Optimal Linear Estimation Fusion— Part VII: Dynamic Systems *

X. Rong Li

Department of Electrical Engineering, University of New Orleans

New Orleans, LA 70148, USA

Tel: (504) 280-7416, Fax: (504) 280-3950, Email: xli@uno.edu

Abstract – In this paper, we first present a general data model for discretized asynchronous multisensor systems and show that errors in the data model are correlated across sensors and with the state. This coupling renders most existing “optimal” linear fusion rules suboptimal. While our fusion rules of Part I are valid and optimal for this general model, we propose a general, exact technique to decouple the two types of correlation of the errors so that other existing rules can be applied after decoupling. Then, we discuss several theoretically important issues unique to fusion for dynamic systems. The first is the role of prior information in the static case versus that of prediction in the dynamic case. We present two general, best linear unbiased estimation fusers with and without prior information respectively. Other issues discussed include optimality of existing linear fusion rules as well as two commonly used fusion schemes, and the effect of feedback.

Keywords: asynchronism, BLUE, fusion, recursive estimation, linear estimation, error correlation, linear MMSE.

1 Introduction

This is part VII of the series of papers on optimal linear estimation fusion. The first six parts focus on the static case where either the estimatee is not a process or the fuser is not recursive. More specifically, part I presents a unified fusion model and unified batch fusion rules for centralized, distributed, and hybrid estimation fusion with complete, incomplete, or no prior information; part II illustrates the rules by examples; part III presents concrete formulas for computing crosscovariance of the local estimation errors for general linear systems; part IV presents conditions for centralized and distributed fusers to be identical, as well as efficiency of the distributed fusion; part V clarifies relationships among the various fusion rules available; part VI

presents optimal rules for each sensor to compress its measurements to the required dimension. In this part, we focus on various issues unique to fusion for dynamic systems.

2 Discretized Multisensor Systems

2.1 Sampling of Asynchronous Multisensor Systems

Consider a fusion system with n sensors and a fusion center. Measurements are made by the n sensors at the time instants $t_1, \dots, t_n$, possibly different from one another because of asynchronism. Suppose that the i th sensor has the following continuous-time dynamic system model

$$dx(t) = A^i(t)x(t)dt + dw^i(t) \quad (1)$$

$$dz^i(t) = C^i(t)x(t)dt + dv^i(t) \quad (2)$$

where $w^i(t)$ and $v^i(t)$ are Wiener processes. For simplicity of presentation, the deterministic control input $u(t)$ is not included in this model and $w^i(t)$ and $v^i(t)$ are mutually independent and have zero mean.

Note that different sensors may have different dynamic models of the same state process $x(t)$ as well as different observation models. As such, the dynamic modeling error processes of the i th sensor $w^i(t)$ and of the j th sensor $w^j(t)$ are correlated in general.

Let

$$x_p = x(t_p), \quad x_q = x(t_q), \quad x_r = x(t_r), \quad x_s = x(t_s)$$

for arbitrary times t_p, t_q, t_k . Then, the sampling of the dynamic equation at the i th sensor yields

$$x_p = F_{p,q}^i x_q + w_{p,q}^i \quad (3)$$

where $F_{p,q}^i = \Phi^i(t_p, t_q)$ is the state transition matrix that transforms state x_q to state x_p and

$$w_{p,q}^i = w^i(t_p, t_q) = \int_{t_q}^{t_p} \Phi^i(t_p, \tau) dw^i(\tau) \quad (4)$$

*Supported in part by NSF via Grant ECS-9734285 and LEQSF via Grant (1996-99)-RD-A-32.

Note that it follows from $x_p = F_{p,q}^j x_q + w_{p,q}^j$ that the interdependence of $w^i(t)$ and $w^j(t)$ is such that the following equation holds for all t_p and t_q :

$$F_{p,q}^j x_q + w_{p,q}^j = F_{p,q}^i x_q + w_{p,q}^i$$

Suppose that the state x_k at time t_k is to be estimated using all data $y_k^1, \dots, y_k^n$ received from all sensors about x_k . Suppose also that for obtaining the fused estimate of x_k sensor i sends to the fusion center either unprocessed (but sampled) data z_r^i , the estimate $\hat{x}_{r|r}^i$ of x_r using data through time t_r , or some other linearly or nonlinearly processed data $g_i(z_r^i)$. For the centralized fusion, $y_k^i = z_r^i$; for the standard distributed fusion, $y_k^i = \hat{x}_{r|r}^i$; and for non-standard distributed fusion, $y_k^i = g_i(z_r^i)$, where $g_i(\cdot)$ is such that $g_i(z_r^i) \neq \hat{x}_{r|r}^i$ and $g_i(z_r^i) \neq z_r^i$. The nonstandard distributed fusion is the topic of Part VI and will not be discussed in this part. The narrow-sense hybrid fusion means that $y_k^i = z_r^i$ or $\hat{x}_{r|r}^i$ for any i and $y_k^1, \dots, y_k^n$ consist not only of one type. The case where $y_k^1, \dots, y_k^n$ have two or more of these three types of information can be defined as hybrid fusion in a wide sense.

Define for $t_r \leq t_s$

$$z_{s,r}^i = z^i(t_s) - z^i(t_r) = \int_{t_r}^{t_s} dz^i(\tau) = H_{s,r}^i x_r + \tilde{v}_{s,r}^i$$

where

$$H_{s,r}^i = \int_{t_r}^{t_s} C^i(\tau) \Phi^i(\tau, t_r) d\tau,$$

$$\tilde{v}_{s,r}^i = \int_{t_r}^{t_s} [C^i(\tau) w^i(\tau, t_r) d\tau + dv^i(\tau)]$$

Sampling of an observation process is sometimes defined as the observation integrated over the sampling interval $[kT_i, (k+1)T_i)$, given by $z_k^i = z^i((k+1)T_i) - z^i(kT_i)$ or $z_k^i = z^i(kT_i) - z^i((k-1)T_i)$ for uniform sampling with constant sampling interval T_i , but more often as instantaneous observation given by $z_k^i = z^i(t_k)$. In reality, each observation is made by a sensor that integrates over a time interval Δ_i necessary to cumulate signal energy to make an observation. Δ_i is not necessarily related to the sampling interval T_i (and in fact usually significantly smaller than T_i). For example, the measurement z_k provided by a scanning radar is obtained by integrating over an interval much smaller than the interval for each scan, which includes possibly many measurements in one scan. In view of this, we define

$$z_r^i = z_{s,r}^i|_{t_s=t_r+\Delta_i} = H_r^i x_r + v_r^i \quad (5)$$

where

$$H_r^i = H_{s,r}^i|_{t_s=t_r+\Delta_i} = \int_{t_r}^{t_r+\Delta_i} C^i(\tau) \Phi^i(\tau, t_r) d\tau$$

$$v_r^i = \tilde{v}_{s,r}^i|_{t_s=t_r+\Delta_i} = \int_{t_r}^{t_r+\Delta_i} [C^i(\tau) w^i(\tau, t_r) d\tau + dv^i(\tau)]$$

Note that even for uniform sampling with T_i , t_r can be anywhere in a sampling interval (i.e., $z_r^i \neq z_k^i$ or z_{k+1}^i in general) and z_r^i is made over the interval Δ_i that is entirely within one sampling interval, for example, $(t_r, t_r + \Delta_i) \in [kT_i, (k+1)T_i)$ or $(t_r, t_r + \Delta_i) \in ((k-1)T_i, kT_i]$. For some sensors, z_r^i is reported as the observation made at time t_r ; for others, z_r^i is reported as the observation made at some sampling time t_k , for example, when $(t_r, t_r + \Delta_i) \in [t_k, t_{k+1})$.

We define synchronism with respect to sampling and data, respectively. **Synchronous sampling** (with a constant sampling interval) is one such that $[kT_i, (k+1)T_i) = [kT_j, (k+1)T_j)$ for all i, j , and k , while **synchronous data** means that the observations within each sampling interval from all sensors are made at the same time. **Asynchronous sampling** and **asynchronous data** are defined otherwise in an obvious way. It is unrealistic to assume synchronous sampling, let alone synchronous data.

2.2 Synchronization of Asynchronous Data

For the centralized fusion, z_r^i can be treated as an observation of the state x_k at time t_k in general different from t_r as follows

$$z_r^i = H_r^i x_r + v_r^i = H_r^i (F_{r,k}^i x_k + w_{r,k}^i) + v_r^i$$

That is,

$$y_k^i = H_{r,k}^i x_k + v_{r,k}^i$$

where

$$y_k^i = z_r^i, \quad H_{r,k}^i = H_r^i F_{r,k}^i, \quad v_{r,k}^i = H_r^i w_{r,k}^i + v_r^i$$

For the standard distributed fusion, the estimate $\hat{x}_{r|r}^i$ of x_r using data through time t_r can also be treated as an observation of x_k :¹

$$\hat{x}_{r|r}^i = F_{r,k}^i x_k + \hat{x}_{r|r}^i - F_{r,k}^i x_k$$

Thus, we have

$$y_k^i = H_{r,k}^i x_k + v_{r,k}^i$$

where

$$y_k^i = \hat{x}_{r|r}^i, \quad H_{r,k}^i = F_{r,k}^i, \quad \hat{x}_{r|r}^i = x_r - \hat{x}_{r|r}^i$$

$$v_{r,k}^i = \hat{x}_{r|r}^i - F_{r,k}^i (F_{k,r}^i x_r + w_{k,r}^i) = -\hat{x}_{r|r}^i - F_{r,k}^i w_{k,r}^i$$

More generally, y_k^i can be a stacked vector of several z_r^i 's and/or $\hat{x}_{p|r}^i$'s for different r and p .

Taking $t_r = t_i$ for the i th sensor, we thus have the **unified data model**

$$y_k = H_k x_k + v_k \quad (6)$$

¹For simplicity, we do not consider the more general case

$$\hat{x}_{p|r}^i = F_{p,k}^i x_k + \hat{x}_{p|r}^i - F_{p,k}^i x_k$$

which will make the notation very heavy.

where

$$y_k = [(y_1^1)', \dots, (y_n^n)']', \quad H_k = [(H_{1,k}^1)', \dots, (H_{n,k}^n)']', \\ v_k = [(v_{1,k}^1)', \dots, (v_{n,k}^n)']'$$

Note that this model is valid for asynchronous as well as synchronous cases. It also includes the so-called out-of-sequence measurements as special cases [1, 13].

2.3 Correlation in Data Model

Centralized Fusion. The covariance of the noise in the unified data model is $R_k = \text{cov}(v_k) = [R_k^{ij}]$, where $R_k^{ij} = \text{cov}(v_{r,k}^i, v_{s,k}^j)$ and

$$R_k^{ij} = E[(H_r^i w_{r,k}^i + v_r^i)(H_s^j w_{s,k}^j + v_s^j)'] \\ = H_r^i Q_{(r,k),(s,k)}^{ij} (H_s^j)' + R_{r,s}^{ij} \\ + H_r^i E[w_{r,k}^i (v_s^j)'] + E[v_r^i (w_{s,k}^j)'] (H_s^j)'$$

As explained before, $dw^i(t)$ and $dw^j(t)$ are correlated in general for the same time. They are the same, known as the *common process noise*, if the same system dynamics model is used in sensors i and j . Since $w^i(t)$ and $w^j(t)$ are Wiener processes, $dw^i(t_1)$ and $dw^j(t_2)$ are independent for all $t_1 \neq t_2$. This implies the following. (a) $Q_{(r,k),(s,k)}^{ij} = E[w_{r,k}^i (w_{s,k}^j)'] \neq 0$ in general unless t_r and t_s are on two sides of t_k (i.e., $t_r < t_k < t_s$ or $t_s < t_k < t_r$). For example for the case of common process noise with $t_k \leq t_r \leq t_s$ or $t_s \leq t_r \leq t_k$, we have

$$Q_{(r,k),(s,k)}^{ij} = E[w_{r,k}^i w_{s,k}^j] \\ = \int_{t_k}^{t_r} \int_{t_k}^{t_r} \Phi(t_r, \tau) E[dw(\tau) dw(\tau)'] \Phi(t_s, \tau)' \\ = \int_{t_k}^{t_r} \Phi(t_r, \tau) Q \Phi(t_s, \tau)' d\tau \quad (7)$$

where $Q(t) = \text{cov}[w(t)] = Qt$ since $w(t)$ is a Wiener process. (b) $E[w_{r,k}^i (v_s^j)'] \neq 0$ in general unless $t_r < t_k < t_s$ or $t_s + \Delta_j < t_k < t_r$. (c) $E[v_r^i (w_{s,k}^j)'] \neq 0$ in general unless $t_r + \Delta_i < t_k < t_s$ or $t_s < t_k < t_r$. (d) $R_{r,s}^{ij} = E[v_r^i (v_s^j)'] \neq 0$ in general unless $t_r + \Delta_i < t_s$ or $t_s + \Delta_j < t_r$.

As a result, *the measurement errors are correlated across sensors in general unless $t_r + \Delta_i < t_k < t_s$ or $t_s + \Delta_j < t_k < t_r$ for all $i \neq j$.*

The crosscovariance of the state x_k and the noise v_k in the unified data model is $C_{x_k v_k} = \text{cov}(x_k, v_k) = [\text{cov}(x_k, v_{1,k}^1), \dots, \text{cov}(x_k, v_{n,k}^n)]$, where

$$\text{cov}(x_k, v_{r,k}^i) = E[(x_k - \bar{x}_k)(w_{r,k}^i)'] (H_r^i)' + E[(x_k - \bar{x}_k)(v_r^i)']$$

Since x_k is dependent on $dw^i(\tau)$ for all $\tau < t_k$ in general but independent of $dw^i(\tau)$ for all $\tau > t_k$, it follows from an analysis similar to the one above that (a)

$E[(x_k - \bar{x}_k)(w_{r,k}^i)'] \neq 0$ in general unless $t_r > t_k$ and (b) $E[(x_k - \bar{x}_k)(v_r^i)'] \neq 0$ in general unless $t_r + \Delta_i > t_k$. Consequently, *the state and measurement noise are correlated (i.e., $C_{x_k v_k} \neq 0$) in general unless $t_r + \Delta_i > t_k$ for all i .*

Combining the two cases for R_k and $C_{x_k v_k}$, we conclude that in the unified data model for the centralized fusion, *either the measurement errors are correlated across sensors or they are correlated with the state*. This is true even for the case with synchronous data (i.e., $t_r = t_k = t_s$ for all sensors).

Standard Distributed Fusion. We have

$$R_k^{ij} = \text{cov}(v_{r,k}^i, v_{s,k}^j) \\ = P_{r,s|r,s}^{ij} + F_{r,k}^i Q_{(k,r),(k,s)}^{ij} (F_{s,k}^j)' \\ + F_{r,k}^i E[w_{k,r}^i (\tilde{x}_{s|s}^j)'] + E\{\tilde{x}_{r|r}^i (w_{k,s}^j)'\} (F_{s,k}^j)' \quad (8)$$

Assume $t_r \leq t_s$ (without loss of generality) and

$$\tilde{x}_{s|s}^j = x_s - \hat{x}_{s|s}^j \\ = x_s - [\hat{x}_{s|r}^j + K_{s,r}^j (z_{s,r}^j - \hat{z}_{s,r|s}^j)] \\ = F_{s,r}^j (x_r - \hat{x}_{r|r}^j) + w_{s,r}^j \\ - K_{s,r}^j (H_{s,r}^j x_r + \tilde{v}_{s,r}^j - H_{s,r}^j \hat{x}_{r|r}^j) \\ = (F_{s,r}^j - K_{s,r}^j H_{s,r}^j) \tilde{x}_{r|r}^j + w_{s,r}^j - K_{s,r}^j \tilde{v}_{s,r}^j$$

Then

$$P_{r,s|r,s}^{ij} = E[\tilde{x}_{r|r}^i (\tilde{x}_{s|s}^j)'] = P_{r|r}^{ij} (F_{s,r}^j - K_{s,r}^j H_{s,r}^j)' \quad (9)$$

where

$$P_{r|r}^{ij} = E[\tilde{x}_{r|r}^i (\tilde{x}_{r|r}^j)'] \\ = (I - K_r^i H_r^i) (F_{r,r-1}^i P_{r-1|r-1}^{ij} (F_{r,r-1}^j)' \\ + Q_{(r,r-1),(r,r-1)}^{ij}) (I - K_r^j H_r^j)' + K_r^i R_{r,r}^{ij} (K_r^j)'$$

since

$$\tilde{x}_{r|r}^i = x_r - \hat{x}_{r|r-1}^i - K_r^i [H_r^i x_r + v_r^i - H_r^i \hat{x}_{r|r-1}^i] \\ = (I - K_r^i H_r^i) (x_r - \hat{x}_{r|r-1}^i) - K_r^i v_r^i \\ = (I - K_r^i H_r^i) (F_{r,r-1}^i \tilde{x}_{r-1|r-1}^i + w_{r,r-1}^i) - K_r^i v_r^i$$

(8) is a generalization of (8.4.2-3) of [2] for the general case of possible asynchronism and different dynamic models. It reduces to (8.4.2-3) of [2] with $r = s = k$ and common dynamic model for i and j .

For $t_k > t_r$, we have $E[\tilde{x}_{r|r}^i (w_{k,s}^j)'] = 0$ and

$$E[w_{k,r}^i (\tilde{x}_{s|s}^j)'] = E[w_{k,r}^i (w_{s,r}^j - K_{s,r}^j \tilde{v}_{s,r}^j)'] \\ = Q_{(k,r),(s,r)}^{ij} - E[w_{k,r}^i (\tilde{v}_{s,r}^j)'] (K_{s,r}^j)'$$

where $E[w_{k,r}^i(\tilde{v}_{s,r}^j)']$ can be obtained similarly as (7), and for $t_k < t_r$, we have $E[w_{k,r}^i(\tilde{x}_{s|s}^j)'] = 0$ and

$$\begin{aligned} & E[\tilde{x}_{r|r}^i(w_{k,s}^j)'] \\ &= (I - K_r^i H_r^i)(F_{r,r-1}^i E[\tilde{x}_{r-1|r-1}^i(w_{k,s}^j)'] \\ &+ Q_{(r,r-1),(k,s)}^{ij}) - K_r^i E[v_r^i(w_{k,s}^j)'] \end{aligned}$$

where $E[v_r^i(w_{k,s}^j)']$ can be obtained similarly as (7).

The crosscovariance of the state x_k and the noise v_k in the unified data model is $C_{x_k v_k} = \text{cov}(x_k, v_k) = [\text{cov}(x_k, v_{1,k}^1), \dots, \text{cov}(x_k, v_{n,k}^n)]$, where

$$\begin{aligned} \text{cov}(x_k, v_{r,k}^i) &= -E[(x_k - \bar{x}_k)(\tilde{x}_{r|r}^i)'] \\ &\quad - E[(x_k - \bar{x}_k)(w_{k,r}^i)'](F_{r,k}^i)' \quad (10) \end{aligned}$$

Clearly both terms are in general nonzero.

In summary, for the centralized, distributed, or hybrid fusion, the measurement errors in the unified data model are in general correlated across sensors and/or with the state. This is true even for the case with synchronous sampling or synchronous data.

It is also easy to see that the process and data error sequences of the discretized systems obtained by sampling and synchronization of asynchronous multisensor continuous-time systems are in general colored and mutually correlated.

3 Decoupling of Correlated Data Model

The correlation shown above in the data model for the discretized multisensor systems have been ignored or overlooked so far, even after we pointed it out in [7]. This is the case for virtually all prior fusion results for linear systems. Because of this, prior “optimal” fusion rules for discretized systems are actually not globally optimum.

The fusion rules presented in the preceding parts of this series (e.g., Parts I and VI) can be used straightforwardly since they are directly valid for the general data model, including the correlated models discussed here.

We now present an alternative, decoupling-based method of dealing with the above correlation optimally. It converts the original data model to an equivalent one in which errors are uncorrelated with the state and across sensors.

3.1 Decorrelation of Data Errors with State

Consider the unified linear data model (6) $y_k = H_k x_k + v_k$. We drop the time index k in this section for simplicity.

Theorem 3.1. Any linear model $y = Hx + v$ with known $\bar{v} = E[v]$, $C_v = \text{cov}(v)$, $\bar{x} = E[x]$, $C_x = \text{cov}(x)$, and $C_{xv} = \text{cov}(x, v) \neq 0$ can be converted to the following equivalent linear model

$$y = \tilde{H}x + \tilde{v} \quad (11)$$

where

$$\begin{aligned} \tilde{H} &= H + C_{xv}' C_x^+, \quad C_{x\tilde{v}} = 0 \\ E[\tilde{v}] &= \bar{v} - C_{xv}' C_x^+ \bar{x}, \quad C_{\tilde{v}} = C_v - C_{xv}' C_x^+ C_{xv} \end{aligned}$$

where A^+ is the Moore-Penrose pseudoinverse of A , which is equal to inverse A^{-1} whenever A^{-1} exists.

Note that the equivalent noise $\tilde{v}$ are coupled across sensors and has nonzero mean.

Proof. Consider

$$y = (H + C_{xv}' C_x^+)x + v - C_{xv}' C_x^+ x = \tilde{H}x + \tilde{v}$$

where $\tilde{v} = v - C_{xv}' C_x^+ x$. Let $u = (I - C_x C_x^+)(x - \bar{x})$. Then

$$\text{cov}(u) = (I - C_x C_x^+)C_x(I - C_x C_x^+)' = 0$$

implies $u = E[u] = 0$ almost surely. Note that $u = 0$ is the same as $(x - \bar{x}) = C_x C_x^+(x - \bar{x})$. This implies

$$\begin{aligned} C_{xv} &= E[(x - \bar{x})(v - \bar{v})'] \\ &= E[C_x C_x^+(x - \bar{x})(v - \bar{v})'] = C_x C_x^+ C_{xv} \end{aligned}$$

That is,

$$C_{xv} = C_x C_x^+ C_{xv}$$

Thus, we have

$$\begin{aligned} E[\tilde{v}] &= \bar{v} - C_{xv}' C_x^+ \bar{x} \\ C_{\tilde{v}} &= \text{cov}(v - C_{xv}' C_x^+ x) = C_v - C_{xv}' C_x^+ C_{xv} \\ C_{x\tilde{v}} &= \text{cov}(x, \tilde{v}) = C_{xv} - C_x C_x^+ C_{xv} = 0 \end{aligned}$$

That is, $y = \tilde{H}x + \tilde{v}$ is decoupled.

3.2 Decorrelation of Data Errors

We consider now decorrelating the components of the data error $\tilde{v}$ in the linear model $y = \tilde{H}x + \tilde{v}$ with

$$\text{cov}(x, \tilde{v}) = 0, \quad \text{cov}(\tilde{v}) \neq \text{diag}$$

This can be done by converting the model to

$$\check{y} = \check{H}x + \check{v} \quad (12)$$

where

$$\begin{aligned} \check{y} &= By, \quad \check{H} = B\tilde{H}, \quad \check{v} = B\tilde{v} \\ \check{C} &= \text{cov}(\check{v}) = B\text{cov}(\tilde{v})B' \end{aligned}$$

It is well known that any covariance (symmetric) matrix $\text{cov}(\tilde{v})$ can be diagonalized by an orthogonal matrix $B^{-1} = B'$. For example, we may possibly use $B = \{[\text{cov}(\tilde{v})]^{-1}\}^{1/2}$ and thus $\check{C} = I$ if $\text{cov}(\tilde{v})$ is nonsingular.

After such decorrelation, the observation errors v_k in the equivalent data model $\check{y}_k = \check{H}_k x_k + \check{v}_k$ are uncorrelated with each other and with the state x_k . Then, prior optimal fusion rules are applicable to this model directly and the results are optimal.

4 Fusion for Dynamic Systems

4.1 LMMSE: Static Case vs. Dynamic Case

Given data model $y = Hx + v$, the LMMSE estimator of x with prior information using y is

$$\hat{x} = E^*[x|y] = \bar{x} + C_{xy}C_y^+(y - \bar{y}) \quad (13)$$

This leads to the following batch estimator in the dynamic case:

$$\hat{x}_{k|k} = E^*[x_k|y^k] = \bar{x}_k + C_{x_k y^k}C_{y^k}^+(y^k - \bar{y}^k)$$

With few exceptions, it is unrealistic for most dynamic cases since its computational burden increases rapidly with time. A recursive estimator would be more useful. Replacing all priors with the corresponding predictions yields

$$\hat{x}_{k|k} = E^*[x_k|y_k, y^{k-1}] = \hat{x}_{k|k-1} + K_k \tilde{y}_{k|k-1} \quad (14)$$

where

$$K_k = C_{\hat{x}_{k|k-1} \tilde{y}_{k|k-1}} C_{\tilde{y}_{k|k-1}}^+, \quad \tilde{y}_{k|k-1} = y_k - E^*[y_k|y^{k-1}]$$

This estimator appears to be recursive (but in fact is not necessarily so).

One may argue that the only difference between priors and predictions is time—prediction $\hat{x}_{k|k-1}$ is to x_k , as prior mean $\bar{x}$ is to x —since prior information is simply “old” information (it comes only from past data). While we may or may not agree with this philosophically controversial view, nobody can deny the following mathematical fact: In LMMSE estimation prior mean $\bar{x}$ is nonrandom while prediction $\hat{x}_{k|k-1}$ is random. This has important implications. For example, $\bar{x}$ is uncorrelated with anything but this is not the case for $\hat{x}_{k|k-1}$. It is thus clear that (13) does not imply (14) (at least not directly).

In fact, (14) holds true because it can be shown (see, e.g., [9]) that LMMSE has the following *quasi-recursive* form

$$E^*[x|y_1, y_2] = \hat{x}_1 + C_{(x-\hat{x}_1)(y_2-\hat{y}_{2|1})} C_{(y_2-\hat{y}_{2|1})}^+ (y_2 - \hat{y}_{2|1}) \quad (15)$$

where

$$\hat{x}_1 = E^*[x|y_1], \quad \hat{y}_{2|1} = E^*[y_2|y_1]$$

(15) does imply (14) directly with $y_1 := y^{k-1}$ and $y_2 := y_k$. Note, however, that for data model $y_k = H_k x_k + v_k$,

$$\hat{y}_{k|k-1} = E^*[y_k|y^{k-1}] = H_k \hat{x}_{k|k-1} + \bar{v}_k$$

is not generally true, although it is exactly one of the Kalman filter formulas. Generally, we have

$$\hat{y}_{k|k-1} = E^*[y_k|y^{k-1}] = H_k \hat{x}_{k|k-1} + \hat{v}_{k|k-1} \quad (16)$$

where

$$\begin{aligned} \hat{v}_{k|k-1} &= E^*[v_k|y^{k-1}] \\ &= \bar{v}_k + C_{v_k y^{k-1}} C_{y^{k-1}}^+ (y^{k-1} - \bar{y}^{k-1}) \neq \bar{v}_k \end{aligned} \quad (17)$$

It follows from (16)–(17) that (14) is not always truly recursive because $\hat{v}_{k|k-1}$ in general needs all past data y^{k-1} directly. For systems that satisfy the Kalman filter assumption (i.e., process and measurement noises are white, mutually uncorrelated, and uncorrelated with the initial state), $C_{v_k y^{k-1}} = 0$ and thus $\hat{v}_{k|k-1} = \bar{v}_k$.

4.2 Further Discussion on Priors in Static Case vs. Predictions in Dynamic Case

For linear estimation of x using data y , the prior information consists of the prior means $\bar{x}$ and $\bar{y}$, and the prior covariances $C_x = \text{cov}(x)$, $C_y = \text{cov}(y)$, and $C_{xy} = \text{cov}(x, y)$. For linear data model $y = Hx + v$, the prior information consists of $\bar{x}$ and $\bar{v}$, and $C_x = \text{cov}(x)$, $C_v = \text{cov}(v)$, and $C_{xv} = \text{cov}(x, v)$. As shown in Part I [12], LMMSE estimator either does not exist or is meaningless if $\bar{v}$ or C_v (or $\bar{y}$ or C_y) is unknown. Consequently, by prior information we mean prior information about x , that is, $\bar{x}$, C_x , and possibly C_{xy} or C_{xv} .

For convenience of discussion, we define

$$E^*[x|y] = \text{BLUE with prior}$$

$$E^*[x_k|y^k] = (\text{batch}) \text{ BLUE with prior}$$

$$E^*[x_k|y_k, \hat{x}_{k|k-1}] = \text{recursive BLUE with prior using } \bar{x}_k, \text{cov}(x_k), \text{cov}(x_k, v_k)$$

$$E^+[x_k|y_k, \hat{x}_{k|k-1}] = \text{recursive BLUE without prior}$$

$$E^*[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] = \text{distributed BLUE with prior}$$

$$E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] = \text{distributed BLUE without prior}$$

where y_k is the data received from all sensors for estimating x_k , while y^{k-1} is the data received from all sensors for estimation of $x_1, \dots, x_{k-1}$ before. For standard distributed fusion, $y_k = [(\hat{x}_{k|k}^1)', \dots, (\hat{x}_{k|k}^n)']'$. Note that BLUE with prior is also known as linear MMSE (LMMSE) and BLUE without prior reduces to the optimal weighted least squares whenever the latter exists (see [10]).

Treating prior/prediction as data. As shown in [10], the following data model

$$\bar{x} = x + (\bar{x} - x)$$

captures all prior information for BLUE, meaning that $E^*[x|y] \equiv E^+[x|\bar{x}, y]$, that is, LMMSE estimator using data y is always identical to BLUE without prior using data $\bar{x}$ and y . As shown in [8], however, in general

$$E^*[x_k|y_k, y^{k-1}] \succeq E^*[x_k|y_k, \hat{x}_{k|k-1}]$$

$$\succeq E^+[x_k|y_k, \hat{x}_{k|k-1}] \succeq E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] \quad (18)$$

where $\hat{x}_1 \succeq \hat{x}_2$ denotes that $\hat{x}_1$ is never inferior and possibly superior to $\hat{x}_2$ in the BLUE sense. $E^*[x_k|y_k, y^{k-1}] \succeq E^*[x_k|y_k, \hat{x}_{k|k-1}]$ implies that the model

$$\hat{x}_{k|k-1} = x_k + (\hat{x}_{k|k-1} - x_k)$$

does not capture all information contained in y^{k-1} in general. This is a main difference between prior and prediction.

4.3 Optimal Recursive BLUE Fusers with and without Prior Information

While $E^*[x_k|y_k, \hat{x}_{k|k-1}]$ may be inferior to the globally optimum batch fuser $E^*[x_k|y_k, y^{k-1}]$, it is the best recursive fuser and recursion is a necessity for estimation fusion for almost all dynamic systems. It will never be worse than and may outperform the best recursive BLUE without prior $E^+[x_k|y_k, \hat{x}_{k|k-1}]$.

Let

$$\tilde{y}_k = \begin{bmatrix} \hat{x}_{k|k-1} \\ y_k \end{bmatrix} = \begin{bmatrix} x_k + (\hat{x}_{k|k-1} - x_k) \\ H_k x_k + v_k \end{bmatrix} = \tilde{H}_k x_k + \tilde{v}_k$$

where $\tilde{v}_k = [(\hat{x}_{k|k-1} - x_k)', v_k']'$ and

$$E[\tilde{v}_k] = \begin{bmatrix} 0 \\ \bar{v}_k \end{bmatrix}, \quad C_{\tilde{v}_k} = \begin{bmatrix} P_{k|k-1} & -C_{\tilde{x}_{k|k-1} v_k} \\ -C_{\tilde{x}_{k|k-1} v_k}' & \text{cov}(v_k) \end{bmatrix}$$

Then $E^*[x_k|y_k, \hat{x}_{k|k-1}]$ and $E^+[x_k|y_k, \hat{x}_{k|k-1}]$ for the dynamic case become static BLUE estimators of x_k with and without prior, respectively, using data $\tilde{y}_k$; that is, $E^*[x_k|\tilde{y}_k]$ and $E^+[x_k|\tilde{y}_k]$, respectively. Note that the Kalman filter is BLUE **without** prior $E^+[x_k|y_k, \hat{x}_{k|k-1}]$ since it uses y_k and $\hat{x}_{k|k-1}$ only without prior information about x_k directly.

It follows from Theorem 1 of [12, 11] that the optimal recursive BLUE with prior (LMMSE) fuser is given by

$$\begin{aligned} \hat{x}_{k|k}^1 &\triangleq E^*[x_k|y_k, \hat{x}_{k|k-1}] \\ &= \bar{x}_k + K_k(\tilde{y}_k - \tilde{H}_k \bar{x}_k - E[\tilde{v}_k]) \\ P_{k|k}^1 &= \text{MSE}(\hat{x}_{k|k}^1) = C_{x_k} - K_k C_{\tilde{y}_k} K_k' \\ K_k &= (C_{x_k} \tilde{H}_k' + C_{x_k} \tilde{v}_k) C_{\tilde{y}_k}^+ \\ C_{\tilde{y}_k} &= \tilde{H}_k C_{x_k} \tilde{H}_k' + C_{\tilde{v}_k} + \tilde{H}_k C_{x_k} \tilde{v}_k + (\tilde{H}_k C_{x_k} \tilde{v}_k)' \end{aligned}$$

where with a dynamic model similar to (3) at the center,

$$\bar{x}_k = F_{k,k-1} \bar{x}_{k-1} + \bar{w}_{k,k-1} \quad (19)$$

$$C_{x_k} = F_{k,k-1} C_{x_{k-1}} F_{k,k-1}' + Q_{k,k-1} \quad (20)$$

On the other hand, from Theorem 2 of [12, 11] it follows that

$$\begin{aligned} \hat{x}_{k|k}^2 &\triangleq E^+[x_k|y_k, \hat{x}_{k|k-1}] = \tilde{K}_k(\tilde{y}_k - E[\tilde{v}_k]) \\ P_{k|k}^2 &= \text{MSE}(\hat{x}_{k|k}^2) = \tilde{K}_k C_{\tilde{v}_k} \tilde{K}_k' \\ \tilde{K}_k &= \tilde{H}_k^+ [I - C_{\tilde{v}_k} (T_k C_{\tilde{v}_k} T_k)^+] \\ T_k &= I - \tilde{H}_k \tilde{H}_k^+ \end{aligned}$$

Corollary 4.3 of [10] gives the improvement of $\hat{x}_{k|k}^1$ over $\hat{x}_{k|k}^2$ by using prior information:

$$\begin{aligned} &[\text{MSE}(\hat{x}_{k|k}^1)]^{-1} - [\text{MSE}(\hat{x}_{k|k}^2)]^{-1} \\ &= (I + C_{x_k} \tilde{v}_k C_{\tilde{v}_k}^{-1} \tilde{H}_k') (C_{x_k} - C_{x_k} \tilde{v}_k C_{\tilde{v}_k}^{-1} C_{x_k} \tilde{v}_k)^{-1} \\ &\quad \cdot (I + C_{x_k} \tilde{v}_k C_{\tilde{v}_k}^{-1} \tilde{H}_k) \end{aligned}$$

if the inverses involved exist. There is no improvement if

$$I + C_{x_k} \tilde{v}_k C_{\tilde{v}_k}^{-1} \tilde{H}_k = 0$$

This condition is satisfied for systems satisfying the Kalman filter assumption, which agrees with the optimality of the Kalman filter and treating prediction as data.

It follows from Theorem 4.3 of [10] that $E^*[x_k|y_k, \hat{x}_{k|k-1}] = E^+[x_k|y_k, \hat{x}_{k|k-1}]$ holds (no need for $C_{\tilde{v}_k}^{-1}$ or $(C_{x_k} - C_{x_k} \tilde{v}_k C_{\tilde{v}_k}^{-1} C_{x_k} \tilde{v}_k)^{-1}$ to exist) if and only if any of the following conditions are satisfied

- The gain matrix $\tilde{K}_k$ of $E^+[x_k|y_k, \hat{x}_{k|k-1}]$ satisfies $\tilde{K}_k C_{\tilde{y}_k} = C_{x_k} \tilde{y}_k$.
- The gain matrix $\tilde{K}_k$ of $E^+[x_k|y_k, \hat{x}_{k|k-1}]$ satisfies $\tilde{K}_k C_{\tilde{v}_k} \tilde{y}_k = 0$.
- $(I - C_{\tilde{v}_k} \tilde{y}_k C_{\tilde{v}_k}^+) \tilde{H}_k$ has full column rank.
- The following condition holds

$$\begin{aligned} I - C_{x_k} \tilde{y}_k C_{\tilde{y}_k}^+ \tilde{H}_k &= [I - C_{x_k} \tilde{y}_k C_{\tilde{y}_k}^+ \tilde{H}_k] \\ &\quad \cdot [(I - C_{\tilde{y}_k} C_{\tilde{y}_k}^+) \tilde{H}_k]^+ [(I - C_{\tilde{y}_k} C_{\tilde{y}_k}^+) \tilde{H}_k] \end{aligned}$$

4.4 Optimality of Fusion Rules

It was shown in [14] that previous optimal linear fusion rules for the dynamic case are special cases of the fusion rules presented in [14], which was shown to be BLUE fusion rules **without** prior information. Thus, previous optimal linear fusion rules for the dynamic case are special cases of $E^+[x_k|y_k, \hat{x}_{k|k-1}]$ or $E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n]$. The question is now: On one hand, the previous rules have been shown to be globally optimum because they are identical to the centralized fusion. On the other hand, (18) suggests the possible existence of superior linear fusion rules. The answer is: The global optimality of the previous rules was shown only for systems satisfying the Kalman filter assumption, in which case we have

$$\begin{aligned} E^*[x_k|y_k, y^{k-1}] &= E^*[x_k|y_k, \hat{x}_{k|k-1}] \\ &= E^+[x_k|y_k, \hat{x}_{k|k-1}] = E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] \end{aligned}$$

It will be clear later that the Kalman filter is BLUE **without** prior $E^+[x_k|y_k, \hat{x}_{k|k-1}]$.

As explained in Part I [12], most fusion systems, such as the one presented in Section 2, do not satisfy the Kalman filter assumption. In fact, these systems satisfy neither the necessary and sufficient conditions for $E^*[x_k|y_k, y^{k-1}] = E^*[x_k|y_k, \hat{x}_{k|k-1}]$ to hold, derived in [8], nor the necessary and sufficient conditions for $E^*[x_k|y_k, y^{k-1}] = E^*[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n]$ to hold, presented in [9]. As a result, previous linear fusion rules, except for the BLUE with prior rules of Part I [12, 11], are not really globally optimum since $E^*[x_k|y_k, y^{k-1}] \succ E^*[x_k|y_k, \hat{x}_{k|k-1}] \succeq E^+[x_k|y_k, \hat{x}_{k|k-1}]$ for these systems. Globally optimum performance can be obtained for these systems directly from the BLUE rules of [12, 11] or by other optimal fusion rules using the decoupled data model of Section 3.

4.5 Optimality of Fusion Schemes

Two fusion schemes are commonly used: sensor to system and sensor to sensor. In the sensor-to-system fusion scheme, fused estimates (known as *system track* in target tracking) are maintained at the fusion center. Whenever data (say, y_k^i) from a sensor is received the system track is updated by fusing y_k^i with the estimate $\hat{x}_{k|k-1}$ of the state $x(t_k)$ at time t_k using received data $\{y_1, \dots, y_{k-1}\}$ through time t_{k-1} , where t_{k-1} is the previous update time. More generally, there is no need to update every time data is received. Instead, the system track can be updated after multiple pieces of data from one or more sensors are received in the interval $(t_{k-1}, t_k]$. This is a de facto recursive filtering scheme. It differs from the standard recursive filtering only in that each time the received data y_k^i may come from a different model and the data y_k^i itself may be an estimate $\hat{x}_{k|k}^i$ in distributed fusion.

In the sensor-to-sensor fusion scheme, the fusion center fuses data received from all sensors in the time interval $(t_{k-1}, t_k]$ without using any history information (e.g., $\hat{x}_{k|k-1}$), where t_{k-1} and t_k are the time instants for the previous fusion and current fusion, respectively. The sensor-to-sensor scheme is clearly more suitable for the standard distributed fusion than for the centralized fusion since history information is not used directly at the fusion center and is better contained in estimates rather than the observations.

While both schemes are natural, the sensor-to-system scheme appears more appealing in several aspects. With little need for data synchronization, it is more natural for the asynchronous case than the sensor-to-sensor scheme, which needs data synchronization. Also, history information is better and more explicitly accounted for in the sensor-to-system scheme than in the sensor-to-sensor scheme.

More refined versions of these schemes have been proposed, such as those based on the information graph [4, 3] and tracklets [6, 5].

Two related, important, theoretical questions are: Are optimally fused estimates within these two schemes globally optimum within the linear class? If not, under what conditions are they globally optimum? We answer these questions now for BLUE fusion.

As explained above, the sensor-to-system fusion actually amounts to recursive filtering, given by either $E^*[x_k|y_k, \hat{x}_{k|k-1}]$ or $E^+[x_k|y_k, \hat{x}_{k|k-1}]$. Then the two questions become: Is optimal linear recursive filtering $E^*[x_k|y_k, \hat{x}_{k|k-1}]$ or $E^+[x_k|y_k, \hat{x}_{k|k-1}]$ identical to the optimum linear batch filtering $E^*[x_k|y_k, y^{k-1}]$? If not, under what condition are they identical? We have answered these questions in the previous subsection. Simply put, optimal linear recursive filtering is globally optimum if and only if some necessary and sufficient conditions are satisfied, which can be found in [8] but are beyond the scope of this paper.² This implies that the sensor-to-system fu-

sion is in general not globally optimum, because the data model [e.g., (6)] in general does not satisfy the condition required for global optimality due to the correlation of the data errors across sensors and with the state. It is, however, globally optimum if the equivalent decorrelated data model (12) is used.

Since use of the sensor-to-sensor scheme with the centralized fusion does not appear reasonable, we consider only the standard distributed fusion, which is actually $E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n]$. It should be clear from previous discussion that

$$\begin{aligned} E^*[x_k|y_k, \hat{x}_{k|k-1}] &\succeq E^+[x_k|y_k, \hat{x}_{k|k-1}] \\ &\succeq E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] \\ E^*[x_k|y_k, \hat{x}_{k|k-1}] &\succeq E^*[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] \\ &\succeq E^+[x_k|\hat{x}_{k|k}^1, \dots, \hat{x}_{k|k}^n] \end{aligned}$$

Consequently, the sensor-to-sensor scheme usually has inferior performance relative to the sensor-to-system fusion; use of prior information about x_k as in (19)–(20) may possibly improve performance of both schemes.

This discussion reveals that filtering as well as prediction and smoothing are special fusion problems. Their batch versions are centralized fusion, while their recursive versions are the corresponding sensor-to-system centralized fusion.

4.6 Effect of Feedback

Consider the case where fused estimates are sent back to each sensor such that, say, $\hat{x}_{k|k}^i := \hat{x}_{k|k}$. This has effect on both sensor estimates and fused estimates at the next time.

For sensor i , fusing $\hat{x}_{k|k}^i := \hat{x}_{k|k}$ and observations made after t_k is a standard filtering problem. But this amounts to a common reinitialization of all sensor estimators. Their subsequent estimation errors are thus correlated further by this common reinitialization and should be accounted for in the fusion rules if better performance is desired.

It is also possible that sensor i has information from measurements $z_{p_1}^i, \dots, z_{p_m}^i$ made at time $t_{p_1}, \dots, t_{p_m} < t_k$ respectively that is not accounted for by the fused estimate $\hat{x}_{k|k}$ from the fusion center. How to fuse $\hat{x}_{k|k}$ and $z_p = [(z_{p_1}^i)', \dots, (z_{p_m}^i)']'$ is a problem of update with out-of-sequence-measurements. It differs from the standard Kalman filter update due to the coupling between the measurement error v_p and estimation error $x_k - \hat{x}_{k|k}$. Two computationally efficient, general algorithms that yield the global optimum $E^*[x_k|z_p, z^k]$ and recursive BLUE without prior $E^+[x_k|z_p, \hat{x}_{k|k}]$ are available from [13] and several other algorithms were presented elsewhere (see references of [13]). Alternatively, an algorithm that yields recursive

filter (i.e., the process and measurement noises are white, uncorrelated with each other and with the initial state). Here the whiteness condition can be relaxed to a certain type of Markov condition and the uncorrelatedness condition can also be relaxed to certain limited cases of correlation.

²A well-known sufficient condition is the one assumed by the Kalman

BLUE with prior $E^*[x_k|z_p, \hat{x}_{k|k}]$ can be easily obtained based on the discussion in the previous subsections.

5 Summary

We have presented the following:

- A general model for discretized asynchronous multisensor systems.
- A general technique that annihilates all correlation of the data errors.
- General recursive BLUE fusers with and without prior for dynamic systems.

We have shown or discussed the following:

- The data errors in the above model presented are correlated with the state and process noise, correlated across sensors, and colored.
- The Kalman filter is in essence a recursive BLUE estimator *without* prior at each time, which is in general not the globally optimum (batch or recursive) LMMSE estimator. It becomes globally optimum under the standard assumption that the process and measurement noises are white, mutually uncorrelated, and uncorrelated with the initial state.
- For dynamic systems, it is better to use the recursive LMMSE fuser (i.e., BLUE with prior) at each time, which is easily implementable and may outperform all previous Kalman filter based “optimal” fusers.
- Previous “optimal” linear fusion rules are globally optimum in the linear class only for systems satisfying KF assumption.
- Predictions in the dynamic case may or may not correspond to prior information in the static case.
  - For batch LMMSE (i.e., BLUE with prior), they correspond to each other exactly.
  - For recursive BLUE (with or without prior), they do not correspond to each other in general, but they do for systems for which the Kalman filter is LMMSE.
- Necessary and sufficient conditions are available for a batch LMMSE to have a recursive form, for BLUE with and without prior to be identical, and for a BLUE (with or without prior) centralized and distributed fusers to be identical.
- Sensor-to-system centralized fusion amounts to recursive filtering. Its optimality relies on that of the centralized Kalman filter for the same system.
- Sensor-to-sensor (distributed) fusion is in essence BLUE without prior and has an inferior performance than the sensor-to-system fusion in general.
- Feedback provides a common reinitialization and introduces further coupling among local estimates. It may lead to out-of-sequence measurement problems with the local estimation.

References

- [1] Y. Bar-Shalom. Update with Out-of-Sequence Measurements in Tracking: Exact Solution. In *Proc. 2000 SPIE Conf. on Signal and Data Processing of Small Targets*, vol. 4048, Apr. 2000.
- [2] Y. Bar-Shalom and X. R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*. YBS Publishing, Storrs, CT, 1995.
- [3] C. Y. Chong, S. Mori, W. H. Barker, and K. C. Chang. Architectures and Algorithms for Track Association and Fusion. *IEEE AES Systems Magazine*, pages 5–13, Jan. 2000.
- [4] C. Y. Chong, S. Mori, and K. C. Chang. Distributed Multitarget Multisensor Tracking. In Y. Bar-Shalom, editor, *Multitarget-Multisensor Tracking: Advanced Applications*, pages 247–295. Artech House, Norwood, MA, 1990.
- [5] O. E. Drummond. A Hybrid Sensor Fusion Algorithm Architecture and Tracklets. In *Proc. 1997 SPIE Conf. on Signal and Data Processing of Small Targets*, vol. 3163, pages 512–524, 1997.
- [6] O. E. Drummond. Tracklets and a Hybrid Fusion With Process Noise. In *Proc. 1997 SPIE Conf. on Signal and Data Processing of Small Targets*, vol. 3163, pages 512–524, 1997.
- [7] C. Z. Han, J. Wang, and X. R. Li. Fused State Estimation of Linear Continuous-Time System with Multisensor Asynchronous Measurements. In *Proc. 2001 International Conf. on Information Fusion*, pages FrB1.3–FrB1.9, Montreal, QC, Canada, Aug. 2001.
- [8] X. R. Li. Applied Estimation and Filtering. Course Notes, 2003.
- [9] X. R. Li and K. S. Zhang. Optimal Linear Estimation Fusion—Part IV: Optimality and Efficiency of Distributed Fusion. In *Proc. 2001 International Conf. on Information Fusion*, pages WeB1.19–WeB1.26, Montreal, QC, Canada, Aug. 2001.
- [10] X. R. Li, K.-S. Zhang, J. Zhao, and Y.-M. Zhu. Optimal Linear Estimation Fusion—Part V: Relationships. In *Proc. 2002 International Conf. on Information Fusion*, pages 497–504, Annapolis, MD, USA, July 2002.
- [11] X. R. Li, Y. M. Zhu, and C. Z. Han. Unified Optimal Linear Estimation Fusion—Part I: Unified Models and Fusion Rules. In *Proc. 2000 International Conf. on Information Fusion*, pages MoC2.10–MoC2.17, Paris, France, July 2000.
- [12] X. R. Li, Y. M. Zhu, J. Wang, and C. Z. Han. Optimal Linear Estimation Fusion—Part I: Unified Fusion Rules. *IEEE Trans. Information Theory* (to appear), 49, 2003.
- [13] K.-S. Zhang, X. R. Li, and Y.-M. Zhu. Optimal Update with Out-of-Sequence Observations for Distributed Filtering. In *Proc. 2002 International Conf. on Information Fusion*, pages 1519–1526, Annapolis, MD, USA, July 2002.
- [14] Y. M. Zhu and X. R. Li. Best Linear Unbiased Estimation Fusion. In *Proc. 1999 International Conf. on Information Fusion*, Sunnyvale, CA, USA, July 1999.