Optimal estimation in networked control systems subject to random delay and packet drop

Luca Schenato

Department of Information Engineering, University of Padova, Italy

schenato@dei.unipd.it

Abstract

In this paper we study optimal estimation design for sampled linear systems where the sensors measurements are transmitted to the estimator site via a generic digital communication network. Sensor measurements are subject to random delay or might even be completely lost. We show that the minimum error covariance estimator is time-varying, stochastic, and it does not converge to a steady state. Moreover, the architecture of this estimator is independent of the communication protocol and can be implemented using a finite memory buffer if the delivered packets have a finite maximum delay. We also present two alternative estimator architectures which are more computationally efficient and provide upper and lower bounds for the performance of the time-varying estimator. The stability of these estimators does not depend on packet delay but only on the packet loss probability. Finally, algorithms to compute critical packet loss probability and estimators performance in terms of their error covariance are given and applied to some numerical examples.

Index Terms

Networked Control Systems, random time delay, packet drop, minimum variance estimator, Kalman filter, stability

I. Introduction

Recent technological advances in MEMS, DSP capabilities, computing, and communication technology are revolutionizing our ability to build large scale distributed networked control systems (NCS) as recently surveyed in [1]. In particular, among NSCs, the class represented by wireless sensor networks (WSNs), which are large networks of spatially distributed electronic devices – called nodes – capable of sensing, computation and wireless communication, can offer access to an unprecedented quality and quantity of information. This information can revolutionize our ability in controlling of the environment, such as fine-grain building environmental control [2], vehicular networks and traffic control [3], surveillance [4], habitat monitoring [5] [6], and manufacturing automation [7].

However, NCS also pose challenging problems in different research areas such as information theory, signal processing, communication theory, and control theory, mainly caused the fact that sensors, actuators and controllers are not physically co-located and need to exchange information via a digital communication network. These problems are particularly harsh in WSNs [8], where measurements are sampled at each node and then routed through a multi-hop network to a remote location for data analysis and decision-making process. As a consequence, measurements arrive at the decision-making location with a non-deterministic delay or can be totally lost along the way, thus possibly undermining the effectiveness of the decision-making part. As a result, it is important to simultaneously 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 [9], distributed signal processing [10], in-network data compression [11], and packet routing protocols [12], but not completely avoided due to inherent unreliable nature of wireless communication. For example, communication scheduling protocols based on time division medium access (TDMA) can reduce packet loss and power consumption at the price of longer delay [13], while event-triggered protocols which adopt broadcasting of messages can reduce packet delay at the price of larger packet loss [14]. This tradeoff is qualitatively visualized in Fig. 1, where each point on the solid curve represents the performance achieved by a specific communication protocol.

From a data-collection application perspective the best protocols are those sitting on the bottom of the curve since delay is unimportant. Differently, from a real-time application perspective, both parameters are important, therefore it is not easy to determine which is the optimal operating point on the curve.

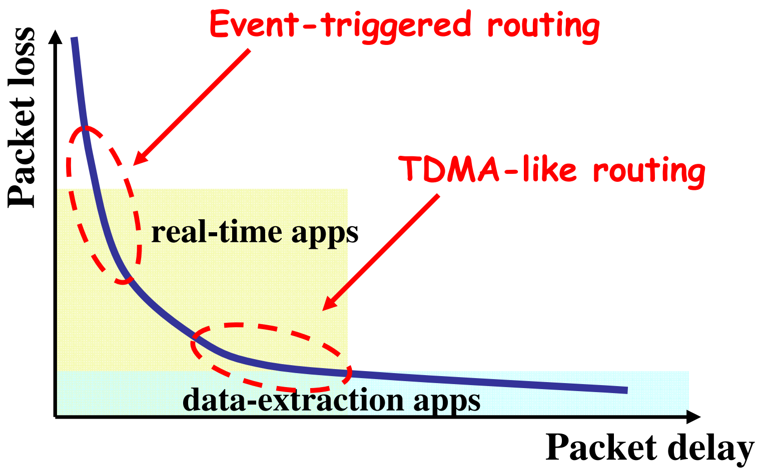

Fig. 1. Pictorial representation of a typical tradeoff curve (solid line) that is achieved by different communication protocols such as event-triggered and TDMA-based routing protocols. The shaded areas indicated constraint regions for real-time and data-collection applications.

Currently, communications protocols and networked control systems are designed separately. In particular, protocols are designed 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 estimator design for systems subject to random packet delay and packet loss. Estimator design is in fact a fundamental ingredient for optimal control design and the estimator error covariance directly affects the closed loop performance of the control system [15]. Moreover, the ability to exactly quantify estimator error covariance based on packet arrival statistics, can also be employed to select the more appropriate communication protocol strategy.

The paper is organized as follows. The next section presents an overview of relevant previous work and the contribution of this paper. Section III formalizes the minimum variance estimation problem and the packet arrival process modeling. In Section IV the minimum variance estimation problem is solved and optimality conditions on memory requirements are given. In Section V we derive stability conditions and quantify performance in terms of expected estimation error covariance for the minimum variance estimator with constant gains under i.i.d. packet arrival with known statistics. This suboptimal estimator provides an upper bound for the estimation error of the time-varying optimal filter proposed in Section IV. Section VI shows how estimation performance can be improved if sensors have computational resources and are physically co-located. This estimator architecture also provides a lower bound for error covariance of the time-varying optimal filter of Section IV. Section VII gives some numerical examples to illustrate the use of the tools derived in the previous sections. Finally, Section VIII summarizes the conclusions and suggest directions for future work.

II. Previous Work and Contribution

Classical control has mainly focused on systems with constant delay [16] or with small delay perturbation known as jitter [17]. Recently several groups have looked at networked control systems with large random delay or packet loss. Two recent survey papers presents several results in the context of packet-switched networks [18], and in the context of bit-rate limited feedback control [19], respectively. Without entering a detailed discussion, the bit-rate limited approach provides tighter analytical performance bounds as compared to the packet-switched framework, but the practical implementation and coding schemes proposed in former framework are not available in today's communication protocols.

In this work, we focus on packet-switched network. Within this class of systems, results can be divided into two main groups: the first group focusing on variable delay but no packet drop, while the second group focusing on packet loss but no delay.

Within the first group, some authors derived stability conditions in terms of LMIs for closed loop continuous time linear systems with stochastic sampling time [20][21], and Nesic at al. [22] obtained Lyapunov-like stability conditions for continuous time nonlinear systems with unknown but bounded sampling time. These works simply determine stability for a given closed loop system, and there is no controller synthesis specifically designed to take into account delay. With this respect, Yue et al. [23] proposed an LMI approach for the design of stabilizing controllers for bounded delay, while Nilsson at al. [24] 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. Another relevant line of research is data fusion of measurements obtained from sensors with different delays. For example, Alexander [25] and Larsen et al. [26] 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. [27] 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 is an upper bound on the possible 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 [28][29][30] the authors proposed compensation techniques for i.i.d Bernoulli packet-drop communication networks and derived stability conditions for closed loop discrete time system. Elia et al. [31][32] proposed a stochastic perturbation approach for general MIMO LTI discrete time systems and showed that the optimal controller design is equivalent to solving a convex LMI optimization problem. Sinopoli at al. [33], extending results that can be traced back to Nahi [34], looked specifically at minimum variance estimation design for packet-drop networks and showed that the optimal estimator is necessarily time-varying, including also stability conditions. These results have been independently extended to LQG optimal control in [35] and [36]. Finally, a number of researches has explored specific mechanisms to improve estimation performance by exploiting local computation at the sensor location [37][38], controlled communication [39][38], and network topology [40].

The previous two groups of results suffer from some limitations. In fact, even with retransmission mechanisms present in all current digital communication networks, and in particular in the wireless ones, it is impossible to guarantee that all packets are correctly delivered to the destination. On the hand, in wireless sensor networks which implement multi-hop communication, delay is not negligible and is subject to large variations. Therefore, none of the modelings considered so far, i.e. random delay but no packet loss and packet loss but no delay, fully represent control systems interconnected by digital communication networks. 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 [41] [42].

In this paper we propose a probabilistic framework to analyze estimation where observation packets 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. We derive the optimal estimator in mean square sense and we show that the minimum error covariance estimator is equivalent to a time-varying Kalman filter with a buffer, for which the optimal gain does not converge to a steady state. Moreover, this estimator structure is independent of the packet arrival statistics and can be implemented using a 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 received packets. We also present two alternative estimator architectures which are computationally more efficient. In the first architecture, the estimator gains are constrained to be constant rather than stochastic as in the time-varying optimal estimator. In the second architecture, smart sensors are used to preprocess data and improve estimation performance. Also necessary and sufficient conditions for stability for these estimators are shown to depend only on the over packet loss probability and not on delay. We also provide numerical algorithms to compute the expected error covariances 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:

$$x_{t+1} = Ax_t + w_t \quad (1)$$
$$y_t = Cx_t + v_t, \quad (2)$$

where $t \in \mathbb{N} = \{0, 1, 2, \dots\}$, $x, w \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $A \in \mathbb{R}^{n\times n}$, $y \in \mathbb{R}^m$, $C \in \mathbb{R}^{m\times n}$, (x_0, w_t, v_t) are Gaussian, uncorrelated, white, with mean $(\bar{x}_0, 0, 0)$ and covariance (P_0, Q, R) respectively. We also assume that the pair (A, C) is observable, $(A, Q^{1/2})$ is reachable, and $R > 0$[1].

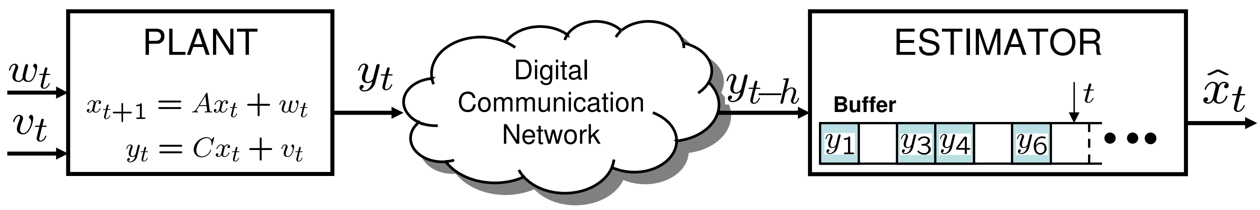

Fig. 2. Networked systems modeling. Sampled observations at the plant site are transmitted to the estimator site via a digital communication network. Due to retransmission and packet loss, observation packets arrive at the estimator site with possibly random delay.

Measurements are time-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 (see Fig. 2). Time-stamping of measurements is necessary to reorder packets at the receiver side as they can arrive out of order. Maintaining global clock synchronization in distributed systems is a challenging task, but recent work has shown remarkably small synchronization errors even in large sensor networks [43]. 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 plant and the estimator which delivers observation measurements to the estimator with possibly random delays. This model allows also for 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

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

buffer, as shown in Fig. 2. The arrival process is modeled via the random variable γ_k^t defined as follows:

$$\gamma_k^t = \begin{cases} 1 & \textit{if } y_k \text{ has arrived at the estimator before or at time } t, \ t \geq k \\ 0 & \textit{otherwise} \end{cases} \tag{3}$$

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 $\tau_k \in \{\mathbb{N}, \infty\}$ for observation y_k as follows:

$$\tau_k = \begin{cases} \infty & \textit{if } \gamma_k^t = 0, \forall t \geq k \\ t_k - k & \textit{otherwise}, \text{ where } t_k \triangleq \min\{t \mid \gamma_k^t = 1\} \end{cases} \tag{4}$$

where t_k is the arrival time of observation y_k at the estimator site . Since the packet delay can be random, observation measurements can arrive out of order at the estimator site (see Fig. 3, $t = 5$). Also it is possible that between two consecutive sampling periods no packet (see Fig. 3, $t = 4$) or multiple packets (see Fig. 3, $t = 6$) 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.

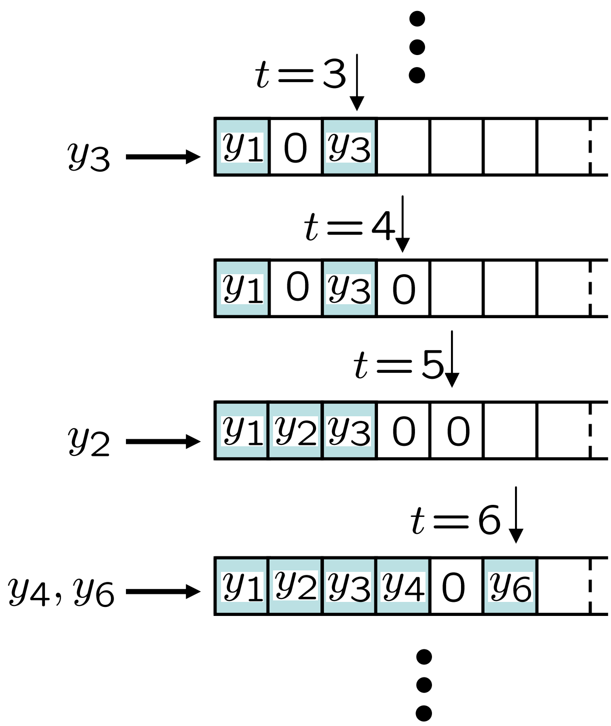

Fig. 3. Packet arrival sequence and buffering at the estimator location. Shaded squares correspond to observation packets that have been successfully received by the estimator. Cursor indicates current time.

If observation y_k is not yet arrived at the estimator at time t, we assume that a zero[3] is stored in the

[2]For example, ATM communication protocols adopts packets with 384-bit data field, Ethernet IEEE 802.3 packets allows for at least 368 bits for data payload, Bluetooth for 499 bits [18] and IEEE 802.15.4 for up to 1000 bits. This assumption might not hold for multimedia signal like audio and video signals, which however are not in the scope of this work.

[3]In practice, any arbitrary value can be stored in the buffer slots corresponding to the packets which have not arrived, since as it will be shown later, the optimal estimator does not use those values as they do not convey any information about the state x_t. Our choice of storing a zero simply reduces some mathematical burden.

k-slot of the buffer, as shown in Fig. 3. More formally, the value stored in the k-slot of the estimator buffer at time t can be written as follows:

$$\tilde{y}_k^t = \gamma_k^t y_k = \gamma_k^t C x_k + \gamma_k^t v_k \tag{5}$$

Our goal to compute the optimal mean square estimator $\hat{x}_{t|t}$ which is given by:

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

where $\tilde{\mathbf{y}}_t = (\tilde{y}_1^t, \tilde{y}_2^t, \ldots, \tilde{y}_t^t)$ and $\boldsymbol{\gamma}_t = (\gamma_1^t, \gamma_2^t, \ldots, \gamma_t^t)$. It is important to remark that the estimator above has the information whether a packet has been delivered or not, and it is not equivalent to computing $\hat{x}_{t|t} \neq \check{x}_{t|t} \triangleq \mathbb{E}[x_t \mid \tilde{\mathbf{y}}_t, \bar{x}_0, P_0]$. The latter estimator would in fact consider the zero entries of the buffer as true measurements and not as dummy variables, thus providing a lower performance. It is also useful to design the estimator error and error covariance as follows:

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

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

The estimate $\hat{x}_{t|t}$ is optimal in the sense that it minimizes the error covariance, i.e. given any estimator $\tilde{x}_{t|t} = f(\tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t)$, 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 is necessary to give rise to the separation principle for the LQG optimal control, which is one of the most widely used tool in control system design [15] [36].

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:

$$\hat{x}_{k|h}^t \triangleq \mathbb{E}[x_k \,|\, \gamma_h^t, \ldots, \gamma_1^t, \tilde{y}_h^t, \ldots, \tilde{y}_1^t, \bar{x}_0, P_0]$$

$$P_{k|h}^t \triangleq \mathbb{E}[(x_k - \hat{x}_{k|h}^t)(x_t - \hat{x}_{k|h}^t)^T \,|\, \gamma_h^t, \ldots, \gamma_1^t, \tilde{y}_h^t, \ldots, \tilde{y}_1^t, \bar{x}_0, P_0]$$

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:

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

$$\tilde{y}_{k+1}^t = C_{k+1}^t x_{k+1} + \tilde{v}_{k+1}^t, \quad k = 0, \ldots, t-1 \tag{10}$$

where the random variables $\tilde{v}_k^t \triangleq \gamma_k^t v_k$ are uncorrelated, zero mean white noise with covariance $R_k^t = \mathbb{E}[\tilde{v}_k^t (\tilde{v}_k^t)^T] = \gamma_k^t R$, and $C_k^t \triangleq \gamma_k^t C$. 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_k^t and measurement noise covariance R_k^t.

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 $R > 0$. Also consider the arrival process defined by Equation (3), and the mean square estimator defined in Equation (6). Then we have:

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

$$\begin{aligned}
\hat{x}^t_{0|0} &= \bar{x}_0, \quad P^t_{1|0} = P_0 & (11)\\
\hat{x}^t_{k|k} &= A\hat{x}^t_{k-1|k-1} + \gamma^t_k K^t_k(\tilde{y}^t_k - CA\hat{x}^t_{k-1|k-1}), \qquad k = 1, \ldots, t & (12)\\
K^t_k &= P^t_{k|k-1}C^T(CP^t_{k|k-1}C^T + R)^{-1} & (13)\\
P^t_{k+1|k} &= AP^t_{k|k-1}A^T + Q - \gamma^t_k AP^t_{k|k-1}C^T(CP^t_{k|k-1}C^T + R)^{-1}CP^t_{k|k-1}A^T & (14)
\end{aligned}$$

(b) The optimal estimator $\hat{x}_{t|t}$ can be computed iteratively using a buffer of finite length N if $\gamma^t_k = \gamma^{t-1}_k, \forall k \geq 1, \forall t \geq k + N$. 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 $t > N$:

$$\begin{aligned}
\hat{x}^t_{t-N|t-N} &= \hat{x}^{t-1}_{t-N|t-N}, & (15)\\
P^t_{t-N+1|t-N} &= P^{t-1}_{t-N+1|t-N} & (16)\\
\text{Equations (12),(13),(14)} & \qquad k = t - N + 1, \ldots, t & (17)
\end{aligned}$$

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:

$$\begin{aligned}
\hat{x}^t_{k|k} &= A\hat{x}^t_{k-1|k-1} + K^t_k(\tilde{y}^t_k - C^t_k A\hat{x}^t_{k-1|k-1})\\
K^t_k &= P^t_{k|k-1}C^{t^T}_k(C^t_k P^t_{k|k-1}C^{t^T}_k + R^t_k)^\dagger\\
P^t_{k+1|k} &= AP^t_{k|k-1}A^T + Q - AP^t_{k|k-1}C^{t^T}_k(C^t_k P^t_{k|k-1}C^{t^T}_k + R^t_k)^\dagger C^t_k P^t_{k|k-1}A^T\\
\hat{x}^t_{0|0} &= \bar{x}_0, \quad P^t_{1|0} = P_0
\end{aligned}$$

whose derivation can be found in any standard textbook on stochastic control [15] [44], and the symbol $\dagger$ denotes the matrix pseudoinverse[4]. By substituting $C^t_{k+1} = \gamma^t_{k+1}C$ and $R^t_k = \gamma^t_k R$ into the previous equations and after performing some straightforward simplifications we obtain the optimal estimator Equations (11)-(14).

(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 $P^t_{k+1|k} = P^{t-1}_{k+1|k}$ and $\hat{x}^t_{k|k} = \hat{x}^{t-1}_{k|k}$ hold under the same conditions on the indices. In particular it holds for $k = t - N$ which implies $P^t_{t-N+1|t-N} = P^{t-1}_{t-N+1|t-N}$ and $\hat{x}^t_{t-N|t-N} = \hat{x}^{t-1}_{t-N|t-N}$. Therefore, it not necessary to compute $P^t_{t+1|t}$ and $\hat{x}^t_{t|t}$ at any time step t starting from $k = 1$, but it is sufficient to use the values $\hat{x}^{t-1}_{t-N|t-N}$ and $P^{t-1}_{t-N+1|t-N}$ 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.

■

Note that the previous theorem does not necessarily implies that any optimal estimator needs to have an infinite buffer when packet delay cannot be bounded. Indeed, this poses the interesting problem of proving or disproving the existence of a filter that never requires an infinite buffer.

If there is no packet loss and no packet delay, i.e. $\gamma^t_k = 1, \forall(k, t)$, then Equations (11)-(14) reduce to the standard Kalman filter equations for a time-invariant system. In practice, the optimal estimator correspond to a time-varying Kalman filter with a buffer. Differently from the Kalman filter for a time-invariant system, neither the optimal gain not the error covariance converge to a steady state value. Therefore, it is not possible adopt a static filter the with the steady-state Kalman gain $K_t = K_\infty$, but in general it is necessary to store all past measurements and to invert of up to t matrices at any time step t to calculate the optimal estimate $\hat{x}_{t|t}$, which is particularly onerous for on-line implementation.

It is well known that the optimal estimator subject to measurement delay can be obtained by augmenting the state by the maximum measurement delay [25]. Since packet loss is equivalent to infinite delay, this might suggests that the optimal estimator should have an infinite buffer. However, the optimal estimator

[4]We employed the fact that the Kalman filter equations can be generalized to the case $R \geq 0$ by using pseudoinverse rather than the inverse in the computation of the posterior error covariance.

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. 4). 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.

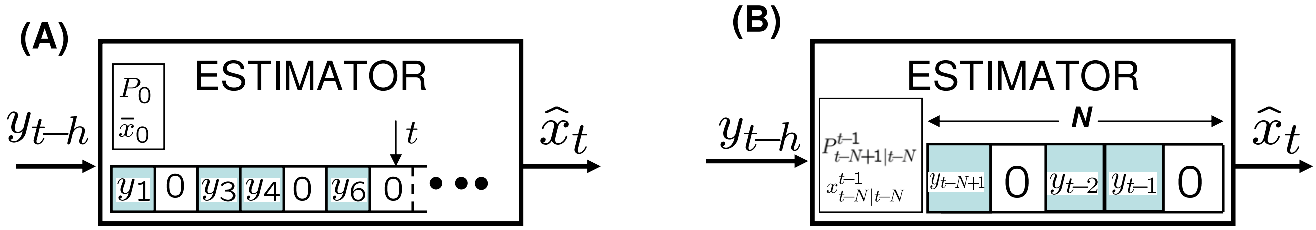

Fig. 4. Optimal estimator for general packet arrival processes (*left*). Optimal estimator with finite memory buffer for packet arrival processes with bounded delay *(right)*.

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. However, other metrics can be considered, such as the probability that the error covariance exceeds a certain threshold, i.e. $\mathbb{P}[P_{t+1|t} > P_{max}]$ [45]. In this work we focus on the expected error covariance $\mathbb{E}_\gamma[P_{t+1|t}]$. 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 be obtained [33]. Rather than trying to bound performance of the time-varying optimal estimator, in the next section we will focus on filters with constant gains and with a finite buffer dimension, i.e. we will consider $K_{t-h}^t = K_h$ for all $t \in \mathbb{N}, h = 0, \ldots, N{-}1$. The gains K_h will then be optimized to achieve the smallest error covariance at steady-state. The advantage of using constant gains is that, differently from the optimal time-varying filter, it is not necessary to invert any matrix at all 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 estimation with constant gains

In this section we will study minimum error covariance 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:

$$\lambda_{loss} \stackrel{\Delta}{=} 1 - \sup\{\lambda_h | \; h \geq 0\} \tag{19}$$

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:

$$\tau_{max} \triangleq \begin{cases} \min\{H \,|\, \lambda_H = \lambda_{H+1}\} & \textit{if } \exists H \textit{ such that } \lambda_h = \lambda_H, \forall h \geq H \\ \infty & \textit{otherwise} \end{cases} \tag{20}$$

Fig. 5 shows some typical scenarios that can be modeled under the previous hypotheses. Scenario (A) corresponds to a deterministic process where all packets are successfully delivered to the estimator with a constant delay. This scenario is typical of wired systems. Scenario (B) models a DCN that guarantees delivery of all packets within a finite time window τ_{max}, but the delay is not deterministic. This is a common scenario in drive-by-wire systems. Scenario (C) represents a DCN which drops packets that are older than τ_{max} and consequently a fraction $\lambda_{loss} > 0$ of observations is lost. This scenario is often encountered in wireless sensor networks. Scenario (D) corresponds to a DCN with no packet loss but with unbounded random packet delay. One example of such a scenario is a DCN that continues to retransmit a packet till it is not delivered.

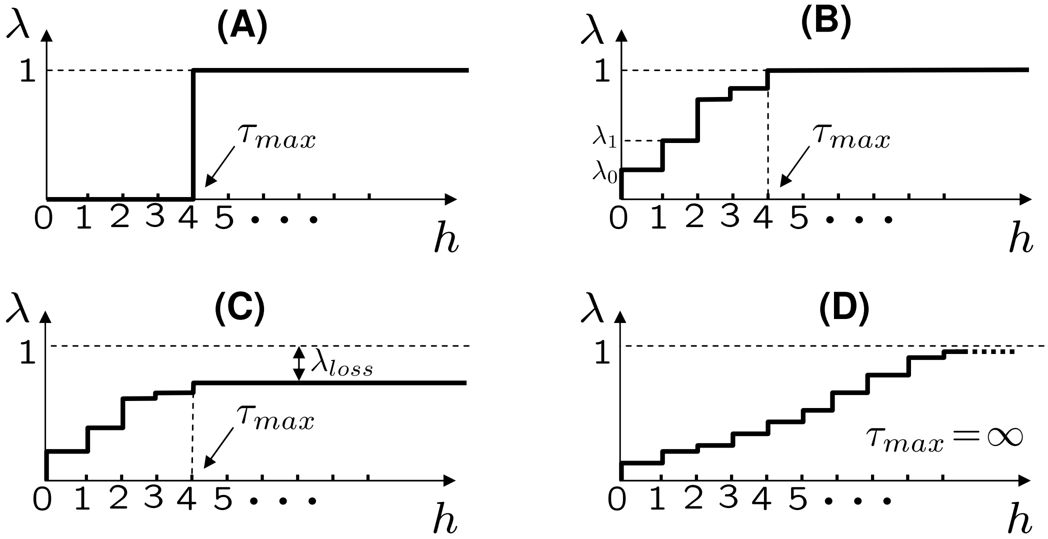

Fig. 5. Probability function of arrival process $\lambda_h = \mathbb{P}[\tau_k \leq h]$ for different scenarios: deterministic packet arrival with fixed delay (A); bounded random packet delay with no packet loss (B); bounded random packet delay with packet loss (C); unbounded random packet delay with no packet loss (D).

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 a generic estimator, where f is a measurable function, 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 and unbiased 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 $\mathbb{E}[\,||\tilde{e}_{t|t}||^2] \leq \mathrm{trace}(M)$.

Let us consider the following constant-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:

$$\tilde{x}^t_{t-k|t-k} = A\tilde{x}^t_{t-k-1|t-k-1} + \gamma^t_{t-k} K_k(\tilde{y}^t_{t-k} - CA\tilde{x}^t_{t-k-1|t-k-1}), \quad k = N-1, \ldots, 0 \tag{21}$$

$$\tilde{x}^t_{t-N|t-N} = \tilde{x}^{t-1}_{t-N|t-N} \tag{22}$$

$$\tilde{x}^t_{-k|-k} = \bar{x}_0, \quad \gamma^t_{-k} = 0, \quad \tilde{y}^t_{-k} = 0 \tag{23}$$

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 only if the observation has arrived, i.e. $\gamma_{t-k}^t = 1$, otherwise only the open loop prediction is considered. However, differently from Equation (12), the gains $K_k, k = 0, \ldots, N-1$ are constant and independent of t, and the computation of the estimate $\tilde{x}_{t|t}$ does not require any on-line matrix inversion differently from the estimator in the previous section.

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

$$\tilde{x}_{k+1|k}^t = A\tilde{x}_{k|k}^t \tag{24}$$
$$\tilde{e}_{k+1|k}^t = x_{k+1} - \tilde{x}_{k+1|k}^t \tag{25}$$
$$\widetilde{P}_{k+1|k}^t = \mathbb{E}[\tilde{e}_{k+1|k}^t \tilde{e}_{k+1|k}^{tT} \,|\, \tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t] \tag{26}$$
$$\overline{P}_{k+1|k}^t = \mathbb{E}[\tilde{e}_{k+1|k}^t \tilde{e}_{k+1|k}^{tT}] = \mathbb{E}[\widetilde{P}_{k+1|k}^t] \tag{27}$$

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

$$\begin{aligned}\tilde{e}_{k+1|k}^t &= Ax_k + w_k - A\big(\tilde{x}_{k|k-1} + \gamma_k^t K_{t-k}(\gamma_k^t C x_k + v_k - C\tilde{x}_{k|k-1})\big)\\ &= A(I - \gamma_k^t K_{t-k} C)\tilde{e}_{k|k-1}^t + w_k - \gamma_k^t A K_{t-k} v_k\end{aligned} \tag{28}$$
$$\widetilde{P}_{k+1|k}^t = A(I - \gamma_k^t K_{t-k} C)\widetilde{P}_{k|k-1}^t (I - \gamma_k^t K_{t-k} C)^T A^T + Q + \gamma_k^t A K_{t-k} R K_{t-k}^T A^T \tag{29}$$
$$\overline{P}_{k+1|k}^t = \lambda_{t-k} A(I - K_{t-k} C)\overline{P}_{k|k-1}^t (I - K_{t-k} C)^T A^T + (1-\lambda_{t-k}) A\overline{P}_{k|k-1}^t A^T + Q + \lambda_{t-k} A^T K_{t-k} R K_{t-k}^T A^T \tag{30}$$

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 $\tilde{e}_{k|k-1}^t$, and the fact that v_k and w_k are zero mean. For ease of notation let us define the following operator:

$$\mathcal{L}_\lambda(K, P) = \lambda A(I - KC)P(I - KC)^T A^T + (1-\lambda)APA^T + Q + \lambda AKRK^T A^T \tag{31}$$

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

$$\overline{P}_{t-N+2|t-N+1}^t = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, \overline{P}_{t-N+1|t-N}^{t-1}) \tag{32}$$
$$\overline{P}_{t-k+1|t-k}^t = \mathcal{L}_{\lambda_k}(K_k, \overline{P}_{t-k|t-k-1}^t), \qquad k = N-2, \ldots, 0 \tag{33}$$

Observe that Equations (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

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

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 $0 \leq \lambda \leq 1$. Also consider the following operator:

$$\Phi_\lambda(P) = APA^T + Q - \lambda\, APC^T(CPC^T + R)^{-1}CPA^T \tag{35}$$

and the gain $K_P = PC^T(CPC^T + R)^{-1}$.

Then the following statements are true:

(a) $\mathcal{L}_\lambda(K, P) = \Phi_\lambda(P) + \lambda A(K - K_P)(CPC^T + R)(K - K_P)^T A^T$.
(b) $\mathcal{L}_\lambda(K, P) \geq \Phi_\lambda(P) = \mathcal{L}_\lambda(K_P, P), \quad \forall K$
(c) $(P_1 \geq P_2) \Longrightarrow \big(\Phi_\lambda(P_1) \geq \Phi_\lambda(P_2)\big)$.
(d) $(\lambda_1 \geq \lambda_2) \Longrightarrow \big(\Phi_{\lambda_1}(P) \leq \Phi_{\lambda_2}(P)\big), \quad \forall P$.
(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 $P^* \leq S^*$.

(i) If A is strictly stable, then $P^* = \Phi_\lambda(P^*)$ has always a solution. Otherwise, there exist λ_c such that $P^* = \Phi_\lambda(P^*)$ has a solution if and only if $\lambda > \lambda_c$. Also $\lambda_{min} \leq \lambda_c \leq \lambda_{max}$, where $\lambda_{max} = 1 - \frac{1}{\prod_i |\sigma_i^u|^2}$, $\lambda_{min} = 1 - \frac{1}{\max_i |\sigma_i^u|^2}$, and $|\sigma_i^u| \geq 1$ are the unstable eigenvalues of A. In particular $\lambda_c = \lambda_{max}$ if *rank*$(C) = 1$, and $\lambda_c = \lambda_{min}$ if C is square and invertible.

(j) The critical probability λ_c and the fixed point $P^* = \Phi_\lambda(P^*)$ for $\lambda > \lambda_c$ can be obtained as the solutions of the following semi-definite programming (SDP) problems: $\lambda_c = \inf\{\lambda \,|\, \Psi_\lambda(Y, Z) > 0, 0 \leq Y \leq I, \text{for some } Z, Y \in \mathbb{R}^{n\times n}\}$, and $P^* = \text{argmax}\{\textit{trace}(P) \,|\, \Theta_\lambda(P) \geq 0, P \geq 0\}$ where:

$$\Psi_\lambda(Y, Z) = \begin{bmatrix} Y & Y & \sqrt{\lambda} Z R^{\frac{1}{2}} & \sqrt{\lambda}(YA + ZC) & \sqrt{1-\lambda} YA \\ Y & Q^\dagger & 0 & 0 & 0 \\ \sqrt{\lambda} R^{\frac{1}{2}} Z' & 0 & I & 0 & 0 \\ \sqrt{\lambda}(A'Y + C'Z') & 0 & 0 & Y & 0 \\ \sqrt{1-\lambda} A'Y & 0 & 0 & 0 & Y \end{bmatrix} \tag{36}$$

$$\Theta_\lambda(P) = \begin{bmatrix} APA' - P & \sqrt{\lambda} APC' \\ \sqrt{\lambda} CPA' & CPC' + R \end{bmatrix} \tag{37}$$

(k) If there exist $P^* > 0$ and K such that $P^* = \mathcal{L}_\lambda(K, P^*)$, then the matrix $A_c = A(I - \lambda KC)$ is strictly stable.

Proof: See Appendix. ■

The previous lemma 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 in closed form, from fact (j) follows that they can be efficiently computed 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.

The following theorem shows how compute the optimal 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 $R > 0$. Also consider the arrival process defined by Equations (18)-(20), and the set of estimators with constant gains $\{K_k\}_{k=0}^N$ defined in Equations (21)-(23). If A is not strictly stable and $\lambda_{loss} \geq 1 - \lambda_c$, where λ_c is defined in Lemma 1(j), then there exist no stable estimator with constant gains. Otherwise, let N such that $\lambda_N > \lambda_c$ and consider the optimal gains $\{K_k^N\}_{k=0}^N$ defined as follows:

$$K_k^N = V_k^N C^T (C V_k^N C^T + R)^{-1}, \quad k = 0, \ldots, N \tag{38}$$

$$V_{N-1}^N = \Phi_{\lambda_{N-1}}(V_{N-1}^N) \tag{39}$$

$$V_k^N = \Phi_{\lambda_k}(V_{k+1}^N), \quad k = N-1, \ldots, 0 \tag{40}$$

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:

$$T_N^N = \mathcal{L}_{\lambda_N}(K_N, T_N^N) \tag{41}$$

$$T_k^N = \mathcal{L}_{\lambda_k}(K_k, T_{k+1}^N), \quad k = N-1, \ldots, 0 \tag{42}$$

then $\lim_{t\to\infty} \overline{P}^t_{t-k+1|t-k} = T^N_k$, and $V^N_k \leq T^N_k$ for $k = 0, \ldots, N$. Also $V^{N+1}_0 \leq V^N_0$. Finally, if $\tau_{max} < \infty$, then $V^N_0 = V^{\tau_{max}}_0$ for all $N \geq \tau_{max}$.

Proof: See Appendix. ∎

The previous theorem shows that the optimal gains can be obtained by finding the fixed point of a modified algebraic Ricatti Equation (39) and then iterating N time an operator with the same structure but with different λ_k. The theorem also demonstrates that a stable 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. If the system to be estimated is unstable, then the estimator is stable if and only if the packet loss probability λ_{loss} is sufficiently small, independently of the packet delay τ_{max}, thus recovering the same stability conditions derived in [33]. Moreover, the performance of the estimator, i.e. its steady state error covariance $\lim_{t\to\infty} P_{t+1|t} = \lim_{t\to\infty} \mathbb{E}[e_{t+1|t}e^T_{t+1|t}] = V^N_0$, 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^N_0 and buffer length N which is directly related to computational requirements.

VI. Optimal estimation with co-located smart sensors

In this section we describe an alternative data pre-processing at the sensor location which improves the overall performance of the estimator at the receiver side. This scheme was independently proposed in [46] and [37] in a scenario where no delay was present but only packet loss. The authors suggested to compute and transmit the state estimate rather than the raw measurement. As will be shown shortly, this approach gives an estimator that has a better performance as compared to the time-varying Kalman filter of Section IV. However, it is applicable only if some computational resources are available on the sensor, commonly known as "smart sensor", and if all entries of the observation vector y_t are collected from sensors which are co-located. For example, this scenario is not suitable in applications running over sensor networks where sensors are distributed and have very limited computation resources [47]. Adding specific coding schemes and data processing at the sensor side could further improve performance, but this would require the modification of the lower layers of the communication protocols, differently from the architecture proposed here. Besides, this estimator architecture proposed here is useful per se since it provides a computable lower bound for the performance of the optimal time-varying filter proposed in Section IV.

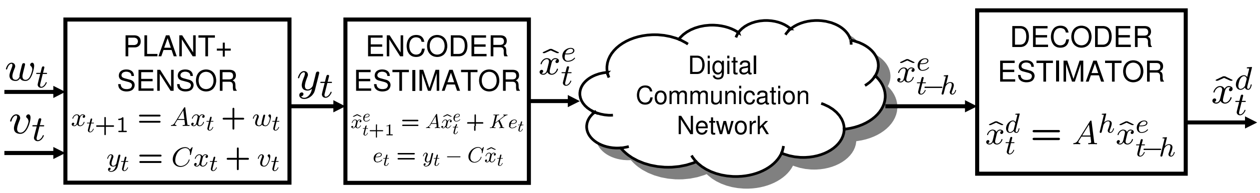

Fig. 6. Smart sensor with state estimator at encoder site before transmission.

Rather than sensing the raw measurements y_t over the DCN the sensor compute the optimal state estimate as follows:

$$\hat{x}^e_t = A\hat{x}^e_{t-1} + K^e_t(y_t - A\hat{x}^e_{t-1}) \quad (43)$$

$$K^e_t = P^e_t C^T(CP^e_tC^T + R)^{-1} \quad (44)$$

$$P^e_{t+1} = AP^e_tA^T + Q - AP^e_tC^T(CP^e_tC^T + R)^{-1}CP^e_tA^T = \Phi_1(P_t) \quad (45)$$

$$P^e_0 = P_0, \quad \hat{x}^e_0 = \bar{x}_0 \quad (46)$$

These are the equations for the standard Kalman filter, i.e. the minimum error covariance estimator $\hat{x}^e_t = \mathbb{E}[x_t \,|\, y_t, \ldots, y_1]$ whose estimation error $e^e_t = x_{t+1} - A\hat{x}^e_t$ has covariance $cov(e^e_t) = \mathbb{E}[e^e_t e^{e^T}_t \,|\, y_t, \ldots, y_1] = P^e_t$. The state estimate computed by the sensor encoder is then transmitted over the DCN to the decoder estimator. Using the same notation of Equation (5) the value stored at the buffer can be written as follows:

$$\tilde{y}^t_k = \gamma^t_k \hat{x}^e_k \quad (47)$$

Let us define the delay of the most recent packet arrived at the decoder estimator as $\kappa_t = t - \max\{k \,|\, \gamma^t_k = 1\}$ if $\exists \gamma^t_k = 1$, or $\kappa_t = t$ otherwise. The estimate of current state at the decoder estimator $\hat{x}^d_t$ is computed as follows:

$$\hat{x}^d_t = A^{\kappa_t}\tilde{y}^t_{t-\kappa_t} = A^{\kappa_t}\hat{x}^e_{t-\kappa_t} \quad (48)$$

Note that the decoder estimate is equivalent to $\hat{x}^d_t = \mathbb{E}[x_t \,|\, y_{t-\kappa_t}, \ldots, y_1]$ and that the its error $e^d_t = x_{t+1} - A\hat{x}^d_t$ has covariance:

$$cov(e^d_t) = \mathbb{E}[e^d_t e^{d^T}_t \,|\, y_{t-\kappa_t}, \ldots, y_1] = \Phi_0^{t-\kappa_t}(\mathbb{E}[e^d_{t-\kappa_t} e^{d^T}_{t-\kappa_t} \,|\, y_{t-\kappa_t}, \ldots, y_1]) = \Phi_0^{t-\kappa_t}(P^e_{t-\kappa_t}) = \Phi_0^{t-\kappa_t} \circ \Phi_1^{\kappa_t}(P_0),$$

where the superscript of $\Phi^n_\lambda(P)$ indicates $\Phi_\lambda \circ \cdots \circ \Phi_\lambda(P)$ composed n-times. Therefore, the decoder estimator error at any time step t is equivalent to the optimal estimator that one would obtain if all observations up to time $t - \kappa_t$ were successfully delivered. This estimation architecture is superior to the estimation architecture proposed in Section IV, in fact the estimator obtained in Theorem 1 has error covariance $cov(x_{t+1} - A\hat{x}^t_{t|t}) = P^t_{t+1|t}$, where $P^t_{t+1|t}$ is given by Equations (13)-(14) and can be written as:

$$\begin{aligned}
P^t_{t+1|t} &= \Phi_{\gamma^t_t} \circ \ldots \circ \Phi_{\gamma^t_1}(P_0) \\
&= \Phi_0^{t-\kappa_t} \circ \Phi_{\gamma^t_{\kappa_t}} \circ \ldots \circ \Phi_{\gamma^t_1}(P_0) \\
&\geq \Phi_0^{t-\kappa_t} \circ \Phi_{\gamma^t_{\kappa_t}} \circ \ldots \circ \Phi_1(P_0) \\
&\geq \Phi_0^{t-\kappa_t} \circ \Phi_1 \circ \ldots \circ \Phi_1(P_0) \\
&= \Phi_0^{t-\kappa_t} \circ \Phi_1^{\kappa_t}(P_0) \\
&= cov(e^d_t)
\end{aligned}$$

where we used the facts $\gamma^t_k = 0$ for $k > \kappa_t$, $\gamma^t_k \leq 1$ for $k \leq \kappa_t$, and Lemma 1(d). Therefore, the error covariance of the estimator proposed in this section is smaller than the error covariance of estimator proposed in Section IV. We can summarize the previous result in the following theorem:

Theorem 3: Let us consider the stochastic linear system given in Equations (1)-(2), where $R > 0$. Also consider the packet arrival process defined by Equation (3). Let $\hat{x}^y_t = \hat{x}^t_{t|t}$ the optimal estimator given by Equations (11)-(14) when raw measurements y_t are transmitted over the network. Let $\hat{x}^d_t$ the estimator given by Equation (48) where the state estimate $\hat{x}^e_t$ defined by Equations (43)-(46) is pre-computed by the sensor and then transmitted over the network. Then the estimation error covariance of $\hat{x}^e_t$ is always smaller than the estimation error covariance of $\hat{x}^t_t$, i.e.

$$cov(x_t - \hat{x}^d_t) \leq cov(x_t - \hat{x}^y_t), \quad \forall t.$$

Besides having a better performance, the estimator proposed in this section requires very limited computational requirements at the receiver side, in fact it suffices to store the most recent packet arrived at the receiver and then to compute the best state estimate at current time by pre-multiplying the packet data with a matrix which depend on the packet delay. Moreover, as for the estimator of Section IV, also the estimator based on co-located smart sensors does not require any statistical a-priori knowledge of the arrival process.

However, if the packet arrival statistics are stationary and i.i.d, then it is possible to give stability criteria and to compute the expected error covariance as shown in the following theorem:

Theorem 4: Let us consider the stochastic linear system given in Equations (1)-(2), where (A, C) is observable, $(A, Q^{1/2})$ is reachable, and $R > 0$. Also consider the arrival process defined by Equations (18)-(20), and the estimator architecture given by Equations (43)-(48). Then the estimator is stable if and only if A is stable, or $\lambda_{loss} < \frac{1}{|\sigma^u_{max}(A)|^2}$, where $\sigma^u_{max}(A)$ is the largest eigenvalue of the matrix A. If the estimator is stable then the covariance of the estimation error defined as $e^d_t = x_{t+1} - Ax^d_t$ has the following property:

$$\lim_{t\to\infty} \mathbb{E}[e^d_t e^{dT}_t] = D^\infty = \lim_{N\to\infty} D^N_0 \tag{49}$$

where the matrix D^N_0 is computed as follows:

$$D^N_N = (1-\lambda_N)AD^N_N A^T + (1-\lambda_N)Q + \lambda_N P^e_\infty \tag{50}$$

$$D^N_k = (1-\lambda_k)AD^N_{k+1} A^T + (1-\lambda_k)Q + \lambda_k P^e_\infty, \quad k = N-1, \ldots, 0 \tag{51}$$

and P^e_∞ is the unique positive definite solution of the Ricatti Equation $P^e_\infty = \Phi_1(P^e_\infty)$. If $\tau_{max} < \infty$, then $D^\infty = D^{\tau_{max}}_0 = D^N_0$, for all $N \geq \tau_{max}$.

Proof: See Appendix. ■

The previous theorem shows that performance of the smart optimal estimator under the assumption of i.i.d. packet arrival process, can be obtained by solving the Lyapunov Equation (50) and then iterating $N = \tau_{max}$ linear equations (51), if τ_{max} is finite. Otherwise if $\tau_{max} = \infty$, then D^∞ can be computed to any arbitrary precision through the sequence D^N_0.

VII. Numerical Examples

Here we illustrate the use of the tools developed in the previous sections with the aid of some numerical examples.

Let us consider the following probability function of packet delay:

$$\lambda_h = \begin{cases} 0.05\,h, & h = 0, \ldots, 15 \\ 0.75, & h > 15 \end{cases} \tag{52}$$

which is depicted in Fig. 7.

Let us consider the following discrete time system:

$$A = \begin{bmatrix} 1.001 & 0.05 \\ 0.05 & 1.001 \end{bmatrix}, \quad C = [\ 1 \quad 0\ ], \quad R = 0.01, \quad Q = \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix} \tag{53}$$

which corresponds to the discretization with sampling period $T = 0.05$ of the continuous time system $\ddot{x} - x = 0$. This system has one stable pole and one unstable pole, and it is the model for the discrete time dynamics of an inverted pendulum. The discrete time eigenvalues of the matrix A are $eig(A) = (1.05, 0.95)$, which give the critical probability $\lambda_c = 1 - 1/1.05^2 = 0.095$, as stated in Lemma 1(i). According to Theorem 2 and Theorem 4 both estimators presented in Section V and Section VI are stable if and only if $N \geq 2$, in fact $\lambda_1 = 0.05 < \lambda_c$ and $\lambda_2 = 0.01 > \lambda_c$.

The trace of the covariance of the estimator error with constant gains, V^N_0, and the estimator error for smart sensors, D^N_0 are shown in Fig. 8. As mentioned in Section IV, the error covariance for time-varying optimal estimator of Theorem 1 cannot be computed explicitly but it is upperbounded and lowerbounded by V^N_0 and by D^N_0, respectively. It is interesting to compare the performance of these estimators with the error covariance $P^e_\infty = \Phi_1(P^e_\infty)$, shown in the same figure, corresponding to the ideal case when there is no packet loss and no delay. In fact, P^e_∞ gives an idea of the degradation due to the communication

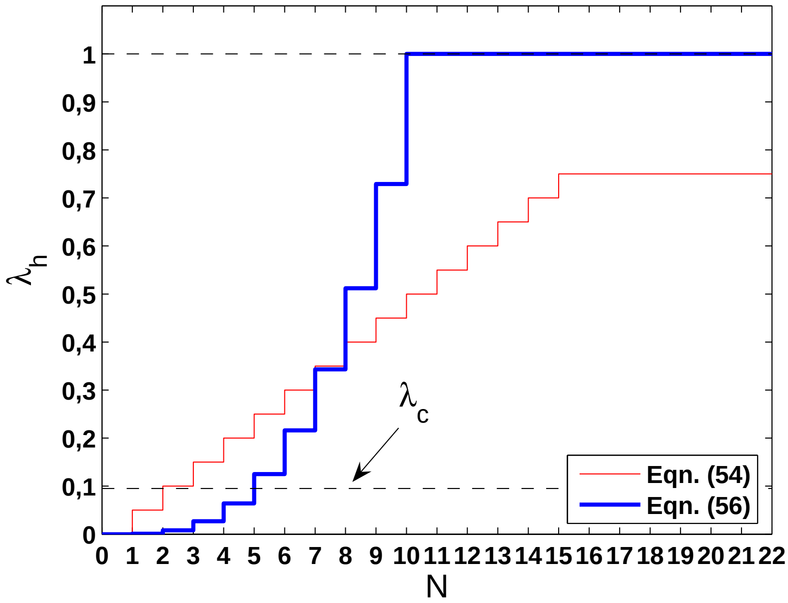

Fig. 7. Probability function of packet delay for different scenarios and critical probability λ for dynamical systems (53).

network. It is also relevant to evaluate the performance of an estimator with constant gains designed without exploiting the prior knowledge about the packet arrival statistics. A natural choice is to use the standard Kalman gain $K^e_\infty = P^e_\infty C^T(CP^e_\infty C^T + R)^{-1}$, i.e. $K_k = K^e_\infty, k = 0, \ldots, N$ rather than the optimal constant gains K^N_k defined in Theorem 2. The corresponding expected error covariance T^N_0 can be obtained by Equations (41)-(42) and it is shown in Fig. 8. From this example it is clear that the tools developed in this paper can help to substantially reduce the degradation of performance when statistics of packet arrival are available.

As a second example, we show empirical computation of the expected error covariance for the different estimator designs presented in this paper. We consider the same dynamical system above given by Equations (53) and we assume that packet delay is generated according to the following arrival statistics:

$$\lambda^1 = (\lambda^1_0, \lambda^1_1, \lambda^1_2, \lambda^1_3) = (0.2, 0.4, 0.7, 1), \quad 0 \leq t \leq 50 \tag{54}$$

$$\lambda^2 = (\lambda^2_0, \lambda^2_1, \lambda^2_2, \lambda^2_3) = (0, 0.3, 0.5, 0.7), \quad t > 50 \tag{55}$$

which models an increase in packet delay due, for example, to network traffic congestion. We consider four different estimators: the time-varying filter of Section IV, the filter with smart sensor of Section VI, and two filters with constant gains optimized for the two different arrival statistics, λ^1 and λ^2. Fig. 9 shows the trace of the empirical expected estimator error covariance, i.e. $P^{emp}(t) = \frac{1}{N}\sum_{i=1}^N P_{t|t}$, averaged over $N = 1000$ Monte Carlo runs. The performance of the time-varying filter seems to be closer to the optimal filter with static gains then the filter with the smart sensor, although a non negligible gap is visible, thus suggesting that none of the two bounds are tight. Moreover, this example clearly shows that the time-varying filter and the filter with smart sensor are optimal regardless the underlaying packet arrival statistics. Differently, the filters with static gain can achieve good performance only if the underlying arrival statistics is well known. In fact, filter optimized for λ^1 (thick solid line) performs better than the filter optimized for λ^2 (thick dashed line) within the first 50 time steps, but then this situation reverses when the arrival statistics change. Therefore, besides giving better performance of any static gain filter, the time-varying filter optimally cope with time-varying packet arrival statistics.

As a final example, we illustrate how these tools can be also used to compare two different communication protocols. Let us consider a protocol giving rise to arrival statistics of Equation (52) and a protocol

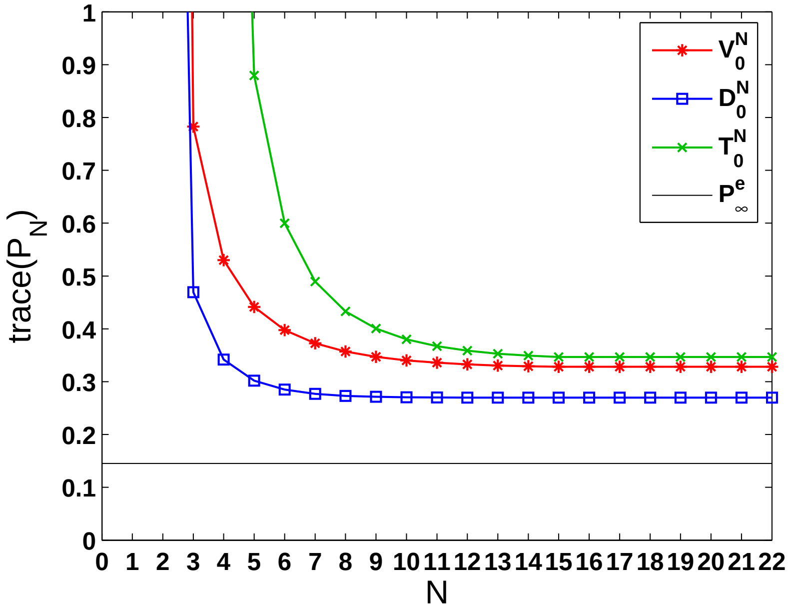

Fig. 8. Trace of the steady state error covariance for the optimal estimator with constant gains (V_0^N), for the optimal estimator with a smart sensor (D_0^N). The horizontal line P_∞^e corresponds to the trace of the error covariance in the ideal scenario with no delay and no packet loss, i.e. $\lambda_h = 1$ for all h, while T_0^N is the actual steady state error when using the Kalman gain K_∞^e. The error covariances V_0^N, D_0^N are unbounded for $N < 2$, while the covariance P_∞^e is unbounded for $N < 4$, and they are all cosnstant constant for $N \geq \tau_{max} = 15$.

giving rise to the following arrival statistics:

$$\lambda_h = \begin{cases} (\frac{h}{10})^3, & h = 0, \dots, 10 \\ 1, & h > 10 \end{cases} \tag{56}$$

for which $\tau_{max} = 10$, and it is graphically shown in Fig. 7. These two protocols are substantially different: the first protocol has larger packet delivery with small delay, but also larger overall packet loss than the first protocol, therefore it is difficult to evaluate which one is better suited for a real-time control application. In Fig. 10 it is shown the trace of the error covariance V_0^N for the two protocols with respect to the system dynamics of Equation (53). For a buffer with a short memory the first protocol performs better, but for a buffer of length $N = 10$ the second protocol starts performing better as the larger packet delivery can compensate for a larger delay of arrived packets. If buffer length is further increased, then the first protocol returns to perform better. This example clearly shows how optimal estimation design can be used to evaluate and compare the performance of different communication protocols with respect to a specific real-time application, which is otherwise very difficult if based on intuition or empirical rule-of-thumbs.

VIII. Conclusions

In this work we proposed a framework to optimally design and analyze the performance of estimators in networked control system subject to simultaneous random packet delay and packet dropped. We showed that the optimal estimator is time-varying, stochastic, and does not depend on the specific communication protocol adopted as long as measurements are time-stamped and can be re-ordered at the estimator site. Also two alternative optimal estimator designs based on finite memory buffers and constant gains were described and it was shown that if packet arrival is i.i.d., then the estimators are mean square stable if and only if the packet loss probability is below a critical value. Therefore, implicitly we also provided necessary and sufficient conditions about existence of stable estimators. Finally, we presented numerical algorithms for the computation of the expected estimator error covariance of all the proposed estimators.

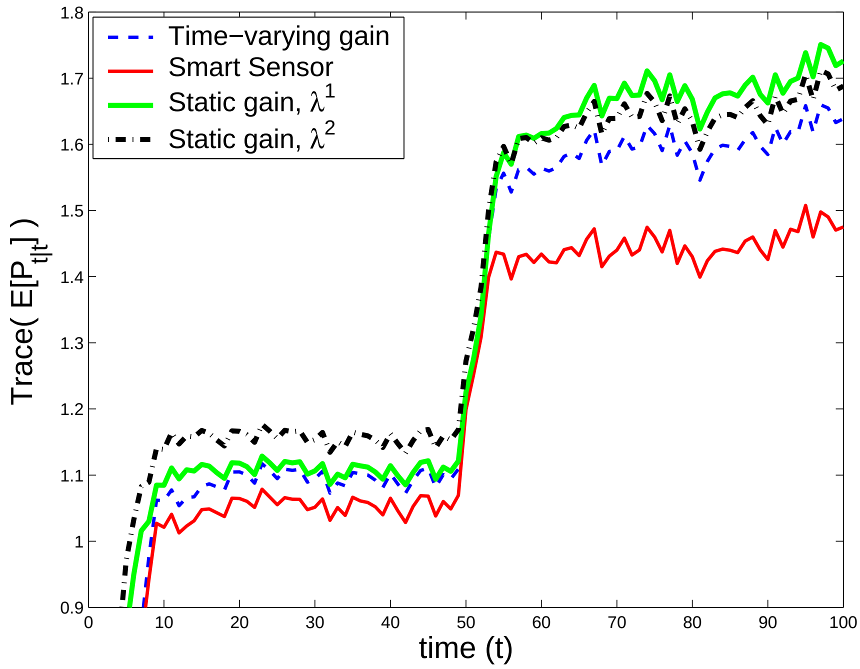

Fig. 9. Time series of trace of trace of empirical error covariance $P^{emp}(t) \cong \mathbb{E}[P_{t|t}]$ averaged over 1000 Monte Carlo runs for four different estimators under the same conditions. The packet delay arrival probability distribution changes from λ^1 to λ^2 at time $t = 50$. The filters with static gains are optimized for λ^1 and λ^2.

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 $\{K_k^N\}_{k=0}^N$ 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.

A possible future avenue of research is the extension of this work to the design of optimal LQG-like controller design. This is not a trivial step as many important assumptions in standard LQG control, like the separation principle, do not always hold for NCSs [36]. Another research direction is the implementation and testing of these tools in real-time control applications for wireless sensor networks. A preliminary attempt has already been successfully applied to multiple target tracking [48], but extensive experimental work is still needed.

IX. Appendix

Proof of Lemma 1. Some of these statements can be found in [33] or can be derived along similar lines, therefore only a brief sketch is reported here.

(a) This fact can be verified by direct substitution

(b) This statement follows from previous fact and $\lambda A(K - K_P)(CPC^T + R)(K - K_P)^T A^T \geq 0$.

(c) From previous fact $\Phi_\lambda(P_1) = \mathcal{L}_\lambda(K_{P_1}, P_1) \geq \mathcal{L}_\lambda(K_{P_1}, P_2) \geq \mathcal{L}_\lambda(K_{P_2}, P_2) = \Phi_\lambda(P_2)$.

(d) From Equation (35) we have $\Phi_{\lambda_1}(P) - \Phi_{\lambda_2}(P) = -(\lambda_1 - \lambda_2)APC^T(CPC^T + R)^{-1}CPA^T \leq 0$.

(e) Uniqueness and strictly positive definiteness of P^* follows from the assumption that $(A, Q^{1/2})$ is reachable [33].

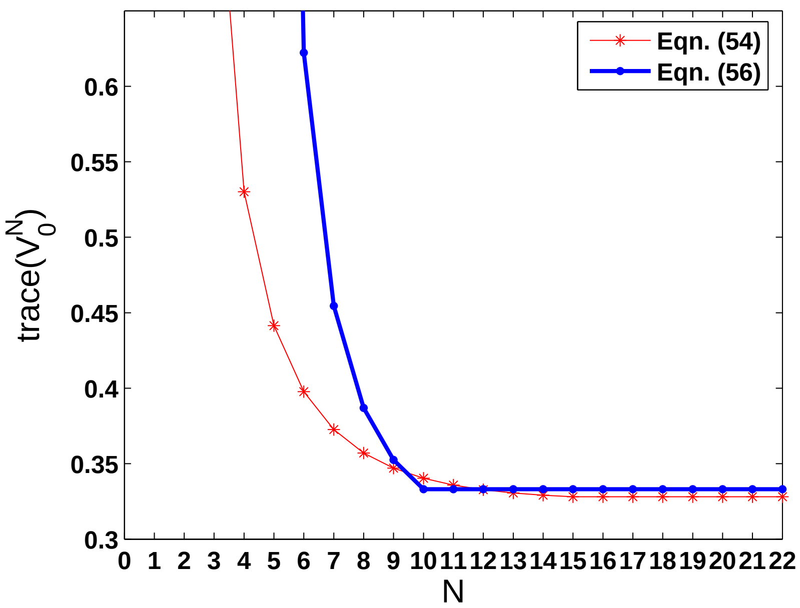

Fig. 10. Trace of the steady state error covariance for the optimal estimator with constant gains (V_0^N) for two different communication protocols whose packet arrival statistics is given by Equations (52) and (56).

(f) Consider $P_{t+1} = \Phi_{\lambda_1}(P_t)$ and $S_{t+1} = \Phi_{\lambda_2}(S_t)$ where $P_0 = S_0 = 0$. From fact (c) and (e) it follows that $P_t \leq S_t$. Also $P_t \leq P_1^*$ and $S_t \leq P_2^*$, therefore $\lim_{t\to\infty} P_t = \bar{P}$, $\lim_{t\to\infty} S_t = \bar{S}$, and $\bar{P} \leq \bar{S}$. From fact (e) it follows that $\bar{P} = P_1^*$ and $\bar{S} = P_2^*$, and thus $P_1^* \leq P_2^*$

(g-h) Let consider $P_{t+1} = \Phi_\lambda(P_t)$ and $S_{t+1} = \mathcal{L}_\lambda(K, S_t)$ where $P_0 = S_0 = 0$. From fact (c) and monotonicity of operator $\mathcal{L}_\lambda(K, P)$ with respect to P we have $P_{t+1} \geq P_t$, $S_{t+1} \geq S_t$, and $P_t \leq S_t \leq S^*$ for all t. Since both sequences are monotonically increasing and bounded, then $\lim_{t\to\infty} P_t = \bar{P}$, $\lim_{t\to\infty} S_t = \bar{S}$, $\bar{P} = \Phi_\lambda(\bar{P})$, $\bar{S} = \mathcal{L}_\lambda(K, \bar{S})$, and $\bar{P} \leq \bar{S}$. From fact (e) it follows that $\bar{P} = P^*$ and $\bar{S} = S^*$. A complete proof for convergence from any initial condition can be obtained along the lines of Theorem 1 in [33], thus it is not reported here.

(i) The proof for existence of a critical probability λ_c was given in [33] and it is based on observability of (A, C) and monotonicity of $\Phi_\lambda(P)$ with respect to λ. The proof for $\lambda_c = \lambda_{max}$ when $rank(C) = 1$ can be found in [49][32] although it was not explicitly derived for the operator Φ_λ. The proof for $\lambda_c = \lambda_{min}$ when C is square and invertible was first proved in [50].

(j) The proof can be found in [33].

(k) Let us consider the linear operator $\mathcal{F}(P) = \lambda A(I-KC)P(I-KC)^T A^T + (1-\lambda)APA^T$. Clearly $\mathcal{L}_\lambda(K, P) = \mathcal{F}(P) + D$, where $D = Q + \lambda AKRK^T A^T \geq 0$. Consider the sequences $S_{t+1} = \mathcal{L}_\lambda(K_{P^*}, S_t)$, $T_{t+1} = \mathcal{L}_\lambda(K_{P^*}, T_t)$ with initial condition $S_0 = 0$, then $T_0 \geq 0$. Note that $S_t = \sum_{k=0}^{t-1} \mathcal{F}^k(D)$ and $T_t = \mathcal{F}^t(T_0) + \sum_{k=0}^{t-1} \mathcal{F}^k(D)$ for $t \geq 1$, where we define $\mathcal{F}^0(D) = D$ and $\mathcal{F}^{k+1}(D) = \mathcal{F} \circ \mathcal{F}^k(D)$. Therefore $\mathcal{F}^t(T_0) = T_t - S_t$. From fact (g) it follows $\lim_{t\to\infty} S_t = \lim_{t\to\infty} T_t = P^*$, therefore $\lim_{t\to\infty} \mathcal{F}^t(T_0) = 0$, for all $T_0 \geq 0$, i.e. the linear operator $\mathcal{F}()$ is strictly stable. Now consider the system $A_c = A(I - \lambda KC)$. The system is strictly stable if and only if $\lim_{t\to\infty} A_c^t x_0 = 0$, for all x_0. This is equivalent to $\lim_{t\to\infty} A_c^t x_0 x_0^T (A_c^T)^t = \mathcal{G}^t(X_0) = 0$, where $X_0 = x_0 x_0^T \geq 0$ and $\mathcal{G}^t(X_0) = A_c^t X_0 (A_c^T)^t$. Note that $\mathcal{G}(X_0) = AX_0A^T - 2\lambda AX_0(AKC)^T + \lambda^2 AKCX_0(AKC)^T = \mathcal{F}(X_0) + \lambda(\lambda-1)AKCX_0(AKC)^T \leq \mathcal{F}(X_0)$ since $\lambda(\lambda-1)AKCX_0(AKC)^T \leq 0$. Since we just proved that $\lim_{t\to\infty} \mathcal{F}^t(X_0) = 0$ for all $X_0 \geq 0$, then also $\lim_{t\to\infty} \mathcal{G}^t(X_0) \leq \mathcal{F}^t(X_0) = 0$ for $X_0 = x_0 x_0^T$, i.e. the system A_c is strictly stable. ∎

Proof of Theorem 2. First we prove by contradiction that there is no stable estimator with constant gains if A is not strictly stable and $\lambda_{loss} \geq 1 - \lambda_c$. 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 $\overline{P}^t_{t+1|t} = A\overline{P}^t_{t|t}A^T + Q$ also $\overline{P}^t_{t+1|t}$ must be bounded for all t. From Equations (32) and (33) it follows that $\overline{P}^t_{t+1|t}$ is bounded if and only if $\overline{P}^t_{t-k+1|t-k}$ for $k = 0, \ldots, N-1$ are bounded for all t. Therefore, since the bounded sequence $S_t = \overline{P}^t_{t-N+1|t-N}$ 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, by hypothesis $\lambda_{N-1} \leq \sup\{\lambda_h \mid h \geq 0\} = 1 - \lambda_{loss} \leq \lambda_c$. Consequently, according to Lemma 1(i), $P^* = \Phi_{\lambda_{N-1}}(P^*)$ cannot have a solution, which contradicts the hypothesis that a stable estimator exists.

Consider now the case when N is such that $\lambda_N > \lambda_c$. From Theorem 1(h) it follows that Equations (38)-(40) are well defined and have a solution. From Lemma 1(g) it follows that $\lim_{t\to\infty} \overline{P}^t_{t-k+1|t-k} = V^N_k$ for the optimal gains $\{K^N_k\}_{k=0}^{N-1}$, and $\lim_{t\to\infty} \overline{P}^t_{t-k+1|t-k} = T^N_k$ when using generic gains $\{K_k\}_{k=0}^{N-1}$. From Lemma 1(h) it follows that $V^N_{N-1} \leq T^N_{N-1}$. From Lemma 1(c) we have $V^N_{N-2} = \Phi_{\lambda_{N-2}}(V^N_{N-1}) \leq \mathcal{L}_{\lambda_{N-2}}(K_{N-2}, V^N_{N-1}) \leq \mathcal{L}_{\lambda_{N-2}}(K_{N-2}, T^N_{N-1}) = T^N_{N-2}$. Inductively, it is easy to show that $V^N_k \leq T^N_k$ for all $k = 0, \ldots, N-1$.

Now we want to show that $V^{N+1}_0 \leq V^N_0$. 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 $V^{N+1}_N = \Phi_{\lambda_N}(V^{N+1}_{N+1}) \leq \Phi_{\lambda_N}(V^N_N) = V^N_N$ and inductively $V^{N+1}_k \leq V^N_k$ 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 $V^N_N = \Phi_{\lambda_N}(V^N_N) = \Phi_{\lambda_{N-1}}(V^N_N) = V^N_{N-1} = \Phi_{\lambda_{N-1}}(V^N_{N-1}) = \Phi_{\lambda_{N-2}}(V^N_{N-1}) = V^N_{N-2} = \ldots = V^N_{\tau_{max}} = \Phi_{\lambda_{\tau_{max}}}(V^N_{\tau_{max}})$. Since $V^{\tau_{max}}_{\tau_{max}} = \Phi_{\lambda_{\tau_{max}}}(V^{\tau_{max}}_{\tau_{max}})$, then by Lemma 1(e) we have that $V^{\tau_{max}}_{\tau_{max}} = V^N_{\tau_{max}}$. According to Equation (40) we also have $V^{\tau_{max}}_k = V^N_k$ for $k = \tau_{max}, \ldots, 0$, which concludes the theorem. ■

Proof of Theorem 4. The proof follows along the same lines of Theorems 1 and 2. Let us consider the following estimator:

$$\begin{aligned}
\breve{x}^t_{t-N|t-N} &= (1-\gamma^t_{t-N})A\breve{x}^{t-1}_{t-N-1|t-N-1} + \gamma^t_{t-N}\hat{x}^s_{t-N} \\
\breve{x}^t_{t-k|t-k} &= (1-\gamma^t_{t-k})A\breve{x}^t_{t-k-1|t-k-1} + \gamma^t_{t-k}\hat{x}^s_{t-k}, \quad k = N-1, \ldots, 0 \\
\breve{x}^t_{-k|-k} &= \bar{x}_0, \;\; \gamma^t_{-k} = 0, \;\; \hat{x}^s_{-k} = 0
\end{aligned}$$

It should be clear that by construction $\hat{x}^d_t = \breve{x}^t_{t|t}$ if and only if $N \geq \tau_{max}$. If $N < \tau_{max}$, then the estimator $\hat{x}^d_t$ cannot be optimal. Let us consider the estimator error defined as $\breve{e}^t_{k+1|k} = x_{k+1} - A\breve{x}^t_{k|k}$ that can be written as:

$$\begin{aligned}
\breve{e}^t_{k+1|k} &= x_{k+1} - A\big((1-\gamma^t_k)A\breve{x}^t_{k-1|k-1} + \gamma^t_k\hat{x}^s_k\big) = (1-\gamma^t_k)(x_{k+1} - AA\breve{x}^{t-1}_{k-1|k-1}) + \gamma^t_k(x_{k+1} - A\hat{x}^s_k) \\
&= (1-\gamma^t_k)(A(x_k - A\breve{x}^{t-1}_{k-1|k-1}) + w_k) + \gamma^t_k e^s_k = (1-\gamma^t_k)(A\breve{e}^t_{k|k-1} + w_k) + \gamma^t_k e^s_k
\end{aligned}$$

and its error covariance $\breve{P}^t_{k+1|k} = \mathbb{E}[\breve{e}^t_{k+1|k}\breve{e}^{tT}_{k+1|k} \mid \gamma^t_k, \ldots, \gamma^t_1]$ is then given by:

$$\begin{aligned}
\breve{P}^t_{t-N+1|t-N} &= (1-\gamma^t_{t-N})(A\breve{P}^{t-1}_{t-N|t-N-1}A^T + Q) + \gamma^t_{t-N}P^e_{t-N} \\
\breve{P}^t_{t-k+1|t-k} &= (1-\gamma^t_{t-k})(A\breve{P}^t_{t-k|t-k-1}A^T + Q) + \gamma^t_{t-k}P^e_{t-k}, \quad k = N-1, \ldots, 0 \\
\breve{P}^t_{-k|-k} &= P_0, \;\; \gamma^t_{-k} = 0
\end{aligned}$$

The error covariance $\breve{P}^t_{t+1|t}$ is then stochastic and depends on the arrival sequence. However since it is linear in the arrival sequence γ^t_k, it is possible to compute the expected error covariance $\mathbb{E}[\breve{P}^t_{k+1|k}] = \widehat{P}^t_{k+1|k}$ as follows:

$$\begin{aligned}
\widehat{P}^t_{t-N+1|t-N} &= (1-\lambda_N)A\widehat{P}^{t-1}_{t-N|t-N-1}A^T + (1-\lambda_N)Q + \lambda_N P^s_{t-N} \\
\widehat{P}^t_{t-k+1|t-k} &= (1-\lambda_k)A\widehat{P}^t_{t-k|t-k-1}A^T + (1-\lambda_k)Q + \lambda_k P^s_{t-k}, \quad k = N-1, \ldots, 0 \\
\widehat{P}^t_{-k|-k} &= P_0, \;\; \gamma^t_{-k} = 0
\end{aligned}$$

Since $\lim_{t\to\infty} P^e_{t-k} = P^e_\infty$ where $P^e_\infty = \Phi_1(P^e_\infty)$, then $\lim_{t\to\infty} \widehat{P}^t_{t-N+1|t-N} = D^N_N$ exists and it is finite if and only if $\sqrt{1-\lambda_N}A$ is stable, i.e. $\sqrt{1-\lambda_N}\,|\sigma^u_{max}(A)| < 1$. This is equivalent to $\lambda_N > 1 - \frac{1}{|\sigma^u_{max}(A)|^2}$. Such λ_N exists if and only if $\lambda_{loss} < \frac{1}{|\sigma^u_{max}(A)|^2}$. If this condition holds then Equations (50)-(51) follow. Also it is simple to show that $\{D^N_0\}^\infty_{N=0}$ is a decreasing function of N and bounded from below, therefore $\lim_{N\to\infty} D^N_0 = D^\infty$. Moreover, since by construction $\lim_{N\to\infty} \hat{x}^d_t = \breve{x}^t_{t|t}$, then also $\mathbb{E}[e^d_t e^{d^T}_t] = \lim_{N\to\infty} \widehat{P}^t_{t+1|t}$, and therefore it follows $\lim_{t\to\infty} \mathbb{E}[e^d_t e^{d^T}_t] = D^\infty$. Following Theorem 2, it is easy to show that if $\tau_{max} < \infty$, then $D^\infty = D^{\tau_{max}}_0 = D^N_0$, for all $N \geq \tau_{max}$, which concludes the theorem. ∎

REFERENCES

[1] "Technology of networked control systems," *Proceedings of the IEEE, Special issue*, vol. 95, no. 1, pp. 5–312, January 2007.

[2] M. Kintner-Meyer and R. Conant, "Opportunities of wireless sensors and controls for building operation," *Energy Engineering Journal*, vol. 102, no. 5, pp. 27–48, 2005.

[3] M. Nekovee, "Ad hoc sensor networks on the road: the promises and challenges of vehicular ad hoc networks," in *Workshop on Ubiquitous Computing and e-Research*, Edinburgh, UK, May 2005.

[4] 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.*

[5] 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.

[6] 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.

[7] A. Willig, K. Matheus, and A.Wolisz, "Wireless technology in industrial networks," *Proceedings of the IEEE*, vol. 93, no. 6, pp. 1130–1151, June 2005.

[8] "Sensor networks and applications," *Proceedings of the IEEE, Special issue*, vol. 91, August 2003.

[9] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, 2000.

[10] R. Cristescu and M. Vetterli, "On the optimal density for real-time data gathering of spatio-temporal processes in sensor networks," in *Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks, IPSN05*, 2005, pp. 159–164.

[11] H. Durrant-Whyte, M. Stevens, and E. Nettleton, "Data fusion in decentralised sensing networks," in *Proceeding on Information Fusion 2001*, Montreal, Canada, 2001, pp. 302–307.

[12] K. Akkaya and M. Younis, "A survey of routing protocols in wireless sensor networks," *Ad Hoc Network Journal*, vol. 3, no. 3, pp. 325–349, 2005.

[13] 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.

[14] 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.

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

[16] K. Gu, V. Kharitonov, and J. Chen, *Stability of time-delay systems*, 1st ed. Birkhauser, 2003.

[17] A. Cervin, D. Henriksson, B. Lincoln, J. Eker, and K.-E. Årzén, "How does control timing affect performance?" *IEEE Control Systems Magazine*, vol. 23, no. 3, pp. 16–30, June 2003.

[18] 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.

[19] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: an overview," *Proceedings of the IEEE, special issue*, vol. 95, no. 1, pp. 108–137, January 2007.

[20] L. Montestruque and P. Antsaklis, "Stability of model-based networked control systems with time-varying transmission times," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1562–1572, September 2004.

[21] W. Zhang and M. S. Branicky, "Stability of networked control systems with time-varying transmission period," *Allerton Conf. of Communication, Control and Computing*, October 2001.

[22] D. Nesic and A. Teel, "Input-output stability properties of networked control systems," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1650–1667, October 2004.

[23] D. Yue, Q. Han, and C. Peng, "State feedback controller design for networked control systems," *IEEE Transactions on Circuits and Systems*, vol. 51, no. 11, pp. 640–644, November 2004.

[24] 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.

[25] 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.

[26] 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.

[27] 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,*.

[28] 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.

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

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

[31] N. Elia and J. Eisembeis, "Limitation of linear control over packet drop networks," in *Proceedings of IEEE Conference on Decision and Control*, vol. 5, Bahamas, December 2004, pp. 5152–5157.

[32] N. Elia, "Remote stabilization over fading channels," *Systems and Control Letters*, vol. 54, pp. 237–249, 2005.

[33] 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.

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

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

[36] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE, special issue*, vol. 95, no. 1, pp. 163–187, January 2007.

[37] V. Gupta, D. Spanos, B. Hassibi, and R. M. Murray, "Optimal LQG control across a packet-dropping link," *Submitted to Systems and Control Letters*, 2005.

[38] Y. Xu and J. Hespanha, "Optimal communication logics for networked control systems," in *Proceedings of the IEEE Conf. on Decision and Control*, December 2004, pp. 3527–3532.

[39] J. Yook, D. Tilbury, and N. Soparkar, "Trading computation for bandwidth: reducing communication in distributed control systems using state estimators," *IEEE Transactions on Control Systems Technology*, vol. 10, no. 4, pp. 503–518, July 2002.

[40] A. Dana, V. Gupta, J. Hespanha, B. Hassibi, and R. Murray, "Estimation over communication networks: Performance bounds and achievability results," *Submitted for publication*, 2006.

[41] 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.

[42] 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.

[43] M. Maròti, B. Kusy, G. Simon, and Àkos Ldeczi, "The flooding time synchronization protocol," in *SenSys '04: Proceedings of the 2nd international conference on Embedded networked sensor systems*, 2004, pp. 39–49.

[44] 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.

[45] L. Shi, M. Epstein, A. Tiwari, and R.M.Murray, "Estimation with information loss: Asymptotic analysis and error bounds," in *Proceedings of IEEE CDC-ECC 05*, Seville, Spain, December 2005.

[46] Y. Xu and J. Hespanha, "Estimation under controlled and uncontrolled communications in networked control systems," in *To be presented at the IEEE Conference on Decision and Control*, Sevilla, Spain, December 2005.

[47] B. Sinopoli, C. Sharp, L. Schenato, S. Schaffert, and S. Sastry, "Distributed control applications within sensor networks," *Proceedings of the IEEE, Special Issue on Distributed Sensor Networks*, vol. 91, no. 8, pp. 1235–1246, August 2003.

[48] S. Oh, L. Schenato, P. Chen, and S. Sastry, "Tracking and coordination of multiple agents using sensor networks: system design, algorithms and experiments," *Proceedings of the IEEE, special issue*, vol. 95, no. 1, pp. 234–254, January 2007.

[49] N. Elia and S. Mitter, "Stabilization of linear systems with limited information," *IEEE Transaction on Automatic Control*, vol. 46, no. 9, pp. 1384–1400, 2001.

[50] T. Katayama, "On the matrix Riccati equation for linear systems with a random gain," *IEEE Transactions on Automatic Control*, vol. 21, no. 2, pp. 770–771, October 1976.