SCIENCE CHINA
Information Sciences

• **RESEARCH PAPER** •
Special Focus on Analysis and Synthesis for Stochastic Systems

December 2017, Vol. 60 120205:1–120205:14
doi: 10.1007/s11432-017-9140-x

Optimal fusion estimation for stochastic systems with cross-correlated sensor noises

Liping YAN*, Yuanqing XIA & Mengyin FU

School of Automation, Key Laboratory of Intelligent Control and Decision of Complex Systems, Beijing Institute of Technology, Beijing 100081*, China*

Received April 14, 2017; accepted June 14, 2017; published online November 9, 2017

Abstract This paper is concerned with the optimal fusion of sensors with cross-correlated sensor noises. By taking linear transformations to the measurements and the related parameters, new measurement models are established, where the sensor noises are decoupled. The centralized fusion with raw data, the centralized fusion with transformed data, and a distributed fusion estimation algorithm are introduced, which are shown to be equivalent to each other in estimation precision, and therefore are globally optimal in the sense of linear minimum mean square error (LMMSE). It is shown that the centralized fusion with transformed data needs lower communication requirements compared to the centralized fusion using raw data directly, and the distributed fusion algorithm has the best flexibility and robustness and proper communication requirements and computation complexity among the three algorithms (less communication and computation complexity compared to the existed distributed Kalman filtering fusion algorithms). An example is shown to illustrate the effectiveness of the proposed algorithms.

Keywords optimal estimation, distributed fusion, Kalman filter, cross-correlated noises, linear transformation

Citation Yan L P, Xia Y Q, Fu M Y. Optimal fusion estimation for stochastic systems with cross-correlated sensor noises. Sci China Inf Sci, 2017, 60(12): 120205, doi: 10.1007/s11432-017-9140-x

1 Introduction

Estimation fusion, or data fusion for estimation, is the problem of how to best utilize useful information contained in multiple sets of data for the purpose of estimating a quantity, e.g., a parameter or process [1, 2]. It originated in the military field, and is now widely used in military and civilian fields, e.g., target tracking and localization, guidance and navigation, surveillance and monitoring, due to its improved estimation accuracy, enhanced reliability and survivability.

Most of the earlier researches were based on the assumption of cross-independent sensor noises. In the practical applications, most of multisensor systems often have correlated noises when the dynamic process is observed in a common noisy environment [3]. In this case, the traditional centralized fusion is also applicable and is still optimal in the sense of LMMSE (linear minimum mean square error). However, the computation complexity is huge.

* Corresponding author (email: ylp@bit.edu.cn)

For most of practical cases, there is correlative between process and measurement noises (Correlation I) and among measurement noises (Correlation II) [4]. Some people are working on the optimal Kalman filtering fusion with cross-correlated sensor noises in recent years. In [3], Song et al. presented the distributed fusion algorithms both without feedback and with feedback. While, augmentations of some system parameters are existed. In [5], distributed sequential estimation of a nonrandom parameter over noisy communication links was considered, where the observations are correlated spatially across the sensor field, and a recursive algorithm for updating the sequential estimator was derived. By reconstructing the measurements, Ref. [6] also deduced the distributed Kalman filtering fusion with feedback and without feedback, in which the noise is decoupled in sequence by the method given in [7]. By these methods, the augmentation of measurements and measurement matrices are needed. The fusion algorithm given in [8] is similar. These methods have more complex form compared to [3], but they have the advantage of capability of handling the correlation of measurement noise and system noise besides decoupling the measurement noises. Ref. [9] studied the distributed fusion when the measurement noises are correlated across sensors and with the system noise at the same time step. When the noise of different sensors are cross-correlated and also coupled with the system noise of the previous step, we derive the optimal sequential fusion and optimal distributed fusion algorithm in [10], and generate this result to the fusion of multirate sensor cases [11, 12]. When there is correlation between the process noise and the measurement noise and among measurement noises, a distributed weighted robust Kalman filter fusion algorithm is derived for uncertain systems with multiple sensors in [13]. A novel decentralized cubature Kalman fusion algorithm is presented in [4] for nonlinear systems with noise Correlations I and II.

For fusion of the sensors with correlated noises, there are some literatures using linear transforms. Li et al. first propose this idea in [1], where the optimal batch fusion and distributed fusion have been obtained in a unified form. In [14, 15], correlated measurement fusion Kalman filtering algorithms are obtained based on orthogonal transformation. By using the Cholesky factorization, the coupled noises can be decoupled and the optimal state estimations were derived in [16–18]. In [19], the authors proved that the sufficient condition of the lossless transform for distributed estimation with cross-correlated measurement noises is the transformation matrix of full column rank.

To summarize, there are some literatures deal with optimal Kalman filtering fusion of the dynamic systems with coupled noises. When the measurement noises are correlated, the centralized fusion can still be used and it is optimal in the sense of LMMSE. However, the computation and power requirements are too huge to be practical. Many people intend to generate the distributed fusion algorithm. While, the distributed fusion algorithms are not given explicitly in [1, 17, 18]. The explicit distributed fusion algorithms are presented in [3, 6, 7], but augmentations of some system parameters are existed, and therefore it is not optimal when the computation performance is concerned.

With the development of internet, networked fusion become the hot topic in the field of information fusion, in which the measurements or the local estimations are transmitted to the fusion center through the communication channel. Constrained by the limited resources and communication bandwidth, the computation and communication efficiencies are a problem to which we should pay much attention. For networked state estimation, there are many algorithms could be effective, including the study conceren random delay of the observations [20], faulty of the measurements [21], mismatch of the measurement noise covariance [22]. Ref. [23] analyzed the observable degree of a mobile target tracking system for wireless sensor networks, and pointed out that it is potentially hopeful to achieve an effective function to study estimation performance of tracking estimators by directly using the observable degree. There has been discussion about compression in the hope of reducing the communication requirements in the literatures [24–27]. A sufficient condition and a necessary and sufficient condition were given in [24, 28], respectively, for lossless of performance for distributed fusion with compressed data. A suboptimal and optimal compression rules were derived in [25, 26], respectively, for estimation fusion. In WSNs, the energy-saving related estimation algorithms are researched recently, the literatures include data compression methods [27, 29], quantization based methods [30–32], and the methods to slow down the information transmitting rate in the sensors [33–35]. For distributed Kalman filtering fusion of measurements with cross-correlated noises, the simple form in considering of both optimality in accuracy as well as in computation and communication complexity, is still an open problem, and is the motivation of this paper.

This paper is organized as follows. Section 2 describes the problem formulation. In Section 3, the Cholesky decomposition and the linear transformation are introduced. Section 4 provides the optimal state fusion estimation algorithms, including the centralized fusion based on raw data, the centralized and the distributed fusion estimation algorithms based on the transformed measurement models. Performances, including the global optimization and the computation complexity and the communication requirements of each algorithms, are also analyzed in Section 4. Section 5 is the simulation and Section 6 draws the conclusion.

2 Problem formulation

Consider the following generic linear dynamic system:

$$x(k+1) = A(k)x(k) + w(k), \quad k = 0, 1, \ldots; \tag{1}$$

$$z_i(k) = C_i(k)x(k) + v_i(k), \quad i = 1, 2, \ldots, N, \tag{2}$$

where $x(k) \in \mathbb{R}^n$ is the system state, $A(k) \in \mathbb{R}^{n\times n}$ is the state transition matrix. $w(k)$ is the system noise and is assumed to be Gaussian distributed with

$$\begin{cases} \mathrm{E}\{w(k)\} = 0; \\ \mathrm{E}\{w(k)w^{\mathrm{T}}(j)\} = Q(k)\delta_{ij}, \end{cases} \tag{3}$$

where $Q(k) \geqslant 0$, and

$$\delta_{kj} = \begin{cases} 1, \ k = j; \\ 0, \ k \neq j. \end{cases} \tag{4}$$

$z_i(k) \in \mathbb{R}^{m_i}$ is the measurement of sensor i at time k. $C_i(k) \in \mathbb{R}^{m_i\times n}$ is the measurement matrix. $v_i(k)$ is the measurement noise, and is assumed to be white Gaussian distributed, and is independent of $w(k)$, i.e., for $k, l = 1, 2, \ldots, i, j = 1, 2, \ldots, N$, we have

$$\begin{cases} \mathrm{E}\{v_i(k)\} = 0; \\ \mathrm{E}\{v_i(k)v_j^{\mathrm{T}}(l)\} = R_{ij}(k)\delta_{kl}. \end{cases} \tag{5}$$

From the above formula we can see that the measurement noises of different sensors are correlated, i.e., $v_i(k)$ and $v_j(k)$ are cross-correlated for $i \neq j$ at time k with $\mathrm{E}\{v_i(k)v_j^{\mathrm{T}}(k)\} = R_{ij}(k) \neq 0$. For simplicity, we denote $R_i(k) \underset{=}{\triangle} R_{ii}(k) > 0$ for $i = 1, 2, \ldots, N$.

The initial state $x(0)$ is independent of $w(k)$ and $v_i(k)$ for $k = 1, 2, \ldots$, and $i = 1, 2, \ldots, N$, and is assumed to be Gaussian distributed with

$$\begin{cases} \mathrm{E}\{x(0)\} = x_0; \\ \mathrm{cov}\{x(0)\} = \mathrm{E}\{[x(0) - x_0][x(0) - x_0]^{\mathrm{T}}\} = P_0, \end{cases} \tag{6}$$

where $\mathrm{cov}\{x(0)\}$ means the covariance of $x(0)$.

The objective of this paper is to generate the optimal estimation of state $x(k)$ by use of the measurements $z_i(k)$ based on the above description.

3 Linear transformation

For the systems given in (1) and (2), let

$$z(k) = [z_1^{\mathrm{T}}(k), z_2^{\mathrm{T}}(k), \ldots, z_N^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{7}$$

$$C(k) = [C_1^{\mathrm{T}}(k), C_2^{\mathrm{T}}(k), \ldots, C_N^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{8}$$
$$v(k) = [v_1^{\mathrm{T}}(k), v_2^{\mathrm{T}}(k), \ldots, v_N^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{9}$$

then we have

$$\begin{cases} \mathrm{E}\{v(k)\} = 0_{m\times 1},\ m = \sum\limits_{i=1}^{N} m_i; \\ \mathrm{E}\{v(k)v^{\mathrm{T}}(j)\} = R(k)\delta_{kj}; \\ \mathrm{E}\{w(k)v^{\mathrm{T}}(j)\} = 0, \end{cases} \tag{10}$$

where

$$R(k) = \mathrm{cov}\{v(k)\} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1N} \\ R_{21} & R_{22} & \cdots & R_{2N} \\ \vdots & \vdots & & \vdots \\ R_{N1} & R_{N2} & \cdots & R_{NN} \end{bmatrix} \tag{11}$$

is a symmetric positive definite matrix.

To decouple the measurement noises and to decrease the computation complexity, we now use Cholesky decomposition to decompose $R(k)$. Namely,

$$R(k) = L^{\mathrm{T}}(k)L(k), \tag{12}$$

where $L(k)$ is a lower triangular matrix with strictly positive diagonal entries.

Let

$$T(k) = L^{-\mathrm{T}}(k), \tag{13}$$
$$\bar{z}(k) = T(k)z(k), \tag{14}$$
$$\bar{C}(k) = T(k)C(k), \tag{15}$$
$$\bar{v}(k) = T(k)v(k). \tag{16}$$

It follows that

$$\bar{z}(k) = \bar{C}(k)x(k) + \bar{v}(k), \tag{17}$$

where

$$\begin{cases} \mathrm{E}\{\bar{v}(k)\} = 0_{m\times 1},\ m = \sum\limits_{i=1}^{N} m_i; \\ \bar{R}(k) = \mathrm{cov}\{\bar{v}(k)\} = T(k)R(k)T^{\mathrm{T}}(k) = I_{m\times m}; \\ \mathrm{E}\{w(k)\bar{v}^{\mathrm{T}}(j)\} = 0. \end{cases} \tag{18}$$

Hence, the cross-correlated sensor noises is transformed into uncorrelated sensor noises, whose covariance is identity matrix.

4 The optimal state fusion estimation algorithms

4.1 The centralized state fusion estimation with raw data

Based on systems (1) and (2), the optimal state estimation can be generated by use of the centralized fusion algorithm.

Theorem 1. Given the centralized fusion estimation $\hat{x}_c(k-1|k-1)$ and $P_c(k-1|k-1)$ at $k-1$, the optimal state fusion estimation at time k can be computed as follows:

$$\begin{cases} \hat{x}_c(k|k-1) = A(k-1)\hat{x}_c(k-1|k-1); \\ P_c(k|k-1) = A(k-1)P_c(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1); \\ \hat{x}_c(k|k) = \hat{x}_c(k|k-1) + K_c(k)[z(k) - C(k)\hat{x}_c(k|k-1)]; \\ K_c(k) = P_c(k|k-1)C^{\mathrm{T}}(k)[C(k)P_c(k|k-1)C^{\mathrm{T}}(k) + R(k)]^{-1}; \\ P_c(k|k) = [I - K_c(k)C(k)]P_c(k|k-1), \end{cases} \tag{19}$$

where subscript "c" denotes the centralized algorithm, and $z(k)$, $C(k)$ and $R(k)$ are computed by (7), (8) and (11), respectively.

Proof. From the problem formulation and formula (10), it can be easily shown that $v(k)$ is white Gaussian noise and is uncorrelated with system noise $w(k)$. Therefore, to generate the optimal state estimation, the traditional centralized fusion can be used.

4.2 The centralized fusion with transformed data

From Section 3, systems (1) and (2) could be rewritten as

$$\begin{cases} x(k+1) = A(k)x(k) + w(k); \\ \bar{z}(k) = \bar{C}(k)x(k) + \bar{v}(k), \end{cases} \tag{20}$$

where $x(k) \in \mathbb{R}^n$ is the system state, $A(k) \in \mathbb{R}^{n\times n}$ is the state transition matrix, $w(k)$ is zero-mean white Gaussian process noises with covariance being $Q(k) \geqslant 0$, and the initial state meets $\mathrm{E}\{x(0)\} = x_0$ and $\mathrm{cov}\{x(0)\} = P_0$. $\bar{z}(k) \in \mathbb{R}^m$ is the measurement at time k, and $\bar{C}(k) \in \mathbb{R}^{m\times n}$ is the measurement matrix, where $m = \sum_{i=1}^{N} m_i$. Measurement noise $\bar{v}(k)$ is zero-mean white Gaussian distributed and is uncorrelated with $w(k)$ and $x(0)$, and $\mathrm{cov}\{\bar{v}(k)\} = \bar{R}(k) = I$.

The optimal state estimation of $x(k)$ could be generated by use of the following theorem.

Theorem 2. Based on system (20), given the centralized fusion estimation $\hat{x}_{ct}(k-1|k-1)$ and the estimation error covariance $P_{ct}(k-1|k-1)$ at time $k-1$, the optimal estimation at time k can be computed by

$$\begin{cases} \hat{x}_{ct}(k|k-1) = A(k-1)\hat{x}_{ct}(k-1|k-1); \\ P_{ct}(k|k-1) = A(k-1)P_{ct}(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1); \\ \hat{x}_{ct}(k|k) = \hat{x}_{ct}(k|k-1) + K_{ct}(k)[\bar{z}(k) - \bar{C}(k)\hat{x}_{ct}(k|k-1)]; \\ K_{ct}(k) = P_{ct}(k|k-1)\bar{C}^{\mathrm{T}}(k)[\bar{C}(k)P_{ct}(k|k-1)\bar{C}^{\mathrm{T}}(k) + \bar{R}(k)]^{-1}; \\ P_{ct}(k|k) = [I - K_{ct}(k)\bar{C}(k)]P_{ct}(k|k-1), \end{cases} \tag{21}$$

where subscript "ct" denotes the centralized algorithm with the transformed data, and where $\bar{z}(k)$, $\bar{C}(k)$ and $\bar{R}(k)$ are computed by (14), (15) and (18), respectively.

Proof. For the linear system (20), it is obvious that the system noise $w(k)$ is zero-mean white Gaussian distributed. The measurement noise $\bar{v}(k)$ is also zero-mean white Gaussian distributed, and is uncorrelated with $w(k)$ and $x(0)$. So, the optimal state fusion estimation can be generated by use of the centralized fusion algorithm.

Theorem 3. The state estimation of the centralized fusion given in Theorem 2 with transformed data is equivalent to the centralized fusion given in Theorem 1 with the original observations in the sense of LMMSE.

Proof. From (19), we can rewrite $P_c(k|k)$ and $P_c(k|k-1)$ by use of the information form of the Kalman filter as follows:

$$P_c(k|k-1) = A(k-1)P_c(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1), \tag{22}$$

$$P_c^{-1}(k|k) = P_c^{-1}(k|k-1) + C^{\mathrm{T}}(k)R^{-1}(k)C(k). \tag{23}$$

Similarly, from (21), we have

$$P_{ct}(k|k-1) = A(k-1)P_{ct}(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1), \tag{24}$$

$$P_{ct}^{-1}(k|k) = P_{ct}^{-1}(k|k-1) + \bar{C}^{\mathrm{T}}(k)\bar{R}^{-1}(k)\bar{C}(k). \tag{25}$$

Obviously, $P_c(0|0) = P_{ct}(0|0) = P_0$. Assume that $P_c(k-1|k-1) = P_{ct}(k-1|k-1)$, then due to (22) and (24), we obtain

$$P_{ct}(k|k-1) = P_c(k|k-1). \tag{26}$$

So due to (23) and (25), in order to prove $P_{ct}(k|k) = P_c(k|k)$, we just need to prove $\bar{C}^{\mathrm{T}}(k)\bar{R}^{-1}(k)\bar{C}(k) = C^{\mathrm{T}}(k)R^{-1}(k)C(k)$.

In fact,

$$\begin{aligned}\bar{C}^{\mathrm{T}}(k)\bar{R}(k)\bar{C}(k) &= [T(k)C(k)]^{\mathrm{T}}[T(k)R(k)T^{\mathrm{T}}(k)]^{-1}[T(k)C(k)]\\ &= C^{\mathrm{T}}(k)T^{\mathrm{T}}(k)T^{-T}(k)R^{-1}(k)T^{-1}(k)T(k)C(k)\\ &= C^{\mathrm{T}}(k)R^{-1}(k)C(k).\end{aligned} \tag{27}$$

Hence, we yield $P_{ct}(k|k) = P_c(k|k)$.

From Theorem 3, it can be seen that the state estimation based on system (20) and sytems (1) and (2) is equivalent as far as the centralized fusion is concerned.

4.3 The optimal state estimation by distributed fusion

Rewriting $\bar{z}(k), \bar{C}(k)$ and $\bar{v}(k)$ in (17) to block form, we have

$$\bar{z}(k) = [\bar{z}_1^{\mathrm{T}}(k), \bar{z}_2^{\mathrm{T}}(k), \ldots, \bar{z}_N^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{28}$$

$$\bar{C}(k) = [\bar{C}_1^{\mathrm{T}}(k), \bar{C}_2^{\mathrm{T}}(k), \ldots, \bar{C}_N^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{29}$$

$$\bar{v}(k) = [\bar{v}_1^{\mathrm{T}}(k), \bar{v}_2^{\mathrm{T}}(k), \ldots, \bar{v}_N^{\mathrm{T}}(k)]^{\mathrm{T}}, \tag{30}$$

where $\bar{z}_i(k) \in \mathbb{R}^{m_i \times 1}$, $\bar{C}_i(k) \in \mathbb{R}^{m_i \times n}$ and $\bar{v}_i(k) \in \mathbb{R}^{m_i \times 1}$. Then from (17), we have

$$\bar{z}_i(k) = \bar{C}_i(k)x(k) + \bar{v}_i(k), \quad i = 1, 2, \ldots, N. \tag{31}$$

From formula (18), it can be easily obtained that $\bar{v}_i(k)$ is uncorrelated across different sensors, which is zero mean Gaussian distributed and is independent of system noise $w(k)$ and the initial state $x(0)$. Therefore, based on (1) and (31), the optimal state estimation could be generated by use of the distributed fusion algorithm.

Theorem 4. By use of the linear systems (1) and (31), the optimal state estimation can be generated by use of the distributed fusion algorithm. For simplicity, we only give the one without feedback here,

$$P_d^{-1}(k|k) = P_d^{-1}(k|k-1) + \sum_{i=1}^{N}[P_{d,i}^{-1}(k|k) - P_{d,i}^{-1}(k|k-1)], \tag{32}$$

$$\begin{aligned}P_d^{-1}(k|k)\hat{x}_d(k|k) = P_d^{-1}(k|k-1)\hat{x}_d(k|k-1) + \sum_{i=1}^{N}[&P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k)\\ &-P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1)],\end{aligned} \tag{33}$$

where the fused state prediction at time k is

$$\hat{x}_d(k|k-1) = A(k-1)\hat{x}_d(k-1|k-1), \tag{34}$$

$$P_d(k|k-1) = A(k-1)P_d(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1), \tag{35}$$

and the local state estimations $\hat{x}_{d,i}(k|k)$ and $P_{d,i}(k|k)$ are computed by Kalman filter by use of

$$\begin{cases} \hat{x}_{d,i}(k|k) = \hat{x}_{d,i}(k|k-1) + K_{d,i}(k)[\bar{z}_i(k) - \bar{C}_i(k)\hat{x}_{d,i}(k|k-1)]; \\ P_{d,i}(k|k) = [I - K_{d,i}(k)\bar{C}_i(k)]P_{d,i}(k|k-1); \\ \hat{x}_{d,i}(k|k-1) = A(k-1)\hat{x}_{d,i}(k-1|k-1); \\ P_{d,i}(k|k-1) = A(k-1)P_{d,i}(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1); \\ K_{d,i}(k) = P_{d,i}(k|k-1)\bar{C}_i^{\mathrm{T}}(k)[\bar{C}_i(k)P_{d,i}(k|k-1)\bar{C}_i^{\mathrm{T}}(k) + \bar{R}_i(k)]^{-1}; \\ \bar{R}_i(k) = I. \end{cases} \tag{36}$$

Proof. Use the information form of Kalman filter, from (36), the update equations of sensor i at time k can be rewritten as [36]

$$\begin{cases} \hat{x}_{d,i}(k|k) = \hat{x}_{d,i}(k|k-1) + P_{d,i}(k|k)\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)[\bar{z}_i(k) - \bar{C}_i(k)\hat{x}_{d,i}(k|k-1)]; \\ P_{d,i}^{-1}(k|k) = P_{d,i}^{-1}(k|k-1) + \bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{C}_i(k). \end{cases} \tag{37}$$

Multiplying the first equation by the second equation in (37), we have

$$\begin{aligned} P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k) &= [P_{d,i}^{-1}(k|k-1) + \bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{C}_i(k)]\hat{x}_{d,i}(k|k-1) + P_{d,i}^{-1}(k|k)P_{d,i}(k|k) \\ &\quad \cdot\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)[\bar{z}_i(k) - \bar{C}_i(k)\hat{x}_{d,i}(k|k-1)] \\ &= P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1) + \bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{z}_i(k). \end{aligned} \tag{38}$$

Then

$$\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{z}_i(k) = P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k) - P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1). \tag{39}$$

On the other hand, from (21), we have

$$\begin{cases} \hat{x}_{ct}(k|k) = \hat{x}_{ct}(k|k-1) + P_{ct}(k|k)\bar{C}^{\mathrm{T}}(k)\bar{R}^{-1}(k)[\bar{z}(k) - \bar{C}(k)\hat{x}_{ct}(k|k-1)]; \\ P_{ct}^{-1}(k|k) = P_{ct}^{-1}(k|k-1) + \bar{C}^{\mathrm{T}}(k)\bar{R}^{-1}(k)\bar{C}(k). \end{cases} \tag{40}$$

Since $\bar{R}(k)$ is a diagonal matrix, we can rewrite it as

$$\bar{R}(k) = \mathrm{diag}\{\bar{R}_1(k), \bar{R}_2(k), \ldots, \bar{R}_N(k)\}, \tag{41}$$

where $\bar{R}_i(k) = I_{m_i \times m_i}$.

Substituting (28), (29) and (41) into (40), we have

$$\begin{cases} \hat{x}_{ct}(k|k) = \hat{x}_{ct}(k|k-1) + P_{ct}(k|k)\sum\limits_{i=1}^{N}\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)[\bar{z}_i(k) - \bar{C}_i(k)\hat{x}_{ct}(k|k-1)]; \\ P_{ct}^{-1}(k|k) = P_{ct}^{-1}(k|k-1) + \sum\limits_{i=1}^{N}\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{C}_i(k). \end{cases} \tag{42}$$

Multiplying the first equation by the second of (42), after reorganization, yields

$$P_{ct}^{-1}(k|k)\hat{x}_{ct}(k|k) = P_{ct}^{-1}(k|k-1)\hat{x}_{ct}(k|k-1) + \sum_{i=1}^{N}\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{z}_i(k). \tag{43}$$

Substituting (39) to (43), we have

$$P_{ct}^{-1}(k|k)\hat{x}_{ct}(k|k) = P_{ct}^{-1}(k|k-1)\hat{x}_{ct}(k|k-1) + \sum_{i=1}^{N}[P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k) - P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1)]. \tag{44}$$

From the second equation of (37), we have

$$\bar{C}_i^{\mathrm{T}}(k)\bar{R}_i^{-1}(k)\bar{C}_i(k) = P_{d,i}^{-1}(k|k) - P_{d,i}^{-1}(k|k-1). \tag{45}$$

Substituting (45) to the second equation of (42), we have

$$P_{ct}^{-1}(k|k) = P_{ct}^{-1}(k|k-1) + \sum_{i=1}^{N} [P_{d,i}^{-1}(k|k) - P_{d,i}^{-1}(k|k-1)]. \tag{46}$$

Combined (44) and (46), we have

$$\begin{cases} P_{ct}^{-1}(k|k)\hat{x}_{ct}(k|k) = P_{ct}^{-1}(k|k-1)\hat{x}_{ct}(k|k-1) + \sum\limits_{i=1}^{N} [P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k) \\ \qquad\qquad -P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1)]; \\ P_{ct}^{-1}(k|k) = P_{ct}^{-1}(k|k-1) + \sum\limits_{i=1}^{N} [P_{d,i}^{-1}(k|k) - P_{d,i}^{-1}(k|k-1)]. \end{cases} \tag{47}$$

Comparing (32)–(35) with (21) and (47), deductively, we have

$$\begin{cases} \hat{x}_d(k|k) = \hat{x}_{ct}(k|k); \\ P_d(k|k) = P_{ct}(k|k); \\ \hat{x}_d(k|k-1) = \hat{x}_{ct}(k|k-1); \\ P_d(k|k-1) = P_{ct}(k|k-1). \end{cases} \tag{48}$$

According to the derivation process of Theorem 4, the distributed fusion algorithm is generated from the centralized fusion. So, we have the following corollary directly.

Corollary 1. The state estimation by use of the distributed fusion algorithm with Theorem 4 is equivalent to the estimation by use of the centralized fusion algorithm given by Theorem 2.

Theorem 5. The state estimation by use of the distributed fusion algorithm of Theorem 4 with the transformed measurements of (31) is equivalent to the estimation by use of the centralized fusion algorithm with the original measurements of (2) given in Theorem 1.

Proof. From Theorem 3 and Corollary 1, we draw the conclusion.

From Theorem 5, we can see that by applying the linear transformation to the measurement equations and then by use of the distributed fusion algorithm given in Theorem 4, we can get the optimal state estimation, which is equivalent to the centralized fusion estimation based on the original observations in estimation precision. While, the distributed algorithm with transformed data is more flexible and applicable, and has much less computation complexity compared to the centralized fusion algorithm with raw data. We will show this in the next subsection.

Remark 1. In Theorem 4, to generate the state estimation, the distributed fusion without feedback is introduced. In fact, we should only change formula (36) to get the distributed fusion with feedback by use of the following formulae:

$$\begin{cases} \hat{x}_{d,i}(k|k) = \hat{x}_{d,i}(k|k-1) + K_{d,i}(k)[\bar{z}_i(k) - \bar{C}_i(k)\hat{x}_{d,i}(k|k-1)]; \\ P_{d,i}(k|k) = [I - K_{d,i}(k)\bar{C}_i(k)]P_{d,i}(k|k-1); \\ \hat{x}_{d,i}(k|k-1) = A(k-1)\hat{x}_d(k-1|k-1); \\ P_{d,i}(k|k-1) = A(k-1)P_d(k-1|k-1)A^{\mathrm{T}}(k-1) + Q(k-1); \\ K_{d,i}(k) = P_{d,i}(k|k-1)\bar{C}_i^{\mathrm{T}}(k)[\bar{C}_i(k)P_{d,i}(k|k-1)\bar{C}_i^{\mathrm{T}}(k) + \bar{R}_i(k)]^{-1}; \\ \bar{R}_i(k) = I. \end{cases} \tag{49}$$

For the distributed fusion algorithm with feedback, it can be shown that it is equivalent to the distributed fusion without feedback in estimation precision, and is therefore globally optimal in the sense of LMMSE, which has better robustness [3, 13].

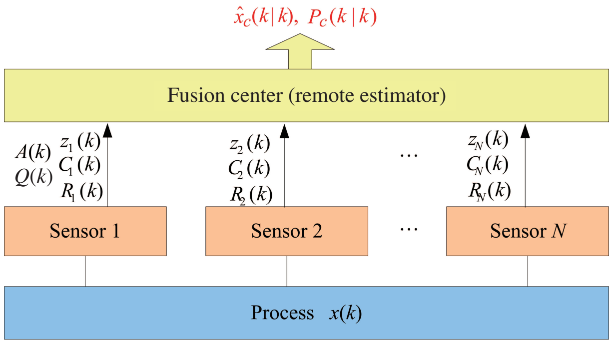

Figure 1 (Color online) Architecture of the centralized fusion by use of the raw measurements.

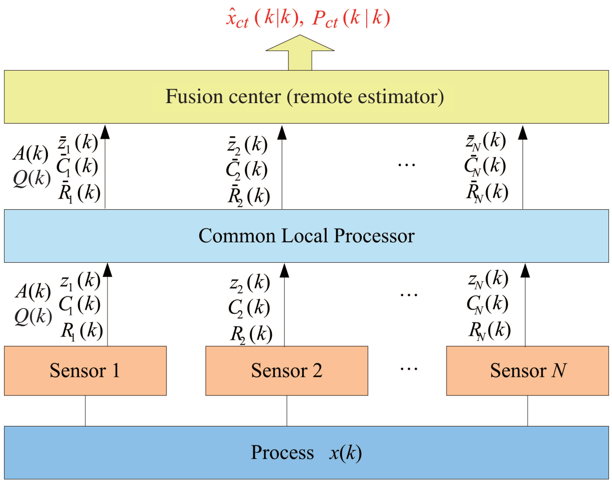

Figure 2 (Color online) Architecture of the centralized fusion by use of the transformed measurements.

4.4 The complexity analysis

The architectures of the centralized Kalman filtering fusion, the centralized fusion by use of the transformed measurements, and the optimal distributed fusion algorithms are shown in Figures 1–3, respectively. For the centralized fusion, the system parameters $A(k)$, $Q(k)$, $C_i(k)$, $R_{ij}(k)$ and the measurements $z_i(k)$ for sensors $i = 1, 2, \dots, N$ are sent to the remote fusion center directly. For the centralized fusion with transform, the system parameters and the measurements are sent to the local processor first, and after proper process (linear transform), the transformed data are sent to the remote fusion center. For the distributed fusion, the transformed data are sent to the local estimators, and the local estimations are sent to the remote fusion center. In the sequel, we will compare the computation and communication complexity of the three algorithms.

First, let us compare the centralized fusion algorithms by use of the original system and the system with transformed measurement equations.

It is known that to use the centralized fusion algorithm, besides the observations, it is necessary to send the state transition matrix $A(k)$, the measurement matrix $C(k)$, the covariance of the system noise $Q(k)$ and the covariance of measurement noise $R(k)$ to the fusion center. Nothing changes of $A(k)$ and $Q(k)$ after the transformation, neither does the dimension of $C(k)$. Let us check $R(k)$. Before transformation, the communication requirement to send the data of $R(k)$ to the fusion center is $m(m+1)/2$, where $m = \sum_{i=1}^{N} m_i$. However, as has been proven in (18), after the linear transformation, the measurement noises covariance becomes an identity matrix. So there is no need to send $R(k)$. This nice property certainly help to reduce the transmission burden to the fusion center. Particularly, when m is very large, the property will greatly reduce the communication demands and the computation complexity.

Next, let us compare the distributed fusion with the centralized fusion algorithm based on the system

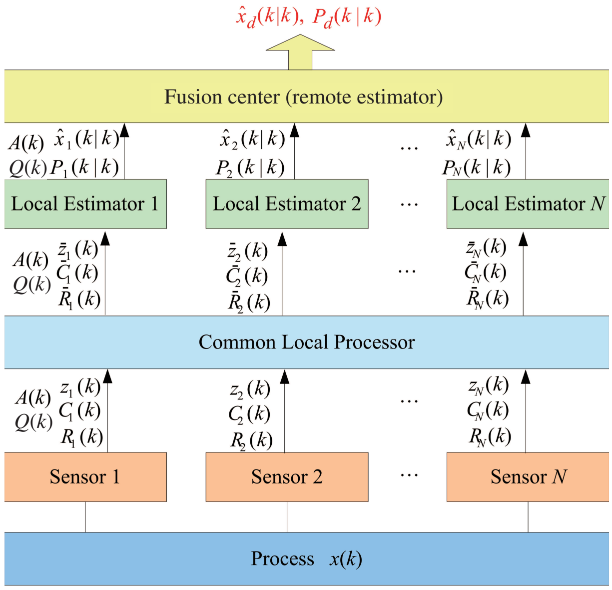

Figure 3 (Color online) Architecture of the distributed fusion.

with transformed form. The centralized fusion algorithm should transmit $A(k)$, $\bar{C}(k)$, $Q(k)$, $\bar{z}(k)$ to the fusion center at each moment k, and the communication requirement is $\frac{3}{2}n^2 + \frac{1}{2}n + mn + m$, where $m = \sum_{i=1}^{N} m_i$. For the distributed fusion however, the system model parameters and the local state estimations should be transmitted to the fusion center, and the communication requirement is $N(n^2+3n)$. It can be seen that when the state and the measurements dimensions are high, the centralized fusion will occupy larger transmission bandwidth. As far as the computation complexity is concerned, it is well known that the distributed fusion is more efficient compared to the centralized fusion algorithm, because augmentation of measurements, measurement matrix, and measurement noises covariances are avoided.

Finally, let us compare the presented distributed algorithm with the distributed algorithm presented in [3]. To make it clear, we list the distributed fusion formula (21) of paper [3] here by use of the same notations as in this paper:

$$\begin{aligned} P_d^{-1}(k|k)\hat{x}_d(k|k) = {} & P_d^{-1}(k|k-1)\hat{x}_d(k|k-1) + C^{\mathrm{T}}(k)\sum_{i=1}^{N} R^{-1}(k,i*)R_i^{-1}(k)C_i^{+}(k) \\ & \cdot[P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k) - P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1)], \end{aligned} \tag{50}$$

where subscript "+" denotes the pseudo-inverse, and $R^{-1}(k,i*)$ denotes the i-th submatrix column of $R^{-1}(k)$.

While, in the current presented distributed fusion algorithm, from (33), we have

$$\begin{aligned} P_d^{-1}(k|k)\hat{x}_d(k|k) = {} & P_d^{-1}(k|k-1)\hat{x}_d(k|k-1) + \sum_{i=1}^{N}[P_{d,i}^{-1}(k|k)\hat{x}_{d,i}(k|k) \\ & - P_{d,i}^{-1}(k|k-1)\hat{x}_{d,i}(k|k-1)]. \end{aligned} \tag{51}$$

The two algorithms are both proven to be equivalent to the centralized fusion in estimation precision, and therefore are both optimal in the sense of LMMSE. While, from (50), it can be seen that to generate the state fusion estimation, besides the necessary quantities of the presented distributed algorithm, $C_i(k)$ and $R(k)$ are also required by the fusion center, which means it requires $\frac{1}{2}m(m+1) + mn$ more data transmission at each moment compared to the presented distributed fusion algorithm. Moreover, comparing (51) with (50), it can be seen that the computation complexity of the current presented algorithm is much less than [3], since there is an augmentation of $C_i(k)$, i.e., $C(k)$, and the inverse of $R(k)$ needed to compute and to multiply in (50).

To summarize, the centralized fusion with the transformed data reduced the communication burden compared to the centralized fusion with raw data, and the distributed algorithm has even less computation complexity compared to the centralized fusion algorithm when the transformed system is concerned. While, as has been proven, the three algorithms are all globally optimal in the sense of LMMSE. Compared to the distributed fusion algorithm presented in [3], the presented distributed algorithm has the same estimation precision but less computation complexity and requires less communication bandwidth.

5 Numerical example

To illustrate the effectiveness of the proposed algorithms, a numerical example is provided in this section.

A target is observed by three sensors, which could be described by (1) and (2) with

$$A(k) = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \cdot 0.95; \tag{52}$$

$$Q(k) = \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix} \cdot q, \tag{53}$$

where $T = 1$ s is the sampling rate, and $q = 0.01$ is the disturbance parameter. Sensors 1 and 2 observe the first dimension of position, and sensor 3 observes the second dimension of velocity, i.e., the measurement matrices are

$$C_1(k) = [1 \quad 0], \tag{54}$$

$$C_2(k) = [1 \quad 0], \tag{55}$$

$$C_3(k) = [0 \quad 1]. \tag{56}$$

The measurement noises covariances is given by

$$R(k) = \begin{bmatrix} 0.25 & 0.125 & 0.00125 \\ 0.125 & 0.25 & 0.00125 \\ 0.00125 & 0.00125 & 0.0025 \end{bmatrix}. \tag{57}$$

The initial conditions are

$$x_0 = \begin{bmatrix} 10 \\ 0.5 \end{bmatrix}, \quad P_0 = \begin{bmatrix} 1 & \frac{1}{T} \\ \frac{1}{T} & \frac{2}{T^2} \end{bmatrix}. \tag{58}$$

The Monte Carlo simulation results are shown in Figures 4–6, where “CF” denotes the centralized fusion algorithm by using the original observations as shown in Theorem 1, “CTF” denotes centralized fusion by use of the transformed data as shown in Theorem 2, “DF” denotes the distributed fusion presented by [3] and “DTF” denotes the presented distributed fusion by use of the transformed data given in Theorem 4.

In Figure 4, the measurements of sensors 1 through 3 are shown in red dotted line, where sensors 1 and 2 observe the position dimension, and sensor 3 observes the velocity. For comparison, we use blue line to show the original signal of the corresponding dimensions. It can be seen that the measurements are corrupted by noises.

Figure 5 is the state estimation errors of different algorithms. The red dotted line in Figure 5(a) and (c) is the estimation errors of CTF, DT and DTF, respectively, and the estimation error of CT is shown in each subfigures in blue real line for comparison. It can be seen that the estimation errors of CTF, DT, and DTF are exactly the same as that of CT, which means the four algorithms are equivalent in state fusion estimation.

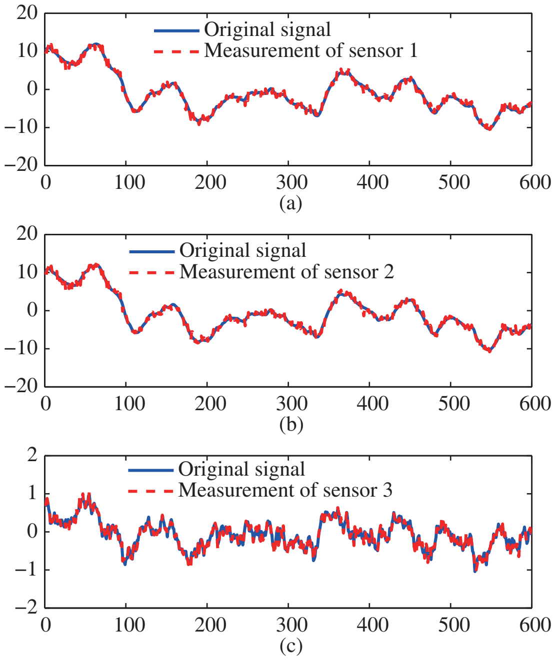

Figure 4 (Color online) First dimension of the original signal and the measurements.

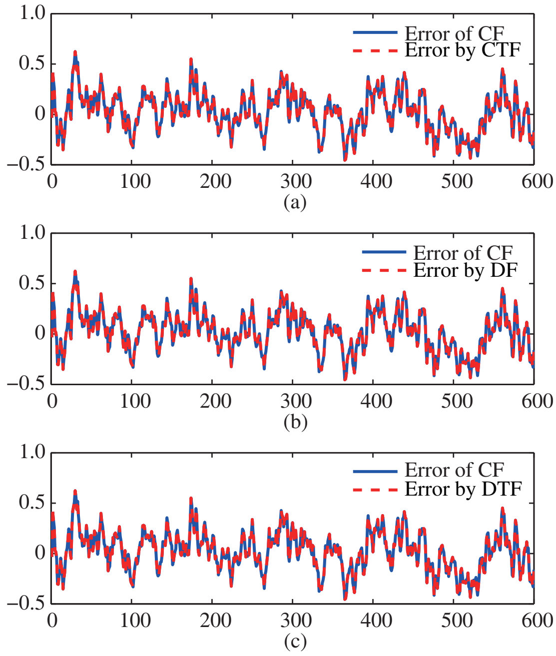

Figure 5 (Color online) State estimation errors.

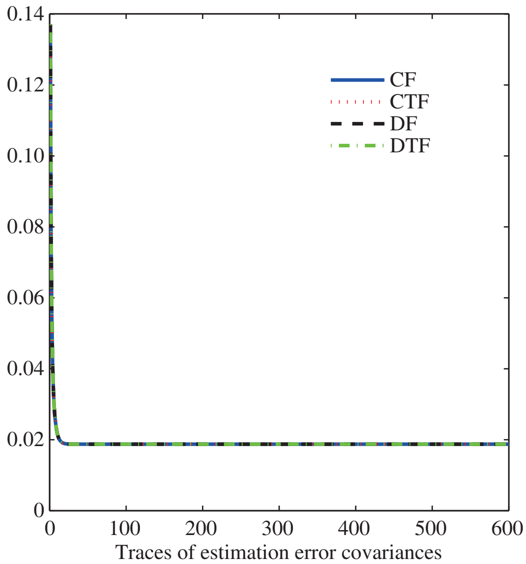

Figure 6 (Color online) Trace of the covariances of the state estimation errors.

Figure 6 shows the trace of the covariances of the state estimation errors, it can be seen that they are coincided, which further verify the conclusion drawn from Figure 5.

The simulation in this section shows that in state estimation precision, the presented distributed fusion algorithm based on linear transform is equivalent to the centralized fusion with and without linear transformation, and is also equivalent to the distributed fusion algorithm presented by [3]. In fact, the MMSEs of each algorithms are 0.0192 in this simulation. As far as the running time is concerned, the average time for 100 runs for each algorithms, namely, CF, CTF, DF and DTF are $0.0468, 0.0156, 0.0780$ and 0.0624, respectively. It can be seen that the centralized fusion after transformation is faster than the centralized fusion with raw data. The distributed fusion by use of the transformed data is more efficient than the distributed fusion presented in [3]. However, the running time of the distributed fusion algorithms are longer than the centralized fusion algorithms. By our analysis, it may rely on the following reasons: (1) The dimensions of states and measurements in this simulation are not large, the advantages of the distributed fusion algorithms are not obvious. (2) The computing of the inverse of the state estimation error covariances of the local estimations and the state prediction error covariances are really time cost. By comparing the distributed or the centralized fusion with and without linear transform, respectively, it can be seen that linear transformation is really an efficient way to reduce computation complexity, let alone the communication or network transmission of data is concerned, which can show even more advantages of linear transformation.

Briefly, the simulation results in this section illustrate the effectiveness of the presented algorithms. When the dimensions of the measurements and the states are not high, the centralized fusion is fine. While, when the dimensions of the measurements are high, the distributed fusion with transformed data has potential advantages.

6 Conclusion

When the sensor noises are cross-correlated, a linear transformation is presented to decouple the noises into cross-independent. Then, the centralized fusion and the distributed fusion with transformed data are presented and proven to have the same optimality as the centralized fusion with the original observations, while have lower computation complexity in computing and are more efficient in communication or data transmission, which is very useful in networked environments. A numerical example is shown to verify the effectiveness of the proposed algorithms.

Acknowledgements The work was supported by National Natural Science Foundation of China (Grant Nos. 61225015, 61473040), Beijing Natual Science Foundation (Grant No. 4161001), and Innovative Research Groups of the National Nature Science Foundation of China (Grant No. 61321002).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Li X R, Zhu Y M, Wang J, et al. Optimal linear estimation fusion-part 1: unified fusion rules. IEEE Trans Inform Theory, 2003, 49: 2192–2208
2 BarShalom Y, Li X R, Kirubarajam T. Estimation With Application to Tracking and Navigation. New York: John Wiley and Sons, 2001
3 Song E B, Zhu Y M, Zhou J, et al. Optimal Kalman filtering fusion with cross-correlated sensor noises. Automatica, 2007, 43: 1450–1456
4 Ge Q B, Xu D X, Wen C L. Cubature information filters with correlated noises and their applications in decentralized fusion. Signal Process, 2014, 94: 434–444
5 Jayaweera S, Mosquera C. Distributed sequential estimation with noisy, correlated observations. IEEE Signal Process Lett, 2008, 15: 741–744
6 Feng J X, Zeng M. Optimal distributed Kalman filtering fusion for a linear dynamic system with cross-correlated noises. Int J Syst Sci, 2012, 43: 385–398
7 Li Y, Wen C. An optimal sequential decentralized filter of discrete-time systems with crosscorrelated noises. In: Proceedings of the 17th World Congress of the International Federation of Automatic Control, Seoul, 2008. 7560–7565
8 Feng X. Fusion estimation for sensor network systems with correlated measurements and oosms. Dissertation for Ph.D. Degree. Hangzhou: Hangzhou Dianzi University, 2009
9 Sun S, Deng Z. Multi-sensor optimal information fusion Kalman filter. Automatica, 2004, 40: 1017–1023
10 Yan L P, Li X R, Xia Y Q, et al. Optimal sequential and distributed fusion for state estimation in cross-correlated noise. Automatica, 2013, 49: 3607–3612
11 Liu Y L, Yan L P, Xia Y Q, et al. Multirate multisensor distributed data fusion algorithm for state estimation with cross-correlated noises. In: Proceedings of the 32th Chinese Control Conference, Xi'an, 2013. 4682–4687
12 Yan L P, Liu J, Jiang L, et al. Optimal sequential estimation for multirate dynamic systems with unreliable measurements and correlated noise. In: Proceedings of the 35th Chinese Control Conference, Chengdu, 2016. 4900–4905
13 Feng J X, Wang Z D, Zeng M. Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises. Inf Fusion, 2013, 14: 78–86
14 Ran C J, Deng Z L. Correlated measurement fusion Kalman filters based on orthogonal transformation. In: Proceedings of the 21st Annual International Conference on Chinese Control and Decision Conference (CCDC 2009), Guilin, 2009.

1193–1198
15 Ran C J, Deng Z L. Two correlated measurement fusion kalman filtering algorithms based on orthogonal transformation and their functional equivalence. In: Proceedings of Joint 48th IEEE Conference on Decision and Control and the 28th Chinese Control Conference, Shanghai, 2009. 2351–2356
16 Duan Z, Han C, Tao T. Optimal multi-sensor fusion target tracking with correlated measurement noises. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Hague, 2004. 2: 1272–1278
17 Duan Z, Li X R. The optimality of a class of distributed estimation fusion algorithm. In: Proceedings of the 11th International Conference on Information Fusion, Cologne, 2008. 16: 1–6
18 Duan Z, Li X R. Lossless linear transformation of sensor data for distributed estimation fusion. IEEE Trans Signal Process, 2011, 59: 362–372
19 Liu X, Li Z, Liu X, et al. The sufficient condition for lossless linear transformation for distributed estimation with cross-correlated measurement noises. Int J Process Control, 2013, 23: 1344–1349
20 Zhang Y G, Huang Y L. Gaussian approximate filter for stochastic dynamic systems with randomly delayed measurements and colored measurement noises. Sci China Inf Sci, 2016, 59: 092207
21 Wang Y Q, Zhao D, Li Y Y, et al. Unbiased minimum variance fault and state estimation for linear discrete time–varying two–dimensional systems. IEEE Trans Autom Control, 2017, 62: 5463–5469
22 Ge Q, Shao T, Duan Z, et al. Performance analysis of the Kalman filter with mismatched measurement noise covariance. IEEE Trans Autom Control, 2016, 61: 4014–4019
23 Ge Q B, Ma J Y, Chen S D, et al. Observation degree analysis on mobile target tracking for wireless sensor networks. Asian J Control, 2017, 19: 1259–1270
24 Zhang K S, Li X R, Zhang P, et al. Optimal linear estimation fusion-part vi: sensor data compression. In: Proceedings of the 2nd International Conference of Information Fusion, Cairns, 2003. 221–228
25 Song E B, Zhu Y M, Zhou J. Sensors' optimal dimensionality compression matrix in estimation fusion. IEEE Trans Signal Process, 2005, 41: 2131–2139
26 Zhu Y M, Song E B, Zhou J, et al. Information fusion strategies and performance bounds in packet-drop networks. IEEE Trans Signal Process, 2005, 53: 1631–1639
27 Schizas I D, Giannakis G B, Luo Z Q. Distributed estimation using reduced-dimensionality sensor observations. IEEE Trans Signal Process, 2007, 55: 4284–4299
28 Li X R, Zhang K S. Optimal linear estimation fusion–part iv: optimality and efficiency of distributed fusion. In: Proceedings of the 4th International Conference Information Fusion, Montreal, 2001. 19–26
29 Li J, AlRegib G. Distributed estimation in energy-constrained wireless sensor networks. IEEE Trans Signal Process, 2009, 57: 3746–3758
30 Msechu J J, Roumeliotis S I, Ribeiro A, et al. Decentralized quantized kalman filtering with scalable communication cost. IEEE Trans Signal Process, 2008, 56: 3727–3741
31 Xu J, Li J X, Xu S. Data fusion for target tracking in wireless sensor networks using quantized innovations and Kalman filtering. Sci China Inf Sci, 2012, 55: 530–544
32 Zhang Z, Li J, Liu L. Distributed state estimation and data fusion in wireless sensor networks using multi-level quantized innovation. Sci China Inf Sci, 2016, 59, 022316
33 Liang Y, Chen T, Pan Q. Multi-rate optimal state estimation. Int J Control, 2009, 82: 2059–2076
34 Zhang W, Liu S, Yu L. Fusion estimation for sensor networks with nonuniform estimation rates. IEEE Trans Circ Syst I: Regular Papers, 2014, 61: 1485–1498
35 Yan L P, Jiang L, Xia Y Q, et al. State estimation and data fusion for multirate sensor networks. Int J Adapt Control Signal Process, 2016, 30: 3–15
36 Simon D. Optimal State Estimation. New York: John Wiley and Sons, Inc. Publication, 2006