Optimal sensor fusion for distributed sensors subject to random delay and packet loss

Luca Schenato

Abstract— In this paper we study optimal information fusion for sampled linear systems where the sensors are distributed and measurements are collected to central unit via a wireless network. Every sensor measurement is subject to random delay or might even be completely lost. We show that optimal sensor fusion consist in a time-varying Kalman filter with bufferized measurements. We also propose a suboptimal but computationally efficient fusion architecture based on a bank of static gains that can be optimally designed if packet delay statics are known. Finally, algorithms to check for the existence of stable estimators and to evaluate their error covariance are given and some special cases are analyzed.

Index Terms— Sensor fusion, packet drop, random delay, remote estimation, stability, Kalman filtering

I. INTRODUCTION

Continuing improvements in IC and wireless technologies are promoting the appearance of inexpensive devices provided with sensing, communication and data processing capabilities, thus revolutionizing our ability to build massively distributed networked systems [1]. These systems, commonly referred as wireless sensor networks (WSNs), can offer access to an unprecedented quality and quantity of information which can revolutionize our ability to monitor and control the environment [2] [3]. However, they also pose challenging problems such as sensor power scheduling, packet routing, packet loss compensation, data compression and coding, just to name a few, arising from the fact that sensors, actuators and controllers are not physically co-located, are often battery-powered and need to exchange information via a wireless communication network. These are just few problems that appear in data-collection applications, whose goal is to accurately reconstruct the sensor measurements. Real-time monitoring and control applications add another level of difficulty since not only packet loss but also delay of gathered data can undermine the effectiveness of the decision-making part. Therefore, it is important to evaluate the impact of random packet delay and packet loss in the overall system performance. Obviously, packet loss and delay can be reduced by network coding, distributed signal processing, in-network data compression, and packet routing protocols, but not completely avoided due to inherent unreliable nature of wireless communication. Indeed, different communication protocol architectures can reduce packet loss at the price of larger packet delay and viceversa. For example, fixed time-division communication scheduling can reduce packet loss and power consumption at the price of longer delivery [4], while randomized geographical routing can reduce packet delay at the price of larger packet loss [5].
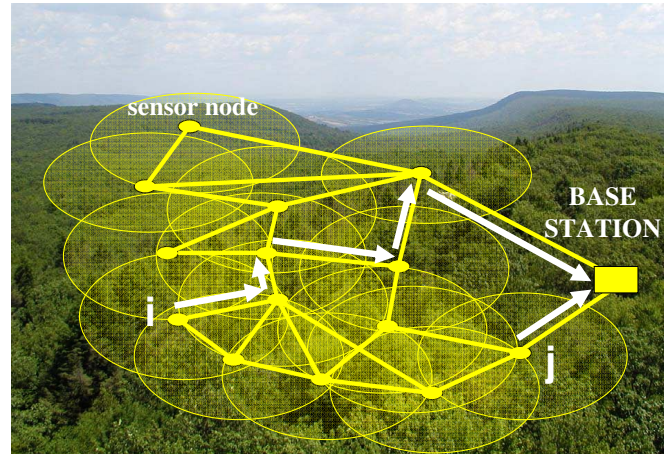
Fig. 1. Pictorial representation of Wireless Sensor Network for environment monitoring. The small dots indicate the location of the sensing nodes, the shaded circles indicate the sensing regions and the segments the communication links. Sensor network topology comprising M sensors showing routing path from nodes i and node j to sensor fusion location.

Currently, communications protocols and networked control systems are designed separately. In particular, protocols are design based on conservative heuristics which specify what the maximum time delay and maximum packet loss should be, but with no clear understanding of their impact on the overall application performance. On the application side, control systems are not specifically designed to exploit information about packet loss and delay statistics of the communication protocols over which they will run.

Motivated by these considerations, the goal of this paper is to study optimal sensor fusion where measurements are relayed via a wireless communication network to a single location center and then fused to obtain the estimate of the system state. Each sensor measurement is subject to its own packet delay and packet loss probability, thus allowing to associate different packet arrival distributions depending on the distance of each sensor from the centralized fusion node.

II. PREVIOUS WORK AND CONTRIBUTION

Recently several groups have looked at networked control systems with large random delay or packet loss. The survey paper [6] nicely reviews several results in this area. These results can be divided into two main groups: the first group focuses on variable delay but no packet drop, while the second group focuses on packet loss but no delay.

Within the first group, Nilsson at al. [7] extended LQG optimal control design to sampled linear systems subject to stochastic measurement and control packet delay, and showed how the optimal controller gains are time-delay dependent. However, their work was limited to co-located sensors and actuators. Differently, other authors studied data

fusion of measurements obtained from senors with different delays. For example, Alexander [8] and Larsen et al. [9] derived suboptimal but computationally efficient Kalman-like filters to account for random delay and they tested their performance through Monte-Carlo simulations. Julier at al. [10] studied the estimation problem when measurement time-stamping is uncertain. All the previous works rely on the major assumption that there is no packet loss or there are at most m consecutive packet drops.

In the second group of results, there has been a considerably effort to apply optimal control and estimation to discrete time systems where measurements and control packets can be dropped with some probability, but have otherwise no delay. This framework is equivalent of saying that all packets have either no delay or infinite delay. For example, in [11][12][13] the authors proposed compensation techniques for i.i.d Bernoulli packet-drop communication networks and derived stability conditions for closed loop discrete time system. Sinopoli at al. [14] [15], and Imer et al. [16], extending results that can be traced back to Nahi [17], looked specifically at minimum variance estimation and optimal control. However, all these works where related to single sensor and single actuator, and only Liu et al. [18]. Very little work has been done to take into account simultaneous packet drop and packet delay, leading to somewhat conservative results as they are based on worst-case scenarios [19] [20].

In this paper we propose a probabilistic framework to analyze sensor fusion where observation packets from multiple sensors are subject to arbitrary random delay and packet loss. In this framework sensor measurements need to be time-stamped at the sensor side, but packets can arrive in burst or even out of order at the receiver side. This paper extends a previous work of the author [21] where only a single sensor was considered. This estimator structure is independent of the packet arrival statistics and can be implemented using a Kalman filter with finite memory buffer if the delivered packets have a finite maximum delay. In particular, the memory length is equal to the maximum packet delay of the arrived packets. We also present an alternative sensor fusion architecture which is computationally more efficient since it does not require any matrix inversion, but simply associate a constant gain to every slot of the measurement buffer at the estimator location. Necessary and sufficient conditions for stability for these estimators are shown to depend only on the packet loss probability of each sensor. We also provide quantitative measures for the expected error covariance of such estimators which turns out to be the solution of modified algebraic Riccati equations and Lyapunov equations. These metrics can be used to compare different communication protocols for real-time control applications as long as the packet arrival statistics are known, i.i.d and stationary. Very importantly, these results do not depend on the specific implementation of the digital communication network (fieldbuses, Bluetooth, ZigBee, Wi-Fi, etc ..) in the sense that it is not necessary to modify the communication stack to implement the estimators.

III. PROBLEM FORMULATION

Consider the following discrete time linear stochastic plant:

$$x_{t+1} = Ax_t + w_t \qquad (1)$$
$$y_t^i = C_i x_t + v_t^i, \quad i = 1, \ldots, M \qquad (2)$$

where M is the total sensor number, $t \in \mathbb{N} = \{0, 1, 2, \ldots\}$, $x, w \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, $y \in \mathbb{R}^{m_i}$, $C_i \in \mathbb{R}^{m_i \times n}$, $(x_0, w_t, v_t^1, ..., v_t^M)$ are Gaussian, uncorrelated, white, with mean $(\bar{x}_0, 0, 0, ..., 0)$ and covariance $(P_0, Q, R_1, ..., R_M)$ respectively. Let also consider the full observation matrix $C \in \mathbb{R}^{m \times n}$ where $m = \sum_i m_i$ defined as:

$$C = \begin{bmatrix} C_1 \\ \vdots \\ C_M \end{bmatrix}$$

We also assume that the pair (A, C) is observable, $(A, Q^{1/2})$ is reachable, and $R_i > 0, \forall i$ [1]. Measurements are time-
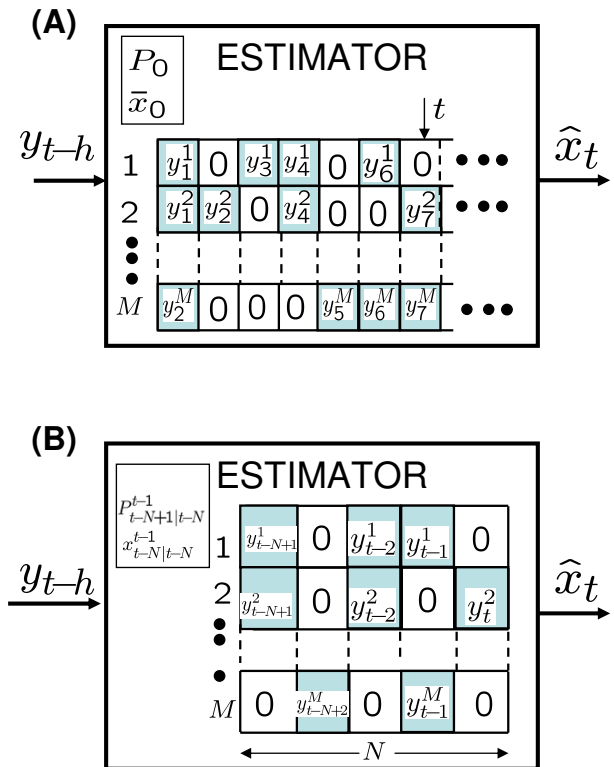

Fig. 2. Optimal Kalman estimator memory requirements for general arrival processes (A). Optimal Kalman estimator with finite memory buffer (B).

stamped, encapsulated into packets, and then transmitted through a digital communication network (DCN), whose goal is to deliver packets from a source to a destination. Time-stamping of measurements is necessary to reorder packets at the receiver side as they can arrive out of order. Modern DCNs are in general very complex and can greatly differ in their architecture and implementation depending on the medium used (wired, wireless, hybrid), and on the applications they are meant to serve (real-time monitoring, data extraction, media-related, etc ..). In our work we model a DCN as a module between the sensors and the fusion center which delivers observation measurements to the estimator with possibly random delays. This model allows also for

---

[1] These assumptions can be relaxed to (A, C) detectable, $(A, Q^{1/2})$ stabilizable, and $R_i \geq 0$, however the proofs of the following theorems would be more convoluted, therefore we decided to adopt the stronger hypotheses.

packets with infinite delay which corresponds to a packet loss. We assume that all observation packets correctly delivered to the estimator site are stored in an infinite buffer, as shown in top panel of Fig. 2. The arrival process is modeled via the random variables $\gamma_{t,k}^i$ defined as follows:

$$\gamma_{t,k}^i = \begin{cases} 1 & \text{if } y_k^i \text{ arrived before or at time } t, \ t \geq k \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

From this definition it follows that $(\gamma_k^t = 1) \Rightarrow (\gamma_k^{t+h} = 1, \forall h \in \mathbb{N})$, which simply states that if packet y_k is present in the receiver buffer at time t, then it will be present for all future times. We also define the packet delay $\tau_k^i \in \{\mathbb{N}, \infty\}$ for observation y_k^i of sensor node i as follows:

$$\tau_k^i = \begin{cases} \infty & \text{if } \gamma_{t,k}^i = 0, \forall t \geq k \\ t_k^i - k & \text{otherwise, } t_k^i \triangleq \min\{t \mid \gamma_{t,k}^i\} \end{cases} \tag{4}$$

where t_k^i is the arrival time of observation y_k^i of i-th sensor at the sensor fusion site. This measurement model is very general since it allows for observation measurements to arrive out of order at the estimator site. Also it is possible that between two consecutive sampling periods no packet or multiple packets are delivered. In our work we do not consider quantization distortion due to data encoding/decoding since we assume that observation noise is much larger then quantization noise, as it is the case in most DCNs where each packet allocates hundreds of bits for measurement data[2]. Also we do not consider channel noise since we assume that if any bit error incurred during packet transmission is detected at the receiver, then the packet is dropped.

If observation y_k^i is not yet arrived at the estimator at time t, we assume that a zero is stored in the i, k-slot of the buffer, as shown in Fig. 2[3]. More formally, the value stored in the i, k-slot of the estimator buffer at time t can be written as follows:

$$y_{t,k}^i = \gamma_{t,k}^i y_k^i = \gamma_{t,k}^i C_i x_k + \gamma_{t,k}^i v_k^i \tag{5}$$

We can also define the total measurement vector $\tilde{y}_{t,k} = [\tilde{y}_{t,k}^1, \ldots, \tilde{y}_{t,k}^M]^T \in \mathbb{R}^m$ at the sensor fusion location that can be written as:

$$y_{t,k} = \begin{bmatrix} \gamma_{t,k}^1 C_1 \\ \gamma_{t,k}^2 C_2 \\ \vdots \\ \gamma_{t,k}^M C_M \end{bmatrix} x_k + \begin{bmatrix} \gamma_{t,k}^1 v_k^1 \\ \gamma_{t,k}^2 v_k^2 \\ \vdots \\ \gamma_{t,k}^M v_k^M \end{bmatrix} = C_{t,k} x_k + v_{t,k}$$

Our goal to compute the optimal mean square estimator $\hat{x}_{t|t}$ based on all the sensor measurements arrived at the sensor fusion location up to time t which is given by:

$$\hat{x}_{t|t} \triangleq \mathbb{E}[x_t \mid \mathbf{y}_t, \boldsymbol{\gamma}_t, \bar{x}_0, P_0] \tag{6}$$

where $\mathbf{y}_t = (y_{t,1}, y_{t,2}, \ldots, y_{t,t})$, $\boldsymbol{\gamma}_t = (\gamma_{t,1}, \gamma_{t,2}, \ldots, \gamma_{t,t})$, and $\gamma_{t,k} = [\gamma_{t,k}^1, \ldots, \gamma_{t,k}^M]^T \in \mathbb{R}^M$. It is important to

remark that the estimator above has the information weather a packet has been delivered or not, and it is not equivalent to computing $\hat{x}_{t|t} \neq \tilde{x}_{t|t} \triangleq \mathbb{E}[x_t \mid \mathbf{y}_t, \bar{x}_0, P_0]$. The latter estimator would in fact consider the zero entries of the buffer as true measurements and not as dummy variables, thus giving rise to a lower performance. It is also useful to design the estimator error and error covariance as follows:

$$e_{t|t} \triangleq x_t - \hat{x}_{t|t} \tag{7}$$

$$P_{t|t} \triangleq \mathbb{E}[e_{t|t} e_{t|t}^T \mid \mathbf{y}_t, \boldsymbol{\gamma}_t, \bar{x}_0, P_0] \tag{8}$$

The estimate $\hat{x}_{t|t}$ is optimal in the sense that it minimizes the error covariance, i.e. given any estimator $\tilde{x}_{t|t} = f(\mathbf{y}_t, \boldsymbol{\gamma}_t)$, where f is a measurable function, we always have

$$\mathbb{E}[(x_t - \tilde{x}_{t|t})(x_t - \tilde{x}_{t|t})^T \mid \tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t, \bar{x}_0, P_0] \geq P_{t|t}.$$

Another property of the mean square optimal estimator is that $\hat{x}_{t|t}$ and its error $e_{t|t} \triangleq x_t - \hat{x}_{t|t}$ are uncorrelated, i.e. $\mathbb{E}[e_{t|t} \hat{x}_{t|t}^T] = 0$. This is a fundamental property since it gives rise to the separation principle for the LQG optimal control, which is of the most widely used tool in control system design [22] [15].

IV. MINIMUM ERROR COVARIANCE ESTIMATOR DESIGN

In this section we want to compute the optimal estimator given by Equation (6). First, it is convenient to define the following variables:

$$\hat{x}_{k|h}^t \triangleq \mathbb{E}[x_k \mid \gamma_{t,h}, \ldots, \gamma_{t,1}, y_{t,h}, \ldots, y_{t,1}, \bar{x}_0, P_0]$$

$$P_{k|h}^t \triangleq \mathbb{E}[(x_k - \hat{x}_{k|h}^t)(x_t - \hat{x}_{k|h}^t)^T \mid \gamma_{t,h}, .., \gamma_{t,1}, y_{t,h}, .., y_{t,1}, \bar{x}_0, P_0]$$

from which it follows that, with a little abuse of notation, $\hat{x}_{t|t} = \hat{x}_{t|t}^t$ and $P_{t|t} = P_{t|t}^t$.

It is also useful to note that at time t the information available at the estimator site, given by Equation (5), can be written as the output of the following system:

$$x_{k+1} = A x_k + w_k \tag{9}$$

$$y_{t,k} = C_{t,k} x_k + \tilde{v}_{t,k}, \quad k = 0, \ldots, t \tag{10}$$

where the random vector $\tilde{v}^{t,k}$ is zero mean white noise with covariance $R_{t,k} = \mathbb{E}[\tilde{v}_k^t (\tilde{v}_k^t)^T] \in \mathbb{R}^{m \times m}$ given as

$$R_{t,k} = \begin{bmatrix} \gamma_{t,k}^1 R_1 & 0 & \cdots \\ 0 & \ddots & \\ \vdots & & \gamma_{t,k}^M R_M \end{bmatrix} = \text{diag}(\gamma_{t,k}^1 R_1, \ldots, \gamma_{t,k}^M R_M)$$

For any fixed t this system can be seen as a linear time-varying system with respect to time k, where the only time-varying elements are the observation matrix $C_{t,k}$ and measurement noise covariance $R_{t,k}$.

We can now state the main theorem of this section:

Theorem 1: Let us consider the stochastic linear system given in Equations (1)-(2), where $R_i > 0$. Also consider the arrival process defined by Equation (3), and the mean square estimator defined in Equation (6). Let $\tilde{y}_{t,k} \in \mathbb{R}^{m_{t,k}}$, $m_{t,k} = \sum_i \gamma_{t,k}^i m_i$ the vector including only the measurements y_k^i which are present at the sensor fusion location at time t, i.e. the vector obtained from $y_{t,k}$ by removing the dummy zero-measurements. Similarly consider $\widetilde{C}_{t,k}$ and $\widetilde{R}_{t,k}$ the matrices from which all the zeros rows corresponding to the dummy zero measurements have been removed. Then we have:

(a) The optimal mean square estimator is given by $\hat{x}_{t|t} = \hat{x}_{t|t}^t$ where:

$$\hat{x}_{0|0}^t = \bar{x}_0, \quad P_{1|0}^t = P_0 \tag{11}$$

$$\hat{x}_{k|k}^t = A\hat{x}_{k-1|k-1}^t + \widetilde{K}_{t,k}(\tilde{y}_{t,k} - \widetilde{C}_{t,k}A\hat{x}_{k-1|k-1}^t) \tag{12}$$

$$\widetilde{K}_{t,k} = P_{k|k-1}^t\widetilde{C}_{t,k}^T(\widetilde{C}_{t,k}P_{k|k-1}^t\widetilde{C}_{t,k}^T + \widetilde{R}_{t,k})^{-1} \tag{13}$$

$$P_{k+1|k}^t = AP_{k|k-1}^t A^T + Q - $$
$$- AP_{k|k-1}^t\widetilde{C}_{t,k}^T(\widetilde{C}_{t,k}P_{k|k-1}^t\widetilde{C}_{t,k}^T + R)^{-1}\widetilde{C}_{t,k}P_{k|k-1}^t A^T \tag{14}$$

for $k = 1, \ldots, t$.

(b) The optimal estimator $\hat{x}_{t|t}$ can be computed iteratively using a buffer of finite length N if $\gamma_{t,k}^i = \gamma_{t-1,k}^i, \forall k \geq 1, \forall t \geq k + N, \forall i$. If this property is satisfied, then $\hat{x}_{t|t} = \hat{x}_{t|t}^t$ where $\hat{x}_{t|t}^t$ is given by Equations (11)-(14) for $t = 1, \ldots, N$ and as follows for $t > N$:

$$\hat{x}_{t-N|t-N}^t = \hat{x}_{t-N|t-N}^{t-1}, \tag{15}$$

$$P_{t-N+1|t-N}^t = P_{t-N+1|t-N}^{t-1} \tag{16}$$

$$\text{Eqns. (12),(13),(14)} \quad k = t-N+1, \ldots, t \tag{17}$$

Proof: (a) Since the information available at the estimator site at time t is given by the time-varying linear stochastic system of Equations (9)-(10), then the optimal estimator is given by its corresponding time-varying Kalman filter [22] [23]:

$$\hat{x}_{k|k}^t = A\hat{x}_{k-1|k-1}^t + K_{t,k}(y_{t,k} - C_{t,k}A\hat{x}_{k-1|k-1}^t)$$

$$K_{t,k} = P_{k|k-1}^t C_{t,k}^T(C_{t,k}P_{k|k-1}^t C_{t,k}^T + R_{t,k})^{\dagger}$$

$$P_{k+1|k}^t = AP_{k|k-1}^t A^T + Q - $$
$$- AP_{k|k-1}^t C_{t,k}^T(C_{t,k}P_{k|k-1}^t C_{t,k}^T + R_{t,k})^{\dagger}C_{t,k}P_{k|k-1}^t A^T$$

$$\hat{x}_{0|0}^t = \bar{x}_0, \quad P_{1|0}^t = P_0$$

where the symbol $\dagger$ denotes the pseudoinverse. Note that using the properties of the pseudoinverse it is easy to verify that

$$C_{t,k}^T(C_{t,k}P_{k|k-1}^t C_{t,k}^T + R_{t,k})^{\dagger}C_{t,k} = \widetilde{C}_{t,k}^T(\widetilde{C}_{t,k}P_{k|k-1}^t\widetilde{C}_{t,k}^T + \widetilde{R}_{t,k})^{-1}\widetilde{C}_{t,k}$$

which is equivalent to consider only the non-dummy observation vector $\tilde{y}_{t,m}$.

(b) Let us consider $t > N$. If $\gamma_{t,k} = \gamma_{t-1,k}, \forall k \geq 1, \forall t \geq k + N$, then also $P_{k+1|k}^t = P_{k+1|k}^{t-1}$ and $\hat{x}_{k|k}^t = \hat{x}_{k|k}^{t-1}$ hold under the same conditions on the indices. In particular it holds for $k = t - N$ which implies $P_{t-N+1|t-N}^t = P_{t-N+1|t-N}^{t-1}$ and $\hat{x}_{t-N|t-N}^t = \hat{x}_{t-N|t-N}^{t-1}$. Therefore, it not necessary to compute $P_{t+1|t}^t$ and $\hat{x}_{t|t}^t$ at any time step t starting from $k = 1$, but it is sufficient to use the values $\hat{x}_{t-N|t-N}^{t-1}$ and $P_{t-N+1|t-N}^{t-1}$ precomputed at the previous time step $t-1$, as in Equations (15) and (16), and then iterate Equations Equations (12)-(14) for the latest N observations. ∎

If there is no packet loss and no packet delay, i.e. $\gamma_k^t = 1, \forall (k,t)$, then Equations (11)-(14) reduce to the standard Kalman filter equations for a time-invariant system. If we consider only colocated sensors, i.e. M=1, then we recover the same results from [21].

The previous theorem states that in general it is necessary to invert up to t matrices at every time step t of dimension $\mathbb{R}^{m_{t,k} \times m_{t,k}}$. However, the optimal estimator can be implemented incrementally according to Equations (15)-(17)

using a buffer of finite length N if all successfully received observations have a delay smaller than N time steps, i.e. $\gamma_k^t = \gamma_k^{t-1}, \forall k \geq 1, \forall t - k \geq N$ (see Fig. 2). This does not mean that *all* packets arrive at the receiver within N time steps, but only that if a packet arrives then it does within N time steps.

It is important to remark that the Kalman gain $K_{t,k}$ and the estimator error covariance $P_{t|t}$ are time-varying since they depend on the packet arrival history γ_t. Differently from the standard Kalman filter which converges to a steady state, in the context of random packet delay and packet loss this filter is strongly time-varying, thus making it less computationally efficient due to all matrix inversions.

Up to this point we made no assumptions on the packet arrival process which can be deterministic, stochastic or time-varying. However, from an engineering perspective it is important to determine the performance of the estimator, which is evaluated based on the error covariance $P_{t+1|t}$. If the packet arrival process is stochastic, then also the error covariance is stochastic. In this scenario a common performance metric is the expected error covariance, i.e. $\mathbb{E}_\gamma[P_{t+1|t}]$, where the expectation is performed with respect to the arrival process γ_k^t. It is not clear whether is it possible to compute $\mathbb{E}_\gamma[P_{t+1|t}]$ analytically even for a simple Bernoulli arrival process, and so far only upper and lower bounds have been obtained [14]. Rather than trying to bound the performance of the time-varying optimal estimator, we will focus on a sensor fusion filter with a bank of constant gains associated to each slot in the $M \times N$ buffer of Figure 2(b), i.e. K_h^i for all $i = 1, ..., M$ and $h = 0, \ldots, N-1$. The gains K_h^i will then be optimized to achieve the smallest error covariance at steady-state. The advantage of using constant gains is that it is not necessary to invert any matrix at all differently from the optimal time-varying filter, thus making it attractive for on-line applications. Moreover, since filters with constant gains are necessarily suboptimal, the computation of their error covariance is useful per se as it provides an upper bound for the error covariance of the optimal minimum error covariance filter given by Equations (11)-(14).

V. OPTIMAL FILTERING WITH CONSTANT GAINS

In this section we will study minimum error covariance sensor fusion filters with constant gains under stationary i.i.d arrival processes.

Assumption: The packet arrival process at the estimator site is stationary and i.i.d. with the following probability function:

$$\mathbb{P}[\tau_t \leq h] = \lambda_h \tag{18}$$

where $t \geq 0$, and $0 \leq \lambda_h \leq 1$ is a non-decreasing in $h = 0, 1, 2, \ldots$, and τ_t was defined in Equation (4).

Equation (18) corresponds to the probability that a packet sampled h time steps ago has arrived at the estimator. Obviously, λ_h must be non-increasing since $\lambda_h = \mathbb{P}[\tau_t \leq h-1] + \mathbb{P}[\tau_t = h] = \lambda_{h-1} + \mathbb{P}[\tau_t = h]$.

Also, we define the packet loss probability as follows:

$$\lambda_{loss} \triangleq 1 - \sup\{\lambda_h | h \geq 0\} \tag{19}$$

The arrival process defined by Equation (18) can be also be defined with respect to the probability density of packet delay. In fact, by definition we have $\mathbb{P}[\tau_k = 0] = \lambda_0$, $\mathbb{P}[\tau_k = h] = \lambda_h - \lambda_{h-1}$ for $h \geq 1$, and $P[\tau_k = \infty] = \lambda_{loss}$.

Finally, we define the maximum delay of arrived packets as follows:

$$\tau_{max} \triangleq \begin{cases} \min\{H | \lambda_H = \lambda_{H+1}\} & \text{if } \exists H \text{ s.t. } \lambda_h = \lambda_H, \forall h \geq H \\ \infty & \text{otherwise} \end{cases}$$

(20)

We also assume that different sensors can have different packet delay distribution, i.e. $\lambda_h = \lambda_h^i$ for $i = 1, ..., M$. This is very natural assumption for wireless sensor networks. In fact, as indicated in Figure 1 sensor node i which is located far away from the sensor fusion node is subject to longer packet delay and packet loss than sensor node j. In fact the node i needs to multi-hop its packet through the network, while node j is only one-hop apart from the receiver.

In the rest of the paper we will use the following definition of stability for an estimator.

Definition: Let $\tilde{x}_{t|t} = f(\tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t)$ be an estimator, and $\tilde{e}_{t|t} = x_t - \tilde{x}_{t|t}$ and $\tilde{P}_{t|t} = \mathbb{E}[\tilde{e}_{t|t}\tilde{e}_{t|t}^T | \tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t]$ its error and error covariance, respectively. We say that the estimator is mean-square stable stable if and only if $\lim_{t\to\infty} \mathbb{E}[\tilde{e}_{t|t}] = 0$ and $\mathbb{E}[\tilde{P}_{t|t}] \leq M$ for some $M > 0$ and for all $t \geq 1$.

The previous definition can be rephrased in terms of the moments of the estimator error. In fact the conditions above are equivalent to $\lim_{t\to\infty} \mathbb{E}[||\tilde{e}_{t|t}||] = 0$ and $\mathbb{E}[||\tilde{e}_{t|t}||^2] \leq \text{trace}(M)$.

Let us consider the following static-gain estimator $\tilde{x}_{t|t} = \tilde{x}_{t|t}^t$ with finite-buffer of dimension N, where $\tilde{x}_{t|t}^t$ is computed as follows:

$$\tilde{x}_{t-k|t-k}^t = A\tilde{x}_{t-k-1|t-k-1}^t + \sum_{i=1}^{M} \gamma_{t,t-k}^i K_k^i (y_{t,t-k}^i - C_i A\tilde{x}_{t-k-1|t-k-1}^t) \quad (21)$$

$$\tilde{x}_{t-N|t-N}^t = \tilde{x}_{t-N|t-N}^{t-1} \quad (22)$$

$$\tilde{x}_{-k|-k}^t = \bar{x}_0, \quad \gamma_{-k}^t = 0, \quad \tilde{y}_{-k}^t = 0 \quad (23)$$

for $k = N-1, \ldots, 0$, where the last line include some dummy variables necessary to initialize the estimator for $t = 1, \ldots, N$. where the last line includes some dummy variables necessary to initialize the estimator for $t = 1, \ldots, N$. Note that constant-gain estimator structure is very similar to the optimal estimator structure given by Equation (12) as the estimate is corrected using only measurements y_{t-k}^i that have arrived at the sensor fusion location, i.e. using only $y_{t,t-k}^i$ for which $\gamma_{t,t-k}^i = 1$. However, differently from Equation (12) where the gain $K_{t,k}$ is time-varying and requires several matrix inversions for its computation, the gains K_k^i are constant and independent of t, and the computation of the estimate $\tilde{x}_{t|t}$ is obtained through a simple sequence of sum-product operations.

We also define the following variables that will be useful to analyze the performance of the estimator:

$$\tilde{x}_{k+1|k}^t = A\tilde{x}_{k|k}^t \quad (24)$$

$$\tilde{e}_{k+1|k}^t = x_{k+1} - \tilde{x}_{k+1|k}^t \quad (25)$$

$$\tilde{P}_{k+1|k}^t = \mathbb{E}[\tilde{e}_{k+1|k}^t \tilde{e}_{k+1|k}^{t^T} | \tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t] \quad (26)$$

$$\overline{P}_{k+1|k}^t = \mathbb{E}[\tilde{e}_{k+1|k}^t \tilde{e}_{k+1|k}^{t^T}] = \mathbb{E}[\tilde{P}_{k+1|k}^t] \quad (27)$$

where $t \geq k \geq 1$. From these definitions we get:

$$\tilde{e}_{k+1|k}^t = A(I - \sum_{i=1}^{M} \gamma_{t,k}^i K_{t-k}^i C_i)\tilde{e}_{k|k-1}^t + $$
$$+ w_k - \sum_{i=1}^{M} \gamma_{t,k}^i A K_{t-k}^i v_k^i \quad (28)$$

$$\tilde{P}_{k+1|k}^t = A(I - \sum_{i=1}^{M} \gamma_{t,k}^i K_{t-k}^i C_i)\tilde{P}_{k|k-1}^t (I - \sum_{j=1}^{M} \gamma_{t,k}^j K_{t-k}^j C_i)^T A^T + $$
$$+ Q + \sum_{i=1}^{M} \gamma_{t,k}^i A K_{t-k}^i R_i K_{t-k}^{i^T} A^T \quad (29)$$

$$\overline{P}_{k+1|k}^t = A(I - \sum_{i=1}^{M} \lambda_{t-k}^i K_{t-k}^i C_i)\overline{P}_{k|k-1}^t (I - \sum_{j=1}^{M} \lambda_{t-k}^j K_{t-k}^j C_i)^T A^T + $$
$$+ \sum_{i=1}^{M} \lambda_{t-k}^i (1 - \lambda_{t-k}^i) A K_{t-k}^i C_i \overline{P}_{k|k-1}^t K_{t-k}^i C_i^T A^T + $$
$$+ Q + \sum_{i=1}^{M} \lambda_{t-k}^i A K_{t-k}^i R_i K_{t-k}^{i^T} A^T \quad (30)$$

where $I \in \mathbb{R}^{n \times n}$ is the identity matrix. To obtain the previous equations we employed independence of γ_k^t, v_k, w_k, and $\tilde{e}_{k|k-1}^t$, the fact that v_k and w_k are zero mean, and finally that $\mathbb{E}[\gamma_{t,k}^i] = \mathbb{E}[(\gamma_{t,k}^i)^2] = \lambda_{t-k}^i$.

Note that the right had side of Equation (30) is linear in $\overline{P}_{k|k-1}^t$ for fixed $K_{t-k^i}^i$, and that is quadratic in $K_{t-k^i}^i$ for fixed $\overline{P}_{k|k-1}^t$. Therefore we can write Equation (30) more compactly by defining the following operator:

$$\mathcal{L}_\lambda(K, P) = A(I - KC_\lambda)P(I - KC_\lambda)^T A^T + Q + $$
$$+ AKP_\lambda K^T A^T + AKR_\lambda K^T A^T \quad (31)$$

where

$$C_\lambda = \begin{bmatrix} \lambda_1 C_1 \\ \vdots \\ \lambda_M C_M \end{bmatrix}, K = [K^1 \ldots K^M]$$

$$P_\lambda = \text{diag}(\lambda^1(1 - \lambda^1)C_1 P C_1^T, \ldots, \lambda^M(1 - \lambda^M)C_M P C_M^T),$$
$$R_\lambda = \text{diag}(\lambda^1 R_1, \ldots, \lambda_M R_M)$$

If we substitute $k = t - N$ into Equation (30), and noting that from Equation (22) follows that $\tilde{P}_{t-N+1|t-N}^t = \tilde{P}_{t-N+1|t-N}^{t-1}$ and $\overline{P}_{t-N+1|t-N}^t = \overline{P}_{t-N+1|t-N}^{t-1}$, we obtain:

$$\overline{P}_{t-N+2|t-N+1}^t = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, \overline{P}_{t-N+1|t-N}^{t-1}) \quad (32)$$

$$\overline{P}_{t-k+1|t-k}^t = \mathcal{L}_{\lambda_k}(K_k, \overline{P}_{t-k|t-k-1}^t), \quad k = N-2, \ldots, 0 \quad (33)$$

where $\lambda_k = (\lambda^1, \ldots, \lambda^M)$. Observe that Equation (32) and (33) define a set of linear deterministic equations for fixed λ_k and K_k. In particular, if we define $S_t = P_{t-N+1|t-N}^{t-1}$, then Equations (32) can be written as

$$S_{t+1} = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, S_t) \quad (34)$$

Since all matrices $\overline{P}_{t-k+1|t-k}^t, k = 0, \ldots, N-1$ can be obtained from S_t it follows that stability of estimator can be inferred from the properties of the operator $\mathcal{L}_\lambda(K, P)$. The following lemma provides these properties:

Lemma 1: Consider the operator $\mathcal{L}_\lambda(K, P)$ as defined in Equation (31). Assume also that $P \geq 0$, (A, C) is observable, $(A, Q^{1/2})$ is reachable, $R > 0$, and $0 \leq \lambda \leq 1$, where the

inequalities are to be interpreted for each component of the vector $\lambda = (\lambda^1, \ldots, \lambda^M) \in \mathbb{R}^M$. Also consider the following operator:

$$\Phi_\lambda(P) = APA^T + Q - APC_\lambda^T (C_\lambda PC_\lambda^T + P_\lambda + R_\lambda)^{-1} C_\lambda PA^T \tag{35}$$

and the gain $K_P = PC_\lambda^T (C_\lambda PC_\lambda^T + P_\lambda + R_\lambda)^{-1}$.

Then the following statements are true:

(a) $\mathcal{L}_\lambda(K, P) = \Phi_\lambda(P) + A(K - K_P)(C_\lambda PC_\lambda^T + P_\lambda + R_\lambda)(K - K_P)^T A^T$.

(b) $\mathcal{L}_\lambda(K, P) \geq \Phi_\lambda(P) = \mathcal{L}_\lambda(K_P, P), \quad \forall K$

(c) $(P_1 \geq P_2) \implies (\Phi_\lambda(P_1) \geq \Phi_\lambda(P_2))$.

(d) $(\lambda_1 \geq \lambda_2) \implies (\Phi_{\lambda_1}(P) \leq \Phi_{\lambda_2}(P)), \quad \forall P$.

(e) If there exists P^* such that $P^* = \mathcal{L}_\lambda(K, P^*)$, then $P^* > 0$ and it is unique. Consequently this is true also for $K = K_{P^*}$, where $P^* = \Phi_\lambda(P^*)$.

(f) If $(\lambda_1 \geq \lambda_2)$ and there exist P_1^*, P_2^* such that $P_1^* = \Phi_{\lambda_1}(P_1^*)$ and $P_2^* = \Phi_{\lambda_2}(P_2^*)$, then $P_1^* \leq P_2^*$.

(g) Let $S_{t+1} = \mathcal{L}_\lambda(K, S_t)$ and $S_0 \geq 0$. If $S^* = \mathcal{L}_\lambda(K, S^*)$ has a solution, then $\lim_{t \to \infty} S_t = S^*$, otherwise the sequence S_t is unbounded.

(h) If there exists S^*, K such that $S^* = \mathcal{L}_\lambda(K, S^*)$, then also $P^* = \Phi_\lambda(P^*)$ exists and $P^* \leq S^*$.

Proof: The proofs for all the claims of this lemma can ba obtained along the lines of the proofs in [21], therefore in the interest of space they are omitted. ∎

The previous theorem provides all tools necessary to analyze and design the optimal estimator with constant gains. In particular, fact (g) indicates that the constant gain K^* that minimizes the steady state error covariance P^* can be derived from the unique fixed point of the nonlinear operator $P^* = \Phi_\lambda(P^*)$, where $K^* = K_{P^*}$. If the optimal gain K^* is used, then the expected error covariance converges to P^* regardless of the initial conditions $(P_0, \bar{x}_0)$, as follows from fact (f). Fact (i) shows that if the system A is unstable the arrival probability λ needs to be sufficiently large to ensure stability, and that the critical value λ_c is a function of the unstable eigenvalues of A. Finally, although λ_c and the the fixed point $P^* = \Phi_\lambda(P^*)$ cannot be computed analytically, from fact (j) follows that they can be computed efficiently using numerical optimization tools. Finally, fact (k) will be used to show that if the error covariance is bounded then the estimator is also unbiased.

Lemma 2: Let us consider the nonlinear operator $\Psi_\lambda(P) = APA^T - APC_\lambda^T (C_\lambda PC_\lambda^T + P_\lambda)^{-1} C_\lambda PA^T$. Then the operator Φ_λ has a unique fixed point solution $\Phi_\lambda(P) = P$ if and only if

$$\lim_{k \to \infty} \text{trace}(\Psi_\lambda^k(I))^{\frac{1}{k}} < 1 \tag{36}$$

Proof: The proof follows along the lines of Theorem 5 in [24], therefore in the interest of space is omitted. ∎

Note that $\Psi_\lambda(P) = \Phi_\lambda(P)|_{Q=0, R_\lambda=0}$, i.e. it is the error covariance update in the absence of process and measurement noise, therefore stability is independent of the error noise covariances. The previous lemma provides a numerical algorithm to verify whether the operator Φ_λ has a fixed point. In fact, it is sufficient to define the sequence $V_{k+1} = \Psi(V_k), V_0 = I$ and check if it converges to zero exponentially. Obviously, the stability is a complex function of the probability vector λ and the observation matrices C_i, and it is difficult to derive analytical expression even for $M = 1$ [15].

The previous results are sufficient to derive the optimal sensor fusion estimator with constant gains.

Theorem 2: Let us consider the stochastic linear system given in Equations (1)-(2), where (A, C) is observable, $(A, Q^{1/2})$ is reachable, and $R_i > 0, \forall i = 1, .., M$. Also consider the arrival process defined by Equations (18)-(20), and the set of estimators with constant gains $\{K_k\}_{k=0}^N$ where $K_k = [K_k^1 \ldots K_k^M] \in \mathbb{R}^{n \times m}$ are defined in Equations (21)-(23). If condition given by Eqn. (36) is not satisfied for $\lambda = \lambda_N = (\lambda_N^1, \ldots, \lambda_N^M)$, then there exist no stable estimator with constant gains. Otherwise, let consider the optimal gains $\{K_k^*\}_{k=0}^N$ defined as follows:

$$K_k^* = V_k^N C_{\lambda_k}^T (C_{\lambda_k} V_k^N C_{\lambda_k}^T + (V_k^N)_{\lambda_k} + R_{\lambda_k})^{-1} \tag{37}$$

$$V_{N-1}^N = \Phi_{\lambda_{N-1}}(V_{N-1}^N) \tag{38}$$

$$V_k^N = \Phi_{\lambda_k}(V_{k+1}^N), \quad k = N-1, \ldots, 0 \tag{39}$$

Also consider $\overline{P}_{k+1|k}^t$ as defined in Equation (27), then $\lim_{t \to \infty} \overline{P}_{t-k+1|t-k}^t = V_k^N$, independently of initial conditions $(P_0, \bar{x}_0)$. For any other choice of gains $\{K_k\}_{k=0}^N$ for which the solution $\{T_k\}_{k=0}^N$ to the following equations exist:

$$T_N^N = \mathcal{L}_{\lambda_N}(K_N, T_N^N) \tag{40}$$

$$T_k^N = \mathcal{L}_{\lambda_k}(K_k, T_{k+1}^N), \quad k = N-1, \ldots, 0 \tag{41}$$

then $\lim_{t \to \infty} \overline{P}_{t-k+1|t-k}^t = T_k^N$, and $V_k^N \leq T_k^N$ for $k = 0, \ldots, N$. Also $V_0^{N+1} \leq V_0^N$. Finally, if $\tau_{max} = \sup_{i=1}^M (\tau_{max}^i) < \infty$, then $V_0^N = V_0^{\tau_{max}}$ for all $N \geq \tau_{max}$.

Proof: First we prove by contradiction that there is no stable estimator with constant gains if condition Eqn. (36) is not satisfied. Suppose such an estimator exists, i.e. there exist N and $\{K_k\}_{k=0}^{N-1}$ such that $\overline{P}_{t|t}^t$ is bounded for all t. Since $\overline{P}_{t+1|t}^t = A\overline{P}_{t|t}^t A^T + Q$ also $\overline{P}_{t+1|t}^t$ must be bounded for all t. From Equations (32) and (33) it follows that $\overline{P}_{t+1|t}^t$ is bounded if and only if $\overline{P}_{t-k+1|t-k}^t$ for $k = 0, \ldots, N-1$ are bounded for all t. Therefore, since the bounded sequence $S_t = \overline{P}_{t-N+1|t-N}^t$ needs to satisfy Equation (34), from Lemma 1(g) follows that $S^* = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, S^*)$ has a solution. From Lemma 1(h) follows that also $P^* = \Phi_{\lambda_{N-1}}(P^*)$ has a solution. However, according to Lemma 2, $P^* = \Phi_{\lambda_{N-1}}(P^*)$ cannot have a solution, which contradicts the hypothesis that a stable estimator exists.

Consider now the case when N is such that λ_N satisfies condition Eqn. (36) . From Theorem 1(h) it follows that Equations (37)-(39) are well defined and have a solution. From Lemma 1(g) it follows that $\lim_{t \to \infty} \overline{P}_{t-k+1|t-k}^t = V_k^N$ for the optimal gains $\{K_k^N\}_{k=0}^{N-1}$, and $\lim_{t \to \infty} \overline{P}_{t-k+1|t-k}^t = T_k^N$ when using generic gains $\{K_k\}_{k=0}^{N-1}$. From Lemma 1(h) it follows that $V_{N-1}^N \leq T_{N-1}^N$. From Lemma 1(c) we have $V_{N-2}^N = \Phi_{\lambda_{N-2}}(V_{N-1}^N) \leq \mathcal{L}_{\lambda_{N-2}}(K_{N-2}, V_{N-1}^N) \leq \mathcal{L}_{\lambda_{N-2}}(K_{N-2}, T_{N-1}^N) = T_{N-2}^N$. Inductively, it is easy to show that $V_k^N \leq T_k^N$ for all $k = 0, \ldots, N-1$.

Now we want to show that $V_0^{N+1} \leq V_0^N$. From Lemma 1(f) and the property $\lambda_{N+1} \geq \lambda_N$ follow also that $V_{N+1}^{N+1} = \Phi_{\lambda_{N+1}}(V_{N+1}^{N+1}) \leq V_N^N = \Phi_{\lambda_N}(V_N^N)$. Therefore $V_N^{N+1} = \Phi_{\lambda_N}(V_{N+1}^{N+1}) \leq \Phi_{\lambda_N}(V_N^N) = V_N^N$ and inductively $V_k^{N+1} \leq V_k^N$ for all $k = N, \ldots, 0$ which proves the statement.

Finally, if τ_{max} is finite, then $\lambda_k = \lambda_{\tau_{max}}$ for all $k \geq \tau_{max}$. Assume $N > \tau_{max}$, then $V_N^N = \Phi_{\lambda_N}(V_N^N) =$

$\Phi_{\lambda_{N-1}}(V_N^N) = V_{N-1}^N = \Phi_{\lambda_{N-1}}(V_{N-1}^N) = \Phi_{\lambda_{N-2}}(V_{N-1}^N) = V_{N-2}^N = \ldots = V_{\tau_{max}}^N = \Phi_{\lambda_{\tau_{max}}}(V_{\tau_{max}}^N)$. Since $V_{\tau_{max}}^{\tau_{max}} = \Phi_{\lambda_{\tau_{max}}}(V_{\tau_{max}}^{\tau_{max}})$, then by Lemma 1(e) we have that $V_{\tau_{max}}^{\tau_{max}} = V_{\tau_{max}}^N$. According to Equation (39) we also have $V_k^{\tau_{max}} = V_k^N$ for $k = \tau_{max}, \ldots, 0$, which concludes the theorem. ∎

The previous theorems shows that the optimal gains can be obtained by finding the fixed point of a modified algebraic Ricatti Equation (38) and then iterating N time an operator with the same structure but with different probability vector λ_k. The theorem also demonstrates that a stable sensor fusion estimator with constant gains exists if and only if the optimal estimator with constant gains exists, therefore the optimal estimator design implicitly solves the problem of existence of stable estimators.

Note also that the estimator stability does not depend on the complete packet arrival statistics, but only in the overall packet loss of each sensor. In fact, let $\lambda_c = (1 - \lambda_{loss}^1, \ldots, 1 - \lambda_{loss}^M)$ the vector of successful packet transmission probability for all sensor node, then a stable sensor fusion estimator exists if and only if there exist $\epsilon > 0$ such that condition given by Eqn. (36) is satisfied for $\lambda = \lambda_c + \epsilon$.

Moreover, the performance of the estimator, i.e. its steady state error covariance $\lim_{t\to\infty} P_{t+1|t} = \lim_{t\to\infty} \mathbb{E}[e_{t+1|t} e_{t+1|t}^T] = V_0^N$, improves as the buffer length N is increased, which is to be expected since more information is stored. However, if the maximum packet delay is finite $\tau_{max} < \infty$, then the performance of the estimator does not improve for $N > \tau_{max}$. This is consistent with Theorem 1(b) since if a measurement packet has not arrived within τ_{max} time steps after it was sampled, then it will never arrive and it is useless to wait longer.

From a practical perspective, the previous tools can be used by the designer to evaluate the tradeoff between the estimator performance V_0^N and buffer length N which is directly related to computational requirements.

VI. Special Cases

In this section we consider some special cases that gives rise to simpler equation.

A. Colocated sensors (M=1)

If all sensors are colocated, then all measurements are instantaneously collected at the transmitter node and are then sent over the network. This correspond to the scenario for which $M = 1$, and $C = C_1$. In this context the optimal gain and the nonlinear operator in Lemma 2 are given by

$$\begin{aligned} \Phi_\lambda(P) &= APA^T + Q - \lambda APC^T(CPC^T + R)^{-1}CPA^T \quad (42) \\ K_P &= PC^T(CPC^T + R)^{-1} \end{aligned}$$

where λ is a scalar. This scenario was previously studied in [21].

B. Identical sensors ($C_i = C, R_i = R$)

If all sensors are identical, i.e. they have the same observation matrix $C_i = C$ and measurement noise covariance $R_i = R$, but different packet arrival probabilities, i.e. $\lambda_k^i \neq \lambda_k^j$, then also the sensor gain are different, i.e. $K_k^i \neq K_k^j$. At first look this seems unexpected since if there are two sensor measurements from two identical sensor taken at the same step time k corresponding to the same column k in the sensor fusion buffer, then one might expect to give them the same weight. However, this is reasonable at the light of the previous analysis where packet loss distribution affect the design of the optimal gains K_h^i.

C. Identical packet arrival distribution ($\lambda_h^i = \lambda_h$)

If all sensors have identical packet arrival probabilities, i.e. $\lambda_k^i = \lambda_k, \forall i$, then we can restrict the stability analysis to a single parameter λ. Let the observation matrix $C = [C_1^T \ldots C_M^T]^T$ and $R = \text{diag}(R_1, \ldots, R_M)$. In this context the optimal gain and the nonlinear operator in Lemma 2 are given by

$$\begin{aligned} \Phi_\lambda(P) &= APA^T + Q - \\ &\quad - \lambda APC^T(\lambda CPC^T - (1-\lambda)P_C + R)^{-1}CPA^T \\ &\leq APA^T + Q - \lambda APC^T(CPC^T + R)^{-1}CPA^T \\ K_P &= PC^T(\lambda CPC^T - (1-\lambda)P_C + R)^{-1} \\ P_C &= \text{diag}(C_1PC_1^T, \ldots, C_MPC_M^T) \end{aligned}$$

where we employed the fact that $\lambda CPC^T - (1-\lambda)P_C \leq CPC^T$. The matrix $\Phi_\lambda(P)$ considered here is smaller than the corresponding operator in Eqn. (42) for all P, from which it follows that it better to independently send measurements along different paths than sending a single packet which includes all measurements. Indeed, this is a common strategy for most routing protocols adopted in wireless sensor networks.

VII. Conclusions

In this work we proposed a framework to optimally design and analyze the performance of sensor fusion filters based on measurements collected from a set of distributed in a single location. We showed that the optimal filter is a time-varying Kalman filter with a buffer. We also proposed a suboptimal sensor fusion filter which is computationally more efficient since it requires only a finite memory buffer and constant gains. For this class of filters it was shown that if packet arrival of each sensor is i.i.d. but possibly different for each sensor, then the estimators are mean square stable if the packet loss is sufficiently small. Therefore, implicitly we also provided necessary and sufficient conditions about existence of stable sensor fusion filter. Finally, we presented numerical algorithms for the computation of the expected estimator error covariance of all the proposed estimators.

The tools developed in this paper are useful both from a control system design perspective and from a communication design perspective. In fact, from a control perspective they can help to evaluate the tradeoffs between performance (error covariance), memory requirements (buffer length), and the hardware resources ("smart" sensor and fast matrix inversion). In particular, the knowledge of the packet arrival statistics can be used to find the optimal constant gains K_k^i and thus improving performance. From a communication perspective, these tools can be used to aid communication protocol design for real-time applications. In fact, as mentioned in Section I, when designing communication protocols, in particular for wireless systems, there is tradeoff between packet loss and packet delay. At the moment, the choice between favoring reduction of overall packet delay or reduction of packet loss is based on heuristics and experience, and it is not tailored to the specific real-time applications. Therefore, being able to quantitatively measure performance of different protocols can improve cross-layer design of complex networked control systems.

As future work, we are currently implementing the most popular wireless sensor network routing protocols to obtain experimental packet arrival distributions in order to evaluate these routing protocols for on-line monitoring applications using the tools developed in this paper.

REFERENCES

[1] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the physical world with pervasive networks," *IEEE Pervasive Computing*, vol. 1, no. 1, pp. 59–69, January 2002.

[2] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. M. Mainwaring, and D. Estrin, "Habitat monitoring with sensor networks," *Communication of the ACM*, vol. 47, no. 6, pp. 34–40, 2004.

[3] G. Tolle, J. Polastre, R. Szewczyk, N. Turner, K. Tu, P. Buonadonna, S. Burgess, D. Gay, W. Hong, T. Dawson, and D. Culler, "A macroscope in the redwoods," in *Proceedings of the Third ACM Conference on Embedded Networked Sensor Systems (SenSys05)*, San Diego, CA, USA, November 2005, pp. 51–63.

[4] B. Hohlt, L. Doherty, and E. Brewer, "Flexible power scheduling for sensor networks," in *IEEE and ACM International Symposium on Information Processing in Sensor Networks (IPSN'04)*, April 2004.

[5] T. Stankovic, J. C. Lu, and T. Abdelzaher, "SPEED: a stateless protocol for real-time communication in sensor networksapproximate distributed kalman filtering in sensor networks with quantifiable performance," in *Proceedings of the IEEE Conference on Distributed Computing Systems*, 2003, pp. 46–55.

[6] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "Networked control systems: analysis and design," *Submitted for publication*, 2006.

[7] J. Nilsson, B. Bernhardsson, and B. Wittenmark, "Stochastic analysis and control of real-time systems with random time delays," *Automatica*, vol. 34, no. 1, pp. 57–64, January 1998.

[8] H. L. Alexander, "State estimation for distributed systems with sensing delay," in *Proc. SPIE Vol. 1470, Data Structures and Target Classification*, 1991, pp. 103–111.

[9] T. Larsen, N. Andersen, O. Ravn, and N. Poulsen, "Incorporation of time delayed measurements in a discrete-time Kalman filter," in *Proceedings of IEEE Conference on Decision and Control (CDC'98)*, vol. 4, 1998, pp. 3972–3977.

[10] S. Julier and J. Uhlmann, "Fusion of time delayed measurements with uncertain time delays," in *Proceedings of IEEE American Control Conference (ACC'05), YEAR = 2005, volume = 6, pages = 4028-4033,*.

[11] C. N. Hadjicostis and R. Touri, "Feedback control utilizing packet dropping network links," in *Proceedings of the IEEE Conference on Decision and Control*, vol. 2, Las Vegas, NV, Dec. 2002, pp. 1205–1210.

[12] Q. Ling and M. Lemmon, "Optimal dropout compensation in networked control systems," in *IEEE conference on decision and control*, Maui, HI, December 2003.

[13] D. Giorgiev and D. Tilbury, "Packet-based control," in *Proc. of the 2004 Amer. Contr. Conf.*, June 2004, pp. 329–336.

[14] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, September 2004.

[15] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *To appear in Proceedings of IEEE*.

[16] O. C. Imer, S. Yuksel, and T. Basar, "Optimal control of dynamical systems over unreliable communication links," *To appear in Automatica*, July 2006.

[17] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Transaction on Information Theory*, vol. 15, no. 4, pp. 457–462, 1969.

[18] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proceedings of IEEE Conference on Decision and Control*, vol. 4, Bahamas, December 2004, pp. 4180–4186.

[19] N. Yu, L. Wang, T. Chu, and F. Hao, "An LMI approach to networked control systems with data packet dropout and transmission delay," *Journal of Hybrid Systems*, vol. 3, no. 11, November 2004.

[20] P. Naghshtrabizi and J. Hespanha, "Designing observer-based controller for network control systems," in *IEEE Conference on Decision and Control (CDC-ECC05)*, Sivilla, Spain, December 2005.

[21] L. Schenato, "Kalman filtering for networked control systems with random delay and packet loss," in *Conference of Mathematical Theroy of Networks and Systems (MTNS'06)*, Kyoto, Japan, July 2006.

[22] G. Chen, G. Chen, and S. Hsu, *Linear Stochastic Control Systems*. CRC Press, 1995.

[23] P. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*, ser. Information and System Science Series, T. Kailath, Ed. Englewood Cliffs, NJ 07632: Prentice Hall, 1986.

[24] W. De Koning, "Compensability and optimal compensation of systems with white parameters," *IEEE Transactions on Automatic Control*, vol. 37, no. 5, pp. 579–588, 1992.

[25] L. Schenato, "Kalman filtering for networked control systems with random delay and packet loss," Univerity of Padova," submitted for publication, 2006. [Online]. Available: http://www.dei.unipd.it/ṡchenato/PAPERS/KalmRand.pdf

[26] S. Oh, L. Schenato, P. Chen, and S. Sastry, "Tracking and coordination of multiple agents using sensor networks: system design, algorithms and experiments," *To appear in Proceedings of IEEE*.