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 $H_2$-norm of the system with stochastic parameter, the optimal $H_2$ 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 coefficient 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-
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:

\[ x_k = F_{k-1}x_{k-1} + G_{k-1}w_{k-1} \]  

where \( x_k \in \mathbb{R}^n \), \( \langle w_k \rangle \) 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

\[ z_k = H_k x_k + v_k, \quad k = 1, 2, \cdots \]  

where \( z_k \in \mathbb{R}^m \), \( \langle v_k \rangle \) 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 \( \langle v_k \rangle \) 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:

\[ y_k = \gamma_k z_k + (1 - \gamma_k)y_{k-1} \]  

where \( \gamma_k \) has Bernoulli distribution with \( P\{\gamma_k = 1\} = p_k \) and \( P\{\gamma_k = 0\} = 1 - p_k = q_k \). \( \gamma_k \) 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 \( \gamma_k \) 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,

\[ \hat{x}_{k|k}^{\text{LMMSE}} \Delta \arg \min \text{MSE}(\hat{x}_{k|k}) \]

where

\[ y^k = \{ y_1, \cdots, y_k \}, \quad Y_k = \{ y'_1, \cdots, y'_k \} \quad \text{MSE}(\hat{x}_{k|k}) = E[(x_k - \hat{x}_{k|k})(\hat{x}_k - \hat{x}_{k|k})'] \]

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:

\[ X_{k+1} = \tilde{F}_k X_k + \tilde{G}_k W_k \]  

\[ y_k = \bar{H}_k X_k + \gamma_k v_k \]  

where

\[ X_k = [ x'_k \ y'_k ] \in \mathbb{R}^{n+m}, \quad W_k = [ w'_k \ v'_k ]' \]

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

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

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

\[ c_{k+1} = \tilde{F}_k c_k \tilde{F}_k' \]

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

where

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

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

and it is assumed that

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

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:

\[ \tilde{X}_{k+1|k} = \tilde{F}_k \tilde{X}_{k|k} \]

\[ \text{MSE}(\tilde{X}_{k+1|k}) = p_k q_k \begin{bmatrix} 0 & 0 \\ H_k & -I_m \end{bmatrix} c_k \begin{bmatrix} 0 & 0 \\ H_k & -I_m \end{bmatrix}' + \tilde{F}_k \text{MSE}(\tilde{X}_{k|k}) \tilde{F}_k' + U_k \]
The one-step ahead prediction is
\[ \hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\varepsilon_{k+1} \]
\[ \varepsilon_{k+1} = y_{k+1} - \tilde{H}_{k+1}\hat{x}_{k+1|k} \]
\[ K_{k+1} = \text{MSE}(\hat{x}_{k+1|k}H'_{k+1}S_{k+1}^{-1}) \]
\[ \tilde{H}_{k+1} = E[\tilde{H}_{k+1}] = [p_{k+1}H_{k+1} q_{k+1}I_m] \]
\[ S_{k+1} = p_{k+1}q_{k+1} + [H_{k+1} - I_m ]c_{k+1} \]
\[ + \tilde{H}_{k+1}\text{MSE}(\hat{x}_{k+1|k})\tilde{H}_{k+1}' + p_{k+1}R_{k+1} \]
MSE(\hat{x}_{k+1|k+1}) = MSE(\hat{x}_{k+1|k}) - K_{k+1}S_{k+1}K_{k+1}'
with the initial conditions
\[ \hat{x}_{0|-1} = [\bar{x}_0 \ 0]' \text{, MSE}(\hat{x}_{0|-1}) = \begin{bmatrix} P_0 & 0 \\ 0 & 0 \end{bmatrix} \]

Finally,
\[ \hat{x}_{k+1|k+1} = [I_n \ 0]\hat{x}_{k+1|k} \]
\[ P_{k+1|k+1} = [I_n \ 0]\text{MSE}(\hat{x}_{k+1|k})[I_n \ 0]' \]

In [6], it was further assumed that
\[ F_k = F, \ G_k = G, \ H_k = H, \ Q_k = Q, \ R_k = R \]
\[ p_k = p, \ q_k = q = 1 - p \]

By designating the optimal of \( x_{k+1} \) to be of the form
\[ \hat{x}_{k+1|k+1} = \Phi_k\hat{x}_k + K_k^1y_{k+1} + K_k^2y_k \]
(7)
and from the unbiasedness and minimum error covariance properties of \( \hat{x}_{k+1|k+1} \), it was found that
\[ \Phi_k = (I_n - pK_k^1H)F \]
\[ K_k^1 = P_{k+1|k}H'\Lambda_k^{-1} \]
\[ K_k^2 = -qK_k^1 \]
\[ P_{k+1|k} = FP_{k+1|k}F' + GQQ' \]
\[ \Lambda_k = p(HP_{k+1|k}H' + R) + qHGQG'H' \]
\[ + q[I_pH - HF \ qI_m \ c_k[H - H - I_m]' \]
\[ + pq[2[H - I_m ]c_k[H - I_m]' + (1 - p^2)R \]
\[ P_{k+1|k+1} = P_{k+1|k} - pK_k^1\Lambda_k(K_k^1)' \]
The one-step ahead prediction is
\[ \hat{x}_{k+1|k} = F\hat{x}_k \]
with the initial conditions
\[ \hat{x}_{0|0} = \bar{x}_0, \ P_{0|0} = P_0 \]

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 \( \bar{x}_{0|0} = \bar{x}_0 \) and \( P_{0|0} = P_0 \) as well as \( E[x_0y'_{-1}] \) and \( E[y_{-1}y'_{-1}] \), both assumed zero therein, is needed. To initialize the LMMSE estimator in [7], besides \( \bar{x}_{0|1} = \bar{x}_0 \) and \( P_{0|1} = P_0 \) \( E[x_0y'_{-1}], E[y_{-1}y'_{-1}], \) \( E[(x_0 - \bar{x}_{0|1})(y_{-1} - y_{-1|1})]' \), 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 \( \gamma \), if \( \gamma \) 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
\[ E[\hat{x}x'] = E[xx'] - \text{MSE}(\hat{x}) \]

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

Prediction:
\[ \hat{x}_{k|k-1} = E^*[x_k|y_{k-1}'] = F_{k-1}\hat{x}_{k-1|k-1} \]
\[ P_{k|k-1} = \text{MSE}(\hat{x}_{k|k-1}) \]
\[ = F_{k-1}P_{k-1|k-1}F_{k-1}' + G_{k-1}Q_{k-1}G_{k-1}' \]

Update1:
\[ \hat{x}_{k|k} = E^*[x_k|y_{k}] \]
\[ = \hat{x}_{k|k-1} + C\varepsilon_{k|k-1, y_{k-1}}^+C_{y_{k-1}|y_{k-1}}\varepsilon_{k|k-1, y_{k-1}} \]
\[ P_{k|k} = \text{MSE}(\hat{x}_{k|k}) \]
\[ = P_{k|k-1} - C\varepsilon_{k|k-1, y_{k-1}}^+C_{y_{k-1}|y_{k-1}}C_{\varepsilon_{k|k-1, y_{k-1}}^+} \]
\[ \varepsilon_{k|k-1} = y_k - p_kH\hat{x}_{k|k-1} - q_ky_k \]
\[ C_{\varepsilon_{k|k-1, y_{k-1}}} = p_kP_{k|k-1}H'_{k} \]

1In the following \( A^+ \) stands for the unique MP inverse of matrix \( A \). It reduces to \( A^{-1} \) whenever \( A^{-1} \) exists.
Furthermore, with the initial conditions

\[ x_0|0 = \bar{x}_0, \quad P_0|0 = P_0, \quad C_0 = P_0 + \bar{x}_0\bar{x}_0^\top \]

\[ D_0 = E[\tilde{y}_0y_0], \quad E_1 = F_0E[x_0y_0] + G_0E[w_0y_0] \]

**Proof:**

Given \( \tilde{x}_k|k-1 \) and \( P_{k-1}|k-1 \), it follows easily from the property of LMMSE estimation that

\[ \hat{x}_{k|k} = E^*[x_k|y_k] \]

\[ = E^*[F_{k-1}x_{k-1} + G_{k-1}w_{k-1}|y_k] \]

\[ = F_{k-1}\hat{x}_{k-1|k-1} + G_{k-1}w_{k-1} \]

\( P_{k|k} = \text{MSE}(\hat{x}_{k|k}) = F_{k-1}P_{k-1|k-1}F_{k-1}^\top + G_{k-1}Q_{k-1}G_{k-1}^\top \)

The LMMSE estimator \( E^*[x_k|y_k] \) always has the following quasi-recurrent form [21]

\[ \hat{x}_{k|k} = E^*[x_k|y_k] = E^*[x_k|y_{k-1}, y_k] \]

\[ \hat{x}_{k|k} = C^{+}_{\hat{x}_{k|k-1}, \hat{x}_{k|k-1}} C^{+}_{\hat{x}_{k|k-1}, \hat{x}_{k|k-1}} C^{+}_{x_{k|k-1}, y_{k-1}} \]

From Lemma 1 above, we have

\[ \hat{y}_{k|k-1} = y_k - E^*[y_k|y_{k-1}] \]

\[ = y_k - E^*[\gamma_k z_k + (1 - \gamma_k)y_{k-1}|y_{k-1}] \]

\[ = y_k - E^*[\gamma_k z_k] E^*[H_k x_k + v_k|y_{k-1}] \]

\[ = E[1 - \gamma_k] E^*[y_{k-1}|y_{k-1}] \]

\[ = y_k - p_k H_k \tilde{x}_{k|k-1} - q_k y_{k-1} \]

Furthermore, \( \hat{y}_{k|k-1} \) can be rewritten as

\[ \hat{y}_{k|k-1} = y_k - p_k H_k \tilde{x}_{k|k-1} - q_k y_{k-1} \]

\[ = \gamma_k (H_k x_k + v_k) + (1 - \gamma_k) y_{k-1} \]

\[ = p_k H_k \tilde{x}_{k|k-1} - (1 - \gamma_k) y_{k-1} \]

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

\[ C_{\tilde{x}_{k|k-1}, \hat{x}_{k|k-1}} = \text{cov}(\tilde{x}_{k|k-1}, \hat{x}_{k|k-1}) = p_k P_{k-1|k-1} H_k^\top \]

\[ \hat{y}_{k|k-1} = \text{cov}(\tilde{y}_{k|k-1}, \hat{y}_{k|k-1}) = p_k P_{k-1|k-1} H_k^\top \]

\[ = E[y_k|k-1] \]

\[ = E[\gamma_k^2 |H_k P_{k-1|k-1} H_k^\top + R_k] \]

\[ + E[(\gamma_k - p_k)^2 |H_k E[\tilde{x}_{k|k-1} x_{k|k-1}] H_k^\top - H_k E[\tilde{x}_{k|k-1} y_{k-1}] + E[y_{k-1} y_{k-1}]] \]

\[ = p_k (H_k P_{k-1|k-1} H_k^\top + R_k) \]

\[ + (p_k - p_k^2) (H_k E[\tilde{x}_{k|k-1} x_{k|k-1}] H_k^\top - H_k E[\tilde{x}_{k|k-1} y_{k-1}] + E[y_{k-1} y_{k-1}]) \]

\[ = p_k (H_k P_{k-1|k-1} H_k^\top + R_k) + (p_k - p_k^2) \]

\[ (H_k C_k - H_k E[\tilde{x}_{k|k-1} x_{k|k-1}] + p_k R_k + p_k q_k) \]

\[ \cdot (H_k C_k H_k^\top - H_k E[\tilde{x}_{k|k-1} y_{k-1}] + E[y_{k-1} y_{k-1}]) \]

where we have used

\[ E[\tilde{x}_{k|k-1} x_{k|k-1}] = E[x_k x_k^\top] - P_{k-1} = C_k - P_{k-1} \]

\[ C_k = E[x_k x_k^\top] \]

\[ = E[(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})^\top] \]

\[ = F_{k-1} C_{k-1} F_{k-1}^\top + G_{k-1} Q_{k-1} G_{k-1}^\top \]

\[ D_k = E[y_k y_k^\top] \]

\[ = E[(\gamma_k z_k + (1 - \gamma_k) y_{k-1})(\gamma_k z_k + (1 - \gamma_k) y_{k-1})^\top] \]

\[ = E[\gamma_k^2 z_k z_k^\top] + E[\gamma_k (1 - \gamma_k) z_k y_{k-1}^\top] + E[(1 - \gamma_k) y_{k-1} y_{k-1}^\top] \]

\[ = p_k E[z_k z_k^\top] + q_k D_{k-1} \]

\[ = p_k E[(H_k x_k + v_k)(H_k x_k + v_k)^\top] + q_k D_{k-1} \]

\[ = p_k (H_k C_k H_k^\top + R_k) + q_k D_{k-1} \]

\[ E_k = E[x_k y_k^\top] \]

\[ = E[(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})(\gamma_k x_{k-1} + (1 - \gamma_k) y_{k-1})] \]

\[ = E[(F_{k-1} x_{k-1} + G_{k-1} w_{k-1})(\gamma_k x_{k-1} + (1 - \gamma_k) y_{k-1})^\top] \]

\[ = p_k E[H_k x_k + v_k] H_k x_k + v_k)^\top] + q_k D_{k-1} \]

\[ = p_k (H_k C_k H_k^\top + R_k) + q_k D_{k-1} \]

\[ \tilde{x}_{k|k-1} = \tilde{x}_{k|k-1} \]

**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 \( \tilde{y}_{k|k-1} \) and \( C_{\tilde{y}_{k|k-1}} \) and \( C_{\tilde{\theta}_{k|k-1}} \). To initialize the LMMSE estimator of Theorem 1, besides \( \hat{x}_{0|0} = \bar{x}_0 \) and \( P_{0|0} = P_0 \), we need \( E[y_0y_0'] \), \( E[x_0y_0'] \), \( E[w_0y_0'] \) 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 \( \gamma_k \), 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 \( \hat{x}_{k-1|k-1} = E[x_{k-1|k-1}] \), then the MMSE estimate of \( x_k \) for system (1) – (3) is:

Prediction:

\[
\hat{x}_{k|k} = E[x_k|y_k^k] = F_{k-1}\hat{x}_{k-1|k-1}
\]

\[
P_{k|k} = \text{MSE}(\hat{x}_{k|k}) = F_{k-1}P_{k-1|k-1}F_{k-1}' + G_{k-1}Q_{k-1}G_{k-1}'
\]

Update:

If \( y_k = y_{k-1} \), then

\[
\hat{x}_{k|k} = E[x_k|y_k^k] = \hat{x}_{k|k-1}
\]

\[
P_{k|k} = \text{MSE}(\hat{x}_{k|k}) = P_{k|k-1}
\]

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

\[
\hat{x}_{k|k} = E[x_k|y_k^k] = \hat{x}_{k|k-1} + P_{k|k-1}H_k^r(H_kP_{k|k-1}H_k^r + R_k)^+ (y_k - H_k\hat{x}_{k|k-1})
\]

\[
P_{k|k} = \text{MSE}(\hat{x}_{k|k}) = P_{k|k-1} - P_{k|k-1}H_k^r(H_kP_{k|k-1}H_k^r + R_k)^+ H_kP_{k|k-1}
\]

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:

\[
E[x_k|y_k^k] = E[x_k|\gamma_k = 1, y_k^k]P_{\{\gamma_k = 1\}|y_k^k} + E[x_k|\gamma_k = 0, y_k^k]P_{\{\gamma_k = 0\}|y_k^k}
\]

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

\[
P\{\gamma_k = 0|y_k^k\} = 1, \quad P\{\gamma_k = 1|y_k^k\} = 0
\]

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

\[
E[x_k|y_k^k] = E[x_k|\gamma_k = 0, y_k^k] = E[x_k|y_k^k-1]
\]

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

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

\[
y_k = z_k, \quad P\{\gamma_k = 1|y_k^k\} = 1, \quad P\{\gamma_k = 0|y_k^k\} = 0
\]

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

\[
E[x_k|y_k^k] = E[x_k|\gamma_k = 1, y_k^k] = E[x_k|y_k^k-1, z_k = y_k]
\]

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

\[
E[x_k|y_k^k] = E[x_k|y_k^k-1, z_k = y_k]
\]

\[
= \hat{x}_{k|k-1} + P_{k|k-1}H_k^r(H_kP_{k|k-1}H_k^r + R_k)^+ (y_k - H_k\hat{x}_{k|k-1})
\]

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
Kalman filter without packet dropouts

\[ p \] and \[ y \] listed in Table 1 for \[ \gamma_k \]. Instead, it is optimal w.r.t. the raw measurement sequence \( \langle z_k \rangle \) without the unarrived 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 \( \gamma_k \). 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

\[
F_k = 0.95, \quad G_k = 1, \quad w_k \sim \mathcal{N}(0, Q_k) \\
Q_k = 1, \quad k = 0, 1, 2, \ldots, 50 \\
x_0 \sim \mathcal{N}(\bar{x}_0, P_0), \quad \bar{x}_0 = 0, \quad P_0 = 20 \\
u_k \sim \mathcal{N}(0, R_k), \quad R_k = 9, \quad k = 1, 2, \ldots, 50
\]

It is also known that

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

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>
<thead>
<tr>
<th>name</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>KF</td>
<td>Kalman filter without packet dropouts</td>
</tr>
<tr>
<td>MMSE-a</td>
<td>MMSE estimator with ( \sigma_y^2 = 100 )</td>
</tr>
<tr>
<td>MMSE-b</td>
<td>MMSE estimator with ( \sigma_y^2 = 900 )</td>
</tr>
<tr>
<td>LMMSE-a</td>
<td>LMMSE estimator with ( \sigma_y^2 = 100 )</td>
</tr>
<tr>
<td>LMMSE-b</td>
<td>LMMSE estimator with ( \sigma_y^2 = 900 )</td>
</tr>
</tbody>
</table>

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
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 unarrived 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


