Kalman Filtering for networked control systems with random delay and packet loss

Luca Schenato

***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, this estimator is independent of the communication protocol and can be implemented using a finite memory buffer if and only if the delivered packets have a finite maximum delay. We also present an alternative estimator with constant gains and finite buffer memory for which, surprisingly, the stability 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.**

***Keywords*— Networked control systems, packet loss, random delay, optimal estimation, stability**

I. Introduction

Recent technological advances in MEMS, DSP capabilities, computing, and communication technology are revolutionizing our ability to build massively distributed networked control systems (NCS). These networks can offer access to an unprecedented quality and quantity of information which can revolutionize our ability in controlling of the environment. However, they also present challenging problems arising from the fact that sensors, actuators and controllers are not physically collocated and need to exchange information via a digital communication network. In particular, measurement and control packets are subject to random delay and loss. 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 networks. Sensor measurements are subject to random delay or might even be completely lost. We show that the minimum error covariance estimator is time-varying and stochastic which does not converge to a steady state. Moreover this estimator can be implemented using a finite memory buffer if and only if the delivered packets have a finite maximum delay and it is independent of the communication protocol. In particular, the memory length is equal to the maximum packet delay. We also present a suboptimal but simpler to implement estimator which constrains the estimator gains to be constant rather than stochastic as for the true optimal estimator. In particular we show how to compute the optimal static gains if the packet arrival statistic is stationary and known. We derive necessary and sufficient condition for stability of the estimator. Surprisingly we show that stability does not depend on packet delay but only on the packet loss probability which needs to be smaller that a threshold which depend on the unstable eigenvalues of the system to be estimated. We also provide quantitative measures for the expected error covariance of such estimators which turn out to be the solution of some modified algebraic Riccati equations and Lyapunov equations. These measures can be used to compare different communication protocols for real-time control applications. Very importantly, these results do not depend on the specific implementation of the digital communication network (wired bus, Bluetooth, ZigBee, Wi-Fi, etc ..) as long as the packet arrival statistics are stationary and i.i.d. In particular, this paper provides a useful tool for the designer of communication protocols since he/she can compare the performance of different protocols schemes. For example it is possible to quantitatively compare the performance of a protocol that has small average packet delay but a large packet loss probability with a protocol that has large packet delay but a low packet loss probability.

This research is supported in part by the European Community Research Information Society Technologies under Grant No. RECSYS IST-2001-32515 and under Grant Marie Curie N. SENSNET IRG-2005-014815, and by the Italian Ministry of Education, University and Research (MIUR).

L. Schenato is with the Department of Information Engineering, University of Padova, Via Gradenigo, 6/b, 35131 Padova, Italy `schenato@dei.unipd.it`

II. Problem formulation

Consider the following discrete time linear stochastic plant:

$$x_{t+1} = Ax_t + w_t \tag{1}$$

$$y_t = Cx_t + v_t, \tag{2}$$

where $t \in \mathbb{N} = \{0, 1, 2, \ldots\}$, $x, w \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $A \in \mathbb{R}^{n\times n}$, $y \in \mathbb{R}^m$, $C \in \mathbb{R}^{m\times n}$, (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 controllable, and $R > 0$.

Observation packets are then transmitted through a digital communication network (DCN), whose goal is to deliver packets from a source to a destination (see Fig. 1). Modern DSNs 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,

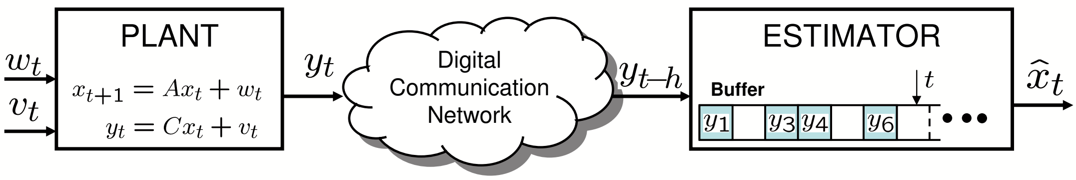

Fig. 1. Networked systems modeling. Sampled observations at the plant site are transmitted to the estimator site via a digital communication network. Due to retransmission and packet loss, observation packets arrive at the estimator site with possibly random delay.

data gathering, media-related, etc ..). In our work we model a DSN 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 observation loss. We assume that all observation packets correctly delivered to the estimator site are stored in an infinite buffer, as shown in Fig. 1. The arrival process is modeled by defining the random variable γ_k^t as follows:

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

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 buffer at time t, 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:

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

where t_k is the arrival time of observation y_k at the estimator site . The packet delay can be random, therefore there can be observation measurements that can arrive out of order at the estimator site (see Fig. 2 for $t = 5$) and there can be no packet or multiple packets delivered at the same time (see Fig. 2 for $t = 4$ and $t = 6$). 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 for most DSNs where packets allocate tens to hundreds of bits for measurement data. Also we do not consider channel noise since we assume that any bit error incurred during packet transmission is detected at the receiver and the packet is dropped.

If observation y_k is not yet arrived at the estimator at time t, we assume that a random number d_k stored in the k-slot of the buffer with zero mean and covariance

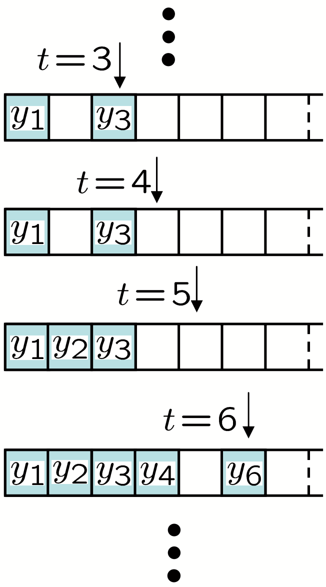

Fig. 2. Packet arrival sequence and buffering at the estimator location. Shaded squares correspond to observation packets that have been successfully received by the estimator. Cursor indicates current time.

$\mathbb{E}[d_k d_k^T] = R$, as shown in Fig. 1[1]. More formally, the value stored in the k-slot of the estimator buffer at time t can be written as follows:

$$\tilde{y}_k^t = \gamma_k^t y_k + (1-\gamma_k^t) d_k = \gamma_k^t C x_k + \gamma_k^t v_k + (1-\gamma_k^t) d_k \tag{5}$$

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 $\boldsymbol{\gamma}_t = (\gamma_1^t, \gamma_2^t, \ldots, \gamma_t^t)$. The optimal mean square estimator is more commonly

[1]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 random number with the same covariance of measurement noise simply reduces some mathematical burden.

known as Kalman filter. It is also useful to design the estimator error and error covariance as follows:

$$e_{t|t} \triangleq x_t - \hat{x}_{t|t} \quad (7)$$
$$P_{t|t} \triangleq \mathbb{E}[\, e_{t|t} e_{t|t}^T \mid \tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t, \bar{x}_0, P_0] \quad (8)$$

The estimate $\hat{x}_{t|t}$ is optimal in the sense that it minimizes the error covariance, i.e. given any estimator $\tilde{x}_{t|t} = f(\tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t)$ 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 $e_{t|t} \triangleq x_t - \hat{x}_{t|t}$ 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.

III. 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:

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

from which it follows that $\hat{x}_{t|t} = \hat{x}_{t|t}^t$ 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:

$$x_{k+1} = Ax_k + w_k \quad (9)$$
$$\tilde{y}_{k+1}^t = C_{k+1}^t x_{k+1} + \tilde{v}_{k+1}^t, \quad k = 0, \ldots, t-1 \quad (10)$$

where $C_k^t = \gamma_k^t C$, and the random variables $\tilde{v}_k^t = \gamma_k^t v_k + (1-\gamma_k^t) d_k$ are uncorrelated, zero mean white noise with covariance $R_k^t = \mathbb{E}[\tilde{v}_k^t (\tilde{v}_k^t)^T] = R$. For any fixed t this system can be seen as a linear time varying stochastic system with respect to time k, where the only time-varying element is the observation matrix C_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 $\hat{x}_{t|t} = \hat{x}_{t|t}^t$:

$$\hat{x}_{k|k}^t = A\hat{x}_{k-1|k-1}^t + \gamma_k^t K_k^t (\tilde{y}_k^t - CA\hat{x}_{k-1|k-1}^t) \quad (11)$$
$$K_k^t = P_{k|k-1}^t C^T (C P_{k|k-1}^t C^T + R)^{-1} \quad (12)$$
$$P_{k+1|k}^t = A P_{k|k-1}^t A^T + Q - \gamma_k^t A K_k^t C P_{k|k-1}^t A^T \quad (13)$$
$$\hat{x}_{0|0}^t = \bar{x}_0, \quad P_{1|0}^t = P_0 \quad (14)$$

where 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 and only if $\gamma_k^t = \gamma_k^{t-1}, \forall k \geq 1, \forall t-k \geq N$. 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 $t > N$:

$$\hat{x}_{k|k}^t = A\hat{x}_{k-1|k-1}^t + \gamma_k^t K_k^t (\tilde{y}_k^t - CA\hat{x}_{k-1|k-1}^t) \quad (15)$$
$$K_k^t = P_{k|k-1}^t C^T (C P_{k|k-1}^t C^T + R)^{-1} \quad (16)$$
$$P_{k+1|k}^t = A P_{k|k-1}^t A^T + Q - \gamma_k^t A K_k^t C P_{k|k-1}^t A^T \quad (17)$$
$$\hat{x}_{t-N|t-N}^t = \hat{x}_{t-N|t-N}^{t-1}, \quad (18)$$
$$P_{t-N+1|t-N}^t = P_{t-N+1|t-N}^{t-1} \quad (19)$$

where $k = t-N+1, \ldots, t$.

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:

$$\hat{x}_{k|k}^t = A\hat{x}_{k-1|k-1}^t + K_k^t (\tilde{y}_k^t - C_k^t A \hat{x}_{k-1|k-1}^t)$$
$$K_k^t = P_{k|k-1}^t C_k^{t^T} (C_k^t P_{k|k-1}^t C_k^{t^T} + R_k^t)^{-1}$$
$$P_{k+1|k}^t = A P_{k|k-1}^t A^T + Q - A K_k^t C_k^t P_{k|k-1}^t A^T$$
$$\hat{x}_{0|0}^t = \bar{x}_0, \quad P_{1|0}^t = P_0$$

whose derivation can be found in any standard text on stochastic control [1] [2]. By substituting $C_{k+1}^t = \gamma_{k+1}^t C$ and $R_{k+1}^t = R$ into the previous equations and after performing some simplifications we obtain the equivalent optimal estimator of Equations (11)-(14).

(b)($\Longrightarrow$) Let us consider $t > N$. If $\gamma_k^t = \gamma_k^{t-1}, \forall k \geq 1, \forall t-k \geq N$, 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 (18) and (19), and then iterate Equations (15)-(17) for the latest N observations.

($\Longleftarrow$) Using a contradiction argument suppose that there exist N for which estimator defined by Equations (15)-(19) is optimal. Now consider an arrival sequence for which $\gamma_1^t = 0$ for $t = 1, \ldots, N$ and $\gamma_1^{N+1} = 1$, and also $P_0 > 0$. Then $P_{2|1}^{N+1} < P_{2|1}^N$ and recursively it follows $P_{k+1|k}^{N+1} < P_{k+1|k}^N$ for all $k = 2, \ldots, N+1$. Therefore, the estimator using Equation (19) cannot be optimal, which concludes the theorem. ■

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. However there some differences that is important to remark. The first difference is that the optimal estimator under our framework jumps between an open loop estimate when the value stored in the buffer is not used and the error covariance increases ($\gamma_k^t = 0$), and a closed loop estimate when the observation measurement is used and the error covariance decreases ($\gamma_k^t = 1$). Therefore, the optimal estimator is strongly time-varying and stochastic. Differently, in standard Kalman filtering the error covariance $P_{t|t}$ and the optimal gain K_t converge to finite steady-state values, P_∞ and K_∞ respectively, as time progresses. Even more remarkably it is possible to show that using the steady-state optimal gain K_∞ it is possible to achieve the same steady-state error, thus not requiring any on-line matrix inversion.

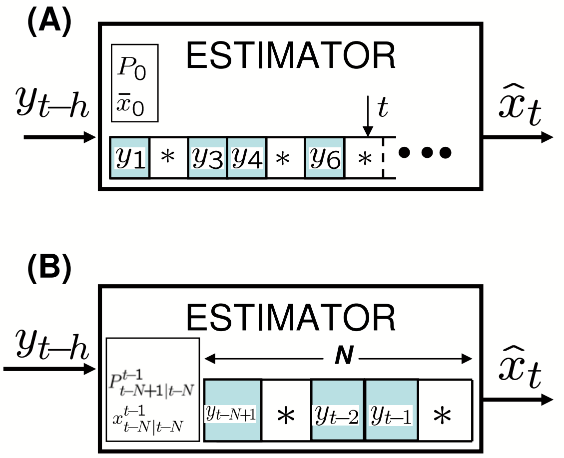

Fig. 3. Optimal Kalman estimator memory requirements for general arrival processes (A). Optimal Kalman estimator with finite memory buffer (B).

The second difference is that the standard Kalman filter does not need to store all past observations and to compute $\hat{x}_{t|t}$ starting from $k = 0$, but the optimal estimate can be computed incrementally by storing only the current observation y_t, the past state estimate $x_{t-1,t-1}$ and the past error covariance $P_{t,t-1}$. Differently, the optimal estimator subject to random packet delay requires the storing of all past packets and the inversion of up to t matrices at any time step t to calculate the optimal estimate, as shown in Theorem 1(b). The optimal estimator can be implemented incrementally according to Equations (15)-(19) using a buffer of finite length N only if all successfully received observations have a delay smaller than N time step, i.e. $\gamma_k^t = \gamma_k^{t-1}, \forall k \geq 1, \forall t-k \geq N$. 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 (see Fig. 3). Equivalently, this condition can be written in terms of the packet delay $\tau_k \in \{0, \dots, N-1, \infty\}, \forall k \geq 1$. This condition is rather common in DSNs since it is very difficult to guarantee correct delivery of all transmitted packets, while it is rather easy to implement mechanisms to drop packets that are too old.

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, 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. $\mathbb{P}[P_{t+1|t} > P_{max}]$ [3]. In this work we will consider only the expected error covariance $\mathbb{E}_\gamma[P_{t+1|t}]$. It has been shown in [4] that computing $\mathbb{E}_\gamma[P_{t+1|t}]$ analytically it is not possible even for a simple Bernoulli arrival process, and only upper and lower bounds can be obtained. Rather than extending those results by trying to bound performance of the time-varying optimal estimator, 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 $t \in \mathbb{N}, h = 0, \dots, N-1$. 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 it is not necessary to invert up to N matrices at any time step t, thus making it difficult to implement for on-line applications, and these gains can be computed off-line. Moreover, filters with constant gains are necessarily suboptimal, therefore their error covariance provide an upper bound for the error covariance of the true optimal minimum error covariance filter given by Equations (11)-(14). Therefore, in the next section we will study minimum error covariance filters with constant gains under stationary i.i.d. packet arrival processes.

IV. Optimal filtering 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:

$$\mathbb{P}[\tau_t \leq h] = \lambda_h \tag{20}$$

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

Equation (20) 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:

$$\lambda_{loss} \triangleq 1 - \sup\{\lambda_h | h \geq 0\} \tag{21}$$

The arrival process defined by Equation (20) 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:

$$\tau_{max} \triangleq \begin{cases} \min\{H | \lambda_H = \lambda_{H+1}\} & \textit{if } \exists H | \lambda_h = \lambda_H, \forall h \geq H \\ \infty & \textit{otherwise} \end{cases} \tag{22}$$

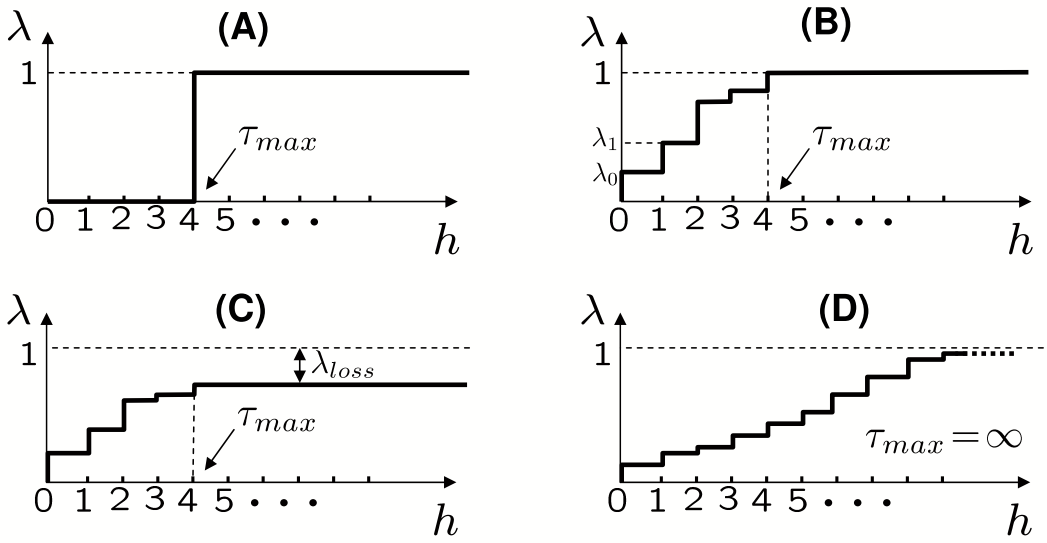

Fig. 4. Probability function of arrival process $\lambda_h = \mathbb{P}[\tau_k \leq h]$ for different scenarios: deterministic packet arrival with fixed delay (A); bounded random packet delay with no packet loss (B); bounded random packet delay with packet loss (C); unbounded random packet delay with no packet loss (D).

Fig. 4 shows some typical scenarios that can be modeled. 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 not delivered and the transmission channel is such that the packet is not delivered correctly with a probability ϵ. Simple calculations show that in this case $\lambda_h = 1 - \epsilon^h$.

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 $\mathbb{E}[\,||\tilde{e}_{t|t}||^2] \leq \mathrm{trace}(M)$.

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:

$$\tilde{x}^t_{t-k|t-k} = A\tilde{x}^t_{t-k-1|t-k-1} + \gamma^t_{t-k}K_k(\tilde{y}^t_{t-k} - CA\tilde{x}^t_{t-k-1|t-k-1}) \quad (23)$$

$$\tilde{x}^t_{t-N|t-N} = \tilde{x}^{t-1}_{t-N|t-N} \quad (24)$$

$$\tilde{x}^t_{-k|-k} = \bar{x}_0, \quad \gamma^t_{-k} = 0, \quad \tilde{y}^t_{-k} = 0 \quad (25)$$

where $k = N-1, \ldots, 0$ and the last line include some dummy variables necessary to initialize the estimator for $t = 1, \ldots, N$. Note that static-gain estimator structure is very similar to the optimal estimator structure given by Equation (11) as the estimate is corrected only if the observation has arrived, i.e. $\gamma^t_{t-k} = 1$, otherwise the open loop estimate is considered. However, differently from Equation (11), 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 $\hat{x}_{t|t}$ as it follows from Equations (11)-(12).

We also define the following variables that will be useful in analyzing the performance of the estimator:

$$\tilde{x}^t_{k+1|k} = A\tilde{x}^t_{k|k} \quad (26)$$

$$\tilde{e}^t_{k+1|k} = x_{k+1} - \tilde{x}^t_{k+1|k} \quad (27)$$

$$\widetilde{P}^t_{k+1|k} = \mathbb{E}[\tilde{e}^t_{k+1|k}\tilde{e}^{t^T}_{k+1|k} \,|\, \tilde{\mathbf{y}}_t, \boldsymbol{\gamma}_t] \quad (28)$$

$$\overline{P}^t_{k+1|k} = \mathbb{E}[\tilde{e}^t_{k+1|k}\tilde{e}^{t^T}_{k+1|k}] = \mathbb{E}[\widetilde{P}^t_{k+1|k}] \quad (29)$$

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

$$\widetilde{P}^t_{k+1|k} = A(I - \gamma^t_k K_{t-k}C)\widetilde{P}^t_{k|k-1}(I - \gamma^t_k K_{t-k}C)^T A^T + +Q + \gamma^t_k A K_{t-k} R K^T_{t-k} A^T \quad (30)$$

$$\overline{P}^t_{k+1|k} = \lambda_{t-k} A(I - K_{t-k}C)\overline{P}^t_{k|k-1}(I - K_{t-k}C)^T A^T + +(1-\lambda_{t-k})A\overline{P}^t_{k|k-1}A^T + Q + \lambda_{t-k}A^T K_{t-k} R K^T_{t-k} A^T \quad (31)$$

where we used independence of γ^t_k, v_k, w_k, and $\tilde{e}^t_{k|k-1}$, and we also used the fact that v_k and w_k are zero mean.

For ease of notation let us define the following operator:

$$\mathcal{L}_\lambda(K,P) = \lambda A(I-KC)P(I-KC)^T A^T + (1-\lambda)APA^T + Q + \lambda AKRK^T A^T \tag{32}$$

If we substitute $k = t-N$ into Equation (31) and noting that from Equation (24) follows that $\widetilde{P}^t_{t-N+1|t-N} = \widetilde{P}^{t-1}_{t-N+1|t-N}$ and $\overline{P}^t_{t-N+1|t-N} = \overline{P}^{t-1}_{t-N+1|t-N}$, we obtain:

$$\overline{P}^t_{t-N+2|t-N+1} = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, \overline{P}^{t-1}_{t-N+1|t-N}) \tag{33}$$

$$\overline{P}^t_{t-k+1|t-k} = \mathcal{L}_{\lambda_k}(K_k, \overline{P}^t_{t-k|t-k-1}) \tag{34}$$

where $k = N-2, \ldots, 0$. Note that Equation (33) and (34) define a set of linear deterministic equations for fixed λ_k and K_k. In particular, if we define $S_t = P^{t-1}_{t-N+1|t-N}$, then Equations (33) can be written as

$$S_{t+1} = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, S_t) \tag{35}$$

Since all matrices $\overline{P}^t_{t-k+1|t-k}, k = 0, \ldots, N-1$ 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 theorem describes these properties:

Theorem 2: Consider the operator $\mathcal{L}_\lambda(K,P)$ as defined in Equation (32). Assume also that $P \geq 0$, (A,C) is observable, $(A, Q^{1/2})$ is controllable, $R > 0$, and $0 \leq \lambda \leq 1$. Also consider the following operator:

$$\Phi_\lambda(P) = APA^T + Q - \lambda\, APC^T(CPC^T + R)^{-1}CPA^T \tag{36}$$

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 $\lambda_{min} = 1 - \frac{1}{\prod_i |\sigma_i^u|^2}$, $\lambda_{max} = 1 - \frac{1}{\max_i |\sigma_i^u|^2}$, and $|\sigma_i^u| \geq 1$ are the unstable eigenvalues of A. In particular $\lambda_c = \lambda_{min}$ if *rank*$(C) = 1$, and $\lambda_c = \lambda_{max}$ 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: $\lambda_c = \inf\{\lambda \,|\, \Psi_\lambda(Y,Z), 0 \leq Y \leq I, \text{for some} Z, Y \in \mathbb{R}^{n\times n}\}$, and $P^* = \operatorname{argmax}\{\textit{trace}(P) \,|\, \Theta_\lambda(P) \geq 0, P \geq 0\}$ where the matrices $\Psi_\lambda(Y,Z)$ and $\Theta_\lambda(P)$ are given in Equations (37) and (38):

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

Proof: Most of these statements can be found in [5] or can be derived along the same 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 (36) 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 controllable [5].

(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 [5], thus it is not reported here.

(i) The proof for existence of a critical probability λ_c was given in [5] and it is based on observability of (A,C) and monotonicity of $\Phi_\lambda(P)$ with respect to λ. The proof for $\lambda_c = \lambda_{min}$ when *rank*$(C) = 1$ can be found in [6][7] although it was not explicitly derived for the operator Φ_λ. The proof for $\lambda_c = \lambda_{max}$ when C is square and invertible was first proved in [8].

(j) The proof can be found in [5].

(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

$$\Psi_\lambda(Y,Z) = \begin{bmatrix} Y & \sqrt{\lambda}(YA+ZC) & \sqrt{1-\lambda}YA \\ \sqrt{\lambda}(A'Y+C'Z') & Y & 0 \\ \sqrt{1-\lambda}A'Y & 0 & Y \end{bmatrix} \quad (37)$$

$$\Theta_\lambda(P) = \begin{bmatrix} APA'-P & \sqrt{\lambda}APC' \\ \sqrt{\lambda}CPA' & CPC'+R \end{bmatrix} \quad (38)$$

$A_c = A(I-\lambda KC)$. 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 $X_0 = x_0x_0^T$, i.e. the system A_c is strictly stable. ■

The previous theorem provides all tools necessary to analyze and design the optimal estimator with static gains. In particular, fact (g) indicates that the static gain K^* that minimizes the steady state error covariance P^* can be derived from the unique fixed point of the nonlinear operator Φ_λ, where $K^* = K_{P^*}$. If the optimal gain K^* is used, then the average error covariance converges to the P^* regardless of the initial conditions $(P_0, \bar{x}_0)$, as stated by 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 asymptotically strictly stable, therefore estimator stability reduces to existence of steady state error covariance.

The following theorem shows how compute the optimal estimator with static gains.

Theorem 3: Let us consider the stochastic linear system given in Equations (1)-(2), where (A,C) is observable, $(A,Q^{1/2})$ is controllable, and $R>0$. Also consider the arrival process defined by Equations (20)-(22), and the set of estimators with constant gains $\{K_k\}_{k=0}^N$ defined in Equations (23)-(25). If A is not strictly stable and $\lambda_{loss} \geq 1-\lambda_c$, where λ_c is defined in Theorem 2(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:

$$K_k^N = V_k^N C^T(CV_k^NC^T+R)^{-1}, \ k=0,..,N \quad (39)$$

$$V_{N-1}^N = \Phi_{\lambda_{N-1}}(V_{N-1}^N) \quad (40)$$

$$V_k^N = \Phi_{\lambda_k}(V_{k+1}^N), \quad k=N-1,\ldots,0 \quad (41)$$

Also consider $\overline{P}_{k+1|k}^t$ as defined in Equation (29), 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 following equations exist:

$$T_N^N = \mathcal{L}_{\lambda_N}(K_N, T_N^N) \quad (42)$$

$$T_k^N = \mathcal{L}_{\lambda_k}(K_k, T_{k+1}^N), \quad k=N-1,\ldots,0 \quad (43)$$

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 $\tau_{max} < \infty$, 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 A is not strictly stable and $\lambda_{loss} \geq 1-\lambda_c$. Suppose one 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 $\overline{P}_{t+1|t}^t = A\overline{P}_{t|t}^tA^T + Q$ also $\overline{P}_{t+1|t}^t$ must be bounded for all t. From Equations (33) and (34) 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 (35), from Theorem 2(g) follows that $S^* = \mathcal{L}_{\lambda_{N-1}}(K_{N-1}, S^*)$ has a solution. From Theorem 2(h) follows that also $P^* = \Phi_{\lambda_{N-1}}(P^*)$ has a solution. However, by hypothesis $\lambda_{N-1} \leq \sup\{\lambda_h \,|\, h \geq 0\} = 1-\lambda_{loss} \leq \lambda_c$. Consequently, according to Theorem 2(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 2(h) it follows that Equations (39)-(41) are well defined and have a solution. From Theorem 2(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 Theorem 2(h) it follows that $V_{N-1}^N \leq T_{N-1}^N$. From Theorem 2(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 Theorem 2(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 Theorem 2(e) we have that $V_{\tau_{max}}^{\tau_{max}} = V_{\tau_{max}}^N$. According to Equation (41) 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 (40) 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 static gains exists if and only if the optimal estimator with static gains exists, therefore the optimal estimator design implicitly solves the problem of finding 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. This is a remarkable result since it implies that stability of estimators does not depend on the packet delay τ_{max} as long as most most of the packets eventually arrive. Another important result is that 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. 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.

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

V. Numerical Examples

In this sections we consider some dynamical systems and we compute the estimator error covariances as shown in the previous section. Let us consider the following probability function of packet delay:

$$\begin{aligned} \lambda_h &= 0.05h, \quad h = 0, \ldots, 15 \\ &= 0.75, \quad h > 15 \end{aligned}$$

which is depicted in Figure 5.

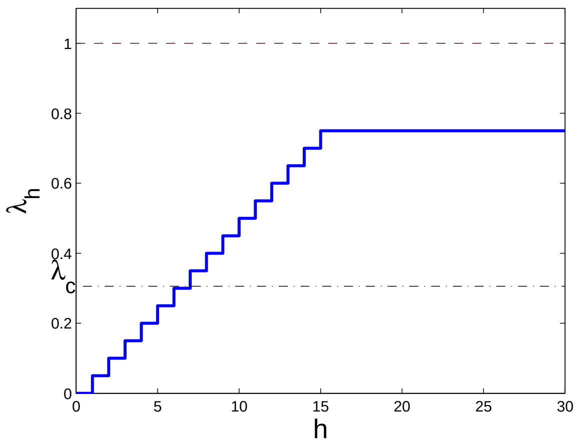

Fig. 5. Probability function of packet delay.

Let us consider the following discrete time system:

$$A = \begin{bmatrix} 1 & 0.1 \\ 0 & 0.8 \end{bmatrix}, \; C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T, R = 1, \; Q = \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 1 \end{bmatrix} \tag{44}$$

which corresponds to the digitalization of a continuous time system with one stable pole and a pole in the origin. This is a dynamical model of an electric motor, for example. The critical probability for this system is $\lambda_c = 0$. Therefore, according to the previous analysis the estimator is stable if and only if the size of the buffer of the estimator is greater than one. The trace of the covariance of the estimator error with constant gains V_0^N is shown in the left panel of Figure 6.

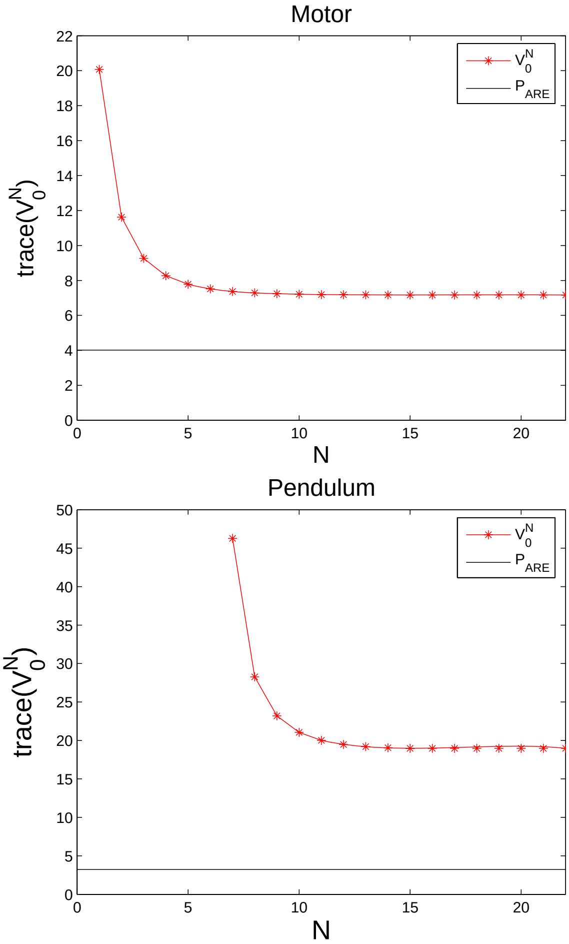

Fig. 6. Trace of the steady state error covariance for the optimal estimator with constant gains (V_0^N) for a discrete time electric motor model (*left*) and a pendulum (*right*). The horizontal line P_{ARE} correspond to the trace of the error covariance in the ideal scenario with no delay and no packet loss, i.e. $\lambda_h = 1$ for all h.

Let us consider the following discrete time system:

$$A = \begin{bmatrix} 1.2 & 0.1 \\ 0 & 0.8 \end{bmatrix}, \; C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T, R = 1, \; Q = \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 1 \end{bmatrix} \tag{45}$$

which corresponds to the digitalization of a continuous time system with one stable pole and one unstable pole. This is a dynamical model of an inverted pendulum, for example. The critical probability for this system is $\lambda_c = 1 - 1/1.2^2 = 0.3056$. According to Theorem 3 the estimator is stable if and only if $N \geq 7$, in fact $\lambda_6 = 0.3 < \lambda_c$ and $\lambda_7 = 0.35 > \lambda_c$. The trace of the covariance of the estimator error with constant gains, V_N is shown in the right panel of Figure 6.

VI. CONCLUSIONS

This numerical examples show how the tools developed in this paper can be readily used to estimate performance of estimator. In particular they can be used to evaluate the tradeoff between performance (the error covariance) and the estimator complexity (the buffer length) and the hardware complexity (the smart sensor). In particular the knowledge of the packet arrival statics can be used to find the optimal static gains $\{K_k^N\}_{k=0}^N$ and thus improving performance. Very importantly, the ability to quantitatively compute the estimator error covariance, can be used to compare different communication protocols that can be hard to compare otherwise. In fact, it would be rather difficult to compare a communication protocol with a small packet loss but overall larger packet delay relative to another communication protocol with a larger packet loss but a small average packet delay. This is particularly important from a technological point of view since most of today's design principles for communication protocols focus on guaranteeing a maximum delay for all packets, while in this work we have shown that an unstable system can be observed effectively even if a fraction of packet is lost. Future work will focus on applying the tools developed in this paper in current communication protocols for control application such as real-time ethernet, CAN, ATMs.

REFERENCES

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

[2] 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.

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

[4] 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.

[5] B. Sinopoli, C. Sharp, S. Schaffert, L. Schenato, 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.

[6] 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.

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

[8] 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.