On Kalman filtering for detectable systems with intermittent observations

Kurt Plarre and Francesco Bullo

Abstract

We consider the problem of Kalman filtering when observations are available according to a Bernoulli process. It is known that there exists a critical probability p_c such that, if measurements are available with probability greater than p_c, then the expected prediction covariance is bounded for all initial conditions; otherwise, it is unbounded for some initial conditions. We show that, when the system observation matrix restricted to the observable subspace is invertible, the known lower bound on p_c is tight. This result is based on a novel decomposition of positive semidefinite matrices.

I. Introduction

We consider the problem of Kalman filtering with intermittent observations, when measurements are available according to a Bernoulli process $\{\gamma_t\}_{t=0}^{\infty}$, with $\Pr[\gamma_t = 1] = p$, rather than at each time instant. Such problem is relevant to sensor networks, robotic networks, and networked control systems, in which the sensor and controller are communicating over an unreliable link or network. Packet losses or excessive delays will cause data to be unavailable at some time instants. When measurements are available at each time instant, the covariance of the prediction error follows an Algebraic Riccati Equation (ARE). When measurements arrive according to a stochastic process, the error covariance becomes a random quantity [1]. We study conditions under which the expected value of the prediction error covariance is bounded.

Kurt Plarre is with the Department of Mechanical Engineering, and the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA, 93106, plarre@engineering.ucsb.edu (corresponding author)

Francesco Bullo is with the Department of Mechanical Engineering, and the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA, 93106, bullo@engineering.ucsb.edu

It is known from [1] that if (A,C) is detectable, and $(\mathrm{A},\mathrm{Q}^{1/2})$ is controllable, where A is the system matrix, C is the observation matrix, and Q is the process noise covariance, then there exists a critical value p_c such that, if $p > p_c$ the expected error covariance is bounded for any initial condition, and if $p \leq p_c$, the expected error covariance is unbounded for some initial conditions. The critical probability p_c is known to be lower bounded by $1-1/\rho(\mathrm{A})^2$, where $\rho(\mathrm{A})$ is the spectral radius of A. In general, however, it is not known whether this bound is tight and an upper bound is given as the solution to a quasi-convex optimization problem. In some cases it is possible to show [1] that the lower bound on p_c is, in fact, the exact critical probability, for example, if C is invertible or if the system matrix A has only one unstable eigenvalue. It is also possible to construct examples in which the upper and lower bounds have different convergence conditions; see [1].

This and similar problems have been considered in the literature, e.g., [2] studies the case when the observations are split into two components, each of which are dropped independently. In [3] a randomized algorithm for sensor selection or scheduling is introduced. At each time instant, the measurement of only one sensor among a number is incorporated into the Kalman filter. The choice of sensor is made according to either a Bernoulli process or a Markov chain. Other variations of this problem have also been considered. For example, the case where a source node estimates the state and transmits such estimate over a packet dropping network is considered in [4], [5]. In [6] the problem of scheduling sensors in a communication bus is considered. It is shown that, under certain conditions, the optimum scheduling policy takes the form of a threshold policy, with the thresholds depending on the a priori distribution of the measurements. In [7] the authors model a packet dropping channel as a two state Markov chain, and find a necessary condition for stability of the Kalman filter. This condition is used to find necessary and sufficient conditions for stability when C is invertible, and in the one dimensional case. Finally, some of these results have also been extended to control with unreliable communications; see, for example, [8], [9], [10].

In this paper we prove that the critical probability is exactly $1-1/\rho(\mathrm{A})^2$ for detectable systems (see, for example, [11] for a definition of detectability) in which C is invertible on the observable subspace, as opposed to the whole space. To prove this fact we introduce two cones of positive semidefinite (PSD) matrices, which we call “observable cone” and “unobservable cone.” We then introduce a novel decomposition of PSD matrices, and use it to prove that any PSD matrix can be decomposed as a sum of two PSD matrices, one in the observable cone, and one in the unobservable cone. We use such decomposition to construct an upper bound on the expected Kalman filter covariance, as a sum of two terms, each of which is bounded above. To illustrate this result, we run simulations on an example.

The paper is organized as follows. In Section II we give the formal setup of the problem. In Section III we give some intermediate results, needed to prove the main result. We introduce the observable and unobservable cones in Section IV, and study some of their properties. In Section V we prove the main result. We give simulations in Section VI. Section VII contains concluding remarks.

II. Problem Setup

We consider the discrete time liner dynamical system

$$\begin{aligned} \mathrm{x}(t+1) &= \mathrm{A}\mathrm{x}(t) + \omega(t), \\ \mathrm{y}(t) &= \mathrm{C}\mathrm{x}(t) + \nu(t), \end{aligned} \tag{1}$$

where $\mathrm{x}(t) \in \mathbb{R}^n$ is the state, $\mathrm{y}(t) \in \mathbb{R}^m$ is the output, $\omega(t) \in \mathbb{R}^n$ is the process noise, and $\nu(t) \in \mathbb{R}^m$ is the measurement noise. $\omega(t)$ and $\nu(t)$ are independent white Gaussian noise processes with $\omega(t) \sim \mathcal{N}(0, \mathrm{Q})$, $\nu(t) \sim \mathcal{N}(0, \mathrm{R})$. We assume that R is invertible. The initial condition is $\mathrm{x}(0) = \mathrm{x}_0$ with $\mathrm{x}_0 \sim \mathcal{N}(0, \mathrm{P}_0)$. The Kalman filter for this system is given by the recursion

$$\begin{aligned} \hat{\mathrm{x}}(t+1|t) &= \mathrm{A}\hat{\mathrm{x}}(t|t), \\ \mathrm{P}(t+1|t) &= \mathrm{A}\mathrm{P}(t|t)\mathrm{A}^T + \mathrm{Q}, \\ \hat{\mathrm{x}}(t+1|t+1) &= \hat{\mathrm{x}}(t+1|t) + \mathrm{K}(t)\left(\mathrm{y}(t) - \mathrm{C}\hat{\mathrm{x}}(t+1|t)\right), \\ \mathrm{K}(t) &= \mathrm{P}(t+1|t)\mathrm{C}^T\left(\mathrm{C}\mathrm{P}(t+1|t)\mathrm{C}^T + \mathrm{R}\right)^{-1}, \\ \mathrm{P}(t+1|t+1) &= \left(\mathrm{I} - \mathrm{K}(t)\mathrm{C}\right)\mathrm{P}(t+1|t), \end{aligned}$$

where $\hat{\mathrm{x}}(t|t)$ and $\hat{\mathrm{x}}(t+1|t)$ are the state estimate and prediction at time t, respectively, and $\mathrm{P}(t|t)$ and $\mathrm{P}(t+1|t)$ are the estimation and prediction error covariances, respectively; $\mathrm{K}(t)$ is the Kalman gain.

We list here the notation that we use in the rest of the paper.

(i) We let $\mathbb{R}^{n_1 \times n_2}$ denote the set of n_1 times n_2 matrices with real entries and $\mathbb{S}^n_{\geq 0}$, (respectively, $\mathbb{S}^n_{>0}$) the set of symmetric positive semidefinite (respectively, positive definite) matrices of size n.

(ii) We let $\mathrm{P}(t) := \mathrm{P}(t+1|t)$, and $\mathrm{E}(t) := \mathrm{E}\left[\mathrm{P}(t)\right]$.

(iii) Given $\mathrm{A} \in \mathbb{R}^{n\times n}$, we define $\rho(\mathrm{A}) := \max\{|\lambda| \mid \lambda$ is an eigenvalue of $\mathrm{A}\}$, i.e., the spectral radius of A.

(iv) For $1 \leq m \leq n$, we define $\mathrm{Block} : \mathbb{R}^{m\times m} \times \mathbb{R}^{m\times(n-m)} \times \mathbb{R}^{(n-m)\times m} \times \mathbb{R}^{(n-m)\times(n-m)} \rightarrow \mathbb{R}^{n\times n}$, by

$$\mathrm{Block}(\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \mathrm{X}_4) := \begin{bmatrix} \mathrm{X}_1 & \mathrm{X}_2 \\ \mathrm{X}_3 & \mathrm{X}_4 \end{bmatrix}.$$

By convention, if $m = n$, $\mathrm{Block}(\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \mathrm{X}_4) = \mathrm{X}_1$.

(v) We define the map $\mathrm{Series} : \mathbb{S}^n_{\geq 0} \times \mathbb{R}^{n\times n} \times (\mathbb{N} \cup \{-1, +\infty\}) \rightarrow \mathbb{S}^n_{\geq 0}$ by $\mathrm{Series}(\mathrm{X}, \Omega, -1) := 0$ and

$$\mathrm{Series}(\mathrm{X}, \Omega, t) := \sum_{\tau=0}^{t} \Omega^\tau \mathrm{X} \Omega^{\tau T}, \quad t \in \mathbb{N} \cup \{+\infty\}.$$

(vi) Finally, we define the "stacking map" $\mathrm{Stack} : \mathbb{R}^{n\times n} \rightarrow \mathbb{R}^{n^2}$, by letting $\mathrm{Stack}(\Omega)$ be the vector obtained by stacking the columns of Ω.

To simplify the presentation, we only indicate the size of a matrix explicitly when it is not clear from the context.

III. Preliminary facts

The evolution of $\mathrm{P}(t)$ is given by the Algebraic Riccati Equation (ARE)

$$\mathrm{P}(t+1) = \mathrm{Q} + \mathrm{A}\mathrm{P}(t)\mathrm{A}^T - \mathrm{A}\mathrm{P}(t)\mathrm{C}^T \left(\mathrm{C}\mathrm{P}(t)\mathrm{C}^T + \mathrm{R}\right)^{-1} \mathrm{C}\mathrm{P}(t)\mathrm{A}^T.$$

When measurements are available randomly, $\mathrm{P}(t)$ becomes a random quantity. In this paper we assume that a measurement at time t is available if $\gamma_t = 1$, where $\{\gamma_t\}_{t=0}^{+\infty}$ is a Bernoulli process with $\Pr\left[\gamma_t = 1\right] = p$. In this case, the ARE can be rewritten as [1]

$$\mathrm{P}(t+1) = \mathrm{Q} + \mathrm{A}\mathrm{P}(t)\mathrm{A}^T - \gamma_t \mathrm{A}\mathrm{P}(t)\mathrm{C}^T \left(\mathrm{C}\mathrm{P}(t)\mathrm{C}^T + \mathrm{R}\right)^{-1} \mathrm{C}\mathrm{P}(t)\mathrm{A}^T. \tag{2}$$

We are interested in sufficient conditions for stability of the expected value of $\mathrm{P}(t)$. To find such conditions, we construct an upper bound on $\mathrm{E}\left[\mathrm{P}(t)\right]$. From [1], [3] we know that an upper bound on $\mathrm{E}(t)$ can be obtained using Jensen's Inequality:

$$\begin{aligned}\mathrm{E}(t+1) &\leq \mathrm{A}\mathrm{E}(t)\mathrm{A}^T + \mathrm{Q} - p\mathrm{A}\mathrm{E}(t)\mathrm{C}^T\left(\mathrm{C}\mathrm{E}(t)\mathrm{C}^T+\mathrm{R}\right)^{-1}\mathrm{C}\mathrm{E}(t)\mathrm{A}^T\\ &= \left(\mathrm{A}\mathrm{E}(t)\mathrm{A}^T - \mathrm{A}\mathrm{E}(t)\mathrm{C}^T\left(\mathrm{C}\mathrm{E}(t)\mathrm{C}^T+\mathrm{R}\right)^{-1}\mathrm{C}\mathrm{E}(t)\mathrm{A}^T\right)p + \mathrm{A}\mathrm{E}(t)\mathrm{A}^T(1-p) + \mathrm{Q}\\ &= \mathrm{A}\mathcal{G}\left(\mathrm{E}(t), \mathrm{C}^T\mathrm{R}^{-1}\mathrm{C}\right)\mathrm{A}^T p + \mathrm{A}\mathrm{E}(t)\mathrm{A}^T(1-p) + \mathrm{Q},\end{aligned} \tag{3}$$

where $\mathrm{E}(0) = \mathbb{E}\left[\mathrm{P}_0\right]$ and where the map $\mathcal{G} : \mathbb{S}^n_{\geq 0} \times \mathbb{S}^n_{\geq 0} \to \mathbb{S}^n_{\geq 0}$ is defined by $\mathcal{G}(\mathrm{X}, \mathrm{M}) := \mathrm{X}\left(\mathrm{I} + \mathrm{M}\mathrm{X}\right)^{-1}$. The last step in equation (3) follows from the matrix inversion lemma or, equivalently, from property (iii) in the following lemma.

Lemma 1 (Properties of $\mathcal{G}$)**.** *Let* $\mathrm{M}, \mathrm{X}, \mathrm{Y} \in \mathbb{S}^n_{\geq 0}$*. Let* $\bar{\mathrm{M}} \in \mathbb{R}^{n\times n}$ *be a matrix, such that* $\bar{\mathrm{M}}\bar{\mathrm{M}}^T = \mathrm{M}$ *(such a matrix always exists). Then the following holds:*

(i) $\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathcal{G}(\mathrm{X}, \mathrm{M})^T$*;*

(ii) $\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathrm{X} - \mathrm{X}(\mathrm{I} + \mathrm{M}\mathrm{X})^{-1}\mathrm{M}\mathrm{X}$*;*

(iii) $\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathrm{X} - \mathrm{X}\bar{\mathrm{M}}\left(\mathrm{I} + \bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}}\right)^{-1}\bar{\mathrm{M}}^T\mathrm{X}$*;*

(iv) *if* $\mathrm{X} \leq \mathrm{Y}$*, then* $\mathcal{G}(\mathrm{X}, \mathrm{M}) \leq \mathcal{G}(\mathrm{Y}, \mathrm{M})$*; and*

(v) *if* M *is invertible, then* $\mathcal{G}(\mathrm{X}, \mathrm{M}) \leq \mathrm{M}^{-1}$.

Proof: To prove (i) we observe that $\mathcal{G}(\mathrm{X}, \mathrm{M})^T = (\mathrm{I} + \mathrm{M}\mathrm{X})^{-T}\mathrm{X}$. We then have that

$$(\mathrm{I} + \mathrm{M}\mathrm{X})^T\left(\mathcal{G}(\mathrm{X}, \mathrm{M}) - \mathcal{G}(\mathrm{X}, \mathrm{M})^T\right)(\mathrm{I} + \mathrm{M}\mathrm{X}) = (\mathrm{I} + \mathrm{M}\mathrm{X})^T\mathrm{X} - \mathrm{X}(\mathrm{I} + \mathrm{M}\mathrm{X}) = 0.$$

It can be proved that $(\mathrm{I}+\mathrm{M}\mathrm{X})$ is invertible. This means that $\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathcal{G}(\mathrm{X}, \mathrm{M})^T$. To prove (ii) we observe that $\mathcal{G}(\mathrm{X}, \mathrm{M})(\mathrm{I} + \mathrm{M}\mathrm{X}) = \mathrm{X}$. This implies that

$$\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathrm{X} - \mathcal{G}(\mathrm{X}, \mathrm{M})\mathrm{M}\mathrm{X} = \mathrm{X} - \mathrm{X}(\mathrm{I} + \mathrm{M}\mathrm{X})^{-1}\mathrm{M}\mathrm{X}. \tag{4}$$

We now prove (iii). We have the following sequence of implications:

$$\begin{aligned}&\mathcal{G}(\mathrm{X}, \mathrm{M})(\mathrm{I} + \mathrm{M}\mathrm{X}) = \mathrm{X} && \Rightarrow \mathcal{G}(\mathrm{X}, \mathrm{M})\bar{\mathrm{M}} + \mathcal{G}(\mathrm{X}, \mathrm{M})\bar{\mathrm{M}}\bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}} = \mathrm{X}\bar{\mathrm{M}}\\ &\Rightarrow \mathcal{G}(\mathrm{X}, \mathrm{M})\bar{\mathrm{M}}\left(\mathrm{I} + \bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}}\right) = \mathrm{X}\bar{\mathrm{M}} && \Rightarrow \left(\mathrm{I} + \bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}}\right)\bar{\mathrm{M}}^T\mathcal{G}(\mathrm{X}, \mathrm{M}) = \bar{\mathrm{M}}^T\mathrm{X}\\ &\Rightarrow \bar{\mathrm{M}}^T\mathcal{G}(\mathrm{X}, \mathrm{M}) = \left(\mathrm{I} + \bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}}\right)^{-1}\bar{\mathrm{M}}^T\mathrm{X} && \Rightarrow \mathrm{X}\mathrm{M}\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathrm{X}\bar{\mathrm{M}}\left(\mathrm{I} + \bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}}\right)^{-1}\bar{\mathrm{M}}^T X.\end{aligned}$$

From (4) we know that $\mathrm{X} - \mathrm{X}\mathrm{M}\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathcal{G}(\mathrm{X}, \mathrm{M})$. This finishes the proof. To prove (iv) we follow [1], [3]. We define a map $\phi : \mathbb{R}^{n\times n} \times \mathbb{S}^n_{\geq 0} \to \mathbb{S}^n_{\geq 0}$ by $\phi(\mathrm{K}, \mathrm{X}) := (\mathrm{I} + \mathrm{K}\bar{\mathrm{M}}^T)\mathrm{X}(\mathrm{I} + \mathrm{K}\bar{\mathrm{M}}^T)^T + \mathrm{K}\mathrm{K}^T$. We rewrite $\phi(\mathrm{K}, \mathrm{X})$ as

$$\phi(\mathrm{K}, \mathrm{X}) = (\mathrm{K} - \mathrm{Z})\left(\bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}} + \mathrm{I}\right)(\mathrm{K} - \mathrm{Z})^T + \mathrm{Z}',$$

where $\mathrm{Z} = -\mathrm{X}\bar{\mathrm{M}}(\bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}} + \mathrm{I})^{-1}$ and $\mathrm{Z}' = \mathrm{X} - \mathrm{Z}(\bar{\mathrm{M}}^T\mathrm{X}\bar{\mathrm{M}} + \mathrm{I})\mathrm{Z}^T$. We can see that for fixed X, $\phi(\mathrm{K}, \mathrm{X})$ is minimized by $\mathrm{K} = \mathrm{K}_\mathrm{X} := \mathrm{Z}$, and that $\phi(\mathrm{K}_\mathrm{X}, \mathrm{X}) = \mathcal{G}(\mathrm{X}, \mathrm{M})$. Also, for any K, if $\mathrm{X} \leq \mathrm{Y}$ are PSD, then $\phi(\mathrm{K}, \mathrm{X}) \leq \phi(\mathrm{K}, \mathrm{Y})$. Let $\mathrm{X} \leq \mathrm{Y}$. Then

$$\mathcal{G}(\mathrm{X}, \mathrm{M}) = \phi(\mathrm{K}_\mathrm{X}, \mathrm{X}) \leq \phi(\mathrm{K}_\mathrm{Y}, \mathrm{X}) \leq \phi(\mathrm{K}_\mathrm{Y}, \mathrm{Y}) = \mathcal{G}(\mathrm{Y}, \mathrm{M}).$$

Finally, to prove (v), we observe that for $\mathrm{X} \in \mathbb{S}^n_{\geq 0}$, $\mathcal{G}(\mathrm{X}, \mathrm{M}) \leq \mathcal{G}(\mathrm{trace}\,(\mathrm{X})\,\mathrm{I}, \mathrm{M})$. But

$$\mathcal{G}(\mathrm{trace}\,(\mathrm{X})\,\mathrm{I}, \mathrm{M})(\mathrm{I} + \mathrm{trace}\,(\mathrm{X})\,\mathrm{M}) = \mathrm{trace}\,(\mathrm{X})\,\mathrm{I}.$$

Given that $\mathcal{G}(\mathrm{trace}\,(\mathrm{X})\,\mathrm{I}, \mathrm{M}) \geq 0$, we have $\mathrm{trace}\,(\mathrm{X})\,\mathcal{G}(\mathrm{trace}\,(\mathrm{X})\,\mathrm{I}, \mathrm{M})\mathrm{M} \leq \mathrm{trace}\,(\mathrm{X})\,\mathrm{I}$, which implies that $\mathcal{G}(\mathrm{trace}\,(\mathrm{X})\,\mathrm{I}, \mathrm{M}) \leq \mathrm{M}^{-1}$. ■

To find an upper bound on the covariance we will also need the following fact.

Lemma 2 (Bound on series)**.** *Let* $\mathrm{X} \in \mathbb{S}^n_{\geq 0}$*, and* $\Omega \in \mathbb{R}^{n \times n}$ *with* $\rho(\Omega) < 1$*. Then* $\mathrm{Series}(\mathrm{X}, \Omega, +\infty)$ *exists, and* $\mathrm{Series}(\mathrm{X}, \Omega, t) \leq \mathrm{Series}(\mathrm{X}, \Omega, +\infty) \leq \mathrm{trace}\,(\mathrm{Series}(\mathrm{X}, \Omega, +\infty))\,\mathrm{I}$*, for all* $t \geq 0$*.*

Proof: Using the stacking map, we can write the following recursion for $\mathrm{Series}(\mathrm{X}, \Omega, t)$:

$$\mathrm{Stack}(\mathrm{Series}(\mathrm{X}, \Omega, t)) = (\Omega \otimes \Omega)\,\mathrm{Stack}(\mathrm{Series}(\mathrm{X}, \Omega, t-1)) + \mathrm{Stack}(\mathrm{X}), \tag{5}$$

where $\otimes$ denotes the Kronecker product. It is known that $\rho(\Omega \otimes \Omega) = \rho(\Omega)^2$. Thus, if $\rho\rho(\Omega) < 1$, equation (5) defines a stable linear system with a constant input, and hence, $\mathrm{Stack}(\mathrm{Series}(\mathrm{X}, \Omega, t))$ converges to $\mathrm{Stack}(\mathrm{Series}(\mathrm{X}, \Omega, +\infty))$, as $t \to +\infty$. This means that $\mathrm{Series}(\mathrm{X}, \Omega, t)$ also converges. Observing that $\mathrm{Series}(\mathrm{X}, \Omega, t)$ is increasing in t, we have the result. ■

IV. CONES OF PSD MATRICES

We observe that, if $\mathrm{X} \in \mathbb{S}^n_{\geq 0}$ is such that $\mathrm{MX} = 0$, then $\mathcal{G}(\mathrm{X}, \mathrm{M}) = \mathrm{X}$. This means that the Riccati equation acts on X as if there was no observation. Motivated by this fact, we seek to decompose a PSD matrix P as a sum of two components, which we denote P_o and P_u, such that $\mathrm{MP}_\mathrm{u} = 0$. This leads to the definition of the following cones of PSD matrices:

Definition 3. *Given any matrix* $\mathrm{M} \in \mathbb{R}^{n_1 \times n_2}$*, we define the* observable cone $\mathcal{K}_\mathrm{o}\,(\mathrm{M})$*, and the* unobservable cone $\mathcal{K}_\mathrm{u}\,(\mathrm{M})$ *by*

$$\mathcal{K}_\mathrm{o}\,(\mathrm{M}) := \{\mathrm{X} \in \mathbb{S}^{n_2}_{\geq 0} \mid \ker\,(\mathrm{MX}) = \ker\,(\mathrm{X})\},$$

$$\mathcal{K}_\mathrm{u}\,(\mathrm{M}) := \{\mathrm{X} \in \mathbb{S}^{n_2}_{\geq 0} \mid \mathrm{MX} = 0\}.$$

Note that, for generality, we have defined the cones for any matrix M. In the next theorem we prove that, if $\mathrm{M} \in \mathbb{S}^n_{\geq 0}$, then any matrix $\mathrm{P} \in \mathbb{S}^n_{\geq 0}$ can be decomposed as $\mathrm{P} = \mathrm{P_o} + \mathrm{P_u}$, with $\mathrm{P_o} \in \mathcal{K}_\mathrm{o}(\mathrm{M})$ and $\mathrm{P_u} \in \mathcal{K}_\mathrm{u}(\mathrm{M})$. The proof uses the results given in the appendix.

Theorem 4 (Decomposition of PSD matrices, induced by cones)**.** *Let* $\mathrm{M}, \mathrm{P} \in \mathbb{S}^n_{\geq 0}$. *Then there exist unique matrices* $\mathrm{P_o} \in \mathcal{K}_\mathrm{o}(\mathrm{M})$ *and* $\mathrm{P_u} \in \mathcal{K}_\mathrm{u}(\mathrm{M})$, *such that* $\mathrm{P_o} + \mathrm{P_u} = \mathrm{P}$.

Proof: If M is full rank, then clearly $\mathrm{P_o} = \mathrm{P}$ and $\mathrm{P_u} = 0$ satisfy the desired properties. Thus, in what follows, we assume that $\mathrm{rank}(\mathrm{M}) = m < n$. We first prove the theorem for $\mathrm{M} = \mathrm{Block}(\mathrm{M}_1, 0, 0, 0)$, with M_1 invertible. In view of Lemma 10 in the appendix, we know that there exist matrices $\mathrm{P}_1 \in \mathbb{S}^m_{\geq 0}$, $\mathrm{L} \in \mathbb{R}^{m \times (n-m)}$, and $\tilde{\mathrm{P}} \in \mathbb{S}^{(n-m)\times(n-m)}_{\geq 0}$, such that $\mathrm{P} = \mathrm{P_o} + \mathrm{P_u}$, with $\mathrm{P_o} := \mathrm{Block}(\mathrm{P}_1, \mathrm{P}_1\mathrm{L}, \mathrm{L}^T\mathrm{P}_1, \mathrm{L}^T\mathrm{P}_1\mathrm{L})$, and $\mathrm{P_u} := \mathrm{Block}(0, 0, 0, \tilde{\mathrm{P}})$. Then, it is easy to see that $\mathrm{P_u} \in \mathcal{K}_\mathrm{u}(\mathrm{M})$, and for any vector $\mathrm{v} \in \mathbb{R}^n$, $\mathrm{MP_o v} = 0$ if and only if $\mathrm{P_o v} = 0$, which means that $\mathrm{P_o} \in \mathcal{K}_\mathrm{o}(\mathrm{M})$. This proves that there exists at least one decomposition satisfying the required properties. To prove uniqueness, suppose that $\mathrm{X_o}$, $\mathrm{X_u}$ is another such decomposition. We can always write $\mathrm{X_o}$ as $\mathrm{X_o} = \mathrm{Block}(\mathrm{X}_1, \mathrm{X}_1\mathrm{N}, \mathrm{N}^T\mathrm{X}_1, \mathrm{N}^T\mathrm{X}_1\mathrm{N}) + \mathrm{Block}(0, 0, 0, \tilde{\mathrm{X}})$. The condition $\mathrm{MX_o v} = 0 \Leftrightarrow \mathrm{X_o v} = 0$ for $\mathrm{v} = [\mathrm{v}_1^T \ \mathrm{v}_2^T]^T$ translates to $\mathrm{X}_1(\mathrm{v}_1 + \mathrm{N}\mathrm{v}_2) = 0 \Leftrightarrow \tilde{\mathrm{X}}\mathrm{v}_2 = 0$. This, in turn, implies that $\tilde{\mathrm{X}} = 0$ (otherwise one can set $\mathrm{v}_1 = -\mathrm{N}\mathrm{v}_2$ with $\tilde{\mathrm{X}}\mathrm{v}_2 \neq 0$). This means that $\mathrm{X_o} = \mathrm{Block}(\mathrm{X}_1, \mathrm{X}_1\mathrm{N}, \mathrm{N}^T\mathrm{X}_1, \mathrm{N}^T\mathrm{X}_1\mathrm{N})$. Using this fact, and $\mathrm{MP_o} = \mathrm{M}(\mathrm{P_o}+\mathrm{P_u}) = \mathrm{M}(\mathrm{X_o}+\mathrm{X_u}) = \mathrm{MX_o}$, it can be proved that $\mathrm{X_o} = \mathrm{P_o}$, and thus, also $\mathrm{X_u} = \mathrm{P_u}$.

Now, let $\mathrm{M} \in \mathbb{S}^n_{\geq 0}$, arbitrary, except for being rank deficient. Then, there exists an orthogonal matrix $\mathrm{U} \in \mathbb{R}^{n \times n}$, such that $\mathrm{M}' := \mathrm{U}^T\mathrm{MU} = \mathrm{Block}(\mathrm{M}'_1, 0, 0, 0)$, with M'_1 diagonal and invertible (this corresponds to the singular value decomposition of M). Let $\mathrm{P}' := \mathrm{U}^T\mathrm{PU}$. We know we can find $\mathrm{P}'_\mathrm{o} \in \mathcal{K}_\mathrm{o}(\mathrm{M}')$ and $\mathrm{P}'_\mathrm{u} \in \mathcal{K}_\mathrm{u}(\mathrm{M}')$, such that $\mathrm{P}'_\mathrm{o} + \mathrm{P}'_\mathrm{u} = \mathrm{P}'$. Let $\mathrm{P_o} := \mathrm{UP}'_\mathrm{o}\mathrm{U}^T$ and $\mathrm{P_u} := \mathrm{UP}'_\mathrm{u}\mathrm{U}^T$. Then $\mathrm{P_o} + \mathrm{P_u} = \mathrm{U}(\mathrm{P}'_\mathrm{o} + \mathrm{P}'_\mathrm{u})\mathrm{U}^T = \mathrm{UP}'\mathrm{U}^T = \mathrm{P}$. Also, $\mathrm{MP_u} = \mathrm{UM}'\mathrm{U}^T\mathrm{UP}'_\mathrm{u}\mathrm{U}^T = \mathrm{UM}'\mathrm{P}'_\mathrm{u}\mathrm{U}^T = 0$. If $\mathrm{MP_o v} = 0$, then $\mathrm{UM}'\mathrm{U}^T\mathrm{UP}'_\mathrm{o}\mathrm{U}^T\mathrm{v} = 0$. Since U is invertible, this means that $\mathrm{M}'\mathrm{P}'_\mathrm{o}\mathrm{U}^T\mathrm{v} = 0$. Then $\mathrm{P}'_\mathrm{o}\mathrm{U}^T\mathrm{v} = 0$. Finally, $\mathrm{UP}'_\mathrm{o}\mathrm{U}^T\mathrm{v} = \mathrm{P_o v} = 0$, and so, $\ker(\mathrm{MP_o}) \subseteq \ker(\mathrm{P_o})$. As above, this means that $\ker(\mathrm{MP_o}) = \ker(\mathrm{P_o})$. The decomposition of P' is unique. This means that also the decomposition of P is unique. This completes the proof. ■

There exists also a recursive algorithm to compute the decomposition. We do not include it here due to space constraints. We conclude this section with a list of some properties of the cones.

Lemma 5 (Separation)**.** *Let* $\mathrm{M} \in \mathbb{S}^n_{\geq 0}$. *Let* $\mathrm{P_o} \in \mathcal{K}_\mathrm{o}(\mathrm{M})$, $\mathrm{P_u} \in \mathcal{K}_\mathrm{u}(\mathrm{M})$. *Then* $\mathcal{G}(\mathrm{P_o} + \mathrm{P_u}, \mathrm{M}) = \mathcal{G}(\mathrm{P_o}, \mathrm{M}) + \mathrm{P_u}$.

Proof: Let $\bar{\mathrm{M}} \in \mathbb{R}^{n\times n}$ such that $\bar{\mathrm{M}}\bar{\mathrm{M}}^T = \mathrm{M}$. Then, it can be proved that $\mathrm{MP_u} = 0$ if and only if $\bar{\mathrm{M}}^T\mathrm{P_u} = 0$. From Lemma 1-(iii) we have

$$\mathcal{G}(\mathrm{P_o} + \mathrm{P_u}, \mathrm{M}) = (\mathrm{P_o} + \mathrm{P_u}) - \mathrm{P_o}\bar{\mathrm{M}}\left(\mathrm{I} + \bar{\mathrm{M}}^T\mathrm{P_o}\bar{\mathrm{M}}\right)^{-1}\bar{\mathrm{M}}^T\mathrm{P_o} = \mathcal{G}(\mathrm{P_o}, \mathrm{M}) + \mathrm{P_u}.$$

■

Lemma 6 (Effect of ARE on observable matrix)**.** *Let* $\mathrm{M} = \mathrm{Block}(\mathrm{M_1}, 0, 0, 0) \in \mathbb{S}^n_{\geq 0}$ *with* $\mathrm{M_1} \in \mathbb{S}^m_{>0}$. *Let* $\mathrm{P} = \mathrm{Block}(\mathrm{P_1}, \mathrm{P_1L}, \mathrm{L}^T\mathrm{P_1}, \mathrm{L}^T\mathrm{P_1L}) \in \mathcal{K}_\mathrm{o}(\mathrm{M})$. *Then*

$$\mathcal{G}(\mathrm{P}, \mathrm{M}) = \mathrm{Block}(\mathcal{G}(\mathrm{P_1}, \mathrm{M_1}), \mathcal{G}(\mathrm{P_1}, \mathrm{M_1})\mathrm{L}, \mathrm{L}^T\mathcal{G}(\mathrm{P_1}, \mathrm{M_1}), \mathrm{L}^T\mathcal{G}(\mathrm{P_1}, \mathrm{M_1})\mathrm{L}).$$

Additionally, if $\mathrm{P} = \theta_P\mathrm{I_o}$, *with* $\theta_P \in \mathbb{R}_{\geq 0}$ *and* $\mathrm{I_o} = \mathrm{Block}(\mathrm{I}, 0, 0, 0) \in \mathcal{K}_\mathrm{o}(\mathrm{M})$, *then* $\mathcal{G}(\mathrm{P}, \mathrm{M}) \leq \mathrm{trace}\left(\mathrm{M}_1^{-1}\right)\mathrm{I_o}$.

Proof: We can explicitly compute $(\mathrm{I} + \mathrm{MP})^{-1}$:

$$\begin{aligned}&\left(\mathrm{I} + \mathrm{Block}(\mathrm{M_1}, 0, 0, 0)\mathrm{Block}(\mathrm{P_1}, \mathrm{P_1L}, \mathrm{L}^T\mathrm{P_1}, \mathrm{L}^T\mathrm{P_1L})\right)^{-1}\\ &\qquad = \mathrm{Block}\left((\mathrm{I} + \mathrm{M_1P_1})^{-1},\ -(\mathrm{I} + \mathrm{M_1P_1})^{-1}\mathrm{M_1P_1}L,\ 0,\ \mathrm{I}\right).\end{aligned}$$

Let $\mathrm{P}(\mathrm{I} + \mathrm{MP})^{-1} =: \mathrm{Block}(z_1(\mathrm{P}, \mathrm{M}), z_2(\mathrm{P}, \mathrm{M}), z_3(\mathrm{P}, \mathrm{M}), z_4(\mathrm{P}, \mathrm{M}))$. Then

$$\begin{aligned}z_1(\mathrm{P}, \mathrm{M}) &= \mathrm{P_1}(\mathrm{I} + \mathrm{M_1P_1})^{-1} = \mathcal{G}(\mathrm{P_1}; \mathrm{M_1}),\\ z_2(\mathrm{P}, \mathrm{M}) &= -\mathrm{P_1}(\mathrm{I} + \mathrm{M_1P_1})^{-1}\mathrm{M_1P_1L} + \mathrm{P_1L} = \mathcal{G}(\mathrm{P_1}, \mathrm{M_1})\mathrm{L},\\ z_3(\mathrm{P}, \mathrm{M}) &= \mathrm{L}^T\mathrm{P_1}(\mathrm{I} + \mathrm{M_1P_1})^{-1} = \mathrm{L}^T\mathcal{G}(\mathrm{P_1}, \mathrm{M_1}),\\ z_4(\mathrm{P}, \mathrm{M}) &= -\mathrm{L}^T\mathrm{P_1}(\mathrm{I} + \mathrm{M_1P_1})^{-1}\mathrm{M_1P_1L} + \mathrm{L}^T\mathrm{P_1L} = \mathrm{L}^T\mathcal{G}(\mathrm{P_1}, \mathrm{M_1})\mathrm{L}.\end{aligned}$$

If $\mathrm{P} = \theta_P\mathrm{I_o}$, then

$$\mathcal{G}(\mathrm{P}, \mathrm{M}) = \mathcal{G}(\theta_P\mathrm{I_o}, \mathrm{M}) = \mathrm{Block}(\mathcal{G}(\mathrm{I}, \mathrm{M_1}), 0, 0, 0) \leq \mathrm{Block}(\mathrm{M}_1^{-1}, 0, 0, 0) \leq \mathrm{trace}\left(\mathrm{M}_1^{-1}\right)\mathrm{I_o}.$$

■

We also observe that if $\mathrm{A} = \mathrm{Block}(\mathrm{A_1}, 0, \mathrm{A_3}, \mathrm{A_4})$ and $\mathrm{X} = \mathrm{Block}(0, 0, 0, \tilde{\mathrm{X}})$, then $\mathrm{AXA}^T = \mathrm{Block}(0, 0, 0, \mathrm{A_4}\tilde{\mathrm{X}}\mathrm{A}_4^T)$. Finally, if $\mathrm{X}, \mathrm{Y} \in \mathcal{K}_\mathrm{u}(\mathrm{M})$, then $\mathrm{X} + \mathrm{Y} \in \mathcal{K}_\mathrm{u}(\mathrm{M})$.

V. Main result

Theorem 7 (Critical arrival probability)**.** *Let* $\{\gamma_t\}_{t=0}^{+\infty}$ *be a Bernoulli process with* $\Pr[\gamma_t = 1] = p$*. Let* (A, C) *be detectable, let* $\rho(\mathrm{A})^2(1-p) < 1$*, and let* $\mathrm{P}(t)$ *be the solution to equation (2) with* $\mathrm{trace}\left(\mathrm{E}\left[\mathrm{P}_0\right]\right) < +\infty$*. If* C *is an invertible linear operator when restricted to the observable subspace of* (A, C)*, then* $\mathrm{E}\left[\mathrm{P}(t)\right]$ *is upper bounded, uniformly in* t*.*

Proof: If C is invertible, we know from [1] that the expected covariance is uniformly bounded. We thus assume that $\mathrm{M} := \mathrm{C}^T\mathrm{R}^{-1}\mathrm{C}$ is rank deficient. We assume without loss of generality that the system is written in Kalman canonical form, that is, $\mathrm{A} = \mathrm{Block}(\mathrm{A}_1, 0, \mathrm{A}_3, \mathrm{A}_4)$, $\mathrm{C} = [\mathrm{C}_1\ 0]$, with A_4 stable and C_1 invertible. Then $\mathrm{M} := \mathrm{Block}(\mathrm{M}_1, 0, 0, 0)$, with $\mathrm{M}_1 := \mathrm{C}_1^T\mathrm{R}^{-1}\mathrm{C}_1$. We define the matrices $\mathrm{I}_\mathrm{o} := \mathrm{Block}(\mathrm{I}, 0, 0, 0) \in \mathcal{K}_\mathrm{o}(\mathrm{M})$, $\mathrm{I}_\mathrm{u} := \mathrm{Block}(0, 0, 0, \mathrm{I}) \in \mathcal{K}_\mathrm{u}(\mathrm{M})$, $\mathrm{S}(t) := \mathrm{Series}(\mathrm{I}, \sqrt{1-p}\,\mathrm{A}, t)$, and $\mathrm{G}(t) := \mathrm{Series}(\mathrm{I}_\mathrm{u}, \mathrm{A}, t)$. By Lemma 2 and the condition on p, we know that $\mathrm{S}(t) \le \theta_\mathrm{S}\mathrm{I}$. Again, using Lemma 2, and the detectability of (A, C), $\mathrm{G}(t) \le \theta_\mathrm{G}\mathrm{I}$. Here θ_S and θ_G satisfy $\theta_\mathrm{S} := \mathrm{trace}\left(\mathrm{S}(+\infty)\right) < +\infty$ and $\theta_\mathrm{G} := \mathrm{trace}\left(\mathrm{G}(+\infty)\right) < +\infty$. Let

$$a := \max\left\{\mathrm{trace}\left(\mathrm{E}\left[\mathrm{P}_0\right]\right), p\,\mathrm{trace}\left(\mathrm{M}_1^{-1}\right)\mathrm{trace}\left(\mathrm{A}\mathrm{I}_\mathrm{o}\mathrm{A}^T\right) + \mathrm{trace}\left(\mathrm{Q}\right)\right\}, \qquad b := pa\theta_\mathrm{S}.$$

We will prove that, for all $t \ge 0$, $\mathrm{E}\left[\mathrm{P}(t)\right] \le a\mathrm{S}(t) + b\mathrm{A}\mathrm{G}(t-1)\mathrm{A}^T \le a\theta_\mathrm{S}\mathrm{I} + b\theta_\mathrm{G}\mathrm{A}\mathrm{A}^T$. The proof is by induction. It is clearly true at time $t = 0$. Assume it is true at time t. At time $t+1$ we have

$$\begin{aligned}
\mathrm{E}(t+1) &\le \mathrm{A}\mathcal{G}\left(\mathrm{E}(t), \mathrm{M}\right)\mathrm{A}^T p + \mathrm{A}\mathrm{E}(t)\mathrm{A}^T(1-p) + \mathrm{Q} \\
&\overset{(a)}{\le} \mathrm{A}\mathcal{G}\left(a\mathrm{S}(t) + b\mathrm{A}\mathrm{G}(t-1)\mathrm{A}^T, \mathrm{M}\right)\mathrm{A}^T p + \mathrm{A}\left(a\mathrm{S}(t) + b\mathrm{A}\mathrm{G}(t-1)\mathrm{A}^T\right)\mathrm{A}^T(1-p) + \mathrm{Q} \\
&\overset{(b)}{=} \mathrm{A}\mathcal{G}\left(a\mathrm{S}(t), \mathrm{M}\right)\mathrm{A}^T p + a\mathrm{A}\mathrm{S}(t)\mathrm{A}^T(1-p) + b\mathrm{A}^2\mathrm{G}(t-1)\mathrm{A}^{2T} + \mathrm{Q} \\
&\overset{(c)}{\le} \mathrm{A}\mathcal{G}\left(a\theta_\mathrm{S}\mathrm{I}, \mathrm{M}\right)\mathrm{A}^T p + a\mathrm{A}\mathrm{S}(t)\mathrm{A}^T(1-p) + b\mathrm{A}^2\mathrm{G}(t-1)\mathrm{A}^{2T} + \mathrm{Q} \\
&\overset{(d)}{=} \mathrm{A}\mathcal{G}\left(a\theta_\mathrm{S}\mathrm{I}_\mathrm{o}, \mathrm{M}\right)\mathrm{A}^T p + a\mathrm{A}\mathrm{S}(t)\mathrm{A}^T(1-p) + pa\theta_\mathrm{S}\mathrm{A}\mathrm{I}_\mathrm{u}\mathrm{A}^T + b\mathrm{A}^2\mathrm{G}(t-1)\mathrm{A}^{2T} + \mathrm{Q} \\
&\overset{(e)}{\le} \mathrm{trace}\left(\mathrm{M}_1\right)^{-1}\mathrm{A}\mathrm{I}_\mathrm{o}\mathrm{A}^T p + \mathrm{Q} + a\mathrm{A}\mathrm{S}(t)\mathrm{A}^T(1-p) + b\mathrm{A}\left(\mathrm{I}_\mathrm{u} + \mathrm{A}\mathrm{G}(t-1)\mathrm{A}^T\right)\mathrm{A}^T \\
&\overset{(f)}{\le} a\left(\mathrm{I} + \mathrm{A}\mathrm{S}(t)\mathrm{A}^T(1-p)\right) + b\mathrm{A}\mathrm{G}(t)\mathrm{A}^T \\
&\overset{(g)}{=} a\mathrm{S}(t+1) + b\mathrm{A}\mathrm{G}(t)\mathrm{A}^T,
\end{aligned}$$

where (a) follows from Lemma 1-(iv) and the induction hypothesis, (b) follows from Lemma 5, (c) follows from the bound on $\mathrm{S}(t)$ and Lemma 1-(iv), (d) also uses Lemma 5, (e) is obtained from Lemma 6 and the definition of b, and finally, (f) and (g) follow from the definitions of a, $\mathrm{S}(t)$, and $\mathrm{G}(t)$. ■

VI. Simulations

In this section we present simulations to illustrate the results. We used the following system in the simulations:

$$\mathrm{A} = \begin{bmatrix} 1.1 & 1 & 0 \\ 0 & 1.25 & 0 \\ 1 & 1 & 0.5 \end{bmatrix}, \quad \mathrm{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathrm{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathrm{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Clearly, A is unstable and (A, C) is detectable. We have $\rho(\mathrm{A}) = 1.25$, which means that the critical probability is $p_c = 1 - 1/\rho(\mathrm{A})^2 = 0.36$. Figure 1 shows the results of the simulations. The thick solid curve corresponds to the average trace of the covariance for $p = 0.38$. For this value of p, the Kalman filter is stable. The dotted curve shows the behavior of the error covariance for $p = 0.30$, i.e., a value for which the Kalman filter is unstable. The bound shown as a horizontal line corresponds to $\mathrm{trace}\left(a\theta_S \mathrm{I} + b\theta_G \mathrm{A}\mathrm{A}^T\right)$, i.e., the trace of the upper bound computed in Section V. The upper bound was computed with $p = 0.38$. The number of simulations used to compute the averages was 10^4. We can see from Figure 1 that the stable expected error covariance is less than the upper bound, for all t. As opposed to this, the average covariance of the unstable filter presents large peaks over the complete simulation interval, reaching above the upper bound.

VII. Conclusions

We have introduced a decomposition of PSD matrices, and used it to prove that the known necessary condition for the stability (in expectation) of the Kalman filter with Bernoulli observations is also sufficient, when the observation matrix, restricted to the observable subspace, defines an invertible linear operator. The decomposition was used to construct an upper bound on the expected Kalman filter covariance. It is worth mentioning that the bounds in Theorem 7 can be made tighter by replacing some of the uses of the trace by the spectral norm.

Acknowledgements The support by ARO Award W911NF-05-1-0219 is gratefully acknowledged. We thank the anonimous reviewers for their comments and suggestions.

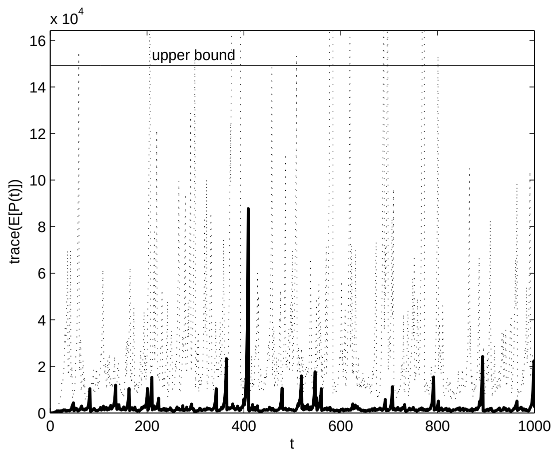

Fig. 1. Trace of expected error covariance for $p = 0.30$ (dotted), and $p = 0.38$ (thick solid). The trace of the upper bound for $p = 0.38$ is shown as an horizontal line and marked "upper bound."

REFERENCES

[1] 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*, vol. 49, no. 9, pp. 1453–63, 2004.

[2] X. Liu and A. J. Goldsmith, "Kalman filtering with partial observation losses," in *IEEE Conf. on Decision and Control*, Paradise Island, Bahamas, Dec. 2004, pp. 4180–4186.

[3] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.

[4] A. F. Dana, V. Gupta, J. Hespanha, R. M. Murray, and B. Hassibi, "Estimation over communication networks: Performance bounds and achievability results," in *American Control Conference*, New York, Jul. 2007, pp. 3450–3455.

[5] V. Gupta, A. F. Dana, J. P. Hespanha, and R. M. Murray, "Data transmission over networks for estimation," in *Mathematical Theory of Networks and Systems*, Kyoto, Japan, Jun. 2006.

[6] O. Imer and T. Başar, "Optimal estimation with scheduled measurements," *Journal of Applied and Computational Mathematics*, vol. 4, no. 2, pp. 1–11, 2005.

[7] M. Huang and S. Dey, "Stability of Kalman filtering with Markov packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

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

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

[10] R. Touri and C. N. Hadjicostis, "Stabilization with feedback control utilizing packet dropping network links," *IET (IEE) Proceedings on Control Theory and Applications*, vol. 1, no. 1, pp. 334–342, 2007.

[11] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd ed., ser. TAM. New York: Springer Verlag, 1998, vol. 6.

APPENDIX

We give here some results that are used in Section IV.

The following is a simple fact. We include its proof for completeness.

Lemma 8 (Kernels). *Let* $\mathrm{X} \in \mathbb{R}^{n_1 \times n_2}$ *and* $\mathrm{Y} \in \mathbb{R}^{n_3 \times n_2}$. *If* $\ker(\mathrm{X}) \subseteq \ker(\mathrm{Y})$, *then there exists* $\mathrm{L} \in \mathbb{R}^{n_3 \times n_1}$ *such that* $\mathrm{Y} = \mathrm{LX}$.

Proof: For each $i = 1, \ldots, n_1$, let $\mathrm{x}_i^T \in \mathbb{R}^{n_2}$ be the $i-$th row of X. Likewise, for each $j = 1, \ldots, n_3$, let $\mathrm{y}_j^T \in \mathbb{R}^{n_2}$ be the $j-$th row of Y. Assume that there exists y_j such that $\mathrm{y}_j \notin \operatorname{span}\{\mathrm{x}_1, \ldots, \mathrm{x}_{n_1}\}$. Then we can decompose y_j as $\mathrm{y}_j = \mathrm{y}_j^{||} + \mathrm{y}_j^{\perp}$, where $\mathrm{y}_j^{||} \in \operatorname{span}\{\mathrm{x}_1, \ldots, \mathrm{x}_{n_1}\}$, $\mathrm{y}_j^{\perp} \perp \operatorname{span}\{\mathrm{x}_1, \ldots, \mathrm{x}_{n_1}\}$, and $\mathrm{y}_j^{\perp} \neq 0$. Then $\mathrm{X}\mathrm{y}_j = 0$, but $\mathrm{Y}\mathrm{y}_j \neq 0$, contradicting $\ker(\mathrm{X}) \subseteq \ker(\mathrm{Y})$. This means that there exists $\mathrm{l}_j \in \mathbb{R}^{n_1}$, such that $\mathrm{y}_j = \mathrm{X}^T \mathrm{l}_j \in \operatorname{span}\{\mathrm{x}_1, \ldots, \mathrm{x}_{n_1}\}$. This is true for any $1 \leq j \leq m$, and so, $\mathrm{Y} = \mathrm{LX}$, where the $j-$th row of L is l_j^T. ■

Lemma 9 (A property of PSD matrices). *Let* $\mathrm{X} := \mathrm{Block}(\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_2^T, \mathrm{X}_4) \in \mathbb{S}^n_{\geq 0}$. *Then* $\mathrm{X}_2 = \mathrm{X}_1 \mathrm{L}$ *for some matrix* L.

Proof: Suppose there exists a vector v such that $\mathrm{X}_1 \mathrm{v} = 0$, but $\mathrm{X}_2^T \mathrm{v} \neq 0$. Then, letting $\mathrm{w} = [\mathrm{v}^T \ -(c\mathrm{X}_2^T \mathrm{v})^T]^T$, with $c \in \mathbb{R}$ to be chosen, we have $\mathrm{w}^T \mathrm{X} \mathrm{w} = -2c\mathrm{v}^T \mathrm{X}_2 \mathrm{X}_2^T \mathrm{v} + c^2 \mathrm{v}^T \mathrm{X}_2 \mathrm{X}_4 \mathrm{X}_2^T \mathrm{v}$. We can choose $c > 0$ small enough, such that $\mathrm{w}^T \mathrm{X} \mathrm{w} < 0$, contradicting the fact that $\mathrm{X} \geq 0$. This means that $\ker(\mathrm{X}_1) \subseteq \ker\left(\mathrm{X}_2^T\right)$. By Lemma 8, there exists a matrix $\bar{\mathrm{L}}$, such that $\mathrm{X}_2^T = \bar{\mathrm{L}} \mathrm{X}_1$. Letting $\mathrm{L} = \bar{\mathrm{L}}^T$, we have $\mathrm{X}_2 = \mathrm{X}_1 \mathrm{L}$. ■

Lemma 10 (Decomposition of PSD matrices). *Let* $\mathrm{X} \in \mathbb{S}^n_{\geq 0}$. *Let* $1 \leq m < n$. *Then there exist* $\mathrm{X}_1 \in \mathbb{S}^m_{\geq 0}$, $\mathrm{L} \in \mathbb{R}^{m \times (n-m)}$, *and* $\tilde{\mathrm{X}} \in \mathbb{S}^{(n-m)\times(n-m)}_{\geq 0}$, *such that*

$$\mathrm{X} = \mathrm{Block}(\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_2^T, \mathrm{X}_4) = \mathrm{Block}(\mathrm{X}_1, \mathrm{X}_1\mathrm{L}, \mathrm{L}^T\mathrm{X}_1, \mathrm{L}^T\mathrm{X}_1\mathrm{L}) + \mathrm{Block}(0, 0, 0, \tilde{\mathrm{X}}),$$

where $\tilde{\mathrm{X}} := \mathrm{X}_4 - \mathrm{L}^T \mathrm{X}_1 \mathrm{L} \in \mathbb{S}^{(n-m)}_{\geq 0}$. *This decomposition is unique.*

Proof: From Lemma 9, we know that there exists L such that $\mathrm{X}_2 = \mathrm{X}_1 \mathrm{L}$. Assume there exist a vector v, such that $\mathrm{v}^T \tilde{\mathrm{X}} \mathrm{v} < 0$. Then, letting $\mathrm{w} = [-(\mathrm{Lv})^T \ \mathrm{v}^T]^T$, we have that $\mathrm{w}^T \mathrm{X} \mathrm{w} = \mathrm{v}^T \tilde{\mathrm{X}} \mathrm{v} < 0$, which contradicts $\mathrm{X} \geq 0$. To prove uniqueness, we notice that, if $\mathrm{X}_1 \mathrm{L} = \mathrm{X}_1 \bar{\mathrm{L}} = \mathrm{X}_2$, then $\mathrm{L} = \bar{\mathrm{L}} + \tilde{\mathrm{L}}$, with $\tilde{\mathrm{L}} \in \ker(\mathrm{X}_1)$. This means that $\mathrm{L}^T \mathrm{X}_1 \mathrm{L} = \bar{\mathrm{L}}^T \mathrm{X}_1 \bar{\mathrm{L}}$, and so, L and $\bar{\mathrm{L}}$ produce the same decomposition. ■