} u_i^* &= \arg\min_{u_i} H_i\left(s_i, \frac{\partial V_i^*}{\partial s_i}, u_i, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*\right) \\ &= -d_i \begin{bmatrix}\mathbf{0} \ \mathbf{I}\end{bmatrix} \frac{\partial V_i^*}{\partial s_i}, \ i \in \mathcal{N}, \qquad (9) \\ v_i^* &= \arg\max_{\mathrm{v_i}} H_i\left(s_i, \frac{\partial V_i^*}{\partial s_i}, u_i^*, u_{\mathcal{N}_i}^*, v_i, v_{\mathcal{N}_i}^*\right) \\ &= \frac{d_i}{\gamma_{\mathrm{ii}}^2} \begin{bmatrix}\mathbf{0} \ \mathbf{I}\end{bmatrix} \frac{\partial V_i^*}{\partial s_i} = -\frac{1}{\gamma_{\mathrm{ii}}^2} u_i^*, \ i \in \mathcal{N}. \qquad (10) \end{aligned}$$

One can see from (9) and (10) that when one player accelerates ($u_i^* > 0$) the other decelerates ($v_i^* < 0$) and vice versa.

Remark 2.2 One could represent the value functions as quadratic in the consensus tracking error, i.e., $V_i^*(s_i) : \mathbb{R}^n \to \mathbb{R}$,

$$V_i^*(s_i) = \frac{1}{2} s_i^T P_i s_i, \ \forall s_i, \forall i \in \mathcal{N},$$

where $P_i \in \mathbb{R}^{n\times n}$, $\forall i \in \mathcal{N}$, are the unique symmetric positive definite matrices that solve the following complicated distributed coupled equations,

$$\begin{aligned} &s_i^T P_i\Bigg(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} s_i - d_i^2 \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\left(1 - \frac{1}{\gamma_{ii}^2}\right) P_i s_i \\ &\quad + \sum_{j\in\mathcal{N}_i} d_j \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\left(1 - \frac{1}{\gamma_{ii}^2}\right) P_j s_j\Bigg) \\ &\quad + \Bigg(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} s_i - d_i^2 \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\left(1 - \frac{1}{\gamma_{ii}^2}\right) P_i \\ &\quad + \sum_{j\in\mathcal{N}_i} d_j \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\left(1 - \frac{1}{\gamma_{ii}^2}\right) P_j s_j\Bigg)^T P_i s_i \\ &\quad + \sum_{j\in\mathcal{N}_i} d_j^2 s_j^T P_j \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\left(1 - \frac{\gamma_{ij}^2}{\gamma_{jj}^4}\right) P_j s_j \\ &\quad + d_i^2 s_i^T P_i \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\left(1 - \frac{1}{\gamma_{ii}^2}\right) P_i s_i + \|s_i\|^2 = 0, \forall i \in \mathcal{N}. \end{aligned}$$

Obviously, in the above equation the coupling of each state s_i with the neighborhood state $s_j,\ j \in \mathcal{N}_i$ makes it difficult to provide a closed-form solution, hereby rendering optimal control techniques of limited use for non-quadratic costs such as the one considered here [15,46].

Remark 2.3 For every fixed $u^*_{\mathcal{N}_i}, v^*_i, v^*_{\mathcal{N}_i}$, we have $\frac{\partial^2 H_i}{\partial^2 u_i} > 0$ and therefore the Hamiltonian is minimized at the stationarity point (9). Conversely, for every fixed $u^*_i, u^*_{\mathcal{N}_i}, v^*_{\mathcal{N}_i}$, we have $\frac{\partial^2 H_i}{\partial^2 v_i} < 0$ and therefore the Hamiltonian is maximized at the stationarity point (10).

Remark 2.4 As we shall see in the following theorem every agent is participating in a zero-sum game, but all the agents together participate in a non-zero-sum game. The controller and the adversarial input of each agent are two players competing with each other, and as explained in [21], the intelligent maximizing adversary plays his strategy to ensure a certain performance. When one player adopts a constant strategy, it becomes an optimal control problem. In general, each player (controller or adversary) tries to make the best outcome taking into account that his opponent also does the same. The two players, controller and adversary, are measuring (3) (due to the presence of $\frac{\partial V_i^*}{\partial s_i}$ in (9)–(10)) which includes the positions and the velocities of the agents in the neighborhood.

The following theorem shows that under the assumption of a smooth solution of the coupled Bellman equations (8) and some other conditions, the Nash equilibrium is attained, with the equality to zero replaced by less than or equal to zero (as in the coupled Hamilton–Jacobi–Isaacs (HJI) inequalities [22,43,47]).

Theorem 2.1 *Suppose that there exist continuously differentiable, positive definite functions $V_i^* \in C^1$, $i \in \mathcal{N}$ that satisfy the following inequalities,*

$$
\begin{aligned}
&{\frac{\partial V_i^*}{\partial s_i}}^T \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + \frac{d_i^2}{2} \frac{1-\gamma_{\mathrm{ii}}^2}{\gamma_{\mathrm{ii}}^2} {\frac{\partial V_i^*}{\partial s_i}}^T \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_i^*}{\partial s_i} \\
&+ \sum_{j \in \mathcal{N}_i} \frac{\gamma_{\mathrm{jj}}^2 - 1}{\gamma_{\mathrm{jj}}^2} {\frac{\partial V_i^*}{\partial s_i}}^T \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_j^*}{\partial s_j} + \frac{1}{2}\Bigg(\|s_i\|^2 \\
&+ \sum_{j \in \mathcal{N}_i} d_j^2 \Big(\frac{\gamma_{\mathrm{jj}}^4 - \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4}\Big) {\frac{\partial V_j^*}{\partial s_j}}^T \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_j^*}{\partial s_j} \Bigg) \leqslant 0, \ \forall s_i \qquad (11)
\end{aligned}
$$

with $V_i^*(0) = 0$, $\forall i \in \mathcal{N}$ and $\gamma_{\mathrm{ii}} > 1$ and $\gamma_{\mathrm{ij}} < \gamma_{\mathrm{jj}}^2$, $\forall i \in \mathcal{N}$, $j \in \mathcal{N}_i$. The closed-loop system (5) with

$$
u_i = u_i^* := -d_i \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_i^*}{\partial s_i}, \qquad v_i = v_i^* := \frac{d_i}{\gamma_{\mathrm{ii}}^2} \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_i^*}{\partial s_i}, \qquad \forall i \in \mathcal{N}, \qquad (12)
$$

is asymptotically stable. Moreover, the controller/adversarial policies (12) *form a Nash equilibrium and*

$$J_i^*(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*) = V_i^*(s_i(0)), \quad \forall i \in \mathcal{N}.$$

Proof of Theorem 2.1 The orbital derivatives of the V_i^* along solutions to (5), (12) are given by

$$\begin{aligned}\dot{V}_i^* = \frac{\partial V_i^*}{\partial s_i}^T \dot{s}_i = \frac{\partial V_i^*}{\partial s_i}^T \Bigg(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_i + v_i) \\ - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_j + v_j) \Bigg) \end{aligned} \tag{13}$$

and, after substituting the policies (12) in (13), one obtains

$$\begin{aligned}\dot{V}_i^* = \frac{\partial V_i^*}{\partial s_i}^T \Bigg(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + \frac{d_i^2}{\gamma_{\text{ii}}^2}(1 - \gamma_{\text{ii}}^2) \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_i^*}{\partial s_i} \\ + \sum_{j \in \mathcal{N}_i} \frac{\gamma_{\text{jj}}^2 - 1}{\gamma_{\text{jj}}^2} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_j^*}{\partial s_j} \Bigg).\end{aligned}$$

By using inequalities (11) we have that

$$\begin{aligned}\dot{V}_i^* &= \frac{\partial V_i^*}{\partial s_i}^T \Bigg(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + \frac{d_i^2}{\gamma_{\text{ii}}^2}(1 - \gamma_{\text{ii}}^2) \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_i^*}{\partial s_i} \\ &\quad + \sum_{j \in \mathcal{N}_i} \frac{\gamma_{\text{jj}}^2 - 1}{\gamma_{\text{jj}}^2} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_j^*}{\partial s_j} \Bigg) \\ &\leqslant -\frac{1}{2} \Bigg(\|s_i\|^2 + \frac{d_i^2}{\gamma_{\text{ii}}^2}(\gamma_{\text{ii}}^2 - 1) \frac{\partial V_i^*}{\partial s_i}^T \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_i^*}{\partial s_i} \\ &\quad + \sum_{j \in \mathcal{N}_i} \frac{d_j^2}{\gamma_{\text{jj}}^4} \big(\gamma_{\text{jj}}^4 - \gamma_{\text{ij}}^2\big) \frac{\partial V_j^*}{\partial s_j}^T \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial V_j^*}{\partial s_j} \Bigg).\end{aligned}$$

Since $\gamma_{\text{ii}} > 1$ and $\gamma_{\text{ij}} < \gamma_{\text{jj}}^2, \forall i \in \mathcal{N}, j \in \mathcal{N}_i$, we further conclude that $\dot{V}_i^* \leqslant -\frac{1}{2}\|s_i\|^2$, from which asymptotic stability follows using the Lyapunov theorem [48].

Next we need to prove that the controller/adversarial policies form a Nash equilibrium. Since the functions V_i^*, $i \in \mathcal{N}$ are smooth, are zero at zero, and converge to zero as $t \to \infty$ (due to the asymptotic stability), we can write (2) as

$$\begin{aligned}J_i(s_i(0), u_i, u_{\mathcal{N}_i}, v_i, v_{\mathcal{N}_i}) &= \frac{1}{2}\int_0^\infty \Bigg(\|s_i\|^2 + \|u_i\|^2\\ &+ \sum_{j\in\mathcal{N}_i}\|u_j\|^2 - \gamma_{\text{ii}}^2\|v_i\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\text{ij}}^2\|v_j\|^2\Bigg)\mathrm{d}t\\ &+ V_i^*(s_i(0)) + \int_0^\infty \dot{V}_i^*(s(t))\mathrm{d}t, \quad i\in\mathcal{N},\end{aligned}$$

and, in view of (5), we can rewrite this equation as

$$\begin{aligned}J_i(s_i(0), u_i, u_{\mathcal{N}_i}, v_i, v_{\mathcal{N}_i}) &= \frac{1}{2}\int_0^\infty \Bigg(\|s_i\|^2 + \|u_i\|^2\\ &+ \sum_{j\in\mathcal{N}_i}\|u_j\|^2 - \gamma_{\text{ii}}^2\|v_i\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\text{ij}}^2\|v_j\|^2\Bigg)\mathrm{d}t\\ &+ V_i^*(s_i(0)) + \int_0^\infty \frac{\partial V_i^*}{\partial s_i}^T\Bigg(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}s_i + d_i\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_i+v_i)\\ &- \sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_j+v_j)\Bigg)\mathrm{d}t, \quad i\in\mathcal{N}.\end{aligned}$$

Writing (8) as

$$\begin{aligned}\frac{\partial V_i^*}{\partial s_i}^T\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}s_i &= -\frac{\partial V_i^*}{\partial s_i}^T\Bigg(d_i\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_i^*+v_i^*)\\ &- \sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_j^*+v_j^*)\Bigg)\\ &- \frac{1}{2}\Bigg(\|s_i\|^2 + \|u_i^*\|^2 - \gamma_{\text{ii}}^2\|v_i^*\|^2 + \sum_{j\in\mathcal{N}_i}\|u_j^*\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\text{ij}}^2\|v_j^*\|^2\Bigg),\end{aligned}$$

we can rewrite the previous equations as

$$\begin{aligned}J_i(s_i(0); u_i, u_{\mathcal{N}_i}, v_i, v_{\mathcal{N}_i}) &= \frac{1}{2}\int_0^\infty \Bigg(\|u_i - u_i^*\|^2\\ &+ \sum_{j\in\mathcal{N}_i}\|u_j - u_j^*\|^2 - \gamma_{\text{ii}}^2\|v_i - v_i^*\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\text{ij}}^2\|v_j - v_j^*\|^2\\ &- 2\frac{\partial V_i^*}{\partial s_i}^T\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_j - u_j^*) + 2\sum_{j\in\mathcal{N}_i}u_j^*(u_j - u_j^*)\\ &- 2\frac{\partial V_i^*}{\partial s_i}^T\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(v_j - v_j^*) - 2\sum_{j\in\mathcal{N}_i}v_j^*\gamma_{\text{ij}}^2(v_j - v_j^*)\Bigg)\mathrm{d}t\\ &+ V_i^*(s_i(0)), \quad i\in\mathcal{N}.\end{aligned}$$

Assuming that all the neighbors of the ith node play at the proposed equilibrium (i.e., $u_j = u_j^*, \forall j \in \mathcal{N}_i$), we obtain

$$J_i(s_i(0); u_i, u_{\mathcal{N}_i}^*, v_i, v_{\mathcal{N}_i}^*) = \frac{1}{2}\int_0^\infty \Big(\|u_i - u_i^*\|^2 - \gamma_{\mathrm{ii}}^2 \|v_i - v_i^*\|^2 \Big) \mathrm{d}t + V_i^*(s_i(0)), \quad i \in \mathcal{N}. \tag{14}$$

Moreover, setting $u_i = u_i^*$ in (14) leads to

$$J_i(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i, v_{\mathcal{N}_i}^*) = \frac{1}{2}\int_0^\infty \Big(- \gamma_{\mathrm{ii}}^2 \|v_i - v_i^*\|^2 \Big) \mathrm{d}t + V_i^*(s_i(0)), \quad i \in \mathcal{N}; \tag{15}$$

setting $v_i = v_i^*$ in (14) leads to

$$J_i(s_i(0); u_i, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*) = \frac{1}{2}\int_0^\infty \Big(\|u_i - u_i^*\|^2 \Big) \mathrm{d}t + V_i^*(s_i(0)), \quad i \in \mathcal{N}; \tag{16}$$

and setting $u_i = u_i^*$ and $v_i = v_i^*$ in (14) leads to

$$J_i^*(s_i(0); u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*) = V_i^*(s_i(0)), \quad i \in \mathcal{N}. \tag{17}$$

The Nash equilibrium condition (4) then follows directly from (15), (16), and (17). □

3 Reaching Consensus with Learning Ideas

Since in general it is not easy to find solutions V_i^* to the Bellman equations (8), based on which one would construct the optimal controllers (9) and adversaries (10), we shall use an actor–critic architecture to compute approximations to these functions. Specifically, N *critic* approximators will provide approximations to the value functions V_i^*, $i \in \mathcal{N}$, in (8) and $2N$ *actor* approximators will provide approximations to the optimal controllers and adversaries u_i^*, v_i^*, $i \in \mathcal{N}$, given by (9) and (10), respectively.

We use linearly parametrized critics with h basis functions, defined using smooth basis functions $\phi_i := [\phi_{\mathrm{i1}}\ \phi_{\mathrm{i2}}\ \dots\ \phi_{\mathrm{ih}}] : \mathbb{R}^{2m} \to \mathbb{R}^h$, $i \in \mathcal{N}$, which allow us to write

$$V_i^*(s_i) = W_i^T \phi_i(s_i) + \epsilon_{\mathrm{c_i}}(s_i), \quad \forall s_i, \quad i \in \mathcal{N}, \tag{18}$$

where the $W_i \in \mathbb{R}^h$ denote ideal weights and the $\epsilon_{c_i}(s_i)$ the corresponding residual errors. Based on this, the optimal controllers in (9) can be rewritten as

$$u_i^*(s_i) = -d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \left(\tfrac{\partial \phi_i}{\partial s_i}(s_i)^T W_i + \tfrac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right), \quad \forall s_i, \quad i \in \mathcal{N}, \tag{19}$$

and adversaries in (10) can be rewritten as

$$v_i^*(s_i) = \frac{d_i}{\gamma_{\mathrm{ii}}^2} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \left(\frac{\partial \phi_i}{\partial s_i}(s_i)^T W_i + \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right), \quad \forall s_i, \quad i \in \mathcal{N}. \tag{20}$$

For brevity from now we will omit the dependence of the basis functions on the state s_i and simply write ϕ_i for $\phi_i(s_i)$.

Remark 3.1 Note that the adversarial input v_i^* in (20) is not the actual input by the adversary, but the worst-case input that she can introduce to our system. The actual adversary has the freedom to do anything she likes, as long as she remains stealthy.

Assumption 3.1 The following three statements must be satisfied $\forall s_i$ and $\forall i \in \mathcal{N}$:

- The basis functions ϕ_i and their derivatives $\frac{\partial \phi_i}{\partial s_i}$ are bounded.
- The ideal weights are bounded by known constants: $\|W_i\| \leqslant W_{\mathrm{imax}}$.
- The residual errors and their derivatives are bounded by known constants: $\|\epsilon_{\mathrm{c_i}}\| \leqslant \epsilon_{\mathrm{cimax}}$ and $\left\| \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right\| \leqslant \nabla \epsilon_{\mathrm{cimax}}$.

Assumption 3.1 is a rather standard assumption in neuro-adaptive control [49–51]. The first two points are satisfied because W_i and ϕ_i provide the value functions [15]. In order to satisfy the third one, one can pick sigmoidal (e.g., hyperbolic tangent) and Gaussian functions as basis functions.

Remark 3.2 According to Weierstrass higher-order approximation theorem [49] as the number of basis sets h increases, the approximation errors $\forall i \in \mathcal{N}$ go to zero, i.e., $\epsilon_{\mathrm{c_i}}$ and $\left\| \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right\|$ as $h \to \infty$. We shall require a form of uniformity in this approximation result that is common in neuro-adaptive control and other approximation techniques [15,49,50].

3.1 Actor–Critic Control Architecture

To find the optimal controller/adversarial policies for every agent one needs to compute the gradient of each optimal value function. While it would seem that a single set of weights would suffice for the approximation since one can easily differentiate (18), we will independently adjust three sets of weights: the critic weights $\hat{W}_{\mathrm{c_i}} \in \mathbb{R}^h$ that approximate the optimal value functions (7) according to

$$\hat{V}_i = \hat{W}_{\mathrm{c_i}}^T \phi_i(s_i), \ i \in \mathcal{N};$$

the controller actor weights $\hat{W}_{\mathrm{u_i}} \in \mathbb{R}^h$ that approximate the optimal controller in (12) according to

$$\hat{u}_i = -d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{u_i}, \quad i \in \mathcal{N}; \tag{21}$$

and the actor weights $\hat{W}_{\mathrm{v_i}} \in \mathbb{R}^h$ that approximate the optimal adversary in (12) according to

$$\hat{v}_i = \frac{d_i}{\gamma_{\mathrm{ii}}^2} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{v_i}, \quad i \in \mathcal{N}. \tag{22}$$

All these approximators share the same set of basis functions ϕ_i. Adjusting three independent sets of weights carries additional computational burden, but the flexibility introduced by this "over-parametrization" will enable us to establish convergence to a Nash equilibrium and guarantee Lyapunov-based stability.

Remark 3.3 Note that the approximated version of the adversarial input $\hat{v}_i$ in (22) is an approximation of the worst-case adversarial input (20) that the controller is using. The actual adversarial inputs have the freedom to do anything they want, as long as they remain stealthy.

Defining the errors $e_i \in \mathbb{R}$, $\forall i \in \mathcal{N}$, between the values of the Hamiltonians in (8) at the optimal value function/policies and their value at the estimated value function/policies leads to

$$\begin{aligned} e_i \equiv \hat{H}_i &\left(s_i, \hat{W}_{\mathrm{c_i}}^T \frac{\partial \phi_i}{\partial s_i}, \hat{u}_i, \hat{u}_{\mathcal{N}_i}, \hat{v}_i, \hat{v}_{\mathcal{N}_i} \right) \\ &- H\left(s_i, \frac{\partial V_i^*}{\partial s_i}, u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^* \right) = \hat{W}_{\mathrm{c_i}}^T \frac{\partial \phi_i}{\partial s_i} \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i \right. \\ &+ \left. d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\hat{u}_i + \hat{v}_i) - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\hat{u}_j + \hat{v}_j) \right) \\ &+ \frac{1}{2} \left(\|s_i\|^2 + \|\hat{u}_i\|^2 + \sum_{j \in \mathcal{N}_i} \|\hat{u}_j\|^2 - \gamma_{\mathrm{ii}}^2 \|\hat{v}_i\|^2 - \sum_{j \in \mathcal{N}_i} \gamma_{\mathrm{ij}}^2 \|\hat{v}_j\|^2 \right) \\ = \hat{W}_{\mathrm{c_i}}^T &\frac{\partial \phi_i}{\partial s_i} \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\hat{u}_i + \hat{v}_i) \right. \\ &- \left. \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\hat{u}_j + \hat{v}_j) \right) + \hat{r}_i, \quad i \in \mathcal{N}, \end{aligned} \tag{23}$$

where $\hat{r}_i := \frac{1}{2}\left(\|s_i\|^2 + \|\hat{u}_i\|^2 + \sum_{j \in \mathcal{N}_i} \|\hat{u}_j\|^2 - \gamma_{\mathrm{ii}}^2 \|\hat{v}_i\|^2 - \sum_{j \in \mathcal{N}_i} \gamma_{\mathrm{ij}}^2 \|\hat{v}_j\|^2 \right)$ and $H(s_i, \frac{\partial V_i^*}{\partial s_i}, u_i^*, u_{\mathcal{N}_i}^*, v_i^*, v_{\mathcal{N}_i}^*) = 0$ from (8). Our goal is to design tuning laws for the critic weights that minimize the squared-norm of the errors e_i, $\forall i \in \mathcal{N}$:

$$K_i(t) = \frac{1}{2} \|e_i\|^2, \ i \in \mathcal{N}. \tag{24}$$

We should note, however, that while it is true that if the functions $s_i \mapsto \hat{W}_{\mathrm{c_i}}^T \phi_i(s_i)$ satisfy the Bellman equations $\forall i \in \mathcal{N}$, $s_i \in \mathbb{R}^{2m}$, then the error $e_i(t)$ and their squared errors $K_i(t)$ are zero $\forall i \in \mathcal{N}, t \geqslant 0$, the converse is not necessarily true. Because of this, it will not suffice to show that the e_i converge to zero and, instead, we will need to prove explicitly that the critics (and actors) asymptotically approximate their optimal values.

3.2 Learning Algorithm

The gradient descent estimate $\hat{W}_{\mathrm{c_i}}(t)$ for the critic's weights can be constructed by differentiating K_i in (24) as follows,

$$\dot{\hat{W}}_{\mathrm{c_i}} = -\alpha_i \frac{\partial K_i}{\partial \hat{W}_{\mathrm{c_i}}} = -\alpha_i \frac{\omega_i}{(1+\omega_i^T \omega_i)^2} e_i^T, \quad i \in \mathcal{N}, \tag{25}$$

where $\omega_i = \frac{\partial \phi_i}{\partial s_i}\Big(\left[\begin{smallmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{smallmatrix}\right] s_i + d_i \left[\begin{smallmatrix}\mathbf{0}\\ \mathbf{I}\end{smallmatrix}\right](\hat{u}_i + \hat{v}_i) - \sum_{j \in \mathcal{N}_i} \left[\begin{smallmatrix}\mathbf{0}\\ \mathbf{I}\end{smallmatrix}\right](\hat{u}_j + \hat{v}_j)\Big)$ and the constant $\alpha_i \in \mathbb{R}_+$ essentially determines the speed of convergence.

We shall see below that the following tuning laws for the actors (controllers and adversaries) guarantee Lyapunov stability of the overall system:

$$\begin{aligned}
\dot{\hat{W}}_{\mathrm{u_i}} = \mathcal{P}r\Bigg[\hat{W}_{\mathrm{u_i}}, \alpha_{ui}\Bigg\{ & d_i^2 \frac{\partial \phi_i}{\partial s_i}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{\mathrm{u_i}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{\mathrm{c_i}} \\
& + \sum_{j \in \mathcal{N}_i} d_j^2 \frac{\partial \phi_j}{\partial s_j}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial \phi_j}{\partial s_j}^T \hat{W}_{\mathrm{u_j}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{c_j} \\
& + \sigma_{ui}(\hat{W}_{\mathrm{c_i}} - \hat{W}_{\mathrm{u_i}})\Bigg\}\Bigg], \ i \in \mathcal{N}
\end{aligned} \tag{26}$$

$$\begin{aligned}
\dot{\hat{W}}_{\mathrm{v_i}} = \mathcal{P}r\Bigg[\hat{W}_{\mathrm{v_i}}, \alpha_{\mathrm{vi}}\Bigg\{ & -\frac{d_i^2}{\gamma_{\mathrm{ii}}^2} \frac{\partial \phi_i}{\partial s_i}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{\mathrm{v_i}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{\mathrm{c_i}} \\
& - \sum_{j \in \mathcal{N}_i} \frac{d_j^2 \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4} \frac{\partial \phi_j}{\partial s_j}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial \phi_j}{\partial s_j}^T \hat{W}_{\mathrm{v_j}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{c_j} \\
& + \sigma_{\mathrm{vi}}(\hat{W}_{\mathrm{c_i}} - \hat{W}_{\mathrm{v_i}})\Bigg\}\Bigg], \quad i \in \mathcal{N},
\end{aligned} \tag{27}$$

where the constants $\alpha_{ui}, \alpha_{\mathrm{vi}} \in \mathbb{R}_+$ determine the speed of convergence; $\sigma_{ui}, \sigma_{\mathrm{vi}} \in \mathbb{R}_+$ are adaptation gains; and $\bar{\omega}_i := \frac{\omega_i^T}{\omega_i^T \omega_i + 1}$ is a bounded signal, with

$$\|\bar{\omega}_i\| \leqslant \omega_{\mathrm{imax}} := \frac{1}{2}, \ \forall i \in \mathcal{N}.$$

The symbol $\mathcal{P}r$ in (26)–(27) denotes the smooth projection operator that is used to modify the adaptation laws. The inclusion of this operator in the tuning laws guarantees that $\hat{W}_{\mathrm{u_i}}$ and $\hat{W}_{\mathrm{v_i}}$ remain inside $\mathcal{S}$. The projection operator used in the update laws (26) and (27) provides an effective way [50,52–55] to eliminate parameter drift and keeps

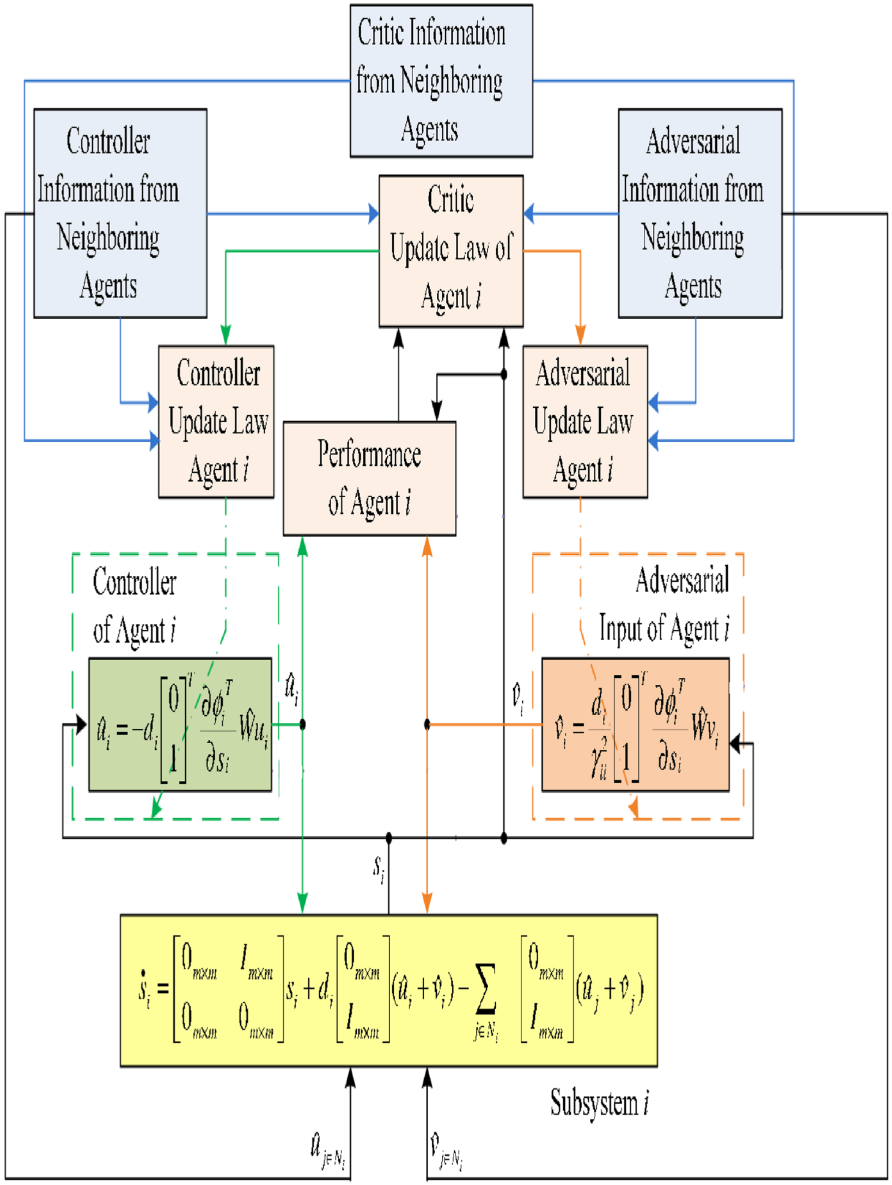
Critic Information from Neighboring Agents
Controller Information from Neighboring Agents
Critic Update Law of Agent i
Adversarial Information from Neighboring Agents
Controller Update Law Agent i
Performance of Agent i
Adversarial Update Law Agent i
Controller of Agent i
$\hat{u}_i = -d_i \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \frac{\partial \phi_i^T}{\partial s_i} \hat{W}u_i$
$\hat{u}_i$
$\hat{v}_i$
Adversarial Input of Agent i
$\hat{v}_i = \frac{d_i}{\gamma_{ii}^2} \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T \frac{\partial \phi_i^T}{\partial s_i} \hat{W}v_i$
s_i
$\dot{s}_i = \begin{bmatrix} 0_{m\times m} & I_{m\times m} \\ 0_{m\times m} & 0_{m\times m} \end{bmatrix} s_i + d_i \begin{bmatrix} 0_{m\times m} \\ I_{m\times m} \end{bmatrix} (\hat{u}_i + \hat{v}_i) - \sum_{j\in N_i} \begin{bmatrix} 0_{m\times m} \\ I_{m\times m} \end{bmatrix} (\hat{u}_j + \hat{v}_j)$
Subsystem i
$\hat{u}_{j\in N_i}$
$\hat{v}_{j\in N_i}$

Fig. 1 Visualization of the proposed control architecture for every agent i. The scheme shown is implemented inside the controller of each agent and uses an approximation of the worst-case adversary

the basis weights within a priori defined bounds (bounded convex set). In other words, as stated in Assumption 3.1, the projection operator makes sure that W_{imax} specifies the boundary and ϵ_{cimax} specifies boundary tolerance.

A block diagram showing the proposed control architecture for agent i is given in Fig. 1, where we can see that the controller and the adversary of every agent i

are driven by the controllers and adversaries, respectively, of her neighborhood. The intuition behind this is that every agent needs to have information of the neighborhood "performance" to see how good the agent does (that is why the neighborhood critic weights appear in both (26) and (27)) and the inputs of the agents in the neighborhood that share a common goal (controllers vs. adversaries).

3.3 Convergence and Stability Analysis

Define the critic and actor estimation errors $\forall i \in \mathcal{N}$ as

$$\tilde{W}_{\mathrm{c_i}} = W_i - \hat{W}_{\mathrm{c_i}},\ \tilde{W}_{\mathrm{u_i}} = W_i - \hat{W}_{\mathrm{u_i}},\ \tilde{W}_{\mathrm{v_i}} = W_i - \hat{W}_{\mathrm{v_i}}.$$

The following lemma proved in "Appendix" provides the critic error dynamics for each agent $i \in \mathcal{N}$ in the form of a nominal system plus a perturbation.

Lemma 3.1 *The critic estimation error dynamics for agent $i \in \mathcal{N}$ can be written in the following form:*

$$\dot{\tilde{W}}_{\mathrm{c_i}} = F_{\mathrm{i0}} + F_{\mathrm{i1}},\ \forall i \in \mathcal{N}, \tag{28}$$

where $F_{\mathrm{i0}} := -\alpha_i \bar{\omega}_i \bar{\omega}_i^T \tilde{W}_{c_i}$ can be viewed as a nominal system*;*

$$\begin{aligned} F_{\mathrm{i1}} := \alpha_i \frac{\omega_i}{(\omega_i^T \omega_i + 1)^2} \Bigg[& - W_i^T \frac{\partial \phi_i}{\partial s_i} \Bigg(d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\tilde{u}_i + \tilde{v}_i) \\ & - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\tilde{u}_j + \tilde{v}_j) \Bigg) + \frac{1}{2} \|\tilde{u}_i\|^2 - \frac{1}{2} \gamma_{\mathrm{ii}}^2 \|\tilde{v}_j\|^2 - \tilde{u}_i^T u_i^* \\ & + \gamma_{\mathrm{ii}}^2 \tilde{v}_i^T v_i^* + \frac{1}{2} \sum_{j \in \mathcal{N}_i} \Big(\|\tilde{u}_j\|^2 - \gamma_{\mathrm{ij}}^2 \|\tilde{v}_j\|^2 - 2\tilde{u}_j^T u_j^* + 2\gamma_{\mathrm{ij}}^2 \tilde{v}_j^T v_j^* \Big) \\ & - \frac{\partial \epsilon_i}{\partial s_i}^T \Bigg(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_i^* + v_i^*) \\ & - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_j^* + v_j^*) \Bigg) \Bigg], \quad \forall i \in \mathcal{N} \end{aligned}$$

as a perturbation*; and $\tilde{u}_i := u_i^* - \hat{u}_i$, $\tilde{v}_i := v_i^* - \hat{v}_i$ with $u_i^*,\ v_i^*,\ \hat{u}_i,\ \hat{v}_i$ given by (19), (20), (21), and (22), respectively.* □

Theorem 3.1 *Let the tuning of the critic be given by (25), for each agent $i \in \mathcal{N}$. Then the nominal system F_{i0} from (28) is globally exponentially stable with its trajectories satisfying $\|\tilde{W}_{\mathrm{c_i}}(t)\| \leqslant \|\tilde{W}_{\mathrm{c_i}}(t_0)\| \kappa_{i1} e^{-\kappa_{i2}(t-t_0)}$, for some $\kappa_{\mathrm{i1}},\ \kappa_{\mathrm{i2}} \in \mathbb{R}_+$, $\forall t > t_0 \geqslant 0$, provided that there exists a constant $T > 0$ such that $\bar{\omega}_i$ is persistently exciting (PE) over every interval of length T (i.e., that the matrix $\bar{\omega}_i \bar{\omega}_i^T$ is positive definite over any finite interval), in the sense that*

$$\int_t^{t+T} \bar{\omega}_i \bar{\omega}_i^T \geqslant \beta_i I, \ \forall i \in \mathcal{N}, \quad t \geqslant t_0$$

with $\beta_i \in \mathbb{R}_+$ and I the identity matrix of appropriate dimensions.

Proof of Theorem 3.1 Consider the following Lyapunov function, $\forall t \geqslant 0$

$$\mathcal{L}_i = \frac{1}{2\alpha_i} \tilde{W}_{\mathrm{c_i}}^T \tilde{W}_{\mathrm{c_i}}, \quad i \in \mathcal{N}, \tag{29}$$

and hence $\mathcal{L}_i$ is decrescent and radially unbounded in the space of $\tilde{W}_{\mathrm{c_i}}$.

By differentiating (29) along the error dynamics nominal system trajectories, one has

$$\dot{\mathcal{L}}_i = -\tilde{W}_{\mathrm{c_i}}^T \bar{\omega}_i \bar{\omega}_i^T \tilde{W}_{\mathrm{c_i}} \leqslant 0, \quad i \in \mathcal{N}.$$

Viewing the nominal system F_{i0} given in (28) as a linear time-varying system, the solution $\tilde{W}_{c_i}$ is given as (the reader is directed to [56] and [57] for the details)

$$\tilde{W}_{\mathrm{c_i}}(t) = \Phi_i(t, t_0) \tilde{W}_{\mathrm{c_i}}(t_0), \ i \in \mathcal{N}, \quad \forall t, t_0 > 0, \tag{30}$$

where the state transition matrix is defined as $\frac{\partial \Phi_i(t,t_0)}{\partial t} := -\alpha_i \bar{\omega}_i \bar{\omega}_i^T \Phi_i(t, t_0)$. We can prove, by following Theorem 1 from [56], that for the nominal system, the equilibrium point is exponentially stable provided that $\bar{\omega}_i$ is PE and therefore for some $\kappa_{\mathrm{i}1}, \ \kappa_{\mathrm{i}2} \in \mathbb{R}_+$ we can write

$$\|\Phi_i(t, t_0)\| \leqslant \kappa_{i1} e^{-\kappa_{i2}(t-t_0)}, \ i \in \mathcal{N}, \quad \forall t, t_0 > 0. \tag{31}$$

Finally, by combining (30) and (31) we have

$$\left\| \tilde{W}_{\mathrm{c_i}}(t) \right\| \leqslant \left\| \tilde{W}_{\mathrm{c_i}}(t_0) \right\| \kappa_{\mathrm{i}1} e^{-\kappa_{i2}(t-t_0)}, \ i \in \mathcal{N}, \quad \forall t, t_0 > 0,$$

from which the result follows. □

Remark 3.4 Note that the parameters $\kappa_{\mathrm{i}1}$ and $\kappa_{\mathrm{i}2}$ depend on the PE condition. The reader is directed to [56,57] for a more detailed explanation. Typically, one can ensure PE of the vector signal $\bar{\omega}_i$ by adding sinusoids of different frequencies to the inputs. This condition is equivalent to state space exploration in reinforcement learning [14] required for convergence to the optimal policies. For linear systems of order n one can guarantee PE by adding $\frac{n(n+1)}{2}$ different frequencies [57]. It has been shown in [15,57] (cf. Chapter 7) that the convergence rate of gradient descent algorithms of the above form converges exponentially fast with the rate that depends on the level of excitation β_i and the size of the time interval T.

The main theorem is presented next and provides a Lyapunov-based stability proof for the proposed game-theoretical learning controller.

Theorem 3.2 *Consider the dynamics given by* (5)*, the controller given by* (21)*, the adversary given by* (22)*, the critic tuning law given by* (25)*, the controller tuning law given by* (26)*, and the adversarial tuning law given by* (27)*. Assume that the coupled Bellman equations* (8) *have locally smooth solutions* $V_i^*(s_i), \forall s_i, \ \forall i \in \mathcal{N}$*. Assume that the signal* $\bar{\omega}_i$ *is persistently exciting, that Assumption* 3.1 *holds, and that* $d_i \neq 0$ *for all* $i \in \mathcal{N}$*. Then, the solution* $\big(s_i(t), \tilde{W}_{\mathrm{c_i}}(t), \tilde{W}_{\mathrm{u_i}}(t), \tilde{W}_{\mathrm{v_i}}(t)\big)$*, for all* $\big(s_i(0), \tilde{W}_{\mathrm{c_i}}(0), \tilde{W}_{\mathrm{u_i}}(0), \tilde{W}_{\mathrm{v_i}}(0)\big)$*,* $\forall i \in \mathcal{N}$*, is uniformly ultimately bounded (UUB).* □

Remark 3.5 By picking some parameters appropriately we can guarantee the asymptotic reduction in the tracking errors $s_i, \forall i \in \mathcal{N}$, and the actor–critic weight estimation errors to a small neighborhood of zero.

Remark 3.6 Although we are considering fixed topology graphs of a networked team, the proof will still go through if the team breaks in multiple disconnected subteams, provided that each agent retains at least one neighbor ($\mathcal{N}_i \neq \emptyset$). In this case, the subteams formed will reach separate consensus values.

Remark 3.7 An online learning framework such as the one in Theorem 3.2 allows the controller to adapt to changing conditions. Each agent might learn defense strategies online, while adversaries perform sequences of simple strategies. The learning algorithm is not known to the adversary, since the adversary can apply any policy. Now if the adversary wants to know which learning framework the controller uses, there are some analyses that could enable the adversary to learn the framework. Our Theorem 3.2 (and the Corollary that follows) shows that in order to achieve a proper approximation one has to use a large number of basis sets and hence if the controller and the controller's estimate of the worst-case adversary do not use a good approximation they will deviate from the Nash equilibrium. The learning framework would also allow the distributed performances to be changed on the fly. The fact of using the same basis sets in (21)–(22) is in order not to use more symbols. The real adversary can conceivably do anything. Finally, it will be difficult for the adversary to have any information regarding the type of approximation used by the controller, since this process could require a number of experimenting that could lead the adversary to perform expensive computations in the state space.

Corollary 3.1 *Suppose that the hypotheses and the statements of Theorem* 3.2 *hold. Then, the policies* $\hat{u}_i, \hat{v}_i, \forall i \in \mathcal{N}$*, form a Nash equilibrium as for a sufficiently large number of basis sets* h*.*

Proof First, we shall compute $\|u_i^* - \hat{u}_i\|$ by using (19), (21) and the fact that $\tilde{W}_{\mathrm{u_i}} = W_i - \hat{W}_{\mathrm{u_i}}$ as

$$\begin{aligned} \|\tilde{u}_i\| := \|u_i^* - \hat{u}_i\| &= \Bigg\| - d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \left(\frac{\partial \phi_i}{\partial s_i}(s_i)^T W_i + \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right) \\ &\quad + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{u_i} \Bigg\| \\ &= \left\| d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \left(\frac{\partial \phi_i}{\partial s_i}(s_i)^T \tilde{W}_{\mathrm{u_i}} + \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right) \right\|. \end{aligned}$$

Similarly for $\|v_i^* - \hat{v}_i\|$ by using (20), (22), and $\tilde{W}_{\mathrm{v_i}} = W_i - \hat{W}_{\mathrm{v_i}}$

$$\|\tilde{v}_i\| := \|v_i^* - \hat{v}_i\| = \left\| \frac{d_i}{\gamma_{ii}^2} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}^T \left(\frac{\partial \phi_i}{\partial s_i}(s_i)^T \tilde{W}_{\mathrm{v_i}} + \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i} \right) \right\|.$$

According to the result of Theorem 3.2, we have proved that $\tilde{W}_{\mathrm{u_i}}$ and $\tilde{W}_{\mathrm{v_i}}$ are UUB. Moreover, after using Assumption 3.1 we have upper bounds on $\frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i}$ and $\frac{\partial \phi_i}{\partial s_i}$ and thus all the quantities on the right-hand side of $\|\tilde{u}_i\|$ and $\|\tilde{v}_i\|$ are bounded. Then after following the results from [36] we can conclude that the approximated policies as one selects a large number of basis sets (c.f. Remark 3.8) satisfy (4). □

3.4 Proof of Theorem 3.2

Consider the following continuously differentiable positive definite candidate Lyapunov function $\mathcal{V}_{\mathcal{L}} : \mathbb{R}^{2m} \times \mathbb{R}^h \times \mathbb{R}^h \times \mathbb{R}^h \to \mathbb{R}$ defined as

$$\mathcal{V}_{\mathcal{L}} := \sum_{i=1}^{N} \mathcal{V}_{\mathcal{L}i}(s_i, \tilde{W}_{\mathrm{c_i}}, \tilde{W}_{\mathrm{u_i}}, \tilde{W}_{\mathrm{v_i}}),$$

with

$$\begin{aligned} \mathcal{V}_{\mathcal{L}i}(s_i, \tilde{W}_{\mathrm{c_i}}, \tilde{W}_{\mathrm{u_i}}, \tilde{W}_{\mathrm{v_i}}) := V_i^*(s_i) + V_{c_i}(\tilde{W}_{c_i}) \\ + \frac{\alpha_{ui}^{-1}}{2} \tilde{W}_{\mathrm{u_i}}^T \tilde{W}_{\mathrm{u_i}} + \frac{\alpha_{\mathrm{vi}}^{-1}}{2} \tilde{W}_{\mathrm{v_i}}^T \tilde{W}_{\mathrm{v_i}}, \end{aligned}$$

where $V_i^*(s_i)$ is the Lyapunov function for (5), (12) used in the proof of Theorem 2.1, $V_{c_i}(\tilde{W}_{\mathrm{c_i}}) := \left\| \tilde{W}_{\mathrm{c_i}} \right\|^2$, and α_{ui} and α_{vi} are positive scalars.

From the positive definiteness and radial unboundedness of the V_i^*, we conclude that there exist class $\mathcal{K}$ functions [48] $k_{\mathrm{i}1}$ and $k_{\mathrm{i}2}$ such that V_i^* satisfies

$$k_{\mathrm{i}1}(\|s_i\|) \leqslant V_i^*(s_i) \leqslant k_{\mathrm{i}2}(\|s_i\|), \ \forall s_i, \ i \in \mathcal{N}.$$

We can write

$$\begin{aligned} k_{\mathrm{i}1}(\|s_i\|) + \|\tilde{W}_{\mathrm{c_i}}\|^2 + \frac{1}{2\alpha_{ui}} \|\tilde{W}_{\mathrm{u_i}}\|^2 + \frac{1}{2\alpha_{\mathrm{vi}}} \|\tilde{W}_{\mathrm{v_i}}\|^2 \leqslant \mathcal{V}_{\mathcal{L}i} \\ \leqslant k_{\mathrm{i}2}(\|s_i\|) + \|\tilde{W}_{\mathrm{c_i}}\|^2 + \frac{1}{2\alpha_{ui}} \|\tilde{W}_{\mathrm{u_i}}\|^2 + \frac{1}{2\alpha_{\mathrm{vi}}} \|\tilde{W}_{\mathrm{v_i}}\|^2. \end{aligned}$$

Computing the time derivative of V_i^*, $\forall i \in \mathcal{N}$, along solutions to the closed-loop system (5) with controller $\hat{u}_i$ and adversary $\hat{v}_i$ and for V_{c_i} along the perturbed system (40) leads to

$$\begin{aligned}\dot{\mathcal{V}}_{\mathcal{L}i} = {} & \frac{\partial V_i^*}{\partial s_i}^T \left(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} s_i + d_i \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} (\hat{u}_i + \hat{v}_i) \right.\\ & \left. - \sum_{j\in\mathcal{N}_i} \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} (\hat{u}_j + \hat{v}_j) \right) + \frac{\partial V_{c_i}}{\partial \tilde{W}_{c_i}} \Big(F_{\mathrm{i}0} + F_{\mathrm{i}1} \Big)\\ & - \dot{\hat{W}}_{\mathrm{u_i}}^T \alpha_{ui}^{-1} \tilde{W}_{\mathrm{u_i}} - \dot{\hat{W}}_{\mathrm{v_i}}^T \alpha_{\mathrm{vi}}^{-1} \tilde{W}_{\mathrm{v_i}}. \end{aligned} \tag{32}$$

We can write the coupled Bellman equations (8) as

$$\begin{aligned} \frac{\partial V_i^*}{\partial s_i}^T \begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} s_i = {} & -\frac{\partial V_i^*}{\partial s_i}^T \left(d_i \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} (u_i^* + v_i^*) \right.\\ & \left. - \sum_{j\in\mathcal{N}_i} \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} (u_j^* + v_j^*) \right) - \frac{1}{2} \left(\|s_i\|^2 + \|u_i^*\|^2 + \sum_{j\in\mathcal{N}_i} \|u_j^*\|^2 \right.\\ & \left. - \gamma_{\mathrm{ii}}^2 \|v_i^*\|^2 - \sum_{j\in\mathcal{N}_i} \gamma_{\mathrm{ij}}^2 \|v_j^*\|^2 \right). \end{aligned} \tag{33}$$

Setting $\tilde{u}_i = u_i^* - \hat{u}_i$, $\tilde{v}_i = v_i^* - \hat{v}_i$, and using (26), (27), and (33), we can bound (32) as

$$\begin{aligned} \dot{\mathcal{V}}_{\mathcal{L}i} \leqslant {} & -\frac{\partial V_i^*}{\partial s_i}^T \left(d_i \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} (\tilde{u}_i + \tilde{v}_i) - \sum_{j\in\mathcal{N}_i} \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} (\tilde{u}_j + \tilde{v}_j) \right)\\ & - \frac{1}{2} \left(\|s_i\|^2 + \|u_i^*\|^2 + \sum_{j\in\mathcal{N}_i} \|u_j^*\|^2 - \gamma_{\mathrm{ii}}^2 \|v_i^*\|^2 - \sum_{j\in\mathcal{N}_i} \gamma_{\mathrm{ij}}^2 \|v_j^*\|^2 \right)\\ & - \alpha_i \left\| \tilde{W}_{\mathrm{c_i}} \right\|^2 + 2 \|\tilde{W}_{\mathrm{c_i}}\| \|F_{i1}\| + \tilde{W}_{\mathrm{u_i}}^T \left(\sigma_{ui} (\tilde{W}_{c_i} - \tilde{W}_{\mathrm{u_i}}) \right.\\ & - d_i^2 \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{\mathrm{u_i}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{\mathrm{c_i}}\\ & \left. - \sum_{j\in\mathcal{N}_i} d_j^2 \frac{\partial \phi_j}{\partial s_j} \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \frac{\partial \phi_j}{\partial s_j}^T \hat{W}_{\mathrm{u_j}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{\mathrm{c_j}} \right)\\ & + \tilde{W}_{\mathrm{v_i}}^T \left(\sigma_{\mathrm{vi}} (\tilde{W}_{\mathrm{c_i}} - \tilde{W}_{\mathrm{v_i}}) + \frac{d_i^2}{\gamma_{\mathrm{ii}}^2} \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T \hat{W}_{\mathrm{v_i}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{\mathrm{c_i}} \right.\\ & \left. + \sum_{j\in\mathcal{N}_i} \frac{d_j^2 \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4} \frac{\partial \phi_j}{\partial s_j} \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \frac{\partial \phi_j}{\partial s_j}^T \hat{W}_{\mathrm{v_j}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} \hat{W}_{c_j} \right). \end{aligned} \tag{34}$$

Using $-d_i \frac{\partial V_i^*}{\partial s_i}^T \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} = u_i^{*T}$ and $\frac{d_i}{\gamma_{\mathrm{ii}}^2} \frac{\partial V_i^*}{\partial s_i}^T \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix} = v_i^{*T}$, (34) can be written as

$$\begin{aligned}
\dot{\mathcal{V}}_{\mathcal{L}i} \leqslant\ & u_i^{*T}\tilde{u}_i - v_i^{*T}\tilde{v}_i - \frac{1}{d_i}u_i^{*T}\sum_{j\in\mathcal{N}_i}\tilde{u}_j + \frac{\gamma_{\mathrm{ii}}^2}{d_i}v_i^{*T}\sum_{j\in\mathcal{N}_i}\tilde{v}_j \\
& -\ \frac{1}{2}\Bigg(\|s_i\|^2 + \|u_i^*\|^2 + \sum_{j\in\mathcal{N}_i}\|u_j^*\|^2 - \gamma_{\mathrm{ii}}^2\|v_i^*\|^2 - \sum_{j\in\mathcal{N}_i}\gamma_{\mathrm{ij}}^2\|v_j^*\|^2\Bigg) \\
& -\ \alpha_i\left\|\tilde{W}_{c_i}\right\|^2 + 2\|\tilde{W}_{\mathrm{c_i}}\| \\
& \Bigg\|\Bigg[-W_i^T\frac{\partial\phi_i}{\partial s_i}\Bigg(d_i\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(\tilde{u}_i+\tilde{v}_i) - \sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(\tilde{u}_j+\tilde{v}_j)\Bigg) \\
& + \frac{1}{2}\|\tilde{u}_i\|^2 - \frac{1}{2}\gamma_{\mathrm{ii}}^2\|\tilde{v}_i\|^2 - \tilde{u}_i^T u_i^* + \gamma_{\mathrm{ii}}^2\tilde{v}_i^T v_i^* \\
& + \frac{1}{2}\sum_{j\in\mathcal{N}_i}\Bigg(\|\tilde{u}_j\|^2 - \gamma_{\mathrm{ij}}^2\|\tilde{v}_j\|^2 - \tilde{u}_j^T u_j^* + \gamma_{\mathrm{ij}}^2\tilde{v}_j^T v_j^*\Bigg) \\
& - \frac{\partial\epsilon_{\mathrm{c_i}}}{\partial s_i}^T\Bigg(\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}s_i + d_i\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_i^*+v_i^*) - \sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}(u_j^*+v_j^*)\Bigg)\Bigg]\Bigg\| \\
& +\ \tilde{W}_{\mathrm{u_i}}^T\Bigg(\sigma_{ui}(\tilde{W}_{c_i} - \tilde{W}_{\mathrm{u_i}}) \\
& -\ d_i^2\frac{\partial\phi_i}{\partial s_i}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial\phi_i}{\partial s_i}^T(W_i - \tilde{W}_{\mathrm{u_i}})\frac{\bar{\omega}_i^T}{(\omega_i^T\omega_i+1)}(W_i - \tilde{W}_{c_i}) \\
& -\ \sum_{j\in\mathcal{N}_i}d_j^2\frac{\partial\phi_j}{\partial s_j}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial\phi_j}{\partial s_j}^T(W_j - \tilde{W}_{\mathrm{u_j}})\frac{\bar{\omega}_i^T}{(\omega_i^T\omega_i+1)}(W_j - \tilde{W}_{c_j})\Bigg) \\
& + \tilde{W}_{\mathrm{v_i}}^T\Bigg(\sigma_{\mathrm{vi}}(\tilde{W}_{c_i} - \tilde{W}_{\mathrm{v_i}}) \\
& +\ \frac{d_i^2}{\gamma_{\mathrm{ii}}^2}\frac{\partial\phi_i}{\partial s_i}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial\phi_i}{\partial s_i}^T(W_i - \tilde{W}_{\mathrm{v_i}})\frac{\bar{\omega}_i^T}{(\omega_i^T\omega_i+1)}(W_i - \tilde{W}_{c_i}) \\
& +\ \sum_{j\in\mathcal{N}_i}\frac{d_j^2\gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4}\frac{\partial\phi_j}{\partial s_j}\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\frac{\partial\phi_j}{\partial s_j}^T(W_j - \tilde{W}_{\mathrm{v_j}}) \\
& \frac{\bar{\omega}_i^T}{(\omega_i^T\omega_i+1)}(W_j - \tilde{W}_{c_j})\Bigg).
\end{aligned} \tag{35}$$

Using Fact A in “Appendix,” and expanding the parentheses we can further upper-bound (35) as

$$\begin{aligned}
\dot{\mathcal{V}}_{\mathcal{L}i} \leqslant & -\sigma_{ui}\|\tilde{W}_{\mathrm{u_i}}\|^2 - \sigma_{\mathrm{vi}}\|\tilde{W}_{\mathrm{v_i}}\|^2 - \frac{1}{2}\|s_i\|^2 \\
& -\ \alpha_i\left\|\tilde{W}_{\mathrm{c_i}}\right\|^2 + b_{i5} + \Bigg(2b_{\mathrm{i4}} + \sigma_{\mathrm{ui}}b_{\mathrm{i2}}
\end{aligned}$$

$$
\begin{aligned}
&+\sigma_{\mathrm{vi}} b_{\mathrm{i}3}+\omega_{\mathrm{imax}} b_{\mathrm{i}2} b_{\mathrm{i}1} d_i^2 W_{\mathrm{imax}} \\
&+\omega_{\mathrm{imax}} b_{\mathrm{i}1} b_{\mathrm{i}2}^2 d_i^2+\frac{\omega_{\mathrm{imax}} b_{\mathrm{i}1} b_{\mathrm{i}3} d_i^2}{\gamma_{\mathrm{ii}}^2} W_{\mathrm{imax}} \\
&\left.+\frac{\omega_{\mathrm{imax}} b_{\mathrm{i}1} b_{\mathrm{i}3}^2 d_i^2}{\gamma_{\mathrm{ii}}^2}\right)\|\tilde{W}_{c_i}\|+\sum_{j \in \mathcal{N}_i}\left(\omega_{\mathrm{imax}} b_{\mathrm{i}2} d_j^2 b_{j1} W_{\mathrm{jmax}}\right. \\
&+\omega_{\mathrm{imax}} b_{\mathrm{i}2} d_j^2 b_{j1} b_{\mathrm{j}2}+\omega_{\mathrm{imax}} b_{\mathrm{i}3} \frac{d_j^2 b_{j1} \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4} W_{\mathrm{jmax}} \\
&\left.+\omega_{\mathrm{imax}} b_{\mathrm{i}3} \frac{b_{\mathrm{j}1} b_{\mathrm{j}3} d_j^2 \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4}\right)\|\tilde{W}_{c_j}\|. \qquad (36)
\end{aligned}
$$

Summing the last term in the equation above over all agents, we conclude that

$$
\begin{aligned}
&\sum_{i=1}^{N}\Bigg(\sum_{j \in \mathcal{N}_i}\Big(\omega_{\mathrm{imax}} b_{\mathrm{i}2} d_j^2 b_{j1} W_{\mathrm{jmax}} \\
&\quad+\omega_{\mathrm{imax}} b_{\mathrm{i}2} d_j^2 b_{j1} b_{\mathrm{j}2}+\omega_{\mathrm{imax}} b_{\mathrm{i}3} \frac{d_j^2 b_{j1} \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4} W_{\mathrm{jmax}} \\
&\quad+\omega_{\mathrm{imax}} b_{\mathrm{i}3} \frac{b_{j1} b_{\mathrm{j}3} d_j^2 \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4}\Big)\|\tilde{W}_{c_j}\|\Bigg) \leqslant \sum_{i=1}^{N} b_{i7}\|\tilde{W}_{c_i}\|
\end{aligned}
$$

for sufficiently large constants b_{i7}. Completing the square in (36), we thus obtain

$$
\begin{aligned}
\dot{\mathcal{V}}_{\mathcal{L}} &:= \sum_{i=1}^{N} \dot{\mathcal{V}}_{\mathcal{L}i} \leqslant \sum_{i=1}^{N}\Big(-\sigma_{\mathrm{ui}}\|\tilde{W}_{\mathrm{u_i}}\|^2 \\
&-\sigma_{\mathrm{vi}}\|\tilde{W}_{\mathrm{v_i}}\|^2-\frac{1}{2}\|s_i\|^2-(1-\delta)\alpha_i\|\tilde{W}_{c_i}\|^2+\mu_i\Big), \qquad (37)
\end{aligned}
$$

where $\delta \in (0, 1)$ and

$$
\begin{aligned}
\mu_i &:= b_{i5}+\frac{1}{4\alpha_i\delta}\big(2b_{\mathrm{i}4}+\sigma_{\mathrm{ui}} b_{\mathrm{i}2}+\sigma_{\mathrm{vi}} b_{\mathrm{i}3}+\omega_{\mathrm{imax}} b_{\mathrm{i}2} b_{\mathrm{i}1} d_i^2 W_{\mathrm{imax}} \\
&+\omega_{\mathrm{imax}} b_{\mathrm{i}1} b_{\mathrm{i}2}^2 d_i^2+\frac{\omega_{\mathrm{imax}} b_{\mathrm{i}1} b_{\mathrm{i}3} d_i^2}{\gamma_{\mathrm{ii}}^2} W_{\mathrm{imax}}+\frac{\omega_{\mathrm{imax}} b_{\mathrm{i}1} b_{\mathrm{i}3}^2 d_i^2}{\gamma_{\mathrm{ii}}^2}+b_{i7}\big)^2. \qquad (38)
\end{aligned}
$$

According to Lemma 4.3 in [48] and by defining $\tilde{Q}_i(t) := [s_i(t)\ \tilde{W}_{\mathrm{c_i}}(t)\ \tilde{W}_{\mathrm{u_i}}(t)\ \tilde{W}_{\mathrm{v_i}}(t)]$, there exist class $\mathcal{K}$ functions $k_{\mathrm{i}5}$, $k_{\mathrm{i}6}$ such that

$$
\begin{aligned}
k_{\mathrm{i}5}(\|\tilde{Q}_i\|) &\leqslant \sigma_{\mathrm{ui}}\|\tilde{W}_{\mathrm{u_i}}\|^2+\sigma_{\mathrm{vi}}\|\tilde{W}_{\mathrm{v_i}}\|^2+\tfrac{1}{2}\|s_i\|^2 \\
&+(1-\delta)\alpha_i\|\tilde{W}_{\mathrm{c_i}}\|^2 \leqslant k_{\mathrm{i}6}(\|\tilde{Q}_i\|),
\end{aligned}
$$

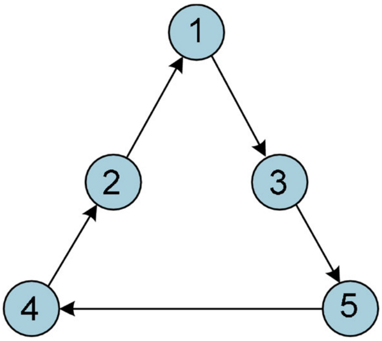

Fig. 2 Graph $\mathcal{G}_1$ describing the topology of a network of 5 agents

which can be used to further upper-bound (37):

$$\dot{\mathcal{V}_{\mathcal{L}}} \leqslant \sum_{i=1}^{N} \Big(- k_{i5}(\|\tilde{Q}_i\|) + \mu_i \Big).$$

From this, we conclude that $\dot{\mathcal{V}_{\mathcal{L}}} < 0$ whenever the combined state $(\tilde{Q}_1, \tilde{Q}_2, \ldots, \tilde{Q}_N)$ lies outside the compact set

$$\Omega_{\tilde{Q}} = \Big\{(\tilde{Q}_1, \tilde{Q}_2, \ldots, \tilde{Q}_N) : \sum_{i=1}^{N} k_{i5}(\|\tilde{Q}_i\|) \leqslant \sum_{i=1}^{N} \mu_i \Big\}.$$

Therefore, there exists a T_u such that for all $t \geqslant T_u$ the state $(\tilde{Q}_1, \tilde{Q}_2, \ldots, \tilde{Q}_N)$ remains inside $\Omega_{\tilde{Q}}$, from which uniformly ultimately boundedness (UUB) follows [48]. □

Remark 3.8 One can make the set $\Omega_{\tilde{Q}}$ small by decreasing the values of the μ_i in (38). To accomplish this, one can select a large value for the critic tuning gain α_i [which decreases the second term in (38)] and select a large number h of basis functions to decrease the errors $\epsilon_{\mathrm{c_i}}$ and consequently the constant $b_{\mathrm{i}5}$ in (38) (see Fact A in "Appendix").

4 Simulation Results

We now illustrate the theoretical developments for a graph $\mathcal{G}_1$ with 5 agents networked as shown in Fig. 2. The gains in the tuning laws are selected as $\sigma_{\mathrm{ui}} = 5, \sigma_{\mathrm{vi}} = 5, \alpha_i = 10, \alpha_{\mathrm{ui}} = 1, \alpha_{\mathrm{vi}} = 1, \forall i \in \mathcal{N}$, and all the weights of the tuning laws are initialized randomly inside $[-1, 1]$. Although Assumption 3.1 requires bounded basis functions, in practice one could pick these functions to be quadratic forms, e.g., $s_i \otimes s_i$, and still obtain satisfactory results. To ensure PE, a probing signal $\rho(t) = \tanh(t)\big(\sin(2t)\cos(t) + 4\sin(t^2) + 10\cos(5t)\sin(11t)\big)$ is added to the input signals for the first 2 s of the simulation.

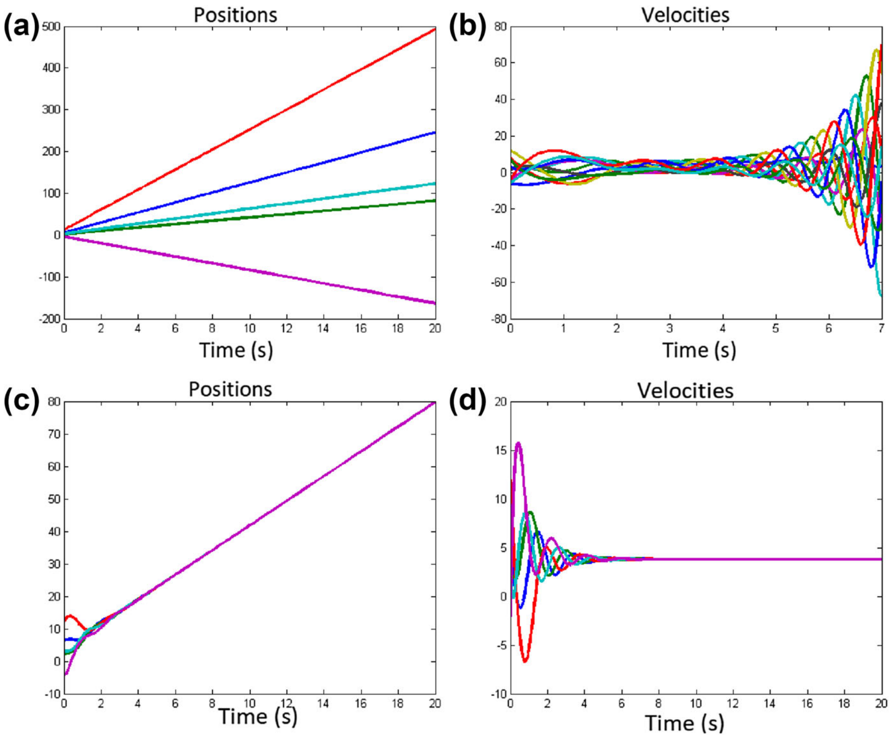

Fig. 3 Agents perturbed by adversaries connected as in $\mathcal{G}_1$ moving in a line. **a** Adversary adds *constant biases*. **b** Adversary adds *time-varying biases*. **c** Positions with the proposed algorithm. **d** Velocities with the proposed algorithm

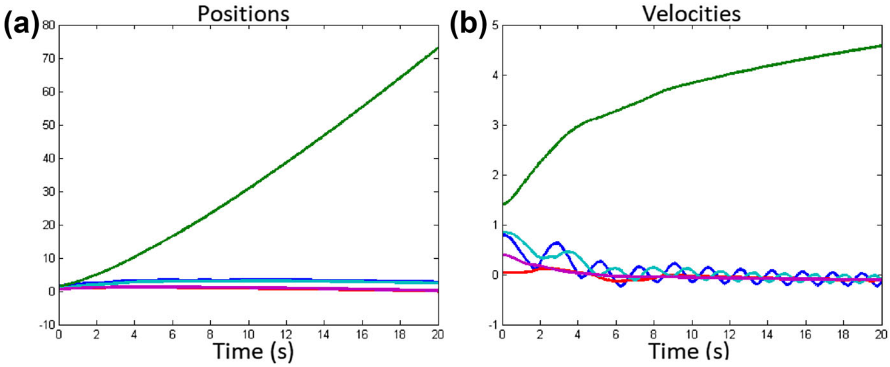

Fig. 4 The proposed algorithm when the agents are perturbed by adversaries that use the complete quadratic basis sets while the controller's adversarial estimate uses a subset of them, connected as in $\mathcal{G}_1$ moving in line. **a** Positions when the adversaries use the complete quadratic basis sets while the controller's adversarial estimate uses a subset of them. **b** Velocities when the adversaries use the complete quadratic basis sets while the controller's adversarial estimate uses a subset of them

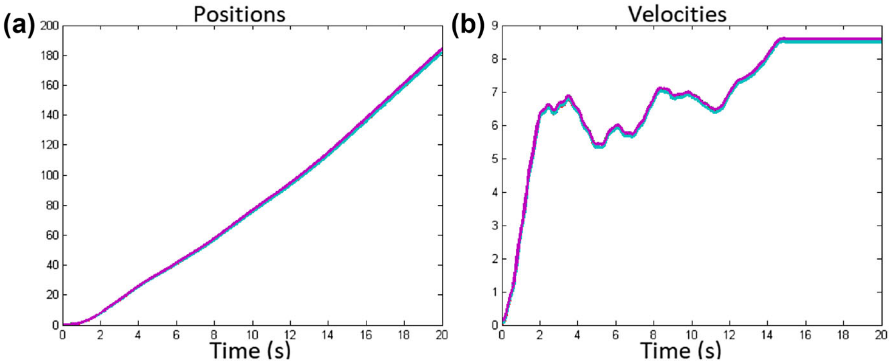

Fig. 5 The proposed algorithm when the agents are perturbed by adversaries that do not let the agents reach a steady-state velocity value, connected as in $\mathcal{G}_1$ moving in line. **a** Positions with the proposed framework while the intelligent adversaries do not let the agents reach a steady-state velocity value. **b** Velocities with the proposed framework, but the intelligent adversaries also do not let the agents reach a steady-state velocity value

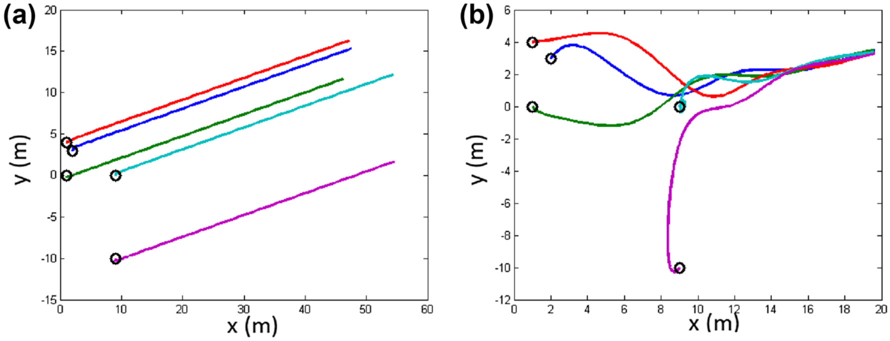

Fig. 6 Agents perturbed by adversaries, connected as in $\mathcal{G}_1$ and moving in the bi-dimensional space. **a** Without the proposed algorithm. **b** With the proposed algorithm

We start by considering agents that move on a line ($m = 1$) with random initial conditions.

Figure 3a, b shows a simulation of the same consensus algorithm [6], but with two different attacking scenarios where a constant bias and a time-varying bias is added to the measurements, which can represent two very basic forms of attacks. The simulation shows that even these simple attacks can result in an unstable behavior. This is in contrast to what can be observed in Fig. 3c, d, that shows the behavior of the algorithm proposed in this paper (with $\gamma_{\text{ii}} = 5, \gamma_{\text{ij}} = 1, \forall i \neq j \in \mathcal{N}$) which results in position and velocity agreement. We will now consider a prototypical case that will test our algorithm to the limits. In this case, the actual adversary uses quadratic basis functions of the form $s_i \otimes s_i$, while the controller's adversarial estimate uses a subset of those basis functions. For the latter case, Fig. 4a, b shows that the actual adversary is able to successfully diverge one agent and oscillate the velocities of the rest. It will be worth noting that when the controller does not use a good set of basis functions, the performance will deviate. On the other hand, another interesting example that is more probable should be when the actual adversary uses less basis functions than the controller's adversarial estimate, then the optimal consensus looks similar to Fig. 3c, d. Another interesting scenario is when the adversaries and the controllers play their Nash solution, but the adversaries also do not let the agents agree on a common velocity value. This case is shown in Fig. 5a, b.

We shall now consider agents moving in the plane ($m = 2$) with random initial conditions. Figure 6a illustrates how even a simple adversarial input consisting of adding a constant bias to the measurements can lead to instability for the regular consensus algorithm. Figure 6b shows that, also here, the algorithm proposed in this paper (with $\gamma_{\mathrm{ii}} = 5$, $\gamma_{\mathrm{ij}} = 1$, $\forall i \neq j \in \mathcal{N}$) results in an asymptotic consensus.

5 Conclusions

We have derived a game-theoretical actor–critic algorithm for reaching consensus of multi-agent systems with guaranteed performance when the dynamics are perturbed by persistent adversaries. It was shown that the proposed architecture is able to optimally reject adversarial inputs using a distributed algorithm. The optimization is performed online and results in agreement among all the agents in the team.

Three approximators are used for each agent: one to approximate the optimal value function, another to approximate the optimal controller, and a final one to approximate the adversary. Each of these approximators uses a specially designed tuning law. A Lyapunov-based argument is used to ensure that all the signals remain bounded. Simulation results show the effectiveness of the proposed approach. This work was focused on adversarial inputs that involve measurement corruption. More sophisticated adversarial inputs that utilize multiple points or multiple methods are important problems for future research. Other problems of interest include coordinated actions, where multiple adversaries cooperate to maximize damage, as well as disguised attacks, where the adversary can mask its actions to induce an erroneous reaction for mitigation.

Acknowledgements This material is based upon work supported in part by NATO under Grant No. SPS G5176, by ONR Minerva under Grant No. N00014-18-1-2160, by an NSF CAREER, by NAWCAD under Grant No. N00421-16-2-0001, and by an US Office of Naval Research MURI Grant No. N00014-16-1-2710.

Appendix

Proof of Lemma 3.1 The error in Eq. (23) after subtracting zero becomes

$$\begin{aligned} e_i = {} & \hat{W}_{\mathrm{c_i}}^T \frac{\partial \phi_i}{\partial s_i} \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\hat{u}_i + \hat{v}_i) - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\hat{u}_j + \hat{v}_j) \right) \\ & - W_i^T \frac{\partial \phi_i}{\partial s_i} \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_i^* + v_i^*) - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_j^* + v_j^*) \right) \\ & + \frac{1}{2} \left(\|\hat{u}_i\|^2 + \sum_{j \in \mathcal{N}_i} \|\hat{u}_j\|^2 - \gamma_{\mathrm{ii}}^2 \|v_i\|^2 - \sum_{j \in \mathcal{N}_i} \gamma_{\mathrm{ij}}^2 \|\hat{v}_j\|^2 \right) \end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{2}\Bigg(\|u_i^*\|^2+\sum_{j\in\mathcal{N}_i}\|u_j^*\|^2-\gamma_{ii}^2\|v_i^*\|^2-\sum_{j\in\mathcal{N}_i}\gamma_{ij}^2\|v_j^*\|^2\Bigg)\\
&-\frac{\partial\epsilon_{\mathrm{c_i}}}{\partial s_i}^T\Bigg(\begin{bmatrix}\mathbf{0}&\mathbf{I}\\\mathbf{0}&\mathbf{0}\end{bmatrix}s_i+d_i\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(u_i^*+v_i^*)\\
&-\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(u_j^*+v_j^*)\Bigg),\ \forall i\in N.
\end{aligned}
$$

Completing the squares we have

$$
\begin{aligned}
e_i=&-\tilde{W}_{\mathrm{c_i}}^T\omega_i-W_i^T\frac{\partial\phi_i}{\partial s_i}\Bigg(d_i\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(\tilde{u}_i+\tilde{v}_i)-\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(\tilde{u}_j+\tilde{v}_j)\Bigg)\\
&+\frac{1}{2}\|\tilde{u}_i\|^2-\frac{1}{2}\gamma_{\mathrm{ii}}^2\|\tilde{v}_j\|^2-\tilde{u}_i^Tu_i^*+\gamma_{\mathrm{ii}}^2\tilde{v}_i^Tv_i^*\\
&+\frac{1}{2}\sum_{j\in\mathcal{N}_i}\Big(\|\tilde{u}_j\|^2-\gamma_{\mathrm{ij}}^2\|\tilde{v}_j\|^2-2\tilde{u}_j^Tu_j^*+2\gamma_{\mathrm{ij}}^2\tilde{v}_j^Tv_j^*\Big)\\
&-\frac{\partial\epsilon_{\mathrm{c_i}}}{\partial s_i}^T\Bigg(\begin{bmatrix}\mathbf{0}&\mathbf{I}\\\mathbf{0}&\mathbf{0}\end{bmatrix}s_i+d_i\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(u_i^*+v_i^*)\\
&-\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(u_j^*+v_j^*)\Bigg),\ \forall i\in N.
\end{aligned}
\tag{39}
$$

The dynamics of the critic estimation error $\tilde{W}_{c_i}$ can be found by substituting (39) in (25) as

$$
\begin{aligned}
\dot{\tilde{W}}_{\mathrm{c_i}}=&-\alpha_i\bar{\omega}_i\bar{\omega}_i^T\tilde{W}_{c_i}\\
&+\alpha_i\frac{\omega_i}{(\omega_i^T\omega_i+1)^2}\Bigg[-W_i^T\frac{\partial\phi_i}{\partial s_i}\Bigg(d_i\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(\tilde{u}_i+\tilde{v}_i)\\
&-\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(\tilde{u}_j+\tilde{v}_j)\Bigg)+\frac{1}{2}\|\tilde{u}_i\|^2-\frac{1}{2}\gamma_{\mathrm{ii}}^2\|\tilde{v}_j\|^2-\tilde{u}_i^Tu_i^*+\gamma_{\mathrm{ii}}^2\tilde{v}_i^Tv_i^*\\
&+\frac{1}{2}\sum_{j\in\mathcal{N}_i}\Big(\|\tilde{u}_j\|^2-2\gamma_{\mathrm{ij}}^2\|\tilde{v}_j\|^2-2\tilde{u}_j^Tu_j^*+2\gamma_{\mathrm{ij}}^2\tilde{v}_j^Tv_j^*\Big)\\
&-\frac{\partial\epsilon_{\mathrm{c_i}}}{\partial s_i}^T\Bigg(\begin{bmatrix}\mathbf{0}&\mathbf{I}\\\mathbf{0}&\mathbf{0}\end{bmatrix}s_i+d_i\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(u_i^*+v_i^*)\\
&-\sum_{j\in\mathcal{N}_i}\begin{bmatrix}\mathbf{0}\\\mathbf{I}\end{bmatrix}(u_j^*+v_j^*)\Bigg)\Bigg],\ \forall i\in\mathcal{N},
\end{aligned}
\tag{40}
$$

from which the result follows. □

The following fact is a consequence of Assumption 3.1 and the properties of the projection operator $\mathcal{P}r[.]$.

Fact A There exist constants $b_{\mathrm{i1}}, b_{\mathrm{i2}}, b_{\mathrm{i3}}, b_{\mathrm{i4}}, b_{\mathrm{i5}} \in \mathbb{R}_+$ for which the following bounds hold, for every agent $\forall i \in \mathcal{N}$ and time $\forall t \geqslant 0$:

$$\left\| \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T \right\| \leqslant b_{\mathrm{i1}}, \quad \|\tilde{W}_{\mathrm{u_i}}\| \leqslant b_{\mathrm{i2}}, \quad \|\tilde{W}_{\mathrm{v_i}}\| \leqslant b_{\mathrm{i3}},$$

$$\begin{aligned}
&\Bigg\| -W_i^T \frac{\partial \phi_i}{\partial s_i} \left(d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\tilde{u}_i + \tilde{v}_i) - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (\tilde{u}_j + \tilde{v}_j) \right) \\
&\quad + \frac{1}{2} \|\tilde{u}_i\|^2 - \frac{1}{2} \gamma_{\mathrm{ii}}^2 \|\tilde{v}_i\|^2 - \tilde{u}_i^T u_i^* + \gamma_{\mathrm{ii}}^2 \tilde{v}_i^T v_i^* \\
&\quad + \frac{1}{2} \sum_{j \in \mathcal{N}_i} \left(\|\tilde{u}_j\|^2 - \gamma_{\mathrm{ij}}^2 \|\tilde{v}_j\|^2 - 2\tilde{u}_j^T u_j^* + 2\gamma_{\mathrm{ij}}^2 \tilde{v}_j^T v_j^* \right) - \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i}^T \\
&\left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} s_i + d_i \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_i^* + v_i^*) - \sum_{j \in \mathcal{N}_i} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} (u_j^* + v_j^*) \right) \Bigg\| \leqslant b_{\mathrm{i4}}
\end{aligned} \tag{41}$$

$$\begin{aligned}
&\Bigg\| u_i^{*T} \tilde{u}_i - v_i^{*T} \tilde{v}_i - \frac{u_i^{*T}}{d_i} \sum_{j \in \mathcal{N}_i} \tilde{u}_j + \frac{\gamma_{\mathrm{ii}}^2}{d_i} v_i^{*T} \sum_{j \in \mathcal{N}_i} \tilde{v}_j \\
&\quad - \frac{1}{2} \left(\|u_i^*\|^2 + \sum_{j \in \mathcal{N}_i} \|u_j^*\|^2 - \gamma_{\mathrm{ii}}^2 \|v_i^*\|^2 - \sum_{j \in \mathcal{N}_i} \gamma_{\mathrm{ij}}^2 \|v_j^*\|^2 \right) \\
&\quad - \tilde{W}_{\mathrm{u_i}}^T \Bigg(d_i^2 \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T W_i \frac{\bar{\omega}_i^T W_i}{(\omega_i^T \omega_i + 1)} \\
&\quad - d_i^2 \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T \tilde{W}_{\mathrm{u_i}} \frac{\bar{\omega}_i^T W_i}{(\omega_i^T \omega_i + 1)} \\
&\quad + \sum_{j \in \mathcal{N}_i} d_j^2 \frac{\partial \phi_j}{\partial s_j} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_j}{\partial s_j}^T W_j \frac{\bar{\omega}_i^T W_j}{(\omega_i^T \omega_i + 1)} \\
&\quad - \sum_{j \in \mathcal{N}_i} d_j^2 \frac{\partial \phi_j}{\partial s_j} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_j}{\partial s_j}^T \tilde{W}_{\mathrm{u_j}} \frac{\bar{\omega}_i^T W_j}{(\omega_i^T \omega_i + 1)} \Bigg) \\
&\quad + \tilde{W}_{\mathrm{v_i}}^T \Bigg(\frac{d_i^2}{\gamma_{\mathrm{ii}}^2} \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T W_i \frac{\bar{\omega}_i^T W_i}{(\omega_i^T \omega_i + 1)} \\
&\quad - \frac{d_i^2}{\gamma_{\mathrm{ii}}^2} \frac{\partial \phi_i}{\partial s_i} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_i}{\partial s_i}^T \tilde{W}_{\mathrm{v_i}} \frac{\bar{\omega}_i^T}{(\omega_i^T \omega_i + 1)} W_i \\
&\quad + \sum_{j \in \mathcal{N}_i} \frac{d_j^2 \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4} \frac{\partial \phi_j}{\partial s_j} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_j}{\partial s_j}^T W_j \frac{\bar{\omega}_i^T W_j}{(\omega_i^T \omega_i + 1)} \\
&\quad - \sum_{j \in \mathcal{N}_i} \frac{d_j^2 \gamma_{\mathrm{ij}}^2}{\gamma_{\mathrm{jj}}^4} \frac{\partial \phi_j}{\partial s_j} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \frac{\partial \phi_j}{\partial s_j}^T \tilde{W}_{\mathrm{v_j}} \frac{\bar{\omega}_i^T W_j}{(\omega_i^T \omega_i + 1)} \Bigg) \Bigg\| \leqslant b_{\mathrm{i5}},
\end{aligned} \tag{42}$$

where $\tilde{u}_i = -d_i \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}^T (\frac{\partial \phi_i}{\partial s_i}^T \tilde{W}_{\mathrm{u_i}} + \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i})$, $\tilde{v}_i = \frac{d_i}{\gamma_{\mathrm{ii}}^2} \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}^T (\frac{\partial \phi_i}{\partial s_i}^T \tilde{W}_{\mathrm{V_i}} + \frac{\partial \epsilon_{\mathrm{c_i}}}{\partial s_i})$, and u_i^*, v_i^*, $\hat{u}_i$, $\hat{v}_i$ are given by (19), (20), (21), and (22), respectively. Inequalities (41) and (42) result from the fact that all the quantities that appear in the left-hand side have known definable bounds. Also all the residual errors $\epsilon_{\mathrm{c_i}}$ that appear in u_i^* [see (19)], v_i^* [see (20)], and $\tilde{u}_i$ and $\tilde{v}_i$ (as shown above) can be reduced by increasing the number of basis functions. □

References

1. Teixeira, A., Sandberg, H., Johansson, K.H.: Networked control systems under cyber attacks with applications to power networks. In: American Control Conference (ACC), 2010, pp. 3690–3696. IEEE (2010)
2. Vamvoudakis, K.G., Hespanha, J.P.: Online optimal operation of parallel voltage-source inverters using partial information. IEEE Trans. Ind. Electron. **64**(5), 4296–4305 (2017)
3. Beard, R.W., McLain, T.W., Nelson, D.B., Kingston, D., Johanson, D.: Decentralized cooperative aerial surveillance using fixed-wing miniature UAVs. Proc. IEEE **94**(7), 1306–1324 (2006)
4. Kunwar, F., Benhabib, B.: Rendezvous-guidance trajectory planning for robotic dynamic obstacle avoidance and interception. IEEE Trans. Syst. Man Cybern. Part B Cybern. **36**(6), 1432–1441 (2006)
5. Lee, D., Spong, M.W.: Stable flocking of multiple inertial agents on balanced graphs. IEEE Trans. Autom. Control **52**(8), 1469–1475 (2007)
6. Olfati-Saber, R., Fax, J.A., Murray, R.M.: Consensus and cooperation in networked multi-agent systems. Proc. IEEE **95**(1), 215–233 (2007)
7. Jadbabaie, A., Lin, J., Morse, A.S.: Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans. Autom. Control **48**(6), 988–1001 (2003)
8. Ren, W., Beard, R.W., Atkins, E.M.: A survey of consensus problems in multi-agent coordination. In: American Control Conference, 2005. Proceedings of the 2005, pp. 1859–1864. IEEE (2005)
9. Tsitsiklis, J.N.: Problems in decentralized decision making and computation. Tech. rep., MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMATION AND DECISION SYSTEMS (1984)
10. Abbeel, P., Coates, A., Ng, A.Y.: Autonomous helicopter aerobatics through apprenticeship learning. Int. J. Robot. Res. **29**(13), 1608–1639 (2010)
11. Shoham, Y., Leyton-Brown, K.: Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations. Cambridge University Press, Cambridge (2008)
12. Vamvoudakis, K.G., Antsaklis, P.J., Dixon, W.E., Hespanha, J.P., Lewis, F.L., Modares, H., Kiumarsi, B.: Autonomy and machine intelligence in complex systems: a tutorial. In: American Control Conference (ACC), 2015, pp. 5062–5079. IEEE (2015)
13. Vamvoudakis, K.G., Modares, H., Kiumarsi, B., Lewis, F.L.: Game theory-based control system algorithms with real-time reinforcement learning: how to solve multiplayer games online. IEEE Control Syst. **37**(1), 33–52 (2017)
14. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction, vol. 1. MIT Press, Cambridge (1998)
15. Vrabie, D., Vamvoudakis, K.G., Lewis, F.L.: Optimal adaptive control and differential games by reinforcement learning principles, vol. 2. IET (2013)
16. Werbos, P.J.: Approximate dynamic programming for real-time control and neural modeling. In: White, D.A., Sofge, D.A. (eds.) Handbook of Intelligent Control. Van Nostrand Reinhold, New York (1992)
17. Bertsekas, D.P., Tsitsiklis, J.N.: Neuro-dynamic programming: an overview. In: Proceedings of the 34th IEEE Conference on Decision and Control, vol. 1, pp. 560–564. IEEE (1995)
18. Cardenas, A., Amin, S., Sinopoli, B., Giani, A., Perrig, A., Sastry, S., et al.: Challenges for securing cyber physical systems. In: Workshop on Future Directions in Cyber-Physical Systems Security, vol. 5 (2009)
19. Cardenas, A.A., Amin, S., Sastry, S.: Secure control: towards survivable cyber-physical systems. In: 28th International Conference on Distributed Computing Systems Workshops, 2008. ICDCS'08, pp. 495–500. IEEE (2008)

20. Pasqualetti, F., Bicchi, A., Bullo, F.: Consensus computation in unreliable networks: a system theoretic approach. IEEE Trans. Autom. Control **57**(1), 90–104 (2012)
21. Alpcan, T., Başar, T.: Network Security: A Decision and Game-Theoretic Approach. Cambridge University Press, Cambridge (2010)
22. Basar, T., Olsder, G.J.: Dynamic noncooperative game theory, vol. 23. Siam (1999)
23. Vamvoudakis, K.G., Hespanha, J.P., Sinopoli, B., Mo, Y.: Detection in adversarial environments. IEEE Trans. Autom. Control **59**(12), 3209–3223 (2014)
24. Holmgren, A.J., Jenelius, E., Westin, J.: Evaluating strategies for defending electric power networks against antagonistic attacks. IEEE Trans. Power Syst. **22**(1), 76–84 (2007)
25. Wang, J., Elia, N.: Distributed averaging algorithms resilient to communication noise and dropouts. IEEE Trans. Signal Process. **61**(9), 2231–2242 (2013)
26. Zhu, M., Martínez, S.: Attack-resilient distributed formation control via online adaptation. In: 2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), pp. 6624–6629. IEEE (2011)
27. Chung, S.J., Slotine, J.J.E.: Cooperative robot control and concurrent synchronization of lagrangian systems. IEEE Trans. Rob. **25**(3), 686–700 (2009)
28. LeBlanc, H.J., Koutsoukos, X.D.: Low complexity resilient consensus in networked multi-agent systems with adversaries. In: Proceedings of the 15th ACM International Conference on Hybrid Systems: Computation and Control, pp. 5–14. ACM (2012)
29. Semsar, E., Khorasani, K.: Optimal control and game theoretic approaches to cooperative control of a team of multi-vehicle unmanned systems. In: 2007 IEEE International Conference on Networking, Sensing and Control, pp. 628–633. IEEE (2007)
30. Semsar-Kazerooni, E., Khorasani, K.: An lmi approach to optimal consensus seeking in multi-agent systems. In: American Control Conference, 2009. ACC'09., pp. 4519–4524. IEEE (2009)
31. Khanafer, A., Touri, B., Başar, T.: Consensus in the presence of an adversary. IFAC Proc. Vol. **45**(26), 276–281 (2012)
32. Bauso, D., Giarre, L., Pesenti, R.: Mechanism design for optimal consensus problems. In: 2006 45th IEEE Conference on Decision and Control, pp. 3381–3386. IEEE (2006)
33. Chung, S.J., Bandyopadhyay, S., Chang, I., Hadaegh, F.Y.: Phase synchronization control of complex networks of lagrangian systems on adaptive digraphs. Automatica **49**(5), 1148–1161 (2013)
34. Carli, R., Zampieri, S.: Networked clock synchronization based on second order linear consensus algorithms. In: 2010 49th IEEE Conference on Decision and Control (CDC), pp. 7259–7264. IEEE (2010)
35. Yucelen, T., Egerstedt, M.: Control of multiagent systems under persistent disturbances. In: American Control Conference (ACC), 2012, pp. 5264–5269. IEEE (2012)
36. Vamvoudakis, K.G., Lewis, F.L., Hudas, G.R.: Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality. Automatica **48**(8), 1598–1611 (2012)
37. Sundaram, S., Hadjicostis, C.N.: Distributed function calculation via linear iterative strategies in the presence of malicious agents. IEEE Trans. Autom. Control **56**(7), 1495–1508 (2011)
38. Zhu, Q., Bushnell, L., Başar, T.: Resilient distributed control of multi-agent cyber-physical systems. In: Control of Cyber-Physical Systems, pp. 301–316. Springer (2013)
39. Chen, L., Roy, S., Saberi, A.: On the information flow required for tracking control in networks of mobile sensing agents. IEEE Trans. Mob. Comput. **10**(4), 519–531 (2011)
40. Peymani, E., Grip, H.F., Saberi, A., Wang, X., Fossen, T.I.: H-∞ almost output synchronization for heterogeneous networks of introspective agents under external disturbances. Automatica **50**(4), 1026–1036 (2014)
41. Bardi, M., Capuzzo-Dolcetta, I.: Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations. Springer, Berlin (2008)
42. Crandall, M.G., Lions, P.L.: Viscosity solutions of Hamilton–Jacobi equations. Trans. Am. Math. Soc. **277**(1), 1–42 (1983)
43. Van Der Schaft, A.J.: L/sub 2/-gain analysis of nonlinear systems and nonlinear state-feedback h/sub infinity/control. IEEE Trans. Autom. Control **37**(6), 770–784 (1992)
44. Rao, V.G., Bernstein, D.S.: Naive control of the double integrator. IEEE Control Syst. **21**(5), 86–97 (2001)
45. Kearns, M., Littman, M.L., Singh, S.: Graphical models for game theory. In: Proceedings of the Seventeenth conference on Uncertainty in Artificial Intelligence, pp. 253–260. Morgan Kaufmann Publishers Inc. (2001)

46. Beard, R.W., Saridis, G.N., Wen, J.T.: Approximate solutions to the time-invariant Hamilton–Jacobi–Bellman equation. J. Optim. Theory Appl. **96**(3), 589–626 (1998)
47. Bryson, A., Ho, Y.C.: Applied Optimal Control. Hemisphere, New York (1975)
48. Khalil, H.K.: Nonlinear Systems, vol. 3. Prentice Hall, Upper Saddle River (2002)
49. Hornik, K., Stinchcombe, M., White, H.: Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Netw. **3**(5), 551–560 (1990)
50. Ioannou, P., Fidan, B.: Adaptive control tutorial. Society for Industrial and Applied Mathematics (2006)
51. Lewis, F., Jagannathan, S., Yesildirak, A.: Neural Network Control of Robot Manipulators and Non-Linear Systems. CRC Press, Boca Raton (1998)
52. Cao, C., Hovakimyan, N.: Novel l_1 neural network adaptive control architecture with guaranteed transient performance. IEEE Trans. Neural Netw. **18**(4), 1160–1171 (2007)
53. Krstic, M., Kanellakopoulos, I., Kokotovic, P.V.: Nonlinear and Adaptive Control Design. Wiley, New York (1995)
54. Lavretsky, E., Wise, K.A.: Robust and adaptive control with aerospace applications. In: Advanced Textbooks in Control and Signal Processing. Springer-Verlag, London (2013)
55. Pomet, J.B., Praly, L.: Adaptive nonlinear regulation: estimation from the Lyapunov equation. IEEE Trans. Autom. Control **37**(6), 729–740 (1992)
56. Anderson, B.: Exponential stability of linear equations arising in adaptive identification. IEEE Trans. Autom. Control **22**(1), 83–88 (1977)
57. Ioannou, P.A., Tao, G.: Dominant richness and improvement of performance of robust adaptive control. Automatica **25**(2), 287–291 (1989)