

# Probability-Guaranteed Distributed Filtering for Nonlinear Systems with Innovation Constraints over Sensor Networks

Lifeng Ma, Zidong Wang, Yun Chen, and Xiaojian Yi

**Abstract**—In this paper, the distributed filtering problem is investigated for a class of nonlinear systems. Each individual sensing node provides the state estimate by using not only its own measurements but also its neighbors' information propagated according to the communication topology. With the purpose of mitigating the effects from possible abnormal data during the signal transmission, an innovation constraint with adaptively determined threshold is imposed on the transmitted innovation during the filter process. The aim of the addressed problem is to design a distributed filtering algorithm which is capable of 1) confining all the estimation errors within certain ellipsoidal regions with prescribed probability; and 2) achieving the required average disturbance attenuation specification. By virtue of convex optimization method, sufficient conditions are derived for the existence of the requested filtering algorithm and the desired filtering parameters are then obtained by iteratively solving the corresponding matrix inequalities. Within the established framework, two optimization problems are put forward to seek locally optimal filtering parameters. Finally, an illustrative numerical example is provided to demonstrate the applicability of the proposed filtering paradigm.

**Index Terms**—Distributed filtering, set-membership filtering, innovation constraints, probability-guaranteed filtering

## I. INTRODUCTION

Sensor networks, built of a group of individual sensing nodes, have recently stirred intensive research attention from both industry and academy in the disciplines including communication, system science and signal processing. These sensing nodes are spatially dispersed with each individual node having the basic abilities of information processing such as collection, computing and transmission. Thanks to their distinctive merits (e.g., low cost, easy deployability, convenient maintenance), sensor networks have found wide applications in many areas such as environment surveillance,

military detection, target location, and so forth [7], [17], [19], [30], [31].

The filtering/estimation problem is well recognized as a fundamental research topic that plays a vitally important role in many engineering applications. Along with the recent development of sensor networks, the distributed filtering/estimation issues have attracted a particular research interest in extracting the true signal from the possibly noisy measurement data collected by sensing nodes. According to their structures, the sensor-network-based filter algorithms can be generally divided into two categories (namely, centralized approach and distributed approach), each with their own merits and demerits. The centralized approach adopts a central unit to calculate the state estimate by using the measurement information from *all* sensing nodes [13], [25]. Apparently, such a technique would impose high demands upon the capabilities of the central unit (e. g. storage capacity and processing speed), especially when the scale of the network becomes large. In contrast, the distributed algorithm deploys local filter at each individual node and provides the state estimate by using not only the local measurements but also the information propagated by the neighbors. In comparison to the centralized approach, the distributed algorithm possesses certain advantages such as flexibility of deployment and robustness against disturbances, etc. Thus, in recent decade, the distributed filtering issues have been garnering considerable interest within systems science and signal processing communities [10], [12], [14], [15].

So far, the distributed filtering problems have been extensively investigated from different perspectives. According to properties of disturbances, several distributed filtering techniques have been proposed in the literature with respective scopes of applicability. For instance, the Kalman filtering technique has been applied in [5], [20], [21] to cope with the distributed state estimation issues for systems with Gaussian noises. As the standard Kalman-like framework is no longer applicable in the case of non-Gaussian disturbances, many alternative methods have been developed with examples including, but are not limited to, the  $H_\infty$  technique presented to deal with energy-bounded noises (see e.g. [24]) and the set-membership filtering theory exploited to handle the unknown-but-bounded (UBB) noises (see e.g. [11]). Recently, a consensus nonlinear information filter has been designed in [9] where the sensors' measurements are subject to a sort of non-Gaussian disturbances. It should be emphasized that, due to the wide appearance of UBB noises in practical applications especially in electrical and electronics engineering, the set-

This work was supported in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20190021, the National Natural Science Foundation of China under Grants 61773209, 61973102, 61873148, 61933007 and 71801196, the Zhejiang Provincial Natural Science Foundation of China under Grant LR16F030003, the Six Talent Peaks Project in Jiangsu Province of China under Grant XYDXX-033, and the Alexander von Humboldt Foundation of Germany. (*Corresponding author: Xiaojian Yi.*)

L. Ma is with the School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China. (Email: malifeng@njjust.edu.cn)

Z. Wang is with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. (Email: Zidong.Wang@brunel.ac.uk)

Y. Chen is with the School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China. (Email: yunchen@hdu.edu.cn)

X. Yi is with the School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China. (Email: yixiaojianbit@sina.cn)

membership filtering technique has recently attracted renewed interest, and some representative publications can be found in [14], [29] and the references therein.

In practical control engineering, it is often unnecessary (and also impossible) to guarantee the performance requirements with probability 1 in a strict yet persistent way. This is particularly true when the target plant undergoes severe stochastic noises. Actually, in certain situations, it is usually impractical to achieve the desired objectives with 100% confidence. For instance, in the maneuvering targets tracking problem, it is generally satisfactory if the radar can track the targets with a 90% success probability. Another quintessential example that should be mentioned is that, in the missile control, we often require the standard deviation error be confined within a 25m-radius-circle with an 80% probability. Such engineering practice gives rise to the idea of designing controllers/filters capable of meeting the performance with an acceptable probability [23], [26], [28]. Obviously, such a probability-guaranteed scheme could reduce the conservatism stemming from unnecessarily stringent design requirements, thereby providing extra freedom for other essential indices as well as reducing the design expense. As a result, it is of significance to establish the probability-guaranteed framework for system analysis and synthesis, and such an issue has not yet received adequate research attention.

In practical systems, it is quite desirable to design filters that satisfy *multiple* performance requirements. Motivated by the above discussions, in this paper, we endeavor to study the probability-guaranteed distributed filtering problem for general time-varying nonlinear systems subject to innovation constraint with simultaneous consideration of multiple performance indices. This appears to be a challenging task because of some fundamental difficulties identified as follows. 1) For general nonlinear time-varying systems, it is invariably arduous to propose appropriate performance indices that quantify transient dynamical characteristics and yet facilitate the subsequent investigation by utilizing existing methodologies. 2) The nonlinear and time-varying nature of the addressed system gives rise to significant difficulties in analysis and synthesis with respect to the proposed multiple requirements. 3) The innovation constraint imposed on the transmitted information among sensing nodes constitutes another kind of nonlinearities that complicates the filter design even further. It is, therefore, the main purpose of this paper to deal with the identified challenges by launching a major study on the addressed probability-guaranteed multi-objective distributed filtering problem.

The novelties of this paper can be summarized as fourfold. *i) The model of the target plant under consideration is comprehensive that caters for nonlinearities, stochasticity and time-varying effects. ii) In order to better characterize the performances in the finite horizon, two transient performance indices, namely, average  $H_\infty$  criterion and probabilistic ellipsoidal constraint are proposed from different perspectives. iii) In order to mitigate the effect of abnormal measurements (e.g., outliers, attacks, etc.), a saturation function is imposed on the innovations and the saturation level is adaptively adjusted at each time step according to previous estimation errors.*

*iv) The proposed algorithm is capable of guaranteeing the ellipsoidal constraint with a predetermined probability (rather than the usual 100% confidence). Such a probabilistically design method could provide much extra flexibility by relaxing certain stringent yet unnecessary performance constraints in real-world applications.*

**Notation** The notation used here is fairly standard except where otherwise stated.  $\mathbb{R}^n$  denotes the  $n$ -dimensional Euclidean space.  $\mathbf{1}_n$  denotes an  $n$ -dimensional column vector with all ones.  $I_n$  and  $0_n$  denote the identity matrix and zero matrix of  $n$  dimensions, respectively. The notation  $X \geq Y$  (respectively  $X > Y$ ), where  $X$  and  $Y$  are symmetric matrices, means that  $X - Y$  is positive semi-definite (respectively positive definite). For matrices  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$ , their Kronecker product is a matrix in  $\mathbb{R}^{mp \times nq}$  denoted as  $A \otimes B$ . The superscript ‘‘T’’ denotes the transpose. The symbol ‘\*’ stands for the corresponding entry of the matrix can be obtained by symmetrical property. For a vector  $a$ ,  $\|a\| = a^T a$ .  $\text{tr}[A]$  means the trace of matrix  $A$  and  $\text{diag}\{F_1, F_2, \dots, F_n\}$  denotes a block diagonal matrix whose diagonal blocks are given by  $F_1, F_2, \dots, F_n$ . The notation  $\text{diag}_n\{A_i\}$  represents the block diagonal matrix  $\text{diag}\{A_1, A_2, \dots, A_n\}$  and  $\text{col}_n\{x_i\}$  denotes the column vector  $[x_1^T \ x_2^T \ \dots \ x_n^T]^T$ .  $\mathbb{P}\{A\}$  means the occurrence probability of the event ‘A’.

## II. PROBLEM FORMULATION

In this paper, it is assumed that the sensor network has  $N$  sensor nodes which are distributed in the space according to a specific interconnection topology characterized by a directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L})$ , where  $\mathcal{V} = \{1, 2, \dots, N\}$  denotes the set of sensing nodes,  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  is the set of edges, and  $\mathcal{L} = [\theta_{ij}]_{N \times N}$  is the nonnegative adjacency matrix associated with the edges of the graph, that is,  $\theta_{ij} > 0$  if and only if  $\text{edge}(i, j) \in \mathcal{E}$  (i.e. there is information transmission from sensor  $j$  to sensor  $i$ ). If  $(i, j) \in \mathcal{E}$ , then node  $j$  is called one of the neighbors of node  $i$ . Also, we assume that  $\theta_{ii} = 1$  for all  $i \in \mathcal{V}$  and, therefore,  $(i, i)$  can be regarded as an additional edge. The set of neighbors of node  $i \in \mathcal{V}$  plus the node itself is denoted by  $\mathcal{N}_i \triangleq \{j \in \mathcal{V} | (i, j) \in \mathcal{E}\}$ .

Consider the following nonlinear system defined on the horizon  $[0, T]$ :

$$\begin{cases} x_{k+1} = f(x_k) + B_k \nu_k + (h(x_k) + D_k \nu_k) \omega_k \\ y_{i,k} = C_{i,k} x_k + E_{i,k} \mu_k \end{cases} \quad (1)$$

where  $x_k \in \mathbb{R}^{n_x}$  and  $y_{i,k} \in \mathbb{R}^{n_y}$  represent, respectively, the system state and measurement output on the  $i$ -th sensing node;  $\omega_k$  is a zero-mean Gaussian white sequence with unitary variance;  $\nu_k \in \mathbb{R}^{n_\nu}$  and  $\mu_k \in \mathbb{R}^{n_\mu}$  are the process and measurement noises;  $B_k, D_k, C_{i,k}$  and  $E_{i,k}$  are known real-valued matrices of compatible dimensions;  $f(x_k) : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x}$  and  $h(x_k) : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x}$  are two nonlinear functions.

*Assumption 1:* The noise sequences  $\nu_k$  and  $\mu_k$  satisfy the constraints

$$\begin{cases} \nu_k \in \mathcal{V}_k \triangleq \{\nu_k : \nu_k^T V_k^{-1} \nu_k \leq 1\} \\ \mu_k \in \mathcal{U}_k \triangleq \{\mu_k : \mu_k^T U_k^{-1} \mu_k \leq 1\} \end{cases} \quad (2)$$

where  $V_k > 0$  and  $U_k > 0$  are positive matrices with suitable dimensions.

For each sensing node  $i$  ( $i = 1, 2, \dots, N$ ), the local filter to be designed is of the following form:

$$\hat{x}_{i,k+1} = F_{i,k}\hat{x}_{i,k} + \sum_{j \in \mathcal{N}_i} \theta_{ij} H_{ij,k} \text{Sat}_{\sigma_{j,k}}(y_{j,k} - \hat{y}_{j,k}) \quad (3)$$

where  $\hat{y}_{i,k} \triangleq C_{i,k}\hat{x}_{i,k}$  is the estimated output on node  $i$ ;  $F_{i,k}$  and  $H_{ij,k}$  are filtering parameters to be designed. Here, for a vector  $a \in \mathbb{R}^{n_a}$  with  $a^{(s)}$  denoting its  $s$ -th entry, the nonlinear function  $\text{Sat}_{\sigma_{j,k}}(a)$  is defined as

$$\text{Sat}_{\sigma_{j,k}}(a) \triangleq \begin{bmatrix} \text{Sat}_{\sigma_{j,k}}(a^{(1)}) \\ \text{Sat}_{\sigma_{j,k}}(a^{(2)}) \\ \vdots \\ \text{Sat}_{\sigma_{j,k}}(a^{(n_a)}) \end{bmatrix}$$

where  $\text{Sat}_{\sigma_{j,k}}(a^{(s)}) \triangleq \text{sign}(a^{(s)}) \cdot \min\{|a^{(s)}|, \sigma_{j,k}\}$ . Furthermore, in this paper, we employ the following function to dynamically govern the saturation level  $\sigma_{i,k}$  for node  $i$ :

$$\sigma_{i,k+1} = \lambda \sigma_{i,k} + (y_{i,k} - \hat{y}_{i,k})^T W_i (y_{i,k} - \hat{y}_{i,k}) \quad (4)$$

where  $\lambda \in [0, 1)$  and  $W_i > 0$  is a given weighting matrix.

*Remark 1:* One of the distinct features of the distributed filtering algorithm is that, at each individual sensing node, not only the local measurements but also the neighboring information will be used to generate the state estimate. This requires the data exchange among sensing nodes, which is realized via network-based communications. The traditional form of distributed filters can be found in [24] and the references therein. In our improved form of distributed filter (3), a saturation function is used to characterize the restraint on the innovation vector by imposing an upper-bound on the absolute value of the transmitted data. Such a mechanism is widely used in many engineering practice. For instance, in electrical and electronics engineering, an amplitude limiter is usually implemented to constrain the signals (e.g., current, voltage, etc) within certain allowable range to protect devices.

*Remark 2:* It should be emphasized that the reason why we propose such a saturation mechanism lies as twofold. *First*, the transmitted data are innovations of neighboring nodes, which we believe should be within certain range if the nodes are in good condition. Accordingly, the saturation function is conducive to alleviate the impacts from possible outliers which usually occur especially to those nodes deployed in harsh environments. *Second*, as is well known, the utilization of networks will face the threat from malicious attacks, and another advantage of the proposed filter structure is the ability of mitigating the effects from possible attacks such as false data injection which aims to deteriorate the performance by injecting false signals to the original ones during the transmission.

*Remark 3:* Different from saturation functions with fixed thresholds in most existing literature (see, e.g. [29]), our filter structure adopts time-varying saturation thresholds (characterized by  $\sigma_{i,k}$ ) that are determined iteratively and adaptively according to the value of innovations, see (4). In comparison to the cases of fixed threshold, our proposed adaptive mechanism enables the saturation threshold  $\sigma_{i,k}$  to be adjusted appropriately along with the estimation performance. Specifically,

when the innovations become larger (which is indicative of larger differences between measurement output and estimated output), we know from (4) that  $\sigma_{i,k}$  will be larger accordingly so as to relax the limitations on the transmitted signals, which is in accordance with the engineering practice.

*Definition 1:* A bounded ellipsoid  $\mathfrak{E}(c, X)$  of  $\mathbb{R}^n$  with a nonempty interior is defined by

$$\mathfrak{E}(c, X) \triangleq \{x \in \mathbb{R}^n : (x - c)^T X^{-1} (x - c) \leq 1\} \quad (5)$$

where  $c \in \mathbb{R}^n$  is the center of  $\mathfrak{E}(c, P)$  and  $P > 0$  is a positive definite matrix that specifies the ellipsoid's shape and orientation.

By defining the estimation error  $\tilde{x}_{i,k} \triangleq x_k - \hat{x}_{i,k}$ , we are now ready to present the design objective. In this paper, it is our aim to determine the filtering parameters  $F_{i,k}$  and  $H_{ij,k}$  in (3) such that the following two requirements *R1*) and *R2*) are met simultaneously.

*R1) Probabilistic ellipsoidal constraint*

$$\mathbb{P}\{x_k \in \mathfrak{E}(\hat{x}_{i,k}, \mathfrak{P}_k)\} \geq \mathbf{p} \quad (6)$$

or, equivalently,

$$\mathbb{P}\{\tilde{x}_{i,k}^T \mathfrak{P}_k^{-1} \tilde{x}_{i,k} \leq 1\} \geq \mathbf{p} \quad (7)$$

where  $\mathfrak{P}_k > 0$  is a prescribed matrix and the prespecified positive scalar  $\mathbf{p}$  satisfies  $0 < \mathbf{p} \leq 1$ .

*R2) Average  $H_\infty$  specification*

$$\begin{aligned} & \frac{1}{N} \mathbb{E} \left\{ \sum_{i=1}^N \sum_{k=0}^T \|\tilde{x}_{i,k}\|^2 \right\} \\ & \leq \gamma^2 \sum_{k=0}^T (\|\nu_k\|^2 + \|\mu_k\|^2) + \gamma^2 \frac{1}{N} \sum_{i=1}^N \tilde{x}_{i,0}^T \Pi_i \tilde{x}_{i,0} \end{aligned} \quad (8)$$

where  $\Pi_i > 0$  are known weighting matrices.

*Remark 4:* In practical systems, it is quite desirable to design filters that satisfy multiple performance requirements. For example, in the scenario of maneuvering target tracking via sensor networks, it is always required the tracking systems show simultaneously good tracking accuracy and disturbance attenuation ability. On the other hand, for time-varying systems, it makes more sense to investigate *transient* performances over a time period of interest than the steady-state characteristics over the infinite horizon. Note that, although the design objectives *R1*) and *R2*) are both proposed to depict transient performances, they are actually put forward from different perspectives. In the first place, inequality (7) proposes the probabilistic ellipsoidal constraint with respect to each *individual* node at each *single* time step. In the second place, inequality (8) measures the disturbance attenuation level in a *collective* way over the *whole* time interval  $[0, T]$ .

### III. MAIN RESULTS

Before presenting our main results, we first introduce the following definition and lemmas that are helpful in subsequent derivations.

*Definition 2:* [1] Let  $\mathfrak{K}_1$  and  $\mathfrak{K}_2$  be some real matrices with  $\mathfrak{K} \triangleq \mathfrak{K}_2 - \mathfrak{K}_1 > 0$ . A nonlinearity  $\phi(\cdot)$  is said to satisfy the sector condition with respect to  $\mathfrak{K}_1$  and  $\mathfrak{K}_2$  if

$$(\phi(y) - \mathfrak{K}_1 y)^\top (\phi(y) - \mathfrak{K}_2 y) \leq 0. \quad (9)$$

In this case, the sector-bounded nonlinearity  $\phi(\cdot)$  is said to belong to the sector  $[\mathfrak{K}_1, \mathfrak{K}_2]$ .

*Lemma 1:* (S-procedure [2]) Let  $\psi_0(\cdot), \psi_1(\cdot), \dots, \psi_p(\cdot)$  be quadratic functions of the variable  $\varsigma \in \mathbb{R}^n$ :  $\psi_j(\varsigma) \triangleq \varsigma^\top X_j \varsigma$  ( $j = 0, \dots, p$ ), where  $X_j^\top = X_j$ . If there exist  $\epsilon_1 \geq 0, \dots, \epsilon_p \geq 0$  such that  $X_0 - \sum_{j=1}^p \epsilon_j X_j \leq 0$ , then the following is true:

$$\psi_1(\varsigma) \leq 0, \dots, \psi_p(\varsigma) \leq 0 \rightarrow \psi_0(\varsigma) \leq 0. \quad (10)$$

*Lemma 2:* (Schur Complement Equivalence [2]) Given constant matrices  $\mathfrak{S}_1, \mathfrak{S}_2, \mathfrak{S}_3$  where  $\mathfrak{S}_1 = \mathfrak{S}_1^\top$  and  $0 < \mathfrak{S}_2 = \mathfrak{S}_2^\top$ , then  $\mathfrak{S}_1 + \mathfrak{S}_3^\top \mathfrak{S}_2^{-1} \mathfrak{S}_3 < 0$  if and only if

$$\begin{bmatrix} \mathfrak{S}_1 & \mathfrak{S}_3^\top \\ \mathfrak{S}_3 & -\mathfrak{S}_2 \end{bmatrix} < 0 \quad \text{or} \quad \begin{bmatrix} -\mathfrak{S}_2 & \mathfrak{S}_3 \\ \mathfrak{S}_3^\top & \mathfrak{S}_1 \end{bmatrix} < 0. \quad (11)$$

*Lemma 3:* [2] Let  $\mathfrak{M} = \mathfrak{M}^\top$ ,  $\mathfrak{H}$  and  $\mathfrak{E}$  be real matrices of appropriate dimensions, and  $\Delta$  satisfies  $\|\Delta\| \leq 1$ , then

$$\mathfrak{M} + \mathfrak{H} \Delta \mathfrak{E} + \mathfrak{E}^\top \Delta \mathfrak{H}^\top \leq 0 \quad (12)$$

if and only if there exists a positive scalar  $\varepsilon$  such that

$$\mathfrak{M} + \varepsilon \mathfrak{H} \mathfrak{H}^\top + \varepsilon^{-1} \mathfrak{E}^\top \mathfrak{E} \leq 0. \quad (13)$$

By resorting to the Taylor expansion technique, we can express the nonlinear functions  $f(x_k)$  and  $g(x_k)$  as follows:

$$f(x_k) = f(\hat{x}_{i,k}) + \Phi_{i,k} \tilde{x}_{i,k} + L_i \Delta_{1i} \tilde{x}_{i,k} \quad (14)$$

$$h(x_k) = h(\hat{x}_{i,k}) + \Psi_{i,k} \tilde{x}_{i,k} + \Sigma_i \Delta_{2i} \tilde{x}_{i,k} \quad (15)$$

where  $L_i \in \mathbb{R}^{n_l}$  and  $\Sigma_i \in \mathbb{R}^{n_\Sigma}$  are known matrices;  $\Delta_{1i} \in \mathbb{R}^{n_l \times n_x}$  and  $\Delta_{2i} \in \mathbb{R}^{n_\Sigma \times n_x}$  are unknown matrices such that  $\|\Delta_{1i}\| \leq 1$  and  $\|\Delta_{2i}\| \leq 1$ ;  $\Phi_{i,k}$  and  $\Psi_{i,k}$  are calculated as follows:

$$\Phi_{i,k} \triangleq \left. \frac{\partial f(x)}{\partial x} \right|_{x=\hat{x}_{i,k}}, \quad \Psi_{i,k} \triangleq \left. \frac{\partial h(x)}{\partial x} \right|_{x=\hat{x}_{i,k}}. \quad (16)$$

Letting  $r_{j,k} \triangleq y_{j,k} - \hat{y}_{j,k}$ , there exist matrices  $0 \leq G_{1j} \leq I \leq G_{2j}$  such that

$$\text{Sat}_{\sigma_{j,k}}(r_{j,k}) = G_{1j} r_{j,k} + \phi_j(r_{j,k}) \quad (17)$$

where  $\phi_j(r_{j,k})$  is a nonlinear vector-valued function which satisfies a sector condition with  $\mathfrak{K}_{1j} = 0$  and  $\mathfrak{K}_{2j} = G_j$  ( $G_j = G_{2j} - G_{1j}$ ), i.e.,  $\phi_j(r_{j,k})$  satisfies the following inequality:

$$\phi_j^\top(r_{j,k}) (\phi_j(r_{j,k}) - G_j r_{j,k}) \leq 0. \quad (18)$$

From system (1) and filter (3), we acquire the following dynamics of estimation error:

$$\begin{aligned} x_{k+1} - \hat{x}_{i,k+1} &= f(\hat{x}_{i,k}) + \Phi_{i,k} \tilde{x}_{i,k} + L_i \Delta_{1i} \tilde{x}_{i,k} + B_k \nu_k - F_{i,k} \hat{x}_{i,k} \\ &\quad + (h(\hat{x}_{i,k}) + \Psi_{i,k} \tilde{x}_{i,k} + \Sigma_i \Delta_{2i} \tilde{x}_{i,k} + D_k \nu_k) \omega_k \\ &\quad - \sum_{j \in \mathcal{N}_i} \theta_{ij} H_{ij,k} G_{1j} C_{j,k} x_k + \sum_{j \in \mathcal{N}_i} \theta_{ij} H_{ij,k} G_{1j} C_{j,k} \hat{x}_{j,k} \end{aligned}$$

$$- \sum_{j \in \mathcal{N}_i} \theta_{ij} H_{ij,k} G_{1j} E_{j,k} \mu_k - \sum_{j \in \mathcal{N}_i} \theta_{ij} H_{ij,k} \phi_j(r_{j,k}) \quad (19)$$

For development brevity, denote

$$\begin{aligned} \xi_k &\triangleq \text{col}_N \{x_k\}, \quad \hat{x}_k \triangleq \text{col}_N \{\hat{x}_{i,k}\}, \quad \tilde{x}_k \triangleq \text{col}_N \{\tilde{x}_{i,k}\}, \\ \hat{f}_k &\triangleq \text{col}_N \{f(\hat{x}_{i,k})\}, \quad \phi_k \triangleq \text{col}_N \{\phi_i(r_{i,k})\}, \\ \hat{h}_k &\triangleq \text{col}_N \{h(\hat{x}_{i,k})\}, \quad \hat{y}_k \triangleq \text{col}_N \{\hat{y}_{i,k}\}, \\ \mathcal{B}_k &\triangleq \mathbf{1}_N \otimes B_k, \quad \mathcal{D}_k \triangleq \mathbf{1}_N \otimes D_k, \\ \mathcal{C}_k &\triangleq \text{diag}_N \{C_{i,k}\}, \quad \mathcal{F}_k \triangleq \text{diag}_N \{F_{i,k}\}, \quad \mathcal{E}_k \triangleq \text{diag}_N \{E_{i,k}\}, \\ \mathcal{G} &\triangleq \text{diag}_N \{G_{1i}\}, \quad \mathcal{L} \triangleq \text{diag}_N \{L_i\}, \quad \Phi_k \triangleq \text{diag}_N \{\Phi_{i,k}\}, \\ \Psi_k &\triangleq \text{diag}_N \{\Psi_{i,k}\}, \quad \Sigma \triangleq \text{diag}_N \{\Sigma_i\}, \\ \Delta_1 &\triangleq \text{diag}_N \{\Delta_{1i}\}, \quad \Delta_2 \triangleq \text{diag}_N \{\Delta_{2i}\}, \\ \Theta_{\iota,i} &\triangleq \text{diag} \{ \underbrace{0, \dots, 0}_{i-1}, \underbrace{I_\iota, 0, \dots, 0}_{N-i} \}, \\ \mathcal{R}_{\iota,i} &\triangleq (\mathbf{1}_N^\top \otimes I_\iota) \Theta_{\iota,i}, \quad \iota = \{n_x, n_q, n_y, n_\Sigma\}. \end{aligned}$$

Then, we have the following compact form of the dynamics of the estimation error:

$$\begin{aligned} \tilde{x}_{k+1} &= \hat{f}_k + \Phi_k \tilde{x}_k + \mathcal{L} \Delta_1 \tilde{x}_k + \mathcal{B}_k \nu_k \\ &\quad + (\hat{h}_k + \Psi_k \tilde{x}_k + \Sigma \Delta_2 \tilde{x}_k + \mathcal{D}_k \nu_k) \omega_k \\ &\quad - \mathcal{F}_k \hat{x}_k - \mathcal{H}_k \mathcal{G} \mathcal{C}_k \xi_k - \mathcal{H}_k \mathcal{G} \mathcal{E}_k (\mathbf{1}_N \otimes I_{n_\mu}) \mu_k \\ &\quad + \mathcal{H}_k \mathcal{G} \mathcal{C}_k \hat{x}_k - \mathcal{H}_k \phi_k \end{aligned} \quad (20)$$

where  $\mathcal{H}_k \triangleq [\theta_{ij} H_{ij,k}]_{N \times N}$ . Obviously, since  $\theta_{ij} = 0$  when  $j \notin \mathcal{N}_i$ ,  $\mathcal{H}_k$  is a sparse matrix which is characterized by

$$\mathcal{H}_k \in \mathcal{T}_{n_x \times n_y} \quad (21)$$

where  $\mathcal{T}_{n_x \times n_y} \triangleq \{\mathcal{T} = [T_{ij}] \in \mathbb{R}^{N n_x \times N n_y} \mid T_{ij} \in \mathbb{R}^{n_x \times n_y}, T_{ij} = 0 \text{ if } j \notin \mathcal{N}_i\}$ .

#### A. Ellipsoidal constraint in probability $\mathbf{p}$

Defining

$$P_k \triangleq \frac{1}{1 - \mathbf{p}} \mathfrak{P}_k,$$

we present the following lemma that will be used in later derivation.

*Lemma 4:* If  $\mathbb{E}\{(x_k - \hat{x}_{i,k})^\top P_k^{-1} (x_k - \hat{x}_{i,k})\} \leq 1$ , then the following holds:

$$\mathbb{P}\{x_k \in \mathfrak{E}(\hat{x}_{i,k}, \mathfrak{P}_k)\} \geq \mathbf{p}. \quad (22)$$

*Proof:* Lemma 4 is easily accessible from Lemma 3 in [23], and its proof is therefore omitted here.  $\blacksquare$

*Assumption 2:* Let

$$P_0 \triangleq \frac{1}{1 - \mathbf{p}} \mathfrak{P}_0$$

be given with  $\mathfrak{P}_0$  being a known positive definite matrix. The system initial value and its estimates satisfy the following condition:

$$(x_0 - \hat{x}_{i,0})^\top P_0^{-1} (x_0 - \hat{x}_{i,0}) \leq 1. \quad (23)$$

*Lemma 5:* Let the filtering parameters  $F_{i,k}$  and  $H_{ij,k}$  be given. For a given sequence of positive definite matrices  $\{P_k\}_{k \geq 0}$  (with a factorization  $P_k = Q_k Q_k^\top$ ) and under

the initial condition (23), if there exist sequences of non-negative scalars  $\{\varrho_{i,k}^{(1)}\}_{k \geq 0}$ ,  $\{\varrho_k^{(2)}\}_{k \geq 0}$ ,  $\{\varrho_k^{(3)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(4)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(5)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(6)}\}_{k \geq 0}$  and sequences of scalars  $\{\beta_{i,k}\}_{k \geq 0}$  satisfying the following  $N$  recursive matrix inequalities:

$$\begin{bmatrix} -\Gamma_k & \bar{\Omega}_k^T \mathcal{R}_{n_x, i}^T & \tilde{\Omega}_k^T \mathcal{R}_{n_x, i}^T \\ * & -P_{k+1} & 0 \\ * & * & -P_{k+1} \end{bmatrix} \leq 0 \quad (24)$$

where

$$\begin{aligned} \Gamma_k \triangleq & \text{diag} \left\{ 1 - \sum_{i=1}^N \varrho_{i,k}^{(1)} - \varrho_k^{(2)} - \varrho_k^{(3)}, \right. \\ & \sum_{i=1}^N \varrho_{i,k}^{(1)} \Theta_{n_q, i} - \sum_{i=1}^N (\varrho_{i,k}^{(4)} + \varrho_{i,k}^{(5)}) \Theta_{n_q, i} \mathcal{Q}_k^T \mathcal{Q}_k, \\ & \sum_{i=1}^N \varrho_{i,k}^{(4)} \Theta_{n_l, i}, \sum_{i=1}^N \varrho_{i,k}^{(5)} \Theta_{n_\Sigma, i}, \\ & \left. \varrho_k^{(2)} V_k^{-1}, \varrho_k^{(3)} U_k^{-1}, 0 \right\} \\ & + \sum_{i=1}^N \varrho_{i,k}^{(6)} \Xi_{i,k} + \sum_{i=1}^N \beta_{i,k} \Upsilon_{i,k}, \end{aligned} \quad (25)$$

$$\Xi_{i,k} \triangleq \frac{1}{2} \begin{bmatrix} 0 & \Xi_{i,k}^{(12)} \\ * & 2I_{Nn_y} \Theta_{n_y, i} \end{bmatrix}, \quad (26)$$

$$\Xi_{i,k}^{(12)} \triangleq \begin{bmatrix} 0 \\ -\text{diag}_N \{ \mathcal{Q}_k^T C_{i,k}^T G_i^T \} \Theta_{n_y, i} \\ 0 \\ 0 \\ 0 \\ -E_{i,k}^T G_i^T \mathcal{R}_{n_y, i} \end{bmatrix},$$

$$\mathcal{W}_{i,k} \triangleq [ 0 \ C_{i,k} \mathcal{Q}_k \mathcal{R}_{n_q, i} \ 0 \ 0 \ 0 \ E_{i,k} \ 0 ], \quad (27)$$

$$\begin{aligned} \Upsilon_{i,k} \triangleq & \mathcal{W}_{i,k}^T \mathcal{W}_i \mathcal{W}_{i,k} \\ & - \text{diag} \{ \sigma_{i,k+1} - \lambda \sigma_{i,k}, 0, 0, 0, 0, 0 \}, \end{aligned} \quad (28)$$

$$\bar{\Omega}_k \triangleq \begin{bmatrix} \bar{\Omega}_{11} & \bar{\Omega}_{12} & \mathcal{L} & 0 & \mathcal{B}_k \\ & & -\mathcal{H}_k \mathcal{G} \mathcal{E}_k (\mathbf{1}_N \otimes I_{n_\mu}) & -\mathcal{H}_k & \end{bmatrix}, \quad (29)$$

$$\begin{aligned} \bar{\Omega}_{11} & \triangleq \hat{f}_k - \mathcal{F}_k \hat{x}_k, \\ \bar{\Omega}_{12} & \triangleq (\Phi_k - \mathcal{H}_k \mathcal{G} \mathcal{C}_k) \mathcal{Q}_k, \\ \tilde{\Omega}_k & \triangleq [ \hat{h}_k \ \Psi_k \mathcal{Q}_k \ 0 \ \Sigma \ \mathcal{D}_k \ 0 \ 0 ], \end{aligned} \quad (30)$$

then the following inequality holds

$$\mathbb{E} \{ (x_{k+1} - \hat{x}_{i,k+1})^T P_{k+1}^{-1} (x_{k+1} - \hat{x}_{i,k+1}) \} \leq 1. \quad (31)$$

*Proof:* See Appendix VI-A. ■

**Theorem 1:** Let the filtering parameters  $F_{i,k}$  and  $H_{ij,k}$  be given. For a prespecified positive scalar  $\mathbf{p}$  and a sequence of positive definite matrices  $\{\mathfrak{P}_k\}_{k \geq 0}$ , the design objective (6) is satisfied if there exist sequences of non-negative scalars  $\{\varrho_{i,k}^{(1)}\}_{k \geq 0}$ ,  $\{\varrho_k^{(2)}\}_{k \geq 0}$ ,  $\{\varrho_k^{(3)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(4)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(5)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(6)}\}_{k \geq 0}$  and sequences of scalars  $\{\beta_{i,k}\}_{k \geq 0}$  satisfying the following set of  $N$  matrix inequalities:

$$\begin{bmatrix} -\Gamma_k & \bar{\Omega}_k^T \mathcal{R}_{n_x, i}^T & \tilde{\Omega}_k^T \mathcal{R}_{n_x, i}^T \\ * & -\frac{1}{1-\mathbf{p}} \mathfrak{P}_{k+1} & 0 \\ * & * & -\frac{1}{1-\mathbf{p}} \mathfrak{P}_{k+1} \end{bmatrix} \leq 0. \quad (32)$$

*Proof:* Theorem 1 can be proved easily from Lemmas 3 and 5 by taking  $P_{k+1} = \frac{1}{1-\mathbf{p}} \mathfrak{P}_{k+1}$  into account. ■

### B. Average $H_\infty$ requirement

For derivation simplicity, we denote the following vectors

$$\vartheta_k \triangleq \begin{bmatrix} \nu_k \\ \mu_k \end{bmatrix}, \quad \varpi_k \triangleq \begin{bmatrix} 1 \\ \hat{x}_k \end{bmatrix}, \quad \zeta_k \triangleq \begin{bmatrix} \varpi_k \\ \vartheta_k \\ \phi_k \end{bmatrix},$$

and the following matrices

$$\tilde{\Pi} \triangleq \gamma^2 \frac{1}{N} \text{diag} \{ 0, \text{diag}_N \{ \Pi_i \} \}, \quad \mathcal{N}_k \triangleq \begin{bmatrix} 0 \\ -\mathcal{H}_k \end{bmatrix},$$

$$\hat{\mathcal{A}}_k \triangleq \begin{bmatrix} 1 & 0 \\ \hat{f}_k - \mathcal{F}_k \hat{x}_k & \Phi_k - \mathcal{H}_k \mathcal{G} \mathcal{C}_k \end{bmatrix},$$

$$\bar{\mathcal{M}}_k \triangleq \begin{bmatrix} 0 & 0 \\ \mathcal{B}_k & -\mathcal{H}_k \mathcal{G} \mathcal{E}_k (\mathbf{1}_N \otimes I_{n_\mu}) \end{bmatrix},$$

$$\check{\mathcal{A}}_k \triangleq \begin{bmatrix} 0 & 0 \\ -\hat{h}_k & \Psi_k \end{bmatrix}, \quad \tilde{\mathcal{M}}_k \triangleq \begin{bmatrix} 0 & 0 \\ \mathcal{D}_k & 0 \end{bmatrix},$$

$$\hat{I} \triangleq [ 0 \ I ], \quad \tilde{\mathcal{I}} \triangleq [ \hat{I} \ 0 \ 0 \ 0 \ 0 ],$$

$$\hat{\mathcal{L}} \triangleq \begin{bmatrix} 0 \\ \mathcal{L} \end{bmatrix}, \quad \hat{\Sigma} \triangleq \begin{bmatrix} 0 \\ \Sigma \end{bmatrix}, \quad \tilde{\mathcal{L}} \triangleq \begin{bmatrix} 0 \\ \hat{\mathcal{L}} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \tilde{\Sigma} \triangleq \begin{bmatrix} 0 \\ \hat{\Sigma} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\hat{\mathcal{A}}_k \triangleq [ \hat{\mathcal{A}}_k \ \bar{\mathcal{M}}_k \ \mathcal{N}_k ], \quad \check{\mathcal{M}}_k \triangleq [ \check{\mathcal{A}}_k \ \tilde{\mathcal{M}}_k \ 0 ],$$

$$\bar{\mathcal{C}}_k \triangleq [ 0 \ \mathcal{G} \mathcal{C}_k ], \quad \bar{\mathcal{E}}_k \triangleq [ 0 \ \mathcal{G} \mathcal{E}_k (\mathbf{1}_N \otimes I_{n_\mu}) ],$$

$$\bar{\mathcal{C}}_{i,k} \triangleq \text{diag} \{ \underbrace{0, \dots, 0}_{i-1}, C_{i,k}, \underbrace{0, \dots, 0}_{N-i} \} \times [ 0 \ I_{Nn_x} ],$$

$$\bar{E}_{i,k} \triangleq [ 0_{n_\nu} \ E_{i,k} ], \quad \tilde{\mathcal{C}}_{i,k} \triangleq [ \bar{\mathcal{C}}_{i,k} \ \bar{E}_{i,k} \ 0 ].$$

**Lemma 6:** Let the filtering parameters  $F_{i,k}$  and  $H_{ij,k}$  be given. Under initial condition  $Y_0 \leq \tilde{\Pi}$ , the average  $H_\infty$  design objective (8) is achieved if there exist a sequence of positive definite matrices  $\{Y_k\}_{k \geq 1}$ , sequences of non-negative scalars  $\{\varepsilon_k\}_{k \geq 0}$ ,  $\{\rho_{1,k}\}_{k \geq 0}$ ,  $\{\rho_{2,k}\}_{k \geq 0}$  and sequences of scalars  $\{\epsilon_{i,k}\}_{k \geq 0}$  satisfying the following set of  $N$  recursive matrix inequalities:

$$\begin{bmatrix} \tilde{\Lambda}_k & \hat{\mathcal{A}}_k^T & \check{\mathcal{M}}_k^T & \mathcal{L} \\ \hat{\mathcal{A}}_k & -Y_{k+1}^{-1} & 0 & 0 \\ \check{\mathcal{M}}_k & 0 & -Y_{k+1}^{-1} & 0 \\ \mathcal{L}^T & 0 & 0 & -J \end{bmatrix} \leq 0 \quad (33)$$

where

$$\tilde{\Lambda}_k \triangleq -\varepsilon_k \mathcal{E}_k - \sum_{i=1}^N \epsilon_{i,k} \mathcal{E}_{i,k} + \bar{\Lambda}_k, \quad (34)$$

$$\mathcal{E}_k \triangleq \begin{bmatrix} 0 & 0 & -\frac{1}{2} \bar{\mathcal{C}}_k^T \\ * & 0 & -\frac{1}{2} \bar{\mathcal{E}}_k^T \\ * & * & I \end{bmatrix}, \quad (35)$$

$$\mathcal{E}_{i,k} \triangleq \tilde{\mathcal{C}}_{i,k}^T W_i \tilde{\mathcal{C}}_{i,k} - \text{diag} \{ \sigma_{i,k+1} - \lambda \sigma_{i,k}, 0_{Nn_x}, 0_{n_\nu + n_\mu}, 0_{Nn_y} \}, \quad (36)$$

$$\bar{\Lambda}_k \triangleq \begin{bmatrix} -Y_k + \frac{1}{N} \text{diag} \{ 0, I \} & 0 & 0 \\ 0 & -\gamma^2 I & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (37)$$

$$\mathcal{L} \triangleq [ \tilde{\mathcal{L}} \quad \tilde{\Sigma} \quad \rho_{1,k} \tilde{\mathcal{L}}^T \quad \rho_{2,k} \tilde{\mathcal{L}}^T ], \quad (38)$$

$$J \triangleq \text{diag}\{\rho_{1,k} I, \rho_{2,k} I, \rho_{1,k} I, \rho_{2,k} I\}. \quad (39)$$

*Proof:* See Appendix VI-B. ■

**Theorem 2:** Let the filtering parameters  $F_{i,k}$  and  $H_{ij,k}$  be given. Under initial condition  $Y_0 \leq \bar{\Pi}$ , the average  $H_\infty$  design objective (8) is achieved if there exist a sequence of positive definite matrices  $\{\mathcal{Y}_k\}_{k \geq 1}$ , sequences of non-negative scalars  $\{\varepsilon_k\}_{k \geq 0}$ ,  $\{\rho_{1,k}\}_{k \geq 0}$  and  $\{\rho_{2,k}\}_{k \geq 0}$  and sequences of scalars  $\{\varepsilon_{i,k}\}_{k \geq 0}$  satisfying the following set of  $N$  recursive matrix inequalities:

$$\begin{bmatrix} \tilde{\Lambda}_k & \hat{\mathcal{A}}_k^T & \hat{\mathcal{M}}_k^T & \mathcal{L} \\ \hat{\mathcal{A}}_k & -\mathcal{Y}_{k+1} & 0 & 0 \\ \hat{\mathcal{M}}_k & 0 & -\mathcal{Y}_{k+1} & 0 \\ \mathcal{L}_k^T & 0 & 0 & -J \end{bmatrix} \leq 0 \quad (40)$$

where the parameter  $Y_k$  contained in  $\tilde{\Lambda}_k$  is updated recursively according to  $Y_k = \mathcal{Y}_k^{-1}$ .

*Proof:* The proof can be easily performed based on Lemma 6 and is therefore omitted here. ■

### C. Filter design

**Theorem 3:** Let the design specifications  $(\gamma^2, \mathbf{p}, \mathfrak{P}_k)$  be given. Under initial condition  $Y_0 \leq \bar{\Pi}$ , if there exist a sequence of positive definite matrices  $\{\mathcal{Y}_k\}_{k \geq 1}$ , sequences of real-valued matrices  $\{\mathcal{F}_k\}_{k \geq 0}$  and  $\{\mathcal{H}_k \in \mathcal{T}_{n_x \times n_y}\}_{k \geq 0}$ , sequences of non-negative scalars  $\{\varrho_{i,k}^{(1)}\}_{k \geq 0}$ ,  $\{\varrho_k^{(2)}\}_{k \geq 0}$ ,  $\{\varrho_k^{(3)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(4)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(5)}\}_{k \geq 0}$ ,  $\{\varrho_{i,k}^{(6)}\}_{k \geq 0}$ ,  $\{\varepsilon_k\}_{k \geq 0}$ ,  $\{\rho_{1,k}\}_{k \geq 0}$ ,  $\{\rho_{2,k}\}_{k \geq 0}$  and sequences of scalars  $\{\varepsilon_{i,k}\}_{k \geq 0}$  and  $\{\beta_{i,k}\}_{k \geq 0}$  such that the inequalities (32) and (40) hold simultaneously, then the design objectives  $R1)$  and  $R2)$  are achieved at the same time, and the desired filtering parameters at each time instant can be computed via solving the corresponding matrix inequalities.

*Proof:* Based on Theorems 1 and 2, we arrive at Theorem 3 directly. Hence the proof is omitted here. ■

In the following, an iterative algorithm is presented to compute the sequences of the filtering parameters  $\{F_{i,k}\}_{k \geq 0}$  and  $\{H_{ij,k}\}_{k \geq 0}$  recursively.

**Algorithm 1: Computational Algorithm for  $\{F_{i,k}\}_{k \geq 0}$  and  $\{H_{ij,k}\}_{k \geq 0}$**

- 1) Initialization: Set  $k = 0$  and the maximum computation step  $k_{\max}$ . Set the triple  $(\mathfrak{P}_k, \mathbf{p}, \gamma, \Pi_i)$  for  $0 \leq k \leq k_{\max}$ . Then, by using  $P_k = \frac{1}{1-\mathbf{p}} \mathfrak{P}_k$ , factorize  $\{P_k\}$  appropriately to obtain the sequence of matrices  $\{Q_k\}$ . Select the initial values of  $x_0$  and  $\hat{x}_{i,0}$  satisfying (23). Then  $\hat{x}_0 = \text{col}_N\{\hat{x}_{i,0}\}$  is known.
- 2) With the obtained  $\hat{x}_k$  and  $Q_k$ , solve the RLMI (32) and (40) for  $\mathcal{F}_k$  and  $\mathcal{H}_k$ . Then  $F_{i,k}$  and  $H_{ij,k}$  can be obtained.
- 3) With the obtained  $\mathcal{F}_k$  and  $\mathcal{H}_k$ , compute  $\hat{x}_{i,k+1}$  according to (3). Then  $\hat{x}_{k+1} = \text{col}_N\{\hat{x}_{i,k+1}\}$  is obtained.
- 4) Set  $k = k + 1$ . If  $k > k_{\max}$ , exit. Otherwise, go to 2).

### D. Optimization problems

Note that the desired distributed filtering parameters acquired by using Theorem 3, if they exist, could form a set

and this provides extra design flexibility by making tradeoff among performance indices. In this subsection, in terms of two corollaries, we will propose two optimization problems (OPs), one is to minimize  $\mathfrak{P}_k$  (in the sense of matrix trace) for locally optimal filtering performance, and the other is to minimize  $\mathbf{p}$  at each time step to guarantee a local threshold probability which indicates the minimal chance with which the errors can be confined within the desired ellipsoid.

Denoting a set by

$$\mathcal{S}_k \triangleq \{\mathcal{F}_k, \mathcal{H}_k, \varrho_{i,k}^{(1)}, \varrho_k^{(2)}, \varrho_k^{(3)}, \varrho_{i,k}^{(4)}, \varrho_{i,k}^{(5)}, \varrho_{i,k}^{(6)}, \varepsilon_k, \varepsilon_{i,k}, \rho_{1,k}, \rho_{2,k}, \beta_{i,k}\}$$

and a function with respect to  $\mathbf{p}$  and  $\mathfrak{P}_{k+1}$  by

$$\mathcal{F}_k(\mathbf{p}, \mathfrak{P}_{k+1}) \triangleq \begin{bmatrix} -\Gamma_k & \tilde{\Omega}_k^T \mathcal{R}_{n_x, i}^T & \tilde{\Omega}_k^T \mathcal{R}_{n_x, i}^T \\ * & -\frac{1}{1-\mathbf{p}} \mathfrak{P}_{k+1} & 0 \\ * & * & -\frac{1}{1-\mathbf{p}} \mathfrak{P}_{k+1} \end{bmatrix},$$

we present the following optimization problems.

**OP1:** Minimization of  $\mathfrak{P}_k$  in the sense of matrix trace with fixed  $\mathbf{p}$  to seek the locally optimal filtering performance subject to prescribed probability constraint.

**Corollary 1:** Let  $\mathbf{p}$  be given. Under the conditions in Theorem 3, a sequence of minimized  $\{\mathfrak{P}_k\}_{k \geq 0}$  (in the sense of matrix trace) is guaranteed if the following optimization problem is solvable:

$$\begin{aligned} & \min_{\{\mathcal{S}_k, \mathfrak{P}_{k+1}\}} \text{trace}[\mathfrak{P}_{k+1}] \\ & \text{subject to (32) \& (40)} \end{aligned} \quad (41)$$

Next, assume that  $\mathbf{p}$  is time-varying and denote by  $\mathbf{p}_k$  the probability constraint at time instant  $k$ . By defining

$$\mathbf{s}_k \triangleq \frac{1}{1 - \mathbf{p}_k},$$

we put forward the following optimization problem.

**OP2:** Minimization of  $\mathbf{s}_k$  with fixed  $\mathfrak{P}_k$  to look for locally lower bound on the probability constraint at each time instant.

**Corollary 2:** Let  $\{\mathfrak{P}_k\}_{k > 1}$  be given. Under the conditions in Theorem 3, the lower bound on probability constraint  $\mathbf{p}_k$  at each time step is ensured if the following optimization problem is feasible:

$$\begin{aligned} & \min_{\{\mathcal{S}_k, \mathbf{s}_k\}} \mathbf{s}_k \\ & \text{subject to } \begin{cases} (40) \\ 1 < \mathbf{s}_k < +\infty \\ \begin{bmatrix} -\Gamma_k & \tilde{\Omega}_k^T \mathcal{R}_{n_x, i}^T & \tilde{\Omega}_k^T \mathcal{R}_{n_x, i}^T \\ * & -\mathbf{s}_k \mathfrak{P}_{k+1} & 0 \\ * & * & -\mathbf{s}_k \mathfrak{P}_{k+1} \end{bmatrix} \leq 0 \end{cases} \end{cases} \quad (42)$$

The proofs of Corollaries 1 and 2 are straightforward based on the previously obtained results and are therefore omitted here.

**Remark 5:** Notice that the RLMI algorithm proposed in this paper is based on LMI approach. As discussed in [2], the computational complexity of an LMI system is bounded by  $O(\mathcal{P} \mathcal{Q} \mathcal{L} \log(\mathcal{U}/\varepsilon))$  where  $\mathcal{P}$  represents the row size,  $\mathcal{Q}$  stands for the number of scalar decision variables,  $\mathcal{U}$  is a

data-dependent scaling factor and  $\varepsilon$  is relative accuracy set for algorithm. For instance, let us now look at the average  $H_\infty$  performance criterion (as proposed in Theorem 2), where the number of sensing nodes is  $N$ , the iteration time is  $T+1$  (since the time interval is  $[0, T]$ ) and the dimensions of variables are known from  $x_k \in \mathbb{R}^{n_x}$ ,  $y_{i,k} \in \mathbb{R}^{n_y}$ ,  $\nu_k \in \mathbb{R}^{n_\nu}$ ,  $\mu_k \in \mathbb{R}^{n_\mu}$  and  $\omega_k \in \mathbb{R}$ . Moreover, we assume that  $q_{i,k} \in \mathbb{R}^{n_q}$ ,  $\delta_{1k} \in \mathbb{R}^{n_{\delta 1}}$  and  $\delta_{2k} \in \mathbb{R}^{n_{\delta 2}}$ . The RLMI-based probabilistic ellipsoidal performance criterion is implemented recursively for  $T+1$  steps and, at each step, we need to solve the LMI (40) with  $\mathcal{P} = N(1 + 2n_x + n_q + n_{\delta 1} + n_{\delta 2} + n_\mu + n_\nu + n_y)$  rows and  $\mathcal{Q} = N(\frac{n_x + n_y}{2} + 7)$  scalar variables. Accordingly, the computational complexity of the proposed RLMI algorithm can be represented by  $O((T+1)\mathcal{P}\mathcal{Q})$ .

#### IV. NUMERICAL EXAMPLE

In this section, a numerical simulation example is presented to show the effectiveness of the algorithm proposed in this paper. We shall estimate the state of the Duffing equation as follows:

$$\ddot{z} + k_0 z(1 + k_d z^2) + c\dot{z} = 0$$

which is usually utilized to describe many practical physical processes, such as nonlinear vibration and nonlinear circuit [6]. In order to apply our proposed algorithm, first, we discretize the above differential equation and obtain the following difference equation:

$$f(x_k) = \begin{bmatrix} x_k^{(1)} + Tx_k^{(2)} \\ x_k^{(2)} - T(k_0 x_k^{(1)}(1 + k_d(x_k^{(1)})^2) + cx_k^{(2)}) \end{bmatrix},$$

where  $T$  is the sampling period;  $x_k^{(1)}$  and  $x_k^{(2)}$  are the 1st and 2nd entries of  $x_k$ , representing the sample values of  $z$  and  $\dot{z}$  at time  $kT$ , respectively.

The measurement matrices are selected as follows:

$$C_{1,k} = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad C_{2,k} = \begin{bmatrix} 0.5 & 0 \end{bmatrix}, \\ C_{3,k} = \begin{bmatrix} 0.6 & 0.5 \end{bmatrix}.$$

Other parameters are chosen as follows:

$$h(x_k) = \begin{bmatrix} 0 \\ -T(k_1 x_k^{(1)} + k_2 (x_k^{(1)})^3) \end{bmatrix}, \\ B_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad E_{1,k} = 0.1, \\ E_{2,k} = 0.15, \quad E_{3,k} = 0.12, \\ T = 0.2, \quad k_0 = 2.1, \\ k_d = 0.7, \quad c = 0.4, \\ k_1 = 0.5, \quad k_2 = 0.3, \\ \lambda = 0.85, \quad W_1 = W_2 = W_3 = 0.01.$$

Moreover, we select  $\nu_k = 0.36 \cos(k)$  and  $\mu_k = 0.4 \sin(2k)$ . Then, set  $V_k = 0.35$  and  $U_k = 0.4$ , we can verify that Assumption 1 is satisfied.

Assuming the communication topology among sensor nodes are shown in Fig. 1, we obtain the corresponding adjacency matrix  $\mathcal{L}$  as follows:

$$\mathcal{L} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Set the filtering performance indices by  $\mathbf{p} = 0.9$  and  $\gamma = 0.7$ , and the initial condition is given as follows:

$$x_0 = \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}, \quad \hat{x}_{1,0} = \begin{bmatrix} 0.1 \\ 0 \end{bmatrix}, \\ \hat{x}_{2,0} = \begin{bmatrix} 0.3 \\ 0 \end{bmatrix}, \quad \hat{x}_{3,0} = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}, \\ \sigma_{1,0} = \sigma_{2,0} = \sigma_{3,0} = 0.25.$$

In order to show the effectiveness of the proposed algorithm to mitigate the possible abnormal data, we here consider a cyber attack scenario with signal injection. For node  $i$  ( $i = 1, 2, 3$ ), the injected attack signals are generated by  $\psi_{i,k} = 3 + \varsigma_{i,k}$  where  $\varsigma_{i,k}$  are uniformly distributed random variables over interval  $[0, 1]$  with the expectations  $\bar{\varsigma}_{1,k} = 0.5$ ,  $\bar{\varsigma}_{2,k} = 0.35$  and  $\bar{\varsigma}_{3,k} = 0.4$ . During the time interval  $[20, 50]$ , the attackers injected  $\psi_{i,k}$  into the transmitted innovations  $y_{i,k} - \hat{y}_{i,k}$  to deteriorate the estimation performance.

By solving Corollary 1, the simulation results are obtained in Figs. 3–6. To be specific, Figs. 3–4 plot the entries  $x_k^{(1)}$  and  $x_k^{(2)}$  of  $x_k$  and their estimates  $\hat{x}_{i,k}^{(1)}$  and  $\hat{x}_{i,k}^{(2)}$ , respectively. The filtering errors of  $x_k^{(1)}$  and  $x_k^{(2)}$  are, respectively, depicted in Figs. 5–6. The occurrence of bias injection attacks are shown in Fig. 2 where the success of injection attack is recorded for each node at corresponding time step. It is easy to see that, from Figs. 3–6, despite the existence of bias injection attacks, the proposed filtering algorithm can effectively estimate the state of the target nonlinear system. Thus, the simulation results demonstrate the effectiveness and correctness of our developed algorithm.

In the following, with the purpose of further illustrating the effectiveness of our algorithm, we carry out a comparative simulation. For the target nonlinear system, under the identical attacks, we apply the traditional algorithm without a saturation constraint (i. e.,  $\sigma_{i,k} = \infty$ ). Moreover, the proposed algorithm and the traditional algorithm are noted by ‘ $\sigma$ -adaptive’ and ‘ $\sigma$ -inf’, respectively. The comparative simulation results are recorded in Figs. 7–10. Specifically, Figs. 7–8 depict  $x_k^{(1)}$  and  $x_k^{(2)}$  of  $x_k$  and their estimates  $\hat{x}_{i,k}^{(1)}$  and  $\hat{x}_{i,k}^{(2)}$ , respectively. The estimation errors of  $x_k^{(1)}$  and  $x_k^{(2)}$  are, respectively, described in Figs. 9–10. It can be seen from Figs. 7–10 that the estima-

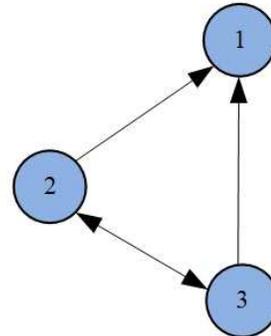


Fig. 1. The communication topology.

tion errors are diverging under attacks, which indicates that the traditional distributed filtering algorithm cannot provide satisfactory performance in such a case.

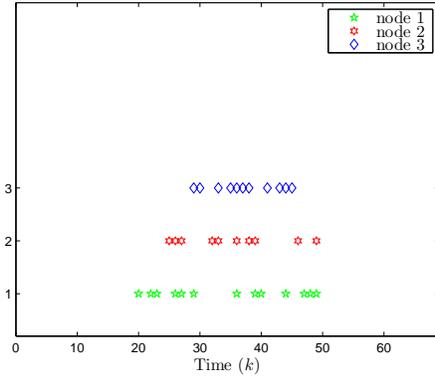


Fig. 2. The bias injection attacks.

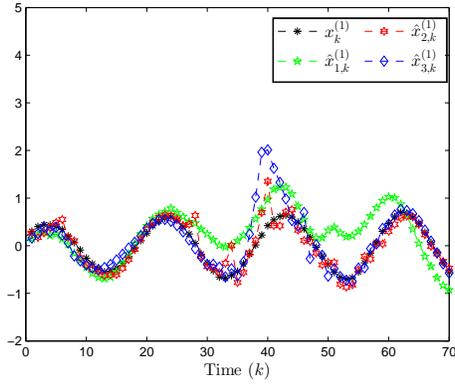


Fig. 3.  $x_k^{(1)}$  and  $\hat{x}_{i,k}^{(1)}$  ( $\sigma$ -adaptive).

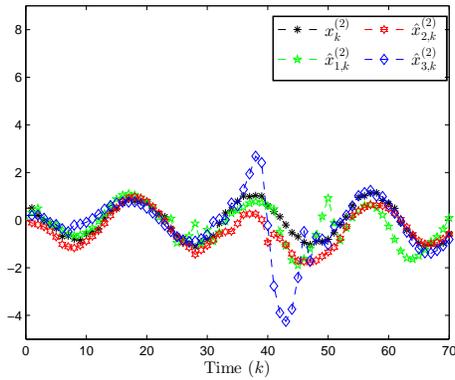


Fig. 4.  $x_k^{(2)}$  and  $\hat{x}_{i,k}^{(2)}$  ( $\sigma$ -adaptive).

## V. CONCLUSION

In this paper, the distributed filtering problem has been discussed for a class of nonlinear systems subject to innovation

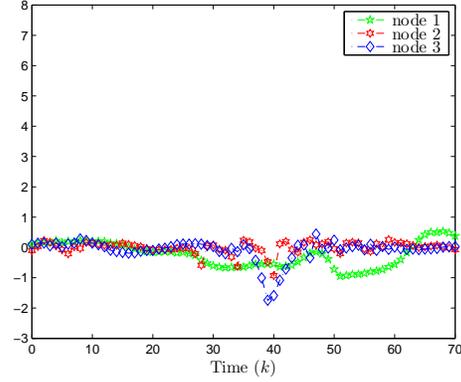


Fig. 5.  $\tilde{x}_k^{(1)}$  ( $\sigma$ -adaptive).

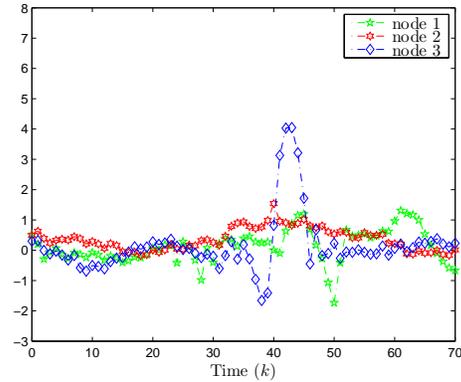


Fig. 6.  $\tilde{x}_k^{(2)}$  ( $\sigma$ -adaptive).

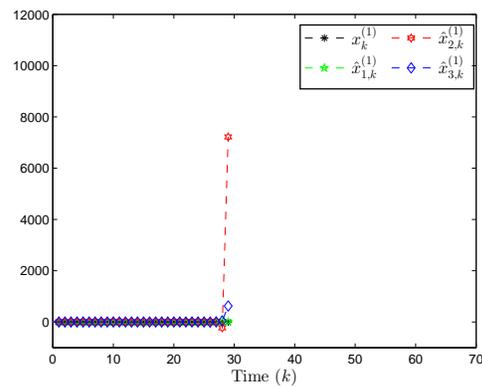
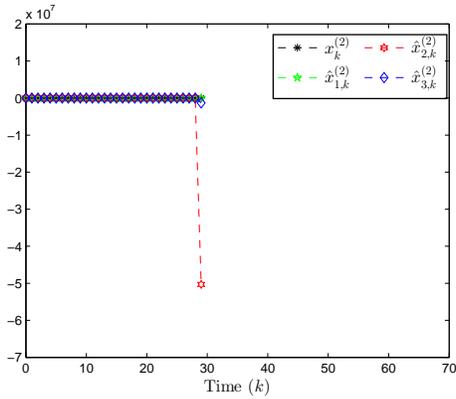
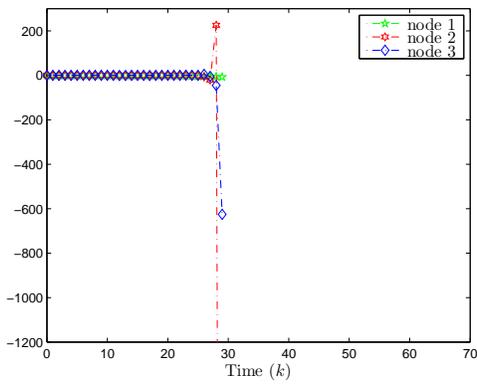
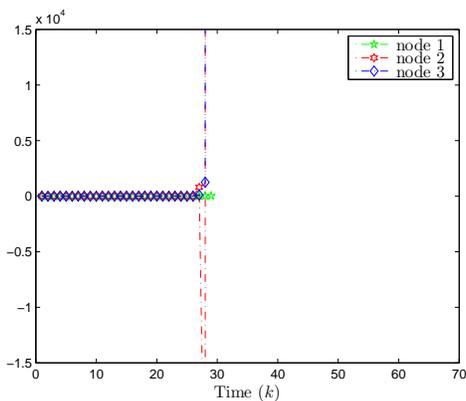


Fig. 7.  $x_k^{(1)}$  and  $\hat{x}_{i,k}^{(1)}$  ( $\sigma$ -inf).


 Fig. 8.  $x_k^{(2)}$  and  $\hat{x}_{i,k}^{(2)}$  ( $\sigma - inf$ ).

 Fig. 9.  $\hat{x}_k^{(1)}$  ( $\sigma - inf$ ).

 Fig. 10.  $\hat{x}_k^{(2)}$  ( $\sigma - inf$ ).

constraint. The innovation constraint has been imposed on the transmitted innovation in order to mitigate the effects from possible abnormal data during the signal transmission among nodes. By resorting to recursive linear matrix inequality approach, sufficient conditions have been established for the existence of the desired distributed filter, ensuring that the estimation error is confined within the prespecified ellipsoidal areas with a guaranteed probability, and meanwhile, the prescribed average  $H_\infty$  criterion is satisfied. The desired filtering parameters can be computed by solving the corresponding set of matrix inequalities recursively from step to step. Within the established framework, two optimization problems have been proposed to look for certain locally suboptimal filtering parameters. Finally, an illustrative numerical example has been provided to verify the applicability of the proposed distributed filtering paradigm. One of our future research topics is to extend the main results to more general systems with more performance requirements [16], [18], [22], [27].

## REFERENCES

- [1] H. K. Khalil, *Nonlinear systems*, Upper Saddle River, Prentice-Hall, NJ, 1996.
- [2] S. Boyd, L. Ghaoui, E. Feron and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, Philadelphia: SIAM Studies in Applied Mathematics, 1994.
- [3] A. Abdessameud, I. G. Polushin and A. Tayebi, Synchronization of nonlinear systems with communication delays and intermittent information exchange, *Automatica*, Vol. 59, pp. 1–8, 2015.
- [4] H. T. Banks, B. M. Lewis and H. T. Tran, Nonlinear feedback controllers and compensators: a state-dependent Riccati equation approach, *Computational Optimization and Applications*, Vol. 37, No. 2, pp. 177–218, 2007.
- [5] R. Carli, A. Chiuso, L. Schenato and S. Zampieri, Distributed Kalman filtering based on consensus strategies, *IEEE Transactions on Selected Areas in Communications*, Vol. 26, No. 4, pp. 622–633, 2010.
- [6] W. Chai and X. Sun, Nonlinear set membership filtering using ellipsoids and its application in fault diagnosis, *Acta Aeronautica Et Astronautica Sinica*, vol. 28, no. 4, pp. 948–952, 2007.
- [7] D. Ciunzo, A. Aubry and V. Carotenuto, Rician MIMO channel- and jamming-aware decision fusion, *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3866–3880, Aug. 2017.
- [8] A. P. Dani, S. J. Chung and S. Hutchinson, Observer design for stochastic nonlinear systems via contraction-based incremental stability, *IEEE Transactions on Automatic Control*, Vol. 60, No. 3, pp. 700–714, 2015.
- [9] P. Dong, Z. Jing, H. Leung, K. Shen and M. Li, Robust consensus nonlinear information filter for distributed sensor networks with measurement outliers, *IEEE Transactions on Cybernetics*, Vol. 49, No. 10, pp. 3731–3743, 2019.
- [10] X. Ge, Q.-L. Han, X.-M. Zhang, L. Ding and F. Yang, Distributed event-triggered estimation over sensor networks: A survey, *IEEE Transactions on Cybernetics*, in press, DOI:10.1109/TCYB.2019.2917179.
- [11] L. Guo and Y.-F. Huang, Frequency-domain set-membership filtering and its applications, *IEEE Transactions on Signal Processing*, Vol. 55, No. 4, pp. 1326–1338, 2007.
- [12] F. Han, G. Wei, D. Ding and Y. Song, Finite-horizon bounded  $H_\infty$  synchronisation and state estimation for discrete-time complex networks: Local performance analysis, *IET Control Theory & Applications*, vol. 11, no. 6, pp. 827–837, Apr. 2017.
- [13] W. Li, G. Wei, F. Han, and Y. Liu, Weighted average consensus-based unscented Kalman filtering, *IEEE Transactions on Cybernetics*, vol. 46, no. 2, pp. 558–567, 2016.
- [14] L. Ma, Z. Wang, H.-K. Lam and N. Kyriakoulis, Distributed event-based set-membership filtering for a class of nonlinear systems with sensor saturations over sensor networks, *IEEE Transactions on Cybernetics*, Vol. 47, No. 11, pp. 3772–3783, 2017.
- [15] L. Ma, Z. Wang, Q.-L. Han and Y. Liu, Distributed filtering for nonlinear time-delay systems over sensor networks subject to multiplicative link noises and switching topology, *International Journal of Robust and Nonlinear Control*, Vol. 29, pp. 2941–2959, 2019.

- [16] S. K. Nguang, W. Assawinchaichote and P. Shi,  $H_\infty$  filter for uncertain Markovian jump nonlinear systems: An LMI approach, *Circuits, Systems and Signal Processing*, Vol. 26, No. 6, pp. 853–874, 2007.
- [17] A. De Paola, S. Gaglio, G. Lo Re, F. Milazzo and M. Ortolani, Adaptive distributed outlier detection for WSNs, *IEEE Transactions on Cybernetics*, Vol. 45, No. 5, pp. 888–899, 2015.
- [18] C. V. Rao, M. Wolf and A. P. Arkin, Control, exploitation and tolerance of intracellular noise, *Nature*, Vol. 420, No. 6912, pp. 231–237, 2002.
- [19] S. Seifzadeh, B. Khaleghi and F. Karray, Distributed soft-data-constrained multi-model particle filter, *IEEE Transactions on Cybernetics*, Vol. 45, No. 3, pp. 384–394, 2015.
- [20] S. P. Talebi and S. Werner, Distributed Kalman filtering and control through embedded average consensus information fusion, *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4396–4403, 2019.
- [21] S. P. Talebi, S. Kanna and D. P. Mandic, 'A Distributed quaternion Kalman filter with applications to smart grid and target tracking,' *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 477–488, 2016.
- [22] H. Tan, B. Shen, Y. Liu, A. Alsaedi and B. Ahmad, Event-triggered multi-rate fusion estimation for uncertain system with stochastic nonlinearities and colored measurement noises, *Information Fusion*, vol. 36, pp. 313–320, Jul. 2017.
- [23] E. Tian, Z. Wang, L. Zou and D. Yue, Probabilistic-constrained filtering for a class of nonlinear systems with improved static event-triggered communication, *International Journal of Robust and Nonlinear Control*, Vol. 29, No. 5, pp. 1484–1498, 2019.
- [24] V. Ugrinovskii, Distributed robust filtering with  $H_\infty$  consensus of estimates, *Automatica*, Vol. 47, No. 1, pp. 1–13, 2011.
- [25] Y.-L. Wang, P. Shi, C.-C. Lim, and Y. Liu, Event-triggered fault detection filter design for a continuous-time networked control system, *IEEE Transactions on Cybernetics*, Vol. 46, No. 12, pp. 3414–3426, 2016.
- [26] G. Wei, S. Liu, Y. Song and Y. Liu, Probability-guaranteed set-membership filtering for systems with incomplete measurements, *Automatica*, vol. 60, pp. 12–16, Oct. 2015.
- [27] Y. Xu, R. Lu, P. Shi, H. Li and S. Xie, Finite-time distributed state estimation over sensor networks with round-robin protocol and fading channels, *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 336–345, Jan. 2018.
- [28] I. Yaesha, S. Boyarskia and U. Shaked, Probability-guaranteed robust  $H_\infty$  performance analysis and state-feedback design, *Systems and Control Letters*, Vol. 48, pp. 351–364, 2003.
- [29] F. Yang and Y. Li, Set-membership filtering for systems with sensor saturation, *Automatica*, Vol. 45, No. 8, pp. 1896–1902, 2009.
- [30] X.-M. Zhang, Q.-L. Han, X. Ge, D. Ding, L. Ding, D. Yue and C. Peng, Networked control systems: A survey of trends and techniques, *IEEE/CAA Journal of Automatica Sinica*, Vol. 7, No. 1, pp. 1–17, 2020.
- [31] W.-A. Zhang, G. Feng and L. Yu, Multi-rate distributed fusion estimation for sensor networks with packet losses, *Automatica*, vol. 48, no. 9, pp. 2016–2028, Sept. 2012.

## VI. APPENDIX

### A. Proof of Lemma 5

*Proof:* The proof of this lemma is carried out according to the principle of mathematical induction. First, we know from the initial condition (23) that

$$\mathbb{E}\{(x_0 - \hat{x}_{i,0})^T P_0^{-1} (x_0 - \hat{x}_{i,0})\} \leq 1. \quad (43)$$

Second, suppose that the following inequality is true at the time step  $k > 0$ :

$$\mathbb{E}\{(x_k - \hat{x}_{i,k})^T P_k^{-1} (x_k - \hat{x}_{i,k})\} \leq 1. \quad (44)$$

Then, it remains to verify that, at time step  $k + 1$ , inequality (31) is true under the condition given in this lemma. To this end, since (44) is true, there exist vectors  $q_{i,k} \in \mathbb{R}^{n_q}$  ( $i = 1, 2, \dots, N$ ) with  $\mathbb{E}\{q_{i,k}^T q_{i,k}\} \leq 1$  such that

$$x_k = \hat{x}_{i,k} + Q_k q_{i,k}. \quad (45)$$

By denoting  $q_k \triangleq \text{col}_N\{q_{i,k}\}$  and  $Q_k \triangleq \text{diag}_N\{Q_k\}$ , (45) is described by

$$\xi_k = \hat{x}_k + Q_k q_k. \quad (46)$$

Subsequently, the dynamics of estimation error (20) is rewritten by

$$\begin{aligned} \tilde{x}_{k+1} = & \hat{f}_k - \mathcal{F}_k \hat{x}_k + \hat{h}_k \omega_k + (\Phi_k + \Psi_k \omega_k - \mathcal{H}_k \mathcal{G} \mathcal{C}_k) \mathcal{Q}_k q_k \\ & + \mathcal{L} \delta_{1k} + \Sigma \delta_{2k} \omega_k + \mathcal{B}_k \nu_k + \mathcal{D}_k \nu_k \omega_k \\ & - \mathcal{H}_k \mathcal{G} \mathcal{E}_k (\mathbf{1}_N \otimes I_{n_\mu}) \mu_k - \mathcal{H}_k \phi_k \end{aligned} \quad (47)$$

where

$$\begin{aligned} \delta_{1k} & \triangleq \text{col}_N\{\delta_{1i,k}\} = \Delta_1 \mathcal{Q}_k q_k, \\ \delta_{2k} & \triangleq \text{col}_N\{\delta_{2i,k}\} = \Delta_2 \mathcal{Q}_k q_k. \end{aligned}$$

By denoting

$$\eta_k \triangleq [1 \quad q_k^T \quad \delta_{1k}^T \quad \delta_{2k}^T \quad \nu_k^T \quad \mu_k^T \quad \phi_k^T]^T,$$

we further express (47) as

$$\tilde{x}_{k+1} \triangleq (\bar{\Omega}_k + \tilde{\Omega}_k \omega_k) \eta_k \quad (48)$$

where  $\bar{\Omega}_k$  and  $\tilde{\Omega}_k$  are defined in (29) and (30), respectively.

It follows from (2) and (45) that the vectors  $q_{i,k}$ ,  $\nu_k$  and  $\mu_k$  satisfy

$$\mathbb{E}\{q_{i,k}^T q_{i,k}\} \leq 1, \quad \nu_k^T V_k^{-1} \nu_k \leq 1, \quad \mu_k^T U_k^{-1} \mu_k \leq 1 \quad (49)$$

which can be rewritten in terms of  $\eta_k$  as follows:

$$\begin{cases} \mathbb{E}\{\eta_k^T \text{diag}\{-1, \Theta_{n_q, i}, 0, 0, 0, 0\} \eta_k\} \leq 0 \\ \eta_k^T \text{diag}\{-1, 0, 0, 0, V_k^{-1}, 0, 0\} \eta_k \leq 0 \\ \eta_k^T \text{diag}\{-1, 0, 0, 0, 0, U_k^{-1}, 0\} \eta_k \leq 0 \end{cases} \quad (50)$$

From definitions of  $\delta_{1i,k}$ ,  $\delta_{2i,k}$  and by noting that  $\|\Delta_{1i}\| \leq 1$  and  $\|\Delta_{2i}\| \leq 1$ , we have

$$\begin{cases} \delta_{1i,k}^T \delta_{1i,k} - q_{i,k}^T Q_k^T Q_k q_{i,k} \leq 0 \\ \delta_{2i,k}^T \delta_{2i,k} - q_{i,k}^T Q_k^T Q_k q_{i,k} \leq 0 \end{cases} \quad (51)$$

which can be expressed in terms of  $\eta_k$  as follows:

$$\begin{cases} \eta_k^T \text{diag}\{0, -\Theta_{n_q, i} Q_k^T Q_k, \Theta_{n_i, i}, 0, 0, 0, 0\} \eta_k \leq 0 \\ \eta_k^T \text{diag}\{0, -\Theta_{n_q, i} Q_k^T Q_k, 0, \Theta_{n_\Sigma, i}, 0, 0, 0\} \eta_k \leq 0 \end{cases} \quad (52)$$

Next, we rewrite (18) in terms of  $\eta_k$  by

$$\eta_k^T \Xi_{i,k} \eta_k \leq 0 \quad (53)$$

where  $\Xi_{i,k}$  is defined in (26).

By considering that

$$y_{i,k} - \hat{y}_{i,k} = \mathcal{W}_{i,k} \eta_k$$

where  $\mathcal{W}_{i,k}$  is defined in (27), we describe the innovation constraint (4) in terms of  $\eta_k$  as follows:

$$\eta_k^T \Upsilon_{i,k} \eta_k = 0 \quad (54)$$

with  $\Upsilon_{i,k}$  defined in (28).

On the other hand, by resorting to the Schur complement equivalence (Lemma 2), it can be seen that the set of matrix inequalities (24) hold if and only if

$$\bar{\Omega}_k^T \mathcal{R}_{n_x, i} P_{k+1}^{-1} \mathcal{R}_{n_x, i} \bar{\Omega}_k + \tilde{\Omega}_k^T \mathcal{R}_{n_x, i} P_{k+1}^{-1} \mathcal{R}_{n_x, i} \tilde{\Omega}_k - \Gamma_k \leq 0, \quad (55)$$



Moreover, it is inferred readily from (4) that

$$\zeta_k^T \mathcal{C}_{i,k} \zeta_k = 0 \quad (69)$$

with  $\mathcal{C}_{i,k}$  defined in (36).

Applying Schur Complement Lemma to inequality (33) leads to

$$\begin{bmatrix} \tilde{\Lambda}_k & \tilde{\mathcal{A}}_k^T & \tilde{\mathcal{M}}_k^T \\ \tilde{\mathcal{A}}_k & -Y_{k+1}^{-1} & 0 \\ \tilde{\mathcal{M}}_k & 0 & -Y_{k+1}^{-1} \end{bmatrix} + \text{diag}\{\rho_{1,k}^{-1} \tilde{\mathcal{L}} \tilde{\mathcal{L}}^T + \rho_{1,k} \tilde{\mathcal{L}}^T \tilde{\mathcal{L}} + \rho_{2,k}^{-1} \tilde{\Sigma} \tilde{\Sigma}^T + \rho_{2,k} \tilde{\Sigma}^T \tilde{\Sigma}, 0, 0\} \leq 0. \quad (70)$$

By resorting to Lemma 3, we know that the inequality (70) is true if and only if

$$\begin{bmatrix} \tilde{\Lambda}_k & \tilde{\mathcal{A}}_k^T & \tilde{\mathcal{M}}_k^T \\ \tilde{\mathcal{A}}_k & -Y_{k+1}^{-1} & 0 \\ \tilde{\mathcal{M}}_k & 0 & -Y_{k+1}^{-1} \end{bmatrix} + \tilde{\mathcal{L}} \Delta_1 \tilde{\mathcal{L}} + (\tilde{\mathcal{L}} \Delta_1 \tilde{\mathcal{L}})^T + \tilde{\Sigma} \Delta_2 \tilde{\Sigma} + (\tilde{\Sigma} \Delta_2 \tilde{\Sigma})^T \leq 0. \quad (71)$$

Accordingly, it is obtained from inequality (71) that

$$\begin{bmatrix} \tilde{\Lambda}_k & \tilde{\mathcal{A}}_k^T & \tilde{\mathcal{M}}_k^T \\ \tilde{\mathcal{A}}_k & -Y_{k+1}^{-1} & 0 \\ \tilde{\mathcal{M}}_k & 0 & -Y_{k+1}^{-1} \end{bmatrix} \leq 0, \quad (72)$$

which further implies that

$$\Lambda_k - \varepsilon_k \mathcal{C}_k - \sum_{i=1}^N \varepsilon_{i,k} \mathcal{C}_{i,k} \leq 0. \quad (73)$$

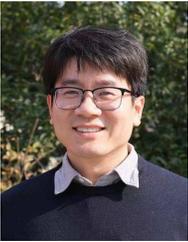
According to Lemma 1, we have

$$\zeta_k^T \Lambda_k \zeta_k \leq 0, \quad (74)$$

which, in combination with (66) and  $Y_0 \leq \tilde{\Pi}$ , results in

$$\frac{1}{N} \mathbb{E} \left\{ \sum_{k=0}^T \|\tilde{x}_k\|^2 \right\} \leq \gamma^2 \sum_{k=0}^T \|\vartheta_k\|^2 + \gamma^2 \frac{1}{N} \tilde{x}_0^T \text{diag}_N \{\Pi_i\} \tilde{x}_0. \quad (75)$$

Therefore, the average  $H_\infty$  performance defined in (8) is achieved. The proof is complete now. ■



**Lifeng Ma** received the B.Sc. degree in Automation from Jiangsu University, Zhenjiang, China, in 2004 and the Ph.D. degree in Control Science and Engineering from Nanjing University of Science and Technology, Nanjing, China, in 2010. From August 2008 to February 2009, he was a Visiting Ph.D. Student in the Department of Information Systems and Computing, Brunel University London, U.K. From January 2010 to April 2010 and May 2011 to September 2011, he was a Research Associate in the Department of Mechanical Engineering, the University of Hong Kong. From March 2015 to February 2017, he was a Visiting Research Fellow at the King's College London, U.K.

He is currently a Professor in the School of Automation, Nanjing University of Science and Technology, Nanjing, China. His current research interests include control and signal processing, machine learning and deep learning. He has published more than 50 papers in refereed international journals. He serves as an editor for *Neurocomputing* and *International Journal of Systems Science*.



**Zidong Wang** (SM'03-F'14) was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 600 papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, the Editor-in-Chief for *Neurocomputing*, and an Associate Editor for 12 international journals including *IEEE Transactions on Automatic Control*, *IEEE Transactions on Control Systems Technology*, *IEEE Transactions on Neural Networks*, *IEEE Transactions on Signal Processing*, and *IEEE Transactions on Systems, Man, and Cybernetics-Part C*. He is a Member of the Academia Europaea, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.



**Yun Chen** was born in Zhejiang Province, China. He received the B.E. degree in thermal engineering in 1999 from Central South University of Technology (Central South University), Changsha, China, and the M.E. degree in engineering thermal physics in 2002 and Ph.D. degree in control science and engineering in 2008, both from Zhejiang University, Hangzhou, China.

From August 2009 to August 2010, he was a visiting fellow with the School of Computing, Engineering and Mathematics, University of Western Sydney, Australia. From December 2016 to December 2017, he was an academic visitor with the Department of Mathematics, Brunel University London, UK. In 2002, he joined Hangzhou Dianzi University, China, where he is currently a Professor. His research interests include stochastic and hybrid systems, robust control and filtering.



**Xiaojian Yi** was born in 1987. He received the B.S. degree in control technology in 2010 from the North University of China, Taiyuan, China, and the M.S. degree in 2012 and Ph.D. degree in 2016 both in reliability engineering from Beijing Institute of Technology, Beijing, China. During 2015-2016, he was a jointly trained PhD student in the University of Ottawa, Canada, to study robot reliability and maintenance. From 2016 to 2020, he was an Associate Professor with the China North Vehicle Research Institute. He is currently an Associate Professor with

the Beijing Institute of Technology, Beijing, China. He is the author of two books and more than 100 articles, and is also the holder of 8 patents. His research interests include system reliability analysis, intelligent control, fault diagnosis and health management.