A New Approach to Linear/Nonlinear Distributed Fusion Estimation Problem

Bo Chen , *Member, IEEE*, Guoqiang Hu , *Member, IEEE*, Daniel W.C. Ho , *Fellow, IEEE*, and Li Yu , *Member, IEEE*

Abstract—In this paper, we study the distributed fusion estimation problem for linear time-varying systems and nonlinear systems with bounded noises, where the addressed noises do not provide any statistical information, and are unknown but bounded. When considering linear time-varying fusion systems with bounded noises, a new local Kalman-like estimator is designed such that the square error of the estimator is bounded as time goes to ∞. A novel constructive method is proposed to find an upper bound of fusion estimation error, then a convex optimization problem on the design of an optimal weighting fusion criterion is established in terms of linear matrix inequalities, which can be solved by standard software packages. Furthermore, according to the design method of linear time-varying fusion systems, each local nonlinear estimator is derived for nonlinear systems with bounded noises by using Taylor series expansion, and a corresponding distributed fusion criterion is obtained by solving a convex optimization problem. Finally, target tracking system and localization of a mobile robot are given to show the advantages and effectiveness of the proposed methods.

Index Terms—Bounded noises, convex optimization, distributed fusion estimation, linear time-varying systems, nonlinear estimation.

I. INTRODUCTION

Multisensor fusion estimation has been one of the most important focuses in the area of information fusion during the past two decades. Since estimation performance and reliability can be improved by the redundant information from multiple sensors, different fusion estimation methods have been found in many application fields such as target tracking and localization, guidance and navigation [1], fault detection [2], sensor networks [3], and cyber-physical systems [4]. Generally, there exist two kinds of fusion estimation structures: centralized fusion structure and distributed fusion structure. Under the centralized fusion structure, all measurement data from different sensors are communicated to the fusion center (FC), and the design of centralized fusion estimator is equivalent to that of state estimator with single sensor. Though the centralized fusion estimation can provide optimal estimation performance, it has poor robustness and reliability when there are faulty sensors and FC. Under the distributed fusion structure, the measurements are first used to derive the local estimates at each sensor, and then these local estimates are sent to the FC to yield optimal or suboptimal state estimate by designing certain fusion criteria. Compared with the centralized fusion structure, the distributed fusion structure is generally more robust, reliable, and fault tolerant [5]. Therefore, most of existing works are focused on how to design the distributed fusion estimation algorithms.

Generally, the physical processes are modeled by linear or nonlinear dynamical systems, while the disturbance noises in multisensor fusion systems are considered as Gaussian or non-Gaussian disturbances. When considering the Gaussian white noise with known covariances, there mainly exist three different distributed fusion estimation methods.

1) Optimal distributed fusion estimation strategies [6]–[8].
2) Suboptimal distributed weighted fusion estimation methods [9]–[11].
3) Suboptimal distributed covariance intersection fusion estimation methods [12]–[15].

Notice that the assumption of Gaussian white noises may not be satisfied in practical systems, particularly, the accurate covariances may not be obtained in practical applications. To overcome this drawback, the energy-bounded noises, which do not require any statistical property of noises, have been considered for multisensor fusion systems, and different distributed H_∞ fusion estimation methods were developed in [16]–[18]. Subsequently, when simultaneously considering the energy-bounded noises and Gaussian white noises with unknown covariances, the distributed mixed H_2/H_∞ fusion estimation algorithms have been developed in [19] and [20] for a class of networked fusion systems. Though the conditions of disturbance noises have been relaxed in [16]–[20], the addressed fusion systems in [16]–[20] were time invariant. Moreover, the energy-bounded noise $w(t)$ in the H_∞ fusion framework must satisfy $\lim_{t\to\infty} w(t) = 0$, which may not be true in some practical systems (e.g., sensor's measurement noise generated from the external environment always exists). To make up for these shortages, a novel fusion method was developed in [21] to solve the networked fusion estimation problem, where a general framework was proposed in [21] to deal with state estimation problem under bounded noises. However, how to find the closest upper bounds and the most suitable optimization objective were not well solved in [21], which still remains challenging.

It should be pointed out that the works in [6]–[21] were concerned with the fusion estimation problem of linear systems. When considering the nonlinear systems with Gaussian white noises, the distributed fusion estimation algorithm was developed in [24] by using the extended Kalman filter (EKF) [25], while the unscented information

fusion filtering algorithm was derived in [26] by using the unscented Kalman filter (UKF) [27], [28]. Meanwhile, the fifth-degree ensemble iterated cubature square-root information filter was introduced in [29] to design nonlinear fusion estimation algorithm, while the support vector regression methodology was proposed in [30] to design nonlinear fusion strategy. Recently, different sequential fusion estimation methods for nonlinear systems were presented in [31] and [32] based on the UKF. Notice that the fusion estimation methods in [24], [26], and [29] –[32] assumed that the system disturbances must be Gaussian white noises with known covariance. On the other hand, when considering non-Gaussian noises in nonlinear fusion estimation framework, the consensus and Rao–Blackwellized fusion particle filtering algorithms were designed in [33] and [34], where the probability density function was required to be known in advance. Meanwhile, the modified Kalman filtering methods in [35] and [36] may also be used to solve nonlinear fusion estimation problem, where the statistical information of noises are required to be known in advance. As mentioned before, the disturbance noises in practical systems are always bounded, and the statistical property of noises is difficult to be accurately obtained in practical applications. However, the above-mentioned methods are not applicable to this case, and few results are focused on the distributed fusion estimation problem for nonlinear systems with bounded noises.

Motivated by the above analysis, we shall study the distributed fusion estimation problem for linear time-varying systems and nonlinear systems with bounded noises. The main contributions of this paper are summarized as follows.

1) For the linear time-varying systems, a new *stable* local Kalman-like estimator, which is different from the estimator structure in [21], is obtained by solving a convex optimization problem at each time step. Subsequently, by constructing a new upper bound of fusion estimation error, an optimal distributed fusion criterion is designed by solving a class of convex optimization problems.

2) Linearizing the nonlinear function using Taylor series expansion, the general nonlinear system reduces to linear time-varying systems, and the linearized errors can be viewed as bounded noises. Under this case, according to the design method of linear time-varying fusion systems, each local nonlinear estimator and a distributed fusion criterion are designed by establishing different convex optimization problems.

3) The proposed fusion estimation method in this paper does not require any statistical information of noises as compared with the classical Kalman fusion estimation methods, and this method can also be applicable to linear time-varying systems and nonlinear systems as compared with the existing H_∞ fusion estimation methods.

Notice that the solutions to the convex optimization problems in this paper can be directly obtained by the standard software packages. In the simulations, the advantages of the linear estimation method in this paper are shown by comparing with the state estimator in [21], the Kalman filter [22], and the H_∞ filter in [23], while the advantages of the nonlinear estimation method in this paper are shown by comparing with the classical EKF [25] and UKF [27], [28].

Notations: The superscript "T" represents the transpose, while "I" represents the identity matrix with appropriate dimension. $X > (<)0$ denotes a positive-definite (negative-definite) matrix, and $\text{diag}\{\cdot\}$ stands for a block diagonal matrix. $\lambda_{\max}(\cdot)$ means the maximum eigenvalue of the corresponding matrix, while $||A||_2$ is the 2-norm of matrix A. The symmetric terms in a symmetric matrix are denoted by "$*$", and $\text{col}\{a_1, \ldots, a_n\}$ means a column vector whose elements are $a_1, \ldots, a_n$. Moreover, if $\tau_1 > \tau_2$, it will be specified that $\prod_{\tau=\tau_1}^{\tau_2} F(\tau) = I_m$, where $F(\tau) \in \mathbb{R}^{m \times m}$ represents a matrix function with respect to the variable τ.

II. PROBLEM STATEMENT

Consider a nonlinear system described by the following state-space model:

$$\mathrm{x}(t + 1) = \mathrm{f}(\mathrm{x}(t)) + B(t)w(t) \tag{1}$$

$$y_i(t) = \mathrm{g}_i(\mathrm{x}(t)) + B_i(t)v_i(t)(i = 1, 2, \ldots, L) \tag{2}$$

where $x(t) \in \mathrm{R}^n$ is the system state, $y_i(t) \in \mathrm{R}^{q_i}$ is the measured output from sensor i, and L is the number of sensors. $\mathrm{f}(x(t)) \in \mathrm{R}^{n \times 1}$ and $\mathrm{g}_i(x(t)) \in \mathrm{R}^{q_i \times 1}$ are nonlinear vector functions that are assumed to be continuously differentiable, while $B(t)$ and $B_i(t)(i = 1, 2, \ldots, L)$ are time-varying bounded matrices with appropriate dimensions. $w(t)$ and $v_i(t)$ are *bounded noises*, i.e.,

$$w^{\mathrm{T}}(t)w(t) \le \delta_w, v_i^{\mathrm{T}}(t)v_i(t) \le \delta_{v_i} \tag{3}$$

where δ_w and δ_{v_i} are *unknown*. At each sensor, based on the measurements $\{y_i(1), y_i(2), \ldots, y_i(t)\}$, the local state estimate (LSE) $\hat{\mathrm{x}}_i(t)$ for nonlinear systems (1)–(2) is given by

$$\begin{cases} \hat{\mathrm{x}}_i^{\mathrm{p}}(t) = \mathrm{f}(\hat{\mathrm{x}}_i(t-1)) \\ \hat{\mathrm{x}}_i(t) = \hat{\mathrm{x}}_i^{\mathrm{p}}(t) + \mathrm{K}_i^N(t)[y_i(t) - \mathrm{g}_i(\hat{\mathrm{x}}_i^{\mathrm{p}}(t))] \end{cases} \tag{4}$$

where $\hat{\mathrm{x}}_i^{\mathrm{p}}(t)$ denotes one-step prediction, and an optimal gain $\mathrm{K}_i^N(t)$ is to be designed in Section III.

When the nonlinear systems (1)–(2) are reduced to the following linear discrete time-varying systems:

$$\mathrm{x}(t + 1) = A(t)\mathrm{x}(t) + B(t)w(t) \tag{5}$$

$$y_i(t) = C_i(t)\mathrm{x}(t) + B_i(t)v_i(t)(i = 1, 2, \ldots, L) \tag{6}$$

where $A(t)$ and $C_i(t)$ are time-varying matrices with appropriate dimensions. Then, each LSE $\hat{\mathrm{x}}_i(t)$ for linear systems (5)–(6) is given by the Kalman-like structure

$$\hat{\mathrm{x}}_i(t) = A(t-1)\hat{\mathrm{x}}_i(t-1)$$
$$+ \mathrm{K}_i^L(t)[y_i(t) - C_i(t)A(t-1)\hat{\mathrm{x}}_i(t-1)] \tag{7}$$

where an optimal gain $\mathrm{K}_i^L(t)$ is to be designed in Section III.

Subsequently, based on the LSEs (4) or (7), the distributed fusion estimate (DFE) of $\mathrm{x}(t)$ is given by

$$\hat{\mathrm{x}}(t) = \sum_{i=1}^{L} \Omega_i(t)\hat{\mathrm{x}}_i(t) \tag{8}$$

where $\sum_{i=1}^{L} \Omega_i(t) = I$, and optimal weighting matrices $\Omega_1(t), \ldots, \Omega_L(t)$ will be designed in Section III. Consequently, the problems to be solved in this paper are described as follows.

1) The first aim is to design optimal gains $\mathrm{K}_i^N(t)$ in (4) and $\mathrm{K}_i^L(t)$ in (7) such that an upper bound of the square error (SE) of the corresponding LSE $\hat{\mathrm{x}}_i(t)$ is minimal at each time, and the SE of $\hat{\mathrm{x}}_i(t)$ for linear time-varying systems is bounded as t goes to ∞.

2) The second aim is to design optimal weighting matrices $\Omega_1(t), \ldots, \Omega_L(t)$ in (8) such that an upper bound of the SE of the DFE $\hat{\mathrm{x}}(t)$ is minimum at each time.

Remark 1: When considering the linear time-varying systems with bounded noises, the LSE in [21] was given by

$$\hat{\mathrm{x}}_i(t) = A(t-1)\hat{\mathrm{x}}_i(t-1)$$
$$+ \mathrm{K}_i(t)(y_i(t-1) - C_i(t-1)\hat{\mathrm{x}}_i(t-1)) \tag{9}$$

where $\mathrm{K}_i(t)$ is the estimator gain of (9). The difference of the estimator structures between the LSE (7) and the LSE (9) is that

the LSE (9) in [21] was designed based on the measurements $\{y_i(1),\ldots,y_i(t-1)\}$, while the LSE (4) is designed based on the measurements $\{y_i(1),\ldots,y_i(t-1),y_i(t)\}$. Notice that $y_i(t)$ can provide important information of $x(t)$, but the design of the LSE (9) did not use the measurement $y_i(t)$. Thus, more available information on the state $x(t)$ is used in this paper to design the estimator. In this sense, under the same criterion of performance assessment, the estimation performance of the LSE (7) can be improved as compared with the LSE in [21].

Remark 2: Compared with the Kalman fusion estimation algorithms in [3]–[15], [24], [26], and [29]–[32], the proposed fusion estimation algorithm in this paper does not require any statistical information of the noises. Since the Gaussian white noises are always bounded in a practical system, the fusion estimation algorithm in this paper is also applicable to the case of Gaussian white noises.

Remark 3: Compared with the H_∞ fusion estimation algorithms in [16]–[20], the proposed fusion estimation algorithm in this paper does not require that the noise $w(t)$ (or $v_i(t)$) is energy bounded (i.e., $\lim_{t\to\infty} w(t) = 0$), and is applicable to time-varying systems and nonlinear systems. Notice that the energy-bounded noise can be viewed as a special case of bounded noises. Moreover, the nonlinear fusion estimation methods based on Taylor series expansion cannot be obtained from the similar ideas in [16]–[20], because the fusion estimation methods in [16]–[20] only dealt with linear time-invariant systems, but the linearized systems must be time varying.

III. MAIN RESULTS

A. DFE Design for Linear Time-Varying Systems

In this section, an optimal local estimator gain $K_i^L(t)$ in (7) and optimal weighting matrices $\Omega_1(t),\ldots,\Omega_L(t)$ in (8) for linear systems (5)–(6) will be presented in Theorem 1. Before deriving the result of Theorem 1, let us define

$$
\begin{cases}
G_{K_i}^L(t) \triangleq I - K_i^L(t)C_i(t) \\
B_{f_i}^L(t) \triangleq [G_{K_i}^L(t)B(t-1) \ -K_i^L(t)B_i(t)] \\
\bar{B}_{f_i}^L(t) \triangleq [G_{K_i}^L(t)B(t-1)\ 0\cdots\ -K_i^L(t)B_i(t)\cdots 0] \\
A_F^L(t) \triangleq \mathrm{diag}\{G_{K_1}^L(t)A(t-1),\ldots,G_{K_L}^L(t)A(t-1)\} \\
B_F^L(t) \triangleq \mathrm{col}\{\bar{B}_{f_1}^L(t),\ldots,\bar{B}_{f_L}^L(t)\}.
\end{cases}
\tag{10}
$$

Theorem 1: An optimal estimator gain $K_i^L(t)$ in (7) can be obtained by solving the following convex optimization problem:

$$
\min_{\vartheta_i(t)>0,P_i(t)>0,\Theta_i(t)>0,K_i^L(t)} \mathrm{Tr}\{\Theta_i(t)\}
$$

$$
\text{s.t.} : \begin{cases}
\begin{bmatrix} -I & G_{K_i}^L(t)A(t-1) & B_{f_i}^L(t) \\ * & -P_i(t) & 0 \\ * & * & -\Theta_i(t) \end{bmatrix} < 0 \\
P_i(t) - \vartheta_i(t)I < 0 \\
\vartheta_i(t) < 1
\end{cases}
\tag{11}
$$

where $G_{K_i}^L(t)$ and $B_{f_i}^L(t)$ are defined in (10). Under this case, the SE of $\hat{x}_i(t)$ will be bounded, i.e., there must exist a positive scalar $p_i > 0$ such that

$$
\lim_{t\to\infty} e_i^T(t)e_i(t) < p_i.
\tag{12}
$$

Moreover, a group of optimal weighting matrices $\Omega_1(t),\ldots,\Omega_L(t)$ in (8) for linear systems (5)–(6) can be obtained by solving the following convex optimization problem:

$$
\min_{\Omega(t),\Upsilon(t),P(t)>0,\Theta(t)>0} \mathrm{Tr}\{P(t)\} + \mathrm{Tr}\{\Theta(t)\}
$$

$$
\text{s.t.} : \begin{bmatrix} -I & \Omega(t)A_F^L(t) & \Omega(t)B_F^L(t) \\ * & -P(t) & -\Upsilon(t) \\ * & * & -\Theta(t) \end{bmatrix} < 0
\tag{13}
$$

where $\Omega(t) \triangleq [\Omega_1(t),\ldots,\Omega_{L-1}(t), I - \sum_{i=1}^{L-1}\Omega_i(t)]$, while $A_F^L(t)$ and $B_F^L(t)$ are defined in (10).

Proof: Define $e_i(t) \triangleq x(t) - \hat{x}_i(t)$ and $\xi_i(t-1) \triangleq \mathrm{col}\{w(t-1),v_i(t)\}$. Then, the estimation error $e_i(t)$ is given by

$$
e_i(t) = G_{K_i}^L(t)A(t-1)e_i(t-1) + B_{f_i}^L(t)\xi_i(t-1)
\tag{14}
$$

where $G_{K_i}^L(t)$ and $B_{f_i}^L(t)$ are defined in (10).

Next, the following performance index is introduced:

$$
J_i(t) \triangleq e_i^T(t)e_i(t) - e_i^T(t-1)P_i(t)e_i(t-1)
$$
$$
- \xi_i^T(t-1)\Theta_i(t)\xi_i(t-1)
\tag{15}
$$

where $P_i(t) > 0$ and $\Theta_i(t) > 0$. Then, it follows from (14) that

$$
J_i(t) = \begin{bmatrix} e_i(t-1) \\ \xi_i(t-1) \end{bmatrix}^T \underbrace{\begin{bmatrix} Z_{i1}(t) & Z_{i2}(t) \\ * & Z_{i3}(t) \end{bmatrix}}_{Z_i(t)} \begin{bmatrix} e_i(t-1) \\ \xi_i(t-1) \end{bmatrix}
\tag{16}
$$

where $Z_{i1}(t) \triangleq A^T(t-1)[G_{K_i}^L(t)]^T G_{K_i}^L(t)A(t-1) - P_i(t)$, $Z_{i2}(t) \triangleq A^T(t-1)[G_{K_i}^L(t)]^T B_{f_i}^L(t)$, and $Z_{i3}(t) \triangleq [B_{f_i}^L(t)]^T B_{f_i}^L(t) - \Theta_i(t)$. According to the Schur complement lemma [37], the first inequality in (11) is equivalent to $Z_i(t) < 0$. This means that $J_i(t) < 0$ under the first inequality in (11), and thus, one has

$$
e_i^T(t)e_i(t) < e_i^T(t-1)P_i(t)e_i(t-1)
$$
$$
+ \xi_i^T(t-1)\Theta_i(t)\xi_i(t-1).
\tag{17}
$$

When the inequality $P_i(t) - \vartheta_i(t)I < 0$ in (11) holds, one has $\lambda_{\max}(P_i(t)) < \vartheta_i(t)$. Then, combining (17) yields that

$$
e_i^T(t)e_i(t) < \vartheta_i(t)e_i^T(t-1)e_i(t-1)
$$
$$
+ \xi_i^T(t-1)\Theta_i(t)\xi_i(t-1).
\tag{18}
$$

Thus, it is derived from (18) that

$$
e_i^T(t)e_i(t) < \left(\prod_{\kappa=0}^{t-1}\vartheta_i(t-\kappa)\right)e_i^T(0)e_i(0)
$$
$$
+ \sum_{\kappa=0}^{t-1}\left\{\left(\prod_{\tau=0}^{\kappa-1}\vartheta_i(t-\tau)\right)\right.
$$
$$
\left. \times \xi_i^T(t-\kappa-1)\Theta_i(t-\kappa)\xi_i(t-\kappa-1)\right\}.
\tag{19}
$$

When the condition "$\vartheta_i(t) < 1$" in (11) holds, one has

$$
\lim_{t\to\infty}\prod_{\kappa=0}^{t-1}\vartheta_i(t-\kappa) = 0, \quad \lim_{\kappa\to\infty}\prod_{\tau=0}^{\kappa-1}\vartheta_i(t-\tau) = 0.
\tag{20}
$$

Then, it is concluded from (19–20) that $\lim_{t\to\infty} e_i^T(t)e_i(t)$ is bounded.

Notice that $\xi_i^{\mathrm{T}}(t-1)\Theta_i(t)\xi_i(t-1) \leq \lambda_{\max}(\xi_i(t-1)\xi_i^{\mathrm{T}}(t-1))\mathrm{Tr}\{\Theta_i(t)\}$ and, thus, it follows from (17) that

$$\mathrm{e}_i^{\mathrm{T}}(t)\mathrm{e}_i(t) < \vartheta_i(t)\mathrm{e}_i^{\mathrm{T}}(t-1)\mathrm{e}_i(t-1)$$
$$+ \lambda_{\max}(\xi_i(t-1)\xi_i^{\mathrm{T}}(t-1))\mathrm{Tr}\{\Theta_i(t)\}. \qquad (21)$$

In this case, the right term of (21) can be viewed as an upper bound of $\mathrm{e}_i^{\mathrm{T}}(t)\mathrm{e}_i(t)$ at each time. Though the estimation error $\mathrm{e}_i(t)$ is generated from the initial error $\mathrm{e}_i(0)$ and the bounded noises $\xi(0), \ldots, \xi(t-1)$, it is concluded from (18) and (19) that when the third condition in (11) holds, the estimation error is independent of the initial value. Based on the above consideration, "$\min \mathrm{Tr}\{\Theta_i(t)\}$" is proposed to be the optimization objective when minimizing this upper bound at each time.

Define $\mathrm{e}_0(t) \triangleq x(t) - \hat{x}(t)$. Then, the fusion estimation error is calculated by

$$\mathrm{e}_0(t) = \sum_{i=1}^{L} \Omega_i(t)\mathrm{e}_i(t) \qquad (22)$$

where $\mathrm{e}_i(t)$ is determined by (14). To design a group of optimal weighting matrices in (8), the following fusion system is constructed from the estimation error (14) and the fusion estimation error (22)

$$\begin{cases} \mathrm{e}_F(t) = A_F^L(t)\mathrm{e}_F(t-1) + B_F^L(t)\xi(t-1) \\ \mathrm{e}_0(t) = \Omega(t)\mathrm{e}_F(t) \end{cases} \qquad (23)$$

where $\mathrm{e}_F(t) \triangleq \mathrm{col}\{\mathrm{e}_1(t), \ldots, \mathrm{e}_L(t)\}$ and $\xi(t) \triangleq \mathrm{col}\{w(t), v_1(t+1), v_2(t+1), \ldots, v_L(t+1)\}$, while $A_F^L(t)$, $B_F^L(t)$ and $\Omega(t)$ are defined in (10) and (13), respectively. Define $\bar{\xi}(t) \triangleq \mathrm{col}\{\mathrm{e}_F(t), \xi(t)\}$, then, we introduce three matrices $\Upsilon(t)$, $P(t) > 0$ and $\Theta(t) > 0$ such that

$$\mathrm{e}_0^{\mathrm{T}}(t)\mathrm{e}_0(t) < \bar{\xi}^{\mathrm{T}}(t-1)\begin{bmatrix} P(t) & \Upsilon(t) \\ * & \Theta(t) \end{bmatrix}\bar{\xi}(t-1). \qquad (24)$$

To guarantee that the right term in (24) is an upper bound of $\mathrm{e}_0^{\mathrm{T}}(t)\mathrm{e}_0(t)$, the following inequality must be satisfied:

$$\bar{\xi}^{\mathrm{T}}(t-1)\underbrace{\begin{bmatrix} \Lambda_1(t) & \Lambda_2(t) \\ * & \Lambda_3(t) \end{bmatrix}}_{\Lambda(t)}\bar{\xi}(t-1) < 0 \qquad (25)$$

where $\Lambda_1(t) \triangleq [A_F^L(t)]^{\mathrm{T}}\Omega^{\mathrm{T}}(t)\Omega(t)A_F^L(t) - P(t)$, $\Lambda_2(t) \triangleq [A_F^L(t)]^{\mathrm{T}}\Omega^{\mathrm{T}}(t)\Omega(t)B_F^L(t) - \Upsilon(t)$, and $\Lambda_3(t) \triangleq [B_F^L(t)]^{\mathrm{T}}\Omega^{\mathrm{T}}(t)\Omega(t)B_F^L(t) - \Theta(t)$. Under this case, the first inequality in (13) is equivalent to $\Lambda(t) < 0$ according to the Schur complement lemma. Notice that $\mathrm{Tr}\{\begin{bmatrix} P(t) & \Upsilon(t) \\ * & \Theta(t) \end{bmatrix}\} = \mathrm{Tr}\{P(t) + \Theta(t)\}$ and, thus, it follows from (24) that

$$\mathrm{e}_0^{\mathrm{T}}(t)\mathrm{e}_0(t) < \lambda_{\max}(\bar{\xi}(t-1)\bar{\xi}^{\mathrm{T}}(t-1))\mathrm{Tr}\{P(t) + \Theta(t)\}. \qquad (26)$$

Since $\bar{\xi}(t-1)$ cannot be changed by the fusion system (23), $\min \mathrm{Tr}\{P(t) + \Theta(t)\}$ can be chosen as the optimization objective when designing optimal weighting matrices. ∎

Based on Theorem 1, the computation procedures for the DFE $\hat{x}(t)$ of linear systems (5)–(6) are summarized as follows.

Remark 4: Different from the constructing method of the upper bounds in [21], the matrices $\Theta_i(t)$ in (15), $\Upsilon(t)$, and $\Theta(t)$ in (24) are introduced to construct the upper bounds of the SEs of the LSEs and DFE in this paper. Notice that the estimation performance of the LSEs and DFE is strongly dependent on the constructed upper bounds. When establishing the convex optimization problems in Theorem 1,

---

Algorithm 1:

1: Determine the local estimator gains $\mathrm{K}_i^L(t)(i = 1, \ldots, L)$ by solving the optimization problem (11);
2: Determine the optimal weighting matrices $\Omega_i(t)(i = 1, \ldots, L)$ by solving the optimization problem (13);
3: Calculate linear LSEs $\hat{x}_i(t)(i = 1, \ldots, L)$ by (7);
4: Calculate the DFE $\hat{x}(t)$ by (8);
5: Return to Step 1 and implement Steps 1–4 for calculating $\hat{x}(t+1)$.

---

these introduced matrices can increase the search space and, thus, may lead to better solutions. On the other hand, the norm inequality was used in [21] to find an upper bound of the SEs and determine an optimization objective by a further inequality relaxation. Different from the relaxation inequality in [21], an upper bound of the SEs in this paper is constructed by the trace inequalities [see (21) and (26)], and the corresponding optimization objectives are also determined by the same trace inequalities. Notice that the constructed optimization objectives do not require further inequality relaxation. From the above analysis, the upper bound after optimizing the objective in Theorem 1 is closer to the real SE at each time. Therefore, the conservatism of the estimation method in Theorem 1 can be less than the estimation method in [21] because of these introduced new matrices, different relaxation inequalities, and optimization objectives.

B. DFE Design for Nonlinear Systems

Based on the DFE design method of linear systems in Section III.A, an optimal nonlinear estimator gain $\mathrm{K}_i^N(t)$ in (4) and optimal weighting matrices $\Omega_1(t), \ldots, \Omega_L(t)$ in (8) for nonlinear systems (1)–(2) will be presented in Theorem 2.

Theorem 2: Define

$$\begin{cases} A_{J_i}(t-1) = \left.\frac{\partial \mathrm{f}(\mathrm{x}(t-1))}{\partial \mathrm{x}(t-1)}\right|_{\mathrm{x}(t-1)=\hat{x}_i(t-1)} \\ C_{J_i}(t) = \left.\frac{\partial \mathrm{g}_i(\mathrm{x}(t))}{\partial \mathrm{x}(t)}\right|_{\mathrm{x}(t)=\mathrm{f}(\hat{x}_i(t-1))} \end{cases}. \qquad (27)$$

An optimal estimator gain $\mathrm{K}_i^N(t)$ in (4) can be obtained by solving the following convex optimization problem:

$$\min_{\Pi_i(t)>0, \Upsilon_i(t)>0, \mathrm{M}_i(t)>0, \eta_i(t)>0, \Psi_i(t), \mathrm{K}_i^N(t)} \mathrm{Tr}\{\Upsilon_i(t) + \mathrm{M}_i(t)\}$$

$$\mathrm{s.t.}: \begin{cases} \begin{bmatrix} -I & \mathrm{G}_{\mathrm{K}_i}^N(t)A_{J_i}(t-1) & \mathrm{X}_{i2}(t) & \mathrm{X}_{i3}(t) \\ * & -\Pi_i(t) & 0 & 0 \\ * & * & -\Upsilon_i(t) & -\Psi_i(t) \\ * & * & * & -\mathrm{M}_i(t) \end{bmatrix} < 0 \\ \Pi_i(t) - \eta_i(t)I < 0 \\ \eta_i(t) \leq 1 \end{cases} \qquad (28)$$

where

$$\begin{cases} \mathrm{G}_{\mathrm{K}_i}^N(t) = I - \mathrm{K}_i^N(t)\mathrm{C}_{J_i}(t) \\ \mathrm{X}_{i2}(t) \triangleq \mathrm{G}_{\mathrm{K}_i}^N(t)B(t-1) \\ \mathrm{X}_{i3}(t) \triangleq -\mathrm{K}_i^N(t)B_i(t) \end{cases}. \qquad (29)$$

Then, a group of optimal weighting matrices $\Omega_1(t), \ldots, \Omega_L(t)$ in (8) for nonlinear systems (1)–(2) can be obtained by solving the following

convex optimization problem:

$$\min_{\Pi(t)>0, \mathrm{M}(t)>0, \Psi(t), \Omega(t)} \mathrm{Tr}\{\Pi(t) + \mathrm{M}(t)\}$$

$$\text{s.t.}: \left\{ \begin{bmatrix} -I & \Omega(t)A_F^N(t) & \Omega(t)B_F^N(t) \\ * & -\Pi(t) & -\Psi(t) \\ * & * & -\mathrm{M}(t) \end{bmatrix} < 0 \right. \quad (30)$$

where $A_F^N(t)$, $B_F^N(t)$ and $\Omega(t)$ are defined by

$$\begin{cases} B_{f_i}^N(t) \triangleq [\mathrm{G}_{\mathrm{K}_i}^N(t)B(t-1) \quad -\mathrm{K}_i^N(t)B_i(t)] \\ A_F^N(t) \triangleq \mathrm{diag}\{\mathrm{G}_{\mathrm{K}_1}^N(t)A_{J_1}(t-1), \\ \qquad\qquad \cdots, \mathrm{G}_{\mathrm{K}_L}^N(t)A_{J_L}(t-1)\}. \\ B_F^N(t) \triangleq \mathrm{diag}\{B_{f_1}^N(t), \ldots, B_{f_L}^N(t)\} \\ \Omega(t) \triangleq [\Omega_1(t), \ldots, \Omega_{L-1}(t), I - \sum_{i=1}^{L-1}\Omega_i(t)] \end{cases} \quad (31)$$

Proof: Define $\tilde{\mathrm{x}}_i^{\mathrm{p}}(t) \triangleq \mathrm{f}(\mathrm{x}(t-1)) - \mathrm{f}(\hat{\mathrm{x}}_i(t-1)) + B(t-1)$ $w(t-1)$ and $\tilde{\mathrm{x}}_i(t) \triangleq \mathrm{x}(t) - \hat{\mathrm{x}}_i(t)$. Then, it follows from (4) that

$$\tilde{\mathrm{x}}_i(t) = \tilde{\mathrm{x}}_i^{\mathrm{p}}(t) - \mathrm{K}_i^N(t)[\mathrm{g}_i(\mathrm{x}(t)) - \mathrm{g}_i(\hat{\mathrm{x}}_i^{\mathrm{p}}(t)) + B_i(t)v_i(t)]. \quad (32)$$

Meanwhile, by expanding $\mathrm{f}(\mathrm{x}(t-1))$ and $\mathrm{g}_i(\mathrm{x}(t))$ in Taylor series about "$\hat{\mathrm{x}}_i(t-1)$" and "$\hat{\mathrm{x}}_i^{\mathrm{p}}(t)$", respectively, one has

$$\begin{cases} \mathrm{f}(\mathrm{x}(t-1)) = \mathrm{f}(\hat{\mathrm{x}}_i(t-1)) + A_{J_i}(t-1)\tilde{\mathrm{x}}_i(t-1) \\ \qquad\qquad + \Delta_{\mathrm{f}}(\tilde{\mathrm{x}}_i^2(t-1)) \\ \mathrm{g}_i(\mathrm{x}(t)) = \mathrm{g}_i(\hat{\mathrm{x}}_i^{\mathrm{p}}(t)) + C_{J_i}(t)\tilde{\mathrm{x}}_i^{\mathrm{p}}(t) + \Delta_{\mathrm{g}_i}([\tilde{\mathrm{x}}_i^{\mathrm{p}}(t)]^2) \end{cases} \quad (33)$$

where $A_{J_i}(t-1)$ and $C_{J_i}(t)$ are defined by (27), while $\Delta_{\mathrm{f}}(\tilde{\mathrm{x}}_i^2(t-1))$ and $\Delta_{\mathrm{g}_i}([\tilde{\mathrm{x}}_i^{\mathrm{p}}(t)]^2)$ represent the high-order terms of the Taylor series expansion.

According to (33), the nonlinear estimation error system (32) is equivalent to

$$\tilde{\mathrm{x}}_i(t) = \mathrm{G}_{\mathrm{K}_i}^N(t)A_{J_i}(t-1)\tilde{\mathrm{x}}_i(t-1)$$

$$+ \mathrm{G}_{\mathrm{K}_i}^N(t)[B(t-1)w(t-1) + \Delta_{\mathrm{f}}(\tilde{\mathrm{x}}_i^2(t-1))]$$

$$- \mathrm{K}_i^N(t)[B_i(t)v_i(t) + \Delta_{\mathrm{g}_i}([\tilde{\mathrm{x}}_i^{\mathrm{p}}(t)]^2)] \quad (34)$$

where $\mathrm{G}_{\mathrm{K}_i}^N(t)$ is defined by (29). Notice that $\Delta_{\mathrm{f}}(\tilde{\mathrm{x}}_i^2(t-1))$ and $\Delta_{\mathrm{g}_i}([\tilde{\mathrm{x}}_i^{\mathrm{p}}(t)]^2)$ in (34) are unknown noises, while $w(t-1)$ and $v_i(t)$ are also unknown noises. Under this case, the terms $B(t-1)\tilde{w}_i(t-1)$ and $B_i(t)\tilde{v}_i(t)$ are introduced to model the affection factors caused by unknown noises. Then, (34) can be written as

$$\tilde{\mathrm{x}}_i(t) = \mathrm{G}_{\mathrm{K}_i}^N(t)A_{J_i}(t-1)\tilde{\mathrm{x}}_i(t-1)$$

$$+ \mathrm{G}_{\mathrm{K}_i}^N(t)[B(t-1)\tilde{w}_i(t-1)] - \mathrm{K}_i^N(t)[B_i(t)\tilde{v}_i(t)]. \quad (35)$$

Since the form of (35) is the same as that of (14), the optimization problems (28) and (30) in Theorem 2 can be obtained by the similar proof in Theorem 1, and the detailed derivation is omitted due to the page limitation.

Based on Theorem 2, the computation procedures for the DFE $\hat{\mathrm{x}}(t)$ of nonlinear systems (1)–(2) are summarized as follows.

Remark 5: The optimization problems (11), (13), (28), and (30) are established in terms of linear matrix inequalities, and thus they can be directly solved by the function "*mincx*" of MATLAB LMI Toolbox [37]. On the other hand, it can be concluded from (22) that when each linear LSE is designed by (11), the SE of the DFE for the linear time-varying systems (5)–(6) must be bounded at each time.

Algorithm 2:

1: Calculate the matrices $A_{J_i}(t-1)$ and $C_{J_i}(t)$ by (27);
2: Determine the local estimator gains $\mathrm{K}_i^N(t)(i = 1, \ldots, L)$ by solving the optimization problem (28);
3: Determine the optimal weighting matrices $\Omega_i(t)(i = 1, \ldots, L)$ by solving the optimization problem (30);
4: Calculate nonlinear LSEs $\hat{\mathrm{x}}_i(t)(i = 1, \ldots, L)$ by (4);
5: Calculate the nonlinear DFE $\hat{\mathrm{x}}(t)$ by (8);
6: Return to Step 1 and implement Steps 1–5 for calculating $\hat{\mathrm{x}}(t+1)$.

IV. SIMULATION EXAMPLES

A. Target Tracking System

Consider a maneuvering target, which is monitored by two sensors, and define the state vector $x(t)$ by $x(t) \triangleq \mathrm{col}\{s(t), \dot{s}(t)\}$, where $s(t)$ is the target's position, and $\dot{s}(t)$ is the target's velocity. Then, the dynamical process of the target's position and velocity can be modeled by [1]

$$x(t+1) = \begin{bmatrix} 1 & f_s(t) \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0.5f_s^2(t) \\ f_s(t) \end{bmatrix} w(t) \quad (36)$$

where $f_s(t)$ is the time-varying sampling period, and $w(t)$ is the process noise. Then, each sensor's measurement is modeled by

$$y_i(t) = C_i(t)x(t) + B_i(t)v(t) \quad (37)$$

where $v(t)$ is the measurement noise, and $C_1(t) = [0.5\ 1], C_2(t) = [1\ 0], B_1(t) = 1.2\cos(f_s(t)), B_2(t) = 2.0\sin(f_s(t))$. In the simulation, three types of disturbance noises $w(t)$ and $v(t)$ in (36) and (37) will be considered

1) **Type I:** When $f_s(t) = 0.5$, $w(t)$ and $v(t)$ are the *energy-bounded noises* given by $w(t) = (2 + 0.2\cos(t))\exp(-t/9)$ and $v(t) = 0.8\sin(t)\exp(-t/6)$.
2) **Type II:** When $f_s(t) = 0.5 + 0.2\sin(t)$, $w(t)$ and $v(t)$ are the uncorrected *Gaussian white noises* with covariances $Q_w = 1.8$ and $Q_v = 0.5$.
3) **Type III:** When $f_s(t) = 0.5 + 0.2\sin(t)$, $w(t)$ and $v(t)$ are the *bounded noises* given by $w(t) = \cos(t) - 0.5$ and $v(t) = 0.7\sin(t) - 0.3$.

Notice that, for the system (36)–(37), each LSE can be obtained by the H_∞ filter in [23] under Type I, the well-known Kalman filtering method (see [22]) can be used to design each LSE under Type II, and each LSE can be obtained by Theorem 1 in [21] under Type III.

To demonstrate the advantages of the designed estimation algorithm, the estimation performances of the first LSEs are shown in Fig. 1 by using different estimation methods. Since energy-bounded noise is a special case of bounded noises, the LSE under Type I can also be obtained by Theorem 1 in this paper. It is seen from Fig. 1(a) that the estimation precision of the LSE in this paper is higher than that of the H_∞ filtering in [23]. Meanwhile, the Gaussian white noise is always bounded in a practical system, thus it can be viewed as a class of bounded noises. Under this case, the designed LSE in this paper can be applicable to Type II. Then, it is seen from Fig. 1(b) that the estimation precision of the LSE given by Theorem 1 is close to that of the Kalman filter [22]. Due to the random noises, the estimation performance of the LSE is assessed by its mean square error (MSE), and the Monte Carlo method is used to approach the theory MSE. Then, the practical MSEs (PMSEs) of the LSE in this paper and the Kalman filter with accurate/inaccurate covariances are shown in Fig. 1(c). It is seen
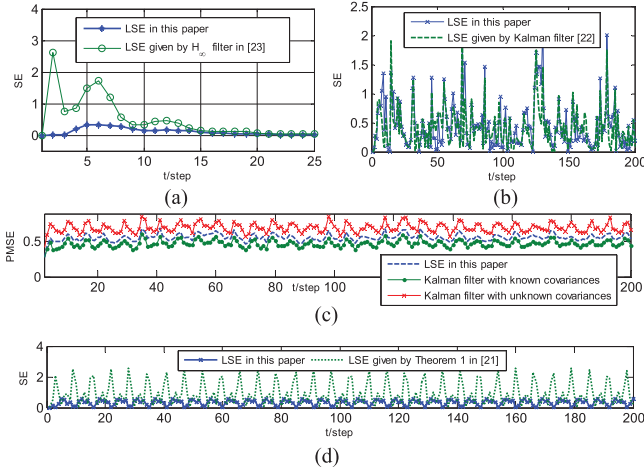
Fig. 1. According to the first sensor measurements, the corresponding LSE is designed by using different methods. (a) Under the Type I, SEs of the H_∞ filter in [23] and the LSE in this paper. (b) Under the Type II, SEs of the Kalman filter [22] and the LSE in this paper. (c) Under the Type II, PMSEs of the LSE in this paper and the Kalman filters with known covariances and unknown covariances. (d) Under the Type III, SEs of the LSE in [21], and the LSE in this paper.

from this figure that when the covariances are known, the estimation performance of Kalman filter is better than that of the LSE in this paper. This is because the Kalman filter is designed in the linear minimum variance sense at each time. On the other hand, when the covariances are unknown or inaccurately known, the estimation precision of the LSE in this paper is higher than that of the Kalman filter. This is because the Kalman filter is required to know the covariance. In contrast, the LSE designed by this paper is not required to know the statistical information of noises. The above discussion implies that the designed LSE in this paper is more robust, and is applicable to a more general case.

Furthermore, when considering the bounded noises, Fig. 1(d) shows that the estimation performance of the LSE in this paper is better than that of the LSE in [21]. This is because more available information is used to design estimator under Kalman-like structure (see Remark 1), and new upper bound and optimization objective constructed in Theorem 1 can reduce the conservatism (see Remark 4). Moreover, when considering the second measurement equation (i.e., $C_2(t) = [1\ 0]$), the optimization problem in Theorem 1 of [21] is unsolvable, while the optimization problem (11) is solvable.

B. Mobile Robot Localization

Consider the localization of an unicycle mobile robot operating in planar environments. Let $p_r(t) \triangleq \mathrm{col}\{s_x(t), s_y(t)\}$ denote the robot's position in X–Y plane, while $\theta(t)$ is to define the angular orientation. Then, the motion model of the robot is given by [38]

$$
\begin{cases}
s_x(t+1) = s_x(t) - \frac{\hat{u}_p}{\hat{u}_r}(\sin\theta(t) - \sin(\theta(t) + T_0\hat{u}_r)) \\
s_y(t+1) = s_y(t) + \frac{\hat{u}_p}{\hat{u}_r}(\cos\theta(t) - \cos(\theta(t) + T_0\hat{u}_r)) \\
\theta(t+1) = \theta(t) + T_0\hat{u}_r + T_0 w_\theta(t) \\
\hat{u}_p = u_p + w_p(t) \\
\hat{u}_r = u_r + w_r(t)
\end{cases}
\tag{38}
$$

where T_0 is the sampling period, $w_\theta(t)$ is the additional rotational noise; u_p is the motion command to control the translational velocity, while u_r is the motion command to control the rotational velocity. As

pointed out in [38], *robot motion is subject to noise in reality*, i.e., the motion commands u_p and u_r may be changed by the unpredictable disturbances. Then, the true velocity control input $\hat{u}_p$ (or $\hat{u}_r$) equals the commanded velocity plus some small, additive noise $w_p(t)$ (or $w_r(t)$). Notice that the motion commands u_p and u_r are constant and known for the robot model (38) and, thus, the model (38) can be written as

$$
\mathrm{x}(t+1) = \mathrm{f}(\mathrm{x}(t)) + \Gamma w(t)
\tag{39}
$$

where $\mathrm{x}(t) \triangleq \mathrm{col}\{s_x(t), s_y(t), \theta(t)\}$, $\Gamma \triangleq \mathrm{diag}\{1, 1, T_0\}$, $w(t) \triangleq \mathrm{col}\{w_1(t), w_2(t), w_3(t)\}$, and

$$
\begin{cases}
\mathrm{f}(\mathrm{x}(t)) = \begin{bmatrix} s_x(t) - \frac{u_p}{u_r}(\sin\theta(t) - \sin(\theta(t) + T_0 u_r)) \\ s_y(t) + \frac{u_p}{u_r}(\cos\theta(t) - \cos(\theta(t) + T_0 u_r)) \\ \theta(t) + T_0 u_r \end{bmatrix} \\
w_1(t) = \left(\frac{u_p}{u_r} - \frac{\hat{u}_p}{\hat{u}_r}\right)\sin\theta(t) + \frac{\hat{u}_p}{\hat{u}_r}\sin(\theta(t) + T_0\hat{u}_r) \\
\qquad\quad - \frac{u_p}{u_r}\sin(\theta(t) + T_0 u_r) \\
w_2(t) = \left(\frac{\hat{u}_p}{\hat{u}_r} - \frac{u_p}{u_r}\right)\cos\theta(t) + \frac{u_p}{u_r}\cos(\theta(t) + T_0 u_r) \\
\qquad\quad - \frac{\hat{u}_p}{\hat{u}_r}\cos(\theta(t) + T_0\hat{u}_r) \\
w_3(t) = w_r(t) + w_\theta(t)
\end{cases}
\tag{40}
$$

Since $w_1(t)$ and $w_2(t)$ are dependent on the state $\theta(t)$, they are state-dependent noises for the system (39).

In X–Y plane, it is considered that four known points, denoted as $(s_{x_i}, s_{y_i})(i = 1, 2, 3, 4)$, are chosen as the landmarks. Then, the distance from the robot's planner Cartesian coordinates $(s_x(t), s_y(t))$ to each landmark (s_{x_i}, s_{y_i}) can be expressed as follows

$$
d_i(t) = \sqrt{(s_{x_i} - s_x(t))^2 + (s_{y_i} - s_y(t))^2}.
\tag{41}
$$

The azimuth $\varphi_i(t)$ at time t can be related to the current system state variables $s_x(t)$, $s_y(t)$, and $\theta(t)$ as follows

$$
\varphi_i(t) = \theta(t) - \arctan\left(\frac{s_{y_i} - s_y(t)}{s_{x_i} - s_x(t)}\right).
\tag{42}
$$

Both the distance $d_i(t)$ and $\varphi_i(t)$ are treated as the measurements. Furthermore, when considering the unpredicted disturbances, the measurement equations for the robotic system (38) can be written as follows

$$
\begin{cases}
y_1(t) = \begin{bmatrix} \mathrm{g}_1(\mathrm{x}(t)) \\ \mathrm{g}_2(\mathrm{x}(t)) \end{bmatrix} + \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \\
y_2(t) = \begin{bmatrix} \mathrm{g}_3(\mathrm{x}(t)) \\ \mathrm{g}_4(\mathrm{x}(t)) \end{bmatrix} + \begin{bmatrix} D_3 \\ D_4 \end{bmatrix} v_3(t)
\end{cases}
\tag{43}
$$

where $v_1(t)$, $v_2(t)$ and $v_3(t)$ are the measurement noises, and

$$
\begin{cases}
\mathrm{g}_i(\mathrm{x}(t)) = \mathrm{col}\{d_i(t), \varphi_i(t)\}(i = 1, 2, 3, 4) \\
D_1 = \mathrm{diag}\{0.5, 0.3\}, D_2 = \mathrm{diag}\{0.3, 0.5\}. \\
D_3 = \mathrm{diag}\{0.2, 0.6\}, D_4 = \mathrm{diag}\{0.5, 0.7\}
\end{cases}
\tag{44}
$$

Then, based on the sensor measurements (43), the localization of this mobile robot can be realized by using different nonlinear estimation methods. On the other hand, by using Taylor series expansions, the linearized matrices $A_{J_j}(t)$ and $C_{J_j}(t)(j = 1, 2)$ for the nonlinear vector functions $\mathrm{f}(\mathrm{x}(t))$ in (39) and $\mathrm{g}_i(\mathrm{x}(t))$ in (43) near the point $x^* \in \mathrm{R}^3$

can be expressed as follows

$$
\begin{cases}
A_{J_j}(t) = \begin{bmatrix} 1 & 0 & -\frac{u_p}{u_r}\cos\theta(t) + \frac{u_p}{u_r}\cos(\theta(t)+T_0 u_r) \\ 0 & 1 & -\frac{u_p}{u_r}\sin\theta(t) + \frac{u_p}{u_r}\sin(\theta(t)+T_0 u_r) \\ 0 & 0 & 1 \end{bmatrix}_{x(t)=x^*} \\[1em]
D_{J_i}(t) = \begin{bmatrix} \frac{-\tilde{s}_{x_i}(t)}{\sqrt{\tilde{s}_{x_i}^2(t)+\tilde{s}_{y_i}^2(t)}} & \frac{-\tilde{s}_{y_i}(t)}{\sqrt{\tilde{s}_{x_i}^2(t)+\tilde{s}_{y_i}^2(t)}} & 0 \\ \frac{-\tilde{s}_{y_i}(t)}{\tilde{s}_{x_i}^2(t)+\tilde{s}_{y_i}^2(t)} & \frac{\tilde{s}_{x_i}(t)}{\tilde{s}_{x_i}^2(t)+\tilde{s}_{y_i}^2(t)} & 1 \end{bmatrix}_{x(t)=x^*} \\[1em]
C_{J_1}(t) \triangleq \mathrm{col}\{D_{J_1}(t), D_{J_2}(t)\} \\
C_{J_2}(t) \triangleq \mathrm{col}\{D_{J_3}(t), D_{J_4}(t)\}
\end{cases}
\tag{45}
$$

where $i = 1, 2, 3, 4$, $\tilde{s}_{x_i}(t) \triangleq s_{x_i} - s_x(t)$, and $\tilde{s}_{y_i}(t) \triangleq s_{y_i} - s_y(t)$.

In the simulation, the parameters T_0, u_p, and u_r are taken as $T_0 = 1$, $u_p = 0.075$, and $u_r = 0.025$, while four landmarks' positions in X–Y plane are $(s_{x_1}, s_{y_1}) = (5, 10)$, $(s_{x_2}, s_{y_2}) = (10, 10)$, $(s_{x_3}, s_{y_3}) = (10, 5)$, and $(s_{x_4}, s_{y_4}) = (5, 5)$. Meanwhile, it is reasonably considered that the initial robot's pose is known in advance, hence the initial estimation errors $\tilde{x}_1(0)$ and $\tilde{x}_2(0)$ can be given by $\tilde{x}_1(0) = \tilde{x}_2(0) = 0$. Furthermore, the disturbance noises $w_p(t), w_r(t)$, $w_\theta(t)$, and $v_k(t)$ $(k = 1, 2, 3)$ in (38)–(39) are taken as follows.

1) **Type IV:** $w_p(t)$, $w_r(t)$, $w_\theta(t)$, and $v_k(t)$ $(k = 1, 2, 3)$ are the bounded noises given by

$$
\begin{cases}
w_p(t) = 0.2\rho_p(t) - 0.1, w_r(t) = 0.3\rho_r(t) - 0.1 \\
w_\theta(t) = 0.2\rho_\theta(t) - 0.1 \\
v_1(t) = \mathrm{col}\{0.05\rho_{v_1}(t) - 0.01, 0.02\rho_{v_2}(t) - 0.01\} \\
v_2(t) = \mathrm{col}\{0.03\rho_{v_3}(t) - 0.01, 0.05\rho_{v_4}(t) - 0.03\} \\
v_3(t) = \mathrm{col}\{0.02\rho_{v_5}(t) - 0.01, 0.06\rho_{v_6}(t) - 0.02\}
\end{cases}
\tag{46}
$$

where $\rho_p(t)(\in [0, 1])$, $\rho_r(t)(\in [0, 1])$, $\rho_\theta(t)(\in [0, 1])$, and $\rho_{v_l}(t)(\in [0, 1])$ $(l = 1, 2, 3, 4, 5, 6)$ are random variables that can be generated by the function "rand" of MATLAB. Under this case, $w(t)$ in (39) is also the bounded noise.

To demonstrate the effectiveness, by implementing Algorithm 2, the actual robot trajectory in the X–Y plane and its position estimation are plotted in Fig. 2(a) under Type IV. It is seen from this figure that the mobile robot can get its position well by using the nonlinear fusion estimation algorithm in this paper. Meanwhile, the SEs of the DFE and LSEs are depicted in Fig. 2(c) under Type IV, which shows that the estimation performance of the DFE is better than that of each LSE. This is as expected for the fusion estimation methods. On the other hand, when the optimization problem (28) is solvable at each time, there must be $J_{d_j}(t) \triangleq \|G_{K_j}^N(t) A_{J_j}(t-1)\|_2 < 1$, which has been illustrated by Fig. 2(b).

To demonstrate the advantages of the developed nonlinear estimation method, it will be compared with the classical EKF method [25] and UKF method [27], [28], where these methods are all applicable to nonlinear systems with *Gaussian white noises*, and the UKF in [27] can also deal with state-dependent noises by using augmentation strategy. Here, the performance of the first LSE is used to show the advantages of the proposed method. Due to the random noises in Type IV, the estimation performance of the LSE is assessed by the PMSEs that are calculated by Monte Carlo method [1] with an average of 500 runs. Then, the PMSEs of the first LSE using different methods are plotted in Fig. 3. It is seen from this figure that the estimation precision of the proposed method is higher than that of EKF or UKF method. This is because the disturbance noises are bounded, and no statistical
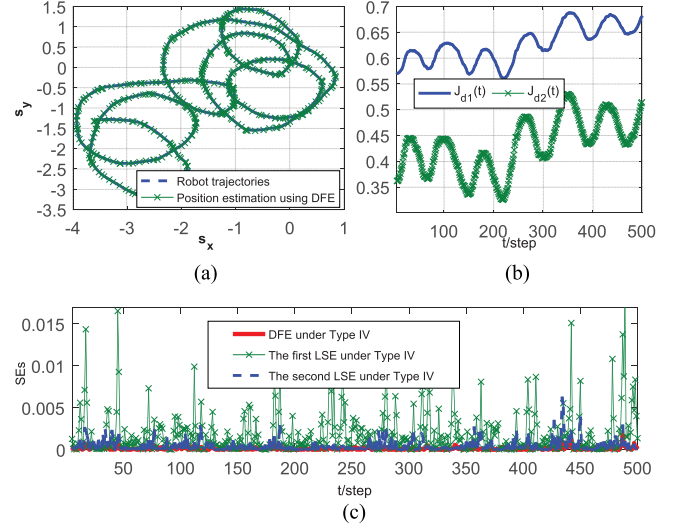

Fig. 2. Nonlinear fusion estimation under Type IV. (a) The robot's trajectory and its position estimation by using Algorithm 2. (b) Trajectories of $J_{d_j}(t)(j = 1, 2)$. (c) Comparison of the estimation precision for the DFE and LSEs under Type IV.
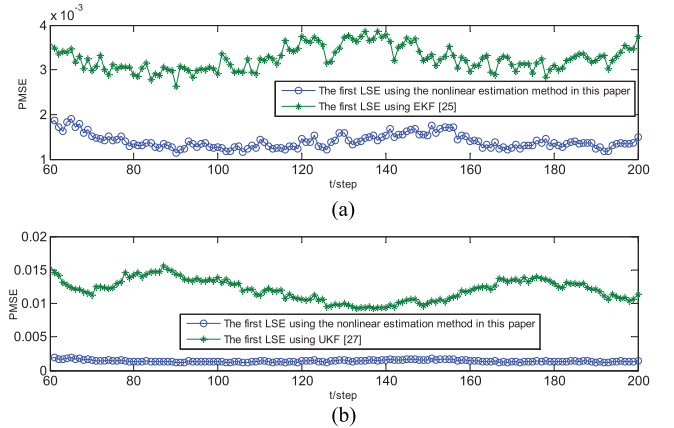

Fig. 3. Under Type IV, the comparison of the first LSE's estimation performance by using the estimation method of this paper, EKF method [25] and UKF method [27], respectively.

information is available, but the EKF and UKF are all used to deal with the case of Gaussian white noises with known covariances. Notice that, as compared with Gaussian white noise with known covariance, the condition "bounded noise" is easier to be satisfied in a practical system. This implies that the nonlinear estimation method in this paper is more robust as compared with the classic EKF and UKF methods.

V. CONCLUSION

In this paper, a new method to distributed fusion estimation problem has been developed for linear time-varying systems and nonlinear systems with bounded noises. When considering linear time-varying fusion systems, each local Kalman-like estimator with time-varying gains was designed such that the SE of each LSE must be bounded as time goes to ∞, and a novel distributed fusion estimation criterion was designed by establishing a convex optimization problem. Furthermore, the general nonlinear systems were transformed to linear time-varying

systems by using Taylor series expansion, and the linearized errors could be viewed as unknown but bounded noises. Under this case, different convex optimization problems on the designs of the nonlinear estimator and distributed fusion criterion were established in terms of linear matrix inequalities for nonlinear fusion systems with bounded noises. Moreover, the solutions to the convex optimization problems in this paper can be directly obtained by using the MATLAB LMI Toolbox. Finally, two illustrative examples were exploited to demonstrate the advantages and effectiveness of the proposed fusion estimation methods.

More recently, a great deal of attention has focused on the networked multisensor fusion estimation problem, where the sensor messages are transmitted to the fusion center through communication networks [4], [19], [21]. Therefore, when considering the communication uncertainties including bandwidth constraints, transmission delays and packet dropouts, one of our future works will focus on how to design the networked nonlinear fusion estimation algorithms based on the developed fusion estimation method in this paper.

REFERENCES

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. Hoboken, NJ, USA: Wiley, 2001.

[2] A. Okatan, C. Hajiyev, and U. Hajiyeva, "Fault detection in sensor information fusion Kalman filter," *Int. J. Electron. Commun.*, vol. 63, 2009, pp. 762–768.

[3] W. Zhang and L. Shi, "Sequential fusion estimation for clustered sensor networks," *Automatica*, vol. 89, 2018, pp. 358–363.

[4] B. Chen, G. Hu, D. W. C. Ho, and L. Yu, "Distributed covariance intersection fusion estimation for cyber-physical systems with communication constraints," *IEEE Trans. Autom. Control*, vol. 61, no. 12, pp. 4020–4026, Dec. 2016.

[5] J. A. Roecker and C. D. McGillen , "Comparison of two-sensor tracking methods based on state vector fusion and measurement fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 447–449, Jul. 1988.

[6] X. Li, Y. Zhu, J. Wang, and C. Han, "Optimal linear estimation fusion-part I: Unified fusion rules," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2192–2208, Sep. 2003.

[7] Z. Duan and X. Li, "Lossless linear transformation of sensor data for distributed estimation fusion," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 362–372, Jan. 2011.

[8] E. Song, J. Xu, and Y. Zhu, "Optimal distributed Kalman filtering fusion with singular covariances of filtering errors and measurement noises," *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1271–1282, May 2014.

[9] N. A. Carlson, "Federated square root filter for decentralized parallel processes," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 3, pp. 517–525, May 1990.

[10] S. Sun and Z. Deng, "Muti-sensor optimal information fusion Kalman filter," *Automatica*, vol. 40, pp. 1017–1023, 2004.

[11] Z. Deng, Y. Gao, L. Mao, Y. Li, and G. Hao, "New approach to information fusion steady-state Kalman filtering," *Automatica*, vol. 41, pp. 1695–1707, 2005.

[12] S. J. Julier and J. K. Uhlmann, "General decentralized data fusion with covariance intersection (CI)," In *Handbook of Multisensor Data Fusion*. Boca Raton, FL, USA: CRC Press, 2001, ch. 12, pp. 1–25.

[13] Y. Wang and X. Li, "Distributed estimation fusion with unavaliable cross correlation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 259–278, Jan. 2012.

[14] J. Sijs and M. Lazar, "State fusion with unknown correlation: Ellipsoidal intersection," *Automatica*, vol. 48, pp. 1874–1878, 2012.

[15] B. Noack, J. Sijs, M. Reinhardt, and U. D. Hanebeck, "Decentralized data fusion with inverse covariance intersection," *Automatica*, vol. 79, pp. 35–42, 2017.

[16] Y. Liang, T. Chen, and Q. Pan, "Multi-rate stochastic H_∞ filtering for networked multi-sensor fusion," *Automatica*, vol. 46, pp. 437–444, 2010.

[17] M. Liu, M. Qiu, S. Zhang, and Z. Lin, "Robust H_∞ fusion filters for discrete-time nonlinear delayed systems with missing measurements," in *Proc. Amer. Control Conf.*, MD, USA, 2010, pp. 6803–6808.

[18] B. Chen, L. Yu, W. Zhang, and H. Wang, "Distributed H_∞ fusion filtering with communication bandwidth constraints," *Signal Process.*, vol. 96, pp. 284–289, 2014.

[19] B. Chen, G. Hu, W. Zhang, and L. Yu, "Distributed mixed H_2/H_∞ fusion estimation with limited communication capacity," *IEEE Trans. Autom. Control*, vol. 61, no. 3, pp. 805–810, Mar. 2016.

[20] B. Chen, G. Hu, D. W. C. Ho, W. Zhang, and L. Yu, "Distributed robust fusion estimation with application to state monitoring systems," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 47, no. 11, pp. 2994–3005, Nov. 2017.

[21] B. Chen, D. W. C. Ho, W. Zhang, and L. Yu, "Networked fusion estimation with bounded noises," *IEEE Trans. Autom. Control*, vol. 62, no. 10, pp. 5415–5421, Oct. 2017.

[22] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME Series D, J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.

[23] R. M. Palhares and P. L. D. Peres, "Optimal filtering schemes for linear discrete-time systems: A linear matrix inequality approach," *Int. J. Syst. Sci.*, vol. 29, no. 6, pp. 587–893, 1998.

[24] A. G. O. Mutambara, *Decentralized Estimation and Control for Multi-Sensor Systems*. Boca Raton, FL, USA: CRC Press, 1998.

[25] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.

[26] D. J. Lee, "Nonlinear estimation and multiple sensor fusion using unscented information filtering," *IEEE Signal Process. Lett.*, vol. 15, pp. 861–864, 2008.

[27] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Proc. 11th Int. Symp. Aerosp./Def. Sens. Simul. Controls*, 1997, vol. 3, no. 26, pp. 182–193.

[28] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[29] Q. Ge, T. Shao, Q. Yang, X. Shen, and C. Wen, "Multi-sensor nonlinear fusion methods based on adaptive ensemble fifth-degree iterated cubature information filter for biomechatronics," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 46, no. 7, pp. 912–925, Jul. 2016.

[30] S. Shamshirband, D. Petković, H. Javidnia, and A. Gani, "Sensor data fusion by support vector regression methodology-a comparative study," *IEEE Sens. J.*, vol. 15, no. 2, pp. 850–854, Feb. 2015.

[31] X. Yang, W. Zhang, M. Z. Q. Chen, and L. Yu, "Hybrid sequential fusion estimation for asynchronous sensor network-based target tracking," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 2, pp. 669–676, Mar. 2017.

[32] X. Yang, W. Zhang, B. Chen, and L. Yu, "A bank of sequential unscented Kalman filters for target tracking in range-only WSNs," in *Proc. 13th IEEE Int. Conf. Control Autom.*, Ohrid, Macedonia, 2017, pp. 1–5.

[33] Y. Ruan, P. Willett, A. Marrs, F. Palmierl, and S. Marano, "Practical fusion of quantization measurements via particle filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 1, pp. 15–29, Jan. 2008.

[34] A. Mohammadi and A. Asif, "Distributed consensus + innovation particle filtering for bearing/range tracking with communication constraints," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 620–635, Feb. 2015.

[35] D. Spinello and D. J. Stilwell, "Nonlinear estimation with state-dependent gaussian observation noise," *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1358–1366, Jun. 2010.

[36] X. Hu, M. Bao, X. P. Zhang, L. Guan, and Y. H. Hu, "Generalized iterated Kalman filter and its performance evaluation," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3204–3217, Jun. 2015.

[37] S. P. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.

[38] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2005.