Performance bounds for information fusion strategies in packet-drop networks

A. Chiuso, L. Schenato

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, where the decision process takes place, to the distributed sensor, under the assumption of perfect communication, i.e. packets arrive with no delay or at worst with a constant delay. In this context, Willsky, Levy et al. [12] [8] showed that it is possible to reconstruct the centralized Kalman filter (CKF) estimate from local Kalman filter (LKF) estimates generated by each sensor. In particular, the CKF can be obtained as the output of a linear filter which uses the LKF estimates as inputs. More recently Wolfe et al. [13] showed that the computational load of the central node can be reduced even further by running on each sensor a local filter which generates a partial estimate of the state so that the central node just need to sum them together to recover the CKF estimate. The main difference between [12] [8] and [13] is that in the latter approach all local sensors need to know the whole system dynamics including the other sensors, while in the former approaches only the central node needs to know the whole system dynamics. There are also dedicated distributed estimation algorithms such as the federated filters proposed by Carlson [4], which require a specific structure in the systems dynamics.

The other class of works is related to 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 tradeoffs 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

L. Schenato is with the Department of Information Engineering, University of Padova, Italy schenato@dei.unipd.it

A. Chiuso is with the Department of Management and Engineering, University of Padova, Vicenza, Italy, chiuso@dei.unipd.it

This work has been partially supported by European Community's Seventh Framework Programme under agreement n. FP7-ICT-223866-FeedNetBack, by CARIPARO Foundation project "Wise-Wai" and by the PRIN Project "New Methods and Algorithms for Identification and Adaptive Control of Technological Systems".

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 nonlinearly 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:

$$\begin{array}{rcl} x_{t+1} &=& Ax_t + w_t \\ y_t^i &=& C_i x_t + v_t^i, \quad i = 1, \ldots, N \end{array} \tag{1}$$

where $x \in \mathbb{R}^n$, $y_i \in \mathbb{R}^{m_i}$, w_t and v_t^i are uncorrelated, zero-mean, white Gaussian noises with covariances $\mathbb{E}[w_t w_t^T] = Q$, and $\mathbb{E}[v_t^i (v_t^j)^T] = R_{ij}$, i.e. we also allow for correlated measurement noise. More compactly, if we define the compound measurement column noise vector $v_t = (v_t^1, \ldots, v_t^N) \in \mathbb{R}^m, m = \sum_i m_i$, we have $\mathbb{E}[v_t v_s^T] = R\delta(t-s)$, 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 $\mathbb{E}[x_0 x_0^T] = P_0$. We also assume that $R > 0$, the pair $(A, Q^{1/2})$ is reachable and (A, C) is observable, where $C^T = [C_1^T \ C_2^T \ \ldots \ C_N^T]$, 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 $\gamma_t^i \in \{0, 1\}$ such that:

$$\gamma_t^i = \begin{cases} 0 & \text{if packet sent at time } t \text{ by node } i \text{ is lost} \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

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 $z_t^i \in \mathbb{R}^\ell$:

$$z_t^i = f_t^i(y_1^i, y_2^i, \ldots, y_t^i) = f_t^i(y_{1:t}^i) \tag{3}$$

where ℓ is bounded and $f_t^i()$ are causal functions of the local measurements. Natural choices are $z_t^i = y_t^i$, i.e. the latest measurement, or the output of a (time varying) linear filter:

$$\begin{array}{c} \xi_t^i = F_t^i \xi_{t-1}^i + G_t^i y_t^i \\ z_t^i = H_t^i \xi_t^i + D_t^i y_t^i \end{array}$$

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

$$\mathcal{I}_t \;\; = \;\; \bigcup_{i=1}^N \mathcal{I}_t^i, \quad \mathcal{I}_t^i = \{z_k^i \,|\, \gamma_k^i = 1, k = 1, \ldots, t\} \tag{4}$$

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

$$\hat{x}_{t|t}^g = g_t(\mathcal{I}_t) \tag{5}$$

such that its error $P_{t|t}^g = var(x_t - \hat{x}_{t|t}^g \,|\, \mathcal{I}_t) = \mathbb{E}[(x_t - \hat{x}_{t|t}^g)(x_t - \hat{x}_{t|t}^g)^T \,|\, \mathcal{I}_t]$ is small. Depending on the choice of the sensor preprocessing functions f_t^i and the estimator functions g_t, we get different strategies. Note that the estimator error covariance $P_{t|t}^g$ 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_t^i 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:

$$z_t^i = y_{1:t}^i$$

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

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

where τ_t^i is the delay of the most recent received packet. This filter is optimal among all possible strategies since it is easy to see that

$$P_{t|t}^{IBF} \leq P_{t|t}^g, \quad \forall f_t^i(), \forall \gamma_t^i, \forall g_t()$$

where $P_{t|t}^{IBF} = var(x_t - \hat{x}_{t|t}^{IBF} \,|\, \mathcal{I}_t)$ 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 same 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 $\hat{x}^{g}_{t|t}$ and $\hat{x}^{IBF}_{t|t}$ defined as above. Then there do not exist (possibly nonlinear) functions $z^i_t = f^i_t(y^i_{1:t}) \in \mathbb{R}^\ell$ with bounded size $\ell < \infty$ such that $P^g_{t|t} = P^{IBF}_{t|t}$ for any possible packet loss sequence, i.e.

$$\nexists f^i_t() \,|\, P^g_{t|t} = P^{IBF}_{t|t}, \ \forall \gamma^i_t$$

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

$$z^i_t = y^i_t$$

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

$$\hat{x}^{MF}_{t|t} = \mathbb{E}[x_t \,|\, \mathcal{I}_t, \, i = 1, \dots, N]$$

where the information set in this case corresponds to $\mathcal{I}^i_t = \{y^i_k \,|\, \gamma^i_k = 1, \, k = 1, \dots, t\}$. 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 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 systems 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:

$$\bar{C}_t = \begin{bmatrix} \gamma^1_t C_1 \\ \vdots \\ \gamma^N_t C_N \end{bmatrix}, \quad \bar{y}_t = \begin{bmatrix} \gamma^1_t y^1_t \\ \vdots \\ \gamma^N_t y^N_t \end{bmatrix},$$
$$\bar{R}_t = diag\{\gamma^1_t R_{11}, \gamma^2_t R_{22}, \dots \gamma^N_t R_{NN}\}$$

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 $y_t = (y^1_t \, y^2_t \, \dots \, y^N_t)^T$ 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:

$$\hat{x}^{MF}_{t|t} = (I - \bar{L}_t \bar{C}_t) A \hat{x}^{MF}_{t-1|t-1} + \bar{L}_t \bar{y}_t \tag{6}$$
$$P^{MF}_{t|t} = P_{t|t-1} - P_{t|t-1} \bar{C}^T_t (\bar{C}_t P_{t|t-1} \bar{C}^T_t + \bar{R}_t)^\dagger \bar{C}_t P_{t|t-1} \tag{7}$$
$$\bar{L}_t = P_{t|t-1} \bar{C}^T_t (\bar{C}_t P_{t|t-1} \bar{C}^T_t + \bar{R}_t)^\dagger \tag{8}$$
$$P_{t+1|t} = A P^{MF}_{t|t} A^T + Q \tag{9}$$

where the symbol † 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 (OLPEF) 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 $x^{CKF}_{t|t} := \mathbb{E}[x_t | y^i_{1:t}, i = 1, .., N]$ the centralized Kalman filter (CKF). Its evolution is governed by the equations:

$$\begin{aligned} \hat{x}^{CKF}_{t|t} &= F_t \hat{x}^{CKF}_{t-1|t-1} + L_t y_t \\ F_t &= (I - L_t C) A \end{aligned}$$

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

$$\begin{aligned} P_{t+1} &= (A - K_t C) P_t (A - K_t C)^T + K_t R K^T_t + Q \\ L_t &= P_t C^T (C P_t C^T + R)^{-1} \\ K_t &= A L_t \end{aligned}$$

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

$$z^i_t = F_t z^i_{t-1} + L^i_t y^i_t, \tag{10}$$

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.:

$$x^{OLPEF}_{t|t} = \sum_{i=1}^N A^{\tau^i_t} z^i_{t-\tau^i_t} \tag{11}$$

where τ_t^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

$$P_{t|t}^{IBF} = P_{t|t}^{OLPEF} < P_{t|t}^{MF}$$

E. Simulations

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

As expected OLPEF and IBF are indistinguishable[1] for small μ_Q , i.e. for small process noise, while OLPEF becomes significantly worse ad μ_Q becomes large. Note also that for small μ_Q, which gives a filter with "long" memory, IBF is better than MF while for large μ_Q, and consequently a short memory filter, MF and IBF become indistinguishable.

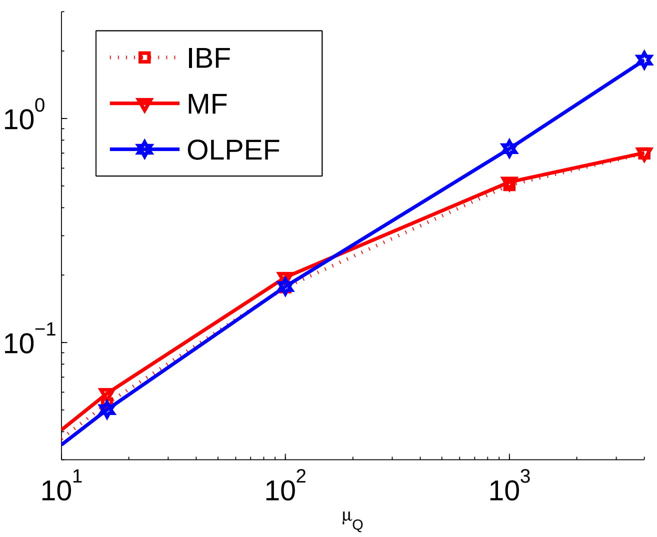

Fig. 1. Error Variance vs. μ_Q.

III. Bounds on Estimation Error Variance

Let us first introduce some notation. In order to simplify the analysis we shall consider the case in which N identical and independent sensors measure the same state, i.e. C_i do not depend upon i and $R = \text{block diag}\{R_{11}, ..., R_{NN}\}$, $R_{ii} = R_{jj}, \forall (i, j)$.

Let us define:

$$\mathcal{L}_f(P, L, \ell) := (I - LC)P(I - LC)^\top + \frac{1}{\ell} LRL^\top \quad (12)$$

This is the state estimation (filtering) error using the gain L when the initial state estimate has variance P and measurements from ℓ sensors are utilized. The optimal (Kalman) gain $K := L_{opt}$ can be obtained by minimizing $\mathcal{L}(P, L, \ell)$ with respect to L, obtaining

$$K = \arg\min_L \mathcal{L}(P, L, \ell) = PC^\top \left(CPC^\top + \frac{1}{\ell}R\right)^{-1} \quad (13)$$

The achieved optimal prediction error is given by

$$\begin{aligned} \Phi_f(P, \ell) &:= \mathcal{L}(P, K, \ell) \\ &= (I - KC)P(I - KC)^\top + \tfrac{1}{\ell} KRK^\top \\ &= P - PC^\top \left(CPC^\top + \tfrac{1}{\ell} R\right)^{-1} CP \end{aligned} \quad (14)$$

For future use we also define the prediction error variance update

$$\Phi(P, \ell) \quad := \quad A\Phi_f(P, \ell)A^\top + Q \quad (15)$$

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 ℓ.

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

As far as convexity in ℓ, the following argument can be used: assume h is (positive) real variable and consider the derivatives $\frac{d\Phi_f(P,\ell)}{d\ell}$ and $\frac{d^2\Phi_f(P,\ell)}{d\ell^2}$. It is easy to verify that $\frac{d\Phi_f(P,\ell)}{d\ell} < 0$ and $\frac{d^2\Phi_f(P,\ell)}{d\ell^2} > 0$. The conclusion for $\Phi(P, \ell)$ follows from the fact that $\Phi(P, \ell)$ 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 $\mathcal{P} := \{P = P^\top : P \geq P_m, P \leq P_M\}$. 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 $\bar{P} \in \mathcal{P}$ one can take $K_{\bar{P}} := \text{argmin}_L \mathcal{L}(\bar{P}, L, \ell)$. Then

$$\mathcal{L}(P, K_{\bar{P}}, \ell) \geq \Phi(P, \ell) \quad \forall P = P^\top \geq 0$$

holds. Among all such $K_{\bar{P}}$ we would like to find that which minimizes the maximal difference

$$S(K_{\bar{P}}) := \max_{P \in \mathcal{P}} \mathcal{L}(P, K_{\bar{P}}, \ell) - \Phi(P, \ell)$$

choosing

$$\begin{aligned} K_{opt} &:= \underset{K_{\bar{P}}}{\text{argmin}}\ S(K_{\bar{P}}) \\ &= \underset{K_{\bar{P}}}{\arg\min} \left[\max_{P \in \mathcal{P}} \mathcal{L}(P, K_{\bar{P}}, \ell) - \Phi(P, \ell)\right]. \end{aligned}$$

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 $\mathcal{L}(P, K_{opt}, \ell) - \Phi(P, \ell) \leq S(K_{opt}),\ \forall P \in \mathcal{P}$

$$\mathcal{L}_f^{LB}(P, \ell) := \mathcal{L}(P, K_{opt}, \ell) - S(K_{opt}) \leq \Phi_f(P, \ell) \quad \forall P \in \mathcal{P}. \quad (16)$$

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 $P_m = \Phi(P_m, N)$, 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

[1] Small differences are due to the sample variability of Montecarlo estimates.

error when no information is available. It is a standard fact to show that, provided $P \in \mathcal{P}$, also $\Phi(P,\ell) \in \mathcal{P}$.

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

We shall also need to consider a linear lower bound $\mathcal{L}^{LB}(P,\ell)$ for $\Phi(P,\ell)$. It is immediate to see that

$$\mathcal{L}^{LB}(P,\ell) = A\mathcal{L}_f^{LB}(P,\ell)A^\top + Q.$$

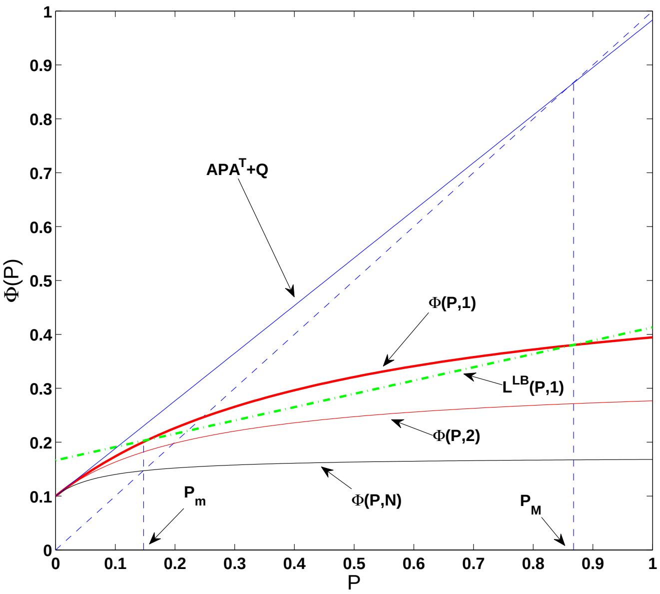

Fig. 2. Graphical representation of bounding functions for scalar system with $A = 0.9, C = 1, Q = 1, R = 10, N = 6$.

The following lemma gives a very simple expression of this lower bound for scalar state space systems. The corresponding functions are graphically portrayed in Fig. 2.

Lemma 2: For system with scalar state spaces, i.e. $n = \dim\{x\} = 1$, the function $\mathcal{L}_f^{LB}(P,\ell)$ admits the very simple closed form expression

$$\mathcal{L}_f^{LB}(P,\ell) = \frac{\Phi_f(P_M,\ell) - \Phi_f(P_m,\ell)}{P_M - P_m}(P - P_m) + \Phi_f(P_m,\ell) \quad (17)$$

Proof: The proof is just based on the observation that (17) is nothing but the line going through the points of coordinates $(P_m, \Phi_f(P_m,\ell))$ and $(P_M, \Phi_f(P_M,\ell))$. Of course concavity of $\Phi_f(P,\ell)$ guarantees that this line is below $\Phi_f(P,\ell)$ for all $P \in \mathcal{P}$. This is indeed the "optimal" approximation from below, i.e. the linear function in P which minimizes the maximal difference $\Phi_f(P,\ell) - \mathcal{L}_f^{LB}(P,\ell)$ $P \in \mathcal{P}$. ■

A. Upper Bound for Measurement Fusion

In [9] it is proposed a suboptimal filter for the measurement fusion where a constant gain $\bar{L}$ is used rather than the optimal one L_t of Eqn. (6)-(9), which is time varying. This suboptimal filter can be written as:

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

It has been shown that the steady state minimum expected error covariance for this filter, i.e. $S = \min_{\bar{L}} \lim_{t\to\infty} \mathbb{E}_\gamma[var(x_{t+1} - \check{x}_{t+1|t}^{MF})]$ is given by the fixed point of the following operator:

$$\begin{aligned}\Psi_\lambda(S) &= ASA^T + Q - \\ &- \lambda ASC^T(\lambda CSC^T + (1-\lambda)S_C + R)^{-1}CSA^T \\ S_C &= \mathrm{diag}(C_1SC_1^T, \ldots, C_NSC_N^T)\end{aligned}$$

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}^f respectively the steady state (w.r.t. the measure induced by the loss mechanism) prediction and filtering state estimation errors. Then

$$\mathbb{E}[P_{MF}] \geq \bar{P}_{MF}^{LB} \quad \mathbb{E}[P_{MF}^f] \geq \bar{P}_{MF}^{f,LB}$$

where $\bar{P}_{MF}^{LB}$ is the unique stationary solution of $\bar{P}_{MF}^{LB} = \mathcal{L}^{LB}(\bar{P}_{MF}^{LB}, \mathbb{E}\ell)$ and $\bar{P}_{MF}^{f,LB} = \mathcal{L}_f^{LB}(\bar{P}_{MF}^{LB}, \mathbb{E}\ell)$.

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 ℓ, it follows that

$$\mathbb{E}[P_{t+1}|P_t] \geq \Phi(P_t, \mathbb{E}\ell_t)$$

where independence of ℓ_t and P_t has been used. Using the lower bound $\Phi(P,\ell) \geq \mathcal{L}^{LB}(P,\ell)$ it follows that

$$\mathbb{E}[P_{t+1}|P_t] \geq \mathcal{L}^{LB}(P_t, \mathbb{E}\ell_t).$$

Since $\mathcal{L}^{LB}$ is linear in P_t, also

$$\mathbb{E}[P_{t+1}] \geq \mathcal{L}^{LB}(\mathbb{E}P_t, \mathbb{E}\ell_t) \quad (18)$$

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

$$\mathbb{E}[P_{MF}] \geq \bar{P}_{MF}^{LB}, \qquad \bar{P}_{MF}^{LB} = \mathcal{L}^{LB}(\bar{P}_{MF}^{LB}, \mathbb{E}\ell)$$

The bound for the filtering solution is easily obtained observing that $P_t^f = \Phi(P_t,\ell_t) \geq \mathcal{L}_f^{LB}(P_t,\ell)$ so that

$$\mathbb{E}P_t^f \geq \mathcal{L}_f^{LB}(\mathbb{E}P_t, \mathbb{E}\ell_t)$$

and therefore

$$\mathbb{E}P_{MF}^f \geq \bar{P}_{MF}^{f,LB} := \mathcal{L}_f^{LB}(\bar{P}_{MF}^{LB}, \mathbb{E}\ell).$$

■

C. Lower Bound for Infinite Bandwidth Filter

The estimator $\hat{x}^{IBF}_{t|t}$ is characterized by the numbers $\tau^1_t, ..., \tau^N_t$, where τ^i_t 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 $h_t, h_{t-1}, h_{t-2}, ...$ defined as:

$$h_{t-i} = \sum_{j=1}^{N} \delta(\tau^j_{t-i} - i)$$

where $\delta(\cdot)$ is the Kronecker delta. As mentioned above the IBF can be though of as a measurement fusion filter where the equivalent numbers of packets arrived at time t are defined, recursively by the relation

$$\begin{cases} \ell_t &= h_t \\ \ell_{t-k} &= \ell_{t-k+1} + h_{t-k} \quad k = 1, 2, ... \end{cases}$$

It is fairly easy to see that the joint probability density function of the variables variables ℓ_{t-k} can be written in terms of the conditional densities $p(\ell_{t-k-1}|\ell_{t-k}, \ell_{t-k+1}, ..., \ell_t) = p(\ell_{t-k-1}|\ell_{t-k})$, which have the expression

$$p(\ell_{t-k-1} = \ell|\ell_{t-k}) = \begin{pmatrix} N - \ell_{t-k} \\ \ell - \ell_{t-k} \end{pmatrix} \lambda^{N-\ell}(1-\lambda)^{\ell-\ell_{t-k}}$$

where λ 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^f_{IBF} and P_{IBF} be the (steady state) state filtering and prediction error variance using the IBF. Let us denote with $P^f_{IBF}(t, k, \ell_{t-k+1}, .., \ell_t)$ the state filtering error variance at time t and with $P_{IBF}(t+1, k, \ell_{t-k+1}, .., \ell_t)$ the state prediction error variance at time $t+1$ conditioned on $\ell_{t-k} = N$ and with subsequent number of arrivals $\ell_{t-k+1}, .., \ell_t$.

It is clear that

$$\begin{aligned} \bar{P}^f_{IBF}(k) &:= \mathbb{E}\left[P^f_{IBF}(t, k, \ell_{t-k+1}, .., \ell_t)|\ell_{t-k} = N\right] \\ \bar{P}_{IBF}(k) &:= \mathbb{E}\left[P_{IBF}(t+1, k, \ell_{t-k+1}, .., \ell_t)|\ell_{t-k} = N\right] \\ &= A\bar{P}^f_{IBF}(k)A^\top + Q \end{aligned} \tag{19}$$

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

$$\begin{aligned} \mathbb{E}[P^f_{IBF}] = \bar{P}^f_{IBF}(\infty) \geq \bar{P}^f_{IBF}(k) \geq \bar{P}^f_{IBF}(1) \\ \mathbb{E}[P_{IBF}] = \bar{P}_{IBF}(\infty) \geq \bar{P}_{IBF}(k) \geq \bar{P}_{IBF}(1) \end{aligned} \quad \forall k.$$

The following proposition holds.

Proposition 2: The matrices $\bar{P}^f_{IBF}(k)$ and $\bar{P}_{IBF}(k)$ in (19) admits the lower bounds as $\bar{P}^f_{IBF}(k) \geq \bar{P}^{f,LB}_{IBF}(k)$ and $\bar{P}_{IBF}(k) \geq \bar{P}^{LB}_{IBF}(k)$ the inequalities:

$$\begin{aligned} \bar{P}^{f,LB}_{IBF}(1) &= \Phi_f(P_m, \mathbb{E}\ell_t) = \Phi_f(P_m, N(1-\lambda)) \\ \bar{P}^{f,LB}_{IBF}(2) &= \mathbb{E}\left[\mathcal{L}^{LB}_f(\Phi(P_m, \mathbb{E}[\ell_{t-1}|\ell_t]), \ell_t)\right] \\ \bar{P}^{f,LB}_{IBF}(k) &= \mathbb{E}\left[\mathcal{L}^{LB}_f \circ \cdots \circ \mathcal{L}^{LB} \circ \Phi(P_m, \mathbb{E}[\ell_{t-k+1}|\ell_{t-k+2}])\right]. \end{aligned}$$

and

$$\begin{aligned} \bar{P}^{LB}_{IBF}(1) &= \Phi(P_m, \mathbb{E}\ell_t) = \Phi(P_m, N(1-\lambda)) \\ \bar{P}^{LB}_{IBF}(2) &= \mathbb{E}\left[\mathcal{L}^{LB}(\Phi(P_m, \mathbb{E}[\ell_{t-1}|\ell_t]), \ell_t)\right] \\ \bar{P}^{LB}_{IBF}(k) &= \mathbb{E}\left[\mathcal{L}^{LB} \circ \cdots \circ \mathcal{L}^{LB} \circ \Phi(P_m, \mathbb{E}[\ell_{t-k+1}|\ell_{t-k+2}])\right]. \end{aligned}$$

where $\mathbb{E}[\ell_{t-k}|\ell_{t-k+1}] = \ell_{t-k+1} + (1-\lambda)(N - \ell_{t-k+1})$ 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 $k = 1$, $P_{IBF}(t+1, 1, \ell_t) = \Phi(P_m, \ell_t)$. Then, using convexity of Φ in ℓ_t it follows that

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

When $k = 2$,

$$\begin{aligned} P_{IBF}(t+1, 2, \ell_{t-1}, \ell_t) &= \Phi(\Phi(P_m, \ell_{t-1}), \ell_t) \\ &\geq \mathcal{L}^{MF}(\Phi(P_m, \ell_{t-1}), \ell_t) \end{aligned}$$

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

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

from which,

$$\begin{aligned} \bar{P}_{IBF}(2) &:= \mathbb{E}[\mathbb{E}[P_{IBF}(t, 2, \ell_{t-1}, \ell_t)|\ell_t]] \\ &\geq \mathbb{E}\left[\mathcal{L}^{MF}(\Phi(P_m, \mathbb{E}\ell_{t-1}|\ell_t), \ell_t)\right] \end{aligned}$$

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

Remark 2: In practice one can compute $P^{f,LB}_{IBF}(k)$ 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.

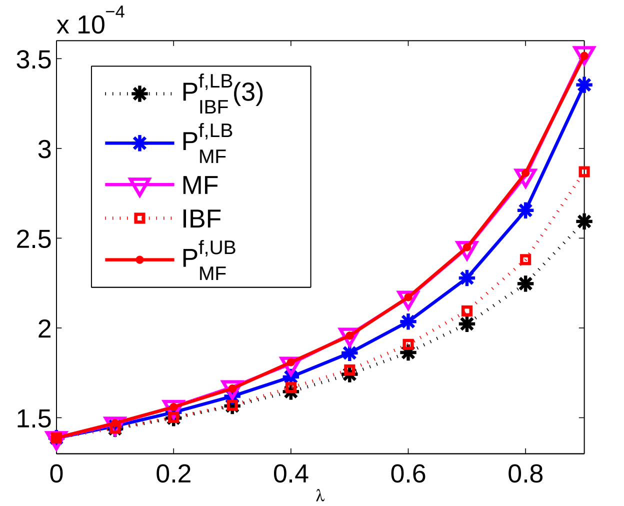

Fig. 3. Error Variance vs. packet loss probability λ: Montecarlo simulations (MF and IBF) vs. analytical bounds.

Remark 3: Note that the lower bound for the IBF provides also a lower bound for MF. Therefore, one can use, as a lower bound for MF whichever is larger among $\bar{P}^{f,LB}_{MF}$ and $\bar{P}^{f,LB}_{IBF}$.

IV. Simulation Results

In order to illustrate our results we consider a simple (scalar) example described by the equations:

$$\begin{cases} x_{t+1} &= 0.9x_t + w_t \\ y^i_t &= x_t + v^i_t \quad i = 1, .., N = 25 \end{cases} \tag{20}$$

where w_t is white and Gaussian zero mean and variance 10^{-4} and the v_t^i's are white, independent of each other and of w_t, Gaussian zero mean and variance 10^{-2}. The packet loss probability λ is varied in the range $\lambda \in [0, 0.9]$.

The results of a simulation are reported in figure 3. The variances for MF and IBF are computed by averaging over a Montecarlo run of 1000 experiments the (filtering) variance $P_{t|t}$ which can be computed, in closed form, for a given packet loss sequence.

In the specific example the performance of the MF algorithm are indistinguishable from the upper bound.

V. Conclusions

In this paper we derived some analytical upper and lower bounds for different strategies addressing the 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 though 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

[1] A. Agnoli, A. Chiuso, P. D'Errico, A. Pegoraro, and L. Schenato, "Sensor fusion and estimation strategies for data traffic reduction in rooted wireless sensor networks," in *Proc. of IEEE ISCCSP'08*, La Valletta, Malta, 2008, pp. 677–682.

[2] Y. Bar-Shalom, H. Chen, and M. Mallick, "One-step solution for the general out-of-sequence measurement problem in tracking," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 40, no. 1, pp. 27–37, 2004.

[3] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., 2001.

[4] N. Carlson, "Federated square root filter for decentralized parallel processors," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 26, no. 3, pp. 517–525, 1990.

[5] A. Chiuso and L. Schenato, "Information fusion strategies from distributed filters in packet-drop networks," in *Proc. of CDC 08*, 2008, pp. 1079–1084.

[6] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "Optimal LQG control across a packet-dropping link," *Systems and Control Letters*, vol. 56, no. 6, pp. 439–446, 2007.

[7] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE, special issue*, vol. 95, no. 1, pp. 138–162, January 2007.

[8] B. Levy, D. Castañon, G. Verghese, and A. Willsky, "A scattering framework for decentralized estimation problems," *Automatica*, vol. 19, no. 4, pp. 373–384, 1983.

[9] L. Schenato, "Optimal sensor fusion for distributed sensors subject to random delay and packet loss," in *Proc. of IEEE Conf. on Decision and Control*, New Orleans, U.S.A., 2007, pp. 1547–1552.

[10] L. Shi, K. Johansson, and R. Murray, "Estimation over wireless sensor networks: Tradeoff between communication, computation and estimation qualities," in *Proc. of IFAC World Congress*, vol. 17, no. 1, Seoul, Korea, 2008.

[11] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, September 2004.

[12] A. Willsky, M. Bello, D. Castanon, B. Levy, and G. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Transactions on Automatic Control*, vol. 27, no. 4, pp. 799–813, 1982.

[13] J. Wolfe and J. Speyer, "A low-power filtering scheme for distributed sensor networks," in *Proceeding of IEEE Conference on Decision and Control (CDC'03)*, 2003, pp. 6325–6326.

[14] K.-S. Zhang, X. R. Li, and Y.-M. Zhu, "Optimal update with out-of-sequence observations for distributed filtering," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 1992–2004, 2005.