COMMUNICATIONS IN INFORMATION AND SYSTEMS
Vol. 12, No. 1, pp. 97-130, 2012
© 2012 International Press
005

FROM KALMAN FILTERING TO SET-VALUED FILTERING FOR DYNAMIC SYSTEMS WITH UNCERTAINTY*

YUNMIN ZHU†

Abstract. In this paper, a brief survey is dedicated to the developments from Kalman filtering for the stochastic dynamic systems to set-valued filtering for the dynamic systems with bounded uncertain model biases. The former has been developed and extended for more than fifty years, and the set-valued estimations have been proposed also for more than forty years but has received more extensive attention than before for the dynamic systems with bounded uncertain model biases just in the past twenty years. They are two types of estimations in terms of completely different optimization criterions to deal with different formulations of dynamic systems with uncertainty. The main focus on this survey is to present some progress in the two filters and compare their own advantage and weakness in order to provide some guidance for people to decide which formulation for dynamic systems with uncertainty and the corresponding filtering method should be chosen in practical applications.

Keywords. Kalman filtering, set-valued estimation, uncertain model.

1. Introduction. The state estimation/filtering of a dynamic system is a very significant problem in many practical applications, such as space technology, military engineering, communications engineering, industrial control, meteorology, biomedical and bioinformatics, economics and sociology, and so on. Obviously, in doing so, there are two fundamental issues to solve: how to formulate the dynamic systems under consideration and provide the corresponding estimation methods in an appropriate optimization criteria. Obviously, estimating a state using all available observations can derive globally optimal state estimate. Unfortunately, it is not realistic in real-world applications since the resulting computational burden becomes heavier and heavier as the observations are getting more and more along with time goes by. Thus, available filtering methods for dynamic systems must be recursive. So far, there are two main formulations for uncertain dynamic systems. The first one, which has been investigated for more than fifty years, is stochastic dynamic systems with random model noises–process and observation noises [1, 2, 9, 14, 15, 19, 20, 21, 25, 41, 46, 67, 80]. The second one, which has been proposed also for more than forty years [5, 6, 7, 8, 12] but has received more extensive attention than before just for recent twenty years [36, 44, 51, 52, 59, 62], is dynamic systems with bounded uncertain model biases. For the stochastic dynamic systems, a natural optimization criterion

*Dedicated to Professor Hanfu Chen on the occasion of his 75th birthday. This work was supported in part by the Natural Science Foundation of China under Grants 60934009 & 61273074.

† The authors are with the College of Mathematics, Sichuan University, Chengdu, Sichuan 610064, China. (E-mail: ymzhu@scu.edu.cn)

to state estimation is the Mean Square Error (MSE) or its simplified version Linear Mean Square Error (LMSE), and the optimal algorithms to the MSE (when the state and observation are jointly Gaussian) or LMSE solutions are Kalman filtering and its various extensions given in the above cited literature, in particular, see books [67, 80]. For the dynamic systems with bounded uncertain model biases, the set-valued filtering based on convex optimization is developed, and then Minimizing Euclidean Error (MEE) filtering based on the set-valued filtering has been proposed in [77, 80].

In this paper, a brief survey is dedicated to the developments form Kalman filtering for the stochastic dynamic systems to set-valued filtering for the dynamic systems with bounded but uncertain model biases. They are actually two types of filtering in terms of completely different optimization criterions to deal with different formulations of dynamic systems with uncertainty. The main focus on this survey is to present some progress in the two filtering and compare their own advantage and weakness in order to provide some guidance for people to decide which formulation for dynamic systems with uncertainty and the corresponding filtering method should be chosen in practical applications.

The rest of the paper is organized as follows. Section 2 presents the formulation and basic results of Kalman filtering. Section 3 is dedicated to some extensions to the dynamic systems with correlated noises, random coefficient matrices, and spurious or negative-time observations and their some applications. In section 4, main limitations to Kalman filtering in practice are briefly listed. Section 5.1 presents set-membership filtering and related researches. A little more details are given here including ideas of minimizing Euclidean estimation errors with illustrations, step-by-step formulation and theoretical conclusions. A numerical example is provided in Section 5.2, and a brief comparison with Kalman filtering is given in Section 5.3. Section 6 concludes the paper.

2. Kalman Filtering Based on LMSE for Stochastic Dynamic Systems. The first type of dynamic systems is stochastic dynamic systems modeled as

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}, \tag{1}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k, \tag{2}$$

where, $k = 1, 2, \ldots$ represents the time instant, $\mathbf{x}_k$ is $r \times 1$ vector-valued state to be estimated at time instant k, $\mathbf{y}_k$ is $n \times 1$ vector-valued observation of state, $\mathbf{F}_k$ and $\mathbf{H}_k$ are deterministic $r \times r$ and $n \times r$ matrices respectively, and $\{\mathbf{v}_k\}$ and $\{\mathbf{w}_k\}$ are, respectively, process and observation noise sequences, and satisfy the following assumptions:

1. The noise sequences $\{\mathbf{v}_k\}$ and $\{\mathbf{w}_k\}$ are white and uncross-correlated, i.e.,

for any $k, l = 0, 1, \ldots,$

$$\begin{aligned}
&\mathbb{E}(\mathbf{v}_k) = \mathbf{0},\ \mathrm{Cov}(\mathbf{v}_k\mathbf{v}_l) = \mathbb{E}(\mathbf{v}_k\mathbf{v}_l^{\mathrm{T}}) = \mathrm{Var}\ (\mathbf{v}_k)\delta_{kl},\\
&\mathbb{E}(\mathbf{w}_k) = \mathbf{0},\ \mathrm{Cov}(\mathbf{w}_k\mathbf{w}_l) = \mathbb{E}(\mathbf{w}_k\mathbf{w}_l^{\mathrm{T}}) = \mathrm{Var}\ (\mathbf{w}_k)\delta_{kl},\\
&\mathrm{Cov}(\mathbf{v}_k\mathbf{w}_l) = \mathbb{E}(\mathbf{v}_k\mathbf{w}_l^{\mathrm{T}}) = \mathbf{0};
\end{aligned}$$

2. The expectation and variance matrix of the initial state $\mathbf{x}_0$ are known, and have the following statistical properties:

$$\mathrm{Cov}(\mathbf{x}_0\mathbf{v}_k) = \mathbb{E}(\mathbf{x}_0\mathbf{v}_k^{\mathrm{T}}) = \mathbf{0},\ \mathrm{Cov}(\mathbf{x}_0\mathbf{w}_k) = \mathbb{E}(\mathbf{x}_0\mathbf{w}_k^{\mathrm{T}}) = \mathbf{0},\ k = 1, 2, \ldots$$

For any instant $j > 0$, we want to use all observations from beginning up to time instant j

$$\mathbf{Y}_j = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_j \end{bmatrix}$$

to optimally estimate the state $\mathbf{x}_k$. It is well-known (see, e.g., [11, Theorem 9.2.1]) that the LMSE estimate of $\mathbf{x}_k$ using $\mathbf{Y}_j$ is given by

$$\mathbf{x}_{k|j} = \mathbb{E}(\mathbf{x}_k) + \mathrm{Cov}(\mathbf{x}_k\mathbf{Y}_j)(\mathrm{Var}\ (\mathbf{Y}_j))^{\dagger}(\mathbf{Y}_j - \mathbb{E}(\mathbf{Y}_j)), \tag{3}$$

and the covariance matrix of estimation error is

$$\begin{aligned}
\mathbf{P}_{k|j} &= \mathbb{E}(\mathbf{x}_k - \mathbf{x}_{k|j})(\mathbf{x}_k - \mathbf{x}_{k|j})^{\mathrm{T}}\\
&= \mathrm{Var}\ \mathbf{x}_k - \mathrm{Cov}(\mathbf{x}_k\mathbf{Y}_j)(\mathrm{Var}\ (\mathbf{Y}_j))^{\dagger}(\mathrm{Cov}(\mathbf{x}_k\mathbf{Y}_j))^{\mathrm{T}}.
\end{aligned} \tag{4}$$

Equation (3) and (4) provides a batch algorithm for the optimal estimate $\hat{\mathbf{x}}$ by invoking all observations of $\mathbf{x}$. For the reductions of computational complexity and storage requirements, in many applications, especially in dynamic system, recursive algorithms must be investigated. Kalman filtering proposed fifty years ago gives an unbiased LMSE recursive algorithm to optimally estimate $\mathbf{x}_k$ in (1) from noisy data in (2) as follows.

THEOREM 1 (Kalman Filtering). *The LMSE estimate $\mathbf{x}_{k|k}$ of the state $\mathbf{x}_k$ using $\mathbf{Y}_k$ ($k = 1, 2, \ldots$) can be recursively computed as follows:*

$$\begin{aligned}
\mathbf{x}_{k|k} &= \mathbf{x}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k-1})\\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{x}_{k|k-1} + \mathbf{K}_k\mathbf{y}_k,
\end{aligned} \tag{5}$$

where, the optimal one-step prediction of state $\mathbf{x}_k$ is

$$\mathbf{x}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{x}_{k-1|k-1}, \tag{6}$$

with the initial state $\mathbf{x}_{0|0} = \mathbb{E}\,\mathbf{x}_0$*, and*

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}} + \mathit{Var}\,(\mathbf{w}_k))^{\dagger}, \tag{7}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1}\mathbf{F}_{k-1}^{\mathrm{T}} + \mathit{Var}\,(\mathbf{v}_{k-1}), \tag{8}$$

$$\begin{aligned}\mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}\\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathrm{T}} + \mathbf{K}_k\,\mathit{Var}\,(\mathbf{w}_k)\mathbf{K}_k^{\mathrm{T}},\end{aligned} \tag{9}$$

with the initial value $\mathbf{P}_0 = \mathit{Var}\,\mathbf{x}_0$*.*

If Var $(\mathbf{w}_k), \mathbf{P}_{k|k-1}$ and $\mathbf{P}_k$ are invertible, then we can easily prove that

$$\mathbf{K}_k = \mathbf{P}_k\mathbf{H}_k^{\mathrm{T}}(\mathrm{Var}\,(\mathbf{w}_k))^{-1}, \tag{10}$$

$$\mathbf{P}_k^{-1} = \mathbf{P}_{k|k-1}^{-1} + \mathbf{H}_k^{\mathrm{T}}(\mathrm{Var}\,(\mathbf{w}_k))^{-1}\mathbf{H}_k. \tag{11}$$

The key of derivation of the above recursive algorithms is the following estimation update lemma (see Theorem 9.2.2 in [11]):

LEMMA 1. *Consider the two successive observations* $\mathbf{y}_1$ *and* $\mathbf{y}_2$*, and denote the LMSE estimates of* $\mathbf{x}$ *using* $\mathbf{y}_1$ *and* $\mathbf{y}$ *respectively by* $\hat{\mathbf{x}}_1$ *and* $\hat{\mathbf{x}}$*, where*

$$\mathbf{y} = \begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}.$$

Then, we have

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_1 + \mathit{Cov}(\widetilde{\mathbf{x}}_1\widetilde{\mathbf{y}}_2)(\mathit{Var}\,(\widetilde{\mathbf{y}}_2))^{\dagger}\widetilde{\mathbf{y}}_2, \tag{12}$$

$$= \hat{\mathbf{x}}_1 + \mathbf{K}(\mathbf{y}_2 - \hat{\mathbf{y}}_2), \tag{13}$$

where $\widetilde{\mathbf{x}}_1 = \mathbf{x} - \hat{\mathbf{x}}_1, \widetilde{\mathbf{y}}_2 = \mathbf{y}_2 - \hat{\mathbf{y}}_2$*, and*

$$\mathbf{K} = \big(\mathit{Cov}(\mathbf{x}\mathbf{y}_2) - \mathit{Cov}(\mathbf{x}\mathbf{y}_1)(\mathit{Var}\,(\mathbf{y}_1))^{\dagger}\,\mathit{Cov}(\mathbf{y}_1\mathbf{y}_2)\big)(\mathit{Var}\,(\widetilde{\mathbf{y}}_2))^{\dagger}.$$

REMARK 1.

- *The quadratic convex optimization criterion plays crucial role in the derivation of analytic solutions in the above LMSE (3)–(4) and Kalman filtering (5)-(11), which give us a recursive analytic solution to be equivalent to batch analytic solution. Hence, the computational burden and storage requirements can be controlled so that they both do not increase with time goes to infinity.*
- *On the other hand, the solution of the quadratic MSE criterion strongly depends on all model knowledge, including model coefficient matrices and noise covariances. If such knowledge is incorrect or biased, the performance of Kalman filtering becomes bad (see the example given in [52]).*

3. Some Extensions of Kalman Filtering. Since Kalman filtering was proposed fifty years ago, its extensions have been extensively derived, for example, see [67, 68]. Here, only a few main extensions will be presented below.

3.1. Kalman filtering with correlated noises. To guarantee the global optimality of Kalman filtering, it was assumed that the process noises and observation noises are uncorrelated temporally, as well as, the both noises are mutually uncorrelated except at the same time instant. However, these restrictive assumptions are not satisfied in many practical systems. For example, in the state estimation of a moving object with own sensors, such as aircraft or missile inertial navigation systems, the process noise and observation noise may be mutually correlated at several time instants. Much more examples for temporally correlated observation noises or process noises are easily found in practice. In these systems, white noise is a crude approximation to real noise owing to keeping the mathematics tractable. In general, the theoretical optimal filtering for white noise driven systems is not optimal, even if quite far from optimality. In fact, the real noises are wide band and white noises are an ideal case of wide band noises (see [3, 13]). In discrete time case, the wide band noise becomes finite-time correlated. Therefore, it is very important to develop a method of handling and working with finitely correlated noises.

In [74] the standard Kalman filtering has been extended to *Globally Optimal Kalman Filtering with Finitely Correlated Noises* in the following directions *(i)* multi-step correlated process noises; *(ii)* multi-step correlated observation noises; and *(iii)* multi-step correlated process noises and observation noises. The key technique to implement the above purpose is to present an exactly recursive version for nonzero observation and process prediction error terms when the noises are finite-time correlated. Of course, by the previous research experience, it seems impossible for the standard Kalman recursive version to play such a role in recursively estimating the observation prediction error and process error terms. Therefore, we must modify the original Kalman gain by decomposing it to two recursively represented factors and increase recursive terms (for more than one-step correlated noises). In fact, this technique can handle more complicated combination of the aforementioned three multi-step correlated process and observation noises (see [79]). In the following, we briefly present the key technique.

Using Lemma 1, we can prove the following theorem:

THEOREM 2. *No matter what the auto-correlations or cross-correlations of the process noises and observation noises are, the optimal estimate* (3) *and its error covariance matrix* (4) *can be equivalently given by*

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{y}_{k|k-1}) = \mathbf{x}_{k|k-1} + \mathbf{J}_k\mathbf{L}_k^{\dagger}\Delta\mathbf{y}_k, \tag{14}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{J}_k^{\mathrm{T}} = \mathbf{P}_{k|k-1} - \mathbf{J}_k\mathbf{L}_k^{\dagger}\mathbf{J}_k^{\mathrm{T}}, \tag{15}$$

where $\mathbf{K}_k = \mathbf{J}_k\mathbf{L}_k^{\dagger}$ *is a gain factor of innovation term* $\Delta\mathbf{y}_{k-1}$, *and*

$$\mathbf{J}_k = \mathbb{E}(\mathbf{x}_k - \mathbf{x}_{k|k-1})(\mathbf{y}_k - \mathbf{y}_{k|k-1})^{\mathrm{T}}, \tag{16}$$

$$\mathbf{L}_k = \mathbb{E}(\mathbf{y}_k - \mathbf{y}_{k|k-1})(\mathbf{y}_k - \mathbf{y}_{k|k-1})^{\mathrm{T}}, \tag{17}$$

$$\mathbf{P}_{k|k-1} = \mathbb{E}(\mathbf{x}_k - \mathbf{x}_{k|k-1})(\mathbf{x}_k - \mathbf{x}_{k|k-1})^{\mathrm{T}}, \tag{18}$$

$$\Delta\mathbf{y}_k = \mathbf{y}_k - \mathbf{y}_{k|k-1}. \tag{19}$$

For the dynamic systems with specific finite-time auto- and cross-correlations of the noises, in [74, 79] the recursions of five terms $\mathbf{x}_{k|k-1}, \Delta\mathbf{y}_k, \mathbf{J}_k, \mathbf{L}_k$ and $\mathbf{P}_{k|k-1}$ in Theorem 2 have been given to develop the recursive algorithm for optimal filtering. In fact, the standard Kalman filtering Theorem 1 can be also directly deduced from recursive formulations (14) and (15) under the assumptions of uncorrelated process and observation noises as given in the two assumptions in Theorem 1. However, for more general correlated noise case, for example, general infinite-time correlated noise or state-dependent noise, the five terms $\mathbf{x}_{k|k-1}, \Delta\mathbf{y}_k, \mathbf{J}_k, \mathbf{L}_k$ and $\mathbf{P}_{k|k-1}$ may not be rewritten as recursive versions. In other words, there is no globally optimal recursive version in terms of the LMSE criterion for more general stochastic dynamic systems.

3.2. Kalman Filtering for Random Coefficient Matrices Dynamic Systems. The standard Kalman filtering require a critical assumption that model coefficient matrices are deterministic. However, in many realistic systems and backgrounds, the model coefficient matrices are actually random. For example, in nonlinear random systems, the extended Kalman filter often linearizes the nonlinear system around the random state estimate, which makes linear coefficient matrices random (see [34, 67]).In addition, discrete systems with random coefficient matrices also arise in many areas such as digital control of chemical processes, systems with human operators, economic systems and stochastically sampled digital control systems (see [18]).

There have been many works involving estimation problems for dynamic systems with random coefficient matrices in various engineering problems (see [18, 22, 32, 34, 42, 47, 63, 69, 71, 78]). Most of them are suboptimal or strongly depend on the probability density functions (pdfs) of the states, noises and random coefficients. Although the recursive formula in [18] is optimal and does not depend on the pdfs of noises and random coefficients, the random coefficient matrices are temporally independent sequences, and the problem in this case can be converted to the standard Kalman filtering. Consider a discrete time dynamic system

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{v}_k, \tag{20}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k\ ,\ \ k = 0, 1, 2, \ldots, \tag{21}$$

where $\mathbf{x}_k \in \mathbb{R}^r$ is the system state, $\mathbf{y}_k \in \mathbb{R}^N$ is the measurement vector, $\mathbf{v}_k \in \mathbb{R}^r$ is the process noise, and $\mathbf{w}_k \in \mathbb{R}^N$ is the measurement noise. The subscript k is the time index. $\mathbf{F}_k \in \mathbb{R}^{r\times r}$ and $\mathbf{H}_k \in \mathbb{R}^{N\times r}$ are random coefficient matrices.

Assume the system has the following statistical properties:

i) $\{\mathbf{F}_k, \mathbf{H}_k, \mathbf{v}_k, \mathbf{w}_k, k = 0, 1, 2, \ldots\}$ are sequences of independent random variables temporally and $\mathbf{x}_0$ is independent of them.

ii) $\mathbf{x}_k$ and $\{\mathbf{F}_k, \mathbf{H}_k, k = 0, 1, 2, \ldots\}$ are uncorrelated.

iii) The initial state $\mathbf{x}_0$, the noises $\mathbf{v}_k$, $\mathbf{w}_k$ and the coefficient matrices $\mathbf{F}_k$, $\mathbf{H}_k$ have the following means and covariances

$$\mathbb{E}(\mathbf{x}_0) = \mu_0, \quad \mathbb{E}(\mathbf{x}_0 - \mu_0)(\mathbf{x}_0 - \mu_0)^{\mathrm{T}} = \mathbf{P}_{0|0}, \tag{22}$$

$$\mathbb{E}(\mathbf{v}_k) = 0, \quad \mathbb{E}(\mathbf{v}_k\mathbf{v}_k^{\mathrm{T}}) = \mathbf{R}_{\mathbf{v}_k}, \ \mathbb{E}(\mathbf{w}_k) = 0, \ \mathbb{E}(\mathbf{w}_k\mathbf{w}_k^{\mathrm{T}}) = \mathbf{R}_{\mathbf{w}_k}, \tag{23}$$

$$\mathbb{E}(\mathbf{F}_k) = \bar{\mathbf{F}}_k, \quad \mathrm{Cov}(f_{ij}^k, f_{mn}^k) = C_{f_{ij}^k f_{mn}^k}, \tag{24}$$

$$\mathbb{E}(\mathbf{H}_k) = \bar{\mathbf{H}}_k, \quad \mathrm{Cov}(h_{ij}^k, h_{mn}^k) = C_{h_{ij}^k h_{mn}^k}, \tag{25}$$

where f_{ij}^k and h_{ij}^k are the (i, j)-th entries of matrices $\mathbf{F}_k$ and $\mathbf{H}_k$ respectively.

Rewrite

$$\mathbf{F}_k = \bar{\mathbf{F}}_k + \tilde{\mathbf{F}}_k, \tag{26}$$

$$\mathbf{H}_k = \bar{\mathbf{H}}_k + \tilde{\mathbf{H}}_k. \tag{27}$$

Substituting (26) , (27) into (20) , (21), the original system is converted to

$$\mathbf{x}_{k+1} = \bar{\mathbf{F}}_k\mathbf{x}_k + \tilde{\mathbf{v}}_k, \tag{28}$$

$$\mathbf{y}_k = \bar{\mathbf{H}}_k\mathbf{x}_k + \tilde{\mathbf{w}}_k, \tag{29}$$

where

$$\tilde{\mathbf{v}}_k = \mathbf{v}_k + \tilde{\mathbf{F}}_k\mathbf{x}_k, \tag{30}$$

$$\tilde{\mathbf{w}}_k = \mathbf{w}_k + \tilde{\mathbf{H}}_k\mathbf{x}_k. \tag{31}$$

System (28) , (29) has deterministic coefficient matrices, but the process noise and observation noise are dependent on the state. However, we can prove that Eqs. (28) , (29) still satisfy the well-known assumptions of the standard Kalman filtering. Readers are referred to [69] for complete proof. Thus, [18] derives the recursive state estimate of the new system as follows:

THEOREM 3. *The LMSE recursive state estimation of system (20), (21) is given by*

$$\begin{aligned}
\mathbf{x}_{k+1|k+1} &= \mathbf{x}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \bar{\mathbf{H}}_{k+1}\mathbf{x}_{k+1|k}),\\
\mathbf{x}_{k+1|k} &= \bar{\mathbf{F}}_k\mathbf{x}_{k|k},\\
\mathbf{P}_{k+1|k} &= \bar{\mathbf{F}}_k\mathbf{P}_{k|k}\bar{\mathbf{F}}_k^{\mathrm{T}} + \mathbf{R}_{\tilde{\mathbf{v}}_k},\\
\mathbf{K}_{k+1} &= \mathbf{P}_{k+1|k}\bar{\mathbf{H}}_{k+1}^{\mathrm{T}}(\bar{\mathbf{H}}_{k+1}\mathbf{P}_{k+1|k}\bar{\mathbf{H}}_{k+1}^{\mathrm{T}} + \mathbf{R}_{\tilde{\mathbf{w}}_{k+1}})^{\dagger},\\
\mathbf{P}_{k+1|k+1} &= (\mathbf{I} - \mathbf{K}_{k+1}\bar{\mathbf{H}}_{k+1})\mathbf{P}_{k+1|k},
\end{aligned}$$

$$\begin{aligned}
\mathbf{R}_{\tilde{\mathbf{v}}_k} &= \mathbf{R}_{\mathbf{v}_k} + \mathbb{E}(\tilde{\mathbf{F}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\tilde{\mathbf{F}}_k^{\mathrm{T}}),\\
\mathbf{R}_{\tilde{\mathbf{w}}_{k+1}} &= \mathbf{R}_{\mathbf{w}_{k+1}} + \mathbb{E}(\tilde{\mathbf{H}}_{k+1}\,\mathbb{E}(\mathbf{x}_{k+1}\mathbf{x}_{k+1}^{\mathrm{T}})\tilde{\mathbf{H}}_{k+1}^{\mathrm{T}}),\\
\mathbb{E}(\mathbf{x}_{k+1}\mathbf{x}_{k+1}^{\mathrm{T}}) &= \bar{\mathbf{F}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\bar{\mathbf{F}}_k^{\mathrm{T}} + \mathbb{E}(\tilde{\mathbf{F}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\tilde{\mathbf{F}}_k^{\mathrm{T}}) + \mathbf{R}_{\mathbf{v}_k},\\
\mathbf{x}_{0|0} &= \mathbb{E}\,\mathbf{x}_0,\ \mathbf{P}_0 = \mathrm{Var}(\mathbf{x}_0),\ \mathbb{E}(\mathbf{x}_0\mathbf{x}_0^{\mathrm{T}}) = \mathbb{E}\,\mathbf{x}_0\,\mathbb{E}\,\mathbf{x}_0{}^{\mathrm{T}} + \mathbf{P}_0.
\end{aligned}$$

Compared with the standard Kalman filtering, the random coefficient matrices Kalman filtering has one more recursion equation $\mathbb{E}(\mathbf{x}_{k+1}\mathbf{x}_{k+1}^{\mathrm{T}})$. In Theorem 3, we eventually have to compute $\mathbb{E}(\tilde{\mathbf{F}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\tilde{\mathbf{F}}_k^{\mathrm{T}})$ and $\mathbb{E}(\tilde{\mathbf{H}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\tilde{\mathbf{H}}_k^{\mathrm{T}}))$. Their analytical expressions are given by

$$\begin{aligned}
\mathbb{E}(\tilde{\mathbf{F}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\tilde{\mathbf{F}}_k^{\mathrm{T}})_{(m,n)} &= \sum_{i=1}^{r} C_{f_{n1}^k f_{mi}^k} X_{i1}^k + \cdots + \sum_{i=1}^{r} C_{f_{nr}^k f_{mi}^k} X_{ir}^k,\\
\mathbb{E}(\tilde{\mathbf{H}}_k\,\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})\tilde{\mathbf{H}}_k^{\mathrm{T}})_{(m,n)} &= \sum_{i=1}^{r} C_{h_{n1}^k h_{mi}^k} X_{i1}^k + \cdots + \sum_{i=1}^{r} C_{h_{nr}^k h_{mi}^k} X_{ir}^k,
\end{aligned}$$

where X_{ij}^k is the (i,j)-th entry of $X^k = \mathbb{E}(\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}})$.

3.2.1. Application to dynamic process with spurious observations . In the classical Kalman filtering problem, the observation is always assumed to contain the signal to be estimated. However, in practice, when the exterior interference is strong, i.e., total covariance of the measurement noise or clutter is large, the estimator will mistake the noise or clutter as the observation sometimes.Usually, the estimator cannot know whether this happens or not, only the probability of two cases known.

Consider a discrete dynamic process

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{v}_k, \tag{32}$$

$$\mathbf{y}_k = \mathbf{A}_k\mathbf{x}_k + \mathbf{w}_k\ ,\ k = 0,1,2,\ldots, \tag{33}$$

where $\{\mathbf{F}_k,\ \mathbf{A}_k,\ \mathbf{v}_k,\ \mathbf{w}_k,\ k = 0,1,2,\ldots\}$ satisfy the assumptions of standard Kalman filtering. $\mathbf{F}_k$ and $\mathbf{A}_k$ are deterministic matrices. Suppose that the probability of observation coming from above model is p_k. Thus, false alarm probability of the observation is $1 - p_k$. Then,we can rewrite the measurement equations as follow

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k\ ,\ k = 0,1,2,\ldots, \tag{34}$$

where the observation matrix $\mathbf{H}_k$ is a binary-valued random with

$$\Pr\{\mathbf{H}_k = \mathbf{A}_k\} = p_k, \tag{35}$$

$$\Pr\{\mathbf{H}_k = 0\} = 1 - p_k. \tag{36}$$

In the false alarm case, the state transition matrix is still deterministic, but the measurement matrix is random. Thus the dynamic process with a false alarm probability

is a special case of model (20), (21). We can use Theorem 3 to deal with the dynamic system with false alarm.

EXAMPLE 1. *The object dynamics and measurement equations are modeled as follows.*

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{F}_k\mathbf{x}_k + \mathbf{v}_k,\\ \mathbf{y}_{k,i} &= \mathbf{H}_{k,i}\mathbf{x}_k + \mathbf{w}_{k,i},\ \ i = 1,2,3,\end{aligned}$$

where $\{\mathbf{F}_k,\ \mathbf{H}_{k,i},\ \mathbf{v}_k,\ \mathbf{w}_{k,i},\ k = 0,1,2,\ldots\}$ *satisfy the assumptions of standard Kalman filtering. The state transition matrix* $\mathbf{F}_k$

$$\mathbf{F}_k = \begin{bmatrix}\cos(2\pi/300) & \sin(2\pi/300)\\ -\sin(2\pi/300) & \cos(2\pi/300)\end{bmatrix}$$

is a constant. The measurement matrix is given by

$$\mathbf{H}_{k,i} = \begin{bmatrix}1 & i\\ 1 & -i\end{bmatrix}, i = 1,2, \quad \mathbf{H}_{k,3} = \begin{bmatrix}1 & 3\\ -1 & 3\end{bmatrix}.$$

The false alarm probability of the i*-th sensor is given by*

$$1 - p_{k,1} = 0.01, \quad 1 - p_{k,2} = 0.02, \quad 1 - p_{k,3} = 0.03.$$

The initial state $\mathbf{x}_0 = \Big[50, 0\Big]$, $\mathbf{P}_{0|0,i} = \mathbf{I}$. *The covariance of the noises are diagonal, given by* $\mathbf{R}_{\mathbf{v}_k} = 1$, $\mathbf{R}_{\mathbf{w}_{k,i}} = 2$, $i = 1,2,3$. *Using Monte-Carlo method of 50 runs, we can evaluate tracking performance of an algorithm by estimating the second moment of the tracking error, given by*

$$E_k^2 = \frac{1}{50}\sum_{i=1}^{50} ||\mathbf{x}_{k|k}^{(i)} - \mathbf{x}_k||^2. \tag{37}$$

Figure 1 shows that the second moments of tracking error for three sensors Kalman filtering fusion without considering the false alarm (i.e. standard Kalman filtering) and three sensors random coefficient matrices Kalman filtering fusion considering the false alarm (i.e. random coefficient matrices Kalman filtering), respectively. It can be shown that even if the false alarm probability is very small, the distributed random coefficient matrices Kalman filtering fusion performs much better than the standard Kalman filtering.

3.2.2. Application to multiple-model dynamic process.

The multiple-model (MM) dynamic process has been considered by many researchers. Although the possible models considered in those papers are quite general and can depend on the state, no optimal algorithm in MSE sense was proposed in the past few decades. On the other hand, when some of the MM systems satisfy the assumptions

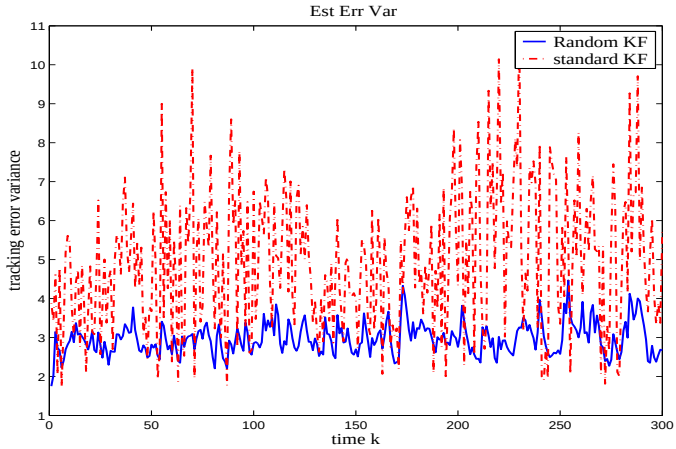

FIG. 1. *Comparison of standard Kalman filtering fusion and random coefficient matrices Kalman filtering fusion*

in this section, they can be reduced to dynamic models with random transition matrix and thus the optimal real-time filter can be given directly according to the random transition matrix Kalman filtering proposed in Theorem 3.

Consider a system

$$\mathbf{x}_{k+1} = \mathbf{F}_{k,i}\mathbf{x}_k + \mathbf{v}_k \quad \text{with probability } p_{k,i},\ i = 1, \ldots, l, \tag{38}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k, \tag{39}$$

where $\mathbf{F}_{k,i}$ and $\mathbf{v}_{k,i}$ are independent sequence, and $\mathbf{H}_k$ is non-random. We use random matrix $\mathbf{F}_k$ to stand for the state transition matrix. The expectation of $\mathbf{F}_k$ can be expressed as:

$$\bar{\mathbf{F}}_k = \sum_{i=1}^{l} p_{k,i}\mathbf{F}_{k,i} \tag{40}$$

$$\tilde{\mathbf{F}}_k = \mathbf{F}_{k,i} - \bar{\mathbf{F}}_k, \quad \text{with probability } p_{k,i}. \tag{41}$$

In MM dynamic process, the measurement matrix is still deterministic, but the state transition matrix is random. Thus it is also a special case of model (20), (21). We can use Theorem 3 to deal with the MM dynamic process.

EXAMPLE 2. *In this simulation, there are three dynamic models, with the corresponding probabilities of occurrence available. The object dynamics and measurement matrix in (38) are given by*

$$\mathbf{F}_{k,1} = \begin{bmatrix} \cos(2\pi/300) & \sin(2\pi/300) \\ -\sin(2\pi/300) & \cos(2\pi/300) \end{bmatrix} \textit{with probability 0.1},$$

$$\mathbf{F}_{k,2} = \begin{bmatrix} \cos(2\pi/250) & \sin(2\pi/250) \\ -\sin(2\pi/250) & \cos(2\pi/250) \end{bmatrix} \textit{ with probability 0.2},$$

$$\mathbf{F}_{k,3} = \begin{bmatrix} \cos(2\pi/100) & \sin(2\pi/50) \\ -\sin(2\pi/50) & \cos(2\pi/100) \end{bmatrix} \textit{ with probability 0.7},$$

$$\mathbf{H}_k = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

The covariance of the noises are diagonal, given by $\mathbf{R}_{\mathbf{v}_k} = 1$, $\mathbf{R}_{\mathbf{w}_k} = 2$. *In the following, we compare our numerical results with the well-known IMM algorithm. Since in this example, the occurrence probability of each model at every time* k *is known and mutually independent, it is also the transition probability in the IMM. Therefore, the transition probability matrix* Π *at each time in the IMM is fixed and given by*

$$\Pi = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.1 & 0.2 & 0.7 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}.$$

$\Pi(i,j)$ *here means the transition probability of model* i *to model* j. *This assumption also implies that the model probability in the IMM is fixed as follows:*

$$\pi_{k,1} = 0.1, \quad \pi_{k,2} = 0.2, \quad \pi_{k,3} = 0.7.$$

Figure 5 shows that the random coefficient matrices Kalman filtering given in this section still works better than the IMM with the fixed transition probability and model probability. This makes sense since the former is optimal in the MSE sense but the latter is not.

3.2.3. Correlated random coefficient matrix sequences. In many realistic systems and practical backgrounds, the model coefficient matrices are not only random, but also correlated. For examples: 1) in the maneuvering target tracking scenario, the targets being tracked undergo occasional maneuvers, so that the motion of such targets is often modeled by one-order Markov random transition matrices (see [22, 32, 34, 42, 47]); 2) in nonlinear random systems, the extended Kalman filter often linearizes the nonlinear system around the random state estimate, which also makes the correlated coefficient matrices (see [34, 67]); 3) [69, 78] proposed that the *combination* step in the interacting multiple model (IMM) and the joint probabilistic data association (JPDA) methods can be transformed to dynamic systems with the correlated measurement matrices.

In paper [76], dynamic systems with only the measurement matrices correlated was consider first. Similar to that in [18], such random dynamic systems are converted to deterministic dynamic systems whose measurement noises can be proved to

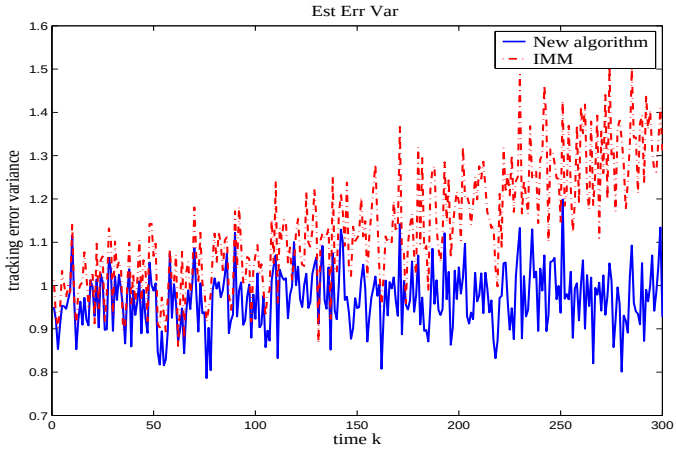

FIG. 2. *Comparison of IMM and random coefficient matrices Kalman filtering*

be generally one-step correlated. More detailed speaking, the original measurement matrix was separated to be the deterministic and random parts and then the later was combined with the original measurement noise to be a modified noise which is proved to be temporally one-step correlated. Thus, the optimal estimation results on finite step correlated measurement noises in [74] can be used to obtain the optimal recursive estimation. Secondly, a class of specific dynamic systems, where both transition matrices and measurement matrices are a one-step moving average (MA) matrix sequences driven by a common independent zero-mean random sequence. The model is one-order Markovian, which is similar to IMM problem discussed in [22, 32, 42, 47]. To derive optimal recursive estimation, the key technique used there is to introduce two auxiliary state variables and decompose the Kalman gain matrix into two factor matrices to find their recursive schemes respectively. The optimal recursive estimation of the state can be obtained by using the first order through the sixth order moments of the driving sequence of the aforementioned MA matrix sequences. If only the transition matrices are temporally correlated, the optimal filter can be given by using the first through the fourth order moments of the driving sequence. Numerical examples show that the optimal estimation is better than the standard Kalman filter and the random Kalman filter [18] with the correlation of coefficient matrices ignored, specially, the new method can get better improvement over them when both the transition matrices and the measurement matrices are temporally correlated.

REMARK 2. *Although the criterion here is still quadratic LMSE, the multiplicative relationship between coefficient matrix and the state, in particular, the recursion in process equation make the high order moments of the driving sequence required to obtain the second moments of Kalman filtering. In practice, it is may be very hard*

to get high order moments of the driving sequences. This shows that the optimality of Kalman filtering in MSE sense not only requires precise model knowledge, but also cannot afford the trouble brought by correlations of model coefficient matrices. Therefore, people may prefer suboptimal Kalman filtering with ignoring some correlations in the model knowledge.

3.3. Kalman filtering with negative time correct/incorrect measurements. In target tracking systems, there often exist various propagation times from sensors to local processors or a central processor, it is clearly possible that some measurements will arrive out of sequence (OOS), as discussed in [34].The early works on the current state update with an OOS measuremnt (OOSM) in [30] and [40] presented an approximate solution called "algorithm B" in [55]. The main result of [55] extended the previous work to create an algorithm with optimal output, "algorithm A". It was also shown by numerical examples in [55] that the "algorithm B" is nearly optimal. Subsequently, the generalized solutions for the ℓ-step-lag case, to be called "algorithm Aℓ1" and "algorithm Bℓ1", were derived in [57]. It was shown by numerical examples that "algorithm Aℓ1" and "algorithm Bℓ1" are only slightly suboptimal and the corresponding numerical results are very close to those obtained by reordering and reprocessing an entire sequence of measurements. The update for the ℓ-step-lag case was discussed in [48] and [54], which are suboptimal and require an iteration back for ℓ steps and considerable computational amount of storage.

For ℓ-step-lag OOSMs, the algorithms in [65] are of globally optimal recursive algorithm. However, they require considerable amount of storage (see [65]). The other algorithms are mostly suboptimal. The readers can also refer to [72, 73] for the optimal centralized and distributed multisensor fusion update for the case of the correlated measurement noises. However, all previous algorithms cannot update the entire state trajectory which is important in some practical applications. For example, an accurate state trajectory can certainly benefit to the identification of the aerodynamic model (see [43]) and the state trajectory association for multi-target multisensor tracking (see [16]). In [75] a globally optimal state trajectory update algorithm for a sequence with arbitrary delayed OOSMs was proposed. It can update the current whole trajectory other than only the current single state with less computationThe resulting object tracking trajectory is a smooth state trajectory.

On the other hand, the counterpart of the OOSM update problem is the need to remove an earlier measurement from the state trajectory (see [70]). This can happen in multi-sensor tracking systems very often. For example, after the central tracker received measurements/estimates from local sensors and obtained a central estimate, a local sensor sometimes may change its mind and send a message to central tracker that an earlier measurement should be removed from tracks because, based on its

information, an earlier measurement is an incorrectly associated one. More practical applications can be referred to [66]. In [75] an optimal trajectory update algorithm with removing early incorrect measurement was also proposed as an extension of the counterpart of the OOSM update problem.

4. Main Limitations to Kalman Filtering in Practice. Obviously, Kalman filtering can achieve the optimal performance in the MSE sense whenever all optimality assumptions and the required precise model knowledge can be given. Besides, the required computational burden is affordable easily. However, while the MSE criterion brings us these advantages, it also yields the following drawbacks in practical applications:

1) **No Robustness.** The optimal solution of the MSE needs much model knowledge: precise model coefficient matrices and noise covariances. If they cannot be known exactly, and one uses nominal model with unknown bias, the filtering performance becomes bad. Unfortunately, in real world there exist great number of uncertain systems with biased model knowledge.

2) **Restrictive Optimality Conditions.** If the system is essentially nonlinear, or the uncorrelation/independence of the random sequences for the model noises and the random coefficient matrices to be violated, it is very hard to analyze and derive the global optimality in terms the MSE although some of nearly optimal algorithms and nonlinear filtering have become widespread and useful in practice, for example, some adaptive Kalman filtering [64], extended forgetting factor recursive Least Squares algorithm [45], and the extended Kalman filtering[61], the unscented Kalman filtering, and particle filtering (see [67]).

3) **Large Error with Non-Zero Probability.** In fact , many of practical dynamic systems only can give us a signal trajectory other than so many repeated trajectories. Hence, Kalman filtering cannot overcome such contradiction between the MSE optimality and practical single implementation and may have arbitrary large error with non-zero probability. Besides, some engineering problems, such as nuclear and airspace engineering cannot tolerate large error even with very small probability.

5. Dynamic Systems with Bounded Uncertain Model Biases. To overcome the aforementioned drawbacks of Kalman filtering, another formulation of uncertain dynamic systems with bounded bias has been investigated in the past two decades, in particular, developed very much based on the recently derived new results on the convex optimization theory and high-efficient algorithms.

5.1. Set-Membership filtering. A dynamic system with bounded uncertain model bias in fact tell us a model set of dynamic systems, where the true dynamic system must be inside the set but one cannot know which member of the set it is. In

this case, to cope with the uncertainty of model, naturally, set-valued filtering is an appropriate choice of the state estimation. Set-valued state estimation–an estimation set that is guaranteed to contain the state vector to be estimated has also been extensively studied in the bounded perturbation framework. In some cases, the assumption of the bounded noises is reasonable indeed. For example, an aircraft often has several tons; it is acceptable that its position, velocity and acceleration have bounded interfered changes. Besides, when the models are uncertain, probability density functions (pdf) of noises are difficult to be obtained in general. Thus, measurements of too far away are usually ignored and the measurement noises are assumed bounded. The set-valued estimation was considered first at end of 1960s and early 1970s (see [5, 6, 8]). These results aim to compute *bounding ellipsoids* which are guaranteed to contain the state vector to be estimated given bounds on the perturbations and noises. The idea of propagating bounding ellipsoids (or other convex sets: boxes, polytopes, simplexes, parallelotopes, and polytopes) for systems with bounded noises has also been investigated by many authors, for example, see papers [17, 23, 36, 44, 51, 52, 60, 62] , the survey papers [29, 28], the special issues of journals [24, 33, 35], the books [12, 31, 37, 39, 49], and references therein. The usual approaches, bounding ellipsoids and bounding boxes, are presented in this section. Specially, [27] proposed a class of "size" measure of an ellipsoid as the optimization objectives of the set-valued estimation, which is in fact a generalized trace measure. Recent results [52] give a general convex optimization method for uncertain linear equations and dynamical systems.

In many practical applications, one needs to know an exact estimation of error upper bound and wants it as small as possible. For example, in some aerospace and nuclear engineering, a large error even with very small probability may not be allowed. However, the estimation methods derived by statistical average error (i.e., MSE) criterion sometimes may not satisfy practical requirements. Fortunately, the set-valued estimate can be easily converted to point estimate by taking the central point of estimate set. Then, the maximal Euclidean distance between such a point estimate and boundary of the estimate set is an Euclidean error bound. Therefore, an important issue is how to minimize the Euclidean error bound. So far, minimizing Euclidean error estimation has not been received enough attention while the statistical average method has been investigated extensively.

Properly using these set-valued estimation methods in terms of convex optimization and the techniques of multisensor and multi-algorithm fusion, the practically and theoretically significant estimation problem–minimizing the Euclidean error of state estimation can be efficiently solved. To our best knowledge, although the ellipsoid/box estimation also provides an Euclidean error estimation (namely, a center of ellipsoid/box as a required estimate, the longest semi-axis of ellipsoid/box as Eu-

clidean estimation error) of the state, one does not figure out that the Euclidean error of state estimation can be further minimized and the previous set-valued estimation methods do have such potential. How to use the complementary advantages of multiple sensors and multiple algorithms to minimize Euclidean estimation error of the state vector is what we present in this section (for more details, see [77]).

Now, we present the mathematical basis of El Ghaoui & Calafiore's method [52] and basic idea of multisensor and multi-algorithm estimation fusion for minimizing Euclidean estimation error of the state vector.

Main Stages of Set-Valued Estimation

- For quite general system, derive a convex constraint condition guaranteeing estimation coverage set must cover true state. It is in general no longer simply quadratic constraint;
- Under such constraint, choose an objective function minimizing "size" of the coverage set. For example, take the sum of squares of all semi-axes lengths as ellipsoid size, or sum of all edges as box size as ... below. They are all convex objective functions;
- Based recent development of optimization theory, there exist some high-efficient algorithms to solve the above convex optimization problem.

Basic Idea of Minimizing Euclidean Error

- Use ***multiple sensors and multiple algorithms*** to get many very distinct estimation sets (ellipsoids or boxes);
- For each algorithm, choose its own objective function $f(\mathbf{P})$ or $g(\mathbf{b})$ (for example, see (48) or (49)) which is a weighted sum and only gives very high weight to a unique term of the sum to yield a prolate estimate coverage set whose estimate error of the unique element of the state vector is as small possible;
- Use set intersection to get smaller-size set;
- Finally, Euclidian estimation error bound can be minimized.

To intuitively understand the aforementioned ideas, see the following three figures.

These results are also illustrated by a numerical example. An interesting phenomenon is that the Euclidean estimation error of the entry of the state vector of the distributed fusion algorithm is smaller than that of the centralized fusion algorithm although the objective of the centralized fusion is generally smaller than that of the distributed fusion. The main reason is that the Euclidean error of the distributed fusion is derived from more estimation ellipsoids/boxes, including the fusion ellipsoid/box and the estimation ellipsoids/boxes (not unprocessed observation sets in the centralized fusion) transmitted from the sensors, but the Euclidean error of the centralized fusion is derived from a fusion ellipsoid/box only. Moreover, using this fact and the complementary advantages of multiple algorithms, we first use mul-

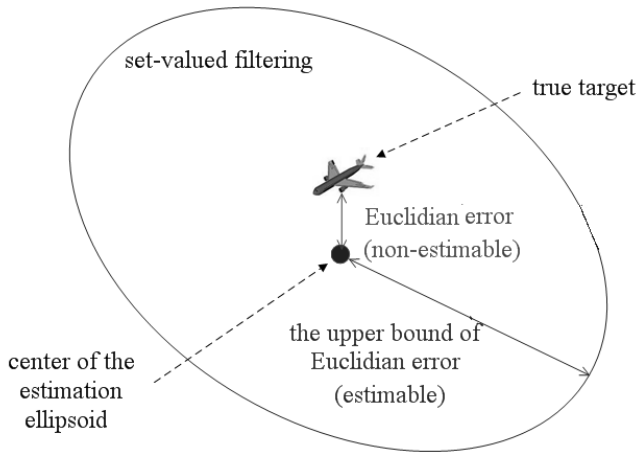

FIG. 3. *Euclidean error bound based set-valued estimation*

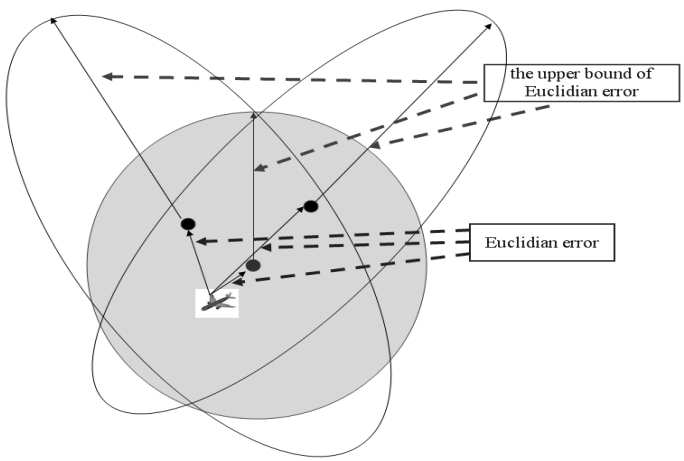

FIG. 4. *Two set-valued estimate fusion from two sensor makes smaller Euclidean error bound*

tiple differently weighted objectives at both the local sensors and the fusion center, where each only emphasizes an individual entry of the estimated state vector, so as to yield multiple estimation ellipsoids/boxes squashed along each entry of the state vector as much as possible respectively. Then *intersection fusion* of these estimation ellipsoids/boxes yields a final Euclidean-error-minimized state estimate. Obviously, the idea in the proposed method can be utilized to other set estimation methods to yield Euclidean-error-minimized state estimate. Numerical examples support the above analysis and show that the new method can significantly reduce the Euclidean estimation error of the state vector.

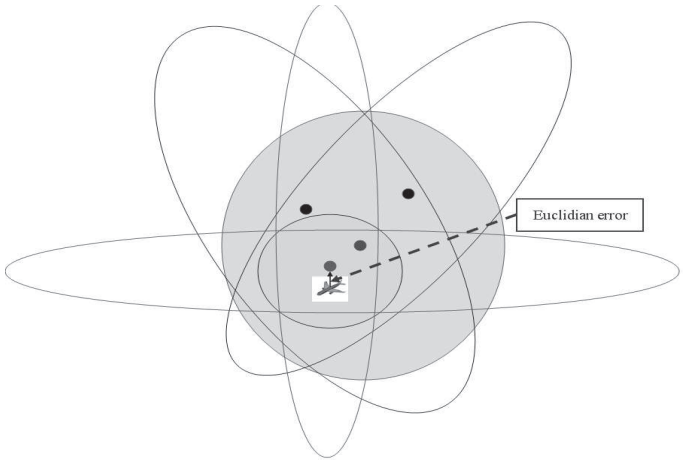

FIG. 5. *Two set-valued estimate fusion from two different algorithms further makes smaller Euclidean error bound*

5.1.1. Problem formulation. system with model uncertainty and bounded noises is formulated as follows:

$$\mathbf{x}_{k+1} = \mathbf{M}_k(\Delta_k) \begin{bmatrix} \mathbf{x}_k \\ \mathbf{v}_k \end{bmatrix} \tag{42}$$

$$\mathbf{y}_{k,i} = \mathbf{M}_{k,i}(\Delta_{k,i}) \begin{bmatrix} \mathbf{x}_k \\ \mathbf{w}_{k,i} \end{bmatrix}, \qquad i = 1, \ldots, L, \tag{43}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state of system at time k; $\mathbf{v}_k \in \mathbb{R}^{n_v}$ is a bounded uncertain process noise taking value in a unit sphere, i.e., $\| \mathbf{v}_k \| \leq 1$; $\mathbf{y}_{k,i} \in \mathbb{R}^{n_i}$ is the measurement of state $\mathbf{x}_k$; $\mathbf{w}_{k,i} \in \mathbb{R}^{n_{w_i}}$ is the bounded uncertain measurement noise of the i-th sensor taking value in a unit sphere, i.e., $\| \mathbf{w}_{k,i} \| \leq 1$. The uncertainty on the system parameter matrices is assumed to be represented in *linear fractional representation* (LFR) form, for any given $\Delta_k \in \mathbb{R}^{n_p,n_q}$, $\Delta_{k,i} \in \mathbb{R}^{n_{pi},n_{qi}}$, $i = 1, \ldots, L$

$$\mathbf{M}_k(\Delta_k) = [\mathbf{F}_k \;\; \mathbf{A}_k] + \mathbf{L}_k \Delta_k (\mathbf{I} - \mathbf{D}_k \Delta_k)^{-1} [\mathbf{R}_{\mathbf{F}_k} \;\; \mathbf{R}_{\mathbf{A}_k}], \tag{44}$$

$$\mathbf{M}_{k,i}(\Delta_{k,i}) = [\mathbf{H}_{k,i} \;\; \mathbf{B}_{k,i}] + \mathbf{L}_{k,i} \Delta_{k,i} (\mathbf{I} - \mathbf{D}_{k,i} \Delta_{k,i})^{-1} [\mathbf{R}_{\mathbf{H}_{k,i}} \;\; \mathbf{R}_{\mathbf{B}_{k,i}}], \tag{45}$$

where $\mathbf{I}$ is an identity matrix with compatible dimensions; $\mathbf{F}_k \in \mathbb{R}^{n,n}$, $\mathbf{A}_k \in \mathbb{R}^{n,n_v}$, $\mathbf{H}_{k,i} \in \mathbb{R}^{n_i,n}$, $\mathbf{B}_{k,i} \in \mathbb{R}^{n_i,n_{w_i}}$, $\mathbf{L}_k \in \mathbb{R}^{n,n_p}$, $\mathbf{L}_{k,i} \in \mathbb{R}^{n_i,n_{pi}}$, $\mathbf{D}_k \in \mathbb{R}^{n_q,n_p}$, $\mathbf{D}_{k,i} \in \mathbb{R}^{n_{qi},n_{pi}}$, $\mathbf{R}_{\mathbf{F}_k} \in \mathbb{R}^{n_q,n}$, $\mathbf{R}_{\mathbf{A}_k} \in \mathbb{R}^{n_q,n_v}$, $\mathbf{R}_{\mathbf{H}_{k,i}} \in \mathbb{R}^{n_{qi},n}$ and $\mathbf{R}_{\mathbf{B}_{k,i}} \in \mathbb{R}^{n_{qi},n_{w_i}}$ are known time-varying matrices. The uncertainty time-varying matrices Δ_k, $\Delta_{k,i}$ are in general structured and bounded $\Delta_k \in \tilde{\Omega}_k \triangleq \{\Delta \in \Omega_k : ||\Delta|| \leq 1\}$, $\Delta_{k,i} \in \tilde{\Omega}_{k,i} \triangleq \{\Delta \in \Omega_{k,i} : ||\Delta|| \leq 1\}$, where $\Omega_k \subset \mathbb{R}^{n_p,n_q}$, $\Omega_{k,i} \subset \mathbb{R}^{n_{pi},n_{qi}}$ are structure subspaces. Let these

LFRs be well-posed over Ω_k and $\Omega_{k,i}$ respectively, meaning that $\det(\mathbf{I}-\mathbf{D}_k\Delta_k)\neq 0$, $\forall\Delta_k\in\tilde{\Omega}_k$, $\det(\mathbf{I}-\mathbf{D}_{k,i}\Delta_{k,i})\neq 0, \forall\Delta_{k,i}\in\tilde{\Omega}_{k,i}$, a sufficient condition for well-posedness can be seen in μ analysis problems [26]. The above LFR of the uncertainty is widely used in control theory and has great generality including the additive uncertainty of the form $\mathbf{M}_k(\Delta_k)=[\mathbf{F}_k+\Delta_F \;\; \mathbf{A}_k+\Delta_A]$ if let $\mathbf{L}_k=\mathbf{I}$, $\mathbf{D}_k=0$, $\Delta_k=\mathrm{diag}(\Delta_F,\Delta_A)$, $[\mathbf{R}_{\mathbf{F}_k} \;\; \mathbf{R}_{\mathbf{A}_k}]=[\mathbf{I} \;\; \mathbf{I}]$ (see [26, 60]).

5.1.2. Measures of size of an ellipsoid or a box. For a given ellipsoid with a center $\hat{\mathbf{x}}$,

$$\mathcal{E}=\{\mathbf{x}\in\mathbb{R}^n:(\mathbf{x}-\hat{\mathbf{x}})^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{x}-\hat{\mathbf{x}})\leq 1\} \tag{46}$$

$$=\{\mathbf{x}\in\mathbb{R}^n:\mathbf{x}=\hat{\mathbf{x}}+\mathbf{E}\mathbf{u},||\mathbf{u}||\leq 1\}, \tag{47}$$

the "size" of the ellipsoid $\mathcal{E}$ is a function of the *shape matrix* $\mathbf{P}\succ 0$, and will be denoted by $f(\mathbf{P})$, $\mathbf{E}$ is Cholesky factorization of $\mathbf{P}$, i.e., $\mathbf{P}=\mathbf{E}\mathbf{E}^{\mathrm{T}}$. The lengths of the semi-axes of $\mathcal{E}$ are given by $\sqrt{\lambda_i}$, where λ_i are the eigenvalues of $\mathbf{P}$. It is well known that $\mathrm{tr}(\mathbf{P})$ corresponds to the sum of squares of semiaxes lengths of $\mathcal{E}$, and $\mathrm{logdet}(\mathbf{P})$ is related to the volume of the ellipsoid (see [58]). Based on $\mathrm{tr}(\mathbf{P})$, [27] proposed a class of objectives, which can emphasize the importance of the state-vector entry of interest. The class of objectives of ellipsoid estimation is

$$f(\mathbf{P})=\omega_1\mathbf{P}_{11}+\omega_2\mathbf{P}_{22}+\cdots+\omega_n\mathbf{P}_{nn}, \tag{48}$$

where ω_j is the weight coefficient, $\omega_j>0$ and $\sum_{j=1}^n\omega_j=1$. If the bound of the j-th entry of the state vector is more important, one can give a larger weight to ω_j, since the projection bound (half of the projection length) of the ellipsoid $\mathcal{E}$ along the j-th direction of the state vector is $\sqrt{\mathbf{P}_{jj}}$ where $\mathbf{P}_{jj}$ is the element in the j-th row and j-th column of the matrix $\mathbf{P}$ (see [58]). In particular, if let $\omega_j=\frac{1}{n}, j=1,2,\ldots,n$ (every entry of the state vector is treated equally), then the corresponding objective is equivalent to the trace objective. Thus, $\mathrm{tr}(\mathbf{P})$ not only corresponds to the sum of squares of semiaxes lengths of $\mathcal{E}$, but also is the sum of squares of projection bounds of $\mathcal{E}$ along all state coordinate directions. Moreover, (48) corresponds to the weighted sum of squares of projection bounds of $\mathcal{E}$ along all state coordinate directions.

For a box $\mathcal{B}$, one can use the perimeter or volume of a box as a measure of the size of the box. To emphasize the importance of the interested entry of the state vector and guarantee the convex optimization, the following weighted sum of squares of side length of $\mathcal{B}$ can also be used to construct the estimation fusion of multiple algorithms for minimizing Euclidean estimation error of the state vector:

$$g(\mathbf{b})=\omega_1(\mathbf{b}(1))^2+\omega_2(\mathbf{b}(2))^2+\cdots+\omega_n(\mathbf{b}(n))^2, \tag{49}$$

where b is the vector of the side length of the box $\mathcal{B}$. The matrix $\mathrm{diag}((\frac{\mathbf{b}(1)}{2})^2, \ldots, (\frac{\mathbf{b}(n)}{2})^2)$, corresponding to the the shape matrix $\mathbf{P}$ of the ellipsoid $\mathcal{E}$, can be taken as the shape matrix of the box $\mathcal{B}$. Moreover, the matrix $\mathrm{diag}(\frac{\mathbf{b}(1)}{2}, \ldots, \frac{\mathbf{b}(n)}{2})$ corresponds to the Cholesky factorization matrix $\mathbf{E}$ of the shape matrix $\mathbf{P}$ of the ellipsoid $\mathcal{E}$. In addition, the free degrees of the box shape and the ellipsoid shape are n and $\frac{n(n+1)}{2}$ respectively, which will be used to compare the number of the decision variables of box and ellipsoid optimization problems.

Throughout this section, $f(\mathbf{P})$ defined by (48) and $g(\mathbf{b})$ defined by (49) will be the objectives of ellipsoid optimization problems and box optimization problems respectively. In addition, they will also be used to construct the estimation fusion of multiple algorithms for minimizing Euclidean estimation error of the state vector.

5.1.3. State bounding box estimation based on centralized fusion. Suppose that the initial state $\mathbf{x}_0$ belongs to a given bounding ellipsoid or box:

$$\mathcal{E}_{\mathbf{x}_0} = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \hat{\mathbf{x}}_0)^{\mathrm{T}} (\mathbf{P}_{\mathbf{x}_0})^{-1} (\mathbf{x} - \hat{\mathbf{x}}_0) \leq 1\}, \tag{50}$$

$$\mathcal{B}_{\mathbf{x}_0} = \{\mathbf{x} \in \mathbb{R}^n :| \mathbf{x}(j) - \hat{\mathbf{x}}_0(j) \mid \leq \frac{\mathbf{b}_{\mathbf{x}_0}(j)}{2}, j = 1, \ldots, n\}, \tag{51}$$

where $\hat{\mathbf{x}}_0$ is the center of ellipsoid $\mathcal{E}_{\mathbf{x}_0}$ or the center of the box $\mathcal{B}_{\mathbf{x}_0}$; $\mathbf{P}_{\mathbf{x}_0}$ is the *shape matrix* of the ellipsoid $\mathcal{E}_{\mathbf{x}_0}$ which is a known symmetric positive-definite matrix; $\mathbf{b}_{\mathbf{x}_0}$ is the vector of the side length of the box $\mathcal{B}_{\mathbf{x}_0}$, whose entries are positive. At time k, given that $\mathbf{x}_k$ belongs to a current bounding ellipsoid or box:

$$\mathcal{E}_{\mathbf{x}_k} = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \hat{\mathbf{x}}_k)^{\mathrm{T}} (\mathbf{P}_{\mathbf{x}_k})^{-1} (\mathbf{x} - \hat{\mathbf{x}}_k) \leq 1\}, \tag{52}$$

$$\mathcal{B}_{\mathbf{x}_k} = \{\mathbf{x} \in \mathbb{R}^n :| \mathbf{x}(j) - \hat{\mathbf{x}}_k(j) \mid \leq \frac{\mathbf{b}_{\mathbf{x}_k}(j)}{2}, j = 1, \ldots, n\}, \tag{53}$$

where $\hat{\mathbf{x}}_k$ is the center of ellipsoid $\mathcal{E}_{\mathbf{x}_k}$ or the center of the box $\mathcal{B}_{\mathbf{x}_k}$; $\mathbf{P}_{\mathbf{x}_k}$ is a known symmetric positive-definite matrix; $\mathbf{b}_{\mathbf{x}_k}$ is the vector of the side length of the box $\mathcal{B}_{\mathbf{x}_k}$. At next time $k+1$, the fusion center can receive the measurements $\mathbf{y}_{k+1,i}$ from the i-th sensor, $i = 1, 2, \ldots, L$.

For the centralized fusion system, at time $k+1$, the goal of the fusion center is to determine a fusion bounding ellipsoid $\mathcal{E}_{\mathbf{x}_{k+1}}$ or bounding box $\mathcal{B}_{\mathbf{x}_{k+1}}$, i.e, look for $\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{\mathbf{x}_{k+1}}$ or $\hat{\mathbf{x}}_{k+1}, \mathbf{b}_{\mathbf{x}_{k+1}}$ such that the state $\mathbf{x}_{k+1}$ belongs to

$$\mathcal{E}_{\mathbf{x}_{k+1}} = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \hat{\mathbf{x}}_{k+1})^{\mathrm{T}} (\mathbf{P}_{\mathbf{x}_{k+1}})^{-1} (\mathbf{x} - \hat{\mathbf{x}}_{k+1}) \leq 1\}, \tag{54}$$

$$\text{or} \quad \mathcal{B}_{\mathbf{x}_{k+1}} = \{\mathbf{x} \in \mathbb{R}^n :| \mathbf{x}(j) - \hat{\mathbf{x}}_{k+1}(j) \mid \leq \frac{\mathbf{b}_{\mathbf{x}_{k+1}}(j)}{2}, j = 1, \ldots, n\}, \tag{55}$$

whenever i) $\mathbf{x}_k$ is in $\mathcal{E}_{\mathbf{x}_k}$ or $\mathcal{B}_{\mathbf{x}_k}$, ii) the process and measurement noises $\mathbf{v}_k, \mathbf{w}_{k+1,i}$ are bounded in a unit sphere, i.e. $\| \mathbf{v}_k \| \leq 1$, $\| \mathbf{w}_{k+1,i} \| \leq 1, i = 1, \ldots, L$, and iii) (42) and (43) hold for some $\Delta_k \in \tilde{\Omega}_k$, $\Delta_{k+1,i} \in \tilde{\Omega}_{k+1,i}$. Moreover, we provide a state

estimation ellipsoid or box by minimizing its "size". At last, Euclidean estimation error of the state vector and Euclidean estimation error of every entry of the state vector are presented.

To minimize Euclidean estimation error, based on $\mathcal{S}$-procedure [10] and Schur complement, quadratic embedding [52], the multisensor centralized fusion is derived in this section. Based on the assumptions and symbols of in this section, the L sensors centralized[1] fusion bounding ellipsoid is given as follows.

THEOREM 4. *At time $k+1$, based on measurements $\mathbf{y}_{k+1,i}$, $i = 1, 2, \ldots, L$, we have*

I. *a centralized fusion bounding ellipsoid $\mathcal{E}_{\mathbf{x}_{k+1}} = \{x : (\mathbf{x} - \hat{\mathbf{x}}_{k+1})^{\mathrm{T}} \ (\mathbf{P}_{\mathbf{x}_{k+1}})^{-1}(\mathbf{x} - \hat{\mathbf{x}}_{k+1}) \leq 1\}$ can be obtained by solving the optimization problem in the variables $\mathbf{P}_{\mathbf{x}_{k+1}}$, $\hat{\mathbf{x}}_{k+1}$, $\mathbf{S}, \mathbf{T}, \mathbf{G}, \mathbf{S}_i, \mathbf{T}_i, \mathbf{G}_i, i = 1, \ldots, L$, nonnegative scalars $\tau^u \geq 0, \tau^v \geq 0, \tau_i^w \geq 0, i = 1, \ldots, L$,*

$$\min \quad f(\mathbf{P}_{\mathbf{x}_{k+1}}) \quad \textit{subject to} \tag{56}$$

$$(\mathbf{S}, \mathbf{T}, \mathbf{G}) \in \mathcal{P}(\Omega_k), \mathbf{S} \succeq 0, \mathbf{T} \succeq 0, \tag{57}$$

$$(\mathbf{S}_i, \mathbf{T}_i, \mathbf{G}_i) \in \mathcal{P}(\Omega_{k+1,i}), \mathbf{S}_i \succeq 0, \mathbf{T}_i \succeq 0, \qquad i = 1, 2 \ldots, L, \tag{58}$$

$$-\tau^u \leq 0, \ \ -\tau^v \leq 0, \ \ -\tau_i^w \leq 0, \qquad i = 1, 2 \ldots, L, \tag{59}$$

$$-\mathbf{P}_{\mathbf{x}_{k+1}} \prec 0 \tag{60}$$

$$\begin{bmatrix} -\mathbf{P}_{\mathbf{x}_{k+1}} & \Phi(\hat{\mathbf{x}}_{k+1})(\Psi)_{\perp} \\ (\Phi(\hat{\mathbf{x}}_{k+1})(\Psi)_{\perp})^{\mathrm{T}} & (\Psi)_{\perp}^{\mathrm{T}}(\Pi - \Xi)(\Psi)_{\perp} \end{bmatrix} \preceq 0. \tag{61}$$

where

$$\begin{aligned} \Phi(\hat{\mathbf{x}}_{k+1}) \triangleq [\quad & \mathbf{F}_k\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k+1} \quad \mathbf{F}_k\mathbf{E}_k \quad \mathbf{A}_k \\ \vdots \quad & \underbrace{0 \quad 0 \ \ldots \ 0}_{L \ \textit{blocks},\ 0\in\mathbb{R}^{n,n_i}} \quad \vdots \ \mathbf{L}_k \ \vdots \quad \underbrace{0 \ 0 \ \ldots \ 0}_{L \ \textit{blocks},\ 0\in\mathbb{R}^{n,n_{pi}}} \quad], \end{aligned} \tag{62}$$

$\mathbf{E}_k$ is the Cholesky factorization of $\mathbf{P}_{\mathbf{x}_k}$, i.e, $\mathbf{P}_{\mathbf{x}_k} = \mathbf{E}_k(\mathbf{E}_k)^{\mathrm{T}}$,

$$\Xi \triangleq \mathrm{diag}(1 - \tau^u - \tau^v - \sum_{i=1}^{L} \tau_i^w, \overbrace{\tau^u\mathbf{I}, \tau^v\mathbf{I}}^{\mathbf{I}\in\mathbb{R}^{n,n}}, \vdots \underbrace{\tau_1^w\mathbf{I}, \tau_2^w\mathbf{I}, \ldots, \tau_L^w\mathbf{I}}_{I\in\mathbb{R}^{n_i,n_i}},$$

[1]Term *centralized* means that the fusion center can receive all $\mathbf{y}_{k,i}$ to derive estimate coverage set at each time k.

$$\vdots \ \overbrace{0}^{0\in\mathbb{R}^{n_p,n_p}} \ , \vdots \ \underbrace{0 \ 0 \ \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_{pi},n_{pi}}} \ ) \tag{63}$$

$$\Pi \triangleq \Upsilon(\mathbf{S},\mathbf{T},\mathbf{G}) + \sum_{i=1}^{L} \Upsilon_i(\mathbf{S}_i,\mathbf{T}_i,\mathbf{G}_i) \tag{64}$$

$$\Upsilon(\mathbf{S},\mathbf{T},\mathbf{G}) \triangleq (\tilde{\Upsilon})^{\mathrm{T}} \begin{bmatrix} \mathbf{T} & \mathbf{G} \\ \mathbf{G}^{\mathrm{T}} & -\mathbf{S} \end{bmatrix} \tilde{\Upsilon} \tag{65}$$

$$\tilde{\Upsilon} \triangleq \begin{bmatrix} \mathbf{R}_{\mathbf{F}_k}\hat{\mathbf{x}}_k & \mathbf{R}_{\mathbf{F}_k}\mathbf{E}_k & \mathbf{R}_{\mathbf{A}_k} & \vdots & \overbrace{0 \ 0 \ \ldots \ 0}^{L \ \ blocks,\, 0\in\mathbb{R}^{n_q,n_i}} & \vdots \\ 0 & 0 & 0 & \vdots & \underbrace{0 \ 0 \ \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_p,n_i}} & \vdots \\ & & & \mathbf{D}_k \vdots & \overbrace{0 \ 0 \ \ldots \ 0}^{L \ \ blocks,\, 0\in\mathbb{R}^{n_q,n_{pi}}} & \\ & & & \mathbf{I} \vdots & \underbrace{0 \ 0 \ \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_p,n_{pi}}} & \end{bmatrix} \tag{66}$$

$$\Upsilon_i(\mathbf{S}_i,\mathbf{T}_i,\mathbf{G}_i) \triangleq (\tilde{\Upsilon}_i)^{\mathrm{T}} \begin{bmatrix} \mathbf{T}_i & \mathbf{G}_i \\ (\mathbf{G}_i)^{\mathrm{T}} & -\mathbf{S}_i \end{bmatrix} \tilde{\Upsilon}_i \tag{67}$$

$$\tilde{\Upsilon}_i \triangleq \begin{bmatrix} \mathbf{R}_{\mathbf{H}_{k+1,i}}\mathbf{F}_k\hat{\mathbf{x}}_k & \mathbf{R}_{\mathbf{H}_{k+1,i}}\mathbf{F}_k\mathbf{E}_k & \mathbf{R}_{\mathbf{H}_{k+1,i}}\mathbf{A}_k & & & \\ 0 & 0 & 0 & & & \\ \vdots \overbrace{0 \ 0 \ \ldots \mathbf{R}_{\mathbf{B}_{k+1,i}} \ldots \ 0}^{L \ \ blocks,\, 0\in\mathbb{R}^{n_{qi},n_i}} \vdots & & \mathbf{R}_{\mathbf{H}_{k+1,i}}\mathbf{L}_k \vdots & \overbrace{0 \ 0 \ \ldots \mathbf{D}_{k+1,i} \ldots \ 0}^{L \ \ blocks,\, 0\in\mathbb{R}^{n_{qi},n_{pi}}} & \\ \vdots \underbrace{0 \ 0 \ \ldots 0 \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_{pi},n_i}} \vdots & & 0 \ \vdots & \underbrace{0 \ 0 \ \ldots \mathbf{I} \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_{pi},n_{pi}}} & \end{bmatrix}; \tag{68}$$

$(\Psi)_{\perp}$ *is the orthogonal complement of* Ψ*, i.e, a basis of the null space of* Ψ*,*

$$\Psi \triangleq [(\Psi_1)^{\mathrm{T}} \ \ (\Psi_2)^{\mathrm{T}} \ \ldots \ (\Psi_L)^{\mathrm{T}}]^{\mathrm{T}} \tag{69}$$

$$\Psi_i \triangleq [\ \mathbf{H}_{k+1,i}\mathbf{F}_k\hat{\mathbf{x}}_k - \mathbf{y}_{k+1,i} \quad \mathbf{H}_{k+1,i}\mathbf{F}_k\mathbf{E}_k \quad \mathbf{H}_{k+1,i}\mathbf{A}_k \\ \vdots \ \underbrace{0 \ 0 \ \ldots \mathbf{B}_{k+1,i} \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_i,n_i}} \vdots \ \ \mathbf{H}_{k+1,i}\mathbf{L}_k \vdots \ \underbrace{0 \ 0 \ \ldots \mathbf{L}_{k+1,i} \ldots \ 0}_{L \ \ blocks,\, 0\in\mathbb{R}^{n_i,n_{pi}}} \ ]; \tag{70}$$

II. *If take different waights in objective $f(\mathbf{P}_{\mathbf{x}_{k+1}})$ given in (56) and (48), the intersection of the derived multiple estimate sets is a multi-algorithm fusion, and make the derived Eucdicen error as minimal as possible.*

The proof of the theorem can be found in [77].

5.2. Numerical examples. In this subsection, we give an example to evaluate the performance of the centralized fusion C-Algorithm (for details, see Algorithm 7.2.1 in our book [80]) and the distributed fusion D-Algorithm (for details, see Algorithm 7.2.2 in our book [80]). Moreover, based on the weighted objective (48) and (49), the fusion of multiple Algorithms is used to reduce Euclidean estimation error.

Consider a two-state tracking system (see [50, 67]) with uncertainty of time stamps and measurements. The states consist of position, velocity, and the measurements consist of position plus noise. Here, we assume that the noises are bounded. The uncertain state and measurement equations are given as

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & T + 0.2\delta_{\mathbf{x}} T \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{A}_k \mathbf{v}_k, \tag{71}$$

$$\mathbf{y}_{k,i} = [1 + 0.001\delta_{\mathbf{y},i} \qquad 0]\mathbf{x}_k + \mathbf{B}_{k,i}\mathbf{w}_{k,i},\ i = 1, 2, \tag{72}$$

where $T = 1s$ is the sample time interval, model uncertainty parameters $||\delta_{\mathbf{x}}|| \leq 1$, $||\delta_{\mathbf{y},i}|| \leq 1, i = 1, 2$, $\mathbf{v}_k$ and $\mathbf{w}_{k,i}$ are the process noise and the measurement noises taking value in a unit spheres, i.e., $\| \mathbf{v}_k \| \leq 1$, $\| \mathbf{w}_{k,i} \| \leq 1$ respectively. If we denote

$$\mathbf{Q}_k = 3^2 \begin{bmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{bmatrix} q, \tag{73}$$

$$\mathbf{R}_{k,i} = 3^2 R, \tag{74}$$

where $q = 0.5\ m^2/s^3$ is the power spectral density of the continuous time process noise, $R = 1\ m^2$, then the matrices $\mathbf{A}_k, \mathbf{B}_{k,i}$ are satisfied with $\mathbf{Q}_k = \mathbf{A}_k \mathbf{A}_k^{\mathrm{T}}$, $\mathbf{R}_{k,i} = \mathbf{B}_{k,i}{\mathbf{B}_{k,i}}^{\mathrm{T}}$. In [**?**, 67], the noises are assumed to be white noise processes with covariance $\mathbf{Q}_k/3^2, \mathbf{R}_{k,i}/3^2$. The LFR uncertainty representation (44)–(45) specializes to $\mathbf{L}_k = [1 \quad 0], \mathbf{D}_k = 0, \mathbf{R}_{\mathbf{F}_k} = [0 \quad 0.2T], \mathbf{R}_{\mathbf{A}_k} = [0 \quad 0]$, $\mathbf{L}_{k,1} = 1, \mathbf{D}_{k,1} = 0, \mathbf{R}_{\mathbf{H}_{k,1}} = [0.001 \quad 0], \mathbf{R}_{B_{k,1}} = 0$, $\mathbf{L}_{k,2} = 1, \mathbf{D}_{k,2} = 0, \mathbf{R}_{\mathbf{H}_{k,2}} = [0.001 \quad 0], \mathbf{R}_{B_{k,2}} = 0$. The scaling subspaces are in this case described by $\mathbf{S} = \mathbf{T} = \lambda$ (a scalar), $\mathbf{G} = 0$ and $\mathbf{S}_i = \mathbf{T}_i = \lambda_i$ (a scalar), $\mathbf{G}_i = 0$, $i = 1, 2$. The uniform distribution in $[-1, 1]$ for uncertain parameter and noises is used in simulation.

The following simulation results are the mean of 200 Monte Carlo runs under Matlab R2007b with YALMIP (see [56]). Figure 6 presents a comparison of the Euclidean error bounds along position direction for sensors 1, 2 using C-Algorithm ($L = 1$) respectively and for the fusion center using the centralized fusion C-Algorithm

($L = 2$) and the distributed fusion D-Algorithm respectively. The objective (48) or (49) with $(w_1, w_2) = (0.5, 0.5)$ is used. Comparisons between the ellipsoid bounding and the box bounding are presented simultaneously in Figure 6. Figures 7–8 present comparisons of the Euclidean error bounds along velocity direction and comparisons of the Euclidean error bounds of the state vectors respectively.

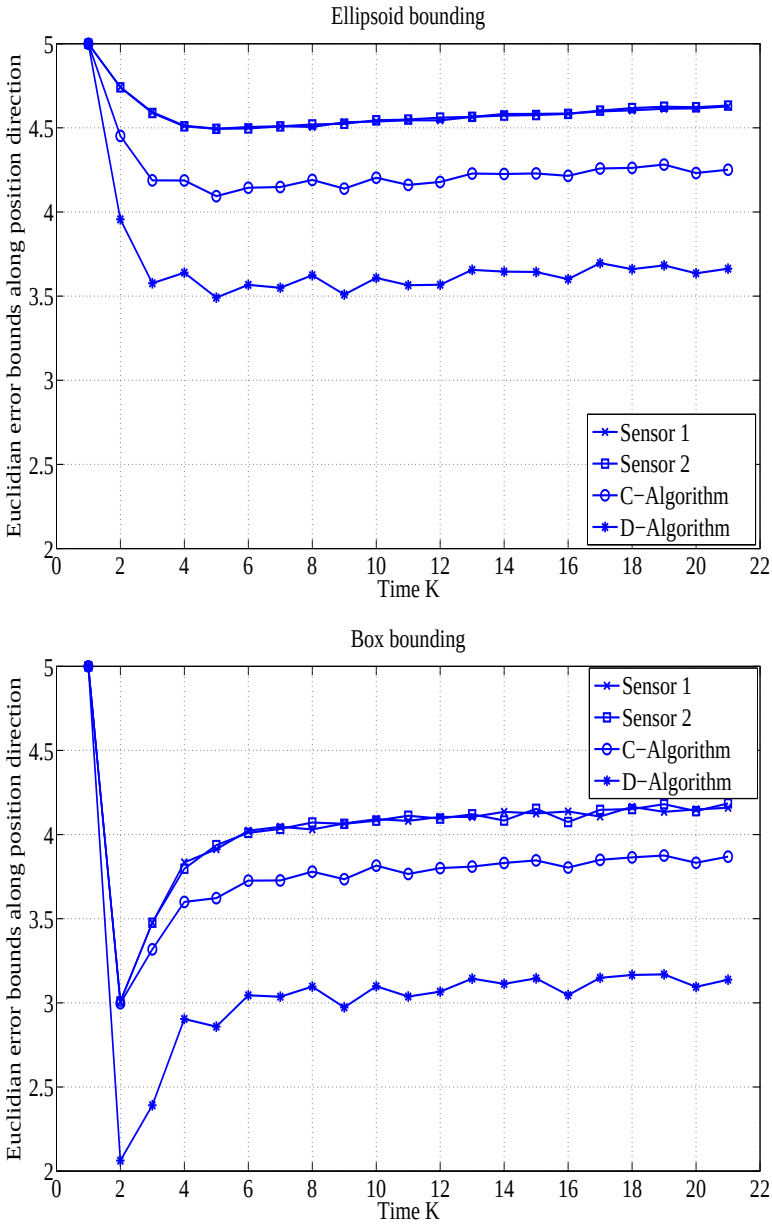

FIG. 6. *Comparison of the Euclidean error bounds along position direction.*

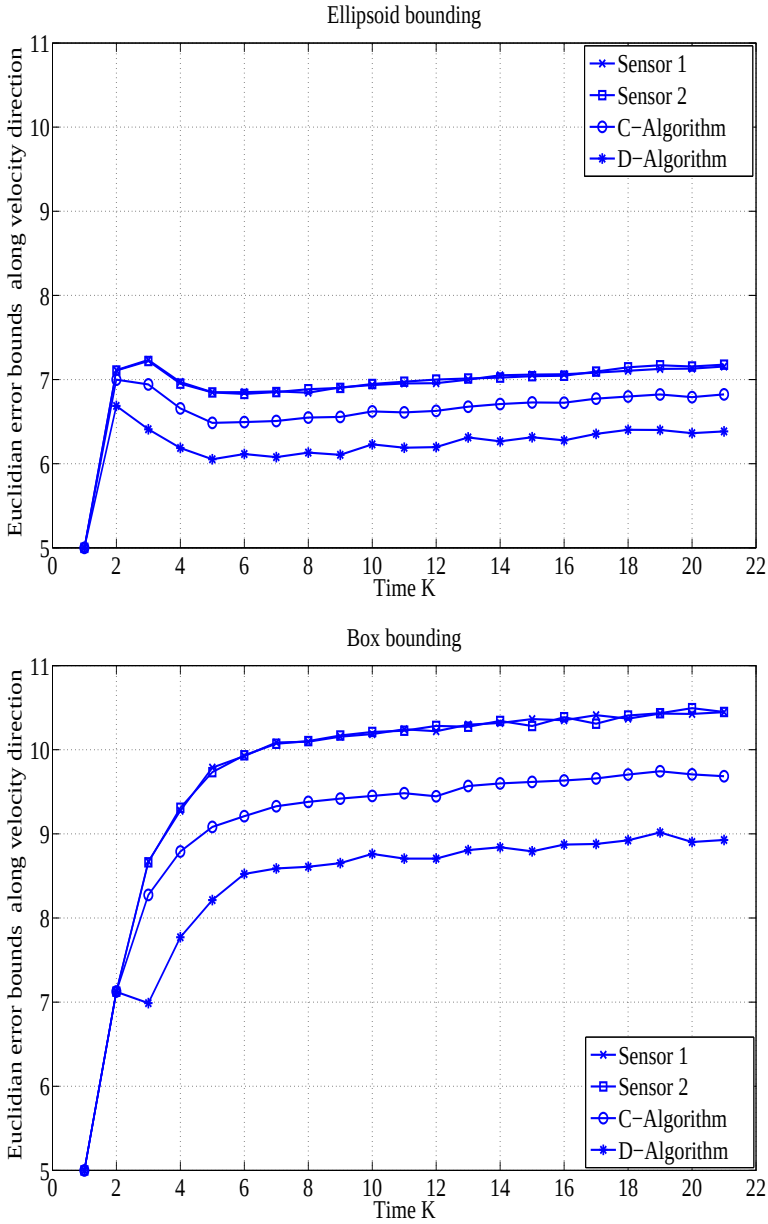

FIG. 7. *Comparison of the Euclidean error bounds along velocity direction.*

From Figures 6–8, the following observations can be seen.

1. Performance of both the centralized fusion C-Algorithm and the distributed fusion D-Algorithm is better than that of sensors.
2. Performance of the distributed fusion D-Algorithm is better than that of the centralized fusion C-Algorithm except the ellipsoid bounding case for

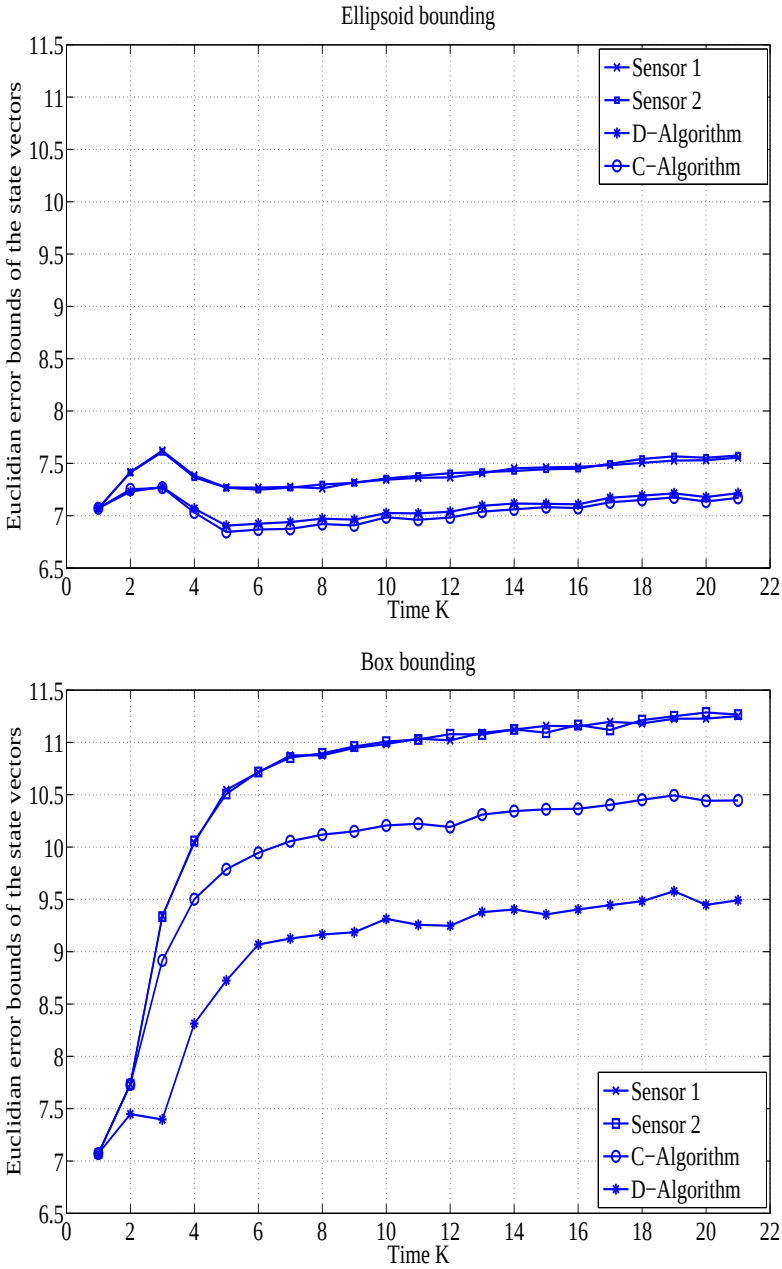

FIG. 8. *Comparisons of the Euclidean error bounds of the state vectors.*

the Euclidean error bounds of the state vectors in Figure 8. The reasons may involve: i) the former is derived from much more estimation ellipsoids or boxes including the fusion ellipsoid and all sensor estimation ellipsoids or boxes (not unprocessed observation sets) with different shapes than the

latter; ii) for the ellipsoid bounding approach, the Euclidean error bound of the state vector is calculated by the maximum eigenvalue of the shape matrix of the bounding ellipsoid but not based on the Euclidean error bounds along position and velocity directions, respectively.

3. Figure 6 shows that the Euclidean error bounds along position direction based on the box bounding are better than those based on the ellipsoid bounding. Figures 7–8 show that the Euclidean error bounds along velocity direction and the Euclidean error bounds of the state vectors based on the ellipsoid bounding are better than those based on the box bounding.

5.2.1. Figures 9–11 for fusion of multiple algorithms. Figure 9 presents a comparison of the Euclidean error bounds along position direction of the fusion of multiple D-Algorithm between the ellipsoid bounding and the box bounding methods. For the sensors 1-2, three objectives (48) or (49) with $(w_1, w_2) = (0.99, 0.01)$, $(w_1, w_2) = (0.01, 0.99)$, $(w_1, w_2) = (0.5, 0.5)$ are used in C-Algorithm ($L = 1$). Then the corresponding ellipsoids/boxes are sent to the fusion center. For the fusion center, based on the received ellipsoids/boxes, three D-Algorithm with using the above three objectives respectively are fused. Similarly, Figures 10–11 present comparisons of the Euclidean error bounds along velocity direction and the Euclidean error bounds of the state vectors of the fusion of multiple D-Algorithm respectively.

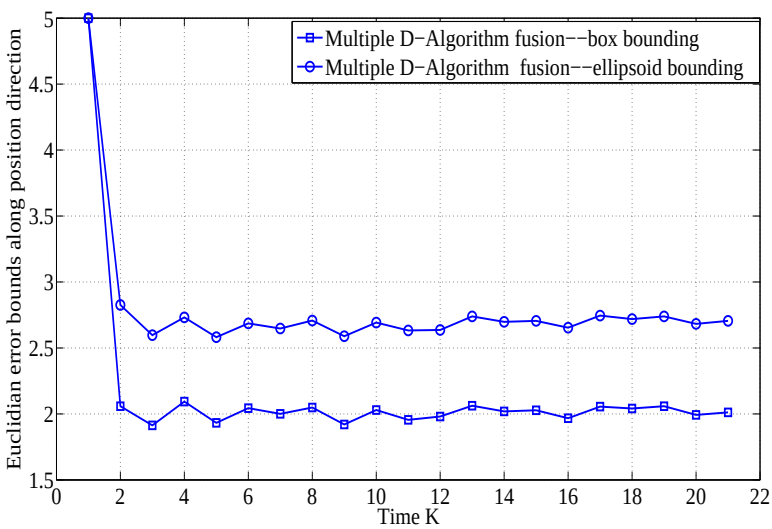

FIG. 9. *Comparisons of the Euclidean error bounds along position direction.*

Figures 9–11 also show that the Euclidean error bounds along position direction based on the box bounding are better than those based on the ellipsoid bounding. On

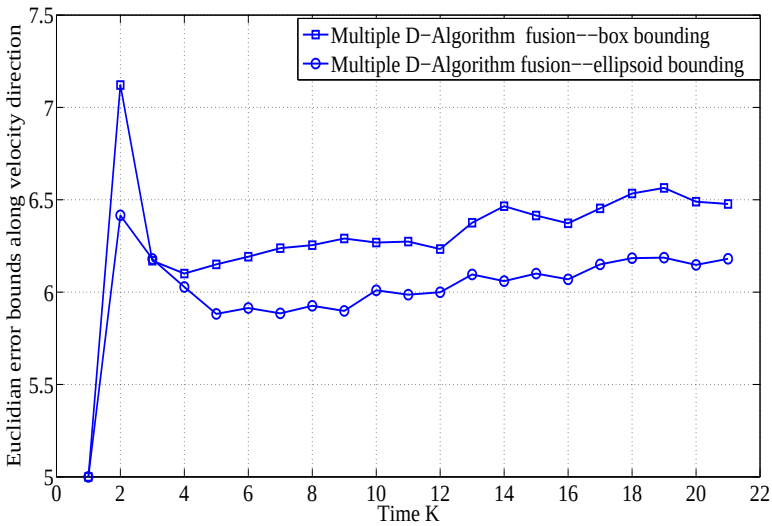

FIG. 10. *Comparisons of the Euclidean error bounds along velocity direction.*

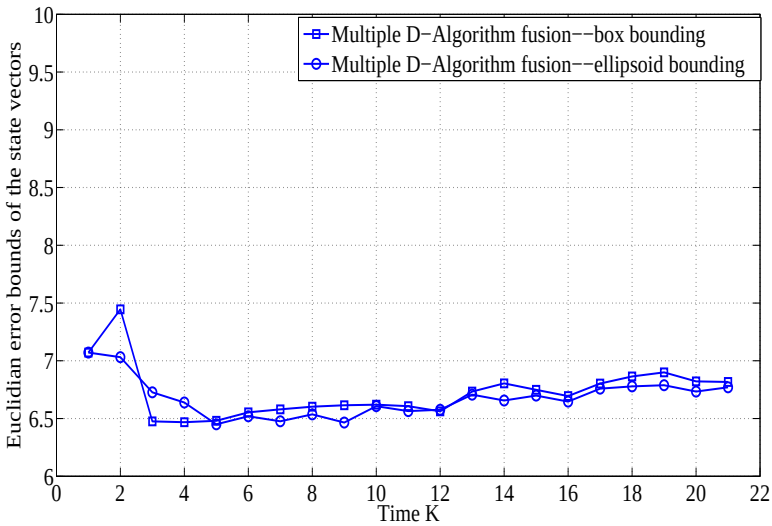

FIG. 11. *Comparisons of the Euclidean error bounds of the state vectors.*

the contrary, the Euclidean error bounds along velocity direction and the Euclidean error bounds of the state vectors based on the ellipsoid bounding are better than those based on the box bounding.

Comparing Figures 6–8 with Figures 9–11, it can be observed that the Euclidean estimation error of every entry of the state vector and the Euclidean estimation error of the state vector can be significantly decreased by the fusion of multiple algorithms.

In a summary, we have proposed the centralized fusion C-Algorithm and the distributed fusion D-Algorithm. Their performances are better than that of sensors. Simulations show that the Euclidean estimation error along the every state direction of the distributed fusion D-Algorithm is smaller than those of the centralized fusion C-Algorithm. Moreover, the Euclidean estimation error of the state vectors can be significantly improved by estimation fusion of multiple algorithms.

5.3. Comparison with Kalman Filtering. We have the following observations in comparison with Kalman filtering:

- No close solution, more computation requirement but still can be efficiently solved;
- No analysis on observability and controllability;
- More robust for biased models;
- Can be converted to Euclidean Error (bound) estimation and further minimize it by ***multisensor and multi-algorithm fusion***.

It follows that when the assumptions on optimality for Kalman filtering can be fulfilled, particularly, model precise knowledge can be known, Kalman filtering is a good choice for state estimation for dynamic systems; otherwise, one should try to find the bias bound of model uncertainty and choose set-valued filtering. Besides, if large error with nonzero probability can be tolerated, choose the former, otherwise, choose the latter.

6. Conclusions. We have given a brief survey to some developments form Kalman filtering for the stochastic dynamic systems to set-valued filtering for the dynamic systems with bounded but uncertain model biases. The main focus on this survey is is to present some progress in the two filters and compare their own advantage and weakness in order to provide some guidance for people to decide which formulation for dynamic systems with uncertainty and the corresponding filtering method should be chosen in practical applications. Since computational burden for Kalman filtering is easily affordable and theoretical properties are clear, we should choose it to handle a stochastic dynamic system while the system model knowledge can be given relevant precisely. Otherwise, we should formulate the model to be a biased dynamic system and adopt the minimizing Euclidean error algorithm to estimate the state. A future work is to study the later theoretical properties although it looks very hard.

REFERENCES

[1] R.E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASMECJournal of Basic Engineering, vol. 82, no. Series D, pp. 35C45, 1960.

[2] R.E. Kalman and R. S. Bucy, *New results in linear filtering and prediction problems*, Transactions of the ASME-Journal of Basic Engineering, vol. 83, no. Series D, pp. 95C108, 1961.

[3] A.E. Bryson, and D. E. Johansen, *Linear filtering for time-varying systems using measurements containing colored noise*, IEEE Transactions on Automatic Control, 10:1(1965), pp. 4–10.

[4] A.E. Bryson, and L. Henrikson, *Estimation using sampled data containing sequentially correlated noise*, Journal of Spacecraft and Rockets, 5:6(1968), pp. 662–665.

[5] F.C. Schweppe, *Recursive state estimation: unknown but bounded errors and system inputs*, IEEE Trans. Automat. Contr., 13:1(1968), pp. 22–28.

[6] H.S. Witsenhausen, *Sets of possible states of linear systems given perturbed observations*, IEEE Trans. Automat. Contr., 13:5(1968), pp. 556-558.

[7] F.C. Schweppe, *Recursive state estimation: unknown but bounded errors and system inputs.* IEEE Transactions on Automatic Control, 13:1(1968), pp. 22–28.

[8] D.P. Bertsekas, and I. B. Rhodes, *Recursive state estimation for a setmembership description of uncertainty*, IEEE Transactions on Automatic Control, 16(1971), pp. 117–128.

[9] B.D.O. Anderson and J.B. Moore, *Linear Optimal Control*, Englewood Cliffs, NJ: Prentice-Hall, 1971.

[10] V.A. Yakubovich, *S-procedure in nonlinear control theory*, Vestnik Leningrad University, pages 62-77, 1971. (English translation in Vestnik Leningrad Univ. 4:73-93, 1977).

[11] A. Albert, *Regression and the Moore-Penrose Pseudoinverse.* Academic Press,1972.

[12] F.C. Schweppe, *Uncertain Dynamic Systems.* Englewood Cliffs: Prentice- Hall, 1973.

[13] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer, 1975.

[14] B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Englewood Cliffs, NJ: Prentice-Hall, 1979.

[15] P.S. Maybeck, *Stochastic Models, Estimation, and Control*, Volume I, II, III. Academic Press, 1979, 1982.

[16] C. Chang, and L. Youens, *Measurement correlation for multiple sensor tracking in a dense target environment*, IEEE Transactions on Automatic Control, 27:6(1982), pp.1250–1252.

[17] E.Fogel, and Y.F. Huang, *On the value of information in system identification-bounded noise case*, Automatica, 18(1982), pp. 229–238.

[18] W.L. De Koning, *Optimal Estimation of linear discrete-time systems with stochastic parameters*, Automatica, 20:1(1984), pp. 113–115.

[19] G.C. Goodwin and K.S. Sin, *Adaptive Filtering*, Prediction and Control. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[20] P.R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*, Prentice-Hall, 1986.

[21] P.E. Caines, *Linear Stochastic Systems*, John Wiley & Sons, 1988.

[22] H.A.P. Blom and Y. Bar-Shalom, *The interacting multiple model algorithm for systems with Markovian switching coefficients*, IEEE Trans. Automat. Contr., 33:8(1988), pp.780-783.

[23] E.Walter, and H. Piet-Lahanier, *Exact recursive polyhedral description of the feasible parameter set for bounded-error models*, IEEE Transactions on Automatic Control, 34(1989), pp. 911–915.

[24] E. Walter, *Parameter identifications with error bound*, Special issue of Mathematics and Computers in Simulation, 32:5-6(1990), pp. 447–638.

[25] H.F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*, Boston: Birkhäuster, 1991.

[26] K.H. Fan, A.L. Tits, and J.C. Doyle, *Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics*,IEEE Transactions on Automatic Control, 36(1991), pp. 25–38.

[27] O.N. Kiselev and B.T. Polyak, *Ellipsoidal estimation with respect to a generalized criterion*, Automation and Remote Control, 52(1992), pp. 1281–1292.

[28] J.Deller, R.M. Nayeri, and S.F. Odeh, *Least-square identification with error bounds for real-time signal processing and control*, Proceeding of IEEE, 81(1993), pp. 813–849.

[29] P.L. Combettes, *The foundations of set-theoretic estimation*, Proceeding of IEEE, 81(1993), pp. 182–208.

[30] R.D. Hilton, D.A. Martin, and W.D. Blair, *Tracking with time-delayed data in multisensor systems*, In: NSWCDD/TR-93/351, Dahlgren, VA, 1993.

[31] F.L. Chernousko, *State estimation of dynamic systems*. Boca Raton, FL, CRC Press, 1993.

[32] X.R. Li and Y. Bar-Shalom, *Design of an interacting multiple model algorithm for air traffic control tracking*, IEEE Trans. Contr. Syst. Technol., 1:3(1993), pp. 186-194.

[33] J.P. Norton, *Special issues on bounded-error estimation: Part 1*, International Journal of Adaptive Control and Signal Processing, 8(1994), pp. 1–118.

[34] Y. Bar-Shalom and X.R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[35] J.P. Norton, *Special issues on bounded-error estimation: Part 2*, International Journal of Adaptive Control and Signal Processing, 9(1995), pp. 1–132.

[36] D.G. Maskarov and J.P. Norton, *State bounding with ellipsoidal set description of the uncertainty*, Int. J. Control, 65:5(1996), pp. 847–866.

[37] A.B. Kurzhanski and I. Valyi, *Ellipsoidal Calculus for Estimation and Control*. Boston,Massachusetts: Birkhäuser, 1997.

[38] L. Vandenberghe, S. Boyd, and S.P. Wu, *Determinant maximization with linear matrix inequality constraints*, SIAM Journal on Matrix Analysis and Applications, 19:2(1998), pp. 499–533.

[39] A.I. Matasov, *Estimators for Uncertain Dynamic Systems*. Kluwer Academic Publishers, 1998.

[40] S.S. Blackman and R.F. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.

[41] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[42] X.R. Li, X.R. Zhi, and Y.M. Zhang, *Multiple-model estimation with variable structure-part III: Model-group switching algorithm*, IEEE Trans. Aerosp. Electron. Syst, 35:1(1999), pp. 225–241.

[43] J.A. Mulder, Q.P. Chu, J.K. Sridhar, J. H. Breeman, and M.Laban, *Non-linear aircraft flight path reconstruction review and new advances*, Progress in Aerospace Sciences, 35(1999), pp. 673–726.

[44] J.S. Shamma and K.Y. Tu, *Set-valued observers and optimal disturbance rejection*, IEEE Transactions on Automatic Control, 44:2(1999), pp. 253–264.

[45] Y. Zhu, *Efficient recursive state estimator for dynamic systems without knowledge of noise covariances*, IEEE Trans. Aerospace and Electronic Systems, No.1, 1999.

[46] T. Kailath, A.H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[47] X.R. Li, *Multiple-model estimation with variable structure-part II: Model-set adaptation*, IEEE Trans. Automat. Contr., 45:11(2000), pp. 2047–2060.

[48] M. Mallick, S. Coraluppi, and C. Carthel, *Advances in asynchronous and decentralized estimation*. In: Proceedings of 2001 IEEE Aerospace Conference, 2001.

[49] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis*. Springer, 2001.

[50] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, New York: Wiley, 2001.

[51] C. Durieu, E. Walter, and B.T. Polyak, *Multi-Input Multi-Output Ellipsoidal State Bounding*, Journal of Optimization Theory and Applications 111:2(2001), pp. 273–303.

[52] L. El Ghaoui and G. Calafiore, *Robust Fltering for Discrete-Time Systems with Bounded Noise and Parametric Uncertainty*, IEEE Trans. Automat. Contr., 46:7(2001), pp. 1084–1089.

[53] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using Matlab,* second edition, John Wiley & Sons, Inc. 2001.

[54] E.W. Nettleton and H. Durrant-Whyte, *Delayed and Asequent Data in Decentralized Sensing Networks*. In: Proceedings of SPIE Conference, vol.4571, 2001.

[55] Y. Bar-Shalom, *Update with Out-of-Sequence Measurements in Tracking: Exact Solution*, IEEE Trans. Aeros. and Electr. Syst., 38:3(2002), pp.769–778.

[56] J. Löfberg, *YALMIP: A toolbox for modelling and optimization in Matlab*, in: Proc. IEEE CACSD Symposium, Taipei, Taiwan, pp. 284-289, Sept. 2004.

[57] Y. Bar-Shalom, M. Mallick, H. Chen, and R. Washburn, *One-Step Solution for the Multistep Out-of-Sequence Measurement Problem in Tracking*, IEEE Trans. Aerosp. Electron. Syst, 40:1(2004), pp. 27–37.

[58] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[59] G. Calafiore, and L. El Ghaoui, *Ellipsoidal bounds for uncertain equations and dynamical systems*, Automatica, 40(2004), pp. 773–787.

[60] G. Calafiore and L. El Ghaoui, *Ellipsoidal bounds for uncertain equations and dynamical systems*, Automatica, 40(2004), pp. 773–787.

[61] S.J. Julier, J.K. Uhlmann, *Unscented filtering and nonlinear estimation*, Proc. IEEE, 92:3(2004), pp. 401–422,

[62] B.T. Polyak, S.A. Nazin, C. Durieu, and E. Walter, *Ellipsoidal coefficient or state estimation under model uncertainty*, Automatica, 40(2004), pp. 1171–1179.

[63] A. El-Keyi, T. Kirubarajan, and A.B. Gershman, *Robust Adaptive Beamforming Based on the Kalman Filter*, IEEE Trans. Signal Process., 53:8(2005).

[64] K.K.C. Yu, *An Adaptive Kalman flter for Dynamic Harmonic State Estimation and Harmonic Injection Tracking*, IEEE Trans. Power Delivery, 20:2(2005).

[65] K. Zhang, X.R. Li, and Y. Zhu, *Optimal update with out-of-sequence measurements*, IEEE Transactions on Signal Processing, 53:6(2005), pp. 1992–2004.

[66] R. Rothrock, *The need to remove out-of-sequence measurements*, Presented at *the Tracking and CID Alternatives Identification and Evaluation Team (TCAIET) Meeting*. Arlington, VA., 2006.

[67] D. Simon. *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches.* Wiley-Interscience, 2006.

[68] M.S. Grewal and A.P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*, 3rd Edition Wiley-IEEE Press, 2008.

[69] Y. Luo, Y. Zhu, D. Luo, J. Zhou, E. Song, and D. Wang, *Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications*, Sensors, 8:12(2008), pp. 8086–8103.

[70] Y. Bar-Shalom and H. Chen, *Removal of out-of-sequence measurements from tracks*, IEEE Transactions on Aerospace and Electronic Systems, 45:2(2009), pp. 612–619.

[71] C.E. Seah and I. Hwang, *State Estimation for Stochastic Linear Hybrid Systems with Continuous-State-Dependent Transitions: An IMM Approach*, IEEE Trans. Aerosp. Electron. Syst, 45:1(2009), pp. 376–392.

[72] X. Shen, Y. Zhu, E. Song, and Y. Luo, *Optimal centralized update with multiple local out-of-sequence measurements*, IEEE Transactions on Signal Processing, 57:4(2009), pp. 1551–1562.

[73] X. Shen, E. Song, Y. Zhu, Y. Luo, *Globally Optimal Distributed Kalman Fusion with Local Out-of-Sequence-Measurement Updates*, IEEE Trans. on, Automatic Control, 54:8(2009), pp. 1928–1934.

[74] P. Jiang, J. Zhou, and Y. Zhu, *Globally optimal Kalman filtering with finite-time correlated noises*, in: The 49th IEEE Conference on Decision and Control, Atlanta, GA, 5007–5012, 2010.

[75] X. Shen, Y. Luo, Y. Zhu, E. Song, and Z. You, *Globally Optimal Flight Path Update with Adding or Removing Out-of-Sequence Measurements*, Automatica, 46:9(2010), pp.1437–1442.

[76] X. Shen, Y. Zhu, and Y. Luo, *Optimal State Estimation of Linear Discrete-Time Systems with Correlated Random Parameter Matrices*, Proceedings of the 30th Chinese Control Conference, pp.1488-1493, Yantai, China July 22-24, 2011.

[77] X. Shen, Y. Zhu, E. Song, and Y. Luo, *Minimizing Euclidean state estimation error for linear uncertain dynamic systems based on multisensor and multi-algorithm fusion*, IEEE Transactions on Information Theory, 57:10(2011), pp. 7131–7146.

[78] Y. Luo, Y. Zhu, X. Shen, and E. Song, *A Novel Data Association Algorithm Based on Integrated Random Coefficient Matrices Kalman Filtering*, IEEE Trans. Aerosp. Electron. Syst, 48:1(2012), pp.144–158.

[79] F. Li, J. Zhou, and D. Wu, *Optimal Filtering for Systems with Finite-Step Autocorrelated Noises and Multiple Packet Dropouts*, Aerospace Science and Technology, 2012.

[80] Y. Zhu, J. Zhou, X. Shen, E. Song, and Y. Luo, *Networked Multisensor Decision and Estimation Fusion: Based on Advanced Mathematical Methods*, CRC Press, 2012.