Distributed Recursive Filtering over Sensor Networks with Non-Logarithmic Sensor Resolution

Hongwei Chen, Zidong Wang, Bo Shen, and Jinling Liang

Abstract—Sensor resolution, which is one of the most important parameters/specifications for almost all kinds of sensors, plays an important role in any signal processing problems. This paper deals with the distributed filtering problem for a class of discrete time-varying stochastic systems subject to non-logarithmic sensor resolution and stochastic nonlinearities. The soft measurement technique is exploited in the filter design to overcome the difficulties resulting from the sensor-resolution-induced (SRI) uncertainty. The aim of the presented filtering problem is to construct the distributed filter over a sensor network such that, in the presence of SRI uncertainty and stochastic nonlinearity, an upper bound on the filtering error covariance is guaranteed and subsequently minimized by appropriately designing the filer parameters at each time instant. Moreover, a matrix simplification method is utilized to tackle the difficulties stemming from the sparsity of sensor networks. Finally, a numerical example is employed to illustrate the effectiveness of the proposed filtering scheme.

Index Terms—Distributed filtering, recursive filtering, sensor resolution, stochastic nonlinearity, wireless sensor networks.

I. INTRODUCTION

In wireless sensor networks (WSNs), the distributed filtering or state estimation problem has been a central topic that has received a persistent research interest, and a great number of distributed filtering/estimation algorithms have been reported in the literature. For example, the distributed filtering problems have been investigated in [9] within the H_∞ filtering framework, and the distributed filtering schemes have been proposed in [3], [7], [16], [18]–[20] based on the Kalman filtering technology. Representatively, in [19], [20], a dynamic average-consensus scheme has been proposed and the consensus-based distributed Kalman filters have been designed. In [18], three novel distributed Kalman filtering algorithms have been proposed over sensor networks by resorting to the dynamic consensus protocols, under which the local filter obtains the state estimate based on the measurements not only from the corresponding sensor node itself but also from its neighboring sensor nodes.

As is well known, the nonlinearity is often encountered in real-world systems and is viewed as one of the essential sources that complicate the system analysis and synthesis [11], [28]. It is worth pointing out that, in some practical systems such as the networked control systems, the nonlinear disturbances may occur in a stochastic fashion due mainly to unpredictable fluctuations of the network load/traffic [21]. Such a randomly occurring nonlinear phenomenon is usually referred to as the *stochastic nonlinearity* which, unfortunately, would invalidate those control/filtering algorithms suitable for linear and/or deterministic systems. As such, it is of theoretical necessity and practical significance to make dedicated efforts in dealing with the impact from various stochastic nonlinearities on the corresponding control/filtering problems. In particular, in [25], a kind of stochastic nonlinearities has been characterized by a stochastic variable with known first- and second-order moments. Subsequently, considerable research attention has been paid to the filtering and control issues with stochastic nonlinearities, see e.g. [12], [15], [21], [24], [26].

It is well known that the resolution is one of the most important specifications for sensors applied in engineering practice. Sensor resolution (SR) is basically understood as the capability of a sensor to sense the smallest change of measurements, where a low resolution means that the information received by sensors is inaccurate to an extent. Clearly, the effect caused by the low resolution should be adequately taken into consideration. For example, in a multi-target tracking problem, the SR is closely related to the capability of distinguishing the individual target from others and thus plays a vitally important role in achieving the ideal tracking performance. In [4], [5], [8], the tracking problem has been investigated for multiple maneuvering targets, where the case of false (and possibly unresolved) measurements induced by the SR is tackled. In [23], the SR-induced effects have been dealt with for an arbitrary number of targets, and a multi-target SR model has been proposed which can be incorporated into the traditional Bayesian tracking filters.

Powered by the advanced sensor technologies, WSNs have been extensively deployed to record and monitor a wide variety of valuable information including seismic, thermal, visual, or other types of measured observations. In WSNs, it is often the case that the sensors are low-cost and low-power disposable devices [27], which implies that these sensors might have limited resolution. As a result, the available measurements obtained by these sensors are inaccurate to a certain extent, which infers that the performance of the distributed filers based on the limited-resolution sensors would be unavoidably deteriorated. A fundamental issue then arises for the WSNs: how to construct the effective distributed filters to estimate the real states of the target plants from the SR-induced inaccurate measurements. For many applications of WSNs such as target tracking, the SR-induced inaccurate measurements, if not felicitously handled, would seriously impair the performance of the distributed filters.

It has been well recognized that ignoring the limited resolution of the sensors in a WSN might result in an incorrect interpretation of the data [23]. In order to cope with the data association problem in tracking, various alternative computational models have been established to provide a reliable description of the resolution phenomenon. One effective way is to represent the capability of the sensor to resolve individual targets in a group by a resolution probability. Following this line of modeling, a grid-based SR model has been proposed in the fundamental paper [8], where the model is integrated into the joint probabilistic data association filter for tracking the multiple targets in a cluttered environment. Such a model has been further extended and applied to a variety of target tracking algorithms [4], [5]. It is worth mentioning that the aforementioned SR models are appropriate

H. Chen and B. Shen are with the College of Information Science and Technology, Donghua University, Shanghai 201620, China, and are also with the Engineering Research Center of Digitalized Textile and Fashion Technology, Ministry of Education, Shanghai 201620, China. (e-mail: hongwei@dhu.edu.cn; bo.shen@dhu.edu.cn)

Z. Wang is with the Department of Computer Science, Brunel University London, Middlesex UB8 3PH, U.K. (e-mail: Zidong.Wang@brunel.ac.uk)

J. Liang is with the School of Mathematics, Southeast University, Nanjing 210096, China. (e-mail: jinlliang@seu.edu.cn)

for the situation where the measurements from different targets are merged into one due to an inherent resolution threshold. As for the SR-induced inaccurate measurements, a suitable model is still lacking in the literature, which motivates the work of this paper.

Indeed, in the context of the distributed filtering problem, it remains challenging as how to model the SRs and how to design SR-resistant filters. In the interesting paper [29], a model of the so-called non-logarithmic SR has been proposed and a robust filter has then been designed. As for the more practical WSNs, to date, the SR has not received adequate research attention yet despite the fact that the SR phenomenon is deemed to be more prevalent in the WSN setting. Clearly, the SR-induced effect would contribute substantially to the complexity of designing distributed filters over WSNs, and the issue would be even more intricate when the stochastic nonlinearity is also involved. In order to meet the challenges mentioned above, the objective of this paper is to design a recursive filter for a class of discrete time-varying systems subject to the non-logarithmic SR and stochastic nonlinearities, where the design of gains is performed offline and requires the global information of the sensors. The main technical contributions of this paper can be highlighted as follows: 1) an appropriate model is proposed for the SR which characterizes the actually available measurements; 2) the SR-induced effects are dedicatedly sorted out by exploiting the soft measurement technique; and 3) a set of distributed filters is designed recursively such that an upper bound of the filtering error covariance is minimized at each time step. Finally, a simulation example is presented to verify the effectiveness of the proposed filtering scheme.

Notations: $\mathbb{N}^+$ denotes the set of positive integers. $[x]_j$ represents the jth component of a vector x. $\mathbb{E}\{x|y\}$ means the mathematical expectation of x conditional on y. For matrices M_i ($i = 1, 2, \ldots, n$), $\mathrm{diag}_n\{M_i\}$ refers to a block-diagonal matrix where matrix M_i is in the ith main diagonal block. $\{M_{ij}\}_{n \times n}$ means a partitioned $n \times n$ block matrix where M_{ij} is the (i, j)th block submatrix. $\mathrm{col}_n\{x_i\}$ is an augmented vector $[x_1^T, x_2^T, \ldots, x_n^T]^T$. $\mathbf{1}_{m \times n}$ means an $m \times n$ matrix with all entries equal to one.

II. MODEL FORMULATION AND PRELIMINARIES

Denote by $y^r \in \mathbb{R}^{n_y}$ the measurement received by a sensor with resolution $S = [s_1, s_2, \ldots, s_{n_y}]^T \in \mathbb{R}^{n_y}$, where the definition of SR is defined as follows.

Definition 1 ([29]): Let $[y^r]_j$ ($j = 1, 2, \ldots, n_y$) be the jth physical quantity of the sensor measurement taking values in the set $\{\kappa s_j | \kappa = 0, \pm 1, \ldots, \pm z\}$, where z is a given positive integer. Then, $S = [s_1, s_2, \ldots, s_{n_y}]^T$ is called the resolution of the sensor.

In this paper, we consider the following class of discrete time-varying systems with stochastic nonlinearity:

$$x_{k+1} = F_k x_k + \tilde{f}(k, x_k, \eta_k) + G_k w_k \qquad (1)$$

where $k \in \mathbb{N}$ represents the discrete time index, $x_k \in \mathbb{R}^{n_x}$ captures the state of the process, $F_k \in \mathbb{R}^{n_x \times n_x}$ and $G_k \in \mathbb{R}^{n_x \times n_w}$ are known deterministic matrices, $w_k \in \mathbb{R}^{n_w}$ denotes the process noise, η_k is a zero-mean Gaussian noise sequence, and $\tilde{f}(k, x_k, \eta_k) \in \mathbb{R}^{n_x}$ describes the stochastic nonlinearity with $\tilde{f}(k, 0, \eta_k) = 0$. The initial state x_0 is a Gaussian random variable with mean $\bar{x}_0$ and covariance $\mathcal{X}_0$. The stochastic nonlinearity under consideration has the first moment $\mathbb{E}\{\tilde{f}(k, x_k, \eta_k)|x_k\} = 0$ and the covariances given by

$$\mathbb{E}\{\tilde{f}(k_1, x_{k_1}, \eta_{k_1})\tilde{f}^T(k_2, x_{k_2}, \eta_{k_2})|x_{k_2}\} = 0, \quad k_1 \neq k_2 \qquad (2)$$

$$\mathbb{E}\{\tilde{f}(k, x_k, \eta_k)\tilde{f}^T(k, x_k, \eta_k)|x_k\} = \sum_{\lambda=1}^{l} \Xi_{\lambda,k} x_k^T \Psi_{\lambda,k} x_k \qquad (3)$$

where l is a given positive integer, $\Xi_{\lambda,k}$ and $\Psi_{\lambda,k}$ ($\lambda = 1, 2, \ldots, l$) are known matrices with compatible dimensions.

Remark 1: The stochastic nonlinearity characterized by (2)–(3) was first proposed in the pioneering work [14] by Jacobson and then considered in [26] for the state estimation problem of uncertain nonlinear stochastic systems. Such kind of nonlinearities covers several well-known nonlinear phenomena as special cases including the linear state-dependent multiplicative white noises and the random variables with their powers depending on the sign of a nonlinear function. We refer the readers to [14], [26] for more details.

The output of system (1) is measured through a WSN consisting of n sensor nodes. The topology of the WSN is represented by a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{H})$ of order n, where $\mathcal{V} = \{1, 2, \ldots, n\}$ stands for the set of sensor nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ means the set of directed edges in digraph $\mathcal{G}$, and $\mathcal{H} = \{a_{ij}\}_{n \times n}$ refers to the weighted adjacency matrix of $\mathcal{G}$. The elements of $\mathcal{H}$ are nonnegative, and $\mathcal{G}$ has a directed edge $(i, j) \in \mathcal{E}$ if and only if $a_{ij} > 0$, which implies that the ith sensor node can receive the information from the jth sensor node. In this case, the jth sensor node is called as a neighbor of the ith sensor node. For notational convenience, the ith sensor node plus its neighbors is denoted by $\mathcal{N}_i \triangleq \{j \in \mathcal{V}|(i, j) \in \mathcal{E}\}$.

The ideal measurement model of the ith ($i = 1, 2, \ldots, n$) sensor is given as follows:

$$y_{i,k}^p = H_{i,k} x_k + v_{i,k} \qquad (4)$$

where $y_{i,k}^p \in \mathbb{R}^{n_y}$ means the perfect measurement of the ith sensor, and $v_{i,k} \in \mathbb{R}^{n_y}$ is the measurement noise. Moreover, the noises w_k and $v_{i,k}$ are mutually independent zero-mean Gaussian white processes with covariances $Q_k > 0$ and $R_{i,k} > 0$, respectively. In addition, they are assumed to be uncorrelated with the other stochastic variables x_0 and η_k.

In practical applications, the available measurements obtained by sensor i ($i = 1, 2, \ldots, n$) with non-logarithmic SR $S_i = [s_{i,1}, s_{i,2}, \ldots, s_{i,n_y}]^T$ satisfy

$$[y_{i,k}^r]_j = Q([y_{i,k}^p]_j) = \left[\frac{[y_{i,k}^p]_j}{[S_i]_j}\right][S_i]_j \qquad (5)$$

where $[\cdot]$ denotes the integer part[1] of "$\cdot$", $y_{i,k}^r \in \mathbb{R}^{n_y}$ represents the available measurement from sensor i with resolution S_i, and $[S_i]_j$ stands for the jth element of the resolution S_i for the ith sensor.

The soft sensor model adopted in this paper is given by

$$[y_{i,k}^s]_j \triangleq \mathbb{E}\left\{[y_{i,k}^p]_j \big| I_{i,j}([y_{i,k}^r]_j)\right\} \qquad (6)$$

where $y_{i,k}^s \in \mathbb{R}^{n_y}$ means the soft measurement of the ith sensor and the interval $I_{i,j}([y_{i,k}^r]_j)$ is defined by

$$I_{i,j}([y_{i,k}^r]_j) = \begin{cases} [[y_{i,k}^r]_j, [y_{i,k}^r]_j + [S_i]_j), & [y_{i,k}^r]_j > 0 \\ (-[S_i]_j, [S_i]_j), & [y_{i,k}^r]_j = 0 \\ ([y_{i,k}^r]_j - [S_i]_j, [y_{i,k}^r]_j], & [y_{i,k}^r]_j < 0. \end{cases} \qquad (7)$$

Remark 2: Note that the perfect measurement $[y_{i,k}^p]_j$ is a Gaussian-like random variable. For the obtained available measurement (5), from (7) one can derive the exact interval that the perfect measurement $[y_{i,k}^p]_j$ falls into, which implies that the soft measurement $[y_{i,k}^s]_j$ has a similar normal-like distribution and lies within the interval $I_{i,j}([y_{i,k}^r]_j)$. That is to say, the perfect measurement $[y_{i,k}^p]_j$ conditional on $I_{i,j}([y_{i,k}^r]_j)$ has a truncated normal-like distribution for the obtained available measurement $[y_{i,k}^r]_j$. Its probability density function is $f(y|I_{i,j}([y_{i,k}^r]_j))$ for $y \in I_{i,j}([y_{i,k}^r]_j)$ and $f(\cdot) \equiv 0$ otherwise, please refer to the proof of Lemma 3 for more details. Fig. 1 gives an illustration of the probability density function of

---

[1]The integer part of a real number r is $\lfloor r \rfloor$ if r is nonnegative, and $\lceil r \rceil$ otherwise, where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor function and the ceiling function, respectively.
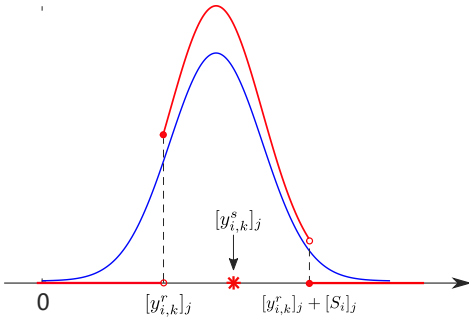
Fig. 1. An illustration of the probability density function for the perfect measurement. The red line and the blue line are, respectively, the probability density functions of $[y^p_{i,k}]_j$ with and without the conditional information concerning $[y^r_{i,k}]_j$.

the perfect measurement. An inspection of Fig. 1 motivates the idea of the proposed soft sensor model (6) with aim of obtaining a better/computational approximation for the perfect measurement.

In this paper, the two-step Kalman-type filtering scheme is utilized to design the distributed filters for the nonlinear time-varying system (1) with imperfect measurement described by (6). Specifically, the structure of the distributed filters is given as follows:

$$\hat{x}^-_{i,k} = F_{k-1}\hat{x}^+_{i,k-1} \quad (8a)$$

$$\hat{x}^+_{i,k} = \hat{x}^-_{i,k} + \sum_{j \in \mathcal{N}_i} a_{ij} K_{ij,k} \left(y^s_{j,k} - H_{j,k}\hat{x}^-_{j,k} \right) \quad (8b)$$

where $\hat{x}^-_{i,k}$ and $\hat{x}^+_{i,k}$ are, respectively, the one-step prediction and the estimate of the system state based on the ith sensor node, and $K_{ij,k}$ stands for the filter gain matrix to be designed. The initial estimate $\hat{x}^+_{i,0}$ $(i = 1, 2, \ldots, n)$ are set to be the expectation of the initial state, i.e., $\hat{x}^+_{i,0} = \bar{x}_0$.

For senor i, define the local one-step prediction error and the local filtering error by $\epsilon^-_{i,k} = x_k - \hat{x}^-_{i,k}$ and $\epsilon^+_{i,k} = x_k - \hat{x}^+_{i,k}$, respectively. Then, we use (1) and (8a) to obtain

$$\epsilon^-_{i,k} = F_{k-1}\epsilon^+_{i,k-1} + \tilde{f}(k-1, x_{k-1}, \eta_{k-1}) + G_{k-1}w_{k-1}. \quad (9)$$

Let $\Delta_{i,k}$ be the difference between the soft measurement and the perfect measurement from sensor i, that is, $\Delta_{i,k} \triangleq y^s_{i,k} - y^p_{i,k}$. The dynamics of the local filtering errors can be described as follows:

$$\epsilon^+_{i,k} = \epsilon^-_{i,k} - \sum_{j \in \mathcal{N}_i} a_{ij} K_{ij,k} \left(H_{j,k}\epsilon^-_{j,k} + v_{j,k} + \Delta_{j,k} \right). \quad (10)$$

For convenience of the later development, we set $\epsilon^-_k \triangleq \text{col}_n\{\epsilon^-_{i,k}\}$ and $\epsilon^+_k \triangleq \text{col}_n\{\epsilon^+_{i,k}\}$. Then, the compact forms of (9) and (10) can be obtained, respectively, as follows:

$$\epsilon^-_k = \mathbf{F}_{k-1}\epsilon^+_{k-1} + \mathbf{f}_{k-1} + \mathbf{G}_{k-1}\mathbf{w}_{k-1} \quad (11)$$

$$\epsilon^+_k = \epsilon^-_k - \sum_{i=1}^{n} \mathbf{I}_i \mathbf{K}_k \mathbf{A}_i \left(\mathbf{H}_k\epsilon^-_k + \mathbf{v}_k + \boldsymbol{\Delta}_k \right) \quad (12)$$

where

$$\mathbf{K}_k = \{K_{ij,k}\}_{n \times n}, \quad \mathbf{G}_{k-1} = I_n \otimes G_{k-1}$$

$$\mathbf{F}_{k-1} = I_n \otimes F_{k-1}, \quad \mathbf{w}_{k-1} = \mathbf{1}_{n \times 1} \otimes w(k-1)$$

$$\mathbf{A}_i = \text{diag}\{a_{i1}I_{n_y}, \ldots, a_{in}I_{n_y}\}, \quad \boldsymbol{\Delta}_k = \text{col}_n\{\Delta_{i,k}\}$$

$$\mathbf{f}_{k-1} = \mathbf{1}_{n \times 1} \otimes \tilde{f}(k-1, x_{k-1}, \eta_{k-1}), \quad \mathbf{v}_k = \text{col}_n\{v_{i,k}\}$$

$$\mathbf{H}_k = \text{diag}_n\{H_{i,k}\}, \quad \mathbf{I}_i = \text{diag}\{\underbrace{0, \ldots, 0}_{i-1}, I_{n_x}, \underbrace{0, \ldots, 0}_{n-i}\}.$$

The main objective of this paper is to design a set of distributed filters with the structure (8), based on the measurements obtained by the sensors with non-logarithmic SR, such that the filtering error covariance is bounded and such a bound is then locally minimized in the trace sense by utilizing a recursive scheme.

III. MAIN RESULTS

In this section, we aim to develop a unified framework to cope with the addressed distributed filtering problem in the presence of non-logarithmic SR and stochastic nonlinearities. The soft measurements of the distributed filters are first obtained to facilitate the filter design. Subsequently, an upper bound of the filtering error covariance is calculated, and appropriate filter gains are designed to guarantee that the obtained upper bound is locally minimized at each time instant.

In the remaining part of this paper, define $X_k \triangleq \mathbb{E}\{x_k x^T_k\}$, $P^-_k \triangleq \mathbb{E}\{\epsilon^-_k (\epsilon^-_k)^T\}$ and $P^+_k \triangleq \mathbb{E}\{\epsilon^+_k (\epsilon^+_k)^T\}$. The following lemma provides a recursive algorithm to calculate the covariance for the one-step prediction error.

Lemma 1: Consider the prediction error dynamic system given by (11). The recursion of the one-step prediction error covariance P^-_k is given as follows:

$$P^-_k = \mathbf{F}_{k-1}P^+_{k-1}\mathbf{F}^T_{k-1} + \boldsymbol{\Gamma}_{k-1} + \mathbf{Q}_{k-1} \quad (13)$$

where $\boldsymbol{\Gamma}_{k-1} = \mathbf{1}_{n \times n} \otimes \sum_{\lambda=1}^{l} \Xi_{\lambda,k-1} \text{Tr}\{X_{k-1}\Psi_{\lambda,k-1}\}$ and $\mathbf{Q}_{k-1} = \mathbf{G}_{k-1}(\mathbf{1}_{n \times n} \otimes Q_{k-1})\mathbf{G}^T_{k-1}$ with X_k satisfying

$$X_{k+1} = F_k X_k F^T_k + \sum_{\lambda=1}^{l} \Xi_{\lambda,k} \text{Tr}\{X_k \Psi_{\lambda,k}\} + G_k Q_k G^T_k. \quad (14)$$

Proof: It is easy to conclude from (3) that

$$\mathbb{E}\{\mathbf{f}_{k-1}\mathbf{f}^T_{k-1}\} = \mathbf{1}_{n \times n} \otimes \sum_{\lambda=1}^{l} \Xi_{\lambda,k-1}\mathbb{E}\{x^T_{k-1}\Psi_{\lambda,k-1}x_{k-1}\}.$$

Noting that $x^T_{k-1}\Psi_{\lambda,k-1}x_{k-1}$ is a scalar, it follows from the property of the trace that $\mathbb{E}\{x^T_{k-1}\Psi_{\lambda,k-1}x_{k-1}\} = \text{Tr}\{X_{k-1}\Psi_{\lambda,k-1}\}$ which, together with (11), yields

$$P^-_k = \mathbf{F}_{k-1}P^+_{k-1}\mathbf{F}^T_{k-1} + \boldsymbol{\Gamma}_{k-1} + \mathbf{Q}_{k-1}$$

which is (13). The proof of (14) follows a similar line as that of (13), and is thus omitted here for the sake of brevity. ∎

Lemma 2: Consider system (1) with initial state x_0. The distributed filters (8) with $\hat{x}^+_{i,0} = \bar{x}_0$ are unbiased, and the corresponding filtering error covariance P^+_k satisfies the following dynamic equation:

$$P^+_k = (I + \mathcal{K}_k \mathbf{H}_k)P^-_k (I + \mathcal{K}_k \mathbf{H}_k)^T$$
$$+ \mathcal{K}_k \mathbb{E}\{\mathbf{v}_k\mathbf{v}^T_k + \boldsymbol{\Delta}_k\boldsymbol{\Delta}^T_k\}\mathcal{K}^T_k$$
$$+ \mathcal{K}_k \mathbb{E}\{\mathbf{v}_k\boldsymbol{\Delta}^T_k + \boldsymbol{\Delta}_k\mathbf{v}^T_k\}\mathcal{K}^T_k \quad (15)$$

where $\mathcal{K}_k = -\sum_{i=1}^{n} \mathbf{I}_i \mathbf{K}_k \mathbf{A}_i$.

Proof: First, let us show the unbiasedness of the distributed filters (8), i.e., $\mathbb{E}\{\epsilon^+_k\} = 0$ for all $k \geq 0$. In view of $\hat{x}^+_{i,0} = \bar{x}_0$, it is easy to verify that $\mathbb{E}\{\epsilon^+_0\} = 0$. Assume, inductively, that this is true for the integers from 0 to $k-1$. By exploiting (12), it is obvious that

$$\epsilon^+_k = (I + \mathcal{K}_k \mathbf{H}_k)\epsilon^-_k + \mathcal{K}_k(\mathbf{v}_k + \boldsymbol{\Delta}_k). \quad (16)$$

Then, it follows from (11) that

$$\epsilon^+_k = (I + \mathcal{K}_k \mathbf{H}_k)\mathbf{F}_{k-1}\epsilon^+_{k-1} + \mathcal{K}_k(\mathbf{v}_k + \boldsymbol{\Delta}_k)$$
$$+ (I + \mathcal{K}_k \mathbf{H}_k)(\mathbf{f}_{k-1} + \mathbf{G}_{k-1}\mathbf{w}_{k-1}). \quad (17)$$

Note that the stochastic nonlinearities and process noise are of zero means. Taking mathematical expectation of both sides of (17)

yields the following recursion (for the mathematical expectation of the filtering error):

$$\mathbb{E}\{\epsilon_k^+\} = (I + \mathcal{K}_k \mathbf{H}_k)\mathbf{F}_{k-1}\mathbb{E}\{\epsilon_{k-1}^+\} + \mathcal{K}_k \mathbb{E}\{\boldsymbol{\Delta}_k\}. \tag{18}$$

On the other hand, it follows from (6) that

$$\begin{aligned}
\mathbb{E}\{[\Delta_{i,k}]_j\} &= \mathbb{E}\big\{\mathbb{E}\{[y_{i,k}^p]_j | I_{i,j}([y_{i,k}^r]_j)\} - [y_{i,k}^p]_j\big\} \\
&= \mathbb{E}\big\{\mathbb{E}\{[y_{i,k}^p]_j | I_{i,j}([y_{i,k}^r]_j)\}\big\} - \mathbb{E}\{[y_{i,k}^p]_j\} \\
&= \mathbb{E}\{[y_{i,k}^p]_j\} - \mathbb{E}\{[y_{i,k}^p]_j\} \\
&= 0
\end{aligned} \tag{19}$$

which, together with the induction hypothesis and (18), implies that $\mathbb{E}\{\epsilon_k^+\} = 0$ holds for all nonnegative integers. Hence, the proposed distributed filters (8) are unbiased.

Next, we are in the position to calculate the covariance matrix P_k^+ of the filtering error. Applying (16), we arrive at

$$\begin{aligned}
P_k^+ =& (I + \mathcal{K}_k \mathbf{H}_k)P_k^-(I + \mathcal{K}_k \mathbf{H}_k)^T + \mathcal{T} + \mathcal{T}^T \\
&+ \mathcal{K}_k \mathbb{E}\{\mathbf{v}_k \mathbf{v}_k^T + \boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^T\}\mathcal{K}_k^T \\
&+ \mathcal{K}_k \mathbb{E}\{\mathbf{v}_k \boldsymbol{\Delta}_k^T + \boldsymbol{\Delta}_k \mathbf{v}_k^T\}\mathcal{K}_k^T
\end{aligned} \tag{20}$$

where $\mathcal{T} = \mathbb{E}\{(I + \mathcal{K}_k \mathbf{H}_k)\epsilon_k^-(\mathbf{v}_k + \boldsymbol{\Delta}_k)^T \mathcal{K}_k^T\}$. Noting that the one-step prediction error ϵ_k^- is uncorrelated with $\mathbf{v}_k$ and $\boldsymbol{\Delta}_k$, one can obtain the recursion (15), which ends the proof. ∎

In the light of (4) and (6), the soft measurement can be further replaced as

$$\begin{aligned}
[y_{i,k}^s]_j &= \mathbb{E}\left\{[H_{i,k}x_k]_j + [v_{i,k}]_j \big| I_{i,j}\left([y_{i,k}^r]_j\right)\right\} \\
&\doteq \mathbb{E}\left\{[\hat{y}_{i,k}^p]_j \big| I_{i,j}\left([y_{i,k}^r]_j\right)\right\}
\end{aligned} \tag{21}$$

where $[\hat{y}_{i,k}^p]_j \triangleq [H_{i,k}\hat{x}_{i,k}^-]_j + [v_{i,k}]_j$.

Remark 3: Note that, when calculating the soft measurement $[y_{i,k}^s]_j$, one needs to know the knowledge of the true state x_k, and this makes the calculation quite difficult. In order to cope with this issue, we replace the real state by its estimate in the calculation of the soft measurement. Actually, such a practice is often employed in the situation that the state is unknown but the state information is still expected, see e.g. [1], [10].

Remark 4: In our developed filter design scheme, the soft measurement model (6) is established to guarantee the *unbiasedness* of the proposed distributed filter (8) (as shown in Lemma 2), which is a desirable property of the filters because it means that, on average, the estimated $\hat{x}_{i,k}^+$ equals to the true value x_k. In the following, we will further show that the distributed filter (8) is *biased* if it is constructed on the basis of the available measurement instead of the soft measurement. Let us first define $\bar{\Delta}_{i,k} \triangleq y_{i,k}^r - y_{i,k}^p$, i.e., the difference between the available measurement and the perfect measurement from sensor i. The local filtering error can be derived as $\epsilon_{i,k}^+ = \epsilon_{i,k}^- - \sum_{j \in \mathcal{N}_i} a_{ij}K_{ij,k}(H_{j,k}\epsilon_{j,k}^- + v_{j,k} + \bar{\Delta}_{j,k})$ and it then follows that $\epsilon_k^+ = \epsilon_k^- - \sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i(\mathbf{H}_k\epsilon_k^- + \mathbf{v}_k + \bar{\boldsymbol{\Delta}}_k)$ where $\bar{\boldsymbol{\Delta}}_k = \text{col}_n\{\bar{\Delta}_{i,k}\}$. Proceeding as in the proof of Lemma 2, the mean of this filtering error can be computed as follows:

$$\mathbb{E}\{\epsilon_k^+\} = (I + \mathcal{K}_k \mathbf{H}_k)\mathbf{F}_{k-1}\mathbb{E}\{\epsilon_{k-1}^+\} + \mathcal{K}_k \mathbb{E}\{\bar{\boldsymbol{\Delta}}_k\}.$$

Next, we fix our attention on the second term in the right-hand side (RHD) of the above equation. Combining (4) and (5), we have

$$\mathbb{E}\{[\bar{\Delta}_{i,k}]_j\} = \mathbb{E}\{Q([y_{i,k}^p]_j) - [y_{i,k}^p]_j\}$$

which is not always identically equal to zero, and this implies that the proposed distributed filter is biased. Therefore, the soft measurement information adopted in the structure of the distributed filter (8) is essential/indispensable.

For the purpose of analysis simplicity, we introduce the following assumption.

Assumption 1: The covariance of the measurement noise $R_{i,k}$ is a diagonal matrix.

Lemma 3: Consider system (1) with measurement described by (5). The soft measurement given by (21) is calculated as follows:

$$[y_{i,k}^s]_j = [H_{i,k}\hat{x}_{i,k}^-]_j - \sqrt{[R_{i,k}]_{j,j}}\frac{\phi(\beta_{i,k}^{(j)}) - \phi(\alpha_{i,k}^{(j)})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})} \tag{22}$$

where $[R_{i,k}]_{j,j}$ is the (j,j)-th entry of $R_{i,k}$, $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the probability density function and the cumulative distribution function of the standard normal distribution, and

$$\beta_{i,k}^{(j)} = \begin{cases} \frac{[y_{i,k}^r]_j + [S_i]_j - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}, & [y_{i,k}^r]_j > 0 \\ \frac{[S_i]_j - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}, & [y_{i,k}^r]_j = 0 \\ \frac{[y_{i,k}^r]_j - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}, & [y_{i,k}^r]_j < 0 \end{cases}$$

$$\alpha_{i,k}^{(j)} = \begin{cases} \frac{[y_{i,k}^r]_j - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}, & [y_{i,k}^r]_j > 0 \\ \frac{-[S_i]_j - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}, & [y_{i,k}^r]_j = 0 \\ \frac{[y_{i,k}^r]_j - [S_i]_j - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}, & [y_{i,k}^r]_j < 0. \end{cases}$$

Proof: For the case $[y_{i,k}^r]_j = 0$, it follows from (7) that

$$I_{i,j}([y_{i,k}^r]_j) = (-[S_i]_j, [S_i]_j).$$

It is easy to see that the random variable $[\hat{y}_{i,k}^p]_j$ obeys a normal distribution $N([H_{i,k}\hat{x}_{i,k}^-]_j, [R_{i,k}]_{j,j})$. Note that the probability that $[\hat{y}_{i,k}^p]_j$ falls into the interval $(-[S_i]_j, [S_i]_j)$ is $\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})$. Then, the conditional probability density function of $[\hat{y}_{i,k}^p]_j$ is

$$f\left(y \big| I_{i,j}([y_{i,k}^r]_j)\right) = \frac{\frac{1}{\sqrt{[R_{i,k}]_{j,j}}}\phi\left(\frac{y - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}\right)}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})} \tag{23}$$

for $-[S_i]_j < y < [S_i]_j$ and $f(\cdot) \equiv 0$ otherwise. The corresponding moment-generating function of $[\hat{y}_{i,k}^p]_j$, denoted by $M(t)$, can be obtained as follows:

$$\begin{aligned}
M(t) &= \mathbb{E}\{\exp\{ty\}|y \in I_{i,j}([y_{i,k}^r]_j)\} \\
&= \int_{-[S_i]_j}^{[S_i]_j} \exp\{ty\}f(y|I_{i,j}([y_{i,k}^r]_j))dy.
\end{aligned} \tag{24}$$

By letting $\mu = [R_{i,k}]_{j,j}t + [H_{i,k}\hat{x}_{i,k}^-]_j$, it is not difficult to verify that (25) (see the bottom of the next page) holds, and it then follows from (24) that

$$\begin{aligned}
M(t) =& \exp\left\{t[H_{i,k}\hat{x}_{i,k}^-]_j + \frac{[R_{i,k}]_{j,j}t^2}{2}\right\} \\
&\times \frac{\Phi(\beta_{i,k}^{(j)} - t\sqrt{[R_{i,k}]_{j,j}}) - \Phi(\alpha_{i,k}^{(j)} - t\sqrt{[R_{i,k}]_{j,j}})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})}.
\end{aligned} \tag{26}$$

By resorting to the properties of moment-generating function, we obtain equality (27) (see the bottom of the next page).

The proof for the case $[y_{i,k}^r]_j \neq 0$ follows the similar lines as above, and is thus omitted here for the sake of brevity. ∎

Based on Lemmas 1-3, the following result provides an upper bound for P_k^+.

Theorem 1: Consider the one-step prediction error covariance P_k^- in (13) and the estimation error covariance P_k^+ in (15). Let γ_k be a positive scalar. If the following difference equations

$$\mathcal{P}_k^- = \mathbf{F}_{k-1}\mathcal{P}_{k-1}^+\mathbf{F}_{k-1}^T + \boldsymbol{\Gamma}_{k-1} + \mathbf{Q}_{k-1} \tag{28}$$

$$\mathcal{P}_k^+ = (I + \mathcal{K}_k \mathbf{H}_k)\mathcal{P}_k^-(I + \mathcal{K}_k \mathbf{H}_k)^T + \mathcal{K}_k \boldsymbol{\Upsilon}_k \mathcal{K}_k^T \tag{29}$$

with initial condition $\mathcal{P}_0^+ = P_0^+$ have positive-definite solutions $\mathcal{P}_k^-$ and $\mathcal{P}_k^+$, then matrix $\mathcal{P}_k^+$ is an upper bound of P_k^+, that is,

$$P_k^+ \le \mathcal{P}_k^+ \tag{30}$$

holds for all $k \ge 0$, where $P_0^+ = \mathbf{1}_{n \times n} \otimes X_0$,

$$\begin{aligned}
\Upsilon_k &\triangleq (1+\gamma_k)\mathrm{diag}_n\{R_{i,k}\} + (1+\gamma_k^{-1})\mathrm{diag}_n\{\Sigma_{i,k}\} \\
\Sigma_{i,k} &\triangleq \mathrm{diag}\{\sigma_{i,k}^{(1)}, \ldots, \sigma_{i,k}^{(n_y)}\} \\
\sigma_{i,k}^{(j)} &\triangleq [R_{i,k}]_{j,j} - [R_{i,k}]_{j,j}\frac{\beta_{i,k}^{(j)}\phi(\beta_{i,k}^{(j)}) - \alpha_{i,k}^{(j)}\phi(\alpha_{i,k}^{(j)})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})} \\
&\quad - [R_{i,k}]_{j,j}\Big[\frac{\phi(\beta_{i,k}^{(j)}) - \phi(\alpha_{i,k}^{(j)})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})}\Big]^2.
\end{aligned} \tag{31}$$

Proof: This theorem is proved by the mathematical induction. The conclusion follows immediately for the case $k = 0$ based on the initial condition. For the induction step, we assume inductively that $P_{k-1}^+ \le \mathcal{P}_{k-1}^+$. Combining (13) and (28), one obtains that

$$P_k^- - \mathcal{P}_k^- = \mathbf{F}_{k-1}(P_{k-1}^+ - \mathcal{P}_{k-1}^+)\mathbf{F}_{k-1}^T \le 0$$

which implies $P_k^- \le \mathcal{P}_k^-$. By resorting to the elementary inequality

$$\mathbf{v}_k\boldsymbol{\Delta}_k^T + \boldsymbol{\Delta}_k\mathbf{v}_k^T \le \gamma_k\mathbf{v}_k\mathbf{v}_k^T + \gamma_k^{-1}\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^T,$$

it follows readily from (15) that

$$\begin{aligned}
P_k^+ \le & (I + \mathcal{K}_k\mathbf{H}_k)P_k^-(I + \mathcal{K}_k\mathbf{H}_k)^T + (1+\gamma_k)\mathcal{K}_k\mathrm{diag}_n\{R_{i,k}\} \\
& \times \mathcal{K}_k^T + (1+\gamma_k^{-1})\mathcal{K}_k\mathbb{E}\{\boldsymbol{\Delta}_k\boldsymbol{\Delta}_k^T\}\mathcal{K}_k^T.
\end{aligned} \tag{32}$$

Now, we are in the position to cope with the third term in the RHD of the above equation, which requires the computation of $\mathbb{E}\{[\Delta_{i,k}]_j^2\}$. From the definition of $\Delta_{i,k}$, noticing the approximation of $y_{i,k}^s$ in (21), we have

$$\begin{aligned}
\mathbb{E}\{[\Delta_{i,k}]_j^2\} &= \mathbb{E}\Big\{\big([y_{i,k}^p]_j - \mathbb{E}\{[y_{i,k}^p]_j | I_{i,j}([y_{i,k}^r]_j)\}\big)^2\Big\} \\
&= \mathbb{E}\{[\hat{y}_{i,k}^p]_j^2 | I_{i,j}([y_{i,k}^r]_j)\} - [y_{i,k}^s]_j^2.
\end{aligned} \tag{33}$$

With the help of the moment-generating function (24), the first term in the RHD of (33) is calculated as

$$\begin{aligned}
\mathbb{E}\{[\hat{y}_{i,k}^p]_j^2 | I_{i,j}([y_{i,k}^r]_j)\} &= \frac{d^2 M(t)}{dt^2}\Big|_{t=0} \\
&= [R_{i,k}]_{j,j} + [H_{i,k}\hat{x}_{i,k}^-]_j^2 + [R_{i,k}]_{j,j}\frac{\phi'(\beta_{i,k}^{(j)}) - \phi'(\alpha_{i,k}^{(j)})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})} \\
&\quad - 2[H_{i,k}\hat{x}_{i,k}^-]_j\sqrt{[R_{i,k}]_{j,j}}\frac{\phi(\beta_{i,k}^{(j)}) - \phi(\alpha_{i,k}^{(j)})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})}.
\end{aligned} \tag{34}$$

In light of the fact that $\phi'(x) = -x\phi(x)$, it follows from (32) that

$$\begin{aligned}
P_k^+ \le & (I + \mathcal{K}_k\mathbf{H}_k)P_k^-(I + \mathcal{K}_k\mathbf{H}_k)^T + \mathcal{K}_k\Upsilon_k\mathcal{K}_k^T \\
\le & (I + \mathcal{K}_k\mathbf{H}_k)\mathcal{P}_k^-(I + \mathcal{K}_k\mathbf{H}_k)^T + \mathcal{K}_k\Upsilon_k\mathcal{K}_k^T = \mathcal{P}_k^+.
\end{aligned}$$

By induction, one concludes that (30) holds for all positive integers. The proof is complete. ∎

In what follows, we will design the filter parameters which are locally optimal in the sense that the trace of $\mathcal{P}_k^+$ is minimized. Since $\mathcal{P}_k^-$ is a square matrix with dimension $n_x n \times n_x n$, it can be partitioned into $n \times n$ submatrices with dimension $n_x \times n_x$. The partitioned matrix can be written as $\mathcal{P}_k^- = \{\mathcal{P}_{ij,k}^-\}_{n \times n}$. Let $\mathbf{K}_k(i)$ and $\mathcal{P}_k^-(i)$ be, respectively, the ith row of the partitioned matrix $\mathbf{K}_k$ and $\mathcal{P}_k^-$, that is, $\mathbf{K}_k(i) \triangleq [K_{i1,k}\ K_{i2,k}\ \cdots\ K_{in,k}]$ and $\mathcal{P}_k^-(i) \triangleq [\mathcal{P}_{i1,k}^-\ \mathcal{P}_{i2,k}^-\ \cdots\ \mathcal{P}_{in,k}^-]$.

For simplicity of presentation, we denote $\mathbf{U}_k^{(i)} \triangleq \mathbf{A}_i(\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T + \Upsilon_k)\mathbf{A}_i$ and $\mathbf{V}_k^{(i)} \triangleq \mathcal{P}_k^-(i)\mathbf{H}_k^T\mathbf{A}_i$. Subsequently, one can partition the above two matrices by $\mathbf{U}_k^{(i)} = \{U_{jl,k}^{(i)}\}_{n \times n}$ and $\mathbf{V}_k^{(i)} = \{V_{l,k}^{(i)}\}_{1 \times n}$. Before proceeding further, define $\bar{\mathbf{K}}_k(i)$ (respectively, $\bar{\mathbf{V}}_k^{(i)}$) to be the matrix derived from $\mathbf{K}_k(i)$ (respectively, $\mathbf{V}_k^{(i)}$) by deleting the column blocks with indexes $l \notin \mathcal{N}_i$, and $\bar{\mathbf{U}}_k^{(i)}$ to be the simplified matrix derived from $\mathbf{U}_k^{(i)}$ by deleting both the row and the column blocks with indexes $l \notin \mathcal{N}_i$.

Theorem 2: Consider the time-varying system (1) with real measurement (5) and distributed filters (8). The trace of the upper bound $\mathcal{P}_k^+$ is locally minimized with filter gain $\mathbf{K}_k = \{K_{ij,k}\}_{n \times n}$ designed as $K_{ij,k} = 0$ for $j \notin \mathcal{N}_i$ and

$$\bar{\mathbf{K}}_k(i) = \bar{\mathbf{V}}_k^{(i)}(\bar{\mathbf{U}}_k^{(i)})^{-1} \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{35}$$

$$\begin{aligned}
\frac{1}{\sqrt{2\pi[R_{i,k}]_{j,j}}} &\int_{-[S_i]_j}^{[S_i]_j} \exp\{ty\}\exp\Big\{ - \frac{1}{2}\Big(\frac{y - [H_{i,k}\hat{x}_{i,k}^-]_j}{\sqrt{[R_{i,k}]_{j,j}}}\Big)^2\Big\}dy \\
&= \exp\Big\{ - \frac{1}{2[R_{i,k}]_{j,j}}\big([H_{i,k}\hat{x}_{i,k}^-]_j^2 - \mu^2\big)\Big\}\frac{1}{\sqrt{2\pi[R_{i,k}]_{j,j}}}\int_{-[S_i]_j}^{[S_i]_j}\exp\Big\{ - \frac{(y-\mu)^2}{2[R_{i,k}]_{j,j}}\Big\}dy \\
&= \exp\Big\{t[H_{i,k}\hat{x}_{i,k}^-]_j + \frac{[R_{i,k}]_{j,j}t^2}{2}\Big\}\Big[\Phi\Big(\frac{[S_i]_j - \mu}{\sqrt{[R_{i,k}]_{j,j}}}\Big) - \Phi\Big(\frac{-[S_i]_j - \mu}{\sqrt{[R_{i,k}]_{j,j}}}\Big)\Big]
\end{aligned} \tag{25}$$

$$\begin{aligned}
[y_{i,k}^s]_j =& \frac{dM(t)}{dt}\Big|_{t=0} \\
=& \exp\Big\{t[H_{i,k}\hat{x}_{i,k}^-]_j + \frac{[R_{i,k}]_{j,j}t^2}{2}\Big\}\Big[\big([H_{i,k}\hat{x}_{i,k}^-]_j + t[R_{i,k}]_{j,j}\big)\frac{\Phi(\beta_{i,k}^{(j)} - t\sqrt{[R_{i,k}]_{j,j}}) - \Phi(\alpha_{i,k}^{(j)} - t\sqrt{[R_{i,k}]_{j,j}})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})} \\
& - \frac{\sqrt{[R_{i,k}]_{j,j}}\phi(\beta_{i,k}^{(j)} - t\sqrt{[R_{i,k}]_{j,j}}) - \sqrt{[R_{i,k}]_{j,j}}\phi(\alpha_{i,k}^{(j)} - t\sqrt{[R_{i,k}]_{j,j}})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})}\Big]\Big|_{t=0} \\
=& [H_{i,k}\hat{x}_{i,k}^-]_j - \sqrt{[R_{i,k}]_{j,j}}\frac{\phi(\beta_{i,k}^{(j)}) - \phi(\alpha_{i,k}^{(j)})}{\Phi(\beta_{i,k}^{(j)}) - \Phi(\alpha_{i,k}^{(j)})}
\end{aligned} \tag{27}$$

Proof: Taking the trace of both sides of (29) yields

$$\text{Tr}\{\mathcal{P}_k^+\} = \text{Tr}\{\mathcal{P}_k^-\} + \text{Tr}\{\mathcal{K}_k\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T\mathcal{K}_k^T\}$$
$$+ 2\text{Tr}\{\mathcal{K}_k\mathbf{H}_k\mathcal{P}_k^-\} + \text{Tr}\{\mathcal{K}_k\Upsilon_k\mathcal{K}_k^T\}. \quad (36)$$

The second term in the RHD of the above equation can be further represented by

$$\text{Tr}\{\mathcal{K}_k\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T\mathcal{K}_k^T\}$$
$$= \text{Tr}\Big\{\Big(\sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i\Big)\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T\Big(\sum_{j=1}^n \mathbf{A}_j\mathbf{K}_k^T\mathbf{I}_j\Big)\Big\}$$
$$= \text{Tr}\Big\{\sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T\mathbf{A}_i\mathbf{K}_k^T\mathbf{I}_i\Big\} \quad (37)$$

where the last equality follows from

$$\text{Tr}\{\mathbf{I}_i\mathbf{K}_k\mathbf{A}_i\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T\mathbf{A}_j\mathbf{K}_k^T\mathbf{I}_j\} = 0, \quad \forall\ i \neq j.$$

Similarly, the last term in the RHD of (36) can be rewritten as

$$\text{Tr}\{\mathcal{K}_k\Upsilon_k\mathcal{K}_k^T\} = \text{Tr}\Big\{\sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i\Upsilon_k\mathbf{A}_i\mathbf{K}_k^T\mathbf{I}_i\Big\}.$$

Taking the partial derivative of $\text{Tr}\{\mathcal{P}_k^+\}$ with respect to $\mathbf{K}_k$, one obtains the following equality

$$\frac{\partial\text{Tr}\{\mathcal{P}_k^+\}}{\partial\mathbf{K}_k} = 2\sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T\mathbf{A}_i - 2\sum_{i=1}^n \mathbf{I}_i\mathcal{P}_k^-\mathbf{H}_k^T\mathbf{A}_i$$
$$+ 2\sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i\Upsilon_k\mathbf{A}_i. \quad (38)$$

In order to minimize the trace of matrix $\mathcal{P}_k^+$, the partial derivative (38) is set as zero, which immediately yields

$$\sum_{i=1}^n \mathbf{I}_i\mathbf{K}_k\mathbf{A}_i(\mathbf{H}_k\mathcal{P}_k^-\mathbf{H}_k^T + \Upsilon_k)\mathbf{A}_i = \sum_{i=1}^n \mathbf{I}_i\mathcal{P}_k^-\mathbf{H}_k^T\mathbf{A}_i.$$

Exploiting the inherent special structure of matrix $\mathbf{I}_i$, we have the following equations:

$$\mathbf{K}_k(i)\mathbf{U}_k^{(i)} = \mathbf{V}_k^{(i)}, \quad i = 1, 2, \ldots, n \quad (39)$$

which give rise to $\sum_{j=1}^n K_{ij,k}U_{jl,k}^{(i)} = V_{l,k}^{(i)}$ for $l = 1, 2, \ldots, n$. Note that $\mathbf{A}_i = \text{diag}\{a_{i1}I, \ldots, a_{in}I\}$, where the entry $a_{il} = 0$ when $l \notin \mathcal{N}_i$. From the definitions of $\mathbf{U}_k^{(i)}$ and $\mathbf{V}_k^{(i)}$, it is easy to see that $U_{jl,k}^{(i)} = 0$ and $V_{l,k}^{(i)} = 0$ hold for $l \notin \mathcal{N}_i$. In this case, the corresponding filter gains can be designed as $K_{il,k} = 0$ because the ith sensor node cannot receive any information from its non-neighbor senor l. Consequently, it is apparent from (39) that

$$\bar{\mathbf{K}}_k(i)\bar{\mathbf{U}}_k^{(i)} = \bar{\mathbf{V}}_k^{(i)}, \quad i = 1, 2, \ldots, n$$

which, together with the fact that matrix $\bar{\mathbf{U}}_k^i$ is full-row rank, yields $\bar{\mathbf{K}}_k(i) = \bar{\mathbf{V}}_k^{(i)}(\bar{\mathbf{U}}_k^{(i)})^{-1}$, $i = 1, 2, \ldots, n$. Therefore, the filter gain $K_{ij,k}$ can be designed as in (35). This accomplishes the proof of Theorem 2. ∎

Remark 5: It is worth pointing out that the distributed recursive filtering algorithm proposed in this paper is an offline algorithm, which means that the filter gains in (35) are firstly computed iteratively offline and then implemented/installed on practical plants for application. On the other hand, the design of the gains in (35) depends on the output matrix $\mathbf{H}_k = \text{diag}_n\{H_{i,k}\}$ of all agents and hence needs the global information of the sensor network when conducting the offline computations.

Remark 6: It is worth mentioning that, due to the influence from the SR and the stochastic nonlinearity, the parameters in the distributed filter (8) are skillfully designed by minimizing the trace of
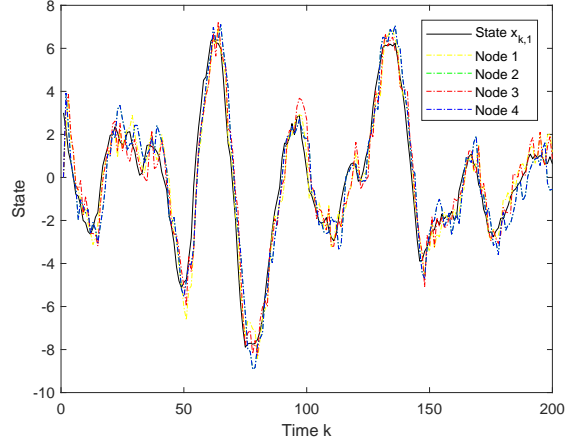

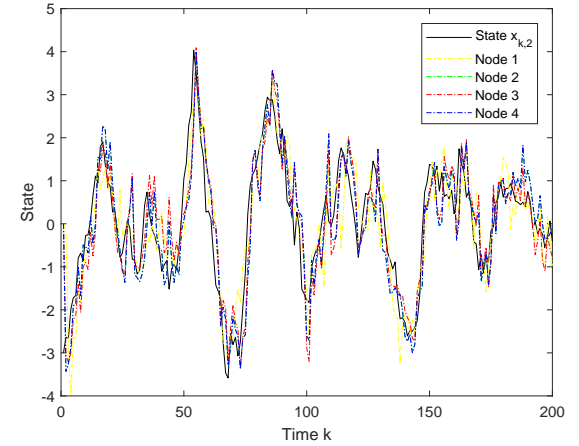Fig. 2. The true state $x_{k,1}$ and its estimate $[\hat{x}_{i,k}^+]_1$ $(i = 1, 2, 3, 4)$.

Fig. 3. The true state $x_{k,2}$ and its estimate $[\hat{x}_{i,k}^+]_2$ $(i = 1, 2, 3, 4)$.

the upper bound for the estimation error covariance. As such, the constructed distributed filter is effectively a locally optimal one, and the possible conservatism stems from the difference between the real estimation error covariance P_k^+ and its upper bound $\mathcal{P}_k^+$. According to (32), it is apparent that the conservatism comes from the utilization of the elementary inequality. One of our future research topics is to further reduce such conservatism in order to improve the performance of our designed locally optimal filter.

IV. AN ILLUSTRATIVE EXAMPLE

In this section, a simulation example is employed to illustrate the validity and applicability of the developed distributed filtering algorithm. Consider a sensor network with four nodes, where the communication topology is described by a directed graph $\mathcal{G}$ with the set of nodes $\mathcal{V} = \{1, 2, 3, 4\}$ and the adjacency matrix

$$\mathcal{H} = \begin{bmatrix} 1 & 0 & 0.9 & 0.6 \\ 0.4 & 1 & 0 & 0.7 \\ 0 & 0.7 & 1 & 0.5 \\ 0.3 & 0.7 & 0 & 1 \end{bmatrix}.$$

The target plant is described by (1) where $x_k = [x_{k,1}\ x_{k,2}]^T \in \mathbb{R}^2$, the initial state x_0 and the process noise obey the zero-mean Gaussian
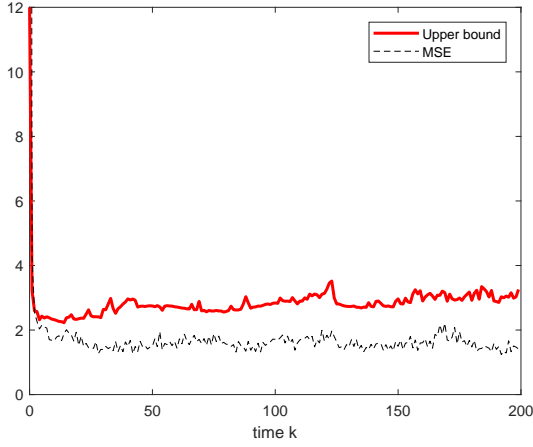
Fig. 4. The MSE and its upper bound $\mathrm{Tr}\{\mathcal{P}_k^+\}$.

distribution with respective covariance $X_0 = \mathrm{diag}\{1.5, 1.5\}$ and $Q_k = \mathrm{diag}\{0.2, 0.1\}$, and

$$F_k = \begin{bmatrix} 1.02 + 0.03\cos(0.12k) & 0.4 \\ -0.1 & 0.88 \end{bmatrix}, \quad G_k = \mathrm{diag}\{1, 1\}.$$

The nonlinearity $f(k, x_k, \eta_k)$ is set as

$$f(k, x_k, \eta_k) = \begin{bmatrix} 0.3 \\ 0.15 \end{bmatrix} \big(0.6[\eta_k]_1 [x_k]_1 \mathrm{sign}([x_k]_1) \\ + 0.8[\eta_k]_2 [x_k]_2 \mathrm{sign}([x_k]_2) \big)$$

where $\eta_k \in \mathbb{R}^2$ represents the zero-mean Gaussian white noise with unitary covariance. It is obvious to confirm that $\mathbb{E}\{f(k, x_k, \eta_k)\} = 0$ and $\mathbb{E}\{f(k, x_k, \eta_k)f^T(k, x_k, \eta_k)|x_k\} = \Xi_k x_k^T \Psi_k x_k$ with $\Xi_k = [0.3\,0.15][0.3\,0.15]^T$ and $\Psi_k = \mathrm{diag}\{0.36, 0.64\}$. The dynamics of the target plant is monitored by sensors with $S_i = 1$ ($i = 1, 2, 3, 4$) and the following measurement parameters:

$$H_{1,k} = [0.85 \quad 0.38 + 0.12\sin(0.12k)], \quad H_{2,k} = [0.5 \quad 1.2]$$
$$H_{3,k} = [0.98 + 0.04\sin(0.1k) \quad 0.435]$$
$$H_{4,k} = [0.75 + 0.03\sin(0.1k) \quad 0.435 + 0.03\sin(0.1k)].$$

In the simulation, the measurement noises $v_{i,k}$ ($i = 1, 2, 3, 4$) are mutually uncorrelated zero-means Gaussian white sequences with $R_{i,k} \equiv 0.4$. Figs. 2-3 display the trajectories of the true states $[x_k]_j$ ($j = 1, 2$) and their corresponding estimates $[\hat{x}_{i,k}^+]_j$ ($i = 1, 2, 3, 4$), which confirm that the proposed filter scheme can perform well to estimate the system states. This is mainly due to the soft measurement technique we have adapted in coping with the SR. In order to quantify the estimation accuracy, the mean square estimation error is defined as $\mathrm{MSE}(k) \triangleq \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\|\epsilon_{i,k}^+\|_2^2$. Fig. 4 shows the upper bound $\mathrm{Tr}\{\mathcal{P}_k^+\}$ as well as the MSE derive from $T = 100$ independent experiments, which illustrates that the MSE stays below its upper bound.

V. CONCLUSION

In this paper, the distributed filtering problem has been addressed for a class of discrete time-varying stochastic systems subject to non-logarithmic SR and stochastic nonlinearities. A soft measurement technique has been exploited to deal with the effect resulting from the SR-induced uncertainty. The upper bound for the filtering error covariance has been calculated by solving certain Riccati-like difference equations. After that, the distributed filter parameters have been

appropriately designed to iteratively minimize such an upper bound in the trace sense. Finally, a numerical example has been provided to illustrate the effectiveness of the developed filtering strategy. One of the future research topics is to extend the main results of this paper to more general systems with more complicated network-induced phenomena [2], [6], [32], [33].

REFERENCES

[1] B. Allik, C. Miller, M. J. Piovoso, and R. Zurakowski, The Tobit Kalman filter: An estimator for censored measurements, *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 1, pp. 365–371, Jan. 2016.

[2] M. V. Basin, A. G. Loukianov, and M. Hernandez-Gonzalez, Joint state and parameter estimation for uncertain stochastic nonlinear polynomial systems, *Int. J. Syst. Sci.*, vol. 44, no. 7, pp. 1200–1208, Jul. 2013.

[3] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, Consensus-based linear and nonlinear filtering, *IEEE Trans. Autom. Control*, vol. 60, no. 5, pp. 1410–1415, May 2015.

[4] H. A. P. Blom and E. A. Bloem, Exact Bayesian filter and joint IMM coupled PDA tracking of maneuvering targets from possibly missing and false measurements, *Automatica*, vol. 42, no. 1, pp. 127–135, Jan. 2006.

[5] H. A. P. Blom and E. A. Bloem, Bayesian tracking of two possibly unresolved maneuvering targets, *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 2, pp. 612–627, Apr. 2007.

[6] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, Distributed fusion filters from uncertain measured outputs in sensor networks with random packet losses, *Inf. Fusion*, vol. 34, pp. 70–79, Mar. 2017.

[7] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, Distributed Kalman filtering based on consensus strategies, *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 622–633, May 2008.

[8] K.-C. Chang and Y. Bar-Shalom, Joint probabilistic data association for multitarget tracking with possibly unresolved measurements and maneuvers, *IEEE Trans. Autom. Control*, vol. 29, no. 7, pp. 585–594, Jul. 1984.

[9] H. Dong, Z. Wang, and H. Gao, Distributed H_∞ filtering for a class of Markovian jump nonlinear time-delay systems over lossy sensor networks, *IEEE Trans. Ind. Electron.*, vol. 60, no. 10, pp. 4665–4672, Oct. 2013.

[10] H. Geng, Z. Wang, L. Zou, A. Mousavi and Y. Cheng, Protocol-based Tobit Kalman filter under integral measurements and probabilistic sensor failures, *IEEE Trans. Signal Process.*, vol. 69, pp. 546–559, 2021.

[11] Q.-L. Han, Absolute stability of time-delay systems with sector-bounded nonlinearity, *Automatica*, vol. 41, no. 12, pp. 2171–2176, Dec. 2005.

[12] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements, *Automatica*, vol. 48, no. 9, pp. 2007–2015, Sep. 2012.

[13] J. Hu, H. Zhang, H. Liu and X. Yu, A survey on sliding mode control for networked control systems, *Int. J. Syst. Sci.*, vol. 52, no. 6, pp. 1129–1147, 2021.

[14] D. H. Jacobson, A general result in stochastic optimal control of nonlinear discrete-time systems with quadratic performance criteria, *J. Math. Anal. Appl.*, vol. 47, no. 1, pp. 153–161, Jul. 1974.

[15] H. J. Kushner, Numerical algorithms for optimal controls for nonlinear stochastic systems with delays, *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2170–2176, Sep. 2010.

[16] Q. Liu, Z. Wang, X. He, and D. H. Zhou, Event-based recursive distributed filtering over wireless sensor networks, *IEEE Trans. Autom. Control*, vol. 60, no. 9, pp. 2470–2475, Sep. 2015.

[17] J. Mao, J. Sun, X. Yi, H. Liu and D. Ding, Recursive filtering of networked nonlinear systems: A survey, *Int. J. Syst. Sci.*, vol. 52, no. 6, pp. 1110–1128, 2021.

[18] R. Olfati-Saber, Distributed Kalman filtering for sensor networks, in *Proc. 46th IEEE Conf. Decis. Control*, New Orleans, LA, Dec. 2007, pp. 5492–5498.

[19] R. Olfati-Saber, Distributed Kalman filter with embedded consensus filters, in *Proc. 44th IEEE Conf. Decis. Control, Eur. Control Conf.*, Seville, Spain, Dec. 2005, pp. 8179–8184.

[20] R. Olfati-Saber and J. Shamma, Consensus filters for sensor networks and distributed sensor fusion, in *Proc. 44th IEEE Conf. Decis. Control, Eur. Control Conf.*, Seville, Spain, Dec. 2005, pp. 6698–6703.

[21] B. Shen, Z. Wang, H. Shu, and G. Wei, Robust H_∞ finite-horizon filtering with randomly occurred nonlinearities and quantization effects, *Automatica*, vol. 46, no. 11, pp. 1743–1751, Nov. 2010.

[22] H. Tan, B. Shen, K. Peng and H. Liu, Robust recursive filtering for uncertain stochastic systems with amplify-and-forward relays, *Int. J. Syst. Sci.*, vol. 51, no. 7, pp. 1188–1199, 2020.

[23] D. Svensson, M. Ulmke, and L. Hammarstrand, Multitarget sensor resolution model and joint probabilistic data association, *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3418–3434, Oct. 2002.

[24] Z. Wang, B. Shen, H. Shu, and G. Wei, Quantized H_∞ control for nonlinear stochastic time-delay systems with missing measurements, *IEEE Trans. Autom. Control*, vol. 57, no. 6, pp. 1431–1444, Jun. 2012.

[25] Y. I. Yaz and E. E. Yaz, On LMI formulations of some problems arising in nonlinear stochastic system analysis, *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 813–816, Apr. 1999.

[26] E. E. Yaz and Y. I. Yaz, State estimation of uncertain nonlinear stochastic systems with general criteria, *Appl. Math. Lett.*, vol. 14, no. 5, pp. 605–610, Jul. 2001.

[27] J. Yick, B. Mukherjee, and D. Ghosal, Wireless sensor network survey, *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.

[28] D. Yue and Q.-L. Han, Delay-dependent exponential stability of stochastic systems with time-varying delay, nonlinearity, and Markovian switching, *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 217–222, Feb. 2005.

[29] J. Zhang, X. He, and D. Zhou, Filtering for stochastic uncertain systems with non-logarithmic sensor resolution, *Automatica*, vol. 89, pp. 194–200, Mar. 2018.

[30] J. Zhang, X. He, and D. Zhou, Distributed filtering over wireless sensor networks with parameter and topology uncertainties, *Int. J. Control*, vol. 93, no. 4, pp. 910–921, Apr. 2020.

[31] Z. Zhao, Z. Wang, L. Zou and J. Guo, Set-membership filtering for time-varying complex networks with uniform quantisations over randomly delayed redundant channels, *Int. J. Syst. Sci.*, vol. 51, no. 16, pp. 3364–3377, 2020.

[32] K. Zhu, J. Hu, Y. Liu, N. D. Alotaibi and F. E. Alsaadi, On ℓ_2-ℓ_∞ output-feedback control scheduled by stochastic communication protocol for two-dimensional switched systems, *Int. J. Syst. Sci.*, in press, DOI: 10.1080/00207721.2021.1914768.

[33] L. Zou, Z. Wang, J. Hu, Y. Liu and X. Liu, Communication-protocol-based analysis and synthesis of networked systems: Progress, prospects and challenges, *Int. J. Syst. Sci.*, in press, DOI: 10.1080/00207721.2021.1917721.