Abstract—In this paper we provide some analytical performance bounds for different distributed estimation schemes for stochastic discrete time linear systems where the communication between the sensors and the estimation center is subject to random packet loss. In particular, we analyze three different strategies. The first, named measurement fusion (MF) optimally fuses the raw measurements received so far from all sensors. The second strategy, named infinite bandwidth filter (IBF), computes the optimal mean square estimator assuming that each node can send all measurements observed up to its current time in a single packet. The last strategy, named open loop partial estimate fusion (OLPEF), simply sums local state estimates received from each sensor and replace the lost ones with their open loop counterpart. In particular, we propose novel mathematical tools to derive analytical upper and lower bounds for expected estimation error covariance the MF and the IBF strategies in the scenario of identical sensors and we compare their values with the empirical performance obtained via simulations.

I. INTRODUCTION

The rapid growth of large wireless sensor networks capable of sensing and computation promises the design of novel applications, but it is also posing several challenges due to the unavoidable lossy nature of the wireless channel. These challenges are particularly evident in control and estimation applications since packet loss and random delay degrade the overall system performance, thus motivating the development of novel tools and algorithms, as illustrated in the survey [7]. In this work we focus on the problem of estimating a stochastic discrete time linear system observed by a number of sensors which can preprocess sensor data and communicate this information to a central node via a wireless lossy channel.

There is a vast literature regarding distributed estimation and sensor fusion with perfect communication links. In particular, there are two classes of problems that are relevant to this work. The first class addresses the problem of distributing computational burden from the central node, while in the former approaches only the central node needs to know the whole system dynamics. There are also dedicated distributed estimation subject to packet loss and variable delay between the sensor and the estimation center. This problem is particularly relevant in moving target tracking applications based on radar and GPS measurements [3]. For example in [2] and [14] the authors showed how to perform optimal estimation with time-varying delay and out-of-order packets without requiring the storage of large memory buffers and the inversion of many matrices. More recently, in [11] the authors provided lower and upper bounds for the optimal mean square estimator subject to random measurement loss, and in [9] those results were extended to multiple distributed sensors subject to simultaneous packet loss and random delay. Finally, the recent papers [10][1] analyze some trade-offs between communication, computation and estimation performance in multi-hop tree networks.

However, there are only few scattered results concerned with distributed estimation subject to packet loss in which sensors are provided with computation capabilities and can preprocess data before transmitting it to the estimation center. A recent result in this direction is given by Gupta et al. [6] who showed that when there is only one sensor, the optimal strategy for the sensor in the presence of packet loss is to send the local Kalman estimate rather than the raw measurement. This is because the local estimate includes the information about all previous measurements, therefore as soon as the central node receives the local estimate it can reconstruct the optimal estimate even if some previous packets were lost. Unfortunately, this result does not generalize to multiple sensors each provided with its own lossy communication channel. We recently explored this problem [5] and we numerically compared different preprocessing strategies at sensor nodes and fusion strategies at the central estimation node. In particular, we showed that the optimal mean square estimation error that can be achieved under packet loss, referred as infinite bandwidth filter (IBF), cannot
be achieved using a limited bandwidth channel, unless in very special scenarios where there is no process noise, such as in estimation of constant parameters. Therefore, in [5] we proposed different suboptimal strategies with different computational and communication requirements and we studied their performance via simulations observing that no strategy was superior to any other, since the performance depended on the packet loss probability and noise scenarios. For convenience, these results are briefly recalled in the following of this paper.

The contribution of this work resides on the derivation of analytical expressions to compute upper and lower bounds of performance of these estimators assuming i.i.d. Bernoulli packet loss probabilities. Finding bounds on performance turns out to be particularly challenging due to the fact that the estimation error covariance of the different estimator at the central node depends non-linearly on the specific packet loss sequence of all sensors, therefore computing expected error covariance a-priori given the packet loss statistics becomes a combinatorial problem that explodes with time. In particular, we derive upper and lower bounds in the scenario where all sensors are identical for two specific strategies: the measurement fusion (MF) strategy and the infinite bandwidth filter (IBF) strategy. The MF strategy is based on optimally fusing the raw measurements received by the central station from the sensors, while the IBF strategy is based on the assumption that each nodes sends to the base station not only the current measurement but also all previous measurements in a single packet. We also show through some simulations that some of these bounds are rather tight and can be used to estimate in advance the expected error of the different strategies. We finally end this paper by discussing possible future extensions.

II. PROBLEM FORMULATION

A. Modeling

We consider a discrete time linear stochastic systems observed by N sensors:

\[ x_{t+1} = Ax_t + w_t \]
\[ y_{i}^t = C_{i}x_t + v_{i}^t, \quad i = 1, \ldots, N \]  

where \( x \in \mathbb{R}^n, y_i \in \mathbb{R}^{m_i}, w_t \) and \( v_i^t \) are uncorrelated, zero-mean, white Gaussian noises with covariances \( \mathbb{E}[w_tw_t^T] = Q \) and \( \mathbb{E}[v_i^tv_i^T] = R_{i} \), i.e. we also allow for correlated measurement noise. More compactly, if we define the compound measurement column noise vector \( v_t = (v_1^t, \ldots, v_N^t) \in \mathbb{R}^m, m = \sum_i m_i \), we have \( \mathbb{E}[v_t^Tv_t^T] = R_{\delta}(t-s) \), where the \((i,j)\)-th block of the matrix \( R \in \mathbb{R}^{m \times m} \) is \( R_{ij} = R_{ij} \in \mathbb{R}^{m_i \times m_j} \). The initial condition \( x_0 \) is again a zero-mean Gaussian random variable uncorrelated with the noises and covariance \( \mathbb{E}[x_0x_0^T] = P_0 \). We also assume that \( R \succ 0 \), the pair \((A, Q^{1/2})\) is reachable and \((A, C)\) is observable, where \( C^T = [C_1^T \ldots C_N^T] \), which are sufficient conditions for the existence of a stable estimator under ideal conditions with no packet loss.

The sensors are not directly connected with each other and can send messages to a common central node through a lossy communication channel, i.e. there is a non zero probability that the message is not delivered correctly. We model the packet dropping events through a binary random variable \( \gamma_i^t \in \{0, 1\} \) such that:

\[ \gamma_i^t = \begin{cases} 0 \quad \text{if packet sent at time } t \text{ by node } i \text{ is lost} \\ 1 \quad \text{otherwise} \end{cases} \]  

Each sensor is provided with computational and memory resources to (possibly) preprocess information before sending it to the central node. More formally, at each time instant \( t \) each sensor \( i \) sends the preprocessed information \( z_i^t \in \mathbb{R}^d \):

\[ z_i^t = f_i^t(y_1^t, y_2^t, \ldots, y_N^t) = f_i^t(y_i^t) \]  

where \( t \) is bounded and \( f_i^t() \) are causal functions of the local measurements. Natural choices are \( z_i^t = y_i^t \), i.e. the latest measurement, or the output of a (time varying) linear filter:

\[ \xi_i^t = F_i^t\xi_{i,t-1} + G_i^ty_i^t \]
\[ z_i^t = H_i^t\xi_i^t + D_i^ty_i^t \]

as for example a local Kalman filter.

The objective is to design a state estimator at the central node given the information arrived up to time \( t \). More formally, let us define the information set available at the central node as:

\[ \mathcal{I}_t = \bigcup_{i=1}^{N} \mathcal{I}_i \]
\[ \mathcal{I}_i = \{ z_k^i | \gamma_k^i = 1, k = 1, \ldots, t \} \]

Based on this set, we want to design an estimator as follows:

\[ \hat{x}_{i,t}^g = g_t(\mathcal{I}_t) \]

such that its error \( P_{i,t}^g = \text{var}(x_t - \hat{x}_{i,t}^g | \mathcal{I}_t) = \mathbb{E}[(x_t - \hat{x}_{i,t}^g)^T(x_t - \hat{x}_{i,t}^g) | \mathcal{I}_t] \) is small. Depending on the choice of the sensor preprocessing functions \( f_i^t \) and the estimator functions \( g_t \), we get different strategies. Note that the estimator error covariance \( P_{i,t}^g \) is a function of \( \mathcal{I}_t \), and therefore also on the specific packet loss sequence, i.e. it is a random variable. In the following of this section, we propose three different strategies, based on natural choices for the functions \( f_i^t \) and \( g_t \). Other choices are obviously possible, as discussed in [5].

B. Infinite Bandwidth Filter

Here we propose the optimal filter in mean square sense that we can obtain if we assume infinite bandwidth in the communication channel when a packet is sent successfully, i.e. each sensor sends to the central node all measurements up to current time:

\[ z_t^i = y_t^i \]

where \( y_t^i = (y_1^t, y_{i-1}^t, \ldots, y_1^t) \). The estimator at the central node is given by

\[ \hat{x}_{i,t}^{IBF} = \mathbb{E}[x_t | \mathcal{I}_t] = \mathbb{E}[x_t | y_{1,t}, y_{1,t-\tau}, \ldots, y_{1,t-\tau}^N] \]

where \( \tau_i^t \) is the delay of the most recent received packet. This filter is optimal among all possible strategies since it is easy to see that

\[ P_{i,t}^{IBF} \leq P_{i,t}^g, \quad \forall j^t, \forall \gamma_i^t, \forall g_t() \]

where \( P_{i,t}^{IBF} = \text{var}(x_t - \hat{x}_{i,t}^{IBF} | \mathcal{I}_t) \) is the error covariance of the infinite bandwidth filter. In other words, this filter sets
a bound on the achievable performance of any other filter. Unfortunately, as shown in [5], there is no hope to find a strategy which achieves the performance with a more parsimonious use of the channel. This finding is formally stated in the next theorem:

**Theorem 1 ([5]):** Let us consider the state estimate $\hat{x}^9_{t|t}$ and $\hat{x}^{IBF}_{t|t}$ defined as above. Then there do not exist (possibly nonlinear) functions $f^9_i(y_{1:t}) \in \mathbb{R}^d$ with bounded size $\ell < \infty$ such that $P^9_{t|t} = P^{IBF}_{t|t}$ for any possible packet loss sequence, i.e.

$$\hat{f}_i^9(y_{1:t}) = P^{IBF}_{t|t}, \quad \forall \gamma^9_i$$

The previous theorem states that there is no hope to find a preprocessing with bounded message size which can achieve the error covariance $P^{IBF}_{t|t}$ of the infinite bandwidth filter (IBF) since it is not possible to know in advance what the packet loss event will be. We will therefore propose two suboptimal estimation strategies which provide the optimal solution in the special case of perfect communication link, i.e. when there is no packet loss.

**C. Measurement Fusion**

The first, referred as measurement fusion (MF), consists in sending the raw measurements

$$z^9_i = y^9_i$$

from each sensor node, and to find the best mean square state estimator with the arrived measurements at the central node:

$$\hat{x}^{MF}_{t|t} = \mathbb{E}[x_t | T^i, i = 1, \ldots, N]$$

where the information set in this case corresponds to $T^i = \{ y^9_k | \gamma^9_k = 1, k = 1, \ldots, t \}$. This strategy has been shown to provide good performance in simulations under different noise regimes [5], however, intuitively, it should provide almost optimal performance in a scenario with high ratio between process noise and measurement noise. In fact, if the process noise is large as compared to the measurement noise only most recent measurements convey relevant information, therefore there is no much gain in filtering the past measurements at the sensors. Although this seems to be the case in many simulations, there are choices for the system dynamics for which the MF strategy is not optimal even under perfect measurements, as stated in the next theorem:

**Theorem 2 ([5]):** Let us consider $R = 0$ and $Q > 0$. Then there exist scenarios in terms of packet loss sequences and system dynamics parameters $A, C$ for which $P^{MF}_{t|t} > P^{IBF}_{t|t}$.

It is possible to explicitly compute the MF filter as follows. Let us first define the following variables:

$$\bar{C}_t = \begin{bmatrix} \gamma^0_1 C_1 \\ \vdots \\ \gamma^N C_N \end{bmatrix}, \quad y_t = \begin{bmatrix} \gamma^0_t y^9_t \\ \vdots \\ \gamma^N_t y^N_t \end{bmatrix},$$

$$\bar{R}_t = \text{diag}\{ \gamma^0_t R_{11}, \gamma^2_t R_{22}, \ldots, \gamma^N_t R_{NN} \}$$

which can be obtained from the centralized matrices $C$ and $R$ under the assumption of uncorrelated measurement noise, i.e. $R_{ij} = 0, i \neq j$, and from the lumped column measurement vector $y_t = (y^9_1 y^9_t \ldots y^N_t)^T$ by replacing the rows and columns corresponding to the lost packet with zeros. It was shown in [9] that the state estimate for the measurement fusion strategy is given by:

$$\hat{x}^{MF}_{t|t} = (I - \bar{L}_t \bar{C}_t) A \hat{x}^{MF}_{t-1|t-1} + \bar{L}_t y_t$$

$$P^{MF}_{t|t} = P_{t|t-1} - P_{t|t-1} \bar{C}_t^T (\bar{C}_t P_{t|t-1} \bar{C}_t^T + \bar{R}_t) \bar{C}_t P_{t|t-1} + R_t$$

$$L_t = P_{t|t-1} \bar{C}_t^T (\bar{C}_t P_{t|t-1} \bar{C}_t^T + R_t)$$

$$P_{t+1|t} = A P^{MF}_{t|t} A^T + Q$$

where the symbol $\dagger$ indicates the Moore-Penrose pseudoinverse. The previous equations correspond to a time-varying Kalman filter which depends on the packet loss sequence. Note that only measurements that have arrived are used to the computation of the estimate $\hat{x}^{MF}_{t|t}$, i.e. the dummy zero measurement in $\bar{y}_t$ are not used as if they were real measurements, but are discarded.

The measurement fusion strategy has the advantage to be computed recursively and exactly with the inversion of one matrix of (at most) the size of the lumped measurement vector $y_t$. On the other hand, if a packet is lost, then the information conveyed by the measurement in that packet is lost forever, while sending filtered version of the output as in the open loop partial estimate fusion (OLP EF) this information might be partially recovered.

**D. Open Loop Partial State Estimate Fusion**

This strategy is suggested by the observation that, in the absence of packet losses, one could compute the gains in a centralized manner and distribute the computations to each sensor. To be more precise, assume all measurements were available to a common location, i.e. that there were no packet losses. We shall denote with $x^{CKF}_{t|t} := \mathbb{E}[x_t | y^9_1, \ldots, y^N_N]$ the centralized Kalman filter (CKF). Its evolution is governed by the equations:

$$\hat{x}^{CKF}_{t|t} = F_t \hat{x}^{CKF}_{t-1|t-1} + L_t y_t$$

$$F_t = (I - L_t C) A$$

where the gain $L_t = [L^1_t L^2_t \cdots L^N_t]$ is the centralized Kalman filter gain computed as

$$P_{t+1} = (A - K_t C) P_t (A - K_t C)^T + K_t R K_t^T + Q$$

$$K_t = P_t C^T (P_t C^T + R)^{-1}$$

Note now that, defining $z^i_t$ to be the solution of

$$z^i_t = F_t z^i_{t-1} + L^i_t y^i_t,$$

the CKF estimate $\hat{x}^{CKF}_{t|t}$ is given by $\hat{x}^{CKF}_{t|t} = \sum_{i=1}^N z^i_t$. For this reason we shall call the $z^i_t$’s “partial estimates”. This strategy has been suggested in [13] for distributed estimation with the purpose of reducing the power consumption. Since $z^i_t$ correspond to a partial estimate, a naive strategy at the central node for compensating the packet loss is to use the open loop partial state estimate based on the latest received packet from each node, i.e.

$$x^{OLP EF}_{t|t} = \sum_{i=1}^N A^T z^i_{t-\tau^i}$$
where \( \tau_i \) is the delay of the most recent packet received from node \( i \) at time instant \( t \). Although this looks like a naive solution, it was shown in [5] to provide optimal performance in the small process noise to measurement noise regime as formally stated in the following theorem:

**Theorem 3 ([5]):** Let \( Q = 0 \) and assume uncorrelated measurement noise among sensors, i.e. \( R = \text{block diag}\{R_{1}, ..., R_{N}\} \). Then

\[
P_{i|i}^{IBF} = P_{i|i}^{OLPEF} < P_{i|i}^{MF}
\]

**E. Simulations**

In Figure 1 we report the empirical performance of MF, IBF, and OLPEF for the system in (20), with \( \mathbb{E}w_i^2 = \mu_Q 10^{-3} \), \( \mathbb{E}(v_i^2) = 10 \) and the packet loss probability \( \lambda = 0.5 \).

As expected OLPEF and IBF are indistinguishable\(^1\) for small \( \mu_Q \), i.e. for small process noise, while OLPEF becomes significantly worse ad \( \mu_Q \) becomes large. Note also that for small \( \mu_Q \), which gives a filter with “long” memory, IBF is better than MF while for large \( \mu_Q \), and consequently a short memory filter, MF and IBF become indistinguishable.

\[
\begin{align*}
K &= \arg \min_{\ell} \mathcal{L}(P, L, \ell) = PC^T \left( CPC^T + \frac{1}{\ell} R \right)^{-1} \\
&= (I - KC)P(I - KC)^\top + \frac{1}{\ell} K R K^\top = P - PC^T \left( CPC^T + \frac{1}{\ell} R \right)^{-1} CP
\end{align*}
\]

For future use we also define the prediction error variance update

\[
\Phi(P, \ell) := A \Phi(P, \ell) A^\top + Q
\]

**Lemma 1:** The functions \( \Phi_f(P, \ell) \) and \( \Phi(P, \ell) \) are concave as a function of \( P \) and convex as a function of \( \ell \).

**Proof:** Let us first consider \( \Phi_f(P, \ell) \). Concavity in \( P \) follows rather easily from the fact that \( \Phi_f(P, \ell) = \min_{P,L,\ell} \mathcal{L}(P, L, \ell) \).

As far as convexity in \( \ell \), the following argument can be used: assume \( h \) is (positive) real variable and consider the derivatives \( \frac{\partial \Phi_f(P,\ell)}{\partial \ell} \) and \( \frac{\partial^2 \Phi_f(P,\ell)}{\partial \ell^2} \). It is easy to verify that \( \frac{\partial \Phi_f(P,\ell)}{\partial \ell} < 0 \) and \( \frac{\partial^2 \Phi_f(P,\ell)}{\partial \ell^2} > 0 \). The conclusion for \( \Phi_f(P, \ell) \) follows from the fact that \( \Phi_f(P, \ell) \) is an affine function of \( \Phi_f(P, \ell) \). This completes the proof.

In the following we shall also make extensive use of a lower bound of the Riccati operator \( \Phi_f(P, \ell) \) as follows. Consider the convex set \( \mathcal{P} := \{P = P^\top : \mathbb{P} \geq P_m, P \leq P_M \} \). We would like to find a linear function of \( P \), say \( G_f(P, \ell) \), such that \( G_f(P, \ell) \leq \Phi_f(P, \ell) \) \( \forall P \in \mathcal{P} \). This linear lower bound can be constructed in this way. For any \( P \in \mathcal{P} \) one can take \( K_P := \arg \min_{P \in \mathcal{P}} \mathcal{L}(P, L, \ell) \). Then

\[
\mathcal{L}(P, K_P, \ell) \geq \Phi(P, \ell) \quad \forall P \in \mathcal{P}, P^\top \geq 0
\]

holds. Among all such \( K_P \) we would like to find that which minimizes the maximal difference

\[
S(K_P) := \max_{P \in \mathcal{P}} \mathcal{L}(P, K_P, \ell) - \Phi(P, \ell)
\]

choosing

\[
K_{opt} := \arg \min_{K_P} S(K_P)
\]

\[
= \arg \min_{K_P} \left[ \max_{P \in \mathcal{P}} \mathcal{L}(P, K_P, \ell) - \Phi(P, \ell) \right].
\]

It is obvious that \( K_{opt} \) is also the optimal (Kalman) gain for a specific value of \( P \in \mathcal{P} \). At this point, since \( \mathcal{L}(P, K_{opt}, \ell) - \Phi(P, \ell) \leq S(K_P), \forall P \in \mathcal{P} \),

\[
\mathcal{L}^B_L(P, \ell) := \mathcal{L}(P, K_{opt}, \ell) - S(K_P) \leq \Phi_f(P, \ell) \quad \forall P \in \mathcal{P}.
\]

The matrices \( P_m \) and \( P_M \) define the set \( \mathcal{P} \) over which the linear lower bounds holds. In the rest of the paper we shall always use \( P_m = P(P_m, N) \), i.e. the lowest achievable (steady state) prediction error variance when all \( N \) sensors are utilized, and \( P_M = AP_M A^\top + Q \), i.e. the steady state variance, which is the upper bound of the state prediction error when no information is available. It is a standard fact to show that, provided \( P \in \mathcal{P} \), also \( \Phi(P, \ell) \in \mathcal{P} \).

\(^1\)Small differences are due to the sample variability of Montecarlo estimates.
Remark 1: The problem of finding \( S(K_F) \) boils down to a convex optimization problem. In general one would like to minimize \( S(K_F) \) over \( K_F \). This however might happen to be difficult. However, finding \( S(K_F) \) for some \( K_F \) is sufficient to the purpose of finding a linear approximation which bounds from below the Riccati update. When \( S(K_F) \) has not been minimized over \( K_F \) a less accurate approximation is found, since \( S(K_F) \) is the maximal difference between \( L(P, K_F, \ell) \) and \( \Phi_f(P, \ell) \).

We shall also need to consider a linear lower bound \( L^{LB}(P, \ell) \) for \( \Phi(P, \ell) \). It is immediate to see that
\[
L^{LB}(P, \ell) = \sum_{k}^{\infty} \left( \Phi_f(P, \ell, k) - \Phi_f(P, \ell, k-1) \right) + \Phi_f(P, \ell, 1)\]

It has been shown that the steady state minimum expected error covariance for this filter, i.e. \( S = \min_{\ell} \lim_{t \to \infty} \mathbb{E}[|x_t - \hat{x}_t|] \) is given by the fixed point of the following operator:
\[
\Psi_{\lambda}(S) = \frac{ASAT + Q}{\lambda} - \lambda ASC(TACST + (1 - \lambda)SC + R)^{-1}CSAT
\]

\[
S_C = \text{diag}(C_1SC_1^T, \ldots, C_NSC_N^T)
\]

B. Lower Bound for Measurement Fusion

The following proposition gives a lower bound on the expected (or average) state estimation error for the measurement fusion approach.

**Proposition 1:** Let \( P_{MF} \) and \( P_{MF}^I \) respectively the steady state (w.r.t. the measure induced by the loss mechanism) prediction and filtering state estimation errors. Then
\[
\mathbb{E}[P_{MF}] \geq P_{MF}^{LB}, \quad \mathbb{E}[P_{MF}^I] \geq P_{MF}^{LB}
\]

where \( P_{MF}^{LB} \) is the unique stationary solution of \( P_{MF}^{LB} = L^{LB}(P_{MF}, \mathbb{E}) \) and \( P_{MF}^{LB} = L^{LB}(P_{MF}, \mathbb{E}) \).

**Proof:** The state estimation error (prediction) using the measurement fusion approach satisfies the recursive equation \( P_{t+1} = \Phi(P_t, \ell_t) \). From convexity of \( \Phi(P, \ell) \) in \( \ell \), it follows that
\[
\mathbb{E}[P_{t+1}|P_t] \geq \Phi(P_t, \mathbb{E}_{\ell_t})
\]

where independence of \( \ell_t \) and \( P_t \) has been used. Using the lower bound \( \Phi(P, \ell) \geq L^{LB}(P, \ell) \) it follows that
\[
\mathbb{E}[P_{t+1}|P_t] \geq L^{LB}(P_t, \mathbb{E}_{\ell_t})
\]

Since \( L^{LB} \) is linear in \( P_t \) also
\[
\mathbb{E}[P_{t+1}] \geq L^{LB}(\mathbb{E}P_t, \mathbb{E}_{\ell_t}) \quad (18)
\]

follows. Using the fact \( L^{LB}(P, \ell) \) is non-decreasing as a function of \( P \), i.e. \( L^{LB}(P_2, \ell) \geq L^{LB}(P_1, \ell) \) whenever \( P_2 \geq P_1 \) and using stationarity of \( \ell_t \) (implying \( \mathbb{E}_{\ell_t} = \mathbb{E} \)), equation (18) can be iterated yielding
\[
\mathbb{E}[P_{MF}] \geq P_{MF}^{LB}, \quad P_{MF}^{LB} = L^{LB}(P_{MF}, \mathbb{E})
\]

The bound for the filtering solution is easily obtained observing that \( P_{MF}^{I} = \Phi(P_t, \ell_t) \geq L^{LB}(P_t, \ell) \) so that
\[
\mathbb{E}[P_{MF}] \geq L^{LB}(\mathbb{E}P_t, \mathbb{E}_{\ell_t})
\]

and therefore
\[
\mathbb{E}[P_{MF}^I] \geq P_{MF}^{LB} := L^{LB}(P_{MF}, \mathbb{E}).
\]

C. Lower Bound for Infinite Bandwidth Filter

The estimator \( x_{t+1}^{IBF} \) is characterized by the numbers \( \ell_1, \ldots, \ell_N \), where \( \ell_i \) is the numbers of steps elapsed since the last packet from node \( i \) has been received. Under the assumption of identical sensors, these are in one to one correspondence with the numbers \( h_t, h_{t-1}, h_{t-2}, \ldots \) defined as:
\[
h_t = \sum_{j=1}^{N} \delta(t_{i,j}^t - i)
\]
where $\delta(\cdot)$ is the Kronecker delta. As mentioned above the IBF can be thought of as a measurement fusion filter where the equivalent numbers of packets arrived at time $t$ are defined, recursively by the relation

$$
\begin{align*}
\ell_t &= h_t \\
\ell_{t-k} &= \ell_{t-k+1} + h_{t-k} & k = 1, 2, \ldots
\end{align*}
$$

It is fairly easy to see that the joint probability density function of the variables $\ell_{t-k}$ can be written in terms of the conditional densities

$$p(\ell_{t-k-1}|\ell_{t-k}, \ell_{t-k+1}, \ldots, \ell_t) = p(\ell_{t-k-1}|\ell_{t-k}),$$

which have the expression

$$p(\ell_{t-k-1}|\ell_{t-k}) = \left( N - \ell_{t-k} \right) \lambda^{N-\ell}(1-\lambda)^{\ell_{t-k}-1}$$

where $\lambda$ is the packet loss probability. Based on this we shall now construct a sequence of lower bounds as follows. Let us assume at time $t-k$ all packets arrived. This means that $\ell_{t-k} = N$. Let $P_{IBF}^f$ and $P_{IBF}^t$ be the (steady state) state prediction error variance at time $t$ and with $P_{IBF}(t+1, k, \ell_{t-k+1}, \ldots, \ell_t)$ the state prediction error variance at time $t+1$ conditioned on $\ell_{t-k} = N$ and with subsequent number of arrivals $\ell_{t-k+1}, \ldots, \ell_t$.

It is clear that

$$P_{IBF}^t(k) := \mathbb{E}[P_{IBF}(t, k, \ell_{t-k+1}, \ldots, \ell_t)|\ell_{t-k} = N]$$

$$P_{IBF}^t(k) := \mathbb{E}[P_{IBF}(t+1, k, \ell_{t-k+1}, \ldots, \ell_t)|\ell_{t-k} = N]$$

are increasing functions of $k$ and provide a sequence of lower bounds for $\mathbb{E}[P_{IBF}^f]$ and $\mathbb{E}[P_{IBF}^t]$, i.e.

$$\mathbb{E}[P_{IBF}^f] = \tilde{P}_{IBF}^f(\infty) \geq \tilde{P}_{IBF}^f(k) \geq P_{IBF}^f \quad \forall k$$

$$\mathbb{E}[P_{IBF}^t] = \tilde{P}_{IBF}^t(\infty) \geq \tilde{P}_{IBF}^t(k) \geq P_{IBF}^t$$

The following proposition holds.

**Proposition 2:** The matrices $P_{IBF}^f(k)$ and $P_{IBF}^t(k)$ in (19) admits the lower bounds as $P_{IBF}^f(k) \geq P_{IBF}^{f, LB}(k)$ and $P_{IBF}^t(k) \geq P_{IBF}^{t, LB}(k)$ the inequalities:

$$P_{IBF}^{f, LB}(1) = \Phi(P_m, \mathbb{E}[\ell_i]) = \Phi(P_m, N(1-\lambda))$$

$$P_{IBF}^{f, LB}(2) = \mathbb{E}[\mathcal{L}^{LB}\Phi(P_m, \mathbb{E}[\ell_{t-1}|\ell_t], \ell_{t-1})]$$

$$P_{IBF}^{f, LB}(k) = \mathbb{E}[\mathcal{L}^{LB}_{1} \cdots \mathcal{L}^{LB}_{k} \Phi(P_m, \mathbb{E}[\ell_{t-k+1}|\ell_{t-k+2}])].$$

and

$$P_{IBF}^{t, LB}(1) = \Phi(P_m, \mathbb{E}[\ell_i]) = \Phi(P_m, N(1-\lambda))$$

$$P_{IBF}^{t, LB}(2) = \mathbb{E}[\mathcal{L}^{LB}\Phi(P_m, \mathbb{E}[\ell_{t-1}|\ell_t], \ell_{t-1})]$$

$$P_{IBF}^{t, LB}(k) = \mathbb{E}[\mathcal{L}^{LB}_{1} \cdots \mathcal{L}^{LB}_{k} \Phi(P_m, \mathbb{E}[\ell_{t-k+1}|\ell_{t-k+2}])].$$

where $\mathbb{E}[\ell_{t-k+1}|\ell_{t-k+1}] = \ell_{t-k+1} + (1-\lambda)(N - \ell_{t-k+1})$ and $P_m$ is the solution of $P_m = \Phi(P_m, N)$, i.e. the optimal (steady state) error variance when measurements from $N$ sensors are received at all times. Similar equations hold for $k > 3$.

**Proof:** For $k = 1$, $P_{IBF}(t+1, 1, \ell_t) = \Phi(P_m, \mathbb{E}[\ell_t])$. Then, using convexity of $\Phi$ in $\ell_t$ it follows that

$$P_{IBF}(1) = \mathbb{E}P_{IBF}(t+1, 1, \ell_t) \geq \Phi(P_m, \mathbb{E}[\ell_t]).$$

When $k = 2$,

$$P_{IBF}(t+1, 2, \ell_{t-1}, \ell_t) = \Phi(P_m, \mathbb{E}[\ell_{t-1}], \ell_t) \geq \mathcal{L}^{MF}(P_m, \mathbb{E}[\ell_{t-1}], \ell_t)$$

holds. Using linearity of $\mathcal{L}^{LB}$ and convexity of $\Phi$ in $\ell_{t-1}$ one obtains that (a.s.)

$$\mathbb{E}[P_{IBF}(t+1, 2, \ell_{t-1}, \ell_t)|\ell_t] \geq \mathcal{L}^{MF}(P_m, \mathbb{E}[\ell_{t-1}], \ell_t)$$

from which,

$$P_{IBF}(2) := \mathbb{E}[P_{IBF}(t, 2, \ell_{t-1}, \ell_t)|\ell_t] \geq \mathcal{L}^{MF}(P_m, \mathbb{E}[\ell_{t-1}], \ell_t)$$

The proofs for $k > 2$ and for $P_{IBF}^f(k)$ follow the same lines and are therefore omitted.

**Remark 2:** In practice one can compute $P_{IBF}^f(k)$ for increasing values of $k$ until convergence. For the example considered (see figure 3) we stopped at $k = 3$. Only marginal improvements could be noticed increasing $k$ further.
V. Conclusions

In this paper we derived some analytical upper and lower bounds for different strategies addressing the problem of distributed estimation subject to random packet loss between the sensors and the central location where the best state estimate is required. We proposed novel mathematical tools to derive these bounds and we showed through numerical simulations the comparison with empirical performance based on simulations. This is just a preliminary work and many problems are still open such as the extension of our results to multidimensional systems with different sensors and the derivation of analytical bounds also for other strategies such as the open loop partial estimate fusion (OLPEF).

References


