2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Kalman Filtering with Intermittent Observations: Bounds on the Error Covariance Distribution

Eduardo Rohr, Damián Marelli and Minyue Fu

***Abstract*— When measurements are subject to random losses, the covariance of the estimation error of a state estimator becomes a random variable. In this paper we present bounds on the cumulative distribution function of the covariance of the estimation error for a discrete time linear system. We also show that the bounds can be arbitrarily tight if sufficient computational power is available. Numerical simulations show that the proposed method provides tighter bounds than the ones available in the literature.**

I. Introduction

The fast development of network (particularly wireless) technology has encouraged its use in control and signal processing applications. From the control systems' perspective, this new technology has imposed new challenges concerning how to deal with the effects of quantisation, delays and loss of packets, leading to the development of a new networked control theory [1]. The study of state estimators, when measurements are subject to random delays and losses, finds applications in both control and signal processing. Most estimators are based on the well-known Kalman filter [2]. In order to cope with network induced effects, the standard Kalman filter paradigm needs to undergo certain modifications.

In the case of missing measurements, the update equation of the Kalman filter depends on whether a measurement arrives or not. When a measurement is available, the filter performs the standard update equation. On the other hand, if the measurement is missing, it must produce open loop estimation, which as pointed out in [3], can be interpreted as the standard update equation when the measurement noise is infinite. If the measurement arrival event is modeled as a binary random variable, the estimator's error covariance (EC) becomes a random matrix. Studying the statistical properties of the EC is important to assess the estimator's performance. Additionally, a clear understanding of how the system's parameters and network delivery rates affect the EC permits a better system design, and the trade-off between conflicting interests must be evaluated.

Studies on the computation of the expected error covariance (EEC) can be dated back at least to [4], where upper and lower bounds for the EEC were obtained using a constant gain on the estimator. In [3], the same upper bound was derived as the limiting value of a recursive equation that computes a weighted average of the next possible error covariances. A similar result which allows partial observation losses was presented in [5]. In [6], it is shown that a system in which the sensor transmits state estimates instead of raw measurements will provide a better error covariance. Most of the available research work is concerned with the expected value of the EC, neglecting higher order statistics. In [7], the present authors introduce tighter lower and upper bounds for the EEC as well introduce the discussion on second order moments. The critical measurement arrival probability for boundedness of general order moments was addressed in [8], where non-degenerate systems are considered.

The study of the complete probability distribution function of the EC is a current research topic. In [9], [10], [11], the authors discuss the existence of a stationary distribution of the EC for Kalman filters with intermittent observations. Using Random Matrix Theory, the authors of [12] present the probability density function for stable systems. In [13], [14], the authors assume that the sensors have the ability to send multiple measurements in a packet. As a consequence, whenever a packet is received, the Kalman filter produces a bounded error covariance. This was used in [13] to provide an upper bound for the CDF for unstable systems and later on extended to any system in [14]. A geometric approach was used by [15] to derive the CDF for a restrictive class of systems. In the present paper, we present upper and lower bounds for the CDF for any system and we assume that only the most recent measurement is sent at each sampling time.

This paper investigates the behavior of the Kalman filter for discrete-time linear systems whose output is intermittently sampled. To this end, we model the measurement arrival event as a binary random variable. In order to keep the approach as general as possible, no assumption is made about the packet dropping model. We derive the solution for the two most popular network packet dropout models: independent and identically distributed (i.i.d.) and Gilbert-Elliott, although the proposed method can be easily extended to more complex models. We assume that the sensor sends only the most recent measurement at each sampling time. The main contribution of this paper is the introduction of a method to obtain lower and upper bounds for the cumulative distribution function (CDF). These bounds can be made arbitrarily tight, at the expense of increased computational complexity. We also present numerical examples in which the performance of the proposed method is compared to existing ones. In particular, we show that the bounds presented here are tighter than the ones in [14].

The authors are with the School of Electrical Engineering and Computer Science, University of Newcastle, University Drive, Callaghan NSW 2308, Australia. E-mail: `eduardo.rohr@uon.edu.au`

II. Problem Statement

Consider the discrete-time linear system:

$$\begin{cases} x_{t+1} &= Ax_t + w_t \\ y_t &= Cx_t + v_t \end{cases} \quad (1)$$

where the state vector $x_t \in \mathbb{R}^n$ has initial condition $x_0 \sim N(0, P_0)$, $y_t \in \mathbb{R}^p$ is the measurement, $w_t \sim N(0, Q)$ is the process noise and $v_t \sim N(0, R)$ is the measurement noise. The goal of the Kalman filter is to obtain the best estimate $\hat{x}_t$ of the state x_t, in the sense that the trace of the a priori covariance matrix P_t of the error $\tilde{x}_t = x_t - \hat{x}_t$ is minimized.

We assume that the measurements y_t are sent to the Kalman estimator through a network subject to random packet losses. The scheme proposed in [6] can be used to deal with delayed measurements. Hence, without loss of generality, we assume that there is no delay in the transmission. Let γ_t be a binary random variable describing the arrival of a measurement at time t. We define that $\gamma_t = 1$ when y_t was received at the estimator and $\gamma_t = 0$ otherwise.

Notice that the update equation of the error covariance is now dependent on the availability of the measurements. When a measurement is available, both measurement and time updates are performed. When a measurement is not available, only the time update can be computed. The update equation of P_t can be written as follows:

$$P_{t+1} = \begin{cases} \Phi_1(P_t), & \gamma_t = 1 \\ \Phi_0(P_t), & \gamma_t = 0 \end{cases} \quad (2)$$

with

$$\begin{aligned} \Phi_1(P_t) &= AP_tA' + Q + \\ &\quad -AP_tC'(CP_tC' + R)^{-1}CP_tA' \quad (3) \\ \Phi_0(P_t) &= AP_tA' + Q. \quad (4) \end{aligned}$$

Refer to [3] for the state update equations.

We point out that when all the measurements are available and the Kalman filter reaches its steady state, the EC is given by the solution of the following algebraic Riccati equation

$$\underline{P} = A\underline{P}A' + Q - A\underline{P}C'(C\underline{P}C' + R)^{-1}C\underline{P}A'. \quad (5)$$

We use the following notation. For given $T \in \mathbb{N}$ and $0 \leq m \leq 2^T - 1$, the symbol S_m^T denotes the binary sequence of length T formed by the binary representation of m. We also use $S_m^T(i)$, $i = 1, \cdots, T$, to denote the i-th entry of the sequence, i.e.,

$$S_m^T = \{S_m^T(1),\ S_m^T(2), \ldots,\ S_m^T(T)\} \quad (6)$$

and

$$m = \sum_{k=1}^{T} 2^{k-1} S_m^T(k). \quad (7)$$

For a given sequence S_m^T, and a matrix $P \in \mathbb{R}^{n\times n}$, we define the map

$$\phi(P, S_m^T) = \Phi_{S_m^T(T)} \circ \Phi_{S_m^T(T-1)} \circ \ldots \Phi_{S_m^T(1)}(P) \quad (8)$$

where $\circ$ denotes the composition of functions (i.e. $f \circ g(x) = f(g(x))$).

Also, for given $T > 0 \in \mathbb{N}$, Γ^T denotes the (random) binary sequence containing the availability of measurements between time 0 and $T-1$, i.e.

$$\Gamma^T = \{\gamma_{T-1},\ \gamma_{T-2},\ \ldots,\ \gamma_0\}. \quad (9)$$

Notice that if m is chosen so that

$$S_m^T = \Gamma^T, \quad (10)$$

then $\phi(P_0, S_m^T)$ updates P_0 according to the measurement arrivals in the last T sampling times, i.e.,

$$P_T = \phi(P_0, S_m^T) = \Phi_{\gamma_{T-1}} \circ \Phi_{\gamma_{T-1}} \circ \ldots \Phi_{\gamma_0}(P_0). \quad (11)$$

The next lemma states the monotonicity of $\phi(\cdot,\cdot)$ with respect to its first argument.

Lemma 2.1: Consider the function $\phi(\cdot,\cdot)$ defined in (8). If $X \leq Y$, then

$$\phi(X, S_m^T) \leq \phi(Y, S_m^T) \quad (12)$$

for any sequence S_m^T.

Proof: The proof follows the argument in [3]. Suppose $X \leq Y$. Since $\Phi_0(\cdot)$ is affine, we have that

$$\Phi_0(X) \leq \Phi_0(Y). \quad (13)$$

Recall that the optimal Kalman gain is

$$\begin{aligned} K_X &= \arg\min_K (A + KC)X(A + KC)' + Q + KRK' \\ &= -AXC'(CXC' + R)^{-1}. \quad (14) \end{aligned}$$

We have that

$$\begin{aligned} &\Phi_1(X) \\ =\ & (A + K_XC)X(A + K_XC)' + Q + K_XRK_X' \\ \leq\ & (A + K_YC)X(A + K_YC)' + Q + K_YRK_Y' \\ \leq\ & (A + K_YC)Y(A + K_YC)' + Q + K_YRK_Y' \\ =\ & \Phi_1(Y). \quad (15) \end{aligned}$$

The proof is completed by observing that $\phi(\cdot,\cdot)$ is formed by the composition of $\Phi_0(\cdot)$ and $\Phi_1(\cdot)$. ■

We use $\mathcal{P}(S_m^T)$ to denote the probability that the sequence of available measurements in the last T sampling times equals S_m^T, i.e.,

$$\mathcal{P}(S_m^T) = \mathcal{P}(\Gamma^T = S_m^T) \quad (16)$$

We also use I_n to denote the identity matrix of size $n \times n$ and, $X > Y$, $(X \geq Y)$ to indicate that $X - Y$ is a positive definite (positive semi-definite) matrix.

III. Main Result

Our goal is to study the CDF of the EC P_T, in the limit when T tends to infinity. We define

$$F(x, P_0) \triangleq \lim_{T\to\infty} F^T(x, P_0) \quad (17)$$

with

$$\begin{aligned} F^T(x, P_0) &= \mathcal{P}\left(\phi(P_0, \Gamma^T) \leq xI_n\right) \quad (18) \\ &= \sum_{m=0}^{2^T-1} \mathcal{P}\left(S_m^T\right) H\left(xI_n - \phi(P_0, S_m^T)\right) \quad (19) \end{aligned}$$

where $H(.)$ is a matrix version of the Heaviside step function, defined for $X \in \mathbb{R}^{n\times n}$ as

$$H(X) = \begin{cases} 1, & X \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

The question of whether $F(x, P_0)$ is independent of the initial EC value P_0 is an open problem [9]. Nevertheless, we can state bounds on all possible values of $F(x, P_0)$. To do so, notice that, a direct consequence of Lemma 2.1 is that, if $X \geq Y$, then

$$F(x, X) \leq F(x, Y). \tag{21}$$

Define,

$$\overline{F}(x) \triangleq \lim_{T\to\infty} F^T(x, \underline{P}) \tag{22}$$

$$\underline{F}(x) \triangleq \lim_{T\to\infty} F^T(x, \infty) \tag{23}$$

where

$$F^T(x, \infty) = \lim_{\alpha\to\infty} F^T(x, \alpha I_n). \tag{24}$$

Then, assuming that $P_0 \geq \underline{P}$, we have that

$$\overline{F}(x) \geq F(x, P_0) \geq \underline{F}(x). \tag{25}$$

Notice that this assumption is reasonable, since in this context $P_T \geq \underline{P}$ for some T and we can always shift the time indexes in order to have $P_0 \geq \underline{P}$.

Equation (25) states bounds on all possible values of $F(x, P_0)$. However, their computation is impractical. To go around this, we develop alternative bounds, which are looser than those in (25), but are suitable for numeric computation. The basic idea is to derive upper and lower bounds on the error covariance, for each arrival sequence that can be observed in the last T sampling instants, and associate each bound with the probability of the observed sequence.

Define the S_m^T-dependant matrix

$$\phi(\infty, S_m^T) \triangleq \lim_{\alpha\to\infty} \phi(\alpha I_n, S_m^T). \tag{26}$$

Notice that for any P_0 we can always find an α such that $\alpha I_n \geq P_0$. Then, from the monotonicity of $\phi(\cdot, S_m^T)$ (Lemma 2.1), it follows that

$$P_T = \phi(P_0, \Gamma^T) \leq \phi(\infty, \Gamma^T). \tag{27}$$

The following lemma provides an expression for computing $\phi(\infty, S_m^T)$.

Lemma 3.1: For a given T, let $0 \leq t_1, t_2, \cdots, t_I \leq T-1$ denote the time indexes where $\gamma_{t_i} = 1$, $i = 1, \cdots, I$. Define

$$O = \begin{bmatrix} CA^{t_1} \\ CA^{t_2} \\ \vdots \\ CA^{t_I} \end{bmatrix}, \quad \Sigma_Q = \begin{bmatrix} \sum_{j=0}^{t_1-1} CA^j QA'^{T-t_1+j} \\ \sum_{j=0}^{t_2-1} CA^j QA'^{T-t_2+j} \\ \vdots \\ \sum_{j=0}^{t_I-1} CA^j QA'^{T-t_I+j} \end{bmatrix}', \tag{28}$$

and the matrix $\Sigma_V \in \mathbb{R}^{pI\times pI}$, whose (i,j)-th submatrix $[\Sigma_V]_{i,j} \in \mathbb{R}^{p\times p}$ is given by

$$[\Sigma_V]_{i,j} = \sum_{k=1}^{\min\{t_i,t_j\}} CA^{t_i-k} QA'^{t_j-k} C' + R\delta(i,j) \tag{29}$$

where

$$\delta(i,j) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases} \tag{30}$$

Then, for any sequence S_m^T, we have that

$$\phi(\infty, S_m^T) = \begin{cases} \infty I_n, & \text{if } O(S_m^T) \text{ is not FCR} \\ \overline{\phi}(S_m^T), & \text{if } O(S_m^T) \text{ is FCR} \end{cases} \tag{31}$$

where FCR stands for Full Column Rank and

$$\begin{aligned} \overline{\phi}(S_m^T) = A^T \left(O'\Sigma_V^{-1}O\right)^{-1} A'^T + \textstyle\sum_{j=0}^{T-1} A^j QA'^j + \\ -A^T(\Sigma_V^{-\frac{1}{2}}O)^\dagger \Sigma_V^{-\frac{1}{2}}\Sigma_Q' - \Sigma_Q\Sigma_V^{-\frac{1}{2}}(\Sigma_V^{-\frac{1}{2}}O)'^\dagger A'^T + \\ -\Sigma_Q\left(\Sigma_V^{-1} - \Sigma_V^{-1}O(O'\Sigma_V^{-1}O)^{-1}O'\Sigma_V^{-1}\right)\Sigma_Q' \end{aligned} \tag{32}$$

with $(\Sigma_V^{-\frac{1}{2}}O)^\dagger$ denoting the Moore-Penrose pseudo-inverse of $\Sigma_V^{-\frac{1}{2}}O$ [16].

Proof: Let Y_T be the vector formed by the available measurements

$$\begin{aligned} Y_T &= \begin{bmatrix} y'_{t_1} & y'_{t_2} & \cdots & y'_{t_I} \end{bmatrix}' && (33) \\ &= Ox_0 + V_T, && (34) \end{aligned}$$

where

$$V_T = \begin{bmatrix} \sum_{j=1}^{t_1-1} CA^{t_1-j} w_{j-1} + v_{t_1} \\ \sum_{j=1}^{t_2-1} CA^{t_2-j} w_{j-1} + v_{t_2} \\ \vdots \\ \sum_{j=1}^{t_I-1} CA^{t_I-j} w_{j-1} + v_{t_I} \end{bmatrix}. \tag{35}$$

From the model (1), it follows that

$$\begin{bmatrix} x_T \\ Y_T \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_x & \Sigma_{xY} \\ \Sigma'_{xY} & \Sigma_Y \end{bmatrix}\right) \tag{36}$$

where

$$\Sigma_x = A^T P_0 A'^T + \sum_{j=0}^{T-1} A^j QA'^j \tag{37}$$

$$\Sigma_{xY} = A^T P_0 O' + \Sigma_Q \tag{38}$$

$$\Sigma_Y = OP_0O' + \Sigma_V. \tag{39}$$

It follows from [2, pp. 39] that the covariance of the estimation error is given by

$$\phi(P_0, S_m^T) = \Sigma_x - \Sigma_{xY}\Sigma_Y^{-1}\Sigma'_{xY}. \tag{40}$$

Substituting (37)-(39) in (40), we have

$$\begin{aligned} \phi(P_0, S_m^T) = \textstyle\sum_{j=0}^{T-1} A^j QA'^j - \Sigma_Q\left(OP_0O' + \Sigma_V\right)^{-1}\Sigma_Q' + \\ A^T\left(P_0 - P_0O'\left(OP_0O' + \Sigma_V\right)^{-1}OP_0\right)A'^T + \\ -A^T P_0 O'\left(OP_0O' + \Sigma_V\right)^{-1}\Sigma_Q' + \\ -\Sigma_Q\left(OP_0O' + \Sigma_V\right)^{-1}OP_0A'^T. \end{aligned} \tag{41}$$

Hence, from (26), we have that

$$\phi(\infty, S_m^T) = P_{T,1} + P_{T,2} + P_{T,3} + P'_{T,3} + P_{T,4}, \quad (42)$$

with

$$\begin{aligned} P_{T,1} &= \lim_{\alpha\to\infty} A^T \left(\alpha I_n - \alpha^2 O' \left(\alpha O O' + \Sigma_V\right)^{-1} O\right) A'^T \\ P_{T,2} &= \sum_{j=0}^{T-1} A^j Q A'^j \\ P_{T,3} &= -\lim_{\alpha\to\infty} \alpha A^T O' \left(\alpha O O' + \Sigma_V\right)^{-1} \Sigma'_Q \\ P_{T,4} &= -\lim_{\alpha\to\infty} \Sigma_Q \left(\alpha O O' + \Sigma_V\right)^{-1} \Sigma'_Q. \end{aligned}$$

Using the matrix inversion lemma, we have that

$$\begin{aligned} P_{T,1} &= A^T \lim_{\alpha\to\infty} \left(\alpha^{-1} I_n + O'\Sigma_V^{-1} O\right)^{-1} A'^T \quad (43) \\ &= A^T \left(O'\Sigma_V^{-1} O\right)^{-1} A'^T \quad (44) \end{aligned}$$

and

$$P_{T,4} = \Sigma_Q \left(\Sigma_V^{-1} - \Sigma_V^{-1} O \left(O'\Sigma_V^{-1} O\right)^{-1} O'\Sigma_V^{-1}\right) \Sigma'_Q. \quad (45)$$

By making $X = \Sigma_V^{-\frac{1}{2}} O$, $P_{T,3}$ can be written as

$$\begin{aligned} P_{T,3} &= -\lim_{\alpha\to\infty} A^T O' \left(O O' + \Sigma_V \alpha^{-1}\right)^{-1} \Sigma'_Q \quad (46) \\ &= -A^T \lim_{\alpha\to\infty} X' \left(X X' + \alpha^{-1} I_n\right)^{-1} X \Sigma'_Q . (47) \end{aligned}$$

From [16, pp. 115], it follows that $\lim_{\alpha\to\infty} X' \left(X X' + \alpha^{-1} I_n\right) = X^\dagger$, for any matrix X.

$$P_{T,3} = -A^T \left(\Sigma_V^{-\frac{1}{2}} O\right)^\dagger \Sigma_V^{-\frac{1}{2}} \Sigma'_Q. \quad (48)$$

The result follows by substituting (44), (48) and (45) in (42). ■

The next theorem is the main result of this paper. It states lower and upper bounds for the CDF $F(x, P_0)$, which can be made arbitrarily tight at the expense of additional computation efforts.

Theorem 3.1: Define

$$\begin{aligned} \underline{F}^T(x) &\triangleq \sum_{m=0}^{2^T-1} \mathcal{P}(S_m^T) H(x I_n - \phi(\infty, S_m^T)) \quad (49) \\ \overline{F}^T(x) &\triangleq \sum_{m=0}^{2^T-1} \mathcal{P}(S_m^T) H(x I_n - \phi(\underline{P}, S_m^T)). \quad (50) \end{aligned}$$

The following properties hold:

$$\begin{aligned} \lim_{T\to\infty} \underline{F}^T(x) &= \underline{F}(x) \quad (51) \\ \lim_{T\to\infty} \overline{F}^T(x) &= \overline{F}(x). \quad (52) \end{aligned}$$

Moreover, these bounds become monotonically tighter as T is increased, i.e.,

$$\begin{aligned} \underline{F}^{T+1}(x) &\geq \underline{F}^T(x) \quad (53) \\ \overline{F}^{T+1}(x) &\leq \overline{F}^T(x). \quad (54) \end{aligned}$$

Proof: Notice that (51) and (52) follow directly from the definitions in (22) and (23). Hence, we only need to show the monotonicity of $\overline{F}^T(x)$ and $\underline{F}^T(x)$ in T.

Notice that when the length of the sequence is increased by one, with the binary variable γ included at the beginning of the sequence, we have

$$\phi(\underline{P}, \{S_m^T, \gamma\}) = \phi(\Phi_\gamma(\underline{P}), S_m^T) \geq \phi(\underline{P}, S_m^T) \quad (55)$$

where the equality holds only for $\gamma = 1$, since $\Phi_1(\underline{P}) = \underline{P}$ and $\Phi_0(\underline{P}) > \underline{P}$. This shows that the step functions in (50) can only by shifted to the right when the sequence length is increased by 1. Hence (54) follows.

Now, from the definition of $\phi(\infty, S_m^T)$, we have that

$$\phi(\infty, \{S_m^T, 0\}) = \phi(\infty, S_m^T). \quad (56)$$

Also, since y_0 and x_T are correlated, the estimate $\hat{x}_T$ given the sequence S_m^T have greater error covariance than the estimate $\hat{x}_T$ given $S_m^T,\ y_0$, that is

$$\phi(\infty, \{S_m^T, 1\}) \leq \phi(\infty, S_m^T). \quad (57)$$

Combining (56) and (57), we have

$$\phi(\infty, \{S_m^T, \gamma\}) \leq \phi(\infty, S_m^T), \quad (58)$$

which shows that the step functions in (49) can only be shifted to the left, as we augment T, showing (53). ■

Corollary 3.1: For any integer $T > 0$, we have the following bounds on the CDF

$$\begin{aligned} \underline{F}^T(x) &\leq \underline{F}(x) \quad (59) \\ \overline{F}^T(x) &\geq \overline{F}(x). \quad (60) \end{aligned}$$

Proof: The relation in (59) follows from (51) and (53), while (60) follows from (52) and (54). ■

IV. The I.I.D. and Gilbert-Elliott Network Models

Theorem 3.1 requires the computation of the probability $\mathcal{P}(S_m^T)$ to observe a given sequence of available measurements in the last T sampling times. The computation of such probability is dependent on the network model considered. We derive an expression for this probability for two popular network models.

The i.i.d. model assumes that γ_t and γ_k are two independent variables whenever $t \neq k$. Let the probability that a given measurement is available be given by $\mathcal{P}(\gamma_t = 1) = \lambda$. We have

$$\mathcal{P}(S_m^T) = \lambda^K (1-\lambda)^{T-K} \quad (61)$$

with K being the number of available measurements, i.e.,

$$K = \sum_{k=1}^{T} S_m^T(k). \quad (62)$$

The Gilbert-Elliott packet dropping model uses a two state Markov chain to describe the probability that a given measurement is available, according to the availability of the previous one. I.e.,

$$\begin{bmatrix} \mathcal{P}(\gamma_{k+1} = 0) \\ \mathcal{P}(\gamma_{k+1} = 1) \end{bmatrix} = M \begin{bmatrix} \mathcal{P}(\gamma_k = 0) \\ \mathcal{P}(\gamma_k = 1) \end{bmatrix} \quad (63)$$

with

$$M = \begin{bmatrix} \alpha & 1-\beta \\ 1-\alpha & \beta \end{bmatrix}$$

Let $[\pi_0 \ \ \pi_1]'$ describe the steady-state distribution of the resulting Markov chain, i.e.,

$$\begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix} = M \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix}.$$

We point out that this distribution is given by the normalized eigenvector associated with the eigenvalue 1 of the probability transition matrix M. The next lemma gives a formula to compute the probability to observe a given sequence.

Lemma 4.1: Suppose that in a networked control system the measurements are randomly dropped according to (63) and that the initial probability distribution is the steady-state distribution of the Markov chain. The probability to observe a sequence S_m^T from time 0 to $T-1$ is given by

$$\begin{aligned} \mathcal{P}(S_m^T) &= \big(\pi_0 + S_m^T(1)(\pi_1 - \pi_0))\big) \\ &\prod_{n=2}^{T} \begin{bmatrix} 1-S_m^T(k) \\ S_m^T(k) \end{bmatrix}' M \begin{bmatrix} 1-S_m^T(k-1) \\ S_m^T(k-1) \end{bmatrix}. \end{aligned} \tag{64}$$

Proof: Define

$$g_n \triangleq \begin{cases} 0, & S_m^T(n) \neq \Gamma^T(n) \\ 1, & S_m^T(n) = \Gamma^T(n). \end{cases} \tag{65}$$

We have that

$$\begin{aligned} \mathcal{P}(S_m^T) &= \mathcal{P}(\Gamma^T = S_m^T) \\ &= \mathcal{P}(g_1 = 1 \cap g_2 = 1 \cap \ldots \cap g_T = 1) \\ &= \mathcal{P}(g_1 = 1) \prod_{n=2}^{T} \mathcal{P}(g_n = 1 | g_{n-1} = 1). \end{aligned} \tag{66}$$

Notice that

$$\begin{aligned} \mathcal{P}(g_1 = 1) &= \begin{cases} \pi_0, & S_m^T(1) = 0 \\ \pi_1, & S_m^T(1) = 1 \end{cases} \\ &= \pi_0 + S_m^T(1)(\pi_1 - \pi_0) \end{aligned} \tag{67}$$

and

$$\mathcal{P}(g_n = 1 | g_{n-1} = 1) =$$

$$\begin{cases} \alpha, & S_m^T(n) = 0, S_m^T(n-1) = 0 \\ 1-\alpha, & S_m^T(n) = 1, S_m^T(n-1) = 0 \\ \beta, & S_m^T(n) = 1, S_m^T(n-1) = 1 \\ 1-\beta, & S_m^T(n) = 0, S_m^T(n-1) = 1 \end{cases} = \tag{68}$$

$$\begin{bmatrix} 1-S_m^T(n) \\ S_m^T(n) \end{bmatrix}' M \begin{bmatrix} 1-S_m^T(n-1) \\ S_m^T(n-1) \end{bmatrix}.$$

Substituting (67) and (68) in (66), we obtain (64). ■

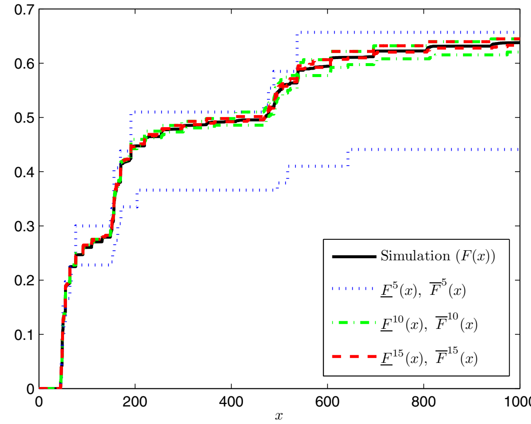

Fig. 1. Example A: Monotonicity of the Bounds.

V. Numerical Examples

In this section we present three simulation results aiming to illustrate the proposed method, and to compare its performance with the method in [14], when applicable. In each case, an estimation of the true CDF $F(x)$ is obtained using Monte Carlo simulations. In each experiment, the initial EC P_0 is updated according to a sequence of 1000 random binary variables, generating the EC P_{1000}. After repeating the experiment for $20,000$ times, we analyze the different EC P_{1000} obtained and summarize the results plotting the CDF.

A. Monotonicity of the bounds

Consider the system below, borrowed from [3],

$$A = \begin{bmatrix} 1.25 & 0 \\ 1 & 1.1 \end{bmatrix} \quad C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}'$$

$$Q = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix} \quad R = 2.5,$$

with $\lambda = 0.3$. In Figure 1 we show the upper bound $\overline{F}^T(x)$ and the lower bound $\underline{F}^T(x)$, for $T = 5$, $T = 10$ and $T = 15$ and assuming the i.i.d. packet drop model. Notice that, as T increases, the bounds become tighter, and for $T = 15$, it is difficult to distinguish between the lower and the upper bound. Note that the critical value (see [3]) of this system is $\lambda_c = 0.36$. This means that with the chosen value $\lambda = 0.3$, the expected value of the error covariance is infinite. Nevertheless, bounds for the CDF exist.

B. Performance comparison

We now compare our result with the one presented in [14]. Notice that in [14], in the case of sensors with limited computational capabilities, the authors assume that each packet contains enough measurements to make the error covariance bounded by a constant matrix. Contrariwise, we assume that only one measurement is sent in each packet, therefore the error covariance is bounded by a constant

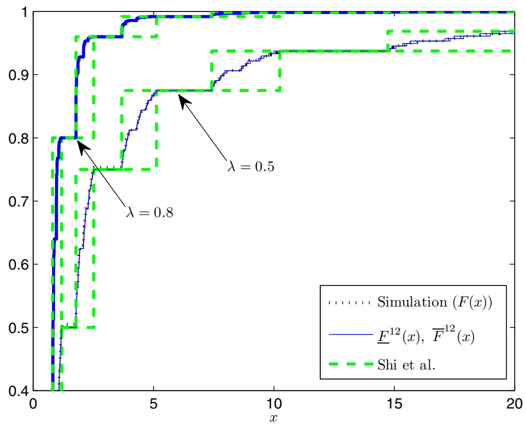

Fig. 2. Example B: Comparison of the CDF bounds for two different probabilities of measurement arrivals.

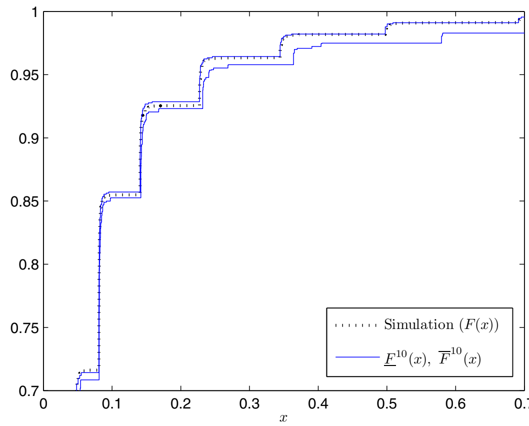

Fig. 3. Example C: CDF bounds for measurement arrivals following the Gilbert-Elliott model.

matrix when a packet is received, only when the matrix C is invertible.

Consider the scalar system below, taken from [14].

$$A = 1.4,\ C = 1,\ Q = 0.2,\ R = 0.5 \tag{69}$$

We consider the i.i.d. packet drop model for two different measurement arrival probabilities (i.e., $\lambda = 0.5$ and $\lambda = 0.8$), and compute the upper and lower bounds for the CDF. We do so using the expressions (50) and (49) with $T = 12$, as well those given in [14]. We see in Figure 2 how our proposed bounds are significantly tighter. Indeed, it is difficult to distinguish between the bounds and the actual CDF obtained from the simulation.

C. Markov Packet Arrivals

The system below describes a vehicle moving in a two dimensional space, according to the standard constant acceleration model. This example was considered in [14].

$$A = \begin{bmatrix} 1 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}'$$

$$Q = 0.01 I_4 \qquad\qquad R = 0.001 I_2,$$

The probability to receive a measurement is governed by the Markov chain in (63), with $\alpha = 0.5$ and $\beta = 0.8$. Figure 3 shows the upper and lower bounds of the CDF.

VI. Conclusion

We studied the Kalman filter for a discrete-time linear system, whose output is intermittently sampled, according to a sequence of binary random variables. We derived lower and upper bounds for the CDF of the EC. These bounds can be made arbitrarily tight, at the expense of increased computational complexity. We presented numerical examples demonstrating that the proposed bounds are tighter than those derived using other available methods.

References

[1] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S.S. Sastry. Foundations of control and estimation over lossy networks. *Proc. IEEE*, 95(1):163, 2007.

[2] B.D.O. Anderson and J.B. Moore. *Optimal filtering*. Prentice-Hall Englewood Cliffs, NJ, 1979.

[3] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9):1453–1464, 2004.

[4] H M Faridani. Performance of Kalman filter with missing measurements. *Automatica*, 22(1):117–120, 1986.

[5] X. Liu and A. Goldsmith. Kalman filtering with partial observation losses. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 4, pages 4180 – 4186 Vol.4, 2004.

[6] L. Schenato. Optimal estimation in networked control systems subject to random delay and packet drop. *IEEE Transactions on Automatic Control*, 53(5):1311–1317, 2008.

[7] Eduardo Rohr, Damian Marelli, and Minyue Fu. Statistical properties of the error covariance in a kalman filter with random measurement losses. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 5881 –5886, 2010.

[8] Y. Mo and B. Sinopoli. Kalman Filtering with Intermittent Observations: Tail Distribution and Critical Value. *IEEE Transaction on Automatic Control, Under review*.

[9] S. Kar, B. Sinopoli, and J.M.F. Moura. Kalman filtering with intermittent observations: Weak convergence to a stationary distribution. *Arxiv preprint arXiv:0903.2890*, 2010.

[10] S. Kar and J. Moura. Moderate Deviations of the Random Riccati Equation. *Arxiv preprint arXiv:0910.4686*, 2010.

[11] A. Censi. Kalman filtering with intermittent observations: convergence for semi-Markov chains and an intrinsic performance measure. *Automatic Control, IEEE Transactions on*, 56(2):376–381, 2011.

[12] A. Vakili and B. Hassibi. On the steady-state performance of Kalman filtering with intermittent observations for stable systems. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 6847–6852. IEEE, 2009.

[13] M. Epstein, L. Shi, A. Tiwari, and R.M. Murray. Probabilistic performance of state estimation across a lossy network. *Automatica*, 44(12):3046–3053, 2008.

[14] Ling Shi, M. Epstein, and R.M. Murray. Kalman filtering over a packet-dropping network: A probabilistic perspective. *Automatic Control, IEEE Transactions on*, 55(3):594 –604, march 2010.

[15] A. Censi. On the performance of Kalman filtering with intermittent observations: a geometric approach with fractals. In *American Control Conference, 2009. ACC'09.*, pages 3806–3812. IEEE, 2009.

[16] A. Ben-Israel and T. N. E. Greville. *Generalized inverses*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 15. Springer-Verlag, New York, second edition, 2003. Theory and applications.