IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 7 August 2023, accepted 22 August 2023, date of publication 27 September 2023, date of current version 2 November 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3320063*

RESEARCH ARTICLE

# Consensus Tracking of Multi-Agent Systems in Presence of Uncertain Dynamics and Communications Faults

**KAUSTAV JYOTI BORAH AND KRISHNA DEV KUMAR, (Senior Member, IEEE)**
Department of Aerospace Engineering, Toronto Metropolitan University, Toronto, ON M5B2K3, Canada

Corresponding author: Kaustav Jyoti Borah (kborah@torontomu.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, in part by Queen Elizabeth II Graduate Scholarships in Science and Technology (QEII-GSST), and in part by Toronto Metropolitan University Scholarships for PhD students.

**ABSTRACT** In this paper, the problem of consensus tracking of uncertain multi-agent systems (MAS) with communication faults is addressed. The communication is assumed to be undirected. A reinforced unscented Kalman filter (RUKF) is employed to adapt the noise covariance matrices and to estimate the uncertain states of MAS as well as to train neural network internal parameters by providing a set of prior measurements. A Chebyshev neural network (CNN) is incorporated to learn the uncertain plant. To avert the neural network approximation errors a hyperbolic tangent function based robust control term is applied. The stability of the RUKF which is running simultaneously with the robust control term has been proven using Lyapunov stability approach. Numerical simulations are presented under different fault conditions to show the effectiveness of the proposed RUKF with 5% less computation power compared to adaptive unscented Kalman filter (AUKF).

**INDEX TERMS** Multi-agent systems, uncertain dynamics, reinforcement learning and non-linear filtering, Chebyshev neural networks, control, communication faults.

## I. INTRODUCTION

The popularity of MAS in academia as well as industry has increased tremendously in recent years [1]. Despite its increased significance in several different areas [2], there are still several challenges that must be explored such as communications issues, uncertain nonlinearity, state estimation, fault diagnosis, health monitoring, smart grids, and consensus [3]. The uncertainty of system dynamic models has been dealt with a wide range of methods, and the theory becomes quite enlightened. On the other hand, in MAS, graph theory, a common communication topology, is used such that every agent can interact and exchange information with each other. Once a definite way to exchange the information is determined, the entire dynamics of the network can be established by knowing the dynamics of the individual agents and their interaction with their neighbors. The form of the information exchange is often called a communication protocol [4]. In real field applications, there are always some obstructions in the communications such as uncertain communication, delay, quantization, and limited channel capacity [5], [6]. Uncertainty in communication is used to study the robust consensus problem. Therefore, it is an interesting work, to design a consensus filter for MAS with more complex dynamics in the presence of communication faults. There are a few works reported regarding uncertain communication links for MAS such as [4], [7], and [8]. In [3], a distributed intermediate observer which can estimate the states and multiple faults is derived based on the time delay closed loop system equation. In [4], a distributed consensus protocol is proposed and designed in terms of linear matrix inequalities to ensure robust consensus over uncertain communications. The communication uncertainties in [7] have been modelled as stochastic uncertainties based on which protocols are validated in the mean square sense. A robust consensus problem over uncertain network graphs of linear feedback protocols has been examined by solving a H$\infty$ control problem for a set of 2D subsystems in [8]. Again,

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng H. Zhu.

it is well known that introducing the communication delay will lead to a reduction of the performance of instability [9] and thus investigating communication time delays in MAS is another interesting work. In [9], a sliding mode control design is proposed for a leader follower MAS to deal with the uncertainties incurred by unknown time varying communication delay and disturbances. In [10], consensus control of MAS with input and communication delay in the frequency domain is discussed. The considered linear dynamic model of each agent in [10] involves multiple input delays. In [11], a distributed coordinated attitude tracking control problem for spacecraft formation with time-varying communication delays under the condition that the dynamic leader spacecraft is a neighbor of only a subset of follower spacecraft is examined. However, the observation made from literature, there are a limited work reported regarding the design of a consensus for nonlinear uncertain MAS in presence of communication faults via the design of an estimator. Hence, there are still many open challenges in this area [12]. Motivated by the preceding discussions, an essential and urgent issue that turns out to be unavoidable is to design a consensus-based estimator for the uncertain non-linear MAS in presence of communication uncertainty and delay.

A basic difficulty in state estimation is to figure out a way to design a state estimator so that the estimation error can be assumed to be bounded. Extensive use of the nonlinear state estimation algorithm, primarily unscented Kalman filter (UKF), has been seen over the last few years for the estimation of states [13], MAS [14], [15], [16], fault diagnosis [17].

Again, another neural network (NN) based approach is applied towards estimation of unknown nonlinear dynamics [18] of second order MAS. Recently, NN based fault tolerant control (FTC) method is developed for each agent using local measurements and suitable information exchange between agents for nonlinear uncertain MAS [19]. Researchers [20] proposed a distributed fixed-time observer which is designed for the follower agent to estimate the leader's state under directed networks. Then, based on the estimate, a fixed-time tracking control protocol is developed where novel approximation and estimation schemes are designed to tackle the nonlinear functions and disturbances. In [21], proposes a finite-time consensus FTC tracking problem in the non-strict feedback form and Lyapunov stability guarantees that the tracking error converges to zero for non-strict MAS. The above-mentioned results are combined with adaptive techniques for adjusting the NNs weight. Reinforcement learning (RL) such as temporal difference and $Q$ learning [22] is a famous machine learning (ML) algorithm, that has been extensively studied from various angles and thoroughly applied for several applications [22]. Achievement of an action selection plan that will make the best use of the collective reward over the term is the primary goal of the agent in $Q$ learning [23]. On the contrary, it is understandable that without accurate statistical knowledge about the process and measurement noise covariance matrices, it is not easy to design an optimal estimator. The inexact values of noise covariance matrices will worsen the estimation performance. In this case, the filter must be made in such a way so that it can continue to get an accurate estimation. This work is driven by the robustness of the $Q-$learning method and its successful implementations [24]. To summarize, the significant advantages and novelty of this paper are as outlined hereby,

(1) Utilizing RL to adjust noise covariance uncertainties in the UKF is demonstrated to enhance the estimation performance for nonlinear uncertain MAS. Additionally, a consensus term is introduced, shedding light on the core principles of multi-agent coordination using graph theory.
(2) The primary innovation in this study lies in the creation of autonomously refining noise covariance uncertainties in the UKF using $Q$–learning. We achieve the optimal noise covariance uncertainties in the UKF within the framework of $Q$–learning, and this approach is referred to as Reinforced UKF. To the best of our knowledge, the issue of $Q$–learning-based noise covariance adaptation in UKF remains unresolved.
(3) We establish the asymptotic stability of the estimation error and state error, incorporating a control law based on Lyapunov stability theorem. This guarantees the leader-follower consensus tracking of MAS.
(4) The outcomes presented in this paper address uncertainties in MAS model, encompassing communication uncertainties and delays, demonstrating the practicality and computational efficiency of our work.

This paper is organized as follows; Section II describes the problem formulation with some other preliminaries. Observer and controller design has been stated in section III. The results and discussion section have been presented in section IV and concluding remarks of this work are drawn in section V.

## II. PROBLEM FORMULATION

### A. MAS DYNAMICS

Consider a group of MAS of $n$ agents (in this paper, $n = 6$) in which the following second-order differential equation describes the $i^{th}$ agent as,

$$\dot{x_i} = v_i, \quad i= 1, 2, \ldots, n \tag{1}$$

$$\dot{v_i} = f_i(x_i, v_i) + g_i(x_i)u_i + d_i, \tag{2}$$

Here, $x_i \in \mathrm{R^m}$ and $v_i \in \mathrm{R^m}$ is the $i^{th}$ agent position and velocity vectors,, $u_i \in \mathrm{R^m}$ is the control input acting at the $i^{th}$ agent, $f_i(x_i,v_i) \in \mathrm{R^m}$ is the uncertain plant, $g_i(x_i) \in \mathrm{R}^{mxm}$ is a non-singular control gain matrix, $d_i \in \mathrm{R^m}$ is the unknown bounded external disturbance acting at the $i^{th}$ agent. The measurement model can be written as,

$$y_i = h_i\,(x_i(t),\, v_i(t),\, t) + \vartheta_i(t), \quad i = 1, 2, \ldots, n \tag{3}$$

where $y_i \in \mathrm{R^m}$ is the measurement vector, $(x_i,\, v_i) \in \mathrm{R^m}$ is the state vectors (position and velocity) that gives all information about the system behavior, $h_i \in \mathrm{R^m}$ is the nonlinear mapping of measurements, $\vartheta_i$ is the additive measurement noise, $t$ is

time. Let's assume that there is a virtual leader, identified as agent 0, whose state $x_0 \in \mathrm{R}^{\mathrm{m}}$ is the time-varying reference trajectory for the MAS.

*Assumption 1:* The reference trajectory $x_0$, its derivatives $\dot{x_0}, \ddot{x_0}$ are in a compact set [25],

$$\Omega_0 \equiv \{(x_0 \dot{x_0} \ddot{x_0}) | \|x_0\|^2 + \|\dot{x_0}\|^2 + \|\ddot{x_0}\|^2 \leq C_1\}$$

where $C_1$ is a positive constant. This assumption states that for a consensus achievement of MAS, the reference trajectory $x_0$, and its derivatives are bounded which is common. A norm $\| x_0 \|$ is a function mapping $x_0$ into nonnegative real.

*Assumption 2:* The disturbances $d_i (i = 1,2, \ldots n)$ are bounded s.t. $\| d_i \| \leq d_{Mi}$, where $d_{Mi}$ is a positive constant. Here, Eqs.(1)-(3) can be used to represent many practical systems such as spacecraft attitude dynamics, satellite orbital control systems, robotic manipulators and so on.

### B. COMMUNICATION BETWEEN THE AGENTS

The communication framework is manifested in the Fig. 2 among the agents in the given MAS $i$ is described by a graph, $G = (\Upsilon, E, A)$, where $\{\Upsilon = \varsigma_1, \varsigma_2, \ldots, \varsigma_{\mathrm{n}}\}$ is the set of interconnected nodes, $E \subseteq \Upsilon \times \Upsilon$ is the set of edges, and $A = [a_{ij}] \in R^{\mathrm{n} \times \mathrm{n}}$ is the weighted adjacency matrix with non-negative elements and $a_{ij=} a_{ji} \geq 0$ and $a_{ii} = 0$. Node $\varsigma_i$ $(i = 1, 2, \ldots n)$ represents the $i^{th}$ agent, and an edge in $G$ is denoted by an unordered pair $(\varsigma_i, \varsigma_j)$. $(\varsigma_i, \varsigma_j) \in E$ if and only if there is an information exchange between the $i^{th}$ agent and the $j^{th}$ agent that is, $(\varsigma_i, \varsigma_j) \in \mathrm{E} \Leftrightarrow (\varsigma_j, \varsigma_i) \in \mathrm{E}$. For any two nodes $\varsigma_i$ and $\varsigma_j$, if there exists a path between them, then $G$ is called a connected graph. In this case, the graph is undirected, $(\varsigma_i, \varsigma_j) \in E \Leftrightarrow (\varsigma_j, \varsigma_i) \in E$. It is assumed that $a_{ij} = a_{ji}$ and $a_{ii} = 0$.

Now, the information flow between the $i^{th}$ agent and agent 0 (virtual leader), where $a_{\mathrm{i}0} > 0$ $(i = 1, 2, \ldots, n)$ is a constant if the $i^{th}$ agent has access to the leader agent that is agent 0, otherwise $a_{\mathrm{i}0} = 0$. Let $\mathrm{B} = \mathrm{diag}\{b_1 \ldots b_n\}$ be the leader adjacency matrix associated with $G$ where $b_i > 0$ $(i = 1,2, 3 \ldots n)$ is a constant if the $i^{th}$ agent has access to the leader and $b_i = 0$ otherwise. Using the adjacency matrices, $A$ and $B$, we define a lumped state error, $\alpha_{xi}$ and $\alpha_{vi}$ with communication fault can be expressed as,

$$\alpha_{xi} = \sum_{j=1}^{n} \{a_{ij}(x_i - x_j) + b_i(x_i - x_0)\}(1 + \Delta_i) \quad (4)$$

$$\alpha_{vi} = \sum_{j=1}^{n} \{a_{ij}(v_i - v_j) + b_i(v_i - \dot{x}_0)\}(1 + \Delta_i) \quad (5)$$

$\Delta_i$ denotes the faults associated with the communication channels where $i = 1,2 \ldots 6$ and $j = 1,2 \ldots 6$. The control input $u_i$ in Eq. (2) is given by,

$$u_i = g_i(x_i, v_i, x_0, \alpha_{xi}, \alpha_{vi}) \quad (6)$$

where $g_i$ is a function and $x_0$ is the reference trajectory of the $i^{th}$ agent.

*Theorem 1:* If $\bar{G}$ is connected, then the matrix L+B associated with $\bar{G}$ is symmetric and positive definite [25].

## III. OBSERVER AND CONTROLLER DESIGN

### A. REINFORCED UNSCENTED KALMAN FILTERING

We have tackled the challenges associated with uncertain noise covariance's in the UKF, specifically related to the $Q$ (process noise covariance matrix) and $R$ (measurement noise covariance matrix). We have effectively mitigated the adverse effects of these uncertainties in the UKF by integrating $Q$–learning. This adaptation strategy, developed within the $Q$–learning framework, enhances the performance of the estimator, which we refer to as the ***Reinforced UKF***. The RUKF is less sensitive to noise covariance uncertainty. The key advantage of the RUKF is its ability to improve estimation accuracy when the actual noise covariance deviates from its nominal value. Another significant benefit is that the RUKF can optimize performance matrices without requiring an explicit optimization algorithm. This is because the RUKF is entirely based on the principles of RL. It's important to note that RL and optimization algorithms, such as genetic algorithms (GA), are essentially addressing the same category of problems, which involve searching for solutions that either maximize or minimize a given function, reward, or cost function. In this regard, they share common characteristics.

To understand RUKF, it is necessary to build the state set as $S$ and corresponding action set as $A$ for $Q-$learning. Inspired by grid wise control [23], we construct a grid as shown in Fig.1(a) in which we have different pairs of noise covariances matrices as states $S$ and their corresponding actions up, down, left, and right as $A$. Let us consider in the grid their C design values for process noise covariance matrices as $\{Q^1_{\mathrm{design}}, Q^2_{\mathrm{design}}, \ldots . Q^M_{\mathrm{design}}\}$ and D design values for measurement noise covariance matrices $\{R^1_{\mathrm{design}}, R^2_{\mathrm{design}}, \ldots . R^N_{\mathrm{design}}\}$ hence the states composed of C by D state pairs. The narrow arrows are actions as up, down, left, and right describes the agents' state transition.

The grid size is set at 4 by 4, which corresponds to the number of states and total agents. In Fig. 1(a), each $S_{ij}$ represents $i^{th}$ row and $j^{th}$ column on the grid, where various $Q_{\mathrm{design}}$ and $R_{\mathrm{design}}$ sets are employed along with actions. It is assumed that the agent can move to adjacent states or remain in its current state based on received actions, with each transition yielding a certain reward. This paper's objective is to determine suitable $Q_{\mathrm{design}}$ and $R_{\mathrm{design}}$ configurations that can enhance the performance of the UKF. The algorithm involves three parallel UKF algorithms. Initially, a standard UKF (UKF) is used to generate filtering results using nominal noise covariance matrices. Subsequently, a grid-based UKF (GUKF) is employed to evaluate the impact of different $Q_{\mathrm{design}}$ and $R_{\mathrm{design}}$ configurations. A scalar reward is calculated based on absolute and relative error matrices (Eqs. 28 and 29), denoted as 'res' (line 14 of Algorithm 1), which are generated by the UKF with the nominal $Q$ and $R$, and the GUKF with $Q_{\mathrm{design}}$ and $R_{\mathrm{design}}$ selected according to the current state on the grid.

The reward is utilized to indicate the estimation performance. The best couple of $Q_{\mathrm{design}}$ and $R_{\mathrm{design}}$ are selected gradually according to the cumulative reward. Generally,

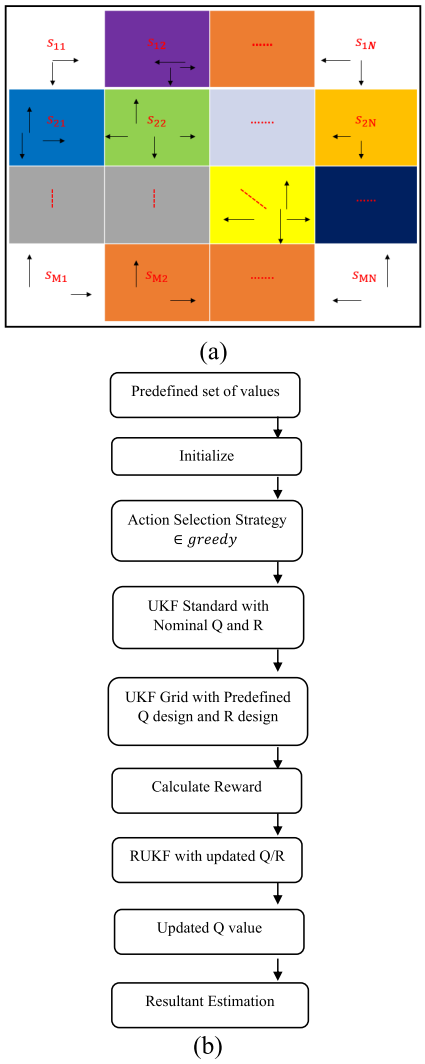


**FIGURE 1.** **(a) States and their corresponding action set in a grid (b) Flowchart of the proposed RUKF.**

a smaller and positive reward indicates that the current setting of the $Q_{\text{design}}$ and $R_{\text{design}}$ in the filter can achieve high accuracy. Based on the action selection strategy and cumulative rewards, the agent attempts to reach the appropriate $Q_{\text{design}}$ and $R_{\text{design}}$ state. The third filter(line 15 of Algorithm 1) which is RUKF because $Q_{\text{design}}$ and $R_{\text{design}}$ has been updated using a $Q$ learning method, is introduced to achieve the optimized final estimates. In Fig. 1(b), we have shown the flowchart of the proposed RUKF algorithm.

In this context, T represents the total time required to attain the reward $R_r$, and $\epsilon$ denotes the probability associated with the $\epsilon-$ greedy strategy for selecting a random action. *Algorithms 2* outlines the process of the $\epsilon-$ greedy approach, which is employed to determine actions based on the received rewards. Within this algorithm, the "*rand*(.)" function generates a random action with a probability of $\epsilon$, while it selects the action that maximizes cumulative rewards with a probability of $(1 - \epsilon)$. This approach is crucial for striking a balance between exploration and exploitation.

**Algorithm 1** Reinforced UKF

1. Initialize $\hat{x}_0^{ukf} = \hat{x}_0^{gukf} = \hat{x_0}$ and $p_0^{ukf} = p_0^{gukf} = P_0$
2. $k \leftarrow 0$
3. $\forall$ s $\in$ $S$ and a $\in$ $A$ , Initialize $Q(s, a) \leftarrow 0$, $V(s) \leftarrow 0$;
4. Initialize the state $S$
5. Initialize $R_r \leftarrow 0$
6. for *each period* do,
7. a $\leftarrow \epsilon-$greedy ($A$, $Q(s,a)$, $s$, $\epsilon$ ) choose $a$ in $s$ based on $\epsilon-greedy$strategy
8. Execute the action $a$ and go to the next state $s'$
9. for $t$ = 0:T do
10. $k \leftarrow_k +1$
11. $[\hat{x}_k^{ukf}, p_k^{ukf}, \text{residue}_k^{ukf}] \leftarrow$ UKF $(\hat{x}_{k-1}^{ukf}, p_{k-1}^{ukf}, y_k, u, Q, R)$

Step 1: Initialize $Q$ and $R$ (nominal)

Step 2: Sigma Points

$$X_{k-1}^a = [\hat{x}_{k-1}^a \hat{x}_{k-1}^a \pm \sqrt{(L+\lambda)\, P_{k-1}^a}]$$

Step 3: Time Update

$X_{k|k-1}^x = F[X_{k-1}^x, X_{k-1}^v]$

$\hat{x}_k^- = \sum_{i=0}^{2L} W_i^{(m)} X_{i,k|k-1}^x$

$P_k^- = \sum_{i=0}^{2L} W_i^{(c)} [X_{i,k|k-1}^x - \hat{x}_k^-][X_{i,k|k-1}^x - \hat{x}_k^-]^T + Q_{k-1}$

$y_{k|k-1} = H[X_{k|k-1}^x, X_{k-1}^n]$

$\hat{y}_k^- = \sum_{i=0}^{2L} W_i^{(m)} y_{i,k|k-1}$

Step 4: Measurement update

$P_{\tilde{y}_k \tilde{y}_k} = \sum_{i=0}^{2L} W_i^{(c)} [y_{i,k|k-1} - \hat{y}_k^-][y_{i,k|k-1} - \hat{y}_k^-]^T + R_k$

$P_{x_k y_k} = \sum_{i=0}^{2L} W_i^{(c)} [X_{i,k|k-1} - \hat{x}_k^-][y_{i,k|k-1} - \hat{y}_k^-]^T$

$K = P_{x_k y_k} P_{\tilde{y}_k \tilde{y}_k}^{-1}$

$\hat{x}_k = \hat{x}_k^- + K(y_k - \hat{y}_k^-) + C_k \sum_{J \in N_i} (\hat{x}_{k-1}^j - \hat{x}_k^-)$

$P_k = P_k^- - K P_{\tilde{y}_k \tilde{y}_k} K^T$

12. $[\hat{x}_k^{gukf}, p_k^{gukf}, \text{residue}_k^{gukf}] \leftarrow$ GUKF $(\hat{x}_{k-1}^{gukf}, p_{k-1}^{gukf}, y_k, u, Q_{\text{design}}, R_{\text{design}})$

Step 1: Initialize $Q_{\text{design}}$ and $R_{\text{design}}$ based on the grid set (Fig. 1a)

Step 2: Repeat step 2 to 4 under line 11 with step 1 under line 12.

13. 13.$Q_{\text{design}}$ and $R_{\text{design}}$ is determined with the current state $s'$

14. Reward $R_r = R_r + [(res_i^{ukf})^T * res_i^{ukf} - (res_i^{gukf})^T * res_i^{gukf}]$

15. $[\hat{x}_k, p_k, \text{residue}_k] \leftarrow$ RUKF $(\hat{x}_{k-1}, p_{k-1}, y_k, u, Q_{\text{design}}, R_{\text{design}})$

Step 1: Initialize $Q_{\text{design}}$ and $R_{\text{design}}$ is determined by the current state $s'$

Step 2: Repeat step 2 to 4 under line 11 with step 1 under line 15.

16. *end* for

17. Update $Q$ value, $Q$ (s,a) $\leftarrow Q$ (s,a) $+\alpha\,(R_r + \gamma V(s') - Q(\text{s,a}))$

18. $V(s) \leftarrow \underset{a}{max}\, Q$ (s,a)

19. s $\leftarrow s'$ and $R_r \leftarrow 0$

20. Reset $\hat{x}_k^{gukf} \leftarrow \hat{x}_k^{ukf}$ and $p_k^{gukf} \leftarrow p_k^{ukf}$

21. *end* for

22. return $\{\hat{x_k}\}$ and $\{p_k\}$, the resultant estimation

The RUKF has both advantages and disadvantages, including:

1) When the actual noise covariance's deviate from their nominal values, the RUKF's performance can be improved. Additionally, unlike the extended Kalman filter, it doesn't rely on linearization for state predictions and covariance's.

2) Typically, process and measurement noise covariance's are manually adjusted through trial and error, often involving

**Algorithm 2** $\epsilon$ - *Greedy* Action Selection

1. *function* $\epsilon-$greedy ($A$, $Q(s, a)$, $s$, $\epsilon$)
2. Choose an *action based onQ value*
3. If rand(.) $< \epsilon$
4. Choose $a$ in $A$ at random
5. else
6. a = arg $\underset{a}{max}$ $Q$ (s,a)
7. *end* if
8. return $a$
9. *end* function

tedious tasks guided by empirical rules, consuming time and yielding only modest results. In contrast, RUKF updates process and measurement noise covariance's without relying on a trial-and-error approach.

3) Because three filters operate concurrently to generate residuals, RUKF is computationally more efficient and capable of achieving optimality in Markov decision processes.

4) RUKF may experience divergence if the design values in the grid are not properly selected. The grid's size and the control of design values are heuristics that fundamentally impact RUKF's performance. The choice of actions within the grid is another critical factor in determining the precise sequence of design values; selecting a random action can degrade the overall performance of RUKF. Additionally, RUKF may fail to converge if the learning rate is set too high or too low.

*Remark 1: Algorithms 1 and 2* depict the structure of the RUKF algorithm. *Line 15* describes the final output of the RUKF with the updated $Q_{\text{design}}$, $R_{\text{design}}$ obtained using $Q$ learning-based approach.

Some initial parameters of *Algorithm 1* can be defined as,

$\hat{x_0}$ : Initial state estimate, $p_0$ : Initial covariance
S: state set, s: a particular state of S
$s'$ : next state, $R_r$ : reward
$A$: action set, a: particular action of $A = \{a^1 = \text{up}, a^2 = \text{down}, a^3 = \text{left}, a^4 = \text{right}\}$
$Q(s, a)$ : $Q-$value with state 's' and action 'a'
$V(s)$: value function
$\alpha$ : learning rate, $\gamma$: discount factor
$Q$ and $R$ : noise covariance matrices
$Q_{\text{design}}$, $R_{\text{design}}$: Predetermined set of values in Grid

*Remark 2:* A linear scalarization function in *line 14* has been used in *algorithm 1* to calculate the reward $R_r$, an impulsive approach to evaluate the state-action pair. If we get a positive reward for the current state-action pair then the $Q_{\text{design}}$ and $R_{\text{design}}$ for that particular state-action pair for GUKF are more appropriate than the UKF, otherwise, the negative reward wont be considered.

### B. CONTROL LAW

Suppose the agents are interconnected by static diffusive-type coupling [25],

$$u_i = g_i^{-1}\left(-W_{\pi_i}\xi_i - K_{1i}s_i - K_{2i}s_i^{\nu} - \psi_i\right) \tag{7}$$

where, $K_{1i}$ and $K_{2i}$ are constant, diagonal positive definite matrices, $\xi_i$ is the Chebyshev polynomial, $s_i^{\nu} = (s_{i1}^{\nu}, s_{i2}^{\nu}, \ldots, s_{im}^{\nu})^T$, the adaptive neural network compensator $W_{\pi_i}$ ($i = 1,2\ldots n$) is used to calculate the unknown nonlinear function, $f_i(x_i, v_i)$ and the robust control term $\psi_i \in \mathrm{R^m}$ ($i = 1,2\ldots n$) is used to prevent the neural network approximation errors as well as the external disturbances. Next, the robust controller $\psi_i$ ($i = 1,2,\ldots n$) in Eq. (7) can be written as [25],

$$\Psi_{ij} = k_i \tan h(\frac{mk_u k_i s_{ij}}{\epsilon}),\ k_u = 0.2785,\ j = 1, 2\ldots n \tag{8}$$

Here, $k_i$ is a positive constant and $\epsilon$ are positive scalar. The terminal sliding manifolds $s_i \in \mathrm{R^m}$ ($i = 1,2\ldots, n$) is expressed as [25],

$$s_i = \alpha_{\hat{v}i} + \sigma_{1i}\alpha_{\hat{x}i} + \sigma_{2i}\beta_i \tag{9}$$

where, $s_i = (s_1^T \ldots . s_n^T)^T \in \mathrm{R^{mn}}$, $\beta_i = (\beta_1^T \ldots \beta_n^T)^T \in \mathrm{R^{mn}}$, $\sigma_{1i} = diag(\sigma_{1i}\ I_m \ldots . \sigma_{1n}\ I_m) \in \mathrm{R^{(mn)x(mn)}}$, and $\sigma_{2i} = diag(\sigma_{2i}\ I_m \ldots . \sigma_{2n} I_m) \in \mathrm{R^{(mn)x(mn)}}$ where, $i = 1,2\ldots . n$. The term $\beta_i$, $\sigma_{1i}$, $\sigma_{2i}$ has been mainly discussed in the stability proof section of this paper. Lumped state errors ($\alpha_{\hat{x}i}$ and $\alpha_{\hat{v}i}$, $i = 1, 2\ldots n$) $\in \mathrm{R^m}$ that includes the absolute and relative state errors are necessary to measure the true tracking errors of the whole system. The $i^{th}$ agent may not obtain all the absolute state errors ($e_{\hat{x}i}$ and $e_{\hat{v}i}$,$i = 1, 2\ldots n$) and all relative state errors ($r_{\hat{x}ij}$ and $r_{\hat{v}ij}$ where, $i = 1, 2\ldots n$ and $j = 1, 2\ldots n$) with the consideration of the above facts, using the weighted adjacency matrices $A$ and $B$, we define lumped state errors as [25],

$$\alpha_{\hat{x}i} = \sum_{J=1}^{n} a_{ij} r_{\hat{x}ij} + b_i e_{\hat{x}i} \tag{10}$$

$$\alpha_{\hat{v}i} = \sum_{J=1}^{n} a_{ij} r_{\hat{v}ij} + b_i e_{\hat{v}i} \tag{11}$$

where, $\alpha_{\hat{x}i} \in \mathrm{R^m}$ and $\alpha_{\hat{v}i} \in \mathrm{R^m}$ where $i = 1,2, \cdots n$. $a_{ij}$ is the element of the weighted adjacency, $k_i$ is the element of the weighted matrix $K$. The lumped state errors are defined as the summation of the absolute and relative state errors, and it is only dependent on the $i^{th}$ neighboring agent's pieces of information.

*Relative Error*: It is defined as the estimated state of an individual agent with respect to another estimated state of an agent. The relative state error between the $i^{th}$ agent and the $j^{th}$ agent ($i = 1,2\ldots n$ and $j = 1, 2\ldots . n$) is defined simply by,

$$r_{\hat{x}ij} = \hat{x}_i - \hat{x}_j \tag{12}$$

$$r_{\hat{v}ij} = \hat{v}_i - \hat{v}_j \tag{13}$$

*Absolute Error*: It is the difference between estimated state of individual agent to the reference trajectory. The absolute errors of the $i^{th}$ agent ($i = 1,2\ldots . n$) are defined by,

$$e_{\hat{x}i} = \hat{x}_i - x_0 \tag{14}$$

$$e_{\hat{v}i} = \hat{v}_i - \dot{x_0} \tag{15}$$

where $e_{\hat{x}i} \in \mathrm{R^m}$ and $e_{\hat{v}i} \in \mathrm{R^m}$ where $i = 1,2,\ldots n$. The commonly reference trajectory (*i.e.,* $x_0$ and $\dot{x}_0$) is available to

only a subset of agents, and each agent in the group has access to information about its neighbor agents that can be expressed by the adjacency matrices $A$ and $B$,

$$B = diag\{a_{10}, \ldots . a_{i0}\} \text{ and } a_{i0} > 0 \tag{16}$$

where $i = 1,2,\ldots 6$ is a constant if the $i^{th}$ agent has access to the leader, otherwise $a_{i0} = 0$.

*Remark 3:* The control law described in Eq.(7), the adaptive neural network compensator $W_{\pi_i}$ $(i = 1,2\ldots n)$ is used to calculate the uncertain plant $f_i(x_i, v_i)$, the nonlinear feedback $K_{1i}s_i$ and $K_{2i}s_i^y$ is used to drive the system states to the sliding manifold in finite time.

### C. STABILITY ANALYSIS

Say Lyapunov candidate [25], [26], [27] for a closed-loop system can be defined as,

$$V(\tilde{x}, s, w) = V(\tilde{x}) + V(s, w) \tag{17}$$

where,

$$V(\tilde{x}) = \tilde{\chi}_{k+1|k}^T \hat{P}_{k+1|k}^{-1} \tilde{\chi}_{k+1|k} \tag{18}$$

$$V(s, w) = \frac{1}{2} s^T P_2 s + \sum_{i=1}^{n} V_{wi} \tag{19}$$

Our goal is to prove $\dot{V}(\tilde{x}, s, w) \leq 0$

*Theorem 2:* To prove RUKF stability, let the following assumptions hold:

*Assumption 3:* There exist real constants $\underline{f}, \underline{h}, \underline{\beta}, \underline{\alpha}, \overline{\alpha} > 0$; $\underline{q}, \overline{q}, \underline{\hat{q}}, \underline{q}^*, \overline{r}, \underline{r}^*, \underline{p}, \overline{p} > 0$ such that the following bounds on various matrices are satisfied for every $K > 0$:

$$\left.\begin{array}{l} \underline{f}^2 I \leq F_k F_k^T, \underline{h}^2 I \leq H_k H_k^T, \underline{\beta} I \leq \beta_k \\ \underline{\alpha} I \leq \alpha_k \leq \overline{\alpha} I, \underline{q} I \leq Q_k \leq \overline{q} I, \underline{q}^* I \leq Q_k^* \\ R_k \leq \overline{r}\, \bar{I}, \underline{r}^* I \leq R_k^*, \underline{p} I \leq \hat{P}_k \leq \overline{p} I \end{array}\right\} \tag{20}$$

There exist real constants overline $\bar{f}, \bar{\phi}, \bar{h}, \bar{\beta}$, such that the matrix norm is bounded via.

$$\|\Upsilon_k F_k\| \leq \bar{f},\ \|\phi_k\| \leq \bar{\phi},\ \|\Upsilon_k H_k\| \leq \bar{h},\ \|\beta_k\| \leq \bar{\beta} \tag{21}$$

Now it can be concluded that the estimation error is exponentially bounded in the mean square. Proof of Theorem 2 can be found in Appendix B. For the consensus tracking control problem for MAS, we consider absolute state errors and relative state errors. Refer to Appendix C to find a more detailed explanation of absolute and relative error matrices to prove the stability using a control law given in Eq. (7).

*Theorem 3:* The MAS described by Eqs. (1) – (3) and assumptions 1,2 and C.1 and C.2 (see Appendix C for assumptions C.1 and C.2) are satisfied. The control law has been described by Eq. (7) where the projection algorithm is given by Eqs. (C.35) and (C.36). For a sufficiently large positive constant $s_{max}$ if the initial conditions satisfy the following,

$$s^T(0)s(0) \leq \frac{2s_{max}}{\lambda_{max}(P_2)} \tag{22}$$

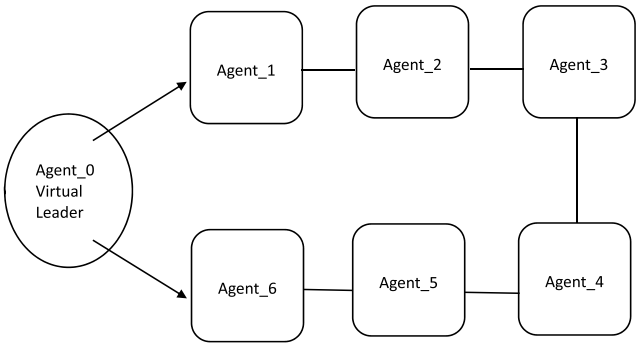


**FIGURE 2.** Leader-follower MAS communication.

Then s is uniformly bounded. Proof can be found in Appendix D.

Based on the derivation provided in Appendix A to D, we get,

$$\dot{V}(\tilde{x}) \leq (-\lambda V_k(\tilde{\chi}_{k|k-1}) + \mu) \tag{23}$$

$$\dot{V}(s, w) \leq -(\lambda_{min}(K_1) - C_2)s^T s + C_3 \tag{24}$$

So,

$$\dot{V}(\tilde{x}, s, w) \leq (-\lambda V_k(\tilde{\chi}_{k|k-1}) + \mu) + (-(\lambda_{min}(K_1) - C_2)s^T s + C_3) \tag{25}$$

In this paper, the estimation error for RUKF is examined. It has been shown that, with some given certain conditions, the estimation error is bounded by the mean square. This fact is embodied in Theorem 2 provided above. We also have proven the stability using control law. It is to be noted that s is uniformly ultimately bounded, and it has been stated in Theorem 3 and thus, one can conclude that lumped state errors and $e_{\hat{x}}$ and $e_{\hat{v}}$ are bounded. At the same time, the boundedness of $x_0, \dot{x_0}$ implies that $x_i$ and $v_i$ $(i = 1,2..6)$ are bounded, and hence, $\xi_i$ and $\|\tilde{w}_{\pi i}\xi_i\|$ are bounded, respectively. Because the controller in Eq.(7) in which all signals associated with it are bounded, so we can say the control input $u_i(i = 1, 2..6)$ is also bounded. Finally, using Eq. (25), we can say that the overall closed-loop system is asymptotically stable, and this concludes the proof. We rewrite the stability proof of the proposed system in Appendix A to D from our paper [28]. In the next sections of the paper, we have shown the numerical simulations.

## IV. RESULTS AND DISCUSSIONS

To assess the execution of the proposed RUKF, the MAS model is described in Eqs. (1)-(3), was examined for various conditions such as 1. uncertainty in the system model, 2. communication uncertainty, and 3. communication delay. MATLAB based numerical integration of the states is conducted using the classical Runge-Kutta method (RK4) with a step size of Ts = 0:01s and the total simulation time = 50s. The machine used for the simulations is a Lenovo ThinkPad 7th generation running on an Intel® Core™ i5-7200U CPU with 2.50 GHz processing power and 8GB RAM. The time-varying reference trajectory of a MAS, as shown in

Fig. 2, is provided by a virtual leader where the reference trajectory is available to a subset of the agents, and each agent in the group has restricted access to information of its neighbor agents only.

The goal is to optimize the reward of RUKF so that the guaranteed convergence of the absolute (AEMest) and relative (REMest) error metric for estimator of the MAS can be obtained. The communication quality between the virtual leader and each agent as shown in Fig.2 is expressed with a symmetric weighted adjacency matrix $A$ and $B$.

The weighted adjacency matrices are given as follows,

$$\left.\begin{aligned} A &= \begin{bmatrix} 0 & 0.7 & 0 & 0 & 0 & 0 \\ 0.7 & 0 & 0.6 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.7 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0.8 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.9 & 0 \end{bmatrix} \\ B &= diag[0.8\ 0.0\ 0.0\ 0.0\ 0.0\ 0.9] \end{aligned}\right\} \quad (26)$$

where, $a_{12} = 0.7 = a_{21}$, $a_{23} = 0.6 = a_{32}$, $a_{34} = 0.7 = a_{43} = a_{12}$, $a_{45} = 0.8 = a_{54}$, $a_{56} = 0.9 = a_{65}$, $b_{11} = 0.8$, $b_{16} = 0.9$

*Unscented Kalman Filter:* The composite scaling parameter is $\lambda = \alpha_{UKF}^2 \; x\,(n + \kappa) - n$ where, $\alpha_{UKF} = 1$, $\beta_{UKF} = 2$, $\kappa = 1$ and $n$ are the dimension of augmented state. The filter weights are calculated according to, $W_0^{(m)} = \frac{\kappa}{n+\kappa}$, $W_0^{(c)} = \frac{\kappa}{n+\kappa} + (1 - \alpha_{UKF}^2 + \beta_{UKF})$ and $W_i^{(m)} = W_i^{(c)} = \frac{1}{2(n+\kappa)}$ ($i$ = 1,2,…6) where, $W_i^{(m)}$, $W_i^{(c)}$ are the weight associated with the mean and covariance matrices accordingly. In the case of the GUKF, the design values of $Q_{\text{design}}$ and $R_{\text{design}}$ will be considered specifically in the range of $10^{-5} \leq Q_{\text{design}} \leq 10^2$ and $10^{-5} \leq R_{\text{design}} \leq 10^2$. The $Q$ learning method, learning rate, discount factor and action selection probability has been taken as $\alpha = 0.2$, $\gamma = 0.9$, and $\epsilon = 0.1$. The window size ($N$) has been taken as 50, adaptive threshold factor ($n$) as 3, adaptive fading factor ($\zeta$) *as* 6000, and measurement error as 0.1. While initiating the filter it is recommended to have large values of diagonal elements of covariance matrix $P$, however as the filter converges, the diagonal elements will slowly decrease to show more confidence in estimation.

*Adaptive Law:* The adaptive law has been proposed by Zou [25]. The adaptive law for $W_i$ ($i = 1, 2, \ldots . n$) is,

$$\dot{W}_i = \delta_i s_i \xi_i^T \quad (27)$$

where $s_i$ is terminal sliding manifold, $\delta_i$ is positive constant and $\xi_i^T$ is polynomial basis function. The reference trajectory is given by $x_0 = 0.5[cos(\text{t})\ sin(\text{t})]^T$ and some other controller parameters are $p = 3$, $q = 5$, $\epsilon = 0.01$, $k_i = 0.5$, $\sigma_{1i} = 2$, $\sigma_{2i} = 0.2$, $\delta_i = 100$, $K_{1i} = \text{diag}\{4,4\}$, $K_{2i} = \text{diag}\{1,1\}$. The AEMest and REMest can be used to evaluate the proposed estimator as,

$$AEM_{est} = \sqrt{\sum_{i=1}^{n} ||e_{\hat{x}i}||^2} \quad (28)$$

**TABLE 1.** Uncertainty in dynamic model.

| Time | Dynamic Model |
|---|---|
| $0 < t \leq 5$ | $f_i(x_i, v_i) = [\,4x_{i2}\ sin\,(\frac{\pi}{4} + \frac{h_{1i}v_{i1}}{2}),\ 4x_{i1}cos\,(\frac{h_{2i}v_{i2}}{2})\,]^T$ |
| $5 < t \leq 10$ | $f_i(x_i, v_i) = [\,12x_{i2}\ sin\,(\frac{\pi}{4} + \frac{h_{1i}v_{i1}}{2}),\ 12x_{i1}cos\,(\frac{h_{2i}v_{i2}}{2})\,]^T$ |
| $10 < t \leq 15$ | $f_i(x_i, v_i) = [\,5x_{i2}\ sin\,(\frac{\pi}{4} + \frac{h_{1i}v_{i1}}{2}),\ -4x_{i1}cos\,(\frac{h_{2i}v_{i2}}{2})\,]^T$ |
| $t > 15$ | $f_i(x_i, v_i) = [\,4x_{i2}\ sin\,(\frac{\pi}{4} + \frac{h_{1i}v_{i1}}{2}),\ 4x_{i1}cos\,(\frac{h_{2i}v_{i2}}{2})\,]^T$ |

**TABLE 2.** Nominal initial system states.

| $i$ | $\boldsymbol{x_i} = [\boldsymbol{x_{i1}}, \boldsymbol{x_{i2}}]^T$ | $\boldsymbol{v_i} = [\boldsymbol{v_{i1}}, \boldsymbol{v_{i2}}]^T$ |
|---|---|---|
| 1 | $x_1\,(0) = [1,\ 0.5]^T$ | $v_1\,(0) = [0,\ 0]^T$ |
| 2 | $x_2\,(0) = [1.8,\ -1.2]^T$ | $v_2\,(0) = [0,\ 0]^T$ |
| 3 | $x_3\,(0) = [2.2,\ -2.5]^T$ | $v_3\,(0) = [0,\ 0]^T$ |
| 4 | $x_4\,(0) = [-1.1,\ 0.2]^T$ | $v_4\,(0) = [0,\ 0]^T$ |
| 5 | $x_5\,(0) = [-4.6,\ 1.8]^T$ | $v_5\,(0) = [0,\ 0]^T$ |
| 6 | $x_6\,(0) = [1,\ -1.7]^T$ | $v_6\,(0) = [0,\ 0]^T$ |

**TABLE 3.** Parameters of agent's dynamics.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $\boldsymbol{h_{1i}}$ | 0.6 | -0.2 | 0.4 | -0.3 | 0.5 | 0.5 |
| $\boldsymbol{h_{2i}}$ | 0.3 | 0.4 | -0.4 | -0.7 | -0.2 | -0.5 |

and the REMest is defined as,

$$REM_{est} = \sqrt{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} ||r_{\hat{x}ij}||^2} \quad (29)$$

*Remark 4:* The controller parameter $K_1$ and $K_2$ determine the size of the parameter $\lambda_{max}(P_2)$ in Eq. (C.40). So, the smaller parameter $\lambda_{max}(P_2)$ the bigger parameter $K_1$ and $K_2$ and bigger $\lambda_{min}(P_2)$ are required. The parameters $\sigma_{1i}$, $\sigma_{2i}$, $p$ and $q$ determine the size of the parameter and the terminal

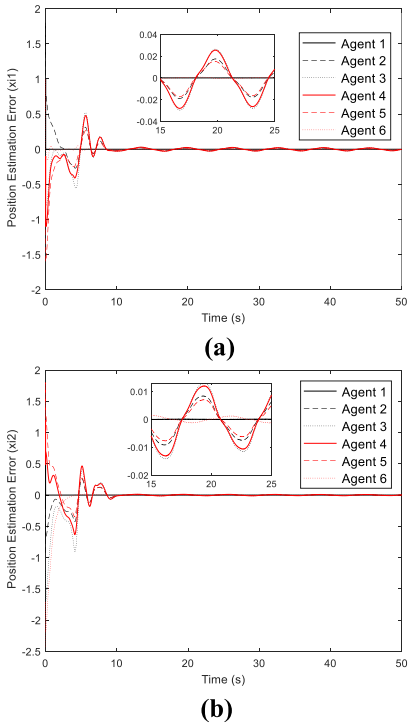


**FIGURE 3.** Estimation error under uncertain dynamics using AUKF: (a) xi1_AUKF (b) xi2_AUKF.

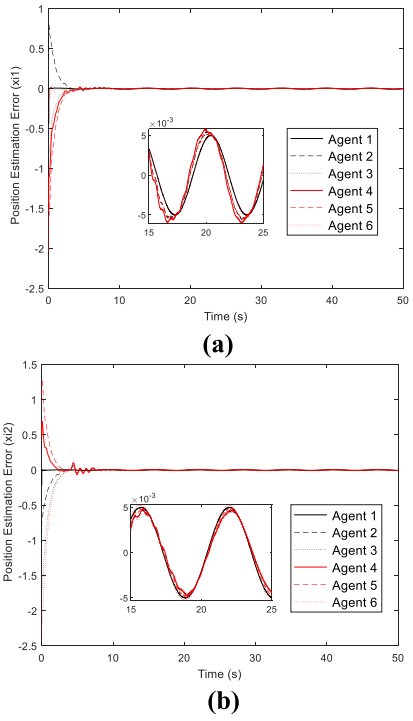


**FIGURE 4.** Estimation error under uncertain dynamics using RUKF: (a) xi1_RUKF (b) xi2_RUKF.

sliding manifolds. So, the smaller parameter for sliding manifold, the higher $\sigma_{1i}$, $\sigma_{2i}$, and $q/p$ are required. The input vector of CNN is normalized as $(x_i^T, v_i^T)^T/$ norm to obtain good control simulation performance where norm = 5 in this study.

*Remark 5:* When $\alpha$ is equal to zero, the agent is unable to acquire new insights from fresh actions. Conversely, when $\alpha$ is set to 1, the agent can disregard previous information and exclusively learn from the most recent data. Therefore, $\alpha$ is consistently maintained within the range of 0 to 1.

In the case of $\gamma$, a value of zero implies that the agent gives no consideration to future rewards and solely focuses on current rewards. Conversely, when $\gamma$ is set to 1, the algorithm prioritizes seeking high rewards over the long term.

The parameter $\epsilon$ introduces an element of randomness into RUKF, encouraging the algorithm to explore different actions and prevent it from becoming trapped in a local optimal solution. A $\epsilon$ value of 0 signifies a lack of exploration, relying solely on existing knowledge. In contrast, a $\epsilon$ value of 1 compels RUKF to consistently select random actions, disregarding any previous information. Typically, $\epsilon$ is chosen as a small value close to 0.

*Remark 6:* When it comes to tuning parameters for the UKF, one of these parameters, $\beta$, has an influence on the estimates of covariance. If the spread is larger, $\beta$ doesn't significantly affect the estimates. This suggests that $\beta$ is not a highly critical parameter that needs adjustment. On the other hand, $\alpha$ doesn't serve to widen the spread of points; its purpose is to bring them closer to the mean. To counteract the scaling effect of a small $\alpha$, a parameter $\kappa$, where $\kappa > 0$, is introduced.

**TABLE 4.** Communication uncertainty and delay profile.

| | Communication Uncertainty | Communication Delay (seconds) |
|---|---|---|
| $\Delta_{ij}$ | $0.4*sin(t)$ | $0.4*sin(t-5)$ |
| $\Delta_{ij}$ | $0.3*\cos(t)$ | $0.3*\cos(t-3)$ |
| $\Delta_{ij}$ | $0.3*sin(2t)$ | $0.3*sin(2t-5)$ |
| $\Delta_{ij}$ | $0.4*\cos(t)$ | $0.4*\cos(t-1)$ |
| $\Delta_{ij}$ | $0.5*sin(t)$ | $0.5*sin(t-5)$ |
| $\Delta_{ij}$ | $0.4*\cos(t)$ | $0.4*\cos(t-6)$ |

### A. UNCERTAINTY IN DYNAMICS MODEL

This section introduces the uncertainty in the MAS model at four different intervals given in Table 1. Numerical simulations are conducted with initial conditions given in Table 2 and parameters given in Table 3 during 50s simulation time, and the external disturbances considered in this case are as $v_i = 0.1\ [sin\ (\frac{t}{2}),\ cos\ (\frac{t}{2})\ ]^T$ throughout the simulation. The simulation reflected changes in the magnitude of the multi-agent dynamics rather than changes in the frequency.

The initial assumption is that no-fault occurs in the time interval between $0 < t \leq 5$ s. The first fault occurs in the time interval between $5s < t \leq 10$ s and the second fault occurs at $10s < t \leq 15$ s. Finally, in the last time interval $t > 15$ s, the MAS is modeled as the same when no-fault occurs between $0 < t \leq 5$ s.

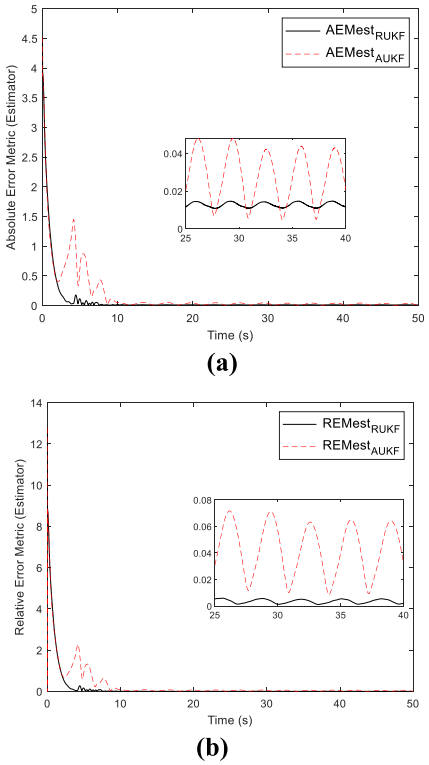


**FIGURE 5.** Performance matrices for estimator under uncertain dynamics: (a) AEMest (b) REMest.

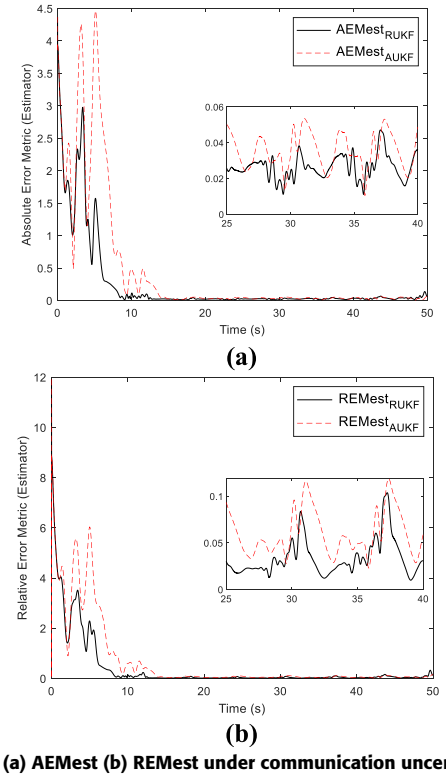


**FIGURE 6.** (a) AEMest (b) REMest under communication uncertainty.

Estimation error for this case is illustrated in Figs. 3(a-b) and 4 (a-b) using AUKF and RUKF, respectively. At the same time, the performance metrices for estimator (AEMest and REMest) can be seen in Fig.5(a-b). Referring to Figs. 3 (a-b) and 4 (a-b), it can be observed that there is an oscillation between 5 to 15s because of the uncertain system model.

It can be noted that the state estimation error is converging to its zero mean after approximately 15s. To illustrate the advantage of the proposed RUKF, the comparison is carried out to compare the performance of our RUKF method with well-known AUKF. The comparison results are shown in Fig.5(a-b). The solid black line represents the RUKF performance, and the red dotted line represents the AUKF performance.

From these figures, it can be concluded that RUKF performances are better leading to better accuracy and estimation performances than AUKF because of the adaptation strategy of UKF. The AEMest and REMest converge to zero after approximately 15s when there is an uncertainty in the system dynamics model, that is, the proposed approach can drive the set of agents to the reference trajectory for the case where only a subset of agents has access to the leader.

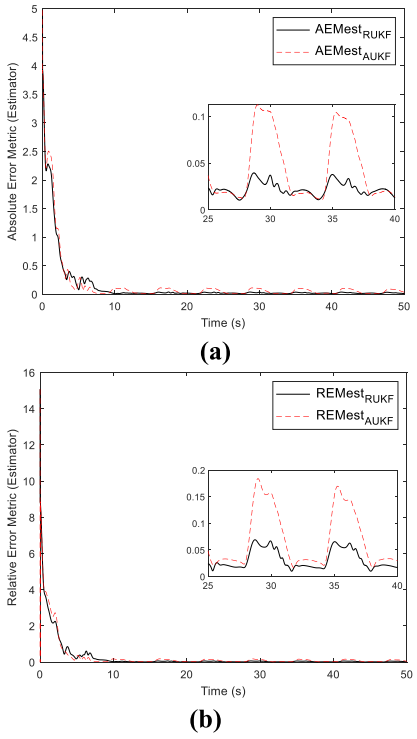


**FIGURE 7.** (a) AEMest (b) REMest under communication delay.

### B. COMMUNICATION UNCERTAINTY AND DELAY

In this case, the uncertainty in the dynamic model is described in Table 1, and the external disturbances are considered as $v_i = 0.1[sin\,(\frac{t}{2}), cos\,(\frac{t}{2})\,]^T$ for this case. The communication topology among the $N$ agents is modeled by a graph $G$, which merely depicts the information flow of the network. The communication uncertainty and delay profile have been stated in Table 4. It has been observed from both Figs. 6 (a-b) and 7 (a-b) that, even in the presence of faults in communication channels, the proposed estimator RUKF outperforms AUKF.

In practice, the information exchange among the agents is conducted over communication channels, which are subject to channel noises and communication constraints. In contrast, RUKF performs well during the simulation, and it converges smoothly.

This case is a general case to see the estimator's performance even in the presence of communication faults. Hence, we can conclude from these figures that RUKF is better compared to AUKF because of the adaptation strategy in UKF which has enhanced the performance in terms of accuracy.

To evaluate the proposed RUKF, comparison results for communication uncertainty have been shown in Fig.6(a-b), and comparison results for communication delay have been shown in Fig.7(a-b). During the simulation, it has been observed that the AUKF oscillates while converging to its zero mean.

## V. CONCLUSION

This paper delves into the challenge of achieving consensus tracking in a second-order uncertain MAS affected by external disturbances and communication faults. The primary innovation of this study lies in the development of an autonomous MAS control framework, which combines RL, nonlinear filtering, CNNs, and Terminal Sliding Mode (TSM). In this proposed control approach, TSM is constructed using collective state errors, and CNNs are leveraged to approximate uncertain nonlinear functions in each agent's dynamics. Furthermore, a RUKF algorithm is applied to estimate the states of uncertain agents. The novel RUKF aims to address the challenging task of tuning of unmanageable noise covariance's in the UKF. The design of the tuning algorithm represents a significant contribution of this research. To mitigate neural network approximation errors, a robust control term based on the hyperbolic tangent function is incorporated.

Of paramount importance is the stability analysis of the RUKF, which operates in tandem with the robust control framework, and this stability is demonstrated using the Lyapunov stability approach. Furthermore, the proposed control scheme is capable of guiding a group of agents to follow a reference trajectory even when faced with uncertain agent dynamics and communication faults. Another noteworthy advantage of the proposed algorithm is its reduced computational time. On average, the RUKF requires 5% less computation time compared to the AUKF. Regardless of the number of states, the RUKF method employs three UKF filters to generate precise estimations, with the standard UKF and GUKF running in parallel. To reduce the computational load, it is advisable to keep the grid size relatively small, as long as it meets the required accuracy criteria. The third filter which is known as RUKF, is utilized based on results obtained from $Q-$learning, the standard UKF, and GUKF.

In summary, it is evident that the use of RUKF leads to improved consensus tracking performance in MAS when compared to AUKF. To conclude, the proposed control framework demonstrates faster convergence and higher consensus tracking capabilities within MAS. In the future, the authors intend to explore autonomous navigation challenges in the context of formation flying spacecraft to further validate the effectiveness of RUKF.

## CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Kaustav Jyoti Borah: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Visualization, Resources, Writing – original draft, Writing – review & editing.

Krishna Dev Kumar: Supervision, review

## DATA AVAILABILITY

There are no data associated with this article.

## DECLARATION OF COMPETING INTEREST

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

The stability analysis of the RUKF is detailed in Appendix A–D, in our recently published paper as referenced in [28].

## APPENDIX A

The mathematical equations of the RUKF can be written as [28],

*Step 1:* Initialize $Q$ ($s$,$a$) arbitrarily, $\forall s \in S$, $a \in A(s)$, and $Q$ (terminal-*state*, $\cdot$ ) = 0

Initialize state estimates, covariance matrices.

Repeat (for each episode):

Initialize$S$

Repeat (for each step of the episode):

Choose $A$ from $S$ using policy derived from $Q$

*Step 2:* Select the sigma points,

Assume that $x_k$ has mean $\overline{x_k}$ and covariance $\overline{p_k}$.

$$X_{0,k} = \overline{x}_k; \tag{A.1}$$

$$X_{i,k} = \overline{x}_k + (a\sqrt{n\overline{p}_k})_i; i = 1, 2...n \tag{A.2}$$

$$X_{i,k} = \overline{x}_k - (a\sqrt{n\overline{p}_k})_i; i = n+1, \ldots, 2n \tag{A.3}$$

where a is proportion parameter, $(\sqrt{n\overline{p}_k})_i$ is the vector of the $i^{th}$ column of the matrix square root.

*Step 3:* Prediction.

$$X_{i,k+1|k} = f(X_{i,k}); \; i = 0, 1, 2...2n \tag{A.4}$$

$$\hat{x}_{k+1|k} = \sum_{i=0}^{2n} \overline{\omega}_i X_{i,k+1|k} \tag{A.5}$$

$$\hat{P}_{k+1|k} = \sum_{i=0}^{2n} \overline{\omega}_i \tilde{X}_{i,k+1|k} \tilde{X}^T_{i,k+1|k} + Q_k \tag{A.6}$$

where, $\tilde{X}_{i,k+1|k} = (X_{i,k+1|k} - \hat{x}_{k+1|k})$, $\overline{\omega}_0 = 1 - \frac{1}{a^2}$; $\overline{\omega}_i = \frac{1}{2na^2}$, $i = 1, 2 \ldots 2n$ are a set of scalar weights, and $\sum_{i=0}^{2n} \overline{\omega}_i = 1$

*Step 4:* Update.

$$z_{i,k+1|k} = h(X_{i,k+1|k}), i = 0, 1, 2...2n \tag{A.7}$$

$$\hat{z}_{i,k+1|k} = \sum_{i=0}^{2n} \overline{\omega}_i z_{i,k+1|k} \tag{A.8}$$

$$\hat{P}_{zz,k+1|k} = \sum_{i=0}^{2n} \overline{\omega}_i \tilde{z}_{i,k+1|k} \tilde{z}^T_{i,k+1|k} + \gamma_{k+1} R_{k+1} + (1-\gamma_{k+1})\sigma^2 I \quad \text{(A.9)}$$

$$\hat{P}_{xz,k+1|k} = \sum_{i=0}^{2n} \overline{\omega}_i \tilde{X}_{i,k+1|k} \tilde{z}^T_{i,k+1|k} \hat{x}_{k+1} = \hat{x}_{k+1|k} + \hat{P}_{xz,k+1|k} [\sum_{i=0}^{2n} \overline{\omega}_i \tilde{z}_{i,k+1|k} \tilde{z}^T_{i,k+1|k} + \gamma_{k+1} R_{k+1} + (1-\gamma_{k+1})\sigma^2 I]^{-1} (z_{k+1|k} - \hat{z}_{k+1|k}) \quad \text{(A.10)}$$

$$\hat{P}_{k+1} = \hat{P}_{k+1|k} - \hat{P}_{xz,k+1} [\sum_{i=0}^{2n} \overline{\omega}_i \tilde{z}_{i,k+1|k} \tilde{z}^T_{i,k+1|k} + \gamma_{k+1} R_{k+1} + (1-\gamma_{k+1})\sigma^2 I]^{-1} \hat{P}^T_{xz,k+1} \quad \text{(A.11)}$$

where,

$\tilde{z}_{i,k+1|k} = z_{i,k+1|k} - \hat{z}_{k+1|k}$. Taking the limit as, $\sigma \to \infty$ then,

$$\hat{x}_{k+1} = \hat{x}_{k+1|k} + \gamma_{k+1} K_{k+1} (z_{k+1|k} - \hat{z}_{k+1|k}) \quad \text{(A.12a)}$$

$$\hat{P}_{k+1} = \hat{P}_{k+1|k} - \gamma_{k+1} K_{k+1} \hat{P}^T_{xz,\,k+1} \quad \text{(A.12b)}$$

$$K_{k+1} = \hat{P}^T_{xz,\,k+1} [\sum_{i=0}^{2n} \overline{\omega}_i \tilde{z}_{i,k+1|k} \tilde{z}^T_{i,k+1|k} + R_{k+1}]^{-1} \quad \text{(A.12c)}$$

We model the measurements by a binary stochastic variable $\gamma_k$. If there are any measurements after $k^{th}$ step, then $\gamma_k = 1$, otherwise $\gamma_k = 0$. So, define the variance of the output noise at time $k$ as $R_k$ if $\gamma_k = 1$ and $\sigma^2 I$ if $\gamma_k = 0$ for some $\sigma^2$ [26].

*Step 5:* Update

$$Q(s,a) \leftarrow Q(\mathrm{s,a}) + \alpha(R_r + \gamma \max_{a'} Q*(s', a') - Q(\mathrm{s,a})), \alpha, \gamma \in [0,1] \quad \text{(A.13)}$$

$\alpha$ regulates the convergence speed, and $\gamma$ determines the future reward.

*Step 6:* Repeat steps 2–5 for the next sample.

Define the error vectors, $\tilde{x}_{k+1}$, $\tilde{x}_{k+1|k}$ and $\tilde{z}_{k+1}$ by

$$\tilde{x}_{k+1} = x_{k+1} - \hat{x}_{k+1}, \tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k}$$
$$\tilde{z}_{k+1} = z_{k+1} - \hat{z}_{k+1|k} \quad \text{(A.14)}$$

For convenience, we use the approach described [26] to simplify the error expression. The unknown matrices $\alpha_k = diag\{\alpha_{1k}, \ldots \alpha_{nk}\}$ and $\beta_k = diag\{\beta_{1k}, \ldots \beta_{nk}\}$ are introduced to model errors due to the first-order linearization technique. Then,

$$\tilde{x}_{k+1|k} = \alpha_k F_k \tilde{x}_{k-1|k} - \alpha_k F_k \gamma_k K_k \tilde{z}_k + \omega_k \quad \text{(A.15)}$$

$$\tilde{z}_{k+1} = \beta_{k+1} H_{k+1} \tilde{x}_{k+1|k} + \nu_{k+1} \quad \text{(A.16)}$$

where,

$F_k = (\frac{\partial f(x)}{\partial x} | x = \hat{x}_k)$ and $H_{k+1} = (\frac{\partial h(x)}{\partial x} | x = \hat{x}_{k+1|k})$ are Jacobian matrices. The actual covariance matrix of the modified UKF is $P_{k+1|k}$. Define, $\partial P_k = \hat{P}_k - P_k$

$$\hat{P}_{k+1|k} = [\alpha_k F_k (I - \gamma_k K_k \beta_k H_k)] \hat{P}_{k|k-1} \times \mathrm{X} [\alpha_k F_k (I - \gamma_k K_k \beta_k H_k)]^T + Q^*_k \quad \text{(A.17)}$$

Where, $Q^*_k = Q_k + \alpha_k F_k \gamma^2_k K_k R_k (\alpha_k F_k \gamma_k K_k)^T + \Delta P_{k+1} + \delta P_{k+1}$ is positive. $\Delta P_k$ is the difference when the mean is removed.

Similarly, the following covariance can be obtained:

$$\hat{P}_{xz,k+1|k} = \hat{P}_{k+1|k} (\beta_{k+1} H_{k+1})^T + \Delta P_{xz,k+1} + \partial P_{xz,k+1} \quad \text{(A.18)}$$

$$\hat{P}_{zz,k+1} = (\beta_{k+1} H_{k+1}) \hat{P}_{k+1|k} (\beta_{k+1} H_{k+1})^T + (1-\gamma_{k+1})\sigma^2 I + R^*_{k+1} \quad \text{(A.19)}$$

$R^*_{k+1} = \gamma_{k+1} R_{k+1} + \Delta P_{zz,k+1} + \partial P_{zz,k+1}$ is positive. Furthermore, we introduce the stochastic matrix $\phi_{k+1} \in R^{\mathrm{m}}$. Then, $\hat{P}_{xz,k+1} = \hat{P}_{k+1|k} \phi_{k+1} (\beta_{k+1} H_{k+1})^T$ and it is easy to get that,

$$\hat{P}_{k+1} = \hat{P}_{k+1|k} - \hat{P}_{k+1|k} \phi_{k+1} H^T_{k+1} \beta_{k+1} [\beta_{k+1} H_{k+1}) X \hat{P}_{k+1|k} (\beta_{k+1} H_{k+1})^T + (1-\gamma_{k+1})\sigma^2 I + R^*_{k+1}]^{-1} \mathrm{X} (\hat{P}_{k+1|k} \phi_{k+1} H^T_{k+1} \beta_{k+1})^T \quad \text{(A.20)}$$

Taking the limit as $\sigma \to \infty$, the following can be obtained:

$$\hat{P}_{k+1} = \hat{P}_{k+1|k} - \gamma_{k+1} K_{k+1} (\hat{P}_{k+1|k} \phi_{k+1} H^T_{k+1} \beta_{k+1})^T \quad \text{(A.21)}$$

$$K_{k+1} = \hat{P}_{k+1} \phi_{k+1} H^T_{k+1} \beta_{k+1} [(\beta_{k+1} H_{k+1}) \hat{P}_{k+1|k} \times (\beta_{k+1} H_{k+1})^T + R^*_{k+1}]^{-1} \quad \text{(A.22)}$$

## APPENDIX B [28]

*Proof:* Define the Lyapunov candidate can be described as

$$V_{k+1}(\tilde{\chi}_{k+1|k}) = \tilde{\chi}^T_{k+1|k} \hat{P}^{-1}_{k+1|k} \tilde{\chi}_{k+1|k}. \quad \text{(B.1)}$$

the assumptions in *Theorem 2* imply that $\underline{p}\underline{\alpha}^2 \underline{f}^2 I \le \hat{P}_{k+1|k} \le \overline{p}\,\overline{\alpha}^2 \overline{f}^2 I + \overline{q} I$. Then we have $\frac{\|\tilde{\chi}_{k+1|k}\|^2}{\overline{p}\,\overline{\alpha}^2 \overline{f}^2 I + \overline{q} I} \le V_{k+1}(\tilde{\chi}_{k+1|k}) \le \frac{\|\tilde{\chi}_{k+1|k}\|^2}{\underline{p}\underline{\alpha}^2 \underline{f}^2 I}$. In considering the white noise property, the following statements are in force:

$$\begin{aligned} E\{V_{k+1}(\tilde{\chi}_{k+1|k}) | \tilde{\chi}_{k|k-1}, \Upsilon_{1:k}\} &= \tilde{\chi}^T_{k|k-1} [\alpha_k F_k (1 - \Upsilon_k K_k \beta_k H_k)]^T \hat{P}^{-1}_{k+1|k} \\ &\mathrm{X} [\Upsilon_k F_k (1 - \Upsilon_k K_k \beta_k H_k)] \tilde{\chi}_{k|k-1} \\ &+ E\{v^T_k [\alpha_k F_k \Upsilon_k K_k]^T \hat{P}^{-1}_{k+1|k} [\alpha_k F_k \Upsilon_k K_k] v_k | \tilde{\chi}_{k|k-1}\} \\ &+ E\{\omega^T_k \hat{P}^{-1}_{k+1|k} \omega_k | \tilde{\chi}_{k|k-1}\} \end{aligned} \quad \text{(B.2)}$$

(I) From Eq. (A.17), we obtain,

$$\hat{P}_{k+1|k} = [\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)] \hat{P}_{k|k-1} + [\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)]^{-1} Q^*_k X [\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)]^{-T} \times [\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)]^T \quad \text{(B.3)}$$

Based on Eq.(21), it can be verified that $\|\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k) Q^{*-T}_k [\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)]^T\| \le (\overline{\alpha f} + \overline{\alpha f} K^* \overline{\beta h})^2 \|Q^{*-T}_k\|$

Then, $[\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)]^{-1} Q^*_k [\alpha_k F_k (I - \Upsilon_k K_k \beta_k H_k)]^{-T} \ge \frac{q^* I}{(\overline{\alpha f} + \overline{\alpha f} K^* \overline{\beta h})^2}$, where $\|K_{k+1}\| \le \frac{\overline{p}\,\overline{\phi}\,\overline{h}\,\overline{\beta}}{\underline{p}^2 \underline{h}^2 p + r^*} = K^*$ follows from Eqs. (A.22) and (20). Putting this into Eq. (B.3), and

taking the inverse operation on both sides of it, it is clear that, $[\alpha_k F_k\,(I-\Upsilon_k K_k \beta_k H_k)\,]^T \hat{P}^{-1}_{k+1|k}[\alpha_k F_k\,(I-\Upsilon_k K_k \beta_k H_k)] \leq \left[1+\frac{q^* I}{(\overline{\alpha f}+\overline{\alpha f} K^* \overline{\beta h})^2-\underline{p}}\right]\hat{P}^{-1}_{k|k-1}$.

Let $1-\lambda = \left[1+\frac{q^* I}{(\overline{\alpha f}+\overline{\alpha f} K^* \overline{\beta h})^2-\underline{p}}\right]$; the first term of $E\{V_{k+1}\left(\tilde{\chi}_{k+1|k}\right)|\tilde{\chi}_{k|k-1}, \Upsilon_{1:k}\}$ is bounded.

(II) Under the assumptions, it can be derived that,

$$\begin{aligned} &E\{v_k^T[\alpha_k F_k \Upsilon_k K_k]^T \hat{P}^{-1}_{k+1|k}[\alpha_k F_k \Upsilon_k K_k] v_k \\ &\quad + \omega_k^T \hat{P}^{-1}_{k+1|k}\omega_k|\tilde{\chi}_{k|k-1}\} \\ &\leq E\{\frac{K^{*2}\overline{\alpha}^2\overline{f}^2\overline{q}}{\underline{p}} tr\left\{v_k^T v_k\right\} + \frac{1}{\underline{p}} tr\left\{\omega_k^T \omega_k\right\} \\ &= \frac{K^{*2}\overline{\alpha}^2\overline{f}^2\overline{q}}{\underline{p}} n_v + \frac{\overline{r}}{\underline{p}} n_\omega = \mu \end{aligned} \tag{B.4}$$

Following the above manipulation, we can get that $E\{V_{k+1}\left(\tilde{\chi}_{k+1|k}\right)\Big|\tilde{\chi}_{k|k-1}\} - V_k\left(\tilde{\chi}_{k|k-1}\right) \leq -\lambda V_k\left(\tilde{\chi}_{k|k-1}\right) + \mu$. Applying *Lemma 2.1*from [27], the stochastic process $\tilde{\chi}_{k+1|k}$ is bounded.

(III) In the following, we will show that the stochastic process is bounded by the mean square. By Eq. (A.15), it is yielded that $\tilde{\chi}_{k+1}$ is bounded in mean square. By Eq. (A.15), it is yielded that $\|\tilde{\chi}_k\|^2 \leq 2\underline{\alpha}^2\underline{f}^2(\|\tilde{\chi}_{k+1|k}\|^2 + \|\omega_k\|^2)$. Taking the mean value from both sides yields $E\{\|\tilde{\chi}_k\|^2\} \leq 2\underline{\alpha}^2\underline{f}^2(E\{\|\tilde{\chi}_{k+1|k}\|^2\} + E\{\|\omega_k\|^2\})$.$\omega_k$ is bounded in the mean square with an approach similar to that above. Therefore, ***we conclude that the estimation error*** $\tilde{\chi}_{k+1}$ ***asymptotically stable.***

## APPENDIX C [28]

The absolute state reference error of the $i$ ($i$ = *1,2 3,…6*) agents is defined by,

$$e_{\hat{x}_i} = \hat{x}_i - x_0 \tag{C.1}$$

$$e_{\hat{v}_i} = \dot{\hat{x}}_i - x_0 \tag{C.2}$$

The dynamic equations for $e_{\hat{x}_i}$ and $e_{\hat{v}_i}$ can be obtained using Eqs. (1) and (2) as

$$\dot{e}_{\hat{x}_i} = e_{\hat{v}_i} \tag{C.3}$$

$$\dot{e}_{\hat{v}_i} = -\hat{x}_0 + \hat{f}_i\left(\hat{x}_i, \hat{v}_i\right) + g_i\left(\hat{x}_i\right) u_i + \vartheta_i \tag{C.4}$$

Now, this can be represented in compact form as,

$$\dot{e}_{\hat{x}} = e_{\hat{v}} \tag{C.5}$$

$$\dot{e}_{\hat{v}} = -X_0 + \hat{F} + g\left(\hat{x}\right) u + \vartheta \tag{C.6}$$

where,

$$\begin{aligned} e_{\hat{x}} &= (e_{\hat{x}_1}^T, e_{\hat{x}_2}^T, e_{\hat{x}_3}^T \ldots\ldots\ldots\ldots, e_{\hat{x}_n}^T)^T \in R^{mn}, \\ e_{\hat{v}} &= (e_{\hat{v}_1}^T, e_{\hat{v}_2}^T, e_{\hat{v}_3}^T \ldots\ldots\ldots\ldots, e_{\hat{v}_n}^T)^T \in R^{mn} \\ F &= (\hat{f}_1^T, \hat{f}_2^T, \hat{f}_3^T \ldots\ldots\ldots\ldots, \hat{f}_n^T)^T \in R^{mn}, \\ X_0 &= (\ddot{x}_0^T, \ddot{x}_0^T, \ddot{x}_0^T \ldots\ldots\ldots\ldots, \ddot{x}_0^T)^T \in R^{mn} \\ u &= (u_1^T, u_2^T, u_3^T \ldots\ldots\ldots\ldots, u_n^T)^T \in R^{mn}, \end{aligned}$$

$$\begin{aligned} \hat{v} &= (\hat{v}_1^T, \hat{v}_2^T, \hat{v}_3^T \ldots\ldots\ldots\ldots, \hat{v}_n^T)^T \in R^{mn} \\ g\left(\hat{x}\right) &= diag\,(g_1, g_2, \ldots\ldots\ldots g_n) \in R^{(mn)x(mn)}, \\ \hat{x} &= (\hat{x}_1^T, \hat{x}_2^T, \hat{x}_3^T \ldots\ldots\ldots\ldots, \hat{x}_n^T)^T \in R^{mn} \end{aligned} \tag{C.7}$$

The relative estimated state errors between the $i$ ($i$ = *1,2,3…6*)) and $j$ ($j$ = *1,2,3…6*) follower agents are defined now as,

$$r_{\hat{x}_{ij}} = \hat{x}_i - \hat{x}_j \tag{C.8}$$

$$r_{\hat{v}_{ij}} = \hat{v}_i - \hat{v}_j \tag{C.9}$$

Based on the discussion provided in [25], each agent in the group has access to the neighboring agents the lumped estimated state error can be written as,

$$\left.\begin{aligned} \alpha_{\hat{x}_i} &= \textstyle\sum_{J=1}^n a_{ij} r_{\hat{x}_{ij}} + b_i e_{\hat{x}_i} \\ &= \textstyle\sum_{J=1}^n a_{ij}(e_{\hat{x}_i} - e_{\hat{x}_j}) + b_i e_{\hat{x}_i} \end{aligned}\right\} \tag{C.10}$$

$$\left.\begin{aligned} \alpha_{\hat{v}_i} &= \textstyle\sum_{J=1}^n a_{ij} r_{\hat{v}_{ij}} + b_i e_{\hat{v}_i} \\ &= \textstyle\sum_{J=1}^n a_{ij}(e_{\hat{v}_i} - e_{\hat{v}_j}) + b_i e_{\hat{v}_i} \end{aligned}\right\} \tag{C.11}$$

where, $a_{ij}$ is the element of the weighted adjacency matrix A. From Eqs. (C.10) and (C.11), to ease the succeeding conceptual analysis, the lumped state errors $\alpha_{\hat{x}} \epsilon R^m$ and $\alpha_{\hat{v}} \epsilon R^m$ ($i$ = *1,2,3…6*) can be expressed in terms of the absolute estimate state errors $e_{\hat{x}_i}$ and $e_{\hat{v}_i}$,

$$\alpha_{\hat{x}_i} = \sum_{J=1}^n l_{ij} e_{\hat{x}_j} + b_i e_{\hat{x}_i} \tag{C.12}$$

$$\alpha_{\hat{v}_i} = \sum_{J=1}^n l_{ij} e_{\hat{v}_j} + b_i e_{\hat{v}_i} \tag{C.13}$$

where, $l_{ij}$ is the element of the graph Laplacian matrix L.

Define,

$$\alpha_{\hat{x}} = (\alpha_{\hat{x}_1}^T, \alpha_{\hat{x}_2}^T, \alpha_{\hat{x}_3}^T \ldots\ldots\ldots\ldots, \alpha_{\hat{x}_n}^T)^T \in R^{mn} \tag{C.14}$$

$$\alpha_{\hat{v}} = (\alpha_{\hat{v}_1}^T, \alpha_{\hat{v}_2}^T, \alpha_{\hat{v}_3}^T \ldots\ldots\ldots\ldots, \alpha_{\hat{v}_n}^T)^T \in R^{mn} \tag{C.15}$$

Then the lumped state estimation error Eqs. (C.12) and (C.13) can be expressed as,

$$\alpha_{\hat{x}} = P_1 e_{\hat{x}} \tag{C.16}$$

$$\alpha_{\hat{v}} = P_1 e_{\hat{v}} \tag{C.17}$$

where, $P_1 = (L+B) \otimes I_m \in R^{(mn)x(mn)}$ is the Kronecker product. Using Eqs. C.(5) and C.(6) the dynamic equations for $\alpha_{\hat{x}}$ are $\alpha_{\hat{v}}$ are given by,

$$\Rightarrow \dot{\alpha}_{\hat{x}} = \alpha_{\hat{v}} \tag{C.18}$$

$$P_2 \dot{\alpha}_{\hat{v}} == -X_0 + F + gu + \vartheta \tag{C.19}$$

Respectively, where $P_2 = P_1^{-1}$

The fast terminal sliding manifold $s_i \epsilon$ $R^m$ ($i$ = *1,2,3…6*) is defined based on [25] as,

$$s_i = \alpha_{\hat{v}_i} + \sigma_{1i}\alpha_{\hat{x}_i} + \sigma_{2i}\beta_i(\alpha_{\hat{x}_i}) \tag{C.20}$$

where, $\sigma_{1i}$ and $\sigma_{2i}$ are positive constants and $\beta_i\left(\alpha_{\hat{x}_i}\right) = (\beta_{i1}\left(\alpha_{\hat{x}_{i,1}}\right), \ldots..\beta_{im}\left(\alpha_{\hat{x}_{i,m}}\right))^T \in R^m$

is defined by,

$$\beta_{ij}\left(\alpha_{\hat{x}_{i,j}}\right)=\begin{cases}\alpha_{\hat{x}_{i,j}}^{\frac{P}{q}}, \text{If } \overline{s_{ij}}=0 \text{ or } \overline{s_{\dot{s}j}}\neq 0, \left|\alpha_{\hat{x}_{i,j}}\right|>u\& \\ l_1\alpha_{\hat{x}_{i,j}}+l_2 sgn(\alpha_{\hat{x}_{i,j}})\alpha_{\hat{x}_{i,j}}^2, \text{ if } \overline{s_{ij}}\neq 0, \\ \left|\alpha_{\hat{x}_{i,j}}\right|\leq u\end{cases} \tag{C.21}$$

With $j = 1,2\ldots m$, $l_1 = (2-\frac{p}{q})\mu^{\frac{p}{q}-1}$, $l_2 = (\frac{p}{q}-1)\mu^{\frac{p}{q}-2}$, $\mu$ is const.(small positive), sgn(.) is the signum function, and $\bar{s}_{ij} = \alpha_{\hat{v}_{ij}} + \sigma_{ij}\alpha_{\hat{x}_{ij}} + \sigma_{2i}\alpha_{\hat{x}_{ij}}^{\frac{p}{q}}$. The time derivative of the $s_i$ in Eq. (C.20) is given by,

$$\dot{s}_i = \dot{\alpha}_{\hat{v}_i} + \sigma_{1i}\alpha_{\hat{v}_i} + \sigma_{2i}\dot{\beta}_i, \quad i = 1, 2, \ldots, 6 \tag{C.22}$$

where, $\dot{\beta}_i \in R^{\mathrm{m}}$ can be written as,

$$\dot{\beta}_{ij}=\begin{cases}\frac{P}{q}\alpha_{\hat{x}_{i,j}}^{\frac{p}{q}-1}\alpha_{\hat{v}_{i,j}}, \text{If } \overline{s_{ij}}=0 \text{ or } \overline{s_{\dot{s}j}}\neq 0, \\ \left|\alpha_{\hat{x}_{i,j}}\right|>u\& l_1\alpha_{\hat{v}_{i,j}}+2l_2 sgn(\alpha_{\hat{x}_{i,j}})\alpha_{\hat{x}_{i,j}}\alpha_{\hat{v}_{i,j}} \\ \text{if } \overline{s_{ij}}\neq 0, \left|\alpha_{\hat{x}_{i,j}}\right|\leq u\end{cases} \tag{C.23}$$

with $j = 1,2\ldots 6$. The terminal sliding manifold in Eq. (C.22) can be written as,

$$s = \alpha_{\hat{v}} + \sigma_1\alpha_{\hat{x}} + \sigma_2\beta \tag{C.24}$$

where,

$$\begin{aligned}s &= (s_1^T, s_2^T, s_3^T \ldots\ldots\ldots\ldots, s_{\mathrm{n}}^T)^T \in R^{\mathrm{mn}},\\ \beta &= (\beta_1^T, \beta_2^T, \beta_3^T \ldots\ldots\ldots\ldots, \beta_{\mathrm{n}}^T)^T \in R^{\mathrm{mn}}\\ \sigma_1 &= (\sigma_{11}I_m, \sigma_{12}I_m, \ldots\ldots\ldots\ldots\sigma_{1n}I_m) \in R^{(mn)x(mn)},\\ \sigma_2 &= (\sigma_{21}I_m, \sigma_{21}I_m, \ldots\ldots\ldots\ldots\sigma_{2n}I_m) \in R^{(mn)x(mn)}\end{aligned}$$

*Lemma C.1:* If the sliding manifold $s = \bar{s} = 0$ is reached where $\bar{s} = (\bar{s}_{11}, \bar{s}_{1m}, \ldots, \bar{s}_{n1}, \ldots., \bar{s}_{nm})^T$ then the absolute estimated state error $e_{\hat{x}}$ converges.

*Proof:* Lyapunov candidate,

$$V = 1/2\alpha_{\hat{x}}^T\alpha_{\hat{x}} \tag{C.25}$$

If $s = \bar{s} = 0$ is reached, then.

$$\alpha_{\hat{v}} = -\sigma_1\alpha_{\hat{x}} - \sigma_2\alpha_{\hat{x}}^{v} \tag{C.26}$$

where, $\alpha_x^v \in R^{\mathrm{mn}}$is defined as

$$\alpha_{\hat{x}}^{v} = (\alpha_{\hat{x}1,1}^{v}, \ldots..\alpha_{\hat{x}1,\mathrm{m}}^{v}, \ldots..\alpha_{\hat{x}\mathrm{n},1}^{v}, \ldots\ldots\ldots.., \alpha_{\hat{x}\mathrm{n},\mathrm{m}}^{v})^T$$

and $v = p/q$

The time derivative of $V$ is defined by Eqs. (C.25) along (C.18) and (C.26) results in,

$$\dot{V} = -\alpha_{\hat{x}}^T\sigma_1\alpha_{\hat{x}} - \alpha_{\hat{x}}^T\sigma_2\alpha_{\hat{x}}^{v} \tag{C.27}$$

which is less than zero, so the lumped estimated state error converges [25]. Note that,

$$\alpha_{\hat{x}} = P_1 e_{\hat{x}} \tag{C.28}$$

and it implies, $\mathbf{e}_{\hat{x}}$ ***converges to zero in finite time.***

*Remark 7:* This ***Lemma C.1*** proves that when the dynamic system reaches the manifold then it is stable. It will be stable if it is in the sliding manifold. Thus, one can design a control law that will steer the system to the sliding manifold. Again, the time derivative of the sliding manifold can be rewritten as,

$$\dot{s} = \dot{\alpha}_{\hat{v}} + \sigma_1\alpha_{\hat{v}} + \sigma_2\dot{\beta} \tag{C.29}$$

where $\dot{\beta} = (\dot{\beta}_1^T, \dot{\beta}_2^T, \ldots., \dot{\beta}_n^T)^{\mathrm{T}} \epsilon R^{\mathrm{mn}}$. In Eq. (C.29) both sides we would like to multiply $P_2$ and apply Eq. (C.19) we obtain,

$$\begin{aligned}P_2\dot{s} = P_2(\dot{\alpha}_{\hat{v}} + \sigma_1\alpha_{\hat{v}} + \sigma_2\dot{\beta}) = P_2\left(\sigma_1\alpha_{\hat{v}} + \sigma_2\dot{\beta}\right)\\ - X_0 + \hat{F} + gu + \vartheta\end{aligned} \tag{C.30}$$

If all the follower agent states approach the reference trajectory then,

$$P_2(\sigma_1\alpha_{\hat{v}} + \sigma_2\dot{\beta}) = 0$$

Now applying CNN approximation property [25],

$$\hat{f}_i\left(\hat{x}_i, \hat{v}_i\right) = W_i^*\xi_i\left(\hat{x}_i, \hat{v}_i\right) + \varepsilon_i \tag{C.31}$$

where, $W_i^*\epsilon R^{m\times N_1}$ with $N_1 = 2\mathrm{m}N_2 + 1$ and $N_2$ is Chebyshev polynomials order is the optimal weight matrix in the sense that it will optimize the weight of the CNN by minimizing the residual of the estimated plant in terms of accuracy, the, $\varepsilon_i \epsilon$ $\mathrm{R}^{\mathrm{m}}$ is the CNN approximation error and $\xi_i(X_i) \epsilon R^{N_1}$ is the Chebyshev polynomial basis function where, $X_i = (\hat{x}_i^T, \hat{v}_i^T)^{\mathrm{T}}$

$$\begin{aligned}&\xi_i(X_i)\\ &= (1, T_1\left(X_{i,1}\right), \ldots, T_{N_2}\left(X_{i,1}\right), \ldots, T_1\left(X_{i,2m}\right), \ldots T_{N_2}\left(X_{i,2m}\right))^T\end{aligned} \tag{C.32}$$

where, $T_i\left(X_{i,1}\right)(i = 1, 2\ldots.N_2, J = 1, 2\ldots 2m)$ represent Chebyshev polynomials, obtained by using the so-called two-term recursive formula as,

$$T_{i+1}\left(X_{i,j}\right) = 2X_{i,j}T_i\left(X_{i,j}\right) - T_{i-1}\left(X_{i,j}\right), T_0\left(X_{i,1}\right) = 1 \tag{C.33}$$

$T_i\left(X_{i,j}\right)$ has several definitions such as $X_{i,j}, 2X_{i,j}, 2X_{i,j} - 1$ and $2X_{i,j} + 1$. $T_1\left(X_{i,j}\right)$ is $\left(X_{i,j}\right)$.The following assumptions C.1 and C.2 holds.

*Assumption C.1:* The optimal weight matrix $W_i^*(i = 1, 2, \ldots\ldots., 6)$ is in a known bounded set $\Omega_{\omega_i}$ that is defined by,

$$\begin{aligned}W_i^*\epsilon\Omega_{W_i} = \Big\{W_i^* : W_{imin} \leq W_{i,jk}^* \leq W_{imax},\\ j = 1, 2, \ldots., 6, k = 1, 2, \ldots., N_1\}\end{aligned} \tag{C.34}$$

where, $W_{imin}$and $W_{imax}$ are constants (known).

*Assumption C.2:*

The CNN approximation error $\varepsilon_i(i = 1, 2, \ldots\ldots., 6)$ is bounded such that $\|\varepsilon_i\| \leq \varepsilon_{M_i}$ ($\varepsilon_{M_i}$ is positive constant).

To update CNN weight and get the estimated weights within some bounded sets, a smooth projection algorithm

[25] is implemented. Let the estimated weight matrix for $W_i^*(i = 1, 2, \ldots\ldots, 6)$ be $W_i$ defined as,

$$\pi_i(W_i) = W_{\pi i} = (\pi_{i,jk}(W_{i,jk})) \tag{C.35}$$

Where, $j = 1, 2, \ldots, 6, k = 1, 2, \ldots, N_1$ and the projection operator $\pi_{i,Jk}: R \to R$ is a real value smooth nondecreasing function defined by,

$$\pi_{i,jk}(W_{i,jk}) = W_{i,jk}, \forall W_{i,jk} \epsilon [W_{imin}, W_{imax}]$$
$$\pi_{i,jk}(W_{i,jk}) \epsilon [W_{imin} - \varepsilon w_i, W_{imax} + \varepsilon w_i], \forall W_{i,jk} \epsilon R \tag{C.36}$$

Where, $\varepsilon w_i$ is positive constant (small) which is determined by the designer of the controller. The control law has been defined in Eq. (7)

Define,

$\tilde{w}_i = W_i^* - \hat{w}_i (i = 1, 2, \ldots\ldots, 6)$, $\tilde{w}_{\pi i} = W_i^* - \hat{W}_{\pi i}$, and

$$V_{W_i} = -1/\delta_i \sum_{j=1}^{m} \sum_{k=1}^{N_1} \int_0^{\tilde{W}_{i,jk}} (W_{i,jk}^* - \pi_{i,jk}(W_{i,jk}^* - \omega_{i,jk})) \times d\omega_{i,jk} \tag{C.37}$$

Hence, the time derivative

$$\dot{V}_{W_i} = -1/\delta_i \sum_{j=1}^{m} \sum_{k=1}^{N_1} \tilde{W}_{\pi i,jk} \dot{W}_{i,jk} \tag{C.38}$$

where the adaptive law [25] $\dot{W}_{i,jk} = \delta_{i,jk} s_{i,jk} \xi_{i,jk}^T$ and $\delta_i$ is a positive constant.

Substituting the control law defined in Eq. (7) in Eq. (C.30), we get,

$$P_2 \dot{s} = P_2(\sigma_1 \alpha_{\hat{\gamma}} + \sigma_2 \dot{\beta}) - X_0 + F - W_\pi \xi - K_1 s - K_2 s^\nu - \psi + \vartheta$$
$$= \chi + \tilde{W}_\pi \xi - K_1 s - K_2 s^\nu - \Psi + \vartheta + \varepsilon - X_0 \tag{C.39}$$

where $\chi = P_2(\sigma_1 \alpha_{\hat{\gamma}} + \sigma_2 \dot{\beta}) \epsilon R^{\text{mn}}$ and $W^* = \text{diag}(W_1^*, W_2^*, \ldots, W_n^*)$; $W_\pi = \text{diag}(W_{\pi 1}, W_{\pi 1}, \ldots W_{\pi n})$

$\tilde{W}_\pi = W^* - W_\pi$, $\xi = \{\xi_1^T, \xi_2^T, \ldots\ldots, \xi_n^T\}^{\text{T}}$, $K_1 = \text{diag}\{K_{11}, K_{12}, \ldots, K_{1n}\}$, $K_2 = \text{diag}\{K_{21}, K_{22}, \ldots, K_{2n}\}$, $\psi = (\psi_1^T, \psi_2^T, \ldots \psi_n^T)^{\text{T}}$, $\varepsilon = (\varepsilon_1^T, \varepsilon_2^T, \ldots, \varepsilon_n^T)^{\text{T}}$

To set up a compress set for adequate large positive constant, $s_{max}$,

$$\Omega_s = \{s | s^T s \le \frac{2 s_{max}}{\lambda_{max}(P_2)}\} \tag{C.40}$$

where, $\lambda_{max}(\cdot)$ denotes the maximal eigenvalue of a matrix reason the sets $\Omega_0$and $\Omega_s$ are compact.

Next, the robust controller $\psi_i$(*i=1,2,..6*) can be defined as [25],

$$\Psi_{ij} = k_i \tanh(\frac{m k_u k_i s_{ij}}{\epsilon}), k_u = 0.2785, j = 1, 2 \ldots 6 \tag{C.41}$$

Here, $k_i$ is a positive constant satisfying $k_i \ge \Delta_{Mi} + \varepsilon_{Mi} + \|\ddot{x}_0\|$ and $\epsilon$ are positive scalars. We obtained the following inequality w.r.t. robust controller $\psi_i$ (*i=1,2..6* ) from [25] as,

$$0 \le s_i^T (v_i + \epsilon_i - \ddot{x}_0) - s_i^T \psi_i$$
$$\le \sum_{j=1}^{n} |s_{ij}| (v_{Mi} + \varepsilon_{Mi} - \|\ddot{x}_0\|) - \psi \sum_{j=1}^{m} (s_{ij} \Psi_{ij}) \le \varepsilon \tag{C.42}$$

## APPENDIX D [28]

*Proof:* Let us consider the Lyapunov candidate defined in Eq.(19) as

$$V(s, w) = \frac{1}{2} s^T P_2 s + \sum_{i=1}^{n} V_{wi}$$

The time derivative of V is,

$$\dot{V}(s,w) = s^T x + s^T \tilde{w}_\pi \xi - s^T k_1 s - s^T k_2 s^\nu - s^T \psi + s^T (\vartheta + \epsilon - X_0) - \sum_{i=1}^{n} (1/\delta_i \sum_{j=1}^{m} \sum_{k=1}^{N_1} \tilde{W}_{\pi i,jk} \dot{W}_{i,jk})$$
$$= s^T x - s^T k_1 s - s^T k_2 s^\nu - s^T \psi + s^T (\vartheta + \epsilon - X_0) \tag{D.1}$$

where,

$s^T \tilde{w}_\pi \xi - \sum_{i=1}^{n} (1/\delta_i \sum_{j=1}^{m} \sum_{k=1}^{N_1} \tilde{W}_{\pi i,jk} \dot{W}_{i,jk}) = 0$ is applied.

By using Eq. (C.42), the following inequality,

$$0 \le s^T (\vartheta + \epsilon - X_0) - s^T \psi$$
$$= \sum_{i=1}^{n} (s_i^T (v_i + \varepsilon_i - \ddot{x}_0) - s_i^T \psi_i) \le n\epsilon \tag{D.2}$$

Can be obtained, and by using the well-known inequality,

$$(\sqrt{C}_2 s - \frac{x}{2\sqrt{C}_2})^{\text{T}} (\sqrt{C}_2 s - \frac{x}{2\sqrt{C}_2}) \ge 0 \tag{D.3}$$

The following inequality can be guaranteed,

$$s^T x \le C_2 s^T s + \frac{x_M^2}{4C_2} \tag{D.4}$$

where $C_2$ is a positive constant satisfying $C_2 < \lambda_{min}(\text{K}_1)$ and $\lambda_{min}(\cdot)$ is the eigen value of a matrix (min.).

using inequalities Eqs. (D.2) and (D.4) to (D.1), $\dot{V}(s,w)$ can be upper bounded by,

$$\dot{V}(s,w) \le -(\lambda_{min}(K_1) - C_2) s^T s + C_3 \tag{D.5}$$

where, $C_3 = n\epsilon + X_M^2 / 4C_2$ is a positive constant. Thus $\dot{V}(s,w)$is strictly negative outside the following compact set $\Omega_{s1}$

$$\Omega_{s1} = \{ s(t) \ | \|s(t)\| \le \sqrt{\frac{c_3}{\lambda_{min}(\text{K}_1) - C_2}} \} \tag{D.6}$$

We can now conclude that, when $s$ is outside the compact set $\Omega_{s1}$, $\|s\|$ decreases and $s$ is uniformly bounded.

## REFERENCES

[1] V. Vahidpour, A. Rastegarnia, A. Khalili, and S. Sanei, "Partial diffusion Kalman filtering for distributed state estimation in multiagent networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3839–3846, Dec. 2019.

[2] D.-D. Zhou, B. Hu, Z.-H. Guan, D.-X. Zhang, and X.-M. Cheng, "Consensus tracking control of uncertain multiagent systems with sampled data and time-varying delay," *IEEE Trans. Cybern.*, vol. 51, no. 12, pp. 5681–5691, Dec. 2021.

[3] Y. Zhong, G. Lyu, X. He, Y. Zhang, and S. S. Ge, "Distributed active fault-tolerant cooperative control for multiagent systems with communication delays and external disturbances," *IEEE Trans. Cybern.*, vol. 53, no. 7, pp. 4642–4652, Jul. 2023, doi: 10.1109/TCYB.2021.3133463.

[4] M. Liu and Z. Li, "Robust consensus of Lur'e networks with uncertain communications," *IET Control Theory Appl.*, vol. 11, no. 6, pp. 877–882, Apr. 2017.
[5] W. Zhang, K. I. Kou, J. Lou, and Y. Liu, "Observer based consensus for nonlinear multi-agent systems with communication failures," *Neurocomputing*, vol. 173, pp. 1034–1043, Jan. 2016.
[6] M. Xing and F. Deng, "Scaled consensus for multi-agent systems with communication time delays," *Trans. Inst. Meas. Control*, vol. 40, no. 8, pp. 2651–2659, May 2018.
[7] Z. Li and J. Chen, "Robust consensus for multi-agent systems communicating over stochastic uncertain networks," 2017, *arXiv:1703.09877*.
[8] Z. Li and J. Chen, "Robust consensus of linear feedback protocols over uncertain network graphs," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 4251–4258, Aug. 2017.
[9] J. Yang, "A consensus control for a multi-agent system with unknown time-varying communication delays," *IEEE Access*, vol. 9, pp. 55844–55852, 2021.
[10] Z. Ahmed, M. M. Khan, M. A. Saeed, and Z. Weidong, "Consensus control of multi-agent systems with input and communication delay: A frequency domain perspective," *ISA Trans.*, vol. 101, pp. 69–77, Jun. 2020.
[11] W. Wang, C. Li, Y. Sun, and G. Ma, "Distributed coordinated attitude tracking control for spacecraft formation with communication delays," *ISA Trans.*, vol. 85, pp. 97–106, Feb. 2019.
[12] H. L. Trentelman, K. Takaba, and N. Monshizadeh, "Robust synchronization of uncertain linear multi-agent systems," *IEEE Trans. Autom. Control*, vol. 58, no. 6, pp. 1511–1523, Jun. 2013.
[13] Z. Zheng, J. Zhao, L. Mili, Z. Liu, and S. Wang, "Unscented Kalman filter-based unbiased minimum-variance estimation for nonlinear systems with unknown inputs," *IEEE Signal Process. Lett.*, vol. 26, no. 8, pp. 1162–1166, Aug. 2019.
[14] C. Liu and Y. Wang, "Distributed unscented Kalman filters for nonlinear multi-agent systems with homologous unknown inputs," in *Proc. 39th Chin. Control Conf. (CCC)*, Jul. 2020, pp. 4831–4836.
[15] H. Lin, C. Hu, Z. Deng, and G. Liu, "Distributed Kalman filter with fuzzy noises over multiagent systems," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 7, pp. 2550–2562, Jul. 2022.
[16] C. Liu, Y. Wang, D. Zhou, and X. Shen, "Minimum-variance unbiased unknown input and state estimation for multi-agent systems by distributed cooperative filters," *IEEE Access*, vol. 6, pp. 18128–18141, 2018.
[17] W. Zhou, X. Li, J. Yi, and H. He, "A novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1415–1424, Mar. 2019.
[18] M. Cai, Z. Xiang, and J. Guo, "Adaptive finite-time consensus protocols for multi-agent systems by using neural networks," *IET Control Theory Appl.*, vol. 10, no. 4, pp. 371–380, Feb. 2016.
[19] M. Khalili, X. Zhang, M. M. Polycarpou, T. Parisini, and Y. Cao, "Distributed adaptive fault-tolerant control of uncertain multi-agent systems," *Automatica*, vol. 87, pp. 142–151, Jan. 2018.
[20] Y. Liu and G.-H. Yang, "Neural learning-based fixed-time consensus tracking control for nonlinear multiagent systems with directed communication networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 639–652, Feb. 2021.
[21] G. Dong, H. Li, H. Ma, and R. Lu, "Finite-time consensus tracking neural network FTC of multi-agent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 653–662, Feb. 2021, doi: 10.1109/TNNLS.2020.2978898.
[22] Q. Yang and S. Jagannathan, "Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 377–390, Apr. 2012.
[23] L. Han, P. Sun, Y. Du, J. Xiong, Q. Wang, X. Sun, H. Liu, and T. Zhang, "Grid-wise control for multi-agent reinforcement learning in video game AI," in *Proc. 36th Int. Conf. Mach. Learn.*, Long Beach, CA, USA, 2019, pp. 1–6.
[24] J. Zhang, Z. Wang, and H. Zhang, "Data-based optimal control of multi-agent systems: A reinforcement learning design approach," *IEEE Trans. Cybern.*, vol. 49, no. 12, pp. 4441–4449, Dec. 2019.
[25] A.-M. Zou, K. D. Kumar, and Z.-G. Hou, "Distributed consensus control for multi-agent systems using terminal sliding mode and Chebyshev neural networks," *Int. J. Robust Nonlinear Control*, vol. 23, no. 3, pp. 334–357, Feb. 2013.
[26] L. Li and Y. Xia, "Stochastic stability of the unscented Kalman filter with intermittent observations," *Automatica*, vol. 48, no. 5, pp. 978–981, May 2012.
[27] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.
[28] K. J. Borah and K. D. Kumar, "Reinforced unscented Kalman filter for consensus achievement of uncertain multi-agent systems subject to actuator faults," *Int. J. Robust. Nonlinear Control*, pp. 1–26, 2023, doi: 10.1002/rnc.6913.

**KAUSTAV JYOTI BORAH** received the B.Tech. degree in electronics and communication engineering in India, in 2013, the M.S. degree in electrical engineering in USA, in 2016, and the Ph.D. degree in aerospace engineering from Toronto Metropolitan University, Canada, in 2022. His research interests include the interdisciplinary areas of autonomous and multi-agent systems, fault diagnosis and prognosis, machine learning, aerospace systems, guidance, navigation, and control. Passionate about cross-disciplinary research, he integrates estimation and control theory, deep reinforcement learning, data-driven techniques, and mathematical modeling into his work. He has published two book chapters, four journals, and five conference papers. He received several internal and external scholarships throughout his academic career, including Aerospace Engineering Research Excellence Award and Queen Elizabeth II Graduate Scholarship in Science and Technology. His Ph.D. research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) discovery grant. He serves as a Reviewer for IEEE Transactions on Automatic Control, IEEE Transactions on Cybernetics, IEEE Transactions on Aerospace and Electronic Systems, IEEE Access, *Automatica*, *Neurocomputing*, and *Nonlinear Dynamics*.

**KRISHNA DEV KUMAR** (Senior Member, IEEE) received the Ph.D. degree in aerospace engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 1998. He is currently a Professor in aerospace engineering and the Director of the Artificial Intelligence for Aerospace Systems (AIAS) Laboratory, Toronto Metropolitan University, Toronto, ON, Canada. He has more than 200 research publications in national and international journals and conferences. His research interests include spacecraft dynamics and control, fault diagnosis and prognosis, big data, predictive analytics, and artificial intelligence. He was a recipient of several awards that include member of the International Academy of Astronautics, France, in 2018, Eminent Alumnus Award, Veer Surendra Sai University of Technology, Sambalpur, in 2017, Sarwan Sahota Ryerson Distinguished Scholar Award, in 2015, the Canada Research Chair (2005–2015), an Associate Fellow and a Life Member of AIAA, in 2012, Ontario Early Researcher Award (2006–2011), Japan Society for the Promotion of Science (JSPS) Fellowship (2001–2003), and Science and Technology Agency (STA) Fellowship (1998–2000).