Distributed Multiple-Model Fusion with Transformed Measurements

Yongxin Gao
School of Electronic
and Information Engineering
Xi'an Jiaotong University
Xi'an, Shaanxi 710049, P.R. China
yongxingao@gmail.com

X. Rong Li
Department of Electrical Engineering
University of New Orleans
New Orleans, LA, U.S.A. 70148
xli@uno.edu

Abstract— **This paper deals with estimation fusion for a Markovian jump-linear system (MJLS) and proposes a distributed fusion scheme, in which local sensors send their transformed measurements to the fusion center and the fusion center fuses them with a multiple-model (MM) filter. A specific linear transformation for local measurements is studied and it is shown that the distributed minimum mean-squared error (MMSE) fusion with these transformed measurements has the same performance as the centralized MMSE fusion when full-rate communication is employed. The reduced-rate communication case is also considered for the systems with very limited communication capacity. Moreover, approximate algorithms are presented for practical application. Illustrative numerical results are provided to show the performance of the fusion methods.**

Keywords: Estimation fusion, distributed fusion, multiple-model estimation, minimum mean-squared error.

I. Introduction

Estimation fusion, which is the problem of best utilizing useful information contained in multiple sets of data for the purpose of estimating a quantity [1], has been investigated for decades. There are a lot of results available (see, e.g., [2], [3], [4], [5] and the series of papers [1], [6], [7], [8], [9], [10], [11]) and almost all of them focus on the fusion for a single model system. However, for many applications, especially for target tracking, the underlying system (e.g., a maneuvering target) encounters motion-mode uncertainty, which is usually modeled by a set of motion models. The single model fusion method is not directly applicable for these situations, although it can provide a basis (see, e.g., [12], [1]). In this paper, we consider the estimation fusion for an MJLS, which is a system with Markovian jumping parameters represented by multiple models with given probabilities of jumping between the models.

Consider the following two-sensor (sensor a and b) distributed MJLS:

$$x_{k+1} = F_k^{(l)} x_k + G_k^{(l)} w_k^{(l)} \quad (1)$$

$$z_k^n = H_k^n x_k + v_k^n, \; n = a, b \quad (2)$$

Research supported in part by ARO through grant W911NF-08-1-0409, ONR-DEPSCoR through grant N00014-09-1-1169, and NAVO through Contract N62306-09-P-3S01.

where $x_k \in \mathbb{R}^{n_x}$, $z_k^n \in \mathbb{R}^{n_z}$, superscript l denotes quantities pertinent to model $m^{(l)}$ in model space $\mathbb{M}$. The event that model $m^{(l)}$ is in effect over the sampling period $(t_{k-1}, t_k]$ will be denoted by $m_k^{(l)}$. The switching between the models is governed by a homogeneous finite-state Markovian chain with the transition probabilities $\pi_{ij} = P(m_k^{(j)}|m_{k-1}^{(i)})$. It is assumed that the process noise $w^{(l)}$ and the measurement noises v^n $(n = a, b)$ are mutually uncorrelated zero-mean white Gaussian processes with covariance matrices

$$\text{cov}(w_k^{(l)}) = Q_k^{(l)}, \; \text{cov}(v_k^n) = R_k^n$$

It is further assumed that the initial state x_0 is a Gaussian mixture and uncorrelated with $w^{(l)}$ and v^n.

Note that the MJLS assumed above is not completely general since the measurement of each local sensor has a single model. However, normally in practice (e.g., target tracking) the way in which local sensors make their measurements is not uncertain.

The problem is how to estimate the system state based on the (processed or unprocessed) data sent by local sensors, i.e., fusing the distributed data. The fusion can obviously employ a centralized architecture, in which all sensors send their raw measurements to the fusion center and then the fusion center estimates the system state with a multiple-model filter. This is referred to as centralized multiple-model fusion (CMMF). CMMF can achieve the best performance, but it needs heavy computation and data transmission and thus is impractical for some cases. In this paper, we study the distributed fusion and use CMMF as a benchmark.

We consider the following distributed fusion schemes depending on the processing methods at the local sensors and the fusion center:

1) Local sensors use linear minimum mean-squared error (LMMSE) filters for the MJLS [12] to estimate the system state and transmit the local estimates to the fusion center. Then a global estimate can be obtained at the fusion center by the optimal linear fusion rule [1]. This scheme is feasible but has some disadvantages. Denote by $M = |\mathbb{M}|$ the number of system models, and the LMMSE filter for MJLS would

introduce a $(M * n_x)$-dimensional stacked state vector, so the transmission of local estimates would overload the communication. What is more, for an MJLS, the optimal linear fusion performs in general not as well as some well-known suboptimal nonlinear MM fusion.

2) Local sensors use multiple-model filters and the fusion center employs a linear combination rule. This scheme has some challenging problems. One of the major issues is how to deal with the cross-correlation of local estimation errors due to the common process noise since the local multiple-model filters are in essence nonlinear filters. Optimal dealing with the cross-correlation is harder than in the LMMSE case above. In [13], [14], interacting multiple model (IMM) filters are used in local senors, information increments are sent to the fusion center and then the method of [2] is used to get a global state estimate at the fusion center, which is an approximate method and does not enjoy optimality.

A common ground for the above two schemes is that they both use a linear fuser at the fusion center. In this paper, we adopt a multiple-model estimator, which is nonlinear, to fuse the transformed data sent by local sensors. The IMM algorithm is utilized to implement the fuser, since the optimal approach involves an exponentially increasing number of model-sequences. In addition, both the full-rate and reduced-rate communication cases are considered, making it flexible for application with any communication constraint.

The paper is organized as follows. Section II formulates the multi-sensor multiple-model distributed fusion problem for the system (1)–(2). Section III presents the algorithm of distributed multiple-model fusion in the full-rate communication case. For reduced-rate communication, an approximate distributed multiple-model fusion method is presented in Section IV. Numerical examples are provided in Section V to compare the performance of our fusion methods with the CMMF. Section VI concludes the paper.

II. Background and Problem Formulation

Single-model distributed fusion with transformed data has been studied, especially for linear transformations (see, e.g., [15], [16]). Theoretically, in the multi-sensor environment, fusion can always be viewed as the estimation with transformed data: the centralized fusion is with an identical transformation, while the standard distributed estimation fusion[1] is with a nonidentical transformation.

Thus the adopted transformation should have some nice properties for the fusion, such as reducing the computational load and data transmission. We will use the transformation in [16] and illustrate its advantages.

Let

$$y_k^n = (H_k^n)'(R_k^n)^{-1} z_k^n$$

Then from Eq. (2), it follows that

$$y_k^n = (H_k^n)'(R_k^n)^{-1} H_k^n x_k + (H_k^n)'(R_k^n)^{-1} v_k^n \qquad (3)$$

[1]When only local estimates are available at the fusion center, the fusion is refer to as standard distributed estimation fusion [1].

Furthermore, let

$$\bar{H}_k^n = (H_k^n)'(R_k^n)^{-1} H_k^n$$
$$\bar{v}_k^n = (H_k^n)'(R_k^n)^{-1} v_k^n$$

Then Eq. (3) can be rewritten as

$$y_k^n = \bar{H}_k^n x_k + \bar{v}_k^n, \ n = a, b$$

where

$$E[\bar{v}_k^n] = 0, \ \text{cov}(\bar{v}_k^n) = \bar{H}_k^n$$

and $\bar{v}_k^a$ and $\bar{v}_k^b$ are also uncorrelated.

In our proposed distributed fusion, each local sensor sends the processed data y_k^n to the fusion center at every sampling step and the fusion is to find a state estimate of x_k, i.e.,

$$\hat{x}_k = E[x_k | Y^k] \qquad (4)$$

where $Y^k = \{Y_\kappa, \kappa = 1, 2, \cdots, k\}$ and $Y_\kappa = [(y_\kappa^a)', (y_\kappa^b)']$.

III. Distributed multiple-model Fusion with Transformed Measurements

A. Model Likelihood

A major issue in the multiple-model fusion with transformed measurements is to compute the posterior probability of the system model, which in essence is to obtain the model likelihood. Here the problem is that H_k^n does not necessarily have full column-rank. Thus, y_k^n might have a singular Gaussian distribution (i.e., one with a singular covariance) given x_k and the likelihood is not known analytically. We use the pseudo-inverse and give the likelihood function for transformed measurements in a static case as follows. Let

$$f(x) = \sum_l \mu_l \mathcal{N}(x; \bar{x}_l, C_{x_l})$$
$$z = Hx + v$$

where μ_l is the probability of model l, $\mathcal{N}(x; \bar{x}_l, C_{x_l})$ is the distribution of x when model l is in effect, $v \sim \mathcal{N}(0, R)$ and R is invertible. Denote by y the transformed data

$$y = H' R^{-1} z$$

and then the likelihood function of model l is

$$L(y | \bar{x}_l, C_{x_l}) = (\det(R + H C_{x_l} H'))^{-\frac{1}{2}} \exp(-\frac{1}{2} \tilde{y}' C_y^+ \tilde{y}) \quad (5)$$

where $\tilde{y} = y - H'R^{-1}H\bar{x}_l$, $C_y = H'R^{-1}[HC_{x_l}H' + R]R^{-1}H$ and the superscript '+' stands for the Moore-Penrose pseudo-inverse (MP inverse in short). The expression is quite similar to the Gaussian probability distribution function (pdf) except for the MP inverse in the quadratic form.

Here we provide the model likelihood of transformed measurements directly and later in Theorem 1 we will show that Eq. (5) is indeed the sought-after likelihood function by proving that the same posterior model probabilities are obtained by using transformed measurements and by using raw measurements.

B. Multiple-Model Fusion with Transformed Measurements

Denote $m_{(l^k)}$ as the event that the model sequence l^k through time k is in effect. Assume that the distributed fusion results at time $k-1$ are $\hat{x}_{k-1}^{(l^{k-1})}$, $P_{k-1}^{(l^{k-1})}$ and $\mu_{k-1}^{(l^{k-1})}$, where $P_{k-1}^{(l^{k-1})}$ is the mean-squared error (MSE) matrix of $\hat{x}_{k-1}^{(l^{k-1})}$ and $\mu_{k-1}^{(l^{k-1})}$ is the posterior probability of l^{k-1}. These quantities propagate to time k to be viewed as the prior information of x_k, which is denoted by $\bar{X}_k = \{\hat{x}_{k|k-1}^{(l^k)}, P_{k|k-1}^{(l^k)}, \mu_{k|k-1}^{(l^k)}\}$. Then fusion (4) can be expressed as

$$\begin{aligned}\hat{x}_k &\approx E[x_k|Y_k, \bar{X}_k] \\ &= \sum_{l^k} P(m_{(l^k)}|Y_k, \bar{X}_k) E[x_k|Y_k, \bar{X}_k, m_{(l^k)}]\end{aligned} \tag{6}$$

where

$$\begin{aligned}&P(m_{(l^k)}|Y_k, \bar{X}_k) \\ &= \frac{L(Y_k|\hat{x}_{k|k-1}^{(l^k)}, P_{k|k-1}^{(l^k)}, m_{(l^k)})\mu_{k|k-1}^{(l^k)}}{\sum_{l^k} L(Y_k|\hat{x}_{k|k-1}^{(l^k)}, P_{k|k-1}^{(l^k)}, m_{(l^k)})\mu_{k|k-1}^{(l^k)}}\end{aligned}$$

the model-sequence conditioned estimation is

$$\begin{aligned}\hat{x}_k^{(l^k)} &= E[x_k|Y_k, \bar{X}_k, m_{(l^k)}] \\ &= \hat{x}_{k|k-1}^{(l^k)} + K_k(Y_k - \bar{H}_k \hat{x}_{k|k-1}^{(l^k)}) \\ K_k &= P_{k|k-1}^{(l^k)} \bar{H}_k' S_k^+ \\ P_k^{(l^k)} &= P_{k|k-1}^{(l^k)} - K_k S_k K_k' \\ S_k &= \bar{H}_k P_{k|k-1}^{(l^k)} \bar{H}_k' + \bar{R}_k\end{aligned}$$

and

$$\bar{H}_k = [(\bar{H}_k^a)',\ (\bar{H}_k^b)']',\ \bar{R}_k = \mathrm{diag}(\bar{H}_k^a,\ \bar{H}_k^b)$$

Clearly, like optimal multiple-model filtering, the above fusion involves an exponentially increasing number of model-sequences and is impractical. Approximate algorithms (e.g., interacting multiple model (IMM), generalized pseudo-Baysian algorithms of order n (GPBn)) can be easily obtained and the reader is referred to Table II in Section IV for the distributed IMM fusion with transformed measurements.

Remark: By the above method, we can see that each local sensor needs to transmit y_k^n to the fusion center. Thus, when the measurement dimension n_z is larger than the state dimension n_x, there are communication savings. It is more suitable for the system in which each local station has many sensors. Otherwise, when $n_z < n_x$, the transformation can be carried out along time, that is, the measurements over a time interval are transformed to save communication. This will be discussed in Section IV.

C. Performance Analysis

For the performance of the above fusion, we have the following Theorem.

Theorem 1: If $\hat{x}_{k|k-1}^{(l^k)} = E[x_k|m_{(l^k)}, Z^{k-1}]$ and $\mu_{k|k-1}^{(l^k)} = P(m_{(l^k)}|Z^{k-1})$, where Z^{k-1} is the stacked measurement with respect to all sensors up to time $k-1$, then for system (1)–(2) the above fusion is equivalent to the centralized multiple-model fusion

$$\hat{x}_k = E[x_k|Z^k]$$

Proof: First, it is clear that given the system model the fusion problem is a single model distributed fusion with transformed data, which was studied in [16]. From that work we already have

$$\hat{x}_k^{(l^k)} = E[x_k|Y_k, \bar{X}_k, m_{(l^k)}] = E[x_k|Z_k, \bar{X}_k, m_{(l^k)}] \tag{7}$$

where $Z_k = [(z_k^a)', (z_k^b)']'$. This means that the model-conditioned fusion with the transformed measurements and with the raw measurements is identical. Then, it suffices to prove

$$P(m_{(l^k)}|Y_k) = P(m_{(l^k)}|Z_k)$$

From the extension of matrix inversion lemma [17]

$$(BCB' + A)^+ = A^+ - A^+B(C^{-1} + B'A^+B)^{-1}B'A^+$$

iff $AA^+B = B$, we have

$$\begin{aligned}&(\bar{H}_k P_{k|k-1}^{(l^k)} \bar{H}_k' + \bar{R}_k)^+ \\ &= \bar{R}_k^+ - \bar{R}_k^+ \bar{H}_k [(P_{k|k-1}^{(l^k)})^{-1} + \bar{H}_k' \bar{R}_k^+ \bar{H}_k]^{-1} \bar{H}_k' \bar{R}_k^+\end{aligned} \tag{8}$$

where

$$\begin{aligned}\bar{R}_k \bar{R}_k^+ \bar{H}_k &= \begin{bmatrix} \bar{H}_k^a & \\ & \bar{H}_k^b \end{bmatrix} \begin{bmatrix} \bar{H}_k^a & \\ & \bar{H}_k^b \end{bmatrix}^+ \begin{bmatrix} \bar{H}_k^a \\ \bar{H}_k^b \end{bmatrix} \\ &= \begin{bmatrix} \bar{H}_k^a (\bar{H}_k^a)^+ \bar{H}_k^a \\ \bar{H}_k^b (\bar{H}_k^b)^+ \bar{H}_k^b \end{bmatrix} = \bar{H}_k\end{aligned}$$

Thus, from Eq. (8), we have

$$\begin{aligned}&R_k^{-1} \check{H}_k (\bar{H}_k P_{k|k-1}^{(l^k)} \bar{H}_k' + \bar{R}_k)^+ \check{H}_k' R_k^{-1} \\ &= R_k^{-1} \check{H}_k \bar{R}_k^+ \check{H}_k' R_k^{-1} - R_k^{-1} \check{H}_k \bar{R}_k^+ \bar{H}_k \\ &\cdot [(P_{k|k-1}^{(l^k)})^{-1} + \bar{H}_k' \bar{R}_k^+ \bar{H}_k]^{-1} \bar{H}_k' \bar{R}_k^+ \check{H}_k' R_k^{-1}\end{aligned} \tag{9}$$

where $R_k = \mathrm{diag}(R_k^a, R_k^b)$ and $\check{H}_k = \mathrm{diag}(H_k^a, H_k^b)$. Since

$$\begin{aligned}\mathrm{rank}((R_k^a)^{-1} H_k^a) &= \mathrm{rank}((R_k^a)^{-1/2} (R_k^a)^{-1/2} H_k^a) \\ &= \mathrm{rank}((R_k^a)^{-1/2} H_k^a) \\ &= \mathrm{rank}((H_k^a)' (R_k^a)^{-1/2} (R_k^a)^{-1/2} H_k^a) \\ &= \mathrm{rank}((H_k^a)' (R_k^a)^{-1} H_k^a) \\ \mathrm{rank}((R_k^b)^{-1} H_k^b) &= \mathrm{rank}((H_k^b)' (R_k^b)^{-1} H_k^b)\end{aligned}$$

and by the properties of pseudo-inverse [17]

$$B(AB)^+AB = B,\ \text{if } \mathrm{rank}(B) = \mathrm{rank}(AB)$$

we have

$$\begin{aligned}&R_k^{-1} \check{H}_k \bar{R}_k^+ \bar{H}_k \\ &= \begin{bmatrix} (R_k^a)^{-1} H_k^a & \\ & (R_k^b)^{-1} H_k^b \end{bmatrix} \begin{bmatrix} \bar{H}_k^a & \\ & \bar{H}_k^b \end{bmatrix}^+ \begin{bmatrix} \bar{H}_k^a \\ \bar{H}_k^b \end{bmatrix} \\ &= \begin{bmatrix} (R_k^a)^{-1} H_k^a ((H_k^a)' (R_k^a)^{-1} H_k^a)^+ (H_k^a)' (R_k^a)^{-1} H_k^a \\ (R_k^b)^{-1} H_k^b ((H_k^b)' (R_k^b)^{-1} H_k^b)^+ (H_k^b)' (R_k^b)^{-1} H_k^b \end{bmatrix} \\ &= R_k^{-1} H_k\end{aligned}$$

where $H_k = [(H_k^a)', (H_k^b)']'$. Therefore, Eq. (9) can be rewritten as

$$
\begin{aligned}
&R_k^{-1}\check{H}_k(\bar{H}_k P_{k|k-1}^{(l^k)}\bar{H}_k' + \bar{R}_k)^+\check{H}_k' R_k^{-1} \\
&= R_k^{-1}\check{H}_k\bar{R}_k^+\check{H}_k' R_k^{-1} - R_k^{-1}H_k \\
&\cdot [(P_{k|k-1}^{(l^k)})^{-1} + \bar{H}_k'\bar{R}_k^+\bar{H}_k]^{-1}H_k' R_k^{-1} \\
&= R_k^{-1}\check{H}_k\bar{R}_k^+\check{H}_k' R_k^{-1} - R_k^{-1}H_k \\
&\cdot [(P_{k|k-1}^{(l^k)})^{-1} + H_k' R_k^{-1} H_k]^{-1}H_k' R_k^{-1}
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
&R_k^{-1}\check{H}_k(\bar{H}_k P_{k|k-1}^{(i^k)}\bar{H}_k' + \bar{R}_k)^+\check{H}_k' R_k^{-1} \\
&- R_k^{-1}\check{H}_k(\bar{H}_k P_{k|k-1}^{(j^k)}\bar{H}_k' + \bar{R}_k)^+\check{H}_k' R_k^{-1} \\
&= R_k^{-1}H_k[(P_{k|k-1}^{(j^k)})^{-1} + H_k' R_k^{-1}H_k]^{-1}H_k' R_k^{-1} \\
&- R_k^{-1}H_k[(P_{k|k-1}^{(i^k)})^{-1} + H_k' R_k^{-1}H_k]^{-1}H_k' R_k^{-1} \\
&= R_k^{-1}H_k[(P_{k|k-1}^{(j^k)})^{-1} + H_k' R_k^{-1}H_k]^{-1}H_k' R_k^{-1} - R_k^{-1} \\
&+ R_k^{-1} - R_k^{-1}H_k[(P_{k|k-1}^{(i^k)})^{-1} + H_k' R_k^{-1}H_k]^{-1}H_k' R_k^{-1} \\
&= (R_k + H_k P_{k|k-1}^{(i^k)}H_k')^{-1} - (R_k + H_k P_{k|k-1}^{(j^k)}H_k')^{-1} \quad (10)
\end{aligned}
$$

where i^k and j^k denote two different model sequences up to time k. Let

$$
\begin{aligned}
c(\tilde{Y}_k, l^k) &= \tilde{Y}_k'(\bar{H}_k P_{k|k-1}^{(l^k)}\bar{H}_k' + \bar{R}_k)^+\tilde{Y}_k \\
c(\tilde{Z}_k, l^k) &= \tilde{Z}_k'(H_k P_{k|k-1}^{(l^k)}H_k' + R_k)^{-1}\tilde{Z}_k
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{Z}_k &= Z_k - H_k\hat{x}_{k|k-1}^{(l^k)} \\
\tilde{Y}_k &= Y_k - \bar{H}_k\hat{x}_{k|k-1}^{(l^k)}
\end{aligned}
$$

Since $Y_k = \check{H}_k' R_k^{-1} Z_k$ and

$$
\begin{aligned}
\tilde{Y}_k &= \check{H}_k' R_k^{-1} Z_k - \bar{H}_k\hat{x}_{k|k-1}^{(l^k)} \\
&= \check{H}_k' R_k^{-1}(Z_k - H_k\hat{x}_{k|k-1}^{(l^k)}) \\
&= \check{H}_k' R_k^{-1}\tilde{Z}_k
\end{aligned}
$$

by Eq. (10) we have

$$
\begin{aligned}
&c(\tilde{Y}_k, i^k) - c(\tilde{Y}_k, j^k) \\
&= \tilde{Z}_k' R_k^{-1}\check{H}_k[(\bar{H}_k P_{k|k-1}^{(i^k)}\bar{H}_k' + \bar{R}_k)^+ \\
&- (\bar{H}_k P_{k|k-1}^{(j^k)}\bar{H}_k' + \bar{R}_k)^+]\check{H}_k' R_k^{-1}\tilde{Z}_k \\
&= \tilde{Z}_k'[(R_k + H_k P_{k|k-1}^{(i^k)}H_k')^{-1} - (R_k + H_k P_{k|k-1}^{(j^k)}H_k')^{-1}]\tilde{Z}_k \\
&= c(\tilde{Z}_k, i^k) - c(\tilde{Z}_k, j^k) \quad (11)
\end{aligned}
$$

This yields

$$
\frac{\exp(-\frac{1}{2}c(\tilde{Z}_k, i^k))}{\exp(-\frac{1}{2}c(\tilde{Z}_k, j^k))} = \frac{\exp(-\frac{1}{2}c(\tilde{Y}_k, i^k))}{\exp(-\frac{1}{2}c(\tilde{Y}_k, j^k))}
$$

Thus by definition have

$$
\begin{aligned}
&\frac{\mathcal{N}(Z_k; H_k\hat{x}_{k|k-1}^{(i^k)}, R_k + H_k P_{k|k-1}^{(i^k)}H_k')}{\mathcal{N}(Z_k; H_k\hat{x}_{k|k-1}^{(j^k)}, R_k + H_k P_{k|k-1}^{(j^k)}H_k')} \\
&= \frac{L(Y_k|\hat{x}_{k|k-1}^{(i^k)},\ P_{k|k-1}^{(i^k)})}{L(Y_k|\hat{x}_{k|k-1}^{(j)},\ P_{k|k-1}^{(j^k)})} \quad (12)
\end{aligned}
$$

By Bayes' theorem, it follows that

$$
\begin{aligned}
&\frac{P(m_{(i^k)}|Z_k)}{P(m_{(j^k)}|Z_k)} \\
&= \frac{\mu_{k|k-1}^{(i^k)}}{\mu_{k|k-1}^{(j^k)}}\frac{\mathcal{N}(Z_k; H_k\hat{x}_{k|k-1}^{(i^k)}, R_k + H_k P_{k|k-1}^{(i^k)}H_k')}{\mathcal{N}(Z_k; H_k\hat{x}_{k|k-1}^{(j^k)}, R_k + H_k P_{k|k-1}^{(j^k)}H_k')}
\end{aligned}
$$

where $m_{(i^k)}$ and $m_{(j^k)}$ denote the events that model sequences i^k and j^k are true respectively.

Substituting Eq. (12) into the above yields

$$
\frac{P(m_{(i^k)}|Z_k)}{P(m_{(j^k)}|Z_k)} = \frac{\mu_{k|k-1}^{(i^k)}}{\mu_{k|k-1}^{(j^k)}}\frac{L(Y_k|\hat{x}_{k|k-1}^{(i^k)},\ P_{k|k-1}^{(i^k)})}{L(Y_k|\hat{x}_{k|k-1}^{(j^k)},\ P_{k|k-1}^{(j^k)})} = \frac{P(m_{(i^k)}|Y_k)}{P(m_{(j^k)}|Y_k)}
$$

and since

$$
\sum_{l^k} P(m_{(l^k)}|Y_k) = \sum_{l^k} P(m_{(l^k)}|Z_k) = 1
$$

we have

$$
P(m_{(l^k)}|Y_k) = P(m_{(l^k)}|Z_k)
$$

which completes the proof. ■

IV. Distributed Multiple-Model Fusion with Transformed Measurements over a Time-Interval

When the communication capacity is very limited, the measurements over a time-interval can be transformed. In such a problem of multiple-model fusion with transformed measurements in reduced-rate communication, every N time instants the sensors send their transformed measurements to the fusion center and then the center fuses them accordingly.

A. Sensor Measurement Transformation

Denote $m_{(s)}^{k+1:k+N}$ as the event that model sequence s from time $k+1$ to $k+N$ is in effect. Therefore, there are M^N possible model sequences in the interval $(k+1, \ldots, k+N)$ on which our transformation is based. Thus the number of transformed measurements increases exponentially with N, which is unbearable. An enabling approximation is using the following assumption.

Assumption 1: $f(z_{k+1}^n, \ldots, z_{k+N}^n | m_{k+N}^{(l)}, x_{k+N})$, $n = a, b$, is Gaussian, where $m_{k+N}^{(l)} \in \mathbb{M}$ denotes the event that the model l is in effect at time $k+N$ and n is the sensor index (see Eq. (2)).

With Assumption 1, the number of model sequences in the interval is reduced to M and we have

$$
Z_{k+1:k+N}^n = \tilde{H}_{k+N}^{n,(s)}x_{k+N} + \tilde{V}_{k+N}^{n,(s)},\ n = a, b
$$

where

$$Z^n_{k+1:k+N} = \left[\begin{array}{cccc} (z^n_{k+N})' & (z^n_{k+N-1})' & \cdots & (z^n_{k+1})' \end{array} \right]'$$

$$\tilde{H}^{n,(s)}_{k+N} = \left[\begin{array}{c} H^n_{k+N} \\ H^n_{k+N-1}(F^{(s)}_{k+N-1})^{-1} \\ \vdots \\ H^n_{k+1}\prod_{j=1}^{N-1}(F^{(s)}_{k+j})^{-1} \end{array} \right]$$

$$\tilde{V}^{n,(s)}_{k+N} = \left[\begin{array}{c} v^n_{k+N} \\ -H^n_{k+N-1}(F^{(s)}_{k+N-1})^{-1}G^{(s)}_{k+N-1}w^{(s)}_{k+N-1} + v^n_{k+N-1} \\ \vdots \\ -H^l_{k+1}w^{(s)}_{k+N-1,k+1} + v^n_{k+1} \end{array} \right]$$

and $w^{(s)}_{k+N-1,k+1}$ is the cumulative effect of the process noise pertinent to model $m^{(s)}$ from $k+1$ to $k+N-1$:

$$w^{(s)}_{k+N-1,k+1} = \sum_{r=1}^{N-1}(\prod_{j=1}^{r}(F^{(s)}_{k+j})^{-1}G^{(s)}_{k+r}w^{(s)}_{k+r})$$

Two methods to implement the measurement-transformation are given as follows.

1) Batch Transformation: The transformed measurements processed at local sensors are

$$y^{n,(s)}_{k+N} = (\tilde{H}^{n,(s)}_{k+N})'C^{-1}_{\tilde{V}^{n,(s)}_{k+N}}Z^n_{k+1:k+N}, \ n = a, b \tag{13}$$

where $C_{\tilde{V}^n_{k+N}} = E[\tilde{V}^n_{k+N}(\tilde{V}^n_{k+N})']$. Let

$$\bar{H}^{n,(s)}_{k+N} = (\tilde{H}^{n,(s)}_{k+N})'C^{-1}_{\tilde{V}^{n,(s)}_{k+N}}\tilde{H}^{n,(s)}_{k+N}$$

$$\bar{V}^{n,(s)}_{k+N} = (\tilde{H}^{n,(s)}_{k+N})'C^{-1}_{\tilde{V}^{n,(s)}_{k+N}}\tilde{V}^{n,(s)}_{k+N}$$

Then we have

$$y^{n,(s)}_{k+N} = \bar{H}^{n,(s)}_{k+N}x_{k+N} + \bar{V}^{n,(s)}_{k+N}$$

where $E[\bar{V}^{n,(s)}_{k+N}] = 0$ and $\mathrm{cov}(\bar{V}^{n,(s)}_{k+N}) = \bar{H}^{n,(s)}_{k+N}$.

TABLE I

ONE CYCLE OF INFORMATION FILTER

1. Initialization (restart information filter):
$I^{n,(s)}_{k\|k} = 0_{n_x \times n_x}$, $\hat{y}^{n,(s)}_{k\|k} = 0_{n_x}$
2. Prediction:
$A^{n,(s)}_k = ((F^{(s)}_k)^{-1})'I^{n,(s)}_{k\|k}(F^{(s)}_k)^{-1}$
$T^{n,(s)}_k = G^{(s)}_k[(Q^{(s)}_k)^{-1} + (G^{(s)}_k)'A^{n,(s)}_kG^{(s)}_k]^{-1}(G^{(s)}_k)'$
$I^{n,(s)}_{k+1\|k} = A^{n,(s)}_k - A^{n,(s)}_kT^{n,(s)}_kA^{n,(s)}_k$
$\hat{y}^{n,(s)}_{k+1\|k} = (I - A^{n,(s)}_kT^{n,(s)}_k)((F^{(s)}_k)^{-1})'\hat{y}^{n,(s)}_{k\|k}$
3. Update:
$\hat{y}^{n,(s)}_{k+1\|k+1} = \hat{y}^{n,(s)}_{k+1\|k} + (H^n_{k+1})'(R^n_{k+1})^{-1}z^n_{k+1}$
$I^{n,(s)}_{k+1\|k+1} = I^{n,(s)}_{k+1\|k} + (H^n_{k+1})'(R^n_{k+1})^{-1}H^n_{k+1}$

2) Recursive Transformation: Apparently, a straightforward way of recursive calculating $y^{n,(s)}_{k+N}$ is using information filter [18] initialized with zero information (see, Table I). Its outputs are indeed $y^{n,(s)}_{k+N}$ and $\bar{H}^{n,(s)}_{k+N}$:

$$y^{n,(s)}_{k+N} = \hat{y}^{n,(s)}_{k+N|k+N}, \ \bar{H}^{n,(s)}_{k+N} = I^{n,(s)}_{k+N|k+N}$$

The batch transformation could not be carried out until the measurement z^n_{k+N} is obtained and thus is computationally inefficient, especially for a large N. Thus, the recursive transformation is preferred.

B. Multiple-model Fusion with Transformed Measurements

Given local transformed measurements, the MMSE fusion (under Assumption 1) is

$$\hat{x}_{k+N} = \sum_l\sum_s P(m^{(l,s)}|Y^{(s)}_{k+N}, \hat{x}^{(l,s)}_{k+N|k}) \cdot E[x_{k+N}|Y^{(s)}_{k+N}, \hat{x}^{(l,s)}_{k+N|k}, m^{(l,s)}]$$

where $Y^{(s)}_{k+N} = [(y^{a,(s)}_{k+N})', (y^{b,(s)}_{k+N})']'$. Let $M(k) = |\mathbb{M}(k)|$ where $\mathbb{M}(k)$ denotes the set of possible model sequences up to time k. Here (l, s) denotes a model sequence up to time $k+N$ composed of $m^k_{(l)} \in \mathbb{M}(k)$ and $m^{k+1:k+N}_{(s)} \in \mathbb{M}$.

Assume that the distributed fusion results at time k are $\hat{x}^{(l)}_k$, $P^{(l)}_k$ and $\mu^{(l)}_k$, and predict these quantities to time $k+N$ to get the prediction of x_{k+N}. Denote by $\hat{x}^{(l,s)}_{k+N|k}$, $P^{(l,s)}_{k+N|k}$ and $\mu^{(l,s)}_{k+N|k}$ the predicted quantities. Thus the model-sequence conditioned estimation is

$$\hat{x}^{(l,s)}_{k+N} = \hat{x}^{(l,s)}_{k+N|k} + K^{(s)}_{k+N}(Y^{(s)}_{k+N} - \bar{H}^{(s)}_{k+N}\hat{x}^{(l,s)}_{k+N|k}) \tag{14}$$

$$K^{(s)}_{k+N} = [P^{(l,s)}_{k+N|k}(\bar{H}^{(s)}_{k+N})' + C^{(l,s)}_{k+N}](S^{(s)}_{k+N})^{-1} \tag{15}$$

$$P^{(l,s)}_{k+N} = P^{(l,s)}_{k+N|k} - K^{(s)}_{k+N}S^{(s)}_{k+N}(K^{(s)}_{k+N})' \tag{16}$$

$$\begin{aligned} S^{(s)}_{k+N} &= \bar{H}^{(s)}_{k+N}P^{(l,s)}_{k+N|k}(\bar{H}^{(s)}_{k+N})' + \bar{R}^{(s)}_{k+N} \\ &\quad + \bar{H}^{(s)}_{k+N}C^{(l,s)}_{k+N} + (\bar{H}^{(s)}_{k+N}C^{(l,s)}_{k+N})' \end{aligned} \tag{17}$$

where

$$\bar{H}^{(s)}_{k+N} = [(\bar{H}^{a,(s)}_{k+N})', \ (\bar{H}^{b,(s)}_{k+N})']' \tag{18}$$

$$\bar{R}^{(s)}_{k+N} = \left[\begin{array}{cc} \bar{H}^{a,(s)}_{k+N} & R^{a,b,(s)}_{k+N} \\ (R^{a,b,(s)}_{k+N})' & \bar{H}^{b,(s)}_{k+N} \end{array} \right] \tag{19}$$

$$R^{a,b,(s)}_{k+N} = \mathrm{cov}(\bar{V}^{a,(s)}_{k+N}, \bar{V}^{b,(s)}_{k+N}) \tag{20}$$

$$\begin{aligned} C^{(l,s)}_{k+N} &= E[(x_{k+N} - \hat{x}^{(l,s)}_{k+N|k})(\bar{V}^{(s)}_{k+N})'] \\ &= \left[\begin{array}{cc} C^{a,(l,s)}_{k+N} & C^{b,(l,s)}_{k+N} \end{array} \right] \end{aligned} \tag{21}$$

The cross-covariance $R^{a,b,(s)}_{k+N}$ can be computed recursively as follows (see [19] for a derivation)

$$\begin{aligned} R^{a,b,(s)}_{m+1} &= U^{a,(s)}_m[R^{a,b,(s)}_m + (P^{a,(s)}_{m|m})^{-1}(F^{(s)}_m)^{-1}G^{(s)}_mQ^{(s)}_m \\ &\quad \cdot (G^{(s)}_m)'((F^{(s)}_m)^{-1})'(P^{b,(s)}_{m|m})^{-1}](U^{b,(s)}_m)' \end{aligned} \tag{22}$$

$$m = k, k+1, \ldots, k+N-1$$

where

$$U^{a,(s)}_m = (I - A^{a,(s)}_mT^{a,(s)}_m)((F^{(s)}_m)^{-1})'$$

$$U^{b,(s)}_m = (I - A^{b,(s)}_mT^{b,(s)}_m)((F^{(s)}_m)^{-1})'$$

and $A^{n,(s)}_m$ and $T^{n,(s)}_m$, $n = a, b$, are given in Table I. Since v^a_k and v^b_k are uncorrelated, we have

$$R^{a,b,(s)}_k = \mathbf{0}$$

The cross-covariance $C_{k+N}^{a,(l,s)}$ can be computed recursively as follows (see [19] for a derivation)

$$C_{m+1}^{a,(l,s)} = [F_m^{(s)} C_m^{a,(l,s)} - G_m^{(s)} Q_m^{(s)} (G_m^{(s)})' \cdot ((F_m^{(s)})^{-1})' (P_{m|m}^{a,(s)})^{-1}] (U_m^{a,(s)})' \quad m = k, k+1, \ldots, k+N-1 \tag{23}$$

and

$$C_k^{a,(l,s)} = \mathbf{0}$$

due to the whiteness of measurement and process noises.

The posterior model-sequence probability is

$$P(m^{(l,s)} | Y_{k+N}^{(s)}, \hat{x}_{k+N|k}^{(l,s)}) = \frac{f(Y_{k+N}^{(s)} | m^{(l,s)}, \hat{x}_{k+N|k}^{(l,s)}) \mu_{k+N|k}^{(l,s)}}{\sum_{l,s} f(Y_{k+N}^{(s)} | m^{(l,s)}, \hat{x}_{k+N|k}^{(l,s)}) \mu_{k+N|k}^{(l,s)}} \tag{24}$$

where $f(Y_{k+N}^{(s)} | m^{(l,s)}, \hat{x}_{k+N|k}^{(l,s)})$ is the model-sequence likelihood based on $Y_{k+N}^{(s)}$. When $\bar{H}_{k+N}^{n,(l)}$ for all n and l are invertible, $f(Y_{k+N}^{(s)} | m^{(l,s)}, \hat{x}_{k+N|k}^{(l,s)})$ has a Gaussian distribution. Otherwise, it is singular Gaussian distributed and we can use the following model-sequence "likelihood"

$$L(Y_{k+N}^{(s)} | \hat{x}_{k+N|k}^{(l,s)}, P_{k+N|k}^{(l,s)}) = (c_{k+N}^{(s)})^{-\frac{1}{2}} \exp[-\frac{1}{2} (\tilde{Y}_{k+N}^{(s)})' (S_{k+N}^{(s)})^{+} \tilde{Y}_{k+N}^{(s)}] \tag{25}$$

where

$$c_{k+N}^{(s)} = \det(\tilde{H}_{k+N}^{(s)} P_{k+N|k}^{(l,s)} (\tilde{H}_{k+N}^{(s)})' + \tilde{R}_{k+N}^{(s)})$$
$$\tilde{Y}_{k+N}^{(s)} = Y_{k+N}^{(s)} - \bar{H}_{k+N}^{(s)} \hat{x}_{k+N|k}^{(l,s)}$$
$$\tilde{H}_{k+N}^{(s)} = [(\tilde{H}_{k+N}^{a,(s)})', (\tilde{H}_{k+N}^{b,(s)})']'$$
$$\tilde{R}_{k+N}^{(s)} = \text{cov}([(\tilde{V}_{k+N}^{a,(s)})', (\tilde{V}_{k+N}^{b,(s)})']')$$

Note that unlike the full-rate communication case, here the errors of $y_{k+N}^{a,(s)}$ and $y_{k+N}^{b,(s)}$ are correlated. Thus, Eq. (25) is not necessarily a real likelihood. When $S_{k+N}^{(s)}$ is not invertible in Eq. (25), the distributed multiple-model fusion is an approximate MMSE fusion given local transformed measurements.

Remark: Each sensor needs to transmit $Z_{k+1:k+N}^n$, which has a dimension of Nn_z, to the fusion center for CMMF and $y_{k+N}^{n,(s)}$, which is Mn_x dimensional in total, for the above distributed fusion. Therefore, we can choose a large N such that $Mn_x < Nn_z$ to save transmission. It should be noted that when a large N is employed the Gaussian assumption of the distribution $f(z_{k+1}^n, \ldots, z_{k+N}^n | m_{k+N}^{(l)}, x_{k+N})$ may be seriously violated, which may cause a significant performance degradation.

C. Distributed IMM Fusion with Transformed Measurements

In practice, the above fusion is computationally infeasible, since the number of model-sequence will increase exponentially. Thus we summarize a complete recursion of the distributed IMM fusion with transformed measurements in Table II (see Eqs. (19), (22) and (23) for the computation of $\bar{R}_{k+N}^{(i)}$ and $C_{k+N}^{(i)}$).

TABLE II
ONE CYCLE OF DISTRIBUTED IMM FUSION

1. Model-conditioned reinitialization (for $i = 1, 2, \ldots, M$):

$$\bar{\mu}_{k+N}^{(i)} = \textstyle\sum_j \pi_{ji} \mu_k^{(j)}$$
$$\mu_k^{j|i} = \pi_{ji} \mu_k^{(j)} / \bar{\mu}_{k+N}^{(i)}$$
$$\bar{x}_{k|k}^{(i)} = \textstyle\sum_j \hat{x}_{k|k}^{(j)} \mu_k^{j|i}$$
$$\bar{P}_{k|k}^{(i)} = \textstyle\sum_j [P_{k|k}^{(j)} + (\bar{x}_{k|k}^{(i)} - \hat{x}_{k|k}^{(j)})(\bar{x}_{k|k}^{(i)} - \hat{x}_{k|k}^{(j)})'] \mu_k^{j|i}$$

2. Model-conditioned fusion (for $i = 1, 2, \ldots, M$):

$$\hat{x}_{k+N|k}^{(i)} = \textstyle\prod_{m=k}^{k+N-1} F_m^{(i)} \bar{x}_{k|k}^{(i)}$$
$$\bar{P}_{k+N|k}^{(i)} = F_{k+N-1}^{(i)} \bar{P}_{k+N-1|k}^{(i)} (F_{k+N-1}^{(i)})' + Q_{k+N-1}^{(i)}$$
$$\tilde{Y}_{k+N}^{(i)} = Y_{k+N}^{(s)} - \bar{H}_{k+N}^{(s)} \hat{x}_{k+N|k}^{(l,s)}$$
$$S_{k+N}^{(i)} = \bar{H}_{k+N}^{(i)} \bar{P}_{k+N|k}^{(i)} (\bar{H}_{k+N}^{(i)})' + \bar{R}_{k+N}^{(i)} + \bar{H}_{k+N}^{(i)} C_{k+N}^{(i)} + (\bar{H}_{k+N}^{(i)} C_{k+N}^{(i)})'$$
$$K_{k+N}^{(i)} = [\bar{P}_{k+N|k}^{(i)} (\bar{H}_{k+N}^{(i)})' + C_{k+N}^{(i)}] (S_{k+N}^{(i)})^{-1}$$
$$\hat{x}_{k+N}^{(i)} = \hat{x}_{k+N|k}^{(i)} + K_{k+N}^{(i)} \tilde{Y}_{k+N}^{(i)}$$
$$P_{k+N}^{(i)} = \bar{P}_{k+N|k}^{(i)} - K_{k+N}^{(i)} S_{k+N}^{(i)} (K_{k+N}^{(i)})'$$

3. Mode probability update (for $i = 1, 2, \ldots, M$):

$$c_{k+N}^{(i)} = \det(\tilde{H}_{k+N}^{(i)} P_{k+N|k}^{(i)} (\tilde{H}_{k+N}^{(i)})' + \tilde{R}_{k+N}^{(i)})$$
$$L_{k+N}^{(i)} = (c_{k+N}^{(i)})^{-\frac{1}{2}} \exp(-\frac{1}{2} (\tilde{Y}_{k+N}^{(i)})' (S_{k+N}^{(i)})^{+} \tilde{Y}_{k+N}^{(i)})$$
$$\mu_{k+N}^{(i)} = \frac{\bar{\mu}_{k+N}^{(i)} L_{k+N}^{(i)}}{\sum_j \bar{\mu}_{k+N}^{(j)} L_{k+N}^{(j)}}$$

4. Overall results

$$\hat{x}_{k+N|k+N} = \textstyle\sum_i \hat{x}_{k+N}^{(i)} \mu_{k+N}^{(i)}$$
$$P_{k+N|k+N} = \textstyle\sum_i [P_{k+N}^{(i)} + (\hat{x}_{k+N}^{(i)} - \hat{x}_{k+N|k+N})(\hat{x}_{k+N}^{(i)} - \hat{x}_{k+N|k+N})']$$

V. NUMERICAL EXAMPLES

In this section, we provide several illustrative examples of the fusion algorithms presented in the previous sections. Two kinematic models are adopted: one is (nearly) constant velocity (CV) and the other is (nearly) constant acceleration (CA). All measurements are generated by two sensors. The system models are as follows

$$x_{k+1} = F_k^{(l)} x_k + G_k^{(l)} w_k^{(l)}, \ l = 1, 2$$
$$z_k^n = [\ 1 \quad 0 \quad 0\ ] x_k + v_k^n, \ n = 1, 2$$

where

$$F_k^{(1)} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ F_k^{(2)} = \begin{bmatrix} 1 & T & \frac{T^2}{2} \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}$$
$$G_k^{(1)} = \begin{bmatrix} \frac{T^2}{2} & T & 0 \end{bmatrix}', \ G_k^{(2)} = \begin{bmatrix} \frac{T^2}{2} & T & 1 \end{bmatrix}'$$

sampling interval $T = 1$s, $w_k^{(l)}$, $l = 1, 2$, are zero-mean white Gaussian process noises with variances $q^{(1)} = 0.01$ and $q^{(2)} = 0.1$ respectively. The two synchronous measurements have a fixed rate and the same measurement model, and each being one-dimensional measurement. The measurement noises are zero-mean white Gaussian and mutually independent with variances $R^1 = R^2 = 1$. The Markovian transition matrix of the models is

$$[\pi_{ij}]_{2\times 2} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

The true initial state is generated Gaussian distributed with mean and covariance

$$\bar{x}_0 = \begin{bmatrix} 100 & 30 & 0 \end{bmatrix}'$$
$$P_0 = \text{diag}(25,\ 16,\ 16)$$

which are also used to initialize the multiple-model filter at the fusion center. The true initial mode of the target motion is CV or CA at random with equal probability and the two model probabilities of the multiple-model filter are set to 0.5. For each example, 500 Monte Carlo trails are conducted, each having a time span of 60 seconds.

Two scenarios are provided to show characters of the fusion methods. In the simulation, centralized IMM fusion (CIMMF) and distributed IMM fusion (DIMMF) are used, in which Kalman filter is employed for model-conditioned estimation; their performance is examined by the root mean-squared errors (RMSE).

A. Scenario 1

In this scenario, the target switches its motion-mode according to the above π_{ij} at each time point.

Fig. 1 shows the position, velocity and acceleration RMSE of the DIMMF with full-rate communication ($N = 1$), the reduced-rate DIMMF ($N = 3, 5, 6, 8$, respectively) and CIMMF. It can be seen that the full-rate DIMMF and CIMMF have identical performance, and the DIMMF performance degrades as N increases (i.e., communication rate decreases). This is understandable, since as N increases more information will be lost by the transformation and the fusion is just based on transformed measurements rather than raw data.

B. Scenario 2

In this scenario, the target switches its motion-mode after every 10 seconds.

Fig. 2 shows the position, velocity and acceleration RMSE of the DIMMF with full-rate communication ($N = 1$), reduced-rate DIMMF ($N = 3, 5, 6, 8$ respectively) and CIMMF. Also, it can be seen that the full-rate DIMMF and CIMMF have identical performance.

For the reduced-rate communication case, the DIMMF methods with $N = 3$ and $N = 5$ have a comparable performance. This is because for $N = 5$ Assumption 1 is always satisfied and the DIMMF with $N = 5$ takes advantage of it, although the transformation loses more information than the DIMMF with $N = 3$. Assumption 1, if satisfied, is useful prior information and should be used in the fusion algorithms. Therefore, it is not surprising to see from Fig. 2 that at some time points, the DIMMF with $N = 3$ or $N = 5$ even outperforms the CIMMF.

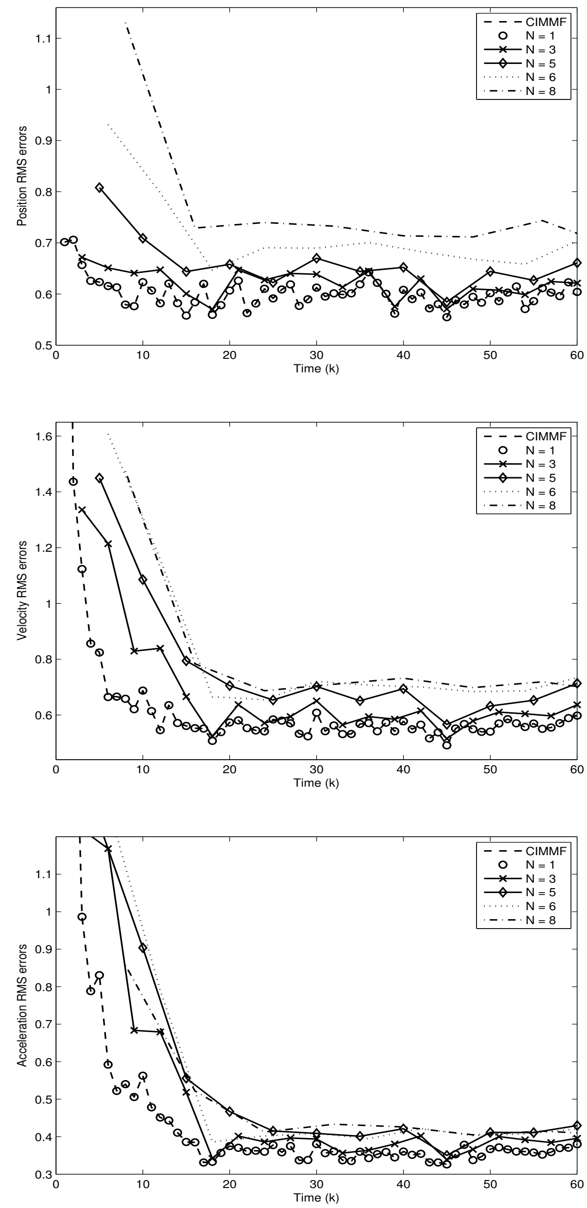

Fig. 1. Scenario 1

VI. Conclusions

In this paper, we have presented a distributed multiple-model fusion with transformed measurements. Both the full-rate and the reduced-rate communication cases are considered. For the full-rate communication case, the proposed distributed MMSE fusion is identical with centralized MMSE fusion; for the reduced-rate communication case, the distributed fusion is developed under a heuristic assumption and performs well compared with centralized fusion. Algorithms for local measurement transformation and the distributed IMM fusion are also provided. They are all in a recursive form convenient for implementation.

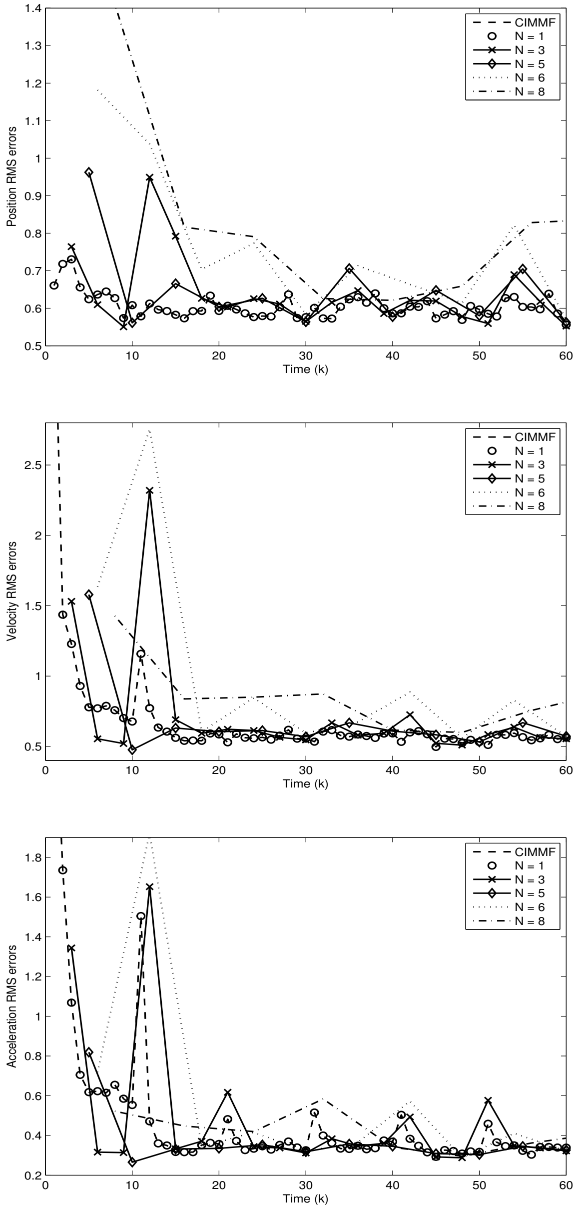

Fig. 2. Scenario 2

REFERENCES

[1] X. R. Li, Y. Zhu, J. Wang, and C. Han, "Optimal linear estimation fusion – Part I: Unified fusion rules," *IEEE Transactions on Information Theory*, vol. 49, no. 9, pp. 2192–2208, 2003.

[2] C. Y. Chong, "Hierarchical estimation," in *Proc. MIT/ONR Workshop on C3, Monterey, CA*, 1979.

[3] Y. Bar-Shalom and L. Campo, "The effect of the common process noise on the two-sensor fused-track covariance," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 6, pp. 803–805, 1986.

[4] S. Mori, W. H. Barker, C. Y. Chong, and K. C. Chang, "Track association and track fusion with nondeterministic target dynamics," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 2, pp. 659–668, 2002.

[5] O. E. Drummond, "On track and tracklet fusion filtering," in *Proceedings of SPIE*, vol. 4728, pp. 176–195, 2002.

[6] X. R. Li and J. Wang, "Optimal linear estimation fusion – Part II: Discussions and examples," in *Proceedings of the 3th International Information Fusion Conference*, vol. 1, pp. MOC2/18 – MOC2/25, July 2001.

[7] X. R. Li and P. Zhang, "Optimal linear estimation fusion – Part III: Cross-correlation of local estimation errors," in *Proceedings of 2001 International Information Fusion Conference*, vol. 1, pp. WeB1.11–Web1.18, Aug. 2001.

[8] X. R. Li and K. S. Zhang, "Optimal linear estimation fusion – Part IV: Optimality and efficiency of distributed fusion," in *Proceedings of the 2001 International Information Fusion Conference*, pp. WeB1.19–Web1.26, Aug. 2001.

[9] X. R. Li, K. S. Zhang, J. Zhao, and Y. Zhu, "Optimal linear estimation fusion – Part V: Relationships," in *Proceedings of the 5th International Information Fusion Conference*, vol. 1, pp. 497–504, 2002.

[10] K. S. Zhang, X. R. Li, P. Zhang, and H. Li, "Optimal linear estimation fusion – Part VI: Sensor data compression," in *Proceedings of the 6th International Information Fusion Conference*, pp. 221–228, June 2003.

[11] X. R. Li, "Optimal linear estimation fusion – Part VII: Dynamic systems," in *Proceedings of the 6th International Information Fusion Conference*, vol. 1, pp. 455–462, 2003.

[12] C. O. L. V., "Linear minimum mean square error estimation for discrete-time markovian jump linear systems," *IEEE Transactions on Automatic Control*, vol. 39, pp. 1685–1689, Aug. 1994.

[13] Z. Ding and L. Hong, "A distributed IMM fusion algorithm for multi-platform tracking," *Signal Processing*, vol. 64, no. 2, pp. 167–176, 1998.

[14] H. Lang, W. W. C., L. M., and D. T., "Multiplatform multisensor fusion with adaptive-rate data communication," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 1, pp. 274–281, 1997.

[15] C. B. Chang and R. B. Holmes, "On linear estimation with transformed measurements," *IEEE Transactions on Automatic Control*, vol. 28, no. 2, pp. 242–244, 1983.

[16] Z. S. Duan and X. R. Li, "Optimal distributed estimation fusion with transformed data," in *Proceedings of the 11th International Information Fusion Conference*, pp. 1–7, June 2008.

[17] D. S. Bernstein, *Matrix Mathematics*. Princeton: Princeton University Press, 2005.

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

[19] Y. X. Gao and X. R. Li, "Quasi-tracklet fusion accounting for cross-correlation," in *Proceedings of the 13th International Information Fusion Conference*, July 2010.