Minimum-Variance Recursive Filtering over Sensor Networks with Stochastic Sensor Gain Degradation: Algorithms and Performance Analysis

Yang Liu, Zidong Wang, Xiao He and D. H. Zhou*

***Abstract*—This paper is concerned with the minimum variance filtering problem for a class of time-varying systems with both additive and multiplicative stochastic noises through a sensor network with a given topology. The measurements collected via the sensor network are subject to stochastic sensor gain degradation, and the gain degradation phenomenon for each individual sensor occurs in a random way governed by a random variable distributed over the interval $[0,1]$. The purpose of the addressed problem is to design a distributed filter for each sensor such that the overall estimation error variance is minimized at each time step via a novel recursive algorithm. By solving a set of Riccati-like matrix equations, the parameters of the desired filters are calculated recursively. The performance of the designed filters is analyzed in terms of both the boundedness and monotonicity. Specifically, sufficient conditions are obtained under which the estimation error is exponentially bounded in mean square. Moreover, the monotonicity property for the error variance with respect to the sensor gain degradation is thoroughly discussed. Numerical simulations are exploited to illustrate the effectiveness of the proposed filtering algorithm and the performance of the developed filter.**

***Index Terms*—Minimum variance filtering; sensor network; recursive algorithm; sensor gain; error boundedness; monotonicity.**

I. Introduction

In the past decade, considerable research efforts have been devoted to sensor networks due to their extensive applications in many fields such as information collection, environmental monitoring, industrial automation and intelligent buildings. Each sensor node has wireless communication capability and some level of intelligence for signal processing. The sensor nodes are usually spatially distributed and coordinated to perform some global tasks by exchanging information with the neighboring nodes. An attractive research focus in relation to sensor networks is the so-called distributed filtering problem whose main idea is to estimate the dynamics of the target plant based on the distributed nodes. Compared to the traditional single sensor leading to the traditional central filtering approaches, each sensor in sensor networks estimates the states of the dynamic process based on not only its own measurement but also the measurements from its neighboring nodes. As such, one of the main difficulties in the distributed filter design problem over a sensor network with given topology is how to take the topology information into account by tackling the complicated couplings between the nodes.

So far, the distributed filtering problem has been gaining an increasing research interest and a wealth of literature has been reported in this topic [2], [6]–[8], [10], [11], [14], [15], [17], [18], [27], [30], [32]. For example, the consensus strategy has been applied in Distributed Kalman Filters (DKFs) [5], [12], [22] that allow the nodes in a sensor network to track the average of the sensor measurements based on consensus filters. Communication complexity and packet-loss issues have been discussed for the performance analysis for DKFs in [21]. The optimal distributed filter has been proposed to minimize the filtering error variance [3], [26]. In these optimal filters, the parameters of estimators have been adjusted at each time step to achieve the minimum mean-square estimation error based on the received signals. The filtering algorithm guaranteeing the desired H_∞ performance has been put forward in [34]. Note that, in most of the reported results, the target plants have been limited to the *time-invariant* systems, where the filter performances (e.g. boundedness and monotonicity) have not been investigated in a quantitative way. On the other hand, a great number of filtering approaches have been proposed to achieve the optimal statistical performance by means of minimum variance. Therefore, there is a practical need to address the distributed filtering problem for *time-varying* systems in sensor networks with detailed analysis on the filter performances including *minimum variance, boundedness and monotonicity*.

In many practical applications, the phenomenon of sensor gain degradation occurs frequently in a random way. This is particularly true for systems which experience unsteady or abnormal working conditions [19], [20], [25], [28], [31], for example, intermittent sensor outages, sensor aging or transmission congestions in networked environments. Note that the filtering problem for systems whose sensor gains are subject to random degradation has received some initial research attention [9]. Unfortunately, despite its practical significance,

This work was supported by the National Natural Science Foundation of China under Grants 61490701, 61210012, 61290324, 61473163, and 61273156, Tsinghua University Initiative Scientific Research Program, and Jiangsu Provincial Key Laboratory of E-business at Nanjing University of Finance and Economics of China under Grant JSEB201301.

Y. Liu and X. He are with the Department of Automation, TNList, Tsinghua University, Beijing 100084, P. R. China.

Z. Wang is with the College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, P. R. China. He is also with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom (email: Zidong.Wang@brunel.ac.uk).

D. H. Zhou is with the College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, P. R. China. He is also with the Department of Automation, TNList, Tsinghua University, Beijing 100084, P. R. China (email: zdh@mail.tsinghua.edu.cn).

* Corresponding author.

the filtering problem with stochastic sensor gain degradation over sensor networks has not been investigated yet for time-varying systems due mainly to the mathematical difficulties, not to mention the case where the filter performance becomes a concern in the design. Those filter structures and design methods in existing literature (e.g. [4], [16], [33]) are not directly applicable to the addressed filtering problem with multiplicative disturbances in the minimum variance sense. *The resulting difficulties stem from the facts that: 1) it is challenging to design an adequate filter gain structure in order to guarantee the minimum error variance at each time step over a sensor network if its topology is not completely connected (i.e. sparse); 2) it is novel to examine how the filter performance is influenced by the statistical law of the sensor gain degradation in a mathematically rigorous way (i.e. monotonicity); 3) it is interesting to establish sufficient conditions under which the estimation error is exponentially bounded in mean square; and 4) it is non-trivial to include the statistical information on the sensor gain degradation in the filter design*. It is, therefore, the main purpose of this paper to handle the challenges mentioned above by launching a major study on algorithm design and performance analysis issues for recursive filter design problems over possibly sparse sensor networks.

In this paper, the minimum variance filtering problem is addressed for a class of time-varying systems through sensor networks with stochastic sensor gain degradation. The gain degradation is allowed to be different for individual sensor. The topology of the sensor network is represented by a directed graph. The minimum-variance distributed filter is designed at each time step using a novel recursive algorithm. The corresponding filtering performances are then analyzed with respect to the boundedness and the monotonicity. Sufficient conditions are obtained under which the estimation error is exponentially bounded in mean square, and the monotonicity property for the error variance with respect to the sensor gain degradation is discussed in the case where all the sensor gains are subject to degradation with the same possibility. Some simulation examples are employed to show the effectiveness of the proposed filtering scheme. *The main contributions of the paper are outlined as follows: 1) a distributed filter is designed that minimizes the filtering error variance in the presence of stochastic sensor gain degradation; 2) the developed filter caters for time-varying systems in sensor networks, and the algorithm is recursive and thus applicable for online computation; and 3) the estimation performance is investigated, including the analysis of the boundedness and monotonicity of the filtering error dynamics*.

Notations. The notation used in the paper is fairly standard except where otherwise stated. $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote, respectively, the n-dimensional Euclidean space and the set of all $n\times m$ real matrices. The superscript "T" denotes the transpose and the notation $X \geq Y$ (respectively, $X > Y$) where X and Y are symmetric matrices, means that $X - Y$ is positive semidefinite (respectively, positive definite). I is the identity matrix with compatible dimension. $M^{\dagger} \in \mathbb{R}^{n\times m}$ denotes the Moore-Penrose pseudo inverse of $M \in \mathbb{R}^{m\times n}$. $\mathbb{E}\{x\}$ stands for the expectation of the stochastic variable x. $\|A\|$ denotes the spectral norm of matrix A, and $\|x\|$ refers to the Euclidean norm of vector x. $\mathrm{diag}\{\cdots\}$ stands for a block-diagonal matrix. $\mathrm{diag}_n\{\star\}$ and $\mathrm{vec}_n\{\bullet\}$ are employed to represent a block-diagonal matrix and a row vector, respectively, whose entries are all $\star$ and $\bullet$.

II. Problem Formulation and Preliminaries

Consider a sensor network whose topology is represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ of order n with the set of nodes $\mathcal{V} = \{1, 2, \dots, n\}$, set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, and a weighted adjacency matrix $\mathcal{A} = [a_{ij}]$ with nonnegative adjacency elements a_{ij}. An edge of $\mathcal{G}$ is denoted by (i, j). The adjacency elements associated with the edges of the graph are positive, i.e., $a_{ij} > 0 \iff (i, j) \in \mathcal{E}$. Moreover, we assume $a_{ii} = 1$ for all $i \in \mathcal{V}$. The set of neighbors of node i plus the node itself are denoted by $\mathcal{N}_i = \{j \in \mathcal{V} : a_{ij} > 0\}$. A communication graph $\mathcal{G}$ is said to be *completely connected* if for any $i, j \in \mathcal{V}$, $(i, j) \in \mathcal{E}$.

Consider the following class of linear discrete time-varying systems

$$x(k+1) = \left[A(k) + \theta(k)\tilde{A}(k)\right] x(k) + w(k), \tag{1}$$

where $x(k) \in \mathbb{R}^{n_x}$ is the state; $A(k)$ and $\tilde{A}(k)$ are known matrices with appropriate dimensions; $w(k) \in \mathbb{R}^{n_x}$ is the additive white noise with $\mathbb{E}\{w(k)\} = 0$ and $\mathbb{E}\left\{w(k)w^T(k)\right\} = S(k)$. $\theta(k) \in \mathbb{R}$ is the multiplicative noise with $\mathbb{E}\{\theta(k)\} = 0$ and $\mathbb{E}\left\{\theta^2(k)\right\} = \xi(k)$. $\mathbb{E}\{x(0)\}$ and $\mathbb{E}\{x(0)x^T(0)\}$ are assumed to be known.

For every sensor node i $(i = 1, 2, \dots, n)$, the measurement is described by

$$y_i(k) = \lambda_i(k)C_i(k)x(k) + v_i(k), \tag{2}$$

where $y_i(k) \in \mathbb{R}^{n_y}$ is the measurement of the ith node; $C_i(k)$s are known matrices with appropriate dimensions for all $i = 1, 2, \dots, n$; $v_i(k) \in \mathbb{R}^{n_y}$ is the additive white noise of the ith node with $\mathbb{E}\{v_i(k)\} = 0$ and $\mathbb{E}\left\{v_i(k)v_i^T(k)\right\} = V_i(k) > 0$. $\lambda_i(k)$, representing the sensor gain degradation in the ith node, is a random variable distributed over the interval $[a_i, b_i]$ $(0 \leq a_i \leq b_i \leq 1)$ with $\mathbb{E}\{\lambda_i(k)\} = m_i(k)$ and $\mathrm{Var}\{\lambda_i(k)\} = l_i(k)$, where $m_i(k)$ and $l_i(k)$ are known scalars. $\theta(k)$, $w(k)$, $\lambda_i(k)$ and $v_i(k)$ are all mutually independent.

Let $\hat{x}_i(k) \in \mathbb{R}^{n_x}$ denote the state estimate of the target plant *from the ith node*. In this paper, the filter to be designed is of the following structure for sensor node i:

$$\begin{aligned} \hat{x}_i(k+1) =& A(k)\hat{x}_i(k) + \sum_{j\in\mathcal{N}_i} H_{ij}(k)a_{ij} \\ & \times \left[y_j(k) - m_j(k)C_j(k)\hat{x}_j(k)\right], \end{aligned} \tag{3}$$

where the matrices $H_{ij}(k)$ are parameters to be determined. The initial values is $\hat{x}_i(0) = \mathbb{E}\{x(0)\}$ for all $1 \leq i \leq n$. Note that the proposed structure (3) reflects how the sensor nodes communicate with their neighbors via $\mathcal{N}_i$ so as to guarantee the unbiased estimation. The unbiasedness can be proven by mathematical induction as follows.

Firstly, one can verify that the unbiasedness assertion is true for $k=0$ according to $\hat{x}_i(0)=\mathbb{E}\{x(0)\}$ for all $i\in\mathcal{V}$. Secondly we assume that it is true for the integers from 0 to k and all the $j\in\mathcal{V}$. Letting $\tilde{x}_i(k)=x(k)-\hat{x}_i(k)$, we have the following system that governs the filtering error dynamics:

$$\begin{aligned}\tilde{x}_i(k+1) =& A(k)\tilde{x}_i(k)-\sum_{j\in\mathcal{N}_i}H_{ij}(k)a_{ij}m_j(k)C_j(k)\tilde{x}_j(k)\\&-\sum_{j\in\mathcal{N}_i}H_{ij}(k)a_{ij}v_j(k)+\Big\{\theta(k)\tilde{A}(k)\\&-\sum_{j\in\mathcal{N}_i}H_{ij}(k)a_{ij}[\lambda_j(k)-m_j(k)]C_j(k)\Big\}x(k)\\&+w(k),\end{aligned}\tag{4}$$

for $i=1,2,\ldots,n$. Then for any $i\in\mathcal{V}$, it follows from (4) that

$$\begin{aligned}\mathbb{E}\{\tilde{x}_i(k+1)\} =& A(k)\mathbb{E}\{\tilde{x}_i(k)\}-\sum_{j\in\mathcal{N}_i}H_{ij}(k)a_{ij}m_j(k)C_j(k)\\&\times\mathbb{E}\{\tilde{x}_j(k)\}-\sum_{j\in\mathcal{N}_i}H_{ij}(k)a_{ij}\mathbb{E}\{v_j(k)\}\\&+\Big\{\mathbb{E}\{\theta(k)\}\tilde{A}(k)-\sum_{j\in\mathcal{N}_i}H_{ij}(k)a_{ij}\\&\times[\mathbb{E}\{\lambda_j(k)\}-m_j(k)]C_j(k)\Big\}\mathbb{E}\{x(k)\}\\&+\mathbb{E}\{w(k)\}.\end{aligned}$$

Then, considering the facts that $w(k)$, $\theta(k)$, $\tilde{x}_i(k)$, and $v_i(k)$ are all zero-mean, and $\mathbb{E}\{\lambda_j(k)\}=m_j(k)$, it can be concluded that $\mathbb{E}\{\tilde{x}_i(k+1)\}=0$ for any $i\in\mathcal{V}$. That concludes the proof.

For notational simplicity, we define

$$\begin{aligned}&\tilde{x}(k)=\mathrm{vec}_n^T\{\tilde{x}_i^T(k)\},\ \ \bar{x}(k)=\mathrm{vec}_n^T\{x^T(k)\},\\&\bar{M}(k)=\mathrm{diag}_n\{m_i(k)I\},\ \ H(k)=[H_{ij}(k)]_{n\times n},\\&\bar{C}(k)=\mathrm{diag}_n\{C_i(k)\},\ \ \bar{A}(k)=\mathrm{diag}_n\{A(k)\},\\&\bar{w}(k)=\mathrm{vec}_n^T\{w^T(k)\},\ \ \bar{v}(k)=\mathrm{vec}_n^T\{v_i^T(k)\},\\&\hat{A}(k)=\mathrm{diag}_n\{\tilde{A}(k)\},\ \ T_i=\mathrm{diag}\{a_{i1}I,\ldots,a_{in}I\},\\&\bar{\Lambda}(k)=\mathrm{diag}_n\{\lambda_i(k)I\},\ \ E_i=\mathrm{diag}\{\underbrace{0,\ldots,0}_{i-1},I,\underbrace{0,\ldots,0}_{n-i}\},\end{aligned}$$

and then (4) can be rewritten in the following form:

$$\begin{aligned}\tilde{x}(k+1)=&\left[\bar{A}(k)-\sum_{i=1}^{n}E_iH(k)T_i\bar{M}(k)\bar{C}(k)\right]\tilde{x}(k)\\&+\Big\{\theta(k)\hat{A}(k)-\sum_{i=1}^{n}E_iH(k)T_i\\&\times\left[\bar{\Lambda}(k)-\bar{M}(k)\right]\bar{C}(k)\Big\}\bar{x}(k)\\&-\sum_{i=1}^{n}E_iH(k)T_i\bar{v}(k)+\bar{w}(k).\end{aligned}\tag{5}$$

Defining the error covariance at the kth time step as

$$P(k)=\mathbb{E}\left\{\tilde{x}(k)\tilde{x}^T(k)\right\},\tag{6}$$

the goal of this paper can be stated as designing a filter of the form (3) for system (1)-(2) so that the filtering error covariance $P(k+1)$ is minimized at each time step k.

Remark 1: In the proposed filter, each sensor node estimates the system states based on the local measurement and measurements of its neighboring sensors. The structure of the filter is set so as to achieve the unbiased estimation at each node. The statistics exploited in the filter (i.e., $m_i(k)$ for all $i\in\mathcal{V}$) are scalars known *a priori*, which facilitates the filter implementation. It is noted that, considering the complicated interconnections between the sensor nodes, there are various filter structures that can be adopted to achieve unbiased distributed state estimation and the proposed structure (3) is among them.

III. Main Results

A. Filter Design

In this subsection, a set of recursive Riccati-like matrix equations is derived to calculate the filter parameters $H_{ij}(k)$ in (3) in order to minimize the error variance for system (1). The following theorem gives the parameterizations of the desired filter gains in the two cases that the network is completely or not completely connected.

For presentation convenience, we denote

$$\Omega(0):=\mathbb{E}\left\{x(0)x^T(0)\right\},\tag{7}$$
$$\Omega(k):=\mathbb{E}\left\{x(k)x^T(k)\right\},\tag{8}$$
$$\bar{\Omega}(k):=\mathbb{E}\left\{\bar{x}(k)\bar{x}^T(k)\right\}=[\Omega(k)]_{n\times n},\tag{9}$$
$$\begin{aligned}U(k):=&\mathbb{E}\Big\{\left[\bar{\Lambda}(k)-\bar{M}(k)\right]\bar{C}(k)\bar{x}(k)\bar{x}^T(k)\bar{C}^T(k)\\&\times\left[\bar{\Lambda}(k)-\bar{M}(k)\right]^T\Big\}\\=&\,\mathrm{diag}_n\Big\{l_i(k)C_i(k)\Omega(k)C_i^T(k)\Big\},\end{aligned}\tag{10}$$
$$W(k):=\mathbb{E}\left\{\bar{w}(k)\bar{w}^T(k)\right\}=[S(k)]_{n\times n},\tag{11}$$
$$V(k):=\mathbb{E}\left\{\bar{v}(k)\bar{v}^T(k)\right\}=\mathrm{diag}_n\{V_i(k)\},\tag{12}$$
$$Y(k):=\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)+V(k)+U(k),\tag{13}$$
$$Z(k):=\bar{M}(k)\bar{C}(k)P(k)\bar{A}^T(k),\tag{14}$$
$$\mathcal{H}(k):=Z^T(k)Y^{-1}(k)=[\mathcal{H}_{ij}(k)]_{n\times n}.\tag{15}$$

Theorem 1: The following statements are true:

a). If the sensor network topology is completely connected, then the parameters of filter (3) achieving the minimum filtering error variance are given by:

$$H_{ij}(k)=\mathcal{H}_{ij}(k)a_{ij}^{-1},\tag{16}$$

and $P(k)$ is calculated as:

$$\begin{aligned}P(k+1)=&-Z^T(k)Y^{-1}(k)Z(k)+\bar{A}(k)P(k)\bar{A}^T(k)\\&+\xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k)+W(k).\end{aligned}\tag{17}$$

b). If the sensor network topology is not completely connected, a practical solution for the parameters of filter (3) is given by:

$$H_{ij}(k)=\begin{cases}\mathcal{H}_{ij}(k)a_{ij}^{-1}, & \text{if}\ \ a_{ij}\neq 0,\\ 0, & \text{if}\ \ a_{ij}=0,\end{cases}\tag{18}$$

and $P(k)$ obeys the following recursion:

$$\begin{aligned} P(k+1) = &\left[\sum_{i=1}^{n} E_i\mathcal{H}(k)T_i - Z^T(k)Y^{-1}(k)\right] Y(k) \\ &\times \left[\sum_{i=1}^{n} E_i\mathcal{H}(k)T_i - Z^T(k)Y^{-1}(k)\right]^T \\ &- Z^T(k)Y^{-1}(k)Z(k) + \bar{A}(k)P(k)\bar{A}^T(k) \\ &+ \xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k) + W(k). \end{aligned} \tag{19}$$

Proof: a). It follows from the definition (6) and the notations (7)-(15) that

$$\begin{aligned} P(k+1) = &\left[\bar{A}(k) - \sum_{i=1}^{n} E_iH(k)T_i\bar{M}(k)\bar{C}(k)\right] P(k) \\ &\times \left[\bar{A}(k) - \sum_{i=1}^{n} E_iH(k)T_i\bar{M}(k)\bar{C}(k)\right]^T \\ &+ \left[\sum_{i=1}^{n} E_iH(k)T_i\right] V(k) \left[\sum_{i=1}^{n} E_iH(k)T_i\right]^T \\ &+ \left[\sum_{i=1}^{n} E_iH(k)T_i\right] U(k) \left[\sum_{i=1}^{n} E_iH(k)T_i\right]^T \\ &+ \xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k) + W(k). \end{aligned} \tag{20}$$

Moreover, $\Omega(k)$ can be recursively calculated as follows:

$$\begin{aligned} \Omega(k+1) &= \mathbb{E}\left\{x(k+1)x^T(k+1)\right\} \\ &= \mathbb{E}\Bigg\{ \left\{\left[A(k) + \theta(k)\tilde{A}(k)\right] x(k) + w(k)\right\} \\ &\quad \times \left\{\left[A(k) + \theta(k)\tilde{A}(k)\right] x(k) + w(k)\right\}^T \Bigg\} \\ &= A(k)\Omega(k)A^T(k) + \xi(k)\tilde{A}(k)\Omega(k)\tilde{A}^T(k) \\ &\quad + S(k). \end{aligned} \tag{21}$$

From (20), it follows that

$$\begin{aligned} P(k+1) = &\left[\sum_{i=1}^{n} E_iH(k)T_i\right] Y(k) \left[\sum_{i=1}^{n} E_iH(k)T_i\right]^T \\ &- Z^T(k) \left[\sum_{i=1}^{n} E_iH(k)T_i\right]^T - \left[\sum_{i=1}^{n} E_iH(k)T_i\right] \\ &\times Z(k) + \xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k) \\ &+ \bar{A}(k)P(k)\bar{A}^T(k) + W(k). \end{aligned} \tag{22}$$

Since $Y(k) = Y^T(k) > 0$, (22) can be rewritten as

$$\begin{aligned} P(k+1) = &\left[\sum_{i=1}^{n} E_iH(k)T_i - Z^T(k)Y^{-1}(k)\right] Y(k) \\ &\times \left[\sum_{i=1}^{n} E_iH(k)T_i - Z^T(k)Y^{-1}(k)\right]^T \\ &- Z^T(k)Y^{-1}(k)Z(k) + \bar{A}(k)P(k)\bar{A}^T(k) \\ &+ \xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k) + W(k). \end{aligned} \tag{23}$$

In view of (15), it is obvious that $P(k+1)$ is minimized if and only if

$$\sum_{i=1}^{n} E_iH(k)T_i = \mathcal{H}(k). \tag{24}$$

To this end, it can be easily seen that, if the network topology is completely connected, then the minimum variance of the filtering error is achieved when $H(k)$ is calculated as in (16), which guarantees

$$[H_{i1}(k), \ldots, H_{in}(k)] = [\mathcal{H}_{i1}(k), \ldots, \mathcal{H}_{in}(k)]\, T_i^{-1}.$$

Furthermore, base on (6), the initial value of P is given by

$$\begin{aligned} P(0) &= \mathbb{E}\left\{\tilde{x}(0)\tilde{x}^T(0)\right\} \\ &= \left[\Omega(0) - \mathbb{E}\left\{x(0)\right\}\mathbb{E}\left\{x^T(0)\right\}\right]_{n\times n}, \end{aligned}$$

and then $P(k)$ can be updated according to (17).

b). In the case that the sensor network topology is not completely connected, it is easily seen that T_i is not invertible for all the $i \in \mathcal{V}$, and therefore the condition (24) no longer holds because of the sparse topology. For such a circumstance, an alternative yet effective way for designing the filter gains is to calculate $H(k)$ as (18). In doing so, it can be guaranteed that

$$[H_{i1}(k), \ldots, H_{in}(k)] = [\mathcal{H}_{i1}(k), \ldots, \mathcal{H}_{in}(k)]\, T_i^{\dagger}, \tag{25}$$

where $T_i^{\dagger}$ is the Moore-Penrose pseudo inverse of T_i. Moreover, in this case, it is straightforward to see that $P(k)$ can be recursively determined as (19). This ends the proof. ■

Remark 2: The filter parameters $H_{ij}(k)$ are calculated at each time step to minimize the filtering error covariance. It is worth mentioning that some statistics of the stochastic sensor gain degradations and noises, the network topology, the state transition matrix, and the measurement matrices are required to determine $P(k)$ and $H_{ij}(k)$. To update $P(k)$ and $H_{ij}(k)$ at each node, it is not necessary to request global measurements from all the sensors. Thus, $P(k)$ and $H_{ij}(k)$ can be updated at each node and the proposed algorithm is applicable in distributed sensor networks. It should also be noted that the filter designed in Part b) of Theorem 1 is obtained without assuming the complete connectedness of the network topology, and the proposed filter is applicable in a sparse sensor network.

In the next subsection, we proceed to deal with the performance evaluation problem of the filter developed in Theorem 1. Let us first discuss the boundedness of the estimation error at each time step.

B. Boundedness

For the dynamics analysis of the estimation error, we will need the following two widely used concepts for the boundedness of stochastic processes [1], [29].

Definition 1: The stochastic process $\zeta(k)$ is said to be exponentially bounded in mean square if there are real numbers $\eta > 0$, $\nu > 0$ and $0 < \vartheta < 1$ such that

$$\mathbb{E}\left\{\|\zeta(k)\|^2\right\} \leq \eta\|\zeta(0)\|^2\vartheta^k + \nu \tag{26}$$

holds for every $k > 0$.

Definition 2: The stochastic process $\zeta(k)$ is said to be bounded with probability one if

$$\sup_{k\geq 0}\|\zeta(k)\| < \infty \tag{27}$$

is true with probability one.

For the boundedness of the estimation error, we first establish sufficient conditions under which the estimation error is exponentially bounded in mean square, and then generalize the results to the case of boundedness with probability one. For this purpose, as in [24], we make the following two assumptions.

Assumption 1: There are positive real numbers $\bar{a}_1, \bar{a}_2, \bar{c}, \underline{c}, \bar{\omega}, \bar{\xi}, \bar{l}, \bar{m}, \underline{m}, \bar{w}, \underline{w}, \underline{v}, \bar{v} > 0$ such that the following bounds on various matrices are fulfilled for every $1 \leq i \leq n$ and $k \geq 0$:

$$\begin{aligned}&\|A(k)\| \leq \bar{a}_1, \ \ \left\|\hat{A}(k)\right\| \leq \bar{a}_2, \ \ \underline{c} \leq \|C_i(k)\| \leq \bar{c},\\ &\operatorname{tr}\left\{\bar{\Omega}(k)\right\} \leq \bar{\omega}, \ \ \xi(k) \leq \bar{\xi}, \ \ \underline{m}I \leq \bar{M}(k) \leq \bar{m}I,\\ &l_i(k) \leq \bar{l}, \ \ \underline{w}I \leq W(k) \leq \bar{w}I, \ \ \underline{v}I \leq V(k) \leq \bar{v}I.\end{aligned} \tag{28}$$

Moreover, the following inequality holds:

$$\bar{a}_1\left(1 + \frac{n\bar{m}^2\bar{c}^2}{\underline{m}^2\underline{c}^2}\right) < 1. \tag{29}$$

Assumption 2: There exists a positive real number $\epsilon > 0$ such that the initial estimation error satisfies

$$\|\tilde{x}(0)\| \leq \epsilon. \tag{30}$$

Theorem 2: Consider the time-varying system (1)-(2) with the minimum-variance filter given in (3) whose parameters are provided in Theorem 1. Under Assumption 1, the estimation error given by (4) is exponentially bounded in mean square.

Proof: Denoting

$$\breve{A}(k) = \bar{A}(k) - \sum_{i=1}^{n} E_i H(k) T_i \bar{M}(k)\bar{C}(k),$$

then (5) can be written as

$$\tilde{x}(k+1) = \breve{A}(k)\tilde{x}(k) + r(k) + s(k), \tag{31}$$

where

$$\begin{aligned}r(k) =&\left\{\theta(k)\hat{A}(k) - \sum_{i=1}^{n} E_i H(k) T_i \big[\bar{\Lambda}(k) - \bar{M}(k)\big]\right.\\ &\left.\times \bar{C}(k)\right\}\bar{x}(k),\\ s(k) =& -\sum_{i=1}^{n} E_i H(k) T_i \bar{v}(k) + \bar{w}(k).\end{aligned} \tag{32}$$

Based on (15), it follows easily from $\|\bar{A}(k)\| \leq \bar{a}_1$ and $\underline{c} \leq \|\bar{C}(k)\| \leq \bar{c}$ that

$$\begin{aligned}\|\mathcal{H}(k)\| =& \left\|\bar{A}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)[\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\right.\\ &\left.\times \bar{M}^T(k) + V(k) + U(k)]^{-1}\right\|\\ <& \frac{\bar{a}_1\bar{m}\bar{c}}{\underline{m}^2\underline{c}^2} := \bar{h}.\end{aligned} \tag{33}$$

According to the fact that $\left\|\sum_{i=1}^{n} E_i H(k) T_i\right\| \leq n\,\|\mathcal{H}(k)\|$, we have

$$\begin{aligned}\|\breve{A}(k)\| \leq& \|\bar{A}(k)\| + \left\|\sum_{i=1}^{n} E_i H(k) T_i\right\| \|\bar{M}(k)\bar{C}(k)\|\\ \leq& \|\bar{A}(k)\| + \|\mathcal{H}(k)\|\,\|\bar{M}(k)\bar{C}(k)\|\\ <& \bar{a}_1 + n\bar{h}\bar{m}\bar{c} := \bar{a}.\end{aligned} \tag{34}$$

Based on (29) and (33), it can be seen that $\bar{a} < 1$.

Denoting $L(k) = \operatorname{diag}_n\{l_i(k)I\}$, it is obvious that $L(k) < \bar{l}I$. Furthermore, we have

$$\begin{aligned}&\left\|\mathbb{E}\left\{\left\{\theta(k)\hat{A}(k) - \sum_{i=1}^{n} E_i H(k) T_i\big[\bar{\Lambda}(k) - \bar{M}(k)\big]\bar{C}(k)\right\}\right.\right.\\ &\left.\left.\times\left\{\theta(k)\hat{A}(k) - \sum_{i=1}^{n} E_i H(k) T_i\big[\bar{\Lambda}(k) - \bar{M}(k)\big]\bar{C}(k)\right\}^T\right\}\right\|\\ =&\left\|\xi(k)\hat{A}(k)\hat{A}^T(k) + \left[\sum_{i=1}^{n} E_i H(k) T_i\right]\left[L(k)\bar{C}(k)\bar{C}^T(k)\right]\right.\\ &\left.\times\left[\sum_{i=1}^{n} E_i H(k) T_i\right]^T\right\|\\ \leq&\left\|\xi(k)\hat{A}(k)\hat{A}^T(k)\right\| + n^2\|\mathcal{H}(k)\|^2\left\|L(k)\bar{C}(k)\bar{C}^T(k)\right\|\\ <&\bar{\xi}\bar{a}_2^2 + n^2\bar{h}^2\bar{l}\bar{c}^2.\end{aligned} \tag{35}$$

Thus, it is clear that

$$\mathbb{E}\left\{r^T(k)r(k)\right\} < \left(\bar{\xi}\bar{a}_2^2 + n^2\bar{h}^2\bar{l}\bar{c}^2\right)\bar{\omega}^2 := \bar{r}^2, \tag{36}$$

and the bounds of $s(k)$ can be calculated as follows:

$$\begin{aligned}&\mathbb{E}\left\{s^T(k)s(k)\right\}\\ =&\mathbb{E}\left\{\bar{v}^T(k)\left[\sum_{i=1}^{n} E_i H(k) T_i\right]^T\left[\sum_{i=1}^{n} E_i H(k) T_i\right]\bar{v}(k)\right.\\ &\left.+\bar{w}^T(k)\bar{w}(k)\right\}\\ <&\mathbb{E}\left\{n^2\bar{h}^2\bar{v}^T(k)\bar{v}(k) + \bar{w}^T(k)\bar{w}(k)\right\}\\ =&\mathbb{E}\left\{\operatorname{tr}\left\{n^2\bar{h}^2\bar{v}(k)\bar{v}^T(k) + \bar{w}(k)\bar{w}^T(k)\right\}\right\}\\ \leq&n^3 n_y\bar{h}^2\bar{v} + n n_x\bar{w} := \bar{s}^2.\end{aligned} \tag{37}$$

$$\mathbb{E}\left\{s(k)s^T(k)\right\} \geq \mathbb{E}\left\{\bar{w}(k)\bar{w}^T(k)\right\} = W(k) \geq \underline{w}I := \underline{s}^2 I. \tag{38}$$

From (31), it follows that

$$P(k+1) = \breve{A}(k)P(k)\breve{A}^T(k) + \breve{R}(k) + \breve{S}(k), \tag{39}$$

where

$$\breve{R}(k) = \mathbb{E}\left\{r(k)r^T(k)\right\}, \quad \breve{S}(k) = \mathbb{E}\left\{s(k)s^T(k)\right\}.$$

It is straightforward to see that $\breve{R}(k) \leq \bar{r}^2 I$ and $\underline{s}^2 I \leq \breve{S}(k) \leq \bar{s}^2 I$.

Consider the following iterative matrix equation with respect to $\Pi(k)$:

$$\Pi(k+1) =\breve{A}(k)\Pi(k)\breve{A}^T(k) + \left[\sum_{i=1}^{n} E_i H(k) T_i\right] V(k)$$

$$\times \left[\sum_{i=1}^{n} E_i H(k) T_i\right]^T + W(k). \quad (40)$$

with initial condition

$$\Pi(0) = \left[\sum_{i=1}^{n} E_i H(0) T_i\right] V(0) \left[\sum_{i=1}^{n} E_i H(0) T_i\right]^T + W(0).$$

With the definition of $\check{S}(k)$, (40) can be rewritten as:

$$\Pi(k+1) = \check{A}(k)\Pi(k)\check{A}^T(k) + \check{S}(k). \quad (41)$$

Then it follows directly that

$$\|\Pi(k+1)\| \leq \|\Pi(k)\| \left\|\check{A}(k)\right\|^2 + \left\|\check{S}(k)\right\| \leq \bar{a}^2 \|\Pi(k)\| + \bar{s}^2. \quad (42)$$

By iteration, we have

$$\|\Pi(k)\| \leq \bar{a}^{2k} \|\Pi(0)\| + \bar{s}^2 \sum_{i=0}^{k-1} \bar{a}^{2i}. \quad (43)$$

Since $0 < \bar{a} < 1$, we arrive at

$$\|\Pi(k)\| < \|\Pi(0)\| + \bar{s}^2 \sum_{i=0}^{\infty} \bar{a}^{2i} = \|\Pi(0)\| + \frac{\bar{s}^2}{1-\bar{a}^2}. \quad (44)$$

Furthermore, since $\Pi(k)$ is positive definite for all k, it is straightforward to see that

$$\|\Pi(k+1)\| \geq \left\|\check{S}(k)\right\| \geq \underline{s}^2. \quad (45)$$

Based on (44) and (45), it can be concluded that there are positive real numbers $\underline{\pi}, \bar{\pi} > 0$ such that the inequality $\underline{\pi} I \leq \Pi(k) \leq \bar{\pi} I$ holds for every $k \geq 0$.

According to (41), we obtain

$$\begin{aligned}
&\check{A}^T(k)\Pi^{-1}(k+1)\check{A}(k) - \Pi^{-1}(k) \\
&= -\left\{\Pi(k) + \Pi(k)\check{A}^T(k)\check{S}^{-1}(k)\check{A}(k)\Pi(k)\right\}^{-1} \\
&= -\left\{\check{A}^T(k)\check{S}^{-1}(k)\check{A}(k)\Pi(k) + I\right\}^{-1} \Pi^{-1}(k) \\
&< -\left(\frac{\bar{a}^2\bar{\pi}}{\underline{s}^2} + 1\right)^{-1} \Pi^{-1}(k). \quad (46)
\end{aligned}$$

Defining $\alpha = \left(\bar{a}^2\bar{\pi}/\underline{s}^2 + 1\right)^{-1}$, $\mu = \left(\bar{r}^2 + \bar{s}^2\right)/\underline{\pi}$, and $V_k\left(\tilde{x}(k)\right) = \tilde{x}(k)^T \Pi^{-1}(k)\tilde{x}(k)$, we obtain from (31) that

$$\begin{aligned}
&\mathbb{E}\left\{V_{k+1}(\tilde{x}(k+1)) | \tilde{x}(k)\right\} - V_k(\tilde{x}(k)) \\
=&\mathbb{E}\left\{\left[\check{A}(k)\tilde{x}(k) + r(k) + s(k)\right]^T \Pi^{-1}(k+1)\right. \\
&\left.\times \left[\check{A}(k)\tilde{x}(k) + r(k) + s(k)\right]\right\} - \tilde{x}^T(k)\Pi^{-1}(k)\tilde{x}(k) \\
=&-\tilde{x}^T(k)\left\{\check{A}^T(k)\check{S}^{-1}(k)\check{A}(k)\Pi(k) + I\right\}^{-1} \Pi^{-1}(k)\tilde{x}(k) \\
&+ \mathbb{E}\left\{r^T(k)\Pi^{-1}(k+1)r(k) + s^T(k)\Pi^{-1}(k+1)s(k)\right\} \\
<&-\alpha V_k(\tilde{x}(k)) + \mu, \quad (47)
\end{aligned}$$

which gives rise to

$$\mathbb{E}\left\{\|\tilde{x}(k)\|^2\right\} \leq \frac{\bar{\pi}}{\underline{\pi}} \|\tilde{x}(0)\|^2 (1-\alpha)^k + \mu\bar{\pi} \sum_{i=0}^{k-1} (1-\alpha)^i. \quad (48)$$

Noticing $0 < \alpha < 1$ and $\mu, \bar{\pi} > 0$, it follows that

$$\begin{aligned}
\mathbb{E}\left\{\|\tilde{x}(k)\|^2\right\} &\leq \frac{\bar{\pi}}{\underline{\pi}} \|\tilde{x}(0)\|^2 (1-\alpha)^k + \mu\bar{\pi} \sum_{i=0}^{\infty} (1-\alpha)^i \\
&= \frac{\bar{\pi}}{\underline{\pi}} \|\tilde{x}(0)\|^2 (1-\alpha)^k + \frac{\mu\bar{\pi}}{\alpha}. \quad (49)
\end{aligned}$$

Therefore, the stochastic process $\tilde{x}(k)$ is exponentially bounded in mean square and the proof is complete. ■

Under Assumption 2, i.e., the initial filtering error is bounded, we have $\|\tilde{x}(0)\|^2 < \epsilon < \infty$. Then, it follows from (30) and (49) that the stochastic process $\tilde{x}(k)$ is bounded with probability 1. In this case, the following corollary is easily accessible.

Corollary 1: Consider the time-varying system (1)-(2) with the minimum-variance filter given in (3) whose parameters are provided in Theorem 1. Under Assumptions 1-2, the estimation error given by (4) is bounded with probability 1.

Remark 3: The main proof of Theorem 2 provides a constructive way to quantify the error bound in (26). Furthermore, in practical systems, because of the energy constraints, it is reasonable to assume that the noise variances and the spectral norms of transfer matrix and measurement matrix are bounded.

Remark 4: Compared with the results in [24], the bound obtained in Theorem 2 is obviously dependent on the stochastic sensor gain degradation and the network topology, which results from the efforts we make on dealing with the stochastic and distributed nature of the system. Therefore, Theorem 2 offers a more applicable error bound to systems over sensor networks with stochastic sensor gain degradations. The consideration of the distributed measurements and the verifiable condition (29) constitute the main differences between our boundedness analysis and those in [13], [24].

C. Monotonicity

In this part we aim to discuss the relationship between the filtering performance and the sensor gain degradation. For demonstration purpose, we assume that all $\lambda_i(k)$ have the same statistics, i.e., $\bar{M}(k) = m(k)I$. We utilize $\mathrm{tr}\left\{P(k)\right\}$ as a standard criterion to measure the filtering performance.

In the following theorem, the influence from $m(k)$ to $\mathrm{tr}\left\{P(k+1)\right\}$ is clearly revealed.

Theorem 3: $\mathrm{tr}\left\{P(k+1)\right\}$ is nonincreasing when $m(k)$ increases.

Proof: Denote $F_i = \mathrm{diag}\left\{f_{i1}I, \ldots, f_{in}I\right\}$ where

$$f_{ij} = \begin{cases} 0, & \text{if} \quad a_{ij} \neq 0, \\ 1, & \text{if} \quad a_{ij} = 0. \end{cases}$$

Based on our assumption in Section II, we have $a_{ii} = 1$ and therefore $I - F_i$ is positive semidefinite and has at least one positive eigenvalue for all $1 \leq i \leq n$.

From the definitions of E_i, T_i, F_i and Theorem 1, it is clear that

$$\sum_{i=1}^{n} E_i H(k) T_i - \mathcal{H}(k) = -\sum_{i=1}^{n} E_i \mathcal{H}(k) F_i, \quad (50)$$

and then it follows from (23) and (50) that

$$\begin{aligned}
P(k+1) = &\left[\sum_{i=1}^{n} E_i \mathcal{H}(k) F_i\right] Y(k) \left[\sum_{j=1}^{n} E_j \mathcal{H}(k) F_j\right]^T \\
&- \left[\sum_{i=1}^{n} E_i\right] Z^T(k) Y^{-1}(k) Z(k) \left[\sum_{j=1}^{n} E_j\right]^T
\end{aligned}$$

$$
\begin{aligned}
&+\bar{A}(k)P(k)\bar{A}^T(k)+\xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k)\\
&+W(k)\\
=&\sum_{i=1}^{n}\sum_{j=1}^{n}E_i\mathcal{H}(k)F_iY(k)F_j^T\mathcal{H}^T(k)E_j^T\\
&-\sum_{i=1}^{n}\sum_{j=1}^{n}E_iZ^T(k)Y^{-1}(k)Z(k)E_j^T+\bar{A}(k)\\
&\times P(k)\bar{A}^T(k)+\xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k)+W(k).
\end{aligned} \tag{51}
$$

To simplify (51), we notice that, when $i \neq j$, the following facts are true: $\mathrm{tr}\left\{E_i\mathcal{H}(k)F_iY(k)F_j^T\mathcal{H}^T(k)E_j^T\right\}=0$, $\mathrm{tr}\left\{E_iZ^T(k)Y^{-1}(k)Z(k)E_j^T\right\}=0$. Then, we have

$$
\begin{aligned}
\mathrm{tr}\left\{P(k+1)\right\}=\mathrm{tr}\Bigg\{&\sum_{i=1}^{n}E_i\mathcal{H}(k)F_iY(k)F_i^T\mathcal{H}^T(k)-\sum_{i=1}^{n}E_i\\
&\times Z^T(k)Y^{-1}(k)Z(k)+\bar{A}(k)P(k)\bar{A}^T(k)\\
&+\xi(k)\hat{A}(k)\bar{\Omega}(k)\hat{A}^T(k)+W(k)\Bigg\}.
\end{aligned} \tag{52}
$$

Furthermore, it follows from (52) that

$$
\begin{aligned}
\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)}=&\frac{\partial}{\partial m(k)}\mathrm{tr}\Bigg\{\sum_{i=1}^{n}E_i\mathcal{H}(k)F_iY(k)F_i^T\mathcal{H}^T(k)\\
&-\sum_{i=1}^{n}E_iZ^T(k)Y^{-1}(k)Z(k)\Bigg\}\\
=&\frac{\partial}{\partial m(k)}\sum_{i=1}^{n}\mathrm{tr}\Bigg\{E_iZ^T(k)Y^{-1}(k)\\
&\times\left[F_iY(k)F_i^T-Y(k)\right]Y^{-1}(k)Z(k)\Bigg\}.
\end{aligned} \tag{53}
$$

With (13), (14) and (53), we have

$$
\begin{aligned}
\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)}=&\sum_{i=1}^{n}\mathrm{tr}\Bigg\{2\bar{C}(k)P(k)\bar{A}^T(k)E_iZ^T(k)\\
&\times Y^{-1}(k)\left[F_iY(k)F_i^T-Y(k)\right]Y^{-1}(k)\\
&-2\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)Y^{-1}(k)\Big\{[F_i\\
&\times Y(k)F_i^T-Y(k)]Y^{-1}(k)Z(k)E_iZ^T(k)\\
&+Z(k)E_iZ^T(k)Y^{-1}(k)[F_iY(k)F_i^T\\
&-Y(k)]\Big\}Y^{-1}(k)+2\bar{C}(k)P(k)\bar{C}^T(k)\\
&\times\bar{M}^T(k)\Big[F_i^TY^{-1}(k)Z(k)E_iZ^T(k)\\
&\times Y^{-1}(k)F_i-Y^{-1}(k)Z(k)E_i\\
&\times Z^T(k)Y^{-1}(k)\Big]\Bigg\}.
\end{aligned} \tag{54}
$$

When $m(k)>0$, we obtain

$$
\begin{aligned}
\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)}=&\sum_{i=1}^{n}\frac{1}{m(k)}\mathrm{tr}\Bigg\{2Z(k)E_iZ^T(k)Y^{-1}(k)\\
&\times\left[F_iY(k)F_i^T-Y(k)\right]Y^{-1}(k)-2\bar{M}(k)\\
&\times\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)Y^{-1}(k)\Big\{F_i\\
&\times Y(k)F_i^TY^{-1}(k)Z(k)E_iZ^T(k)+Z(k)\\
&\times E_iZ^T(k)Y^{-1}(k)F_iY(k)F_i^T\Big\}Y^{-1}(k)\\
&+2\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)F_i^T\\
&\times Y^{-1}(k)Z(k)E_iZ^T(k)Y^{-1}(k)F_i\\
&+2\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)\\
&\times Y^{-1}(k)Z(k)E_iZ^T(k)Y^{-1}(k)\Bigg\}.
\end{aligned}
$$

Using the well-known matrix identity $\mathrm{tr}\left\{\Gamma\Delta\right\}=\mathrm{tr}\left\{\Delta\Gamma\right\}$, where Γ and Δ are appropriately dimensioned, we have

$$
\begin{aligned}
\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)}=&\sum_{i=1}^{n}\frac{2}{m(k)}\mathrm{tr}\Bigg\{Y^{-1}(k)Z(k)E_iZ^T(k)\\
&\times Y^{-1}(k)\Big\{-\left[I-F_iY(k)F_i^TY^{-1}(k)\right]\\
&\times\Big[Y(k)-\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\\
&\times\bar{M}^T(k)\Big]\left[I-F_iY(k)F_i^TY^{-1}(k)\right]^T\\
&-F_i\Big[Y(k)-\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\\
&\times\bar{M}^T(k)\Big]F_i+F_iY(k)F_i^TY^{-1}(k)[Y(k)\\
&-\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)\Big]Y^{-1}(k)\\
&\times F_i^TY(k)F_i\Big\}\Bigg\}.
\end{aligned} \tag{55}
$$

Based on the fact that $\mathrm{eig}(\Xi\Psi)=\mathrm{eig}(\Psi\Xi)$, where Ξ and Ψ are square matrices with the same dimension and $\mathrm{eig}(\cdot)$ denotes the eigenvalues of a matrix, we have that $\mathrm{eig}(Y^{-1}(k)F_iY(k))=\mathrm{eig}(F_i)$. Therefore, the eigenvalues of $Y^{-1}(k)F_iY(k)-I$ are all either 0 or -1. Also notice that $Y(k)-\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\bar{M}^T(k)=U(k)+V(k)>0$. Then it follows from (55) that,

$$
\begin{aligned}
\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)}\leq&\sum_{i=1}^{n}\frac{2}{m(k)}\mathrm{tr}\Bigg\{Y^{-1}(k)Z(k)E_iZ^T(k)\\
&\times Y^{-1}(k)\Big\{-\left[I-F_iY(k)F_i^TY^{-1}(k)\right]\\
&\times\Big[Y(k)-\bar{M}(k)\bar{C}(k)P(k)\bar{C}^T(k)\\
&\times\bar{M}^T(k)\Big]\left[I-F_iY(k)F_i^TY^{-1}(k)\right]^T\Big\}\Bigg\}.
\end{aligned} \tag{56}
$$

Denote

$$
\begin{aligned}
X_i(k)=&-\left[I-F_iY(k)F_i^TY^{-1}(k)\right]\Big[Y(k)-\bar{M}(k)\bar{C}(k)\\
&\times P(k)\bar{C}^T(k)\bar{M}^T(k)\Big]\left[I-F_iY(k)F_i^TY^{-1}(k)\right]^T.
\end{aligned}
$$

It is obvious that $X_i(k)\leq 0$. So there exists a matrix $\mathcal{X}_i(k)$ such that $X_i(k)=-\mathcal{X}_i(k)\mathcal{X}_i^T(k)$ based on the eigenvalue decomposition of $X_i(k)$, and (56) can be written as

$$\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)} \leq -\sum_{i=1}^{n} \frac{2}{m(k)} \mathrm{tr}\Big\{ Y^{-1}(k) Z(k) E_i Z^T(k) \times Y^{-1}(k) \mathcal{X}_i(k) \mathcal{X}_i^T(k) \Big\} = -\sum_{i=1}^{n} \frac{2}{m(k)} \mathrm{tr}\Big\{ \mathcal{X}_i^T(k) Y^{-1}(k) Z(k) E_i \times Z^T(k) Y^{-1}(k) \mathcal{X}_i(k) \Big\}. \quad (57)$$

Noticing that the trace of a symmetric positive semidefinite matrix is always nonnegative, for all $m(k) \in (0,1)$, we have

$$\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m(k)} \leq 0, \quad (58)$$

which means that $\mathrm{tr}\left\{P(k+1)\right\}$ is nonincreasing as $m(k)$ increases. The proof of this theorem is now complete. ■

Remark 5: The finding in Theorem 3 is in accordance with the intuition. Actually, the increase of $m(k)$ can be interpreted that the sensor or the communication channel is under a better working condition. Theorem 3 shows that the filtering performance gets improved when more information about the measurements is received at each sensor node.

To further illustrate the engineering significance of Theorem 3, let us now look at the two interesting extreme cases of $F_i = I$ and $F_i = 0$.

In case of $F_i = I$, each node just estimates the dynamics based on $\hat{x}_i(k+1) = A(k)\hat{x}_i(k)$ (without any measurement signals). In such a situation, the measurements $y_i(k)$ have no effect on the filtering performance, and thus $\partial P(k+1)/\partial m_i(k) = 0$ for all $1 \leq i \leq n$.

In case of $F_i = 0$, we consider each sensor to be subject to individual sensor gain degradation, that is, $\mathbb{E}\left\{\lambda_i(k)\right\} = m_i(k)$. The following result is easily accessible from Theorem 3, and therefore only a sketch of the proof is provided.

Corollary 2: If the sensor topology is completely connected, then $\mathrm{tr}\left\{P(k+1)\right\}$ is nonincreasing as $m_i(k)$ increases for all $1 \leq i \leq n$.

Proof: When the directed graph is completely connected, $F_i = 0$ for all $1 \leq i \leq n$ according to its definition. From (53) we have:

$$\begin{aligned}\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m_i(k)} &= -\frac{\partial}{\partial m_i(k)} \mathrm{tr}\left\{Z^T(k) Y^{-1}(k) Z(k)\right\} \\ &= \mathrm{tr}\Big\{ 2E_i \bar{C}(k) P(k) \bar{C}^T(k) \bar{M}^T(k) Y^{-1}(k) \times Z(k) Z^T(k) Y^{-1}(k) - 2E_i \bar{C}(k) P(k) \times \bar{A}^T(k) Z^T(k) Y^{-1}(k) \Big\} \\ &= \mathrm{tr}\Big\{ 2E_i \bar{C}(k) P(k) \big[\bar{C}^T(k) \bar{M}^T(k) Y^{-1}(k) \times \bar{M}(k) \bar{C}(k) - P^{-1}(k)\big] P(k) \bar{A}^T(k) \bar{A}(k) \times P(k) \bar{C}^T(k) \bar{M}^T(k) Y^{-1}(k) \Big\}. \quad (59)\end{aligned}$$

It follows from (13) that $\bar{C}^T(k)\bar{M}^T(k)Y^{-1}(k)\bar{M}(k)\bar{C}(k) < P^{-1}(k)$ and, subsequently,

$$\frac{\partial \mathrm{tr}\left\{P(k+1)\right\}}{\partial m_i(k)} \leq 0, \quad (60)$$

which concludes the proof. ■

Remark 6: In the main results of this paper, a Kalman filter with proper gain structure is designed to guarantee the minimum error variance at each time step even if the sensor network is not completely connected, and the statistical information on the sensor gain degradation is explicitly reflected in the filter design. Then, sufficient conditions are established under which the estimation error is exponentially bounded in mean square. Furthermore, the relationship between the filter performance and the statistical law of the sensor gain degradation is revealed and analyzed in a mathematically rigorous way. The matrices T_i and F_i reflect the effects of the possibly sparse topology on the filter design and performance analysis, which means that the topology is taken into account and the proposed approach is applicable in distributed settings. The main differences between our work and the decentralized multisensor Kalman filter in [23] are threefold: 1) the consideration of the sparse sensor network; 2) the investigation on stochastic sensor gain degradation phenomenon, and 3) the estimation performance analysis with respect to the boundedness and the monotonicity.

In next section, a numerical example is provided to illustrate our proposed filter design scheme.

IV. Numerical Example

The time-varying target plant is modeled by (1) with the following parameters:

$$A(k) = \begin{bmatrix} 0.1315 + 0.0054\sin(k) & 0.0537 \\ 0.0201 & -0.1007 \end{bmatrix}, \quad \tilde{A}(k) = I.$$

Suppose that $w(k)$ is a zero-mean Gaussian white noise with covariance $2.5 \times 10^{-5} I$, and $\theta(k)$ uniformly distributes over $[-0.001,\ 0.001]$. The initial value of the state $x(0)$ is uniformly distributed over [-0.1, 0], and therefore $\mathbb{E}\left\{x(0)\right\} = [-0.05,\ -0.05]^T$.

The sensor network is represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ with the set of nodes $\mathcal{V} = \{1,2,3,4\}$, the set of edges $\mathcal{E} = \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,1),(3,3),(4,1),(4,4)\}$, and the adjacency elements associated with the edges of the graph are $a_{ij} = 1$.

The dynamics of the sensor nodes are described by (2) with parameters as follows:

$$C_1(k) = [0.82,\ 0.62], \quad C_2(k) = [0.75,\ 0.80],$$
$$C_3(k) = [0.74,\ 0.75], \quad C_4(k) = [0.75,\ 0.70].$$

The additive noises $v_i(k)$ are uncorrelated Gaussian sequences whose covariances are $2.5 \times 10^{-5} I$. $\lambda_i(k)$ is uniformly distributed, respectively, over $[0.45 + 0.1i,\ 0.95 + 0.1i]$ for $i = 1, \ldots, 4$.

Since the sensor network topology is not completely connected, the results of Theorem 1, Part b) are adopted. At each

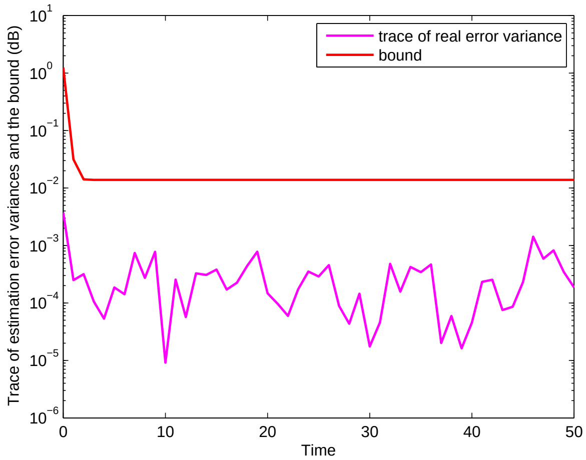

Fig. 1. The estimation error and bound

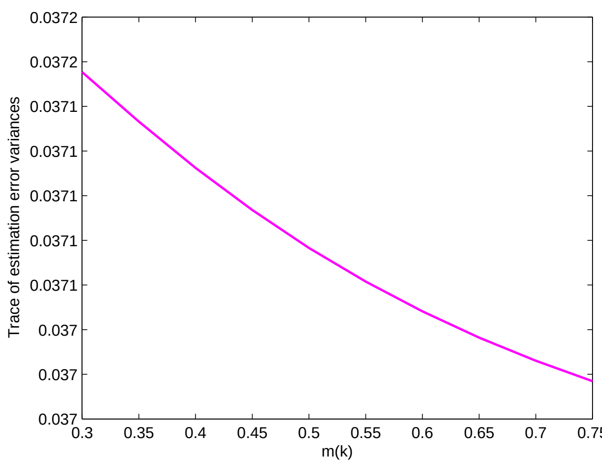

Fig. 2. The estimation error and sensor gain degradation

node, $P(k)$ is updated with (19) and $H_{ij}(k)$ is determined with (18). Fig. 1 plots the real estimation errors and the bound calculated from Theorem 2. In Fig. 2 we assume that all the sensor gain degradations have the same statistics, i.e., variables $\lambda_i(k)$ $(i = 1, \ldots, 4)$ are uniformly distributed over the same interval with the length of 0.5. The relationship between the accumulative estimation error variance in 50 time steps and the value of $m(k)$ is shown in Fig. 2. As pointed out in Theorem 3, the filtering error variance decreases when $m(k)$ increases.

V. Conclusion

In this paper, the minimum variance filtering problem has been investigated for a class of time-varying systems through sensor networks subject to both additive/multiplicative noises and stochastic sensor gain degradation. The gain degradation phenomenon for each individual sensor occurs in a random way governed by a random variable distributed over the interval $[0, 1]$. The distributed filter has been designed recursively by solving a set of Riccati-like matrix equations, such that the overall estimation error variance is minimized at each time step. The performance of the designed filters has been analyzed in terms of both the boundedness and monotonicity. Specifically, sufficient conditions have been obtained under which the estimation error is exponentially bounded in mean square. Moreover, the monotonicity property for the error variance with respect to the sensor gain degradation has been thoroughly discussed in the case where all the sensors are subject to gain degradation with the same probability. Numerical simulations have been exploited to show the effectiveness of the proposed filtering algorithm.

References

[1] R. G. Agnel, and E. I. Jury. Almost sure boundedness of randomly sampled systems. *SIAM Journal on Control*, vol. 9, no. 3, pp. 372–384, 1971.
[2] R. Caballero-Águila, A. Hermoso-Carazo, J. D. Jiménez-López, J. Linares-Pérez, and S. Nakamori. Signal estimation with multiple delayed sensors using covariance information. *Digital Signal Processing*, vol. 20, no. 2, pp. 528–540, 2010.
[3] G. C. Calafiore, and F. Abrate. Distributed linear estimation over sensor networks. *International Journal of Control*, vol. 82, no. 5, pp. 868–882, 2009.
[4] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri. Distributed Kalman filtering based on consensus strategies. *IEEE Journal on Selected Areas in Communication*, vol. 26, no. 4, pp. 622–633, 2008.
[5] F. S. Cattivelli, and A. H. Sayed. Diffusion strategies for distributed Kalman filtering and smoothing. *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.
[6] F. S. Cattivelli, and A. H. Sayed. Diffusion LMS strategies for distributed estimation. *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1035–1048, 2010.
[7] H. Dong, Z. Wang, and H. Gao. Distributed filtering for a class of time-varying systems over sensor networks with quantization errors and successive packet dropouts. *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3164–3173, 2012.
[8] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Distributed moving horizon estimation for nonlinear constrained systems. *International Journal of Robust and Nonlinear Control*, vol. 22, no. 2, pp. 123–143, 2012.
[9] X. He, Z. Wang, and D. H. Zhou. State estimation for time-delay systems with probabilistic sensor gain reductions. *Asia-Pacific Journal of Chemical Engeneering*, vol. 3, no. 6, pp. 712–716, 2008.
[10] S. Kar, and J. M. F. Moura. Gossip and distributed Kalman filtering: weak consensus under weak detectability. *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1766–1784, 2011.
[11] H. R. Karimi. Robust H_∞ filter design for uncertain linear systems over network with network-induced delays and output quantization. *Modeling, Identification and Control*, vol. 30, no. 1, pp. 27–37, 2009.
[12] U. A. Khan, and J. M. F. Moura. Distributing the Kalman filter for large-scale systems. *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.
[13] S. Kluge, K. Reif and M. Brokate, Stochastic stability of the extended Kalman filter with intermittent observations, *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 514–518, 2010.
[14] T. Li, M. Fu, L. Xie, and J. Zhang. Distributed consensus with limited communication data rate. *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 279–292, 2011.
[15] W. Li, and Y. Jia. Consensus-based distributed multiple model UKF for jump markov nonlinear systems. *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 230–236, 2012.
[16] J. Liang, Z.. Wang, and X. Liu. Distributed state estimation for discrete-time sensor networks with randomly varying nonlinearities and missing measurements. *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 486–496, 2011.
[17] J. Linares-Pérez, A. Hermoso-Carazo, R. Caballero-Águila, and J. D. Jiménez-López. Least-squares linear filtering using observations coming from multiple sensors with one- or two-step random delay. *Signal Processing*, vol. 89, no. 10, pp. 2045–2052, 2009.
[18] S. Liu, L. Xie, and H. Zhang. Distributed consensus for multi-agent systems with delays and noises in transmission channels. *Automatica*, vol. 47, no. 5, pp. 920–934, 2011.
[19] Y. Liu, X. He, Z. Wang, and D. H. Zhou. Optimal filtering for networked systems with stochastic sensor gain degradation. *Automatica*, vol. 50, no. 5, pp. 1521–1525, 2014.

[20] A. Manor, A. Osovizky, E. Dolev, E. Marcus, D. Ginzburg, V. Pushkarsky, Y. Kadmon, and Y. Cohen. Compensation of scintillation sensor gain variation during temperature transient conditions using signal processing techniques. In *IEEE Nuclear Science Symposium Conference*, Orlando, FL, USA, October, pp. 2399–2403, 2009.
[21] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, LA, USA, December, pp. 1814–1820, 2007.
[22] R. Olfati-Saber, and J. S. Shamma. Consensus filters for sensor networks and distributed sensor fusion. In *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, Seville, Spain, December, pp. 6698–6703, 2005.
[23] B. Rao, and H. Durrant-Whyte. Fully decentralized algorithm for multisensor Kalman filtering. *IEE Proceedings-D: Control Theory and Applications*, vol. 138, pp. 413–420, 1991.
[24] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen. Stochastic stability of the discrete-time extended Kalman filter. *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 714–728, 1999.
[25] I. S. D. Solomon, and A. J. Knight. Spatial processing of signals received by platform mounted sonar. *IEEE Journal of Oceanic Engineering*, vol. 27, no. 1, pp. 57–65, 2002.
[26] A. Speranzon, C. Fischione, K. H. Johansson, and A. Sangiovanni-Vincentelli. A distributed minimum variance estimator for sensor networks. *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 609–621, 2008.
[27] S. Sun, and Z. Deng. Distributed optimal fusion steady-state Kalman filter for systems with coloured measurement noises. *International Journal of Systems Science*, vol. 36, no. 3, pp. 113–118, 2005.
[28] C. P. Tan, and C. Edwards. Sliding mode observers for detection and reconstruction of sensor faults. *Automatica*, vol. 38, no. 10, pp. 1815–1821, 2002.
[29] T. J. Tarn, and Y. Rasis. Observers for nonlinear stochastic systems. *IEEE Transactions on Automatic Control*, vol. 21, no. 4, pp. 441–448, 1976.
[30] V. Ugrinovskii. Distributed robust filtering with H_∞ consensus of estimates. *Automatica*, vol. 47, no. 1, pp. 1–13, 2011.
[31] H. Yalcin, R. Collins, and M. Hebert. Background estimation under rapid gain change in thermal imagery. *Computer Vision and Image Understanding*, vol. 106, no. 2-3, pp. 148–161, 2007.
[32] W. Yang, X. Wang, and H. Shi. Optimal consensus-based distributed estimation with intermittent communication. *International Journal of Systems Science*, vol. 42, no. 9, pp. 1521–1529, 2011.
[33] W. Yu, G. Chen, Z. Wang, and W. Yang. Distributed consensus filtering in sensor networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 6, pp. 1568–1577, 2009.
[34] H. Yu, Y. Zhuang, and W. Wang. Distributed H_∞ filtering with consensus in sensor networks: a two-dimensional system-based approach. *International Journal of Systems Science*, vol. 42, no. 9, pp. 1543–1557, 2011.

Yang Liu received the Bachelors degree from the department of automation, Tsinghua University, Beijing, China, in 2010. He is currently pursuing the Ph.D. degree in Tsinghua University. His research interests include robust optimal filtering, closed-loop system, and fault detection and diagnosis.

Zidong Wang 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 at Brunel University London in the UK. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. His research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 200 papers in refereed 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.

He is a Fellow of the IEEE. He is serving or has served as 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-Systems. He is also a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Xiao He received the B.E. degree in information technology from Beijing Institute of Technology, Beijing, China, in 2004, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2010.

He is currently an Associate Professor with the Department of Automation, Tsinghua University, Beijing, China. His research interests include networked filter and control, networked fault diagnosis and isolation, and their application. He has published more than twenty papers in refereed international journals. He is now a member of the IEEE.

D. H. Zhou received the B.Eng., M. Sci., and Ph.D. degrees in electrical engineering from Shanghai Jiaotong University, China, in 1985, 1988, and 1990, respectively. He was an Alexander von Humboldt research fellow with the university of Duisburg, Germany, from 1995 to 1996, and a visiting scholar with Yale university, New Haven, CT, USA, from 2001 to 2002. He joined Tsinghua University in 1997, and was a professor and the head of the Department of Automation, Tsinghua University, Beijing, China, from 2008 to 2015. He is now the Vice President of Shandong University of Science and Technology, Qingdao, China. He has authored and coauthored over 130 peer-reviewed international journal papers and 6 monographs in the areas of process identification, fault diagnosis, fault-tolerant control, reliability prediction, and predictive maintenance.

Dr. Zhou is a member of the IFAC Technical Committee on Fault Diagnosis and Safety of Technical Processes, a senior member of IEEE, an associate editor of the Journal of Process Control, the associate Chairman of Chinese Association of Automation (CAA). He was also the NOC Chair of the 6th IFAC Symposium on SAFEPROCESS 2006.