

On Optimal State Estimation with Multiple Packet Dropouts*

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

Abstract – *Due to unreliable communication between local sensors and the processing center, packet dropouts may happen during transmission. Two existing methods for linear minimum mean-squared error (LMMSE) estimation with multiple packet dropouts were obtained completely or partially based on a stochastic parameter system constructed by augmenting the original state and measurement. They have a high computational load, unclear measurement residual characterization and tough requirements on initialization. To overcome these, an alternative form of LMMSE estimation with multiple packet dropouts is derived. Under a Gaussian assumption, the minimum mean-squared error (MMSE) estimation with multiple packet dropouts is also derived. Numerical examples are provided to compare performance of the proposed estimators.*

Keywords: State estimation, LMMSE, MMSE, packet dropout, hard decision.

1 Introduction

With the emerging of sensor networks, traditional estimation problems are facing new challenges. For example, due to unreliable communication between local sensors and the processing center, packet transmission delay [1, 2, 3, 4, 5] and multiple packet dropouts [6, 7, 8, 9, 10] are usually inevitable. Also, the constraints on communication bandwidth, power consumption [11, 12] and computational capability should be considered, which make the research on distributed estimation [13] and estimation problems with compressed [14, 15] or quantized data [16, 17] necessary.

In this paper, we deal with state estimation in the presence of multiple packet dropouts. As in [6, 7, 8, 9, 10], by multiple packet dropouts, it is meant that the received data is either the current raw measurement or the last received data in a probabilistic manner. If the time stamp for the corresponding raw measurement is available for each received packet, then this estimation problem reduces to one

with intermittent measurements [18] or missing data [19] by comparing the time-stamp of the received packet with the current time. Correspondingly, existing methods can be applied directly. What may make the problem harder is that the time-stamp information may not be available in some cases (see the formulation in [6, 7, 8, 9, 10]) and what is known to the estimator is only the data arrival probability.

In [8, 9, 10], in order to have a similar form as the system used by the Kalman filter, the original system with multiple packet dropouts is converted to a stochastic parameter system by augmentation. Then by defining the stochastic \mathcal{H}_2 -norm of the system with stochastic parameter, the optimal \mathcal{H}_2 filters were designed through the linear matrix inequality approach. In [7], by the innovations analysis approach, the LMMSE estimation with multiple packet dropouts was obtained completely from the same stochastic parameter system as in [8, 9, 10]. In [6], for a time-invariant system, by designating the optimal estimator to be of a specific linear form and from the unbiasedness and minimum error covariance properties of the optimal estimator, the associated co-efficient matrices were obtained. This sequential LMMSE estimator is based partially on the same stochastic parameter system as in [8, 9, 10]. The characterization of its measurement residual part is not clear. The relying on a higher dimensional stochastic parameter system increases the computational load of the estimators of [7] and [6]. Also, both of their optimal initializations depend on information about the past two received data before the first physically received data. This dependency seems to be too much for real implementation.

To overcome the shortcomings of the existing LMMSE estimation methods with multiple packet dropouts, an alternative form is derived first in this paper. Then under a Gaussian assumption, the MMSE estimation is also derived by hard decision, which without the Gaussian assumption is LMMSE with respect to (w.r.t.) the raw measurement sequence without the unarrived packets. This MMSE estimation method has two nice properties. First, unlike the proposed LMMSE estimation method, its performance does not depend on the value of the initial data at the processing cen-

*Research supported in part by NSFC grant 60602026, ARO through grant W911NF-08-1-0409, ONR-DEPSCoR through grant N00014-09-1-1169, and NAVO through Contract # N62306-09-P-3S01.

ter. Second, it does not depend on the probability of the data arrival events. Numerical examples are provided to compare performance of the proposed estimators.

The paper is organized as follows. Sec. 2 formulates the problem. Sec. 3 summarizes two existing forms of LMMSE estimation with multiple packet dropouts. Sec. 4 derives an alternative form of the LMMSE estimation. Sec. 5 discusses the MMSE estimation with multiple packet dropouts and its further relaxation. Sec. 6 provides numerical examples to compare performance of the proposed estimators. Sec. 7 gives conclusions.

2 Problem formulation

Consider the following generic dynamic system:

$$x_k = F_{k-1}x_{k-1} + G_{k-1}w_{k-1} \quad (1)$$

where $x_k \in \mathcal{R}^n$, $\langle w_k \rangle$ is zero-mean white noise with covariance $Q_k \geq 0$, $E[x_0] = \bar{x}_0$, $\text{cov}(x_0) = P_0$, and x_0 is uncorrelated with $\langle w_k \rangle$.

The raw measurement observed by a generic local sensor is given by

$$z_k = H_k x_k + v_k, \quad k = 1, 2, \dots \quad (2)$$

where $z_k \in \mathcal{R}^m$, $\langle v_k \rangle$ is zero-mean white noise with covariance $R_k \geq 0$, and $\langle v_k \rangle$ is uncorrelated with x_0 and $\langle w_k \rangle$.

Instead of estimating the state locally by using $\langle z_k \rangle$ directly, they are transmitted to an estimator through an unreliable network where packet dropouts are possible. It is assumed that the data received by the estimator can be modeled by:

$$y_k = \gamma_k z_k + (1 - \gamma_k) y_{k-1} \quad (3)$$

where γ_k has Bernoulli distribution with $P\{\gamma_k = 1\} = p_k$ and $P\{\gamma_k = 0\} = 1 - p_k = q_k$, γ_k is uncorrelated with all the other random variables, and $\langle \gamma_k \rangle$ is a white sequence.

Remark: The data arrival probability p_k is a measure of the reliability and transmission quality of the network. In [6, 7, 8, 9, 10], it was assumed that p_k is known to the estimator. But as shown later, there is no need to make this assumption for some estimators.

Remark: The assumption that γ_k is uncorrelated with all the other random variables seems reasonable for many situations. But as shown later, whether this assumption is valid does not matter for some estimators.

Remark: It is assumed that y_{k-1} is still available when y_k is received. This is also the case in [6, 7, 8, 9, 10].

In this paper, given only the first two moments, we first try to obtain the LMMSE state estimation with multiple packet dropouts. That is,

$$\hat{x}_{k|k}^{\text{LMMSE}} \triangleq E^*[x_k|y^k] = \arg \min_{\hat{x}_{k|k} = a_k + B_k Y_k} \text{MSE}(\hat{x}_{k|k})$$

where

$$y^k = \{y_1, \dots, y_k\}, \quad Y_k = [y_1', \dots, y_k']'$$

$$\text{MSE}(\hat{x}_{k|k}) = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})']$$

and a_k, B_k do not depend on Y_k .

Also we try to obtain the MMSE state estimation with multiple packet dropouts

$$\hat{x}_{k|k}^{\text{MMSE}} = E[x_k|y^k]$$

for some special cases.

3 Summary of two existing forms of LMMSE estimation [6, 7]

First, in order to have a similar form as the system used by the Kalman filter, the original system (1), (2) and (3) is converted to the following stochastic parameter system by augmentation:

$$X_{k+1} = \tilde{F}_k X_k + \tilde{G}_k W_k \quad (4)$$

$$y_k = \tilde{H}_k X_k + \gamma_k v_k \quad (5)$$

where

$$X_k = [x_k' \quad y_{k-1}']' \in \mathcal{R}^{n+m}, \quad W_k = [w_k' \quad v_k']'$$

$$\tilde{F}_k = \begin{bmatrix} F_k & 0 \\ \gamma_k H_k & (1 - \gamma_k) I_m \end{bmatrix}$$

$$\tilde{G}_k = \begin{bmatrix} G_k & 0 \\ 0 & \gamma_k I_m \end{bmatrix}, \quad \tilde{H}_k = [\gamma_k H_k \quad (1 - \gamma_k) I_m]$$

Under the given assumption and from (4), it follows that

$$c_{k+1} = \bar{F}_k c_k \bar{F}_k' \quad (6)$$

$$+ p_k q_k \begin{bmatrix} 0 & 0 \\ H_k & -I_m \end{bmatrix} c_k \begin{bmatrix} 0 & 0 \\ H_k & -I_m \end{bmatrix}' + U_k$$

where

$$c_k = E[X_k X_k'], \quad \bar{F}_k = E[\tilde{F}_k] = \begin{bmatrix} F_k & 0 \\ p_k H_k & q_k I_m \end{bmatrix}$$

$$U_k = \begin{bmatrix} G_k Q_k G_k' & 0 \\ 0 & p_k R_k \end{bmatrix}$$

and it is assumed that

$$c_0 = \begin{bmatrix} P_0 + \bar{x}_0 \bar{x}_0' & 0 \\ 0 & 0 \end{bmatrix}$$

In [7], for the augmented stochastic parameter system (4) and (5), the sequential LMMSE estimate of X_{k+1} was obtained as follows through the innovations approach.

Prediction:

$$\hat{X}_{k+1|k} = \bar{F}_k \hat{X}_{k|k}$$

$$\text{MSE}(\hat{X}_{k+1|k}) = p_k q_k \begin{bmatrix} 0 & 0 \\ H_k & -I_m \end{bmatrix} c_k \begin{bmatrix} 0 & 0 \\ H_k & -I_m \end{bmatrix}'$$

$$+ \bar{F}_k \text{MSE}(\hat{X}_{k|k}) \bar{F}_k' + U_k$$

Update:

$$\begin{aligned}\hat{X}_{k+1|k+1} &= \hat{X}_{k+1|k} + K_{k+1}\varepsilon_{k+1} \\ \varepsilon_{k+1} &= y_{k+1} - \bar{H}_{k+1}\hat{X}_{k+1|k} \\ K_{k+1} &= \text{MSE}(\hat{X}_{k+1|k})\bar{H}'_{k+1}S_{k+1}^{-1} \\ \bar{H}_{k+1} &= E[\tilde{H}_{k+1}] = [p_{k+1}H_{k+1} \quad q_{k+1}I_m] \\ S_{k+1} &= p_{k+1}q_{k+1} \begin{bmatrix} H_{k+1} & -I_m \\ H_{k+1} & -I_m \end{bmatrix} c_{k+1} \\ &\quad \cdot \begin{bmatrix} H_{k+1} & -I_m \end{bmatrix}' \\ &\quad + \bar{H}_{k+1}\text{MSE}(\hat{X}_{k+1|k})\bar{H}'_{k+1} + p_{k+1}R_{k+1}\end{aligned}$$

$$\text{MSE}(\hat{X}_{k+1|k+1}) = \text{MSE}(\hat{X}_{k+1|k}) - K_{k+1}S_{k+1}K'_{k+1}$$

with the initial conditions

$$\hat{X}_{0|-1} = [\bar{x}'_0 \quad 0]', \quad \text{MSE}(\hat{X}_{0|-1}) = \begin{bmatrix} P_0 & 0 \\ 0 & 0 \end{bmatrix}$$

Finally,

$$\begin{aligned}\hat{x}_{k+1|k} &= [I_n \quad 0] \hat{X}_{k+1|k} \\ P_{k+1|k} &= [I_n \quad 0] \text{MSE}(\hat{X}_{k+1|k}) [I_n \quad 0]' \\ \hat{x}_{k+1|k+1} &= [I_n \quad 0] \hat{X}_{k+1|k+1} \\ P_{k+1|k+1} &= [I_n \quad 0] \text{MSE}(\hat{X}_{k+1|k+1}) [I_n \quad 0]'\end{aligned}$$

In [6], it was further assumed that

$$\begin{aligned}F_k &= F, \quad G_k = G, \quad H_k = H, \quad Q_k = Q, \quad R_k = R \\ p_k &= p, \quad q_k = q = 1 - p\end{aligned}$$

By designating the optimal of x_{k+1} to be of the form

$$\hat{x}_{k+1|k+1} = \Phi_k \hat{x}_{k|k} + K_k^1 y_{k+1} + K_k^2 y_k \quad (7)$$

and from the unbiasedness and minimum error covariance properties of $\hat{x}_{k+1|k+1}$, it was found that

$$\begin{aligned}\Phi_k &= (I_n - pK_k^1 H)F \\ K_k^1 &= P_{k+1|k} H' \Lambda_k^{-1} \\ K_k^2 &= -qK_k^1 \\ P_{k+1|k} &= F P_{k|k} F' + G Q G' \\ \Lambda_k &= p(H P_{k+1|k} H' + R) + q H G Q G' H' \\ &\quad + q [pH - HF \quad qI_m] c_k [pH - HF \quad qI_m]' \\ &\quad + p q^2 [H \quad -I_m] c_k [H \quad -I_m]' + (1 - p^2)R \\ P_{k+1|k+1} &= P_{k+1|k} - pK_k^1 \Lambda_k (K_k^1)'\end{aligned}$$

The one-step ahead prediction is

$$\hat{x}_{k+1|k} = F \hat{x}_{k|k}$$

with the initial conditions

$$\hat{x}_{0|0} = \bar{x}_0, \quad P_{0|0} = P_0$$

Remark: As can be clearly seen from the above, the LMMSE estimator of [7] is purely and the LMMSE estimator of [6] is partially based on the augmented stochastic

parameter form (4) and (5) of the original system. Compared with the original system described by (1) – (3), the dimension of the state vector is increased from n to $n + m$, so the computational complexity is increased also.

Remark: The final form (7) of the LMMSE estimator in [6] is not similar to that of the Kalman filter, so it is not easy to analyze the measurement residual of this estimator. For example, how big the measurement residual is and what about its statistical characteristics are not clear. As we know, information about the measurement residual is very helpful for some hypothesis testing problem, e.g., data association in target tracking.

Remark: To initialize the LMMSE estimator in [6], (besides $\hat{x}_{0|0} = \bar{x}_0$ and $P_{0|0} = P_0$) y_0 as well as $E[x_0 y'_{-1}]$ and $E[y_{-1} y'_{-1}]$, both assumed zero therein, is needed. To initialize the LMMSE estimator in [7], (besides $\hat{x}_{0|-1} = \bar{x}_0$ and $P_{0|-1} = P_0$) $E[x_0 y'_{-1}]$, $E[y_{-1} y'_{-1}]$, $\hat{y}_{-1|-1}$, $\text{MSE}(\hat{y}_{-1|-1})$, $E[(x_0 - \hat{x}_{0|-1})(y_{-1} - \hat{y}_{-1|-1})']$, all assumed zero therein, and y_0 are needed.

4 An alternative form of LMMSE estimation

From the appendix of [20], it is easy to verify the following two lemmas for LMMSE estimation.

Lemma 1 For a scalar-valued γ , if γ is uncorrelated with x and z , then

$$E^*[\gamma x | z] = E[\gamma] E^*[x | z]$$

Lemma 2 For the LMMSE estimate \hat{x} of x , we have

$$E[\hat{x} \hat{x}'] = E[x x'] - \text{MSE}(\hat{x})$$

With these two lemmas, besides the two forms of LMMSE estimation obtained in [6] and [7], an alternative form is given in Theorem 1.

Theorem 1 (LMMSE estimation). Given p_k , $\hat{x}_{k-1|k-1} = E^*[x_{k-1} | y^{k-1}]$, $P_{k-1|k-1} = \text{MSE}(\hat{x}_{k-1|k-1})$, an alternative form of the LMMSE estimate of x_k for system (1) – (3) is:

Prediction:

$$\begin{aligned}\hat{x}_{k|k-1} &= E^*[x_k | y^{k-1}] = F_{k-1} \hat{x}_{k-1|k-1} \\ P_{k|k-1} &= \text{MSE}(\hat{x}_{k|k-1}) \\ &= F_{k-1} P_{k-1|k-1} F'_{k-1} + G_{k-1} Q_{k-1} G'_{k-1}\end{aligned}$$

Update¹:

$$\begin{aligned}\hat{x}_{k|k} &= E^*[x_k | y^k] \\ &= \hat{x}_{k|k-1} + C_{\hat{x}_{k|k-1}, \tilde{y}_{k|k-1}} C_{\tilde{y}_{k|k-1}}^+ \tilde{y}_{k|k-1} \\ P_{k|k} &= \text{MSE}(\hat{x}_{k|k}) \\ &= P_{k|k-1} - C_{\hat{x}_{k|k-1}, \tilde{y}_{k|k-1}} C_{\tilde{y}_{k|k-1}}^+ C_{\hat{x}_{k|k-1}, \tilde{y}_{k|k-1}}' \\ \tilde{y}_{k|k-1} &= y_k - p_k H_k \hat{x}_{k|k-1} - q_k y_{k-1}\end{aligned}$$

$$C_{\hat{x}_{k|k-1}, \tilde{y}_{k|k-1}} = p_k P_{k|k-1} H'_k$$

¹In the following A^+ stands for the unique *MP inverse* of matrix A . It reduces to A^{-1} whenever A^{-1} exists.

$$\begin{aligned}
C_{\tilde{y}_{k|k-1}} &= p_k^2 H_k P_{k|k-1} H_k' + p_k R_k + p_k q_k \\
&\quad \cdot (H_k C_k H_k' - H_k E_k - E_k' H_k' + D_{k-1}) \\
C_k &= F_{k-1} C_{k-1} F_{k-1}' + G_{k-1} Q_{k-1} G_{k-1}' \\
D_{k-1} &= p_{k-1} (H_{k-1} C_{k-1} H_{k-1}' + R_{k-1}) \\
&\quad + q_{k-1} D_{k-2} \\
E_k &= p_{k-1} F_{k-1} C_{k-1} H_{k-1}' + q_{k-1} F_{k-1} E_{k-1}
\end{aligned}$$

with the initial conditions

$$\begin{aligned}
\hat{x}_{0|0} &= \bar{x}_0, P_{0|0} = P_0, C_0 = P_0 + \bar{x}_0 \bar{x}_0' \\
D_0 &= E[y_0 y_0'], E_1 = F_0 E[x_0 y_0'] + G_0 E[w_0 y_0']
\end{aligned}$$

Proof:

Given $\hat{x}_{k-1|k-1}$ and $P_{k-1|k-1}$, it follows easily from the property of LMMSE estimation that

$$\begin{aligned}
\hat{x}_{k|k-1} &= E^*[x_k | y^{k-1}] \\
&= E^*[F_{k-1} x_{k-1} + G_{k-1} w_{k-1} | y^{k-1}] \\
&= F_{k-1} \hat{x}_{k-1|k-1} \\
\tilde{x}_{k|k-1} &= x_k - \hat{x}_{k|k-1} = F_{k-1} \tilde{x}_{k-1|k-1} + G_{k-1} w_{k-1} \\
P_{k|k-1} &= \text{MSE}(\hat{x}_{k|k-1}) \\
&= F_{k-1} P_{k-1|k-1} F_{k-1}' + G_{k-1} Q_{k-1} G_{k-1}'
\end{aligned}$$

The LMMSE estimator $E^*[x_k | y^k]$ always has the following quasi-recursive form [21]

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

From Lemma 1 above, we have

$$\begin{aligned}
\tilde{y}_{k|k-1} &= y_k - E^*[y_k | y^{k-1}] \\
&= y_k - E^*[\gamma_k z_k + (1 - \gamma_k) y_{k-1} | y^{k-1}] \\
&= y_k - E[\gamma_k] E^*[H_k x_k + v_k | y^{k-1}] \\
&\quad - E[1 - \gamma_k] E^*[y_{k-1} | y^{k-1}] \\
&= y_k - p_k H_k \hat{x}_{k|k-1} - q_k y_{k-1}
\end{aligned}$$

Furthermore, $\tilde{y}_{k|k-1}$ can be rewritten as

$$\begin{aligned}
\tilde{y}_{k|k-1} &= y_k - p_k H_k \hat{x}_{k|k-1} - q_k y_{k-1} \\
&= \gamma_k (H_k x_k + v_k) + (1 - \gamma_k) y_{k-1} \\
&\quad - p_k H_k \hat{x}_{k|k-1} - (1 - p_k) y_{k-1} \\
&= \gamma_k H_k x_k - p_k H_k \hat{x}_{k|k-1} + \gamma_k v_k + (p_k - \gamma_k) y_{k-1} \\
&= \gamma_k H_k x_k - \gamma_k H_k \hat{x}_{k|k-1} + \gamma_k H_k \hat{x}_{k|k-1} \\
&\quad - p_k H_k \hat{x}_{k|k-1} + \gamma_k v_k + (p_k - \gamma_k) y_{k-1} \\
&= \gamma_k H_k \tilde{x}_{k|k-1} + (\gamma_k - p_k) H_k \hat{x}_{k|k-1} + \gamma_k v_k \\
&\quad + (p_k - \gamma_k) y_{k-1}
\end{aligned}$$

By the principle of orthogonality, $\tilde{x}_{k|k-1}$ is orthogonal to $\hat{x}_{k|k-1}$ and y_{k-1} . Thus

$$C_{\tilde{x}_{k|k-1}, \tilde{y}_{k|k-1}} = \text{cov}(\tilde{x}_{k|k-1}, \tilde{y}_{k|k-1}) = p_k P_{k|k-1} H_k'$$

$$\begin{aligned}
C_{\tilde{y}_{k|k-1}} &= \text{cov}(\tilde{y}_{k|k-1}) = E[\tilde{y}_{k|k-1} \tilde{y}_{k|k-1}'] \\
&= E[\gamma_k^2] (H_k P_{k|k-1} H_k' + R_k) \\
&\quad + E[(\gamma_k - p_k)^2] (H_k E[\hat{x}_{k|k-1} \hat{x}_{k|k-1}'] H_k' \\
&\quad - H_k E[\hat{x}_{k|k-1} y_{k-1}'] \\
&\quad - E[y_{k-1} \hat{x}_{k|k-1}'] H_k' + E[y_{k-1} y_{k-1}']) \\
&= p_k (H_k P_{k|k-1} H_k' + R_k) \\
&\quad + (p_k - p_k^2) (H_k E[\hat{x}_{k|k-1} \hat{x}_{k|k-1}'] H_k' \\
&\quad - H_k E[(x_k - \tilde{x}_{k|k-1}) y_{k-1}'] \\
&\quad - E[y_{k-1} (x_k - \tilde{x}_{k|k-1})'] H_k' + E[y_{k-1} y_{k-1}']) \\
&= p_k (H_k P_{k|k-1} H_k' + R_k) + (p_k - p_k^2) \\
&\quad \cdot (H_k (C_k - P_{k|k-1}) H_k' - H_k E_k - E_k' H_k' + D_{k-1}) \\
&= p_k^2 H_k P_{k|k-1} H_k' + p_k R_k + p_k q_k \\
&\quad \cdot (H_k C_k H_k' - H_k E_k - E_k' H_k' + D_{k-1})
\end{aligned}$$

where we have used

$$E[\hat{x}_{k|k-1} \hat{x}_{k|k-1}'] = E[x_k x_k'] - P_{k|k-1} = C_k - P_{k|k-1}$$

$$\begin{aligned}
C_k &\triangleq E[x_k x_k'] \\
&= E[(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})'] \\
&= F_{k-1} C_{k-1} F_{k-1}' + G_{k-1} Q_{k-1} G_{k-1}'
\end{aligned}$$

$$\begin{aligned}
D_k &\triangleq E[y_k y_k'] \\
&= E[(\gamma_k z_k + (1 - \gamma_k) y_{k-1})(\gamma_k z_k + (1 - \gamma_k) y_{k-1})'] \\
&= E[\gamma_k^2 z_k z_k'] + E[\gamma_k (1 - \gamma_k) z_k y_{k-1}'] + E[\gamma_k (1 - \gamma_k) y_{k-1} z_k'] \\
&\quad + E[(1 - \gamma_k)^2 y_{k-1} y_{k-1}'] \\
&= p_k E[z_k z_k'] + q_k D_{k-1} \\
&= p_k E[(H_k x_k + v_k)(H_k x_k + v_k)'] + q_k D_{k-1} \\
&= p_k (H_k C_k H_k' + R_k) + q_k D_{k-1}
\end{aligned}$$

$$\begin{aligned}
E_k &\triangleq E[x_k y_{k-1}'] \\
&= E[(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})(\gamma_{k-1} z_{k-1} \\
&\quad + (1 - \gamma_{k-1}) y_{k-2})'] \\
&= E[(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})(\gamma_{k-1} (H_{k-1} x_{k-1} + v_{k-1}) \\
&\quad + (1 - \gamma_{k-1}) y_{k-2})'] \\
&= p_{k-1} F_{k-1} C_{k-1} H_{k-1}' + (1 - p_{k-1}) F_{k-1} E_{k-1}
\end{aligned}$$

with the initial conditions

$$\begin{aligned}
C_0 &= E[x_0 x_0'] = \text{cov}(x_0) + E[x_0] E[x_0] = P_0 + \bar{x}_0 \bar{x}_0' \\
D_0 &= E[y_0 y_0'] \\
E_1 &= E[x_1 y_0'] = E[(F_0 x_0 + G_0 w_0) y_0'] \\
&= F_0 E[x_0 y_0'] + G_0 E[w_0 y_0']
\end{aligned}$$

□

Remark: It can be easily seen that when $p_k = 1$, this LMMSE estimator reduces to the Kalman filter.

Remark: It can be clearly seen that to calculate $\hat{x}_{k|k}$, we do need both y_k and y_{k-1} . That is why we assume that y_{k-1} is still available when y_k is received in the problem formulation part. This is similar to what is done in difference measurement method [22, 19] for estimation under autocorrelated measurement noise.

Remark: Since two forms of the LMMSE estimation with multiple packet dropouts are already available in [6] and [7], why derive still another one? There are three sources of motivation. The first is from computational consideration, the second is from the angle of measurement residual characterization and the third is from the initialization perspective. Compared with the two existing forms, a system of a higher dimension like the one in (4) and (5) and matrix operation with a higher dimension like the one in (6) are never used, so the computational burden is reduced. The LMMSE estimator of Theorem 1 is in a similar form as the Kalman filter and its measurement residual is well characterized by $\tilde{y}_{k|k-1}$, $C_{\tilde{x}_{k|k-1}, \tilde{y}_{k|k-1}}$ and $C_{\tilde{y}_{k|k-1}}$. To initialize the LMMSE estimator of Theorem 1, besides $\hat{x}_{0|0} = \bar{x}_0$ and $P_{0|0} = P_0$, we need $E[y_0 y_0']$, $E[x_0 y_0']$, $E[w_0 y_0']$ and y_0 itself, which is clearly easier to obtain than what is required by the two existing forms.

As we know, LMMSE achieves the smallest MSE within the linear class w.r.t. y^k . Given only the first two moments of other random quantities except γ_k , can we do better than the LMMSE estimator of Theorem 1 in terms of MSE? The answer is yes, as shown in the next section.

5 MMSE estimation

For an estimation problem involving intermediate decision, hard decision may be worse than soft decision. That is also one of the reasons why soft decision based algorithms, e.g., interacting multiple model (IMM) algorithm [23], are popular in maneuvering target tracking, as opposed to hard decision based algorithms, e.g., variable dimension filter [24] and input estimation (IE) algorithm [25]. But as will be shown in the following, the best state estimation performance with multiple packet dropouts is achieved by hard decision. Note that in this case hard decision is equivalent to soft decision and hard decision can be done without any decision error.

Consider system (1) – (3). If we further assume that x_0 , $\langle w_k \rangle$ and $\langle v_k \rangle$ are Gaussian distributed, then the MMSE estimation can be simply done as in Theorem 2.

Theorem 2 (MMSE estimation). If x_0 , $\langle w_k \rangle$ and $\langle v_k \rangle$ are Gaussian distributed and given $\hat{x}_{k-1|k-1} = E[x_{k-1}|y^{k-1}]$, $P_{k-1|k-1} = \text{MSE}(\hat{x}_{k-1|k-1})$, then the MMSE estimate of x_k for system (1) – (3) is:

Prediction:

$$\begin{aligned}\hat{x}_{k|k-1} &= E[x_k|y^{k-1}] = F_{k-1}\hat{x}_{k-1|k-1} \\ P_{k|k-1} &= \text{MSE}(\hat{x}_{k|k-1}) \\ &= F_{k-1}P_{k-1|k-1}F_{k-1}' + G_{k-1}Q_{k-1}G_{k-1}'\end{aligned}$$

Update:

If $y_k = y_{k-1}$, then

$$\begin{aligned}\hat{x}_{k|k} &= E[x_k|y^k] = \hat{x}_{k|k-1} \\ P_{k|k} &= \text{MSE}(\hat{x}_{k|k}) = P_{k|k-1}\end{aligned}$$

Otherwise (i.e., if $y_k \neq y_{k-1}$)

$$\begin{aligned}\hat{x}_{k|k} &= E[x_k|y^k] \\ &= \hat{x}_{k|k-1} \\ &\quad + P_{k|k-1}H_k'(H_kP_{k|k-1}H_k' + R_k)^+(y_k - H_k\hat{x}_{k|k-1}) \\ P_{k|k} &= \text{MSE}(\hat{x}_{k|k}) \\ &= P_{k|k-1} - P_{k|k-1}H_k'(H_kP_{k|k-1}H_k' + R_k)^+H_kP_{k|k-1}\end{aligned}$$

Proof:

The prediction follows easily from the property of MMSE estimation.

From the total expectation theorem, the MMSE estimate of x_k can be written as:

$$\begin{aligned}E[x_k|y^k] &= E[x_k|\gamma_k = 1, y^k]P\{\gamma_k = 1|y^k\} \\ &\quad + E[x_k|\gamma_k = 0, y^k]P\{\gamma_k = 0|y^k\}\end{aligned}$$

From the property of a continuous distribution, it follows that

$$P\{z_k = y_{k-1}\} = 0$$

With this, we can easily see that if $y_k = y_{k-1}$, then

$$P\{\gamma_k = 0|y^k\} = 1, P\{\gamma_k = 1|y^k\} = 0$$

and in this case the updated MMSE estimate of x_k simplifies to

$$E[x_k|y^k] = E[x_k|\gamma_k = 0, y^k] = E[x_k|y^{k-1}]$$

which is nothing but the one-step ahead prediction of x_k .

If $y_k \neq y_{k-1}$, then

$$y_k = z_k, P\{\gamma_k = 1|y^k\} = 1, P\{\gamma_k = 0|y^k\} = 0$$

and in this case the updated MMSE estimate of x_k simplifies to

$$E[x_k|y^k] = E[x_k|\gamma_k = 1, y^k] = E[x_k|y^{k-1}, z_k = y_k]$$

And since the LMMSE estimation turns out to be the MMSE estimation under the Gaussian assumption, it follows easily that

$$\begin{aligned}E[x_k|y^k] &= E[x_k|y^{k-1}, z_k = y_k] \\ &= \hat{x}_{k|k-1} + P_{k|k-1}H_k'(H_kP_{k|k-1}H_k' + R_k)^+ \\ &\quad \cdot (y_k - H_k\hat{x}_{k|k-1})\end{aligned}$$

if $y_k \neq y_{k-1}$. □

Remark: From Theorem 2, the test of the equivalence between y_k and y_{k-1} clearly has nothing to do with the distribution of x_0 , $\langle w_k \rangle$ and $\langle v_k \rangle$, so even if only the first two moments of x_0 , $\langle w_k \rangle$ and $\langle v_k \rangle$ are available, we can still use Theorem 2 to obtain an estimate of x_k . The simple hard

decision between $y_k = y_{k-1}$ and $y_k \neq y_{k-1}$ is a nonlinear operation and is not allowed in LMMSE estimation. It can help reduce the uncertainty associated with y_k , so the MMSE estimator should have better performance than the LMMSE estimator of Theorem 1. This is further verified by the illustrative examples in the next section. In this case, the estimator is still optimal in the LMMSE sense, but not w.r.t. y^k any more. Instead, it is optimal w.r.t. the raw measurement sequence $\langle z_k \rangle$ without the unrarrived packets.

Remark: It is clear from Theorem 2 that the test of the equivalence between y_k and y_{k-1} has nothing to do with the distribution of γ_k . That is, we do not need to know the exact value of p_k if $\langle \gamma_k \rangle$ is a Bernoulli distributed white sequence. Also $\langle \gamma_k \rangle$ can be any binary random sequence, e.g., Markov chain.

Remark: It can be easily seen that performance of the LMMSE estimation algorithms in [6, 7] and the LMMSE estimation algorithm of Theorem 1 all depend on the value of y_0 . But for the MMSE estimator of Theorem 2, its performance does not depend on the value of y_0 because the only use of y_0 is for comparison purpose when y_1 is received.

Remark: As is clear from Theorem 2, state estimation with multiple packet dropouts as formulated by system (1) – (3) is easy since the MMSE estimate can be obtained based on the simply hard decision by comparing y_k and y_{k-1} . For problems for which this comparison is legitimate, our simple solution by Theorem 2 largely nullifies the existing work on this problem.

Remark: Taken into account of the finite word length effect, performance loss will occur in the implementation of Theorem 2 when a digital quantity is utilized in place of an analog one. This is due to the many-to-one mapping from analog to digital. However, performance loss should be small for modern digital equipment based implementation.

6 Illustrative examples

In this section, we verify the effect of y_0 and data arrival probability p_k on the performance of our proposed LMMSE and MMSE estimators through numerical examples.

Consider the system (1) – (3), where

$$\begin{aligned} F_k &= 0.95, G_k = 1, w_k \sim \mathcal{N}(0, Q_k) \\ Q_k &= 1, k = 0, 1, 2, \dots, 50 \\ x_0 &\sim \mathcal{N}(\bar{x}_0, P_0), \bar{x}_0 = 0, P_0 = 20 \\ v_k &\sim \mathcal{N}(0, R_k), R_k = 9, k = 1, 2, \dots, 50 \end{aligned}$$

It is also known that

$$y_0 \sim \mathcal{N}(\bar{y}_0, \sigma_y^2), \bar{y}_0 = 0$$

and y_0 is uncorrelated with x_0 and w_0 .

All results in the following are averaged over 1,000 Monte Carlo runs.

Figs. 1, 2 and 3 show comparison results of estimators listed in Table 1 for $p_k = 0.8, 0.5$ and 0.2 , respectively.

Table 1: Estimators used in Figs. 1 to 3

name	explanation
KF	Kalman filter without packet dropouts
MMSE-a	MMSE estimator with $\sigma_y^2 = 100$
MMSE-b	MMSE estimator with $\sigma_y^2 = 900$
LMMSE-a	LMMSE estimator with $\sigma_y^2 = 100$
LMMSE-b	LMMSE estimator with $\sigma_y^2 = 900$

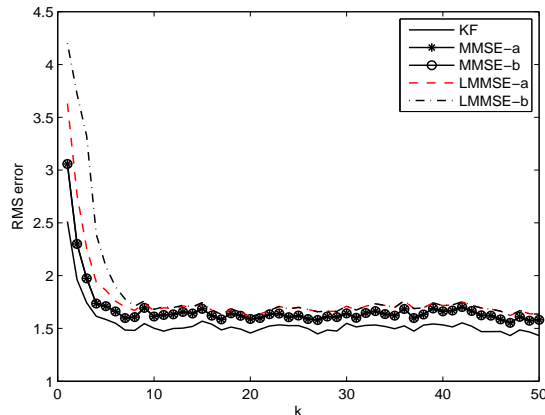


Figure 1: RMS error comparison with $p_k = 0.8$. Note that MMSE-a and MMSE-b overlap with each other.

As is clear from the simulation results, the Kalman filter achieves the best performance since there is no loss to the raw measurement information. The difference between the LMMSE-a filter and LMMSE-b filter verifies that the LMMSE estimator of Theorem 1 depends on the distribution and value of y_0 . The overlap of MMSE-a and MMSE-b verifies that the MMSE estimator does not depend on the value of y_0 and it is only used for comparison. Also, it can be seen that the MMSE estimator outperforms the LMMSE estimator, as expected. The gap between the MMSE estimator and the Kalman filter discloses the effect of packet dropouts over network transmission. As p_k decreases, the gap increases. Also, as p_k decreases, the performance of LMMSE and MMSE estimators becomes worse and their convergence rates become slower. All these are because as p_k decreases, raw measurement packets will arrive at the estimator less frequently. Correspondingly, the estimators will rely on the prediction more and more.

7 Conclusions

In network based applications, communication between local sensors and the processing center is usually not perfectly reliable, so packet dropouts may happen. Two existing forms of LMMSE estimation with multiple packet dropouts are based on a stochastic parameter system constructed by augmenting the original state and measurement. Concerning

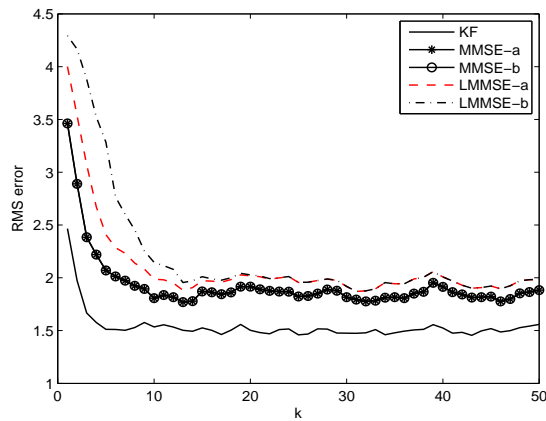


Figure 2: RMS error comparison with $p_k = 0.5$. Note that MMSE-a and MMSE-b overlap with each other.

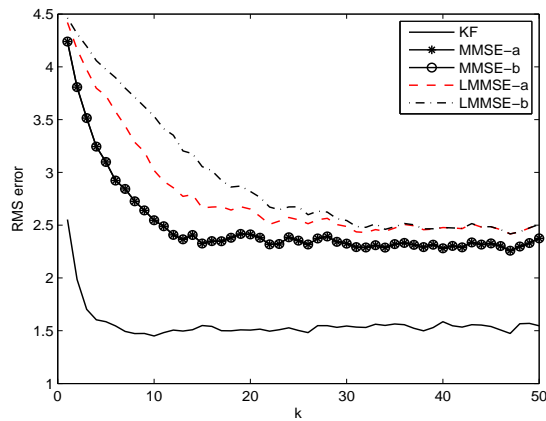


Figure 3: RMS error comparison with $p_k = 0.2$. Note that MMSE-a and MMSE-b overlap with each other.

the computational load, measurement residual characterization, and requirements on initialization, these two forms are not preferred. An alternative form of LMMSE estimator has been derived first. Then under the Gaussian assumption, the MMSE estimator has been also derived, which is, if without the Gaussian assumption, LMMSE-optimal w.r.t. the raw measurement sequence without the unrarried packets. The MMSE estimator has two nice properties. First, unlike the LMMSE estimator, its performance does not depend on the value of the initial data at the processing center. Second, it does not depend on the distribution of the data arrival events. These have been verified by numerical examples. The MMSE estimator is obtained based on a hard decision by comparing the measurement values at two consecutive time instants. If this comparison is legitimate, our simple optimal solution largely nullifies the existing work on this problem.

References

- [1] S. L. Sun, "Linear minimum variance estimators for systems with bounded random measurement delays and packet dropouts," *Signal Processing*, vol. 89, no. 7, pp. 1457–1466, July 2009.
- [2] S. L. Sun, L. H. Xie, W. D. Xiao, and N. Xiao, "Optimal filtering for systems with multiple packet dropouts," *IEEE Transactions on Circuits and Systems—II: Express Briefs*, vol. 55, no. 7, pp. 695–699, July 2008.
- [3] M. Moayed, Y. C. Soh, and Y. K. Foo, "Optimal Kalman filtering with random sensor delays, packet dropouts and missing measurements," in *Proceedings of 2009 American Control Conference*, St. Louis, MO, June 2009, pp. 3405–3410.
- [4] Y. Bar-Shalom, "Update with out-of-sequence measurements in tracking: exact solution," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 769–777, July 2002.
- [5] K. S. Zhang, X. R. Li, and Y. M. Zhu, "Optimal update with out-of-sequence measurements," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 1992–2004, June 2005.
- [6] S. L. Sun, L. H. Xie, and W. D. Xiao, "Optimal full-order and reduced-order estimators for discrete-time systems with multiple packet dropouts," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4031–4038, August 2008.
- [7] S. L. Sun, L. H. Xie, W. D. Xiao, and Y. C. Soh, "Optimal linear estimation for systems with multiple packet dropouts," *Automatica*, vol. 44, no. 5, pp. 1333–1342, May 2008.
- [8] M. Sahebsara, T. Chen, and S. L. Shah, "Optimal H2 filtering in networked control systems with multiple packet dropout," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1508–1513, August 2007.
- [9] —, "Optimal H2 filtering with random sensor delay, multiple packet dropout and uncertain observations," *International Journal of Control*, vol. 80, no. 2, pp. 292–301, February 2007.
- [10] M. Sahebsara, "Optimal filtering in multiple channel networked control systems with multiple packet dropout," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, December 2008, pp. 3367–3371.
- [11] J. L. Li and G. AlRegib, "Distributed estimation in energy-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3746–3758, October 2009.
- [12] J. Fang and H. B. Li, "Power constrained distributed estimation with correlated sensor data," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3292–3297, August 2009.
- [13] X. R. Li, Y. M. Zhu, J. Wang, and C. Z. Han, "Optimal linear estimation fusion - part I: Unified fusion rules," *IEEE Transactions on Information Theory*, vol. 49, no. 9, pp. 2192–2208, September 2003.
- [14] K. S. Zhang, X. R. Li, P. Zhang, and H. F. Li, "Optimal linear estimation fusion - part VI: Sensor data compression," in *Proceedings of the 6th International Conference on Information Fusion*, Cairns, Queensland, Australia, July 2003, pp. 221–228.
- [15] Z. S. Duan and X. R. Li, "Optimal distributed estimation fusion with compressed data," in *Proceedings of the 12th International Conference on Information Fusion*, Seattle, Washington, July 6-9 2009, pp. 563–570.

- [16] V. P. Jilkov, Z. S. Duan, and X. R. Li, "State estimation with quantized measurements: Multiple model approach," in *Proceedings of 2007 International Colloquium on Information Fusion*, Xi'an, China, August 2007, pp. 1–7.
- [17] Z. S. Duan, V. P. Jilkov, and X. R. Li, "State estimation with quantized measurements: Approximate MMSE approach," in *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, July 2008, pp. 1067–1072.
- [18] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poola, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, September 2004.
- [19] X. R. Li, *Applied Estimation and Filtering*. Course Notes, University of New Orleans, February 2006.
- [20] Z. L. Zhao, X. R. Li, and V. P. Jilkov, "Best linear unbiased filtering with nonlinear measurements for target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 4, pp. 1324–1336, October 2004.
- [21] X. R. Li, "Recursibility and optimal linear estimation and filtering," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, December 2004, pp. 1761–1766.
- [22] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley Interscience, 2001.
- [23] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, August 1988.
- [24] Y. Bar-Shalom and K. Birmiwal, "Variable dimension filter for maneuvering target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 18, no. 5, pp. 621–629, September 1982.
- [25] Y. T. Chan, A. G. C. Hu, and J. B. Plant, "A Kalman filter based tracking scheme with input estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 15, no. 2, pp. 237–244, March 1979.