Information Sciences 370–371 (2016) 446–462

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Distributed fusion estimation with square-root array implementation for Markovian jump linear systems with random parameter matrices and cross-correlated noises

Yanbo Yang[a,b,*], Yan Liang[a,b], Quan Pan[a,b], Yuemei Qin[a,b], Feng Yang[a,b]

[a] School of Automation, Northwestern Polytechnical University, Xi'an, 710072, P.R. China
[b] Key Laboratory of Information Fusion Technology, Ministry of Education, Xi'an, 710072, P.R. China

ARTICLE INFO

Article history:
Received 9 February 2016
Revised 25 July 2016
Accepted 8 August 2016
Available online 8 August 2016

Keywords:
Markovian jump systems
Distributed fusion estimation
Linear minimum mean square error estimator
Square-root array implementation
Random parameter matrices
Cross-correlated noises

ABSTRACT

This study presents the distributed fusion estimation of discrete-time Markovian jump linear systems with random parameter matrices and cross-correlated noises in sensor networks. The recursive linear minimum mean square error estimator is proposed based on the Gram-Schmidt orthogonalization procedure under a centralized framework. In order to avoid the loss of positive semidefiniteness and reduce dynamical range, its square-root array implementation is presented by recursively triangularizing the square roots of relevant positive semidefinite matrices. Furthermore, via the information filter form, the distributed fusion estimation with square-root array implementation is derived from the centralized fusion structure, incorporated with consensus strategy. A maneuvering target tracking simulation in a sensor network validates the proposed method.

1. Introduction

State estimation for Markovian jump systems (MJSs), as a class of hybrid systems with the continuous-valued state and discrete-valued mode obeying a finite state Markov chain, has been receiving a great attention due to its wide and successful applications in hybrid system control [20,28], fault detection and isolation [35], statistical signal processing [29] and maneuvering target tracking [9,18]. Different categories of estimators or filters for such systems, including the model estimators [19], H_2/H_∞ filters [23,34] or linear minimum mean square error (LMMSE) estimators [4], have been proposed under different assumptions and performance criterions.

For the LMMSE estimator, it not only provides the recursive estimation of first two moments of the interested vector with an acceptable computation cost, but also has no limitation of Gaussian assumptions on the process and measurement noises. In [4], the LMMSE estimator was derived from geometric augment through estimating $x_k 1_{\{\Theta_k=i\}}$ instead of estimating directly the original state x_k for the Markovian jump linear system (MJLS), where $1_{\{\cdot\}}$ was the Dirac delta function and Θ_k was a discrete-time Markov chain. Then, the convergence of state prediction error covariance in the LMMSE estimator [4] was discussed in [5] for deriving the corresponding time-invariant implementation, where such covariance converged to the unique positive semidefinite solution of an algebraic Riccati equation. Furthermore, the square-root array implementation [31] was given for improving the numerical stability of the estimator. Besides, the LMMSE estimator was

* Corresponding author.
E-mail address: yangyb8911@gmail.com, yangyanbo9@mail.nwpu.edu.cn (Y. Yang).

http://dx.doi.org/10.1016/j.ins.2016.08.020
0020-0255/

proposed for the MJLSs with multiple possible randomly delayed measurements [38]. The corresponding information filter and square-root array implementations for the MJLSs [30] were further developed, which were suitable to the absence of a priori information about initial state.

All in all, above discussed LMMSE estimators and square-root array implementations all have the following two ideal assumptions: system matrices should be known; the measurement noise should be uncorrelated with process noise at each epoch. On the one hand, such model cannot deal with the case that system matrices in MJSs contain random parameters. On the other hand, the un-correlation assumption is unsuitable to practical applications sometimes [33].

Actually, the random parameter existing in the system matrix, which can be regarded as the multiplicative noise, is a hot research issue in estimation and control [8,15,17,26], due to its wide applications, such as target tracking and fusion [22,37]. A Kalman-type filter [27] for the MJLSs with state-dependent noises which is independent of additive noises was proposed as a mean-square filter in a stationary form, based on the solution of some coupled algebraic Riccati equations. In addition, the LMMSE estimator for the MJLSs with multiplicative noises was presented [6], and the convergence of the error covariance matrix of the LMMSE estimator was discussed in the stationary case. Moreover, a robust mode-independent linear filter was presented [7] for the MJLSs with multiplicative noises through the linear matrix inequalities formulation, based on a parameter dependent Lyapunov procedure. Furthermore, the LMMSE estimator for the MJLSs with stochastic coefficient matrices (abbreviated as LMSCE estimator) was derived [37] by the idea of establishing the general filter framework of the joint state estimation and data association in clutters for maneuvering target tracking, and sufficient conditions were given for the stability of the LMSCE estimator.

However, all these methods only focus on the stochastic linear systems with Markov jumps and multiplicative noises, and are also limited to the single sensor or centralized multi-sensor condition. In fact, random parameter matrices (i.e., multiplicative noises) and cross-correlated additive noises may coexist in a complex MJS in the distributed multi-sensor case. Consider the non-cooperative maneuvering target tracking in the sensor network under complex game environment. By the fact that the measurement precision is always dependent on the relative position between the target and corresponding sensor, the multiplicative and additive noises will coexist in each sensor model [39]. At the same time, the additive process noise and measurement noise from each sensor may be also correlated if the external disturbance not only affects the dynamic evolvement but also acts on surroundings to impact upon the measurement noise [1].

To the best of authors' knowledge, the existing studies or researches are limited on the distributed state estimation and fusion for stochastic systems with a single model, along with network-induced constraints or uncertainties, such as limited communication bandwidth [2,3], asynchronous fusion [40] and random delay [10,36]. Although some of these literatures may refer to the estimation and fusion with Markov switching parameters, they mainly focus on random delay obeying a Markov chain during data transmitting [10,36]. Except for several studies about the suboptimal filter and fusion under the multiple model framework with Gaussian posterior density hypotheses [18], there is few research on the distributed estimation fusion for MJSs with the coexistence of multiplicative noises and cross-correlated additive noises. In other words, distributed state estimation for MJLSs with random parameter matrices and cross-correlated noises (MRMCNs) is useful and important, but still open.

This paper presents the distributed fusion estimation for the considered MRMCN with the following technical contributions. Firstly, the recursive LMMSE estimator for the MRMCN (abbreviated as LMRCE estimator) is derived based on the Gram-Schmidt orthogonal innovation sequence in a centralized structure with multiple sensors, which is more practical and general through simultaneously dealing with the multi-mode Markov jump, random parameter matrices and cross-correlation between the process noise and measurement noise. Secondly, the corresponding square-root array implementation of the LMRCE estimator (denoted as SALME) is presented, which guarantees that symmetric second-order moment matrices in the LMRCE estimator are nonnegative definite and reduces the dynamical range of related matrices. Thirdly, the distributed fusion estimation with square-root array implementation is derived from the centralized fusion under the LMMSE criterion, incorporating with consensus strategy to obtain a consistent estimate in each sensor node. The convergent condition and the convergence point are also discussed with respect to consensus strategy.

In this paper, superscripts "−1" and "T" are the inverse and transpose operations of the corresponding matrix, respectively; $(\cdot)$ denotes the same content as that in the previous parenthesis; I and O are the identity and zero matrices with proper dimensions, respectively; $\mathbf{1}_M \in \mathcal{R}^M$ is a column vector with each element being 1; ":=" denotes the definition and "$\otimes$" represents the Kronecker product; $E(X|Z_{1:k})$ is the linear projector of X onto the space spanned by $\{Z_1, \cdots, Z_k\}$; $A(p, q)$ is the element of matrix A at the pth row and qth column.

2. Problem formulation

Consider the discrete-time MRMCNs in sensor networks as follows:

$$x_{k+1} = F_{\Theta_k} x_k + G_{\Theta_k} w_k, \tag{1}$$

$$z_{n,k} = H_{n,\Theta_k} x_k + D_{n,\Theta_k} v_{n,k}, \tag{2}$$

with

$$F_{\Theta_k}=\sum_{i=1}^{M}F_{i,k}1_{\{\Theta_k=i\}},\quad G_{\Theta_k}=\sum_{i=1}^{M}G_{i,k}1_{\{\Theta_k=i\}},\quad H_{n,\Theta_k}=\sum_{i=1}^{M}H_{n,i,k}1_{\{\Theta_k=i\}},\quad D_{n,\Theta_k}=\sum_{i=1}^{M}D_{n,i,k}1_{\{\Theta_k=i\}},$$

where x_k and $z_{n,k}$ represent the system state and measurement from the *nth* sensor node (the subscript n represents the *nth* sensor node in the full text), respectively. Moreover, multiple sensors construct a sensor network, which is denoted by an undirected graph $\mathbb{G}=(\mathbb{V},\mathcal{E})$ with the vertex $n\in\mathbb{V}=\{1,\cdots,N\}$ representing the sensor node and the edge $e_{mn}\in\mathcal{E}=\{e_{mn}|m,n\in\mathbb{V}\}$ meaning the permission about information sharing between neighbor nodes m and n. Θ_k is a discrete-time first-order Markov chain with finite state space $\{1,\cdots,M\}$ and transition probability matrix $P_t=[p_{ij}]$ where $p_{ij}=P\{\Theta_{k+1}=i|\Theta_k=j\}$. $\pi_{i,k}:=P\{\Theta_k=i\}$ denotes the *ith* mode probability at the *kth* sampling instant. $1_{\{\Theta_k=i\}}$ is a binary variable and equals 1 if $\Theta_k=i$ and 0 otherwise. Here, $F_{i,k}$, $G_{i,k}$, $H_{n,i,k}$ and $D_{n,i,k}$ are random matrices with the following known statistics: $E(F_{i,k})=\bar{F}_{i,k}$, $E(G_{i,k})=\bar{G}_{i,k}$, $\mathrm{cov}(f_{i,k}^{pq},f_{j,k}^{rs})=c_{ij,k}^{f_{pq,rs}}$ and $\mathrm{cov}(g_{i,k}^{pq},g_{j,k}^{rs})=c_{ij,k}^{g_{pq,rs}}$; $E(H_{n,i,k})=\bar{H}_{n,i,k}$, $E(D_{n,i,k})=\bar{D}_{n,i,k}$, $\mathrm{cov}(h_{n,i,k}^{pq},h_{n,j,k}^{rs})=c_{n,ij,k}^{h_{pq,rs}}$ and $\mathrm{cov}(d_{n,i,k}^{pq},d_{n,j,k}^{rs})=c_{n,ij,k}^{d_{pq,rs}}$ for $n=1,\cdots,N$, where $f_{i,k}^{pq}:=F_{i,k}(p,q)$ (similarly to $f_{j,k}^{rs}$, $g_{i,k}^{pq}$, $g_{j,k}^{rs}$, $h_{n,i,k}^{pq}$, $h_{n,j,k}^{rs}$, $d_{n,i,k}^{pq}$ and $d_{n,j,k}^{rs}$). Furthermore, $\{w_k,k\geq1\}$ and $\{v_{n,k},k\geq1\}$ are zero-mean noises satisfying

$$E(w_kw_l^T)=Q_k\delta_{k-l},\quad E(w_kv_{n,l}^T)=S_{n,k}\delta_{k-l},\quad E(v_{n,k}v_{n,l}^T)=R_{n,k}\delta_{k-l},$$

where δ_{k-l} is the Dirac delta function that equals 1 if $k=l$ and 0 otherwise. Meanwhile, $E(v_{n,k}v_{m,l}^T)=O$ for different sensors n and m, and this is valid for those sensors that do not have severe interference [11]. The initial value x_0 is a random vector with $E(x_01_{\{\Theta_0=i\}})=\psi_i$ and $E(x_0x_0^T1_{\{\Theta_0=i\}})=V_i$ for $i=1,\cdots,M$. x_0, $\{\Theta_k,k\geq1\}$ and $\{F_{\Theta_k},G_{\Theta_k},H_{n,\Theta_k},D_{n,\Theta_k}\}$ are mutually uncorrelated and uncorrelated with $\{w_k,k\geq1\}$ and $\{v_{n,k},k\geq1\}$. Obviously, the multi-mode switching, parameter randomness, noise cross-correlation and decentralized fusion architecture coexist in the system (1)–(2).

The existing LMMSE estimators for MJLSs [6,37] are only applied to the case with Markov jumps and multiplicative noises. Meanwhile, the multiple model method with the linear optimal filter [22] or Gaussian filters [16,33] are just for MJSs with multiplicative noises or cross-correlated noises, individually. Actually, neither the coexistence of the Markov switching parameter, random parameter matrices and cross-correlated noises nor the corresponding square-root array implementation is considered to develop a more general and better numerical realization. Moreover, the aforementioned methods are all in the view of centralized fusion structure, and hardly extended to the distributed one.

There are two ways to design the distributed estimation fusion for MJSs, i.e., the multiple model method and LMMSE criterion. However, it is inconvenient and inaccurate to design a distributed estimation for the considered MRMCNs in sensor networks based on the multiple model method, by the following considerations. Firstly, it is inaccurate to use Gaussian densities to approximate the posterior densities with the state-dependent noises and cross-correlated noises. Secondly, if the final distributed fusion is a direct synthesis of local estimates from different sensors, the resultant result will be undesirable since the local estimate is suboptimal and the fusion rule is also always suboptimal due to neglecting correlations among different local estimates. Thirdly, if intermediate vectors or matrices with available measurement information are shared to pursue the final consistent estimate, all these time-varying multi-mode items should be shared across sensor nodes, implying intensive computation.

In the following, the distributed fusion estimation for the considered MRMCNs in sensor networks will be derived under the LMMSE criterion in the perspectives of recursive structure design, rapid implementation efficiency and better numerical conditioning, by the following motivations:

- it is designable to derive the centralized LMMSE estimator and further present the distributed fusion for the considered MRMCNs in sensor networks to pursue the scalability and more robustness to sensor failures;
- the corresponding square-root array implementation can give a more stable numerical realization, in the sense that it can provide a better calculation condition and reduced dynamical range.

3. Centralized fusion estimation with square-root array implementation

In this section, we will firstly derive the centralized LMMSE estimator for the considered MRMCN in the multi-sensor case, which can be also used in a single sensor case. Then, aimed at the corresponding symmetric and positive semidefinite second-order moment matrices, its square-root array implementation is presented to pursue a better conditioning and reduced dynamical range.

Denote $z_k:=\mathrm{col}\{z_{1,k},\cdots,z_{N,k}\}$, $H_{\Theta_k}:=[H_{1,\Theta_k}^T,\cdots,H_{N,\Theta_k}^T]^T$, $D_{\Theta_k}:=\mathrm{diag}\{D_{1,\Theta_k},\cdots,D_{N,\Theta_k}\}$, $v_k:=\mathrm{col}\{v_{1,k},\cdots,v_{N,k}\}$, $S_k:=[S_{1,k},\cdots,S_{N,k}]$ and $R_k:=\mathrm{diag}\{R_{1,k},\cdots,R_{N,k}\}$. Eq. (2) can be rewritten as

$$z_k=H_{\Theta_k}x_k+D_{\Theta_k}v_k. \tag{3}$$

Define $\xi_k:=\mathrm{col}\{x_k1_{\{\Theta_k=i\}},i=1,\cdots,M\}$ with $\xi_{i,k}:=x_k1_{\{\Theta_k=i\}}$ and hence $x_k=\sum_{i=1}^{M}x_k1_{\{\Theta_k=i\}}=\sum_{i=1}^{M}\xi_{i,k}$. For the convenience of deriving, $F_{i,k}$, $G_{i,k}$, $H_{n,i,k}$ and $D_{n,i,k}$ will be divided into two parts, where the one is about the mean values of

random variables and the other is about the remained errors, i.e., $F_{i,k}=\bar{F}_{i,k}+\tilde{F}_{i,k}$, $G_{i,k}=\bar{G}_{i,k}+\tilde{G}_{i,k}$, $H_{n,i,k}=\bar{H}_{n,i,k}+\tilde{H}_{n,i,k}$ and $D_{n,i,k}=\bar{D}_{n,i,k}+\tilde{D}_{n,i,k}$. Then, the original system is transformed into:

$$\xi_{k+1}=\bar{F}_k\xi_k+\tilde{F}_k\xi_k+\bar{G}_k w_k+\tilde{G}_k w_k, \tag{4}$$

$$z_k=\bar{H}_k\xi_k+\tilde{H}_k\xi_k+\bar{D}_k\nu_k+\tilde{D}_k\nu_k, \tag{5}$$

where $\bar{F}_k$ is an $M\times M$ block matrix with $\bar{F}_{j,k}1_{\{\Theta_{k+1}=i|\Theta_k=j\}}$ being its $(i,j)th$ sub-block; $\bar{G}_k$ is an $M\times 1$ block matrix with $\sum_{j=1}^{M}\bar{G}_{j,k}1_{\{\Theta_k=j\}}1_{\{\Theta_{k+1}=i|\Theta_k=j\}}$ being its ith sub-block; $\bar{H}_k:=[\bar{H}_{1,k}^T\ \cdots\ \bar{H}_{N,k}^T]^T$ and $\bar{D}_k:=\mathrm{diag}\{\bar{D}_{1,k},\cdots,\bar{D}_{N,k}\}$. Here, $\bar{H}_{n,k}$ is a $1\times M$ block matrix with its jth sub-block being $\bar{H}_{n,j,k}$ and $\bar{D}_{n,k}:=\sum_{i=1}^{M}\bar{D}_{n,i,k}1_{\{\Theta_k=i\}}$ for $n=1,\cdots,N$. $\tilde{F}_k$, $\tilde{G}_k$, $\tilde{H}_k$ and $\tilde{D}_k$ have the same forms as $\bar{F}_k$, $\bar{G}_k$, $\bar{H}_k$ and $\bar{D}_k$ with mean values being replaced by the corresponding errors, respectively. For example, $\tilde{F}_{j,k}$ in the matrix $\tilde{F}_k$ is replaced by $\tilde{F}_{j,k}$ to construct $\tilde{F}_k$. Meanwhile, $1_{\{\Theta_{k+1}=i|\Theta_k=j\}}$ will be 1 if $\Theta_{k+1}=i$ conditioned by $\Theta_k=j$ and 0 otherwise.

After random parameter matrices being divided into determined parts and uncertain parts, in the following derivation, the first-order moment estimation of interested vectors will be mainly dependent on those determined parts, while the related second-order moment estimation will pay much attention on those uncertain parts coupling with unknown state or noises.

3.1. Linear minimum mean square error estimator

In this subsection, the optimal LMMSE estimate $\hat{\xi}_{k|k}$ of the augmented state ξ_k conditioned by $Z_{1:k}:=\{z_1,\cdots,z_k\}$ will be derived recursively. Besides, the LMMSE estimate $\hat{x}_{k|k}$ of x_k can be calculated as: $\hat{x}_{k|l}=\mathcal{I}_t\hat{\xi}_{k|l}$, $P_{k|l}=\mathcal{I}_t\Phi_{k|l}\mathcal{I}_t^T$, where $\Phi_{k|l}:=E(\xi_k-\hat{\xi}_{k|l})(\cdot)^T$, $P_{k|l}:=E(x_k-\hat{x}_{k|l})(\cdot)^T$ and $\mathcal{I}_t:=\mathbf{1}_M^T\otimes I_{n_x}$.

Since $\hat{\xi}_{k|k}$ is the orthogonal projection of ξ_k onto the space $\mathcal{L}(z_1,\cdots,z_k)$ spanned by $\{z_1,\cdots,z_k\}$ and these observations are generally non-orthogonal, the Gram-Schmidt orthogonalized innovation approach will be used. The innovation at the $(k+1)th$ epoch is denoted as $\zeta_{k+1}:=z_{k+1}-\hat{z}_{k+1|k}$, where $\hat{z}_{k+1|k}$ is the one-step linear projector of z_{k+1} onto $\mathcal{L}(z_1,\cdots,z_k)$. $\mathcal{L}(\zeta_1,\cdots,\zeta_k)=\mathcal{L}(z_1,\cdots,z_k)$, i.e., the space spanned by $\{\zeta_1,\cdots,\zeta_k\}$ is the same as the linear space spanned by $\{z_1,\cdots,z_k\}$, and $E(\zeta_k\zeta_l^T)=O$ if $k\neq l$ [21]. In this way, the LMMSE estimate onto $\mathcal{L}(z_1,\cdots,z_k)$ is also the linear projector onto $\mathcal{L}(\zeta_1,\cdots,\zeta_k)$.

Lemma 3.1 [21]. *Let $\hat{\alpha}_{k|l}$ be the LMMSE estimate of α_k onto $\mathcal{L}(\zeta_1,\cdots,\zeta_l)$, then, $\hat{\alpha}_{k|l}=\sum_{t=1}^{l}E(\alpha_k\zeta_t^T)\left(E(\zeta_t\zeta_t^T)\right)^{-1}\zeta_t$.*

Let $\hat{\alpha}_{k|l}$ be replaced by $\hat{\xi}_{k+1|l}$ in Lemma 3.1, we have

$$\hat{\xi}_{k+1|l}=\sum_{t=1}^{l}E(\bar{F}_k\xi_k\zeta_t^T+\bar{G}_k w_k\zeta_t^T)(E(\zeta_t\zeta_t^T))^{-1}\zeta_t.$$

Due to the cross-correlation between w_k and ν_k, w_k is also correlated with ζ_k, i.e., $E(w_k\zeta_k^T)\neq O$. Thus, the LMMSE estimate $\hat{w}_{k|k}$ of w_k onto $\mathcal{L}(\zeta_1,\cdots,\zeta_k)$ also needs to be calculated, which is different from the existing LMMSE estimators in [4,6,37]. Moreover, the second-order moment matrices between ξ_k and w_k also needs to be further estimated.

Denote $\Omega_k:=E(\xi_k\xi_k^T)$ and $\Omega_{i,k}:=E(\xi_{i,k}\xi_{i,k}^T)$. Since $\Theta_k=i$ and $\Theta_k=j$ should not occur simultaneously in the Markov process for $i\neq j$, we have $E(\xi_{i,k}\xi_{j,k}^T)=O$. Thus, $\Omega_k=\mathrm{diag}\{\Omega_{i,k},i=1,\cdots,M\}$. Denote $\Pi_{k|l}:=E(\hat{\xi}_{k|l}\hat{\xi}_{k|l}^T)$, $\mathcal{X}_{k,l}:=E(\xi_k\zeta_l^T)$, $\Psi_k:=E(\zeta_k\zeta_k^T)$, $\mathrm{T}_{k|l}:=E(\hat{\xi}_{k|l}\hat{w}_{k|l}^T)$ and $\Lambda_{k|l}:=E(\hat{w}_{k|l}\hat{w}_{k|l}^T)$. Write $\bar{\mathcal{F}}_k$ as an $M\times M$ block matrix with its $(i,j)th$ sub-block being $p_{ij}\bar{F}_{j,k}$, $\bar{\mathcal{G}}_k$ as an $M\times 1$ block matrix with its $(i,1)th$ sub-block being $\sum_{j=1}^{M}p_{ij}\pi_{j,k}\bar{G}_{j,k}$, and $\bar{\mathcal{D}}_k:=\mathrm{diag}\{\bar{\mathcal{D}}_{n,k},n=1,\cdots,N\}$ with $\bar{\mathcal{D}}_{n,k}:=\sum_{i=1}^{M}\pi_{i,k}\bar{D}_{n,i,k}$ for $n=1,\cdots,N$. Denote the dimension of $\wp_k$ as $n_\wp$ with $\wp_k$ being a vector.

Theorem 1. *The LMMSE estimator for the system (4) and (5), abbreviated as LMRCE estimator, has the following recursive implementations:*

$$\hat{\xi}_{k+1|k+1}=\hat{\xi}_{k+1|k}+\mathcal{X}_{k+1,k+1}\Psi_{k+1}^{-1}\zeta_{k+1}, \tag{6}$$

$$\Phi_{k+1|k+1}=\Omega_{k+1}-\Pi_{k+1|k+1}, \tag{7}$$

$$\Omega_{i,k+1}=\sum_{j=1}^{M}p_{ij}(\bar{F}_{j,k}\Omega_{j,k}\bar{F}_{j,k}^T+\mathcal{C}_{j,k}^F)+\sum_{j=1}^{M}p_{ij}\pi_{j,k}(\bar{G}_{j,k}Q_k\bar{G}_{j,k}^T+\mathcal{C}_{j,k}^G), \tag{8}$$

$$\Pi_{k+1|k+1}=\Pi_{k+1|k}+\mathcal{X}_{k+1,k+1}\Psi_{k+1}^{-1}\mathcal{X}_{k+1,k+1}^T, \tag{9}$$

where

$$\hat{\xi}_{k+1|k} = \bar{\mathcal{F}}_k \hat{\xi}_{k|k} + \bar{\mathcal{G}}_k \hat{w}_{k|k}, \tag{10}$$

$$\mathcal{X}_{k+1,k+1} = \Phi_{k+1|k} \bar{H}_{k+1}^T, \tag{11}$$

$$\Psi_{k+1} = \mathrm{diag}\left\{ \sum_{j=1}^{M} \mathcal{C}_{n,j,k+1}^H + \pi_{j,k+1} (\bar{D}_{n,j,k+1} R_{n,k+1} \bar{D}_{n,j,k+1}^T + \mathcal{C}_{n,j,k+1}^D), n = 1, \cdots, N \right\} + \bar{H}_{k+1} \Phi_{k+1|k} \bar{H}_{k+1}^T, \tag{12}$$

$$\zeta_{k+1} = z_{k+1} - \bar{H}_{k+1} \hat{\xi}_{k+1|k}, \tag{13}$$

$$\Pi_{k+1|k} = \bar{\mathcal{F}}_k \Pi_{k|k} \bar{\mathcal{F}}_k^T + \bar{\mathcal{F}}_k \mathrm{T}_{k|k} \bar{\mathcal{G}}_k^T + \bar{\mathcal{G}}_k \Lambda_{k|k} \bar{\mathcal{G}}_k^T + \bar{\mathcal{G}}_k \mathrm{T}_{k|k}^T \bar{\mathcal{F}}_k^T, \tag{14}$$

$$\Phi_{k+1|k} = \Omega_{k+1} - \Pi_{k+1|k}, \tag{15}$$

with

$$\hat{w}_{k|k} = S_k \bar{\mathcal{D}}_k^T \Psi_k^{-1} \zeta_k, \tag{16}$$

$$\Lambda_{k|k} = S_k \bar{\mathcal{D}}_k^T \Psi_k^{-1} \bar{\mathcal{D}}_k S_k^T, \tag{17}$$

$$\mathrm{T}_{k|k} = \mathcal{X}_{k,k} \Psi_k^{-1} \bar{\mathcal{D}}_k S_k^T, \tag{18}$$

and

$$\mathcal{C}_{j,k}^F(p,q) = \sum_{s=1}^{n_x} \sum_{r=1}^{n_x} c_{jj,k}^{f_{pr,qs}} \Omega_{j,k}(r,s), \quad p,q = 1, \cdots, n_x, \tag{19}$$

$$\mathcal{C}_{j,k}^G(p,q) = \sum_{s=1}^{n_w} \sum_{r=1}^{n_w} c_{jj,k}^{g_{pr,qs}} Q_k(r,s), \quad p,q = 1, \cdots, n_x, \tag{20}$$

$$\mathcal{C}_{n,j,k}^H(p,q) = \sum_{s=1}^{n_x} \sum_{r=1}^{n_x} c_{n,jj,k}^{h_{pr,qs}} \Omega_{j,k}(r,s), \quad p,q = 1, \cdots, n_z, \tag{21}$$

$$\mathcal{C}_{n,j,k}^D(p,q) = \sum_{s=1}^{n_v} \sum_{r=1}^{n_v} c_{n,jj,k}^{d_{pr,qs}} R_{n,k}(r,s), \quad p,q = 1, \cdots, n_z. \tag{22}$$

Moreover, the initial conditions are as follows: $\hat{\xi}_{0|0} = \mathrm{col}\{\psi_i, i = 1, \cdots, M\}$, $\Pi_{0|0} = \hat{\xi}_{0|0} \hat{\xi}_{0|0}^T$, $\Omega_{i,0} = V_i$, $\hat{w}_{0|0} = \mathbf{0}$, $\Lambda_{0|0} = O$ *and* $\mathrm{T}_{0|0} = O$.

Proof. See Appendix A.

As similar to those in [4,6,37], $\Omega_{i,k+1}$, $\Pi_{k+1|k}$ and $\Pi_{k+1|k+1}$ in Theorem 1 are not dependent on the measurement data and can be calculated off-line. Meanwhile, since $c_{jj,k}^{f_{pr,sq}}$, $c_{jj,k}^{g_{pr,sq}}$, $c_{n,jj,k}^{h_{pr,sq}}$ and $c_{n,jj,k}^{d_{pr,sq}}$ are also not dependent on measurements, $\mathcal{C}_{j,k}^F$, $\mathcal{C}_{j,k}^G$, $\mathcal{C}_{n,j,k}^H$ and $\mathcal{C}_{n,j,k}^D$ caused by random parameter matrices and state or additive noises can be also calculated off-line. Moreover, although $f_{i,k}^{pq}$ and $f_{j,k}^{rs}$ coming from different modes are correlated, there is no effect on the calculation of $\mathcal{C}_{j,k}^F$ (as similar to other items $\mathcal{C}_{j,k}^G$, $\mathcal{C}_{n,j,k}^H$ and $\mathcal{C}_{n,j,k}^D$). This is because different modes in the MJS cannot coexist at the same epoch, i.e., there is only one valid mode at each sampling instant.

Without considering the cross-correlation between w_k and v_k, although random parameter matrices in the MRMCN are equivalent to stochastic coefficient matrices in [37], the considered randomness comes from each element of system matrices while it is from the random coefficients of polytopic parameter matrices in [37] as similar to that in [6]. Thus, we give the calculating formula of each element in the related matrix as shown in Eqs. (19)–(22), while it gives the calculating formula about the related matrix directly as Eq. (16) in [37].

Remark 3.1. Due to the coexistence of random matrices, mode uncertainty and cross-correlated additive noises, it is inevitable to calculate the effects of these complex coupling. The process noise has to be estimated due to the cross-correlation, which further leads to the unknown coupling of state and process noise. The high-order moment matrices (i.e., $\mathcal{C}_{j,k}^F$, $\mathcal{C}_{j,k}^G$, $\mathcal{C}_{n,j,k}^H$ and $\mathcal{C}_{n,j,k}^D$) are calculated because of the randomness of system matrices coupled with unknown state or additive noises and mode uncertainty.

In a single sensor case (i.e., $N = 1$), if w_k and v_k are not cross-correlated, i.e., $S_k = O$, the derived LMRCE estimator will degrade to the LMMSE estimator in [4] when system matrices are determined (i.e., $c_{ij,k}^{f_{pq,rs}} = 0$, $c_{ij,k}^{g_{pq,rs}} = 0$, $c_{n,ij,k}^{h_{pq,rs}} = 0$ and $c_{n,ij,k}^{d_{pq,rs}} = 0$); or the filter in [22] when there is only one mode (i.e., $M = 1$).

3.2. Square-root array implementation

In the derived LMRCE estimator, $\Phi_{k+1|k}$ defined in Eq. (15), representing the second-order moment of the state prediction error under the LMMSE criterion, should be positive semidefinite. However, roundoff errors can cause a loss of positive definiteness since it is computed based on Riccati equation [5]. An effective way to improve this situation and get a more stable realization is to develop the square-root array implementation, i.e., $\Phi_{k+1|k} = \Phi_{k+1|k}^{1/2}\Phi_{k+1|k}^{T/2}$, to own the properties of better conditioning and reduced dynamical range [12,30,31].

Definition 1 [13]. For any signature matrix $\mathcal{I}$ (a diagonal matrix with $+1$ and -1 diagonal elements), the matrix $\mathcal{U}_{\mathcal{I}}$ will be called $\mathcal{I}$-unitary if $\mathcal{U}_{\mathcal{I}}\mathcal{I}\mathcal{U}_{\mathcal{I}}^T = \mathcal{I}$.

Define $\mathcal{V}_k := E(\tilde{H}_k\xi_k\xi_k^T\tilde{H}_k^T) + E(\tilde{D}_k v_k v_k^T\tilde{D}_k^T) + E(\tilde{D}_k v_k v_k^T\tilde{D}_k^T)$ as the total covariance of the equivalent measurement noise.

Corollary 1. *$\mathcal{V}_k$ is symmetric and positive semidefinite, and can be decomposed as $\mathcal{V}_k = \mathcal{V}_k^{1/2}\mathcal{V}_k^{T/2}$.*

Denote $\Upsilon_k := \tilde{H}_k^T\mathcal{V}_k^{-1}\tilde{H}_k$, $\Gamma_k := \tilde{H}_k^T\mathcal{V}_k^{-1}\tilde{\mathcal{D}}_k S_k^T$, $\Delta_k := S_k\tilde{\mathcal{D}}_k^T\mathcal{V}_k^{-1}\tilde{\mathcal{D}}_k S_k^T$, $\mathcal{V}_{n,k} := E(\tilde{H}_{n,k}\xi_k\xi_k^T\tilde{H}_{n,k}^T) + E(\tilde{D}_{n,k}v_{n,k}v_{n,k}^T\tilde{D}_{n,k}^T) + E(\tilde{D}_{n,k}v_{n,k}v_{n,k}^T\tilde{D}_{n,k}^T)$, $\Upsilon_{n,k} := \tilde{H}_{n,k}^T\mathcal{V}_{n,k}^{-1}\tilde{H}_{n,k}$, $\Gamma_{n,k} := \tilde{H}_{n,k}^T\mathcal{V}_{n,k}^{-1}\tilde{\mathcal{D}}_{n,k}S_{n,k}^T$ *and* $\Delta_{n,k} := S_{n,k}\tilde{\mathcal{D}}_{n,k}^T\mathcal{V}_{n,k}^{-1}\tilde{\mathcal{D}}_{n,k}S_{n,k}^T$. *Hence, we have* $\Upsilon_k = \sum_{n=1}^N\Upsilon_{n,k}$, $\Gamma_k = \sum_{n=1}^N\Gamma_{n,k}$ *and* $\Delta_k = \sum_{n=1}^N\Delta_{n,k}$. *Denote* $\bar{\Upsilon}_k := \Upsilon_k/N$, $\bar{\Gamma}_k := \Gamma_k/N$ *and* $\bar{\Delta}_k := \Delta_k/N$.

For simplicity, write $\tilde{\mathcal{W}}_{i,k}^F := \left[p_{i1}^{1/2}\tilde{F}_{1,k}\Omega_{1,k}^{1/2} \cdots p_{iM}^{1/2}\tilde{F}_{M,k}\Omega_{M,k}^{1/2}\right]$, $\tilde{\mathcal{W}}_{i,k}^F := \left[p_{i1}^{1/2}\mathcal{C}_{1,k}^{F/2} \cdots p_{iM}^{1/2}\mathcal{C}_{M,k}^{F/2}\right]$, $\tilde{\mathcal{W}}_{i,k}^G := \left[p_{i1}^{1/2}\pi_{1,k}^{1/2}\tilde{G}_{1,k}Q_k^{1/2} \cdots p_{iM}^{1/2}\pi_{M,k}^{1/2}\tilde{G}_{M,k}Q_k^{1/2}\right]$, $\tilde{\mathcal{W}}_{i,k}^G := \left[p_{i1}^{1/2}\pi_{1,k}^{1/2}\mathcal{C}_{1,k}^{G/2} \cdots p_{iM}^{1/2}\pi_{M,k}^{1/2}\mathcal{C}_{M,k}^{G/2}\right]$, $\tilde{\mathcal{W}}_{n,k}^H := \left[\mathcal{C}_{n,1,k}^{H/2} \cdots \mathcal{C}_{n,M,k}^{H/2}\right]$, $\tilde{\mathcal{W}}_{n,k}^D := \left[\pi_{1,k}^{1/2}\tilde{D}_{n,1,k}R_{n,k}^{1/2} \cdots \pi_{M,k}^{1/2}\tilde{D}_{n,M,k}R_{n,k}^{1/2}\right]$, $\tilde{\mathcal{V}}_{n,k}^D := \left[\pi_{1,k}^{1/2}\mathcal{C}_{n,1,k}^{D/2} \cdots \pi_{M,k}^{1/2}\mathcal{C}_{n,M,k}^{D/2}\right]$. $\mathcal{C}_{i,k}^{F/2} = (\mathcal{C}_{i,k}^F)^{1/2}$, *similarly to* $\mathcal{C}_{i,k}^{G/2}$, $\mathcal{C}_{n,i,k}^{H/2}$ *and* $\mathcal{C}_{n,i,k}^{D/2}$. *Without considering the matrix singularity, the square-root array implementation for the proposed LMRCE estimator is given as follows.*

Theorem 2. *The square-root array implementation for the proposed LMRCE estimator in Theorem 1, abbreviated as SALME, is shown as follows:*

$$\begin{bmatrix}\Omega_{1,k+1}^{1/2} & O & \cdots & O & O\\ O & \Omega_{2,k+1}^{1/2} & \cdots & O & O\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ O & O & O & \Omega_{M,k+1}^{1/2} & O\end{bmatrix} = \begin{bmatrix}\mathcal{K}_{1,k+1} & O & \cdots & O\\ O & \mathcal{K}_{2,k+1} & \cdots & O\\ \vdots & \vdots & \ddots & \vdots\\ O & O & \cdots & \mathcal{K}_{M,k+1}\end{bmatrix}U_{\Omega_{k+1}}, \tag{23}$$

$$\begin{bmatrix}\mathcal{Z}_{k+1}^{1/2} & O & O\\ \Xi_{k+1} & \Pi_{k+1|k+1}^{1/2} & O\end{bmatrix} = \begin{bmatrix}\Phi_{k+1|k}^{-T/2} & \sqrt{N}\bar{\Upsilon}_{k+1}^{1/2} & O\\ O & \sqrt{N}\Phi_{k+1|k}\bar{\Upsilon}_{k+1}^{1/2} & \Pi_{k+1|k}^{1/2}\end{bmatrix}U_{\Pi_{k+1|k+1}}, \tag{24}$$

$$\begin{bmatrix}\Phi_{k+1|k+1}^{1/2} & O\end{bmatrix} = \begin{bmatrix}\Omega_{k+1}^{1/2} & \Pi_{k+1|k+1}^{1/2}\end{bmatrix}\mathcal{U}_{\Phi_{k+1|k+1}} \tag{25}$$

where

$$\mathcal{K}_{i,k+1} := \begin{bmatrix}\tilde{\mathcal{W}}_{i,k}^F & \tilde{\mathcal{W}}_{i,k}^F & \tilde{\mathcal{W}}_{i,k}^G & \tilde{\mathcal{W}}_{i,k}^G\end{bmatrix},\ i = 1, \cdots, M, \tag{26}$$

$$\begin{bmatrix}\mathcal{Z}_k^{1/2} & O\end{bmatrix} = \begin{bmatrix}\Phi_{k|k-1}^{-T/2} & \sqrt{N}\bar{\Upsilon}_k^{1/2}\end{bmatrix}U_{\mathcal{Z}_k}, \tag{27}$$

$$\Xi_{k+1} = N\Phi_{k+1|k}\bar{\Upsilon}_{k+1}\mathcal{Z}_{k+1}^{-1/2}, \tag{28}$$

$$\begin{bmatrix}\Phi_{k+1|k}^{1/2} & O\end{bmatrix} = \begin{bmatrix}\Omega_{k+1}^{1/2} & \Pi_{k+1|k}^{1/2}\end{bmatrix}\mathcal{U}_{\Phi_{k+1|k}}, \tag{29}$$

$$\begin{bmatrix}\Pi_{k+1|k}^{1/2} & O\end{bmatrix} = \begin{bmatrix}\bar{\mathcal{F}}_k\Pi_{k|k}^{1/2} + \bar{\mathcal{G}}_k\Upsilon_{k|k}^T\Pi_{k|k}^{-T/2} & \bar{\mathcal{G}}_k\mathcal{B}_k\end{bmatrix}U_{\Pi_{k+1|k}}, \tag{30}$$

$$\begin{bmatrix}\mathcal{B}_k & O\end{bmatrix} = \begin{bmatrix}\Lambda_{k|k}^{1/2} & \Upsilon_{k|k}^T\Pi_{k|k}^{-T/2}\end{bmatrix}\mathcal{U}_{\mathcal{B}_k}, \tag{31}$$

with

$$\begin{bmatrix} \mathcal{Z}_k^{1/2} & O & O \\ \Gamma_k^T \mathcal{Z}_k^{-1/2} & \Lambda_{k|k}^{1/2} & O \end{bmatrix} = \begin{bmatrix} \Phi_{k|k-1}^{-T/2} & \bar{H}_k^T \mathcal{V}_k^{-T/2} & O \\ O & S_k \bar{\mathcal{D}}_k^T \mathcal{V}_k^{-1/2} & O \end{bmatrix} U_{\Lambda_{k|k}}, \tag{32}$$

or

$$\begin{bmatrix} \Lambda_k^{1/2} & O \end{bmatrix} = \begin{bmatrix} \sqrt{N}\bar{\Delta}_k^{1/2} & N\bar{\Gamma}_k^T \mathcal{Z}_k^{-T/2} \end{bmatrix} \mathcal{U}_{\Lambda_{k|k}}, \tag{33}$$

where

$$\begin{bmatrix} \mathcal{V}_{k+1}^{1/2} & O \end{bmatrix} = \mathrm{diag}\{\mathcal{Y}_{n,k+1}, n = 1, \cdots, N\} U_{\mathcal{V}_{k+1}}, \tag{34}$$

$$\mathcal{Y}_{n,k+1} := \begin{bmatrix} \bar{\mathcal{V}}_{n,k+1}^H & \bar{\mathcal{V}}_{n,k+1}^D & \bar{\mathcal{V}}_{n,k+1}^D \end{bmatrix}. \tag{35}$$

Moreover, the initial conditions are as follows: $\Omega_{0|0}^{1/2} = \mathrm{diag}\{V_i^{1/2}, i = 1, \cdots, M\}$ *and* $\Pi_{0|0}^{1/2} = (\xi_{0|0}\xi_{0|0}^T)^{1/2}$. $U_{\Omega_{k+1}}$, $U_{\Pi_{k+1|k+1}}$, $U_{\Pi_{k+1|k}}$, $U_{\mathcal{Z}_k}$, U_{Λ_k} *and* $U_{\mathcal{V}_{k+1}}$ *are corresponding unitary matrices.* $\mathcal{U}_{\Phi_{k+1|k+1}}$, $\mathcal{U}_{\Phi_{k+1|k}}$, $\mathcal{U}_{\mathcal{B}_k}$ *and* $\mathcal{U}_{\Lambda_{k|k}}$ *are corresponding* $\mathcal{I}$*-unitary matrices and* $\mathcal{I} = \mathrm{diag}(I, -I)$.

Proof. See Appendix B.

If A is postive definite, $A^{1/2}$ can be realized by Cholesky decomposition. If it is positive semidefinite, $A^{1/2}$ can be also realized in a similar way [14].

If there is no random parameter matrices, i.e., $\bar{\mathcal{V}}_{k+1}^F$, $\bar{\mathcal{V}}_{k+1}^G$, $\bar{\mathcal{V}}_{n,k+1}^H$ and $\bar{\mathcal{V}}_{n,k+1}^D$ are all zero matrices, then, $\mathcal{K}_{i,k+1} = [\bar{\mathcal{W}}_{i,k}^F \quad \bar{\mathcal{W}}_{i,k}^G]$ and $\mathcal{Y}_{n,k+1} = \bar{\mathcal{W}}_{n,k+1}^D$. The proposed SALME will degrade to the square-root array implementation of LMMSE estimator for MJLS with cross-correlated noises. Moreover, if there is no cross-correlation between w_k and v_k, i.e., $S_k = O$, it is not required to calculate Λ_{k+1} and Eq. (30) will be rewritten as $\Pi_{k+1|k}^{1/2} = \bar{\mathcal{F}}_k \Pi_{k|k}^{1/2} U_{\Pi_{k+1|k}}$. Then, the Theorem 2 under a single sensor condition will degrade to the square-root array implementation of LMMSE estimator for MJLS in [31].

Remark 3.2. Besides considering the system noise correlation and random parameter matrices, the calculations of $\Pi_{k+1|k+1}^{1/2}$ in Theorem 2 and array algorithm in [31] are significantly different. Υ_k is treated as an entirety in Theorem 2, while $\Upsilon_k^{1/2}\Upsilon_k^{T/2} = (\bar{H}_k^T \mathcal{V}_k^{-T/2})(\bar{H}_k^T \mathcal{V}_k^{-T/2})^T$ is adopted in [31], to calculate $\Pi_{k+1|k+1}^{1/2}$ (as similar to Eq. (32)). With the number of sensors increasing, the dimension of $\Upsilon_k^{1/2}$ will not be changed, while the dimension of $\bar{H}_k^T \mathcal{V}_k^{-T/2}$ will increase linearly and this will lead to numerical instability and inaccuracy in matrix calculations. Similarly, it is also better to use Eq. (33) instead of Eq. (32) to calculate $\Lambda_k^{1/2}$ in the multi-sensor case. Meanwhile, similar square-root triangularized formulas should also be adopted for Γ_k and Δ_k.

4. Distributed fusion estimation with square-root array implementation in sensor networks

After deriving the centralized LMRCE estimator with SALME in Section 3, we will present the corresponding distributed method in sensor networks in this section, by considering the sensor decentralized deployment and advantages of distributed realization including scalability, low communication load, fast implementation and more robustness to sensor failures [24,25,36].

The centralized LMRCE estimator has been derived by treating all measurements from different sensors as a whole via measurement augmentation. However, such structure is difficult to split the whole available measurement information into different parts corresponding to different sensor nodes for distributed fusion. A better framework is the information filter form, where the final estimate is dependent on the sum of all measurement information from different sensors. Thus, the corresponding information filter form will be given in Theorem 3.

Define $\mathcal{J}_{k|l} := \Phi_{k|l}^{-1}$, $\hat{\vartheta}_{k|l} := \Phi_{k|l}^{-1}\hat{\xi}_{k|l}$. Denote $\beta_k := \bar{H}_k^T \mathcal{V}_k^{-1} z_k$, $\gamma_k := S_k \bar{\mathcal{D}}_k^T \mathcal{V}_k^{-1} \zeta_k$ and $\varphi_k := \bar{H}_k^T \mathcal{V}_k^{-1} \zeta_k$. And denote $\beta_{n,k} := \bar{H}_{n,k}^T \mathcal{V}_{n,k}^{-1} z_{n,k}$, $\gamma_{n,k} := S_{n,k} \bar{\mathcal{D}}_{n,k}^T \mathcal{V}_{n,k}^{-1} \zeta_{n,k}$ and $\varphi_{n,k} := \bar{H}_{n,k}^T \mathcal{V}_{n,k}^{-1} \zeta_{n,k}$ with $\zeta_{n,k}$ being the corresponding innovation of the *nth* sensor. Hence, we have $\beta_k = \sum_{n=1}^N \beta_{n,k}$, $\gamma_k = \sum_{n=1}^N \gamma_{n,k}$ and $\varphi_k = \sum_{n=1}^N \varphi_{n,k}$. Write $\bar{\beta}_k := \beta_k/N$, $\bar{\gamma}_k := \beta_k/N$ and $\bar{\varphi}_k := \beta_k/N$.

Theorem 3. *The information filter form of the proposed LMRCE estimator, abbreviated as IMRCF, has the following recursive implementations:*

$$\hat{\vartheta}_{k+1|k+1} = \hat{\vartheta}_{k+1|k} + N\bar{\beta}_{k+1}, \tag{36}$$

$$\mathcal{J}_{k+1|k+1} = \mathcal{J}_{k+1|k} + N\bar{\Upsilon}_{k+1}, \tag{37}$$

$$\Pi_{k+1|k+1} = (\Omega_{k+1} - \mathcal{J}_{k+1|k+1}^{-1}), \tag{38}$$

where

$$\hat{\vartheta}_{k+1|k} = \mathcal{J}_{k+1|k}(\bar{\mathcal{F}}_k\hat{\xi}_{k|k} + \bar{\mathcal{G}}_k\hat{w}_{k|k}), \tag{39}$$

$$\mathcal{J}_{k+1|k} = (\Omega_{k+1} - \Pi_{k+1|k})^{-1}, \tag{40}$$

with

$$\hat{w}_{k|k} = N\bar{\gamma}_k - N^2\bar{\Gamma}_k^T(\mathcal{J}_{k|k-1} + N\bar{\Upsilon}_k)^{-1}\bar{\varphi}_k, \tag{41}$$

$$\Lambda_{k|k} = N\bar{\Delta}_k - N^2\bar{\Gamma}_k^T(\mathcal{J}_{k|k-1} + N\bar{\Upsilon}_k)^{-1}\bar{\Gamma}_k, \tag{42}$$

$$\mathrm{T}_{k|k} = N\mathcal{J}_{k|k-1}^{-1}\bar{\Gamma}_k - N^2\mathcal{J}_{k|k-1}^{-1}\bar{\Upsilon}_k(\mathcal{J}_{k|k-1} + N\bar{\Upsilon}_k)^{-1}\bar{\Gamma}_k. \tag{43}$$

Proof. See Appendix C.

In Theorem 3, the remained items Ω_{k+1} and $\Pi_{k+1|k}$, unrelated to multiple sensors, are calculated as the same as those in Theorem 1.

In the view of the information filter form, $\Phi_{k+1|k}$, $\Phi_{k+1|k+1}$ and $\Phi_{k+1|k}^{1/2}$ in the SALME in Theorem 2 can be replaced by $\mathcal{J}_{k+1|k}^{-1}$, $\mathcal{J}_{k+1|k+1}^{-1}$ and $\mathcal{J}_{k+1|k}^{-T/2}$, respectively. Here, $\left[\Pi_{k+1|k+1}^{1/2} \quad O\right] = \left[\Omega_{k+1}^{1/2} \quad \mathcal{J}_{k+1|k+1}^{-T/2}\right]\mathcal{U}_{\Pi_{k+1|k+1}}$ and $\left[\mathcal{J}_{k+1|k+1}^{1/2} \quad O\right] = \left[\mathcal{J}_{k+1|k}^{1/2} \quad \sqrt{N}\bar{\Upsilon}_{k+1}^{1/2}\right]U_{\mathcal{J}_{k+1|k+1}}$ with $\mathcal{J}_{k+1|k}^{1/2} = \Phi_{k+1|k}^{-T/2}U_{\mathcal{J}_{k+1|k}}$, where $U_{\mathcal{J}_{k+1|k}}$ and $U_{\mathcal{J}_{k+1|k+1}}$ are corresponding unitary matrices and $\mathcal{U}_{\Pi_{k+1|k+1}}$ is the corresponding $\mathcal{I}$-unitary matrix with $\mathcal{I} = \mathrm{diag}(I, -I)$.

Remark 4.1. A centralized estimation fusion for the MJLS via data transformation is presented in [9], by using the information filter form. The differences between the proposed IMRCF and the method in [9] are as follows: Firstly, the proposed IMRCF is the linear optimal filter, while the method in [9] is suboptimal based on minimum mean square error criterion with Gaussian posterior density hypotheses and Gaussian noises assumptions; Secondly, Θ_k is only considered in the state evolvement model in [9] while it is also contained in the measurement equation in the considered MRMCNs. Since Θ_k in the measurement equation will affect the transformed measurements, the method in [9] cannot be applied to the considered MRMCNs; Thirdly, due to the coexistence of system matrix randomness and cross-correlation between w_k and v_k, the proposed IMRCF considers simultaneously the calculation of higher-order moment matrices about the stochastic parameters and state or additive noises (i.e., $\mathcal{C}_{j,k}^F$, $\mathcal{C}_{j,k}^G$, $\mathcal{C}_{n,j,k}^H$, $\mathcal{C}_{n,j,k}^D$) and the coupling between state and process noise (i.e., $\mathrm{T}_{k|k}$).

Besides the dynamical state evolvement itself, the proposed IMRCF and its square-root array implementation just depend on average values $\bar{\beta}_k$, $\bar{\gamma}_k$, $\bar{\varphi}_k$, $\bar{\Upsilon}_k$, $\bar{\Gamma}_k$ and $\bar{\Delta}_k$, which synthesize the corresponding values from different sensor nodes. In this way, in the following distributed fusion, only those intermediate vectors and matrices from different nodes need to be shared to gain the consensus values and achieve a balance in the whole sensor network. Based on this, the consensus algorithm, as a common and famous method [24,25], can be adopted here to give realizations for the distributed IMRCF and SALME (abbreviated as D-IMRCF and D-SALME), by obtaining those average values as follows:

$$\mathcal{X}_{n,k}^{i+1} = \mathcal{X}_{n,k}^i + \varepsilon\sum_{m\in\mathcal{N}_n}\left(\mathcal{X}_{m,k}^i - \mathcal{X}_{n,k}^i\right) + \varepsilon\sum_{m\in\mathbb{N}_n}\left(\Im_{m,k} - \mathcal{X}_{n,k}^i\right), \tag{44}$$

where $\mathcal{X}_{n,k}^i$ represents the required average vector or matrix ($\bar{\beta}_k$, $\bar{\gamma}_k$, $\bar{\varphi}_k$, $\bar{\Upsilon}_k$, $\bar{\Gamma}_k$ or $\bar{\Delta}_k$) in the *nth* node at the *kth* epoch in the *ith* iteration, $\Im_{m,k}$ denotes the corresponding original value calculated in the *mth* node and ε is the iteration step size. $\mathcal{N}_n$ represents the set of neighbors of the *nth* node and $\mathbb{N}_n := \mathcal{N}_n \cup \{n\}$. Meanwhile, $\mathcal{X}_{n,k}^1$ is set as $\Im_{n,k}$. The iteration termination condition is $\|\mathcal{X}_{n,k}^i - \mathcal{X}_{n,k}^{i-1}\| \le \hbar$ or $i > \ell_{\max}$, where $\|\cdot\|$ denotes the 2−norm of a vector or matrix, $\hbar$ is a small positive threshold and $\ell_{\max}$ is the maximum iteration step.

For briefness, assume that $\mathcal{X}_{n,k}^i$ is scalar. Denote $\mathcal{X}_k^i := (\mathcal{X}_{1,k}^i, \cdots, \mathcal{X}_{N,k}^i)^T$ and $\Im_k := (\Im_{1,k}, \cdots, \Im_{N,k})^T$. Let $\mathcal{A}$ be the adjacent matrix that specifies the interconnection topology of $\mathbb{G}$, i.e., $\mathcal{A}(m,n) = 1$ if there is an edge e_{mn} and 0 otherwise, and $\mathcal{A}(n,n) = 0$, for $m, n = 1, \cdots, N$. $\Xi := \mathrm{diag}\{\mathcal{A}\cdot\mathbf{1}_N\}$ with $\Xi(n,n) = \sum_{m=1}^N \mathcal{A}(n,m)$. Moreover, a weight a_{mn} can be given to each edge to reconstruct a new adjacent matrix $\mathbb{A}$, i.e., $\mathbb{A}(m,n) = a_{mn}$ if $\mathcal{A}(m,n) > 0$ and $\mathbb{A}(m,n) = 0$ if $\mathcal{A}(m,n) = 0$. Then, the following Theorem is proposed to discuss the convergence of the above consensus method in Eq. (44) and show the condition that the resultant convergent result will reach the equilibrium.

Theorem 4. *If all eigenvalues of $I + \varepsilon\mathcal{A} - 2\varepsilon\Xi - \varepsilon I$ are located strictly inside the unit circle in the complex plane, i.e., $0 < \varepsilon < \frac{2}{1+3\max\{\Xi(n,n), n=1,\cdots,N\}}$, we have $\lim_{i\to\infty}\mathcal{X}_k^i = (I + 2\Xi - \mathcal{A})^{-1}(\mathcal{A} + I)\Im_k$. Moreover, $\lim_{i\to\infty}\mathcal{X}_k{}^i = \frac{1}{N}\mathbf{1}_N\otimes\mathbf{1}_N^T\Im_k$ (i.e., each element of $\mathcal{X}_k^i$ reaches the equilibrium), with $\mathcal{A}$ being replaced by $\mathbb{A}$ and the following condition holding:*

$$C\mu = b, \tag{45}$$

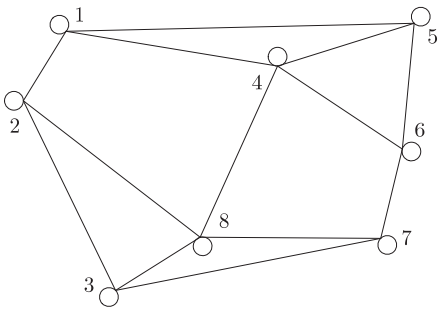

Fig. 1. A sensor network $\mathbb{G}$ observing the moving target.

where $b = \left(\frac{1}{N}\mathbf{1}^T_{N(N-1)/2}, \left(\frac{1}{N}-1\right)\mathbf{1}^T_N\right)^T$, $\mu = \text{col}\{a_{i'j'},\ i', j' = 1, 2, \cdots, N,\ i' > j',\ \mathcal{A}(i', j') > 0\}$ *and*

$$C(s,r) = \begin{cases} 1-\frac{1}{N}, & \frac{i(i-1)}{2} < s \le \frac{i(i+1)}{2}, \quad j = s - \frac{i(i-1)}{2}, \quad j = j' \text{ or } i', \quad i+1 = i', \\ -\frac{1}{N}, & \frac{i(i-1)}{2} < s \le \frac{i(i+1)}{2}, \quad j = s - \frac{i(i-1)}{2}, \quad j = j' \text{ or } i', \quad i+1 \neq i', \\ 0, & \text{else}, \end{cases}$$

with the corresponding rth element of μ *being* $a_{i'j'}$ *for* $s = 1, \cdots, \frac{N(N+1)}{2}$ *and* $r = 1, \cdots, n_\mu$. *The least squares solution for Eq. (45) to construct* $\mathbb{A}$ *is* $\mu_{LS} = (C^TC)^+C^Tb$, *where* $(C^TC)^+$ *represents the Moore-Penrose pseudoinverse.*

Proof. See Appendix D.

In fact, there may exist equivalent equations in Eq. (45). Hence, it is better to remove these equivalent equations and remain only one. Then, Eq. (45) will be simplified as $C_o\mu = b_o$. If C_o is full column rank, the corresponding least squares solution is $\mu_{LS} = (C_o^TC_o)^{-1}C_o^Tb_o$. If C_o is full row rank, the corresponding general solution is $\mu_s = \mu^* + \mu^\perp$ with μ^* being a particular solution and $\mu^\perp$ being a generic vector in the null space, i.e., $C_o\mu^\perp = \mathbf{0}$.

Since that there are only additive and scalar-multiplicative operations in Eq. (44), the above Theorem 4 still holds if $\mathcal{X}^i_{n,k}$ is a vector or matrix, by the fact that Eq. (44) is adopted for each element individually.

5. Numerical example

A maneuvering target tracking example perturbed by multiplicative and cross-correlated additive noises is simulated to validate the proposed method. The target state is $(\varsigma_k, \dot{\varsigma}_k, \eta_k, \dot{\eta}_k)$ in the Cartesian coordinates $o-\varsigma\eta$. Its motion is based on two Markovian states evolving as Eq. (1), i.e., the constant velocity (CV) and constant turn (CT) motion, where $F_1 = I_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$, $F_2 = \begin{bmatrix} 1 & \sin(\omega T)/\omega & 0 & -(1-\cos(\omega T))/\omega \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) \\ 0 & (1-\cos(\omega T))/\omega & 1 & \sin(\omega T)/\omega \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) \end{bmatrix}$ and $G_1 = G_2 = I_2 \otimes \begin{bmatrix} T^2/2 & T \end{bmatrix}^T$ with $T = 1$ and $\omega = 0.1047$. In the first 10 epochs, the movement obeys the CV motion. Then, it switches to the CT motion in the next 30 epochs. Finally, it goes back to the CV motion in the last 20 epochs. w_k is Gaussian white noise and $Q_k = \sigma_w^2 I_2$ with $\sigma_w = 1$. The initial state x_0 is set as $\bar{x}_0 = (100, 3.5, 80, 3)^T$ added with a zero-mean, Gaussian distributed vector with covariance $P_0 = 10^2 \times \text{diag}\{1, 0.01, 1, 0.01\}$.

Both a single sensor and sensor network as shown in Fig. 1 are considered to detect and track the target. All sensors are homogeneous and $z_{n,k} = \begin{bmatrix} 1+\tau^\varsigma_{n,\Theta_k} & 0 & 0 & 0 \\ 0 & 0 & 1+\tau^\eta_{n,\Theta_k} & 0 \end{bmatrix} x_k + \nu_{n,k}$ for $n = 1, 2, \cdots, 8$. $\tau^\varsigma_{n,\Theta_k}$ and τ^η_{n,Θ_k} are both Gaussian white noises with variances being $\sigma^2_{\varsigma,\Theta_k}$ and σ^2_{η,Θ_k}, respectively. $\sigma^2_{\varsigma,CV} = \sigma^2_{\eta,CV} = 0.023^2$ and $\sigma^2_{\varsigma,CT} = \sigma^2_{\eta,CT} = 0.025^2$. $\nu_{n,k}$ is the Gaussian white noise and $R_{n,k} = \sigma_\nu^2 I_2$ with $\sigma_\nu = 5$. $S_{n,k} = \rho\sigma_w\sigma_\nu I_2$.

The simulation is performed with two different numerical implementations: float-point configuration (64-bit) and fixed-point configuration (word length 5 bytes and fraction length 3 bytes, i.e., from $-32,768$ to $32,768$ with the fraction part being ignored and up to 1).

In the following compared methods, the initial estimates and parameters are all equal: $\hat{x}_{0|0} = \bar{x}_0$ with $P_{0|0} = P_0$; $p_{11} = p_{22} = 0.95$; $\pi_{1,k} = 0.9$, $\pi_{2,k} = 0.1$ for the first 10 epochs, $\pi_{1,k} = 0.1$, $\pi_{2,k} = 0.9$ for the middle 30 epochs, and $\pi_{1,k} = 0.9$, $\pi_{2,k} = 0.1$ for the last 20 epochs.

5.1. A single sensor case

In this subsection, compared with the LMMSE estimator in [4] and LMSCE estimator in [6,37], the proposed LMRCE estimator with its SALME in Section 3 is validated in a single sensor case under both float-point configuration and fixed-

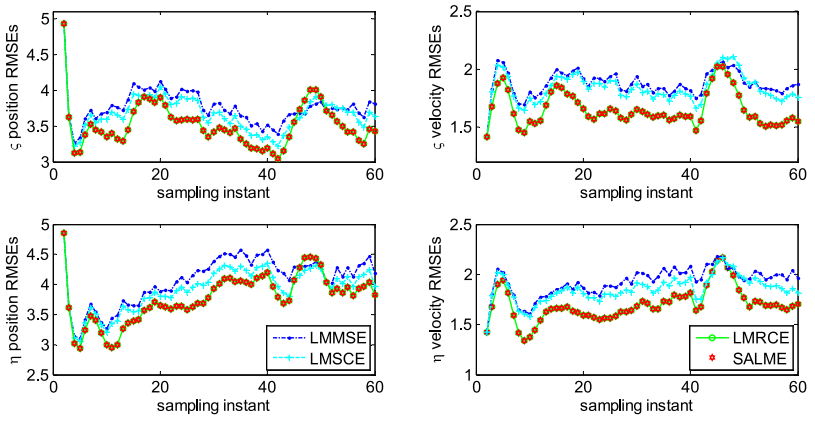

Fig. 2. RMSEs of compared methods with float-point configuration in a single sensor.

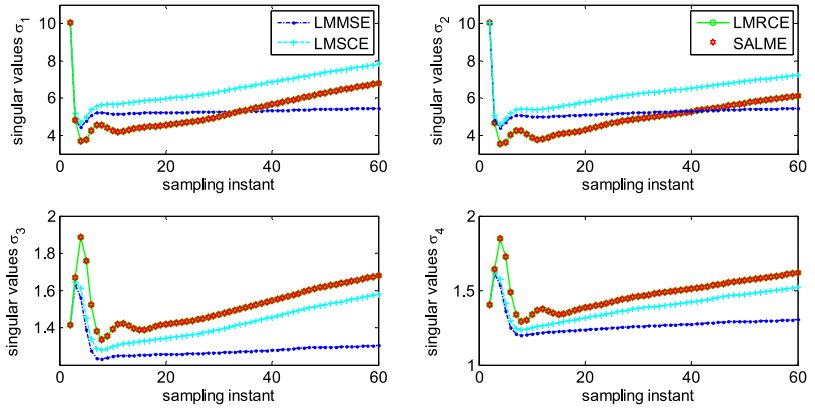

Fig. 3. SVs of compared methods with float-point configuration in a single sensor.

point configuration (abbreviated as F-LMRCE estimator and F-SALME, respectively). $\rho = 0.9$. The root mean square errors (RMSEs) of compared methods based on 1000 Monte Carlo realizations and corresponding singular values (SVs) σ_i of $P_{k|k-1}$ ($\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \sigma_4 > 0$) are shown in Figs. 2–5.

Under float-point configuration, as shown in Fig. 2, RMSEs of the LMRCE estimator and SALME are both smaller than those of the LMMSE and LMSCE estimators, due to considering the effect of random parameter matrices and cross-correlation between w_k and v_k. Meanwhile, in Fig. 3, the main SVs σ_1 and σ_2 in the LMRCE estimator and SALME are also much smaller than those of the LMSCE estimator. Although σ_1 and σ_2 from the LMMSE estimator are smaller than those of the LMRCE estimator or SALME, they cannot reflect the estimate accuracy since that the multiplicative noise which makes SVs increase is not considered. Therefore, in the perspectives of RMSEs and SVs, the estimate performance of the LMRCE estimator and SALME are both more accurate than those of the LMMSE and LMSCE estimators. Additionally, the estimate accuracy of the LMRCE estimator is equivalent to that of the SALME since RMSEs and SVs of both methods are the same.

Moreover, under fixed-point configuration, as shown in Figs. 4 and 5, the accuracy performance of the F-SALME is the same as that of the LMRCE estimator (or SALME), no matter from the view of RMSEs or SVs. However, RMSEs and SVs of the F-LMRCE estimator increase abruptly with target moving far away (i.e., the target position becomes large) due to the existence of roundoff errors. This demonstrates that the proposed SALME has advantages of better conditioning and reduced dynamical range.

5.2. A sensor network case

In this subsection, compared with the LMRCE estimator with a single sensor and centralized multiple sensors (denoted as C-LMRCE estimator), the D-IMRCF and D-SALME are testified in the sensor network case shown in Fig. 1, where both are also implemented under float-point configuration and fixed-point configuration (abbreviated as FD-IMRCF and FD-SALME, respectively). In each sensor node, $\rho = 0.3$. $\hbar = 10^{-6}$, $\ell_{\max} = 50$ and $\varepsilon = 0.05$ in Eq. (44).

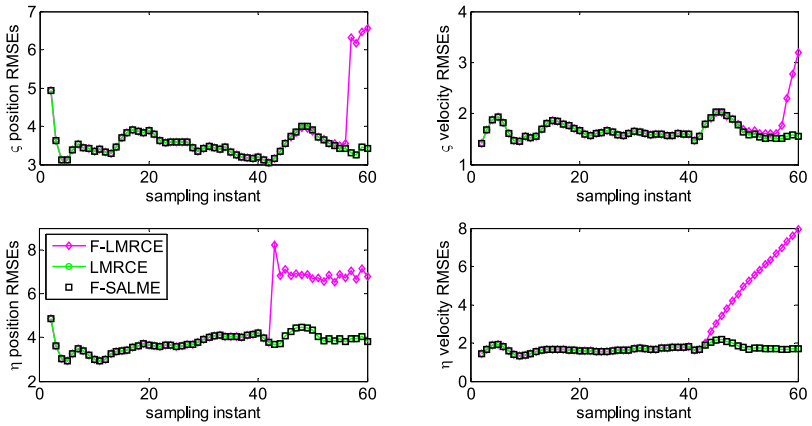

Fig. 4. RMSEs of compared methods with fixed-point configuration in a single sensor.

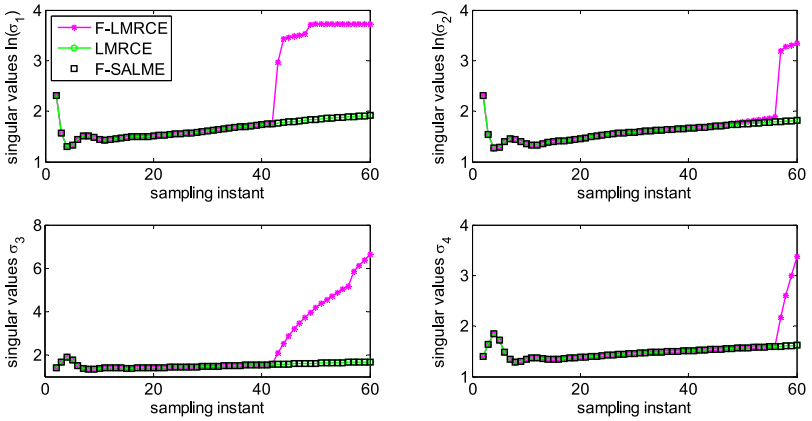

Fig. 5. SVs of compared methods with fixed-point configuration in a single sensor.

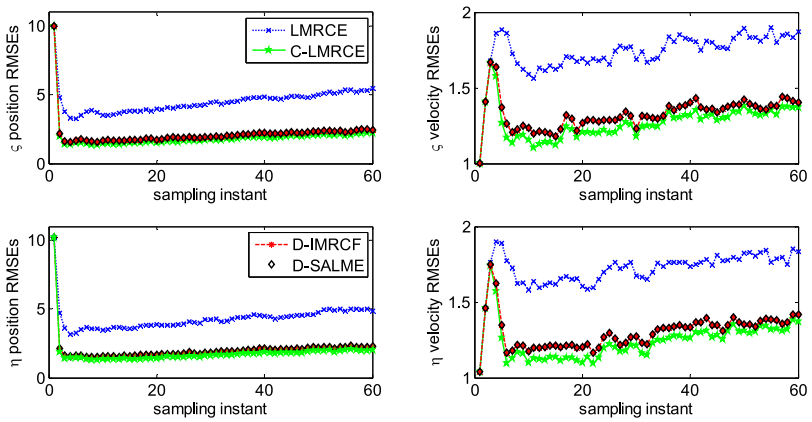

Fig. 6. RMSEs of compared methods with float-point configuration in a sensor network.

5.2.1. Estimate effectiveness analysis

The RMSEs of compared methods based on 1000 Monte Carlo realizations and corresponding SVs of $P_{k|k-1}$ (all from the 8th sensor) are shown in Figs. 6–9.

Under float-point configuration, as shown in Figs. 6 and 7, RMSEs (starting from the 1st epoch) and SVs of the D-IMRCF and D-SALME are both similar to those of the C-LMRCE estimator and all are much smaller than those of the LMRCE estimator, by the fact that more available information from multiple sensors can improve estimate accuracy. This demonstrates that

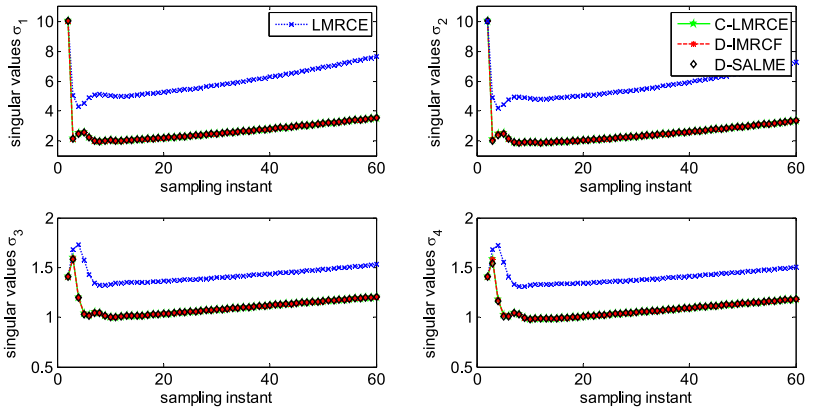

Fig. 7. SVs of compared methods with float-point configuration in a sensor network.

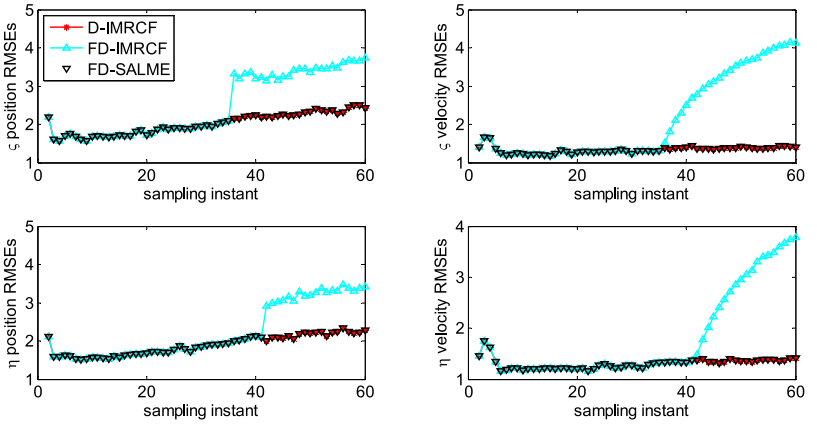

Fig. 8. RMSEs of compared methods with fixed-point configuration in a sensor network.

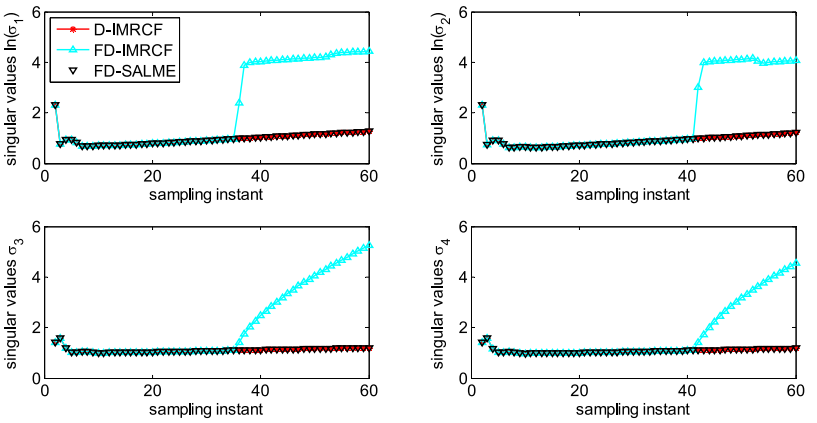

Fig. 9. SVs of compared methods with fixed-point configuration in a sensor network.

the proposed D-IMRCF and D-SALME are effective and have almost the similar accuracy as the centralized implementation. Additionally, in the view of RMSEs and SVs, the accuracy performance of the D-IMRCF is the same as that of the D-SALME.

Moreover, under fixed-point configuration, RMSEs and SVs of the FD-SALME are both as the same as those of the D-IMRCF as shown in Figs. 8 and 9, respectively. However, RMSEs and SVs increase abruptly in the FD-IMRCF due to the existence of roundoff errors when the target position becomes large. This shows that the D-SALME has the advantages of better conditioning and reduced dynamical range.

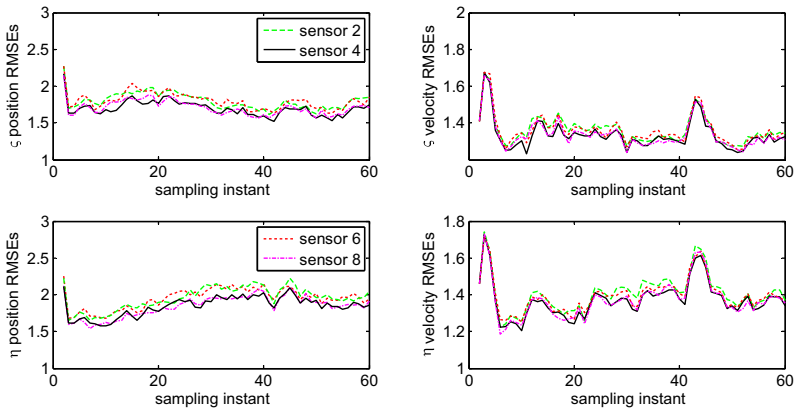

Fig. 10. RMSEs of the FD-SALME from different sensor nodes.

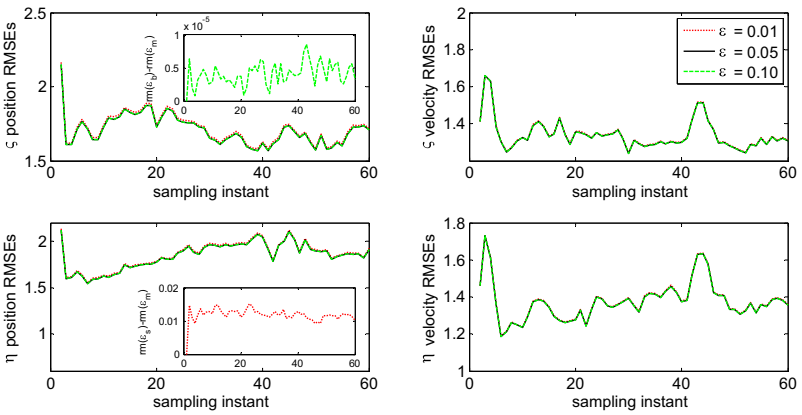

Fig. 11. RMSEs of the FD-SALME with different values of ε.

5.2.2. Parameter robustness analysis

RMSEs of the FD-SALME from the 2nd, 4th, 6th and 8th sensors are shown in Fig. 10, while SVs are the same due to the same second-order statistical properties in all these nodes. As shown in Fig. 10, RMSEs from these four nodes are slightly different. Meanwhile, the results from the 1st, 3th, 5th and 7th nodes are almost coincident with the results shown in Fig. 10. This means that the estimates from different sensor nodes are approximately consistent by using the consensus algorithm in Eq. (44) to realize the distributed fusion.

Moreover, RMSEs of the FD-SALME with different values of ε from the 8th sensor are shown in Fig. 11, where the upper partial enlarged subgraph shows the relative error of RMSEs between $\varepsilon_b = 0.10$ and $\varepsilon_m = 0.05$ and the lower one shows the relative error of RMSEs between $\varepsilon_s = 0.01$ and $\varepsilon_m = 0.05$. In Fig. 11, RMSEs with different values of ε are almost the same, which means that the proposed FD-SALME is robust with different values of ε dropping into the certain range (i.e., $0 < \varepsilon < \frac{2}{1+3\max\{\Xi(n,n),n=1,\cdots,N\}}$).

Furthermore, as shown in Fig. 11, the RMSE of estimated position in ς direction with $\varepsilon_b = 0.10$ is a little bigger than that with $\varepsilon_m = 0.05$, and the RMSE with $\varepsilon_s = 0.01$ is also bigger than that with $\varepsilon_m = 0.05$ in η direction. The reason of this phenomenon is as follows. The iteration step will be increased with the decreasing of ε appropriately when the iteration termination has not reached. However, the decrease of ε will weaken the interaction effect among sensor nodes in each iteration. This leads that the estimate accuracy will not be improved by decreasing ε excessively without increasing $\ell_{\max}$ under a finite iteration step condition. Therefore, when using Eq. (44) to obtain a consistent average value to realize the distributed fusion in practice, since an infinite iteration is impossible, it is required to consider the tradeoff between the value of ε and $\ell_{\max}$.

6. Conclusion

This paper considers the distributed fusion estimation for MRMCNs in sensor networks under the LMMSE criterion. The LMRCE estimator and its square-root array implementation are presented in a centralized structure. Then, an information filter form is proposed, where the final estimate and the corresponding square-root array implementation related to multiple

sensors are only dependent on average values of intermediate vectors and matrices. This allows each sensor node to share these items in the whole sensor network to seek a balance, instead of collecting information from all others directly. Thus, the distributed information filter form with square-root array implementation is proposed via the consensus algorithm to pursue the final consistent estimate.

Acknowledgements

This work has been supported by China National Science Foundation (No. 61135001, 61374023, 61374519) and the Doctorate Foundation of Northwestern Polytechnical University (No. cx201425).

Appendix A. Proof of Theorem 1

Letting $\hat{\alpha}_{k|l}$ be replaced by $\hat{\xi}_{k+1|k+1}$ in Lemma 3.1, $\hat{\xi}_{k+1|k+1} = \sum_{t=1}^{k} \mathcal{X}_{k+1,t}\Psi_t^{-1}\zeta_t + \mathcal{X}_{k+1,k+1}\Psi_{k+1}^{-1}\zeta_{k+1} = \hat{\xi}_{k+1|k} + \mathcal{X}_{k+1,k+1}\Psi_{k+1}^{-1}\zeta_{k+1}$. Since $\tilde{F}_k$ and $\tilde{G}_k$ are uncorrelated with ξ_k, w_k and ζ_t for $t = 1, \cdots, k$, we have $E(\tilde{F}_k\xi_k\zeta_t^T) = O$ and $E(\tilde{G}_k w_k\zeta_t^T) = O$. Thus, $\hat{\xi}_{k+1|k} = \bar{\mathcal{F}}_k\sum_{t=1}^{k}\mathcal{X}_{k,t}\Psi_t^{-1}\zeta_t + \bar{\mathcal{G}}_k\sum_{t=1}^{k}E(w_k\zeta_t^T)\Psi_t^{-1}\zeta_t$, where $\sum_{t=1}^{k}E(w_k\zeta_t^T)\Psi_t^{-1}\zeta_t = \hat{w}_{k|k}$ with $\hat{\alpha}_{k|l}$ being replaced by $\hat{w}_{k|k}$ in Lemma 3.1. Then, we obtain Eq. (10) with $\hat{\xi}_{0|0} = \mathrm{col}\{E(x_0 1_{\{\Theta_0=i\}}), i = 1, \cdots, M\} = \mathrm{col}\{\psi_i, i = 1, \cdots, M\}$. Moreover, the second-order moment matrices of the estimate error and of the estimate are calculated as follows:

$$\Phi_{k+1|k+1} = E(\xi_{k+1}\xi_{k+1}^T) - E(\hat{\xi}_{k+1|k+1}\hat{\xi}_{k+1|k+1}^T) = \Omega_{k+1} - \Pi_{k+1|k+1},$$

$$\Pi_{k+1|k+1} = E(\hat{\xi}_{k+1|k} + \mathcal{X}_{k+1,k+1}\Psi_{k+1}^{-1}\zeta_{k+1})(\cdot)^T = \Pi_{k+1|k} + \mathcal{X}_{k+1,k+1}\Psi_{k+1}^{-1}\mathcal{X}_{k+1,k+1}^T,$$

where $\Pi_{k+1|k} = \bar{\mathcal{F}}_k\Pi_{k|k}\bar{\mathcal{F}}_k^T + \bar{\mathcal{F}}_k\mathrm{T}_{k|k}\bar{\mathcal{G}}_k^T + \bar{\mathcal{G}}_k\Lambda_{k|k}\bar{\mathcal{G}}_k^T + \bar{\mathcal{G}}_k\mathrm{T}_{k|k}^T\bar{\mathcal{F}}_k^T$, with $\Pi_{0|0} = E(\hat{\xi}_{0|0}\hat{\xi}_{0|0}^T) = \hat{\xi}_{0|0}\hat{\xi}_{0|0}^T$.

Since $E(\xi_k w_k^T) = E\big((F_{k-1}\xi_{k-1} + G_{k-1}w_{k-1})w_k^T\big) = O$, we have:

$$\begin{aligned}\Omega_{i,k+1} &= E(\bar{F}_{i\cdot,k}\xi_k\xi_k^T\bar{F}_{i\cdot,k}^T + \bar{F}_{i\cdot,k}\xi_k\xi_k^T\tilde{F}_{i\cdot,k}^T + \bar{G}_{i\cdot,k}w_k w_k^T\bar{G}_{i\cdot,k}^T + \tilde{G}_{i\cdot,k}w_k w_k^T\tilde{G}_{i\cdot,k}^T)\\ &= \sum_{j=1}^{M}p_{ij}\big(\bar{F}_{j,k}\Omega_{j,k}\bar{F}_{j,k}^T + E(\tilde{F}_{j,k}\Omega_{j,k}\tilde{F}_{j,k}^T)\big) + \sum_{j=1}^{M}p_{ij}\pi_{j,k}\big(\bar{G}_{j,k}Q_k\bar{G}_{j,k}^T + E(\tilde{G}_{j,k}Q_k\tilde{G}_{j,k}^T)\big),\end{aligned}$$

where $\bar{F}_{i\cdot,k}$, $\tilde{F}_{i\cdot,k}$, $\bar{G}_{i\cdot,k}$ and $\tilde{G}_{i\cdot,k}$ are the *ith* row entries of $\bar{F}_k$, $\tilde{F}_k$, $\bar{G}_k$ and $\tilde{G}_k$, respectively. For example, $\bar{F}_{i\cdot,k} = [\bar{F}_{1,k}1_{\{\Theta_{k+1}=i|\Theta_k=1\}}, \cdots, \bar{F}_{M,k}1_{\{\Theta_{k+1}=i|\Theta_k=M\}}]$. Assume that $\tilde{A}$ and B are $m \times n$ and $n \times n$ matrices, respectively. The (r, s)*th* sub-block of the general expression $E(\tilde{A}B\tilde{A}^T)$ has the following formula: $E_{(r,s)}(\tilde{A}B\tilde{A}^T) = \sum_{j=1}^{n}\sum_{i=1}^{n}E(\tilde{a}_{ri}\tilde{a}_{sj})b_{ij}$ for $r, s = 1, \cdots, m$, where $\tilde{a}_{ij} := \tilde{A}(i, j)$ and $b_{ij} := B(i, j)$ with b_{ij} being a determined value. Let $\tilde{F}_{j,k}$ be $\tilde{A}$ and $\Omega_{j,k}$ be B, $E(\tilde{F}_{j,k}\Omega_{j,k}\tilde{F}_{j,k}^T)$ is obtained as Eq. (19). Similarly, $E(\tilde{G}_{j,k}Q_k\tilde{G}_{j,k}^T)$ is shown as Eq. (20). Meanwhile, $\Omega_{i,0} = E(x_0x_0^T1_{\{\Theta_0=i\}}) = V_i$ for $i = 1, \cdots, M$.

During the recursion of $\hat{\xi}_{k+1|k+1}$ and $\Pi_{k+1|k+1}$, $\hat{z}_{k+1|k}$, Ψ_{k+1} and $\mathcal{X}_{k+1,k+1}$ are also used. Since v_{k+1} is uncorrelated with ζ_t for $t = 1, \cdots, k$, letting $\hat{\alpha}_{k|l}$ be replaced by $\hat{z}_{k+1|k}$ in Lemma 3.1, we have $\hat{z}_{k+1|k} = \bar{H}_{k+1}\sum_{t=1}^{k}\mathcal{X}_{k+1,t}\Psi_t^{-1}\zeta_t = \bar{H}_{k+1}\hat{\xi}_{k+1|k}$. Substituting it into the definition of the Gram-Schmidt orthogonalized innovation yields Eq. (13). Moreover, $\Psi_{k+1} = E(z_{k+1}z_{k+1}^T) - E(\hat{z}_{k+1|k}\hat{z}_{k+1|k}^T)$ and $\mathcal{X}_{k+1,k+1} = E(\xi_{k+1}z_{k+1}^T) - E(\xi_{k+1}\hat{z}_{k+1|k}^T)$, where $E(z_{k+1}z_{k+1}^T) = \bar{H}_{k+1}\Omega_{k+1}\bar{H}_{k+1}^T + E(\bar{H}_{k+1}\xi_{k+1}\xi_{k+1}^T\tilde{H}_{k+1}^T) + E(\tilde{D}_{k+1}v_{k+1}v_{k+1}^T\tilde{D}_{k+1}^T) + E(\tilde{D}_{k+1}v_{k+1}v_{k+1}^T\tilde{D}_{k+1}^T)$, $E(\xi_{k+1}z_{k+1}^T) = \Omega_{k+1}\bar{H}_{k+1}^T$, $E(\hat{z}_{k+1|k}\hat{z}_{k+1|k}^T) = \bar{H}_{k+1}\Pi_{k+1|k}\bar{H}_{k+1}^T$ and $E(\xi_{k+1}\hat{z}_{k+1|k}^T) = \Pi_{k+1|k}\bar{H}_{k+1}^T$, with

$$E(\tilde{H}_{k+1}\xi_{k+1}\xi_{k+1}^T\tilde{H}_{k+1}^T) = \mathrm{diag}\Big\{\sum_{j=1}^{M}E(\tilde{H}_{n,j,k+1}\Omega_{k+1}\tilde{H}_{n,j,k+1}^T), n = 1, \cdots, N\Big\},$$

$$E(\tilde{D}_{k+1}v_{k+1}v_{k+1}^T\tilde{D}_{k+1}^T) = \mathrm{diag}\Big\{\sum_{j=1}^{M}\pi_{j,k+1}\bar{D}_{n,j,k+1}R_{n,k+1}\bar{D}_{n,j,k+1}^T, n = 1, \cdots, N\Big\},$$

$$E(\tilde{D}_{k+1}v_{k+1}v_{k+1}^T\tilde{D}_{k+1}^T) = \mathrm{diag}\Big\{\sum_{j=1}^{M}\pi_{j,k+1}E(\tilde{D}_{n,j,k+1}R_{n,k+1}\tilde{D}_{n,j,k+1}^T), n = 1, \cdots, N\Big\}.$$

By using $\Phi_{k+1|k} = E(\xi_{k+1}\xi_{k+1}) - E(\hat{\xi}_{k+1|k}\hat{\xi}_{k+1|k}^T) = \Omega_{k+1} - \Pi_{k+1|k}$, we obtain Eqs. (11) and (12). As similar to Eqs. (19) and (20), $E(\tilde{H}_{n,j,k+1}\Omega_{k+1}\tilde{H}_{n,k,k+1}^T)$ and $E(\tilde{D}_{n,j,k+1}R_{n,k+1}\tilde{D}_{n,j,k+1}^T)$ are shown as Eqs. (21) and (22), respectively.

Besides, since w_k is correlated with ζ_k, $\hat{\xi}_{k+1|k}$ and corresponding second-order moment matrices are partially related to the calculations of $\hat{w}_{k|k}$, $\Lambda_{k|k}$ and $\mathrm{T}_{k|k}$. From $E(w_k\zeta_t^T) = S_k\bar{\mathcal{D}}_k^T\delta_{kt}$ for $k \geq t$, by using $\hat{w}_{k|k}$ to replace $\alpha_{k|l}$ in Lemma 3.1, we have $\hat{w}_{k|k} = \sum_{t=1}^{k}E(w_k\zeta_t^T)\Psi_t^T\zeta_t = S_k\bar{\mathcal{D}}_k^T\Psi_k^{-1}\zeta_k$, with $\hat{w}_{0|0} = E(w_0) = \mathbf{0}$. Moreover, $\Lambda_{k|k} = \sum_{t=1}^{k}E(w_k\zeta_t^T)\Psi_t^{-1}\big(E(w_k\zeta_t^T)\big)^T = S_k\bar{\mathcal{D}}_k^T\Psi_k^{-1}\bar{\mathcal{D}}_kS_k^T$ and $\mathrm{T}_{k|k} = E(\hat{\xi}_{k|k-1} + \mathcal{X}_{k,k}\Psi_k^{-1}\zeta_k)(S_k\bar{\mathcal{D}}_k^T\Psi_k^{-1}\zeta_k)^T = \mathcal{X}_{k,k}\Psi_k^{-1}\bar{\mathcal{D}}_kS_k^T$, with $\Lambda_{0|0} = E(\hat{w}_{0|0}\hat{w}_{0|0}^T) = O$ and $\mathrm{T}_{0|0} = E(\hat{\xi}_{0|0}\hat{w}_{0|0}) = O$. □

Appendix B. Proof of Theorem 2

Lemma 1 [31]. *Let A and B be arbitrary $n \times m$ ($n \le m$) matrices. Then, $AA^T = BB^T$ if, and only if, there exists an $m \times m$ unitary matrix U ($UU^T = I = U^TU$) such that $A = BU$.*

Lemma 2 [12,13]. *Let A and B be arbitrary $n \times n$ and $n \times m$ matrices, respectively. Suppose $\mathcal{I} = \mathrm{diag}(\mathcal{I}_1, \mathcal{I}_2)$ with $\mathcal{I}_1$ and $\mathcal{I}_2$ being $n \times n$ and $m \times m$ signature matrices, respectively. Then, $[A\,B]$ can be trangularized by a $\mathcal{U}_{\mathcal{I}}$ as $[A\ B]\mathcal{U}_{\mathcal{I}} = [C\ O]$, with C being a lower (upper) triangular matrix, if and only if, all leading (reversely leading) sub-matrices of $\mathcal{I}_1$ and $A\mathcal{I}_1A^T + B\mathcal{I}_2B^T$ have the same inertia.*

Now, the proof of Theorem 2 is shown as follows.

From Eq. (8), we have $\Omega_{i,k+1} = \left(\begin{bmatrix}\tilde{W}^F_{i,k} & \tilde{W}^F_{i,k} & \tilde{W}^G_{i,k} & \tilde{W}^G_{i,k}\end{bmatrix}\right)(\cdot)^T$. By using Lemma 1, we obtain Eq. (23). Meanwhile, from $\Omega_0^{1/2}\Omega_0^{T/2} = \left(\mathrm{diag}\{V_i^{1/2}, i = 1, \cdots, M\}\right)(\cdot)^T$, we have $\Omega_0^{1/2} = \mathrm{diag}\{V_i^{1/2}, i = 1, \cdots, M\}$.

Substituting $\mathcal{V}_{k+1}$ into Eq. (12), we obtain $\Psi_{k+1} = \tilde{H}_{k+1}\Phi_{k+1|k}\tilde{H}^T_{k+1} + \mathcal{V}_{k+1}$. Moreover, inserting Eqs. (11) and (12) into Eq. (9), we have: $\Pi_{k+1|k+1} = \Pi_{k+1|k} + \Phi_{k+1|k}\Upsilon_{k+1}\Phi_{k+1} - \Phi_{k+1|k}\Upsilon_{k+1}\mathcal{Z}^{-1}_{k+1}\Upsilon_{k+1}\Phi_{k+1|k}$, which is the Schur complement of $\mathcal{Z}_{k+1}$ in $\mathcal{M} := \left(\begin{bmatrix}\Phi^{-T/2}_{k+1|k} & \Upsilon^{1/2}_{k+1} & O\\ O & \Phi_{k+1|k}\Upsilon^{1/2}_{k+1} & \Pi^{1/2}_{k+1|k}\end{bmatrix}\right)(\cdot)^T = \left(\begin{bmatrix}\mathcal{Z}^{1/2}_{k+1} & O & O\\ \Xi_{k+1} & \Pi^{1/2}_{k+1|k+1} & O\end{bmatrix}\right)(\cdot)^T$. Thus, by using Lemma 1, we obtain Eq. (24). From $\Pi^{1/2}_{0|0}\Pi^{T/2}_{0|0} = (\xi_{0|0}\xi^T_{0|0})^{1/2}(\xi_{0|0}\xi^T_{0|0})^{T/2}$, we have $\Pi^{1/2}_{0|0} = (\xi_{0|0}\xi^T_{0|0})^{1/2}$.

Moreover, from Eq. (7), $\Phi_{k+1|k+1} = \left(\begin{bmatrix}\Omega^{1/2}_{k+1} & \Pi^{1/2}_{k+1|k+1}\end{bmatrix}\right)\mathcal{I}_{\Phi_{k+1|k+1}}(\cdot)^T$ with $\mathcal{I}_{\Phi_{k+1|k}} = \mathrm{diag}(I_{n_\xi}, -I_{n_\xi})$. Here, all leading sub-matrices of $\Phi_{k+1|k+1}$ and I_{n_ξ} have the same inertia since they are positive-definite. According to Lemma 2, we obtain Eq. (25). Similarly, $\Phi_{k+1|k} = \left(\begin{bmatrix}\Omega^{1/2}_{k+1} & \Pi^{1/2}_{k+1|k}\end{bmatrix}\right)\mathcal{I}_{\Phi_{k+1|k}}(\cdot)^T$ and we obtain Eq. (29). From Eq. (14), we have

$$\Pi_{k+1|k} = \begin{bmatrix}\bar{\mathcal{F}}_k & \bar{\mathcal{G}}_k\end{bmatrix}\begin{bmatrix}\Pi^{1/2}_{k|k} & O\\ \mathcal{A}_k & \mathcal{B}_k\end{bmatrix}\begin{bmatrix}\Pi^{1/2}_{k|k} & O\\ \mathcal{A}_k & \mathcal{B}_k\end{bmatrix}^T\begin{bmatrix}\bar{\mathcal{F}}^T_k\\ \bar{\mathcal{G}}^T_k\end{bmatrix},$$

where $\mathcal{A}_k = \mathrm{T}^T_{k|k}\Pi^{-T/2}_{k|k}$ and $\mathcal{B}_k\mathcal{B}^T_k = \Lambda_{k|k} - \mathcal{A}_k\mathcal{A}^T_k$. By using Lemma 1, we obtain Eq. (30). Since that $\mathcal{B}_k\mathcal{B}^T_k$ is the Schur complement of $\Pi_{k|k}$ in $E\left((\hat{\xi}^T_{k|k}, w^T_{k|k})^T(\cdot)\right) = \begin{bmatrix}\Pi_{k|k} & \mathrm{T}_{k|k}\\ \mathrm{T}^T_{k|k} & \Lambda_{k|k}\end{bmatrix}$ which is positive-definite (without considering singularity), we know that $\mathcal{B}_k\mathcal{B}^T_k$ is positive-definite. Meanwhile, $\mathcal{B}_k\mathcal{B}^T_k = \left(\begin{bmatrix}\Lambda^{1/2}_{k|k} & \mathrm{T}^T_{k|k}\Pi^{-T/2}_{k|k}\end{bmatrix}\right)\mathcal{I}_{\mathcal{B}_k}(\cdot)^T$, where $\mathcal{I}_{\mathcal{B}_k} = \mathrm{diag}(I_{n_w}, -I_{n_\xi})$. Since $\mathcal{B}_k\mathcal{B}^T_k$ and I_{n_w} are positive-definite, their leading sub-matrices have the same inertia. By using Lemma 2, there exists a $\mathcal{I}_{\mathcal{B}_k}$-unitary matrix $\mathcal{U}_{\mathcal{B}_k}$ to obtain Eq. (31).

Furthermore, inserting Eq. (12) into Eq. (17), we have $\Lambda_{k|k} = S_k\bar{\mathcal{D}}^T_k(\bar{H}_k\Phi_{k|k-1}\bar{H}^T_k + \mathcal{V}_k)^{-1}\bar{\mathcal{D}}_kS^T_k = \Delta_k - \Upsilon^T_k\mathcal{Z}^{-1}_{k+1}\Upsilon_k$, which is the Schur complement of $\mathcal{Z}_k$ in $\mathcal{M}' := \left(\begin{bmatrix}\mathcal{Z}^{1/2}_k & O & O\\ \Gamma^T_k\mathcal{Z}^{-1/2}_k & \Lambda^{1/2}_{k|k} & O\end{bmatrix}\right)(\cdot)^T = \left(\begin{bmatrix}\Phi^{-T/2}_{k|k-1} & \bar{H}^T_k\mathcal{V}^{-T/2}_k & O\\ O & S_k\bar{\mathcal{D}}^T_k\mathcal{V}^{-1/2}_k & O\end{bmatrix}\right)(\cdot)^T$. By using Lemma 1, we obtain Eq. (32). From Corollary 1, $\mathcal{V}_{k+1} = \mathrm{diag}\{\mathcal{V}^{1/2}_{n,k+1}\mathcal{V}^{T/2}_{n,k+1}, n = 1, \cdots, N\}$ with $\mathcal{V}_{n,k+1} = \left(\begin{bmatrix}\tilde{\mathcal{V}}^H_{n,k+1} & \tilde{\mathcal{V}}^D_{n,k+1} & \tilde{\mathcal{V}}^D_{n,k+1}\end{bmatrix}\right)(\cdot)^T$. By using Lemma 1, we obtain Eq. (34). On the other hand, $\Lambda^{1/2}_k\Lambda^{T/2}_k = \left(\begin{bmatrix}\Delta^{1/2}_k & \Gamma^T_kZ^{-T/2}_k\end{bmatrix}\right)\mathcal{I}_{\Lambda_{k|k}}(\cdot)^T$, where $\mathcal{I}_{\Lambda_{k|k}} = \mathrm{diag}(I_{n_w}, -I_{n_\xi})$. Since $\mathcal{V}_k$ is positive-semidefinite, $\mathcal{V}^{-1}_k$ is also positive-semidefinite, i.e., $\varphi^T\mathcal{V}^{-1}_k\varphi \ge 0$ with φ being any a column vector with proper dimension. Since $\bar{\mathcal{D}}_kS^T_k\varphi'$ is a column vector with the same dimension of φ for any a column vector φ' with appropriate dimension, then $(\bar{\mathcal{D}}_kS^T_k\varphi')^T\mathcal{V}^{-1}_k(\bar{\mathcal{D}}_kS^T_k\varphi') \ge 0$, i.e., $(\varphi')^TS_k\bar{\mathcal{D}}^T_k\mathcal{V}^{-1}_k\bar{\mathcal{D}}_kS^T_k\varphi' \ge 0$. Thus, Δ_k is positive-semidefinite. Without considering the singularity, all leading sub-matrices of Δ_k and I_{n_w} have the same inertia. According to Lemma 2, there exists a $\mathcal{I}_{\Lambda_{k|k}}$-unitary matrix to obtain Eq. (33).

Finally, $\mathcal{Z}^{1/2}_k\mathcal{Z}^{T/2}_k = \Phi^{-1}_{k|k-1} + \Upsilon_k = \left(\begin{bmatrix}\Phi^{-T/2}_{k|k-1} & \Upsilon^{1/2}_k\end{bmatrix}\right)(\cdot)^T$. According to Lemma 1, we obtain Eq. (27). □

Appendix C. Proof of Theorem 3

From the definition of $\hat{\vartheta}_{k|l}$ and $\mathcal{J}_{k|l}$, we have

$$\hat{\vartheta}_{k+1|k+1} = (\Phi^{-1}_{k+1|k} + \bar{H}^T_{k+1}\mathcal{V}^{-1}_{k+1}\bar{H}_{k+1})(\hat{\xi}_{k+1|k} + \mathcal{X}_{k+1,k+1}\Psi^{-1}_{k+1}\zeta_{k+1}) = \hat{\vartheta}_{k+1|k} + N\bar{\beta}_{k+1},$$
$$\mathcal{J}_{k+1|k+1} = \Phi^{-1}_{k+1|k} + \Phi^{-1}_{k+1|k}\mathcal{X}_{k+1,k+1}\mathcal{V}^{-1}_{k+1}\mathcal{X}^T_{k+1,k+1}\Phi^{-1}_{k+1|k} = \mathcal{J}_{k+1|k} + N\bar{\Upsilon}_{k+1},$$

where $\hat{\vartheta}_{k+1|k}$ and $\mathcal{J}_{k+1|k}$ are calculated as Eqs. (39) and (40), respectively. Moreover, replace $\Phi_{k+1|k+1}$ by $\mathcal{J}^{-1}_{k+1|k+1}$ in Eq. (7) in Theorem 1, we obtain Eq. (38). Besides, inserting $\Psi^{-1}_k = \mathcal{V}^{-1}_k - \mathcal{V}^{-1}_k\bar{H}_k(\mathcal{J}_{k|k-1} + N\bar{\Upsilon}_k)^{-1}\bar{H}^T_k\mathcal{V}^{-1}_k$ into Eqs. (16), (17) and (18), we obtain Eqs. (41), (42) and (43), respectively. □

Appendix D. Proof of Theorem 4

Denote $\mathcal{P} := (I + \varepsilon\mathcal{A} - 2\varepsilon\Xi - \varepsilon I)$ for brevity. Since $\mathcal{A}$ is real symmetric, $\mathcal{P}$ is a real symmetric matrix and its eigenvalues are all real numbers. From the Gervshgorin circle theorem [32], the eigenvalue λ_n of $\mathcal{P}$ falls into the interval $|\lambda_n - (1 - 2\varepsilon\Xi(n,n) - \varepsilon)| \le |\varepsilon\Xi(n,n)|$ for $n = 1, \cdots, N$ with "$|\cdot|$" representing the corresponding absolute value. Since $\Xi(n, n) \ge 0$, we have $-\varepsilon\Xi(n,n) \le \lambda_n - (1 - 2\varepsilon\Xi(n,n) - \varepsilon) \le \varepsilon\Xi(n,n)$, i.e., $1 - 3\varepsilon\Xi(n,n) - \varepsilon \le \lambda_n \le 1 - \varepsilon\Xi(n,n) - \varepsilon$. From $|\lambda_n| < 1$ for $n = 1, \cdots, N$, we have $\begin{cases} 1 - 3\varepsilon\Xi(n,n) - \varepsilon > -1 \\ 1 - \varepsilon\Xi(n,n) - \varepsilon < 1 \end{cases}$. Thus, $0 < \varepsilon < \frac{2}{3\max\{\Xi(n,n), n=1,\cdots,N\}+1}$.

Write Eq. (44) as a column vector for $n = 1, \cdots, N$, we have $\mathcal{X}_k^{i+1} = \mathcal{X}_k^i + \varepsilon\big((\mathcal{A} - \Xi)\mathcal{X}_k^i + (\mathcal{A} - \Xi)\Im_k\big) + \varepsilon\big((I + \Xi)\Im_k - (I + \Xi)\mathcal{X}_k^i\big)$, i.e., $\mathcal{X}_k^{i+1} = \mathcal{P}\mathcal{X}_k^i + (\varepsilon\mathcal{A} + \varepsilon I)\Im_k$. In this way, $\mathcal{X}_k^{i+1} = \mathcal{P}^i\mathcal{X}_k^1 + (\mathcal{P}^{i-1} + \cdots + \mathcal{P} + I)(\varepsilon\mathcal{A} + \varepsilon I)\Im_k$. Since all eigenvalues of $\mathcal{P}$ are located strictly inside the unit circle in the complex plane, we have that $\lim_{i\to\infty}\mathcal{P}^i = O$ and $\lim_{i\to\infty}\sum_{j=0}^{i}\mathcal{P}^j = (I - \mathcal{P})^{-1}$ with $\mathcal{P}^0 = I$. Therefore, $\lim_{i\to\infty}\mathcal{X}_k^i = (I - \mathcal{P})^{-1}(\varepsilon\mathcal{A} + \varepsilon I)\Im_k = (I + 2\Xi - \mathcal{A})^{-1}(\mathcal{A} + I)\Im_k$.

Moreover, the above proof still holds if $\mathcal{A}$ is placed by $\mathbb{A}$ and $\Xi = \mathrm{diag}\{\mathbb{A}\cdot\mathbf{1}_N\}$. Then, if $(I + 2\Xi - \mathbb{A})^{-1}(\mathbb{A} + I) = \frac{1}{N}\mathbf{1}_N \otimes \mathbf{1}_N^T$, we obtain that each element of the convergent vector of $\mathcal{X}_k^i$ is $\frac{1}{N}\sum_{n=1}^{N}\Im_{n,k}$, i.e., the equilibrium. This means that the following equation holds: $(I - \frac{1}{N}\mathbf{1}_N \otimes \mathbf{1}_N^T)\mathbb{A} = \frac{1}{N}\mathbf{1}_N \otimes \mathbf{1}_N^T - I$, since $\mathbb{A}$ is symmetric. Here, $I - \frac{1}{N}\mathbf{1}_N \otimes \mathbf{1}_N^T$ is singular. By considering $\mathbb{A}(i,j) = 0$ if $\mathcal{A}(i,j) = 0$ for $i, j = 1, \cdots, N$, the above matrix equation is equivalent to the following one: $\begin{cases} -\frac{1}{N}\sum_{n=1,n\neq i}^{N} a_{nj} = \frac{1}{N} - 1, & i = j \\ -\frac{1}{N}\sum_{n=1,n\neq i}^{N} a_{nj} = \frac{1}{N}, & i > j \end{cases}$ for $i = 1, \cdots, N$ and $j = 1, \cdots, i-1$, which can be shown as Eq. (45). And its least squares solution about unknown weights to construct $\mathbb{A}$ is $\mu_{LS} = (C^TC)^+C^Tb$. □

References

[1] G. Chang, Marginal unscented Kalman filter for cross-correlated process and observation noise at the same epoch, IET Radar Sonar Navig. 8 (1) (2014) 54–64, doi:10.1049/iet-rsn.2013.0095.
[2] B. Chen, G. Hu, W. Zhang, L. Yu, Distributed mixed H_2/H_∞ fusion estimation with limited communication capacity, IEEE Trans. Autom. Control 61 (3) (2016) 805–810, doi:10.1109/TAC.2015.2450271.
[3] B. Chen, W. Zhang, L. Yu, G. Hu, H. Song, Distributed fusion estimation with communication bandwidth constraints, IEEE Trans. Autom. Control 60 (5) (2015) 1398–1403, doi:10.1109/TAC.2014.2357113.
[4] O.L.V. Costa, Linear minimum mean square error estimation for discrete-time Markovian jump linear system, IEEE Trans. Autom. Control 39 (8) (1994) 1685–1689, doi:10.1109/9.310052.
[5] O.L.V. Costa, Stationary filter for linear minimum mean square error estimation for discrete-time Markovian jump linear systems, IEEE Trans. Autom. Control 47 (8) (2002) 1351–1356, doi:10.1109/TAC.2002.800745.
[6] O.L.V. Costa, G.R.A.M. Benites, Linear minimum mean square filter for discrete-time linear systems with Markov jumps and multiplicative noises, Automatica 47 (2011) 466–476, doi:10.1016/j.automatica.2011.01.015.
[7] O.L.V. Costa, G.R.A.M. Benites, Robust mode-independent filtering for discrete-time Markov jump linear systems with multiplcative noises, Int. J. Control 86 (5) (2013) 779–793, doi:10.1080/00207179.2012.760047.
[8] O.L.V. Costa, A. de Oliveira, Optimal mean-variance control for discrete-time linear systems with Markovian jumps and multiplicative noises, Automatica 48 (2012) 304–315, doi:10.1016/j.automatica.2011.11.009.
[9] Y. Gao, X. Li, Z. Duan, Estimation fusion for Markovian jump linear system via data transformation, IEEE Trans. Aerosp. Electron. Syst. 50 (1) (2014) 240–253, doi:10.1109/TAES.2013.110740.
[10] M.J. García-Ligero, A. Hermoso-Carazo, J. Linares-Pérez, Distributed and centralized fusion estimation from multiple sensors with Markovian delays, Appl. Math. Comput. 219 (2012) 2932–2948, doi:10.1016/j.amc.2012.09.017.
[11] D. Gu, A game theory approach to target tracking in sensor networks, IEEE Trans. Syst. Man Cybern. Part B 41 (1) (2011) 2–13, doi:10.1109/TSMCB.2010.2040733.
[12] B. Hassibi, T. Kailath, A.H. Sayed, Array algorithms for H_∞ estimation, IEEE Trans. Autom. Control 45 (4) (2000) 702–706, doi:10.1109/9.847105.
[13] B. Hassibi, A.H. Sayed, T. Kailath, Indefinite-Quadratic Estimation and Control: A Unified Approach to H_2 and H_∞ Theories, Philadelphia, PA, USA: SIAM, 1999, doi:10.1137/1.9781611970760.
[14] M.J.R. Healy, Algorithm AS 6: triangular decomposition of a symmetric matrix, J. Roy. Stat. Soc. Ser. C 17 (2) (1968) 195–197, doi:10.2307/2985687.
[15] S. Hu, D. Yue, X. Xie, Z. Du, Event-triggered H_∞ stabilization for networked stochastic systems with multiplicative noise and network-induced delays, Inf. Sci. 299 (2015) 178–197, doi:10.1016/j.ins.2014.11.048.
[16] Y. Huang, Y. Zhang, X. Wang, L. Zhao, Gaussian filter for nonlinear systems with correlated noises at the same epoch, Automatica 60 (2015) 122–126, doi:10.1016/j.automatica.2015.06.035.
[17] Y.G. Kao, C.H. Wang, J. Xie, H.R. Karimi, W. Li, H_∞ sliding mode control for uncertain neutral-type stochastic systems with Markovian jumping parameters, Inf. Sci. 314 (2015) 200–211, doi:10.1016/j.ins.2015.03.047.
[18] W. Li, Y. Jia, Consensus-based distributed multiple model UKF for jump Markov nonlinear systems, IEEE Trans. Autom. Control 57 (1) (2012) 230–236, doi:10.1109/TAC.2011.2161838.
[19] X. Li, V.P. Jilkov, Survey of maneuvering target tracking. part v: multiple-model methods, IEEE Trans. Aerosp. Electron. Syst. 41 (4) (2005) 1255–1321, doi:10.1109/TAES.2005.1561886.
[20] F. Li, L. Wu, P. Shi, C. Lim, State estimation and sliding mode control for semi-Markovian jump systems with mismatched uncertainties, Automatica 51 (2015) 385–393, doi:10.1016/j.automatica.2014.10.065.
[21] J. Linares-Pérez, R. Caballero-Águila, I. García-Garrido, Optimal linear filter design for systems with correlation in the measurement matrices and noises: recursive algorithm and applications, Int. J. Syst. Sci. 45 (7) (2014) 1548–1562, doi:10.1080/00207721.2014.909093.
[22] Y. Luo, Y. Zhu, X. Shen, E. Song, Novel data association algorithm based on integrated random coefficient matrices Kalman filtering, IEEE Trans. Aerosp. Electron. Syst. 48 (1) (2012) 144–158, doi:10.1109/TAES.2012.6129626.
[23] I. Matei, J.S. Baras, Optimal state estimation for discrete-time Markovian jump linear systems, in the presence of delayed output observations, IEEE Trans. Autom. Control 56 (9) (2011) 2235–2240, doi:10.1109/TAC.2011.2160027.
[24] R. Olfati-Saber, J.A. Fax, R.M. Murray, Consensus and cooperation in networked multi-angent systems, Proc. IEEE 95 (1) (2007) 215–233, doi:10.1109/JPROC.2006.887293.
[25] R. Olfati-Saber, J.S. Shamma, Consensus filters for sensor networks and distributed sensor fusion, in: Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005, Seville, Spain, 2005, pp. 6698–6703, doi:10.1109/CDC.2005.1583238.

[26] H. Rezaei, R.M. Esfanjani, M.H. Sedaaghi, Improved robust finite-horizon Kalman filtering for uncertain networked time-varying systems, Inf. Sci. 293 (2015) 263–274, doi:10.1016/j.ins.2014.09.017.
[27] A.M. Stoica, V. Dragan, I. Yaesh, Kalman-type filtering for stochastic systems with state-dependent noise and Markovian jumps, in: Proceedings of the 15th IFAC Symposium on System Identification, Saint-Maio, France, 2009, pp. 1375–1380, doi:10.3182/20090706-3-FR-2004.00229.
[28] Q. Sun, C. Lim, P. Shi, F. Liu, Moving horizon estimation for Markov jump systems, Inf. Sci. 367–368 (2016) 143–158, doi:10.1016/j.ins.2016.05.028.
[29] D. Svensson, L. Svensson, A new multiple model filter with switch time conditions, IEEE Trans. Signal Process. 58 (1) (2010) 11–25, doi:10.1109/TSP.2009.2029720.
[30] M.H. Terra, J.Y. Ishihara, G. Jesus, Information filtering and array algorithms for discrete-time Markovian jump linear systems, IEEE Trans. Autom. Control 54 (1) (2009) 158–162, doi:10.1109/TAC.2008.2007181.
[31] M.H. Terra, J.Y. Ishihara, A.P. Junior, Array algorithm for filtering of discrete-time Markovian jump linear systems, IEEE Trans. Autom. Control 52 (7) (2007) 1293–1296, doi:10.1109/TAC.2007.900833.
[32] R.S. Varga, Geršgorin and His Circles, Berlin, Germany: Springer-Verlag, 2004, doi:10.1007/978-3-642-17798-9.
[33] X. Wang, Y. Liang, Q. Pan, F. Yang, A Gaussian approximation recursive filter for nonlinear systems with correlated noises, Automatica 48 (2012) 2290–2297, doi:10.1016/j.automatica.2012.06.035.
[34] Y. Wei, J. Qiu, H.R. Karimi, M. Wang, Filtering design for two-dimensional Markovian jump systems with state-delays and efficient mode information, Inf. Sci. 269 (2014) 316–331, doi:10.1016/j.ins.2013.12.042.
[35] L. Wu, X. Yao, W.X. Zheng, Generalized H_2 fault detection for two-dimensional Markovian jump systems, Automatica 48 (2012) 1741–1750, doi:10.1016/j.automatica.2012.05.024.
[36] Y. Yang, Y. Liang, Q. Pan, Y. Qin, X. Wang, Gaussian-consensus filter for nonlinear systems with randomly delayed measurements in sensor networks, Inf. Fusion 30 (2016) 91–102, doi:10.1016/j.inffus.2015.12.003.
[37] Y. Yang, Y. Liang, Q. Pan, Y. Qin, F. Yang, Linear minimum-mean-square error estimation of Markovian jump linear systems with stochastic coefficient matrices, IET Control Theory Appl. 8 (12) (2014a) 1112–1126, doi:10.1049/iet-cta.2013.0936.
[38] Y. Yang, Y. Liang, F. Yang, Y. Qin, Q. Pan, Linear minimum-mean-square error estimation of Markovian jump linear systems with randomly delayed measurements, IET Signal Process. 8 (6) (2014b) 658–667, doi:10.1049/iet-spr.2013.0431.
[39] Z. Yang, X. Shi, J. Chen, Optimal coordination of mobile sensors for target tracking under additive and multiplicative noises, IEEE Trans. Ind. Electron. 61 (7) (2014c) 3459–3468, doi:10.1109/TIE.2013.2281157.
[40] W. Zhang, B. Chen, M.Z.Q. Chen, Hierarchical fusion estimation for clustered asynchronous sensor networks, IEEE Trans. Autom. Control 6 (1) (2015) 1–6, doi:10.1109/TAC.2015.2498701. (Available online).