# Kalman Filtering with Intermittent Observations

Bruno Sinopoli, Luca Schenato, Massimo Franceschetti,

Kameshwar Poolla, Michael I. Jordan, Shankar S. Sastry

Department of Electrical Engineering and Computer Sciences

University of California at Berkeley

*{sinopoli, massimof, lusche, sastry}@eecs.berkeley.edu*

*poolla@me.berkeley.edu, jordan@cs.berkeley.edu*

### Abstract

Motivated by navigation and tracking applications within sensor networks, we consider the problem of performing Kalman filtering with intermittent observations. When data travel along unreliable communication channels in a large, wireless, multi-hop sensor network, the effect of communication delays and loss of information in the control loop cannot be neglected. We address this problem starting from the discrete Kalman filtering formulation, and modelling the arrival of the observation as a random process. We study the statistical convergence properties of the estimation error covariance, showing the existence of a critical value for the arrival rate of the observations, beyond which a transition to an unbounded state error covariance occurs. We also give upper and lower bounds on this expected state error covariance.

## I. INTRODUCTION

Advances in VLSI and MEMS technology have boosted the development of micro sensor integrated systems. Such systems combine computing, storage, radio technology, and energy source on a single chip [1] [2]. When distributed over a wide area, networks of sensors can perform a variety of tasks that range from environmental monitoring and military surveillance, to navigation and control of a moving vehicle [3] [4] [5]. A common feature of these systems is the presence of significant communication delays and data loss across the network. From the point of view of control theory, significant delay is equivalent to loss, as data needs to arrive to its destination in time to be used for control. In short, communication and control become tightly coupled such that the two issues cannot be addressed independently.

Consider, for example, the problem of navigating a vehicle based on the estimate from a sensor web of its current position and velocity. The measurements underlying this estimate can be lost or delayed due to the unreliability of the wireless links. What is the amount of data loss that the control loop can tolerate to reliably perform the navigation task? Can communication protocols be designed to satisfy this constraint? At Berkeley, we have faced these kind of questions in building sensor networks for pursuit evasion games. Practical advances in the design of these systems are described in [6]. The goal of this paper is to examine some control-theoretic implications of using sensor networks for control. These require a generalization of classical control techniques that explicitly take into account the stochastic nature of the communication channel.

In our setting, the sensor network provides observed data that are used to estimate the state of a controlled system, and this estimate is then used for control. We study the effect of data losses due to the unreliability of the network links. We generalize the most ubiquitous recursive estimation technique in control—the discrete Kalman filter [7]—modelling the arrival of an observation as a random process whose parameters are related to the characteristics of the communication channel, see Figure 1. We characterize the statistical convergence of the expected estimation error covariance in this setting.

The classical theory relies on several assumptions that guarantee convergence of the Kalman filter. Consider the following discrete time linear dynamical system:

$$
\begin{aligned}
x_{t+1} &= Ax_t + w_t \\
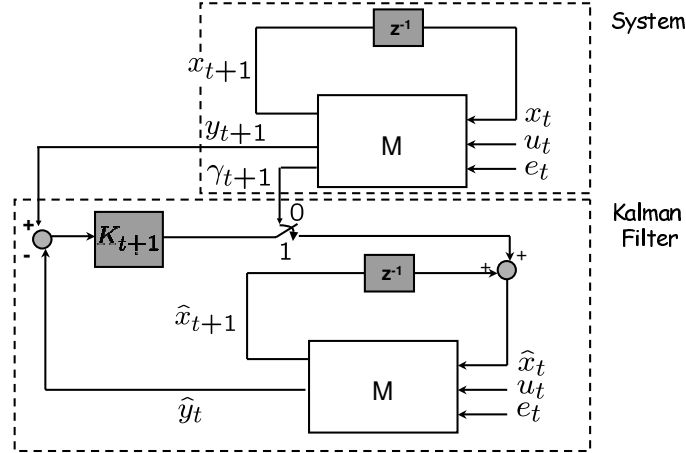y_t &= Cx_t + v_t,
\end{aligned}
\tag{1}
$$

Fig. 1. **Overview of the system.** We study the statistical convergence of the expected estimation error covariance of the discrete Kalman filter, where the observation, travelling over an unreliable communication channel, can be lost at each time step with probability $1 - \lambda$.

where $x_t \in \Re^n$ is the state vector, $y_t \in \Re^m$ the output vector, $w_t \in \Re^p$ and $v_t \in \Re^m$ are Gaussian random vectors with zero mean and covariance matrices $Q \geq 0$ and $R > 0$, respectively. $w_t$ is independent of $w_s$ for $s < t$. Assume that the initial state, $x_0$, is also a Gaussian vector of zero mean and covariance $\Sigma_0$. Under the hypothesis of stabilizability of the pair $(A, Q)$ and detectability of the pair $(A, C)$, the estimation error covariance of the Kalman filter converges to a unique value from any initial condition [8].

These assumptions have been relaxed in various ways [8]. Extended Kalman filtering attempts to cope with nonlinearities in the model; particle filtering is also appropriate for nonlinear models, and additionally does not require the noise model to be Gaussian. Recently, more general observation processes have been studied. In particular, in [9], [10] the case in which observations are randomly spaced in time according to a Poisson process has been studied, where the underlying dynamics evolve in continuous time. These authors showed the existence of a lower bound on the arrival rate of the observations below which it is possible to maintain the estimation error covariance below a fixed value, with high probability. The results were restricted to scalar SISO systems.

We approach a similar problem within the framework of discrete time, and provide results for general $n$-dimensional MIMO systems. In particular, we consider a discrete-time system in which the arrival of an observation is a Bernoulli process with parameter $0 < \lambda < 1$, and, rather than asking for the estimation error covariance to be bounded with high probability, we study the asymptotic behavior (in time) of its average. Our main contribution is to show that, depending on the eigenvalues of the matrix $A$, and on the structure of the matrix $C$, there exists a critical value $\lambda_c$, such that if the probability of arrival of an observation at time $t$ is $\lambda > \lambda_c$, then the expectation of the estimation error covariance is always finite (provided that the usual stabilizability and detectability hypotheses are satisfied). If $\lambda \leq \lambda_c$, then the expectation of the estimation error covariance tends to infinity. We give explicit upper and lower bounds on $\lambda_c$, and show that they are tight in some special cases.

Philosophically this result can be seen as another manifestation of the well known *uncertainty threshold principle* [11], [12]. This principle states that optimum long-range control of a dynamical system with uncertain parameters is possible if and only if the uncertainty does not exceed a given threshold. The uncertainty is modelled as white noise scalar sequences acting on the system and control matrices. In our case, the result applies to optimal estimation, rather than optimal control, and the uncertainty is due to the random arrival of the observation, with the randomness arising from losses in the network.

We can also relate our approach to the theory of jump linear systems [13]. Jump linear systems (JLS) are stochastic hybrid systems characterized by linear dynamics and discrete regime transitions modelled as Markov chains. In the work of Nilsson et al. [14], [15] the Kalman filter with missing observations is modelled as a JLS switching between two discrete regimes: an open loop configuration and a closed loop one. Within this framework these authors obtain a critical loss probability for the divergence of the expected estimation error covariance. However, their JLS formulation is restricted to the steady state Kalman Filter, where the Kalman gain is constant. The resulting process is wide sense stationary [16], and this makes the exact computation of the transition probability and state error covariance possible. Instead, we consider the general case of time-varying Kalman gain, where jump linear systems theory cannot be directly applied. We also show that the resulting filter can tolerate a higher dropping rate than the one obtained with a stationary filter modeled with the JLS approach. In fact, the time-varying Kalman filter is optimal, in the sense that it minimizes the state error covariance, unlike its steady state counterpart.

Considering tracking applications in cluttered environments, Fortmann et al. [17] also study the case of a dynamic Kalman filter with missing or false observations, deriving stochastic equations for the state covariance error. They do not, however, characterize its convergence statistically, providing only numerical evidence of the transition to instability, and leaving a formal characterization of this transition as an open problem—a characterization which is provided in this paper.

Finally, we point out that our problem can also be viewed from the perspective of the Expectation-Maximization (EM) algorithm. EM provides a general framework for finding maximum likelihood estimates in missing-data models [18]. The state-space model underlying Kalman filtering is an instance of a missing-data model, and the E step of the EM algorithm coincides with Kalman filtering and smoothing in this case [19]. More generally, a wide variety of generalizations of state-space models can be unified under the framework of probabilistic graphical models, and Lauritzen [20] shows how to derive EM algorithms for general graphical models. Our case, however, lies beyond Lauritzen's framework, in that the graph structure is a function of the missing data—there is one graph for each pattern of missing data. More fundamentally, EM provides a framework for computing a posteriori parameter estimates; it does not provide the a priori bounds on the state covariance error that are our focus here.

The paper is organized as follows. In section II we formalize the problem of Kalman filtering with intermittent observations. In section III we provide upper and lower bounds on the expected estimation error covariance of the Kalman filter, and find the conditions on the observation arrival probability $\lambda$ for which the upper bound converges to a fixed point, and for which the lower bound diverges. Section IV describes some special cases and provides an intuitive understanding of the results. In section V we compare our approach to previous approaches [14] based on jump linear systems. Finally, in section VI, we state our conclusions and give directions for future work.

## II. PROBLEM FORMULATION

Consider the canonical state estimation problem. We define the arrival of the observation at time $t$ as a binary random variable $\gamma_t$, with probability distribution $p_{\gamma_t}(1) = \lambda$, and with $\gamma_t$ independent of $\gamma_s$ if $t \neq s$. The output noise $v_t$ is defined in the following way:

$$p(v_t|\gamma_t) = \begin{cases} \mathcal{N}(0, R) & : \quad \gamma_t = 1 \\ \mathcal{N}(0, \sigma^2 I) & : \quad \gamma_t = 0, \end{cases}$$

for some $\sigma^2$. Therefore, the variance of the observation at time $t$ is $R$ if $\gamma_t$ is 1, and $\sigma^2 I$ otherwise. In reality the absence of observation corresponds to the limiting case of $\sigma \to \infty$. Our approach is to re-derive the Kalman filter equations using a "dummy" observation with a given variance when the real observation does not arrive, and then take the limit as $\sigma \to \infty$.

First let us define:

$$\hat{x}_{t|t} \triangleq \mathbb{E}[x_t|\mathbf{y_t}, \gamma_{\mathbf{t}}] \tag{2}$$

$$P_{t|t} \triangleq \mathbb{E}[(x_t - \hat{x})(x_t - \hat{x})'|\mathbf{y_t}, \gamma_{\mathbf{t}}] \tag{3}$$

$$\hat{x}_{t+1|t} \triangleq \mathbb{E}[x_{t+1}|\mathbf{y_t}, \gamma_{\mathbf{t+1}}] \tag{4}$$

$$P_{t+1|t} \triangleq \mathbb{E}[(x_{t+1} - \hat{x_{t+1}})(x_{t+1} - \hat{x_{t+1}})'|\mathbf{y_t}, \gamma_{\mathbf{t+1}}] \tag{5}$$

$$\hat{y}_{t+1|t} \triangleq \mathbb{E}[y_{t+1}|\mathbf{y_t}, \gamma_{\mathbf{t+1}}], \tag{6}$$

where we have defined the vectors $\mathbf{y_t} \triangleq [y_0, \ldots, y_t]'$ and $\gamma_{\mathbf{t}} \triangleq [\gamma_0, \ldots, \gamma_t]'$. Using the Dirac delta $\delta(\cdot)$ we have:

$$\mathbb{E}[(y_{t+1} - \hat{y}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})'|\mathbf{y_t}, \gamma_{\mathbf{t+1}}] = CP_{t+1|t} \tag{7}$$

$$\mathbb{E}[(y_{t+1} - \hat{y}_{t+1|t})(y_{t+1} - \hat{y}_{t+1|t})'|\mathbf{y_t}, \gamma_{\mathbf{t+1}}] = CP_{t+1|t}C' + \delta(\gamma_{t+1} - 1)R + \delta(\gamma_{t+1})\sigma^2 I, \tag{8}$$

and it follows that the random variables $x_{t+1}$ and $y_{t+1}$, conditioned on the output $\mathbf{y_t}$ and on the arrivals $\gamma_{\mathbf{t+1}}$, are jointly gaussian with mean

$$\mathbb{E}[x_{t+1}, y_{t+1}|\mathbf{y_t}, \gamma_{\mathbf{t+1}}] = \begin{pmatrix} \hat{x}_{t+1|t} \\ C\hat{x}_{t+1|t} \end{pmatrix},$$

and covariance

$$COV(x_{t+1}, y_{t+1}|\mathbf{y_t}, \gamma_{\mathbf{t+1}}) = \begin{pmatrix} P_{t+1|t} & P_{t+1|t}C' \\ CP_{t+1|t} & CP_{t+1|t}C' + \delta(\gamma_{t+1} - 1)R + \delta(\gamma_{t+1})\sigma^2 I \end{pmatrix}. \tag{9}$$

Hence, the Kalman filter equations are modified as follows:

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} \tag{10}$$

$$P_{t+1|t} = AP_{t|t}A' + Q \tag{11}$$

$$\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + P_{t+1|t}C'(CP_{t+1|t}C' + \delta(\gamma_{t+1} - 1)R + \delta(\gamma_{t+1})\sigma^2 I)^{-1}(y_{t+1} - C\hat{x}_{t+1|t}) \tag{12}$$

$$P_{t+1|t+1} = P_{t+1|t} - P_{t+1|t}C'(CP_{t+1|t}C' + \delta(\gamma_{t+1} - 1)R + \delta(\gamma_{t+1})\sigma^2 I)^{-1}CP_{t+1|t}. \tag{13}$$

Taking the limit as $\sigma \to \infty$, the update equations (12) and (13) can be rewritten as follows:

$$\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + \gamma_{t+1}P_{t+1|t}C'(CP_{t+1|t}C' + R)^{-1}(y_{t+1} - C\hat{x}_{t+1|t}) \tag{14}$$

$$P_{t+1|t+1} = P_{t+1|t} - \gamma_{t+1}P_{t+1|t}C'(CP_{t+1|t}C' + R)^{-1}CP_{t+1|t}. \tag{15}$$

Note that performing this limit corresponds *exactly* to propagating the previous state when there is no observation update available at time t. We also point out the main difference from the standard Kalman filter formulation: Both $\hat{x}_{t+1|t+1}$ and $P_{t+1|t+1}$ are now random variables, being a function of $\gamma_{t+1}$, which is itself random.

Given the new formulation, we now study the Riccati equation of the state error covariance matrix in this generalized setting and provide deterministic upper and lower bounds on its expectation. We then characterize the convergence of these upper and lower bounds, as a function of the arrival probability $\lambda$ of the observation.

## III. CONVERGENCE CONDITIONS AND TRANSITION TO INSTABILITY

It is easy to verify that the modified Kalman filter formulation in Equations (11) and (15) can be rewritten as follows:

$$P_{t+1} = AP_tA' + Q - \gamma_t AP_tC'(CP_tC' + R)^{-1}CP_tA', \tag{16}$$

where we use the simplified notation $P_t = P_{t|t-1}$. Since the sequence $\{\gamma_t\}_0^\infty$ is random, the modified Kalman filter iteration is stochastic and cannot be determined off-line. Therefore, only statistical properties can be deduced. In this section we show the existence of a critical value $\lambda_c$ for the arrival probability of the observation update, such that for $\lambda > \lambda_c$ the mean state covariance $\mathbb{E}[P_t]$ is bounded for all initial conditions, and for $\lambda \le \lambda_c$ the mean state covariance diverges for some initial condition. We also find a lower bound $\underline{\lambda}$, and upper bound $\overline{\lambda}$, for the critical probability $\lambda_c$, i.e., $\underline{\lambda} \le \lambda_c \le \overline{\lambda}$. The lower bound is expressed in closed form; the upper bound is the solution of a linear matrix inequality (LMI). In some special cases the two bounds coincide, giving a tight estimate. Finally, we present numerical algorithms to compute a lower bound $\bar{S}$, and upper bound $\bar{V}$, for $\lim_{t\to\infty}\mathbb{E}[P_t]$, when it is bounded.

First, we define the modified algebraic Riccati equation (MARE) for the Kalman filter with intermittent observations as follows,

$$g_\lambda(X) = AXA' + Q - \lambda AXC'(CXC' + R)^{-1}CXA'. \tag{17}$$

Our results derive from two principal facts: the first is that concavity of the modified algebraic Riccati equation for our filter with intermittent observations allows use of Jensen's inequality to find an upper bound on the mean state covariance; the second is that all the operators we use to estimate upper and lower bounds are monotonically increasing, therefore if a fixed point exists, it is also stable.

We formally state all main results in form of theorems. Omitted proofs appear in the Appendix. The first theorem expresses convergence properties of the MARE.

**Theorem 1.** *Consider the operator* $\phi(K, X) = (1 - \lambda)(AXA' + Q) + \lambda(FXF' + V)$, *where* $F = A + KC$, $V = Q + KRK'$. *Suppose there exists a matrix* $\tilde{K}$ *and a positive definite matrix* $\tilde{P}$ *such that*

$$\tilde{P} > 0 \quad and \quad \tilde{P} > \phi(\tilde{K}, \tilde{P})$$

*Then,*

*(a) for any initial condition* $P_0 \geq 0$, *the MARE converges, and the limit is independent of the initial condition:*

$$\lim_{t \to \infty} P_t = \lim_{t \to \infty} g_\lambda^t(P_0) = \overline{P}$$

*(b)* $\overline{P}$ *is the unique positive semidefinite fixed point of the MARE.*

The next theorem states the existence of a sharp transition.

**Theorem 2.** *If* $(A, Q^{\frac{1}{2}})$ *is controllable,* $(A, C)$ *is detectable, and* $A$ *is unstable, then there exists a* $\lambda_c \in [0, 1)$ *such that*

$$\lim_{t \to \infty} \mathbb{E}[P_t] = +\infty \quad for\ 0 \leq \lambda \leq \lambda_c\ and\ some\ initial\ condition\ P_0 \geq 0 \tag{18}$$

$$\mathbb{E}[P_t] \leq M_{P_0} \ \forall t \quad for\ \lambda_c < \lambda \leq 1\ and\ any\ initial\ condition\ P_0 \geq 0 \tag{19}$$

*where* $M_{P_0} > 0$ *depends on the initial condition* $P_0 \geq 0$[1].

The next theorem gives upper and lower bounds for the critical probability $\lambda_c$.

---

[1]We use the notation $\lim_{t \to \infty} A_t = +\infty$ when the sequence $A_t \geq 0$ is not bounded; i.e., there is no matrix $M \geq 0$ such that $A_t \leq M, \forall t$.

**Theorem 3.** *Let*

$$\underline{\lambda} \;=\; arginf_\lambda[\exists \hat{S} \mid \hat{S} = (1-\lambda)A\hat{S}A' + Q] = 1 - \frac{1}{\alpha^2} \tag{20}$$

$$\overline{\lambda} \;=\; arginf_\lambda[\exists(\hat{K}, \hat{X}) \mid \hat{X} > \phi(\hat{K}, \hat{X})] \tag{21}$$

*where $\alpha = \max_i |\sigma_i|$ and $\sigma_i$ are the eigenvalues of $A$. Then*

$$\underline{\lambda} \leq \lambda_c \leq \overline{\lambda}. \tag{22}$$

Finally, the following theorem gives an estimate of the limit of the mean covariance matrix $\mathbb{E}[P_t]$, when this is bounded.

**Theorem 4.** *Assume that $(A, Q^{\frac{1}{2}})$ is controllable, $(A, C)$ is detectable and $\lambda > \overline{\lambda}$, where $\overline{\lambda}$ is defined in Theorem 4. Then*

$$0 \leq \bar{S} \leq \lim_{t \to \infty} \mathbb{E}[P_t] \leq \bar{V} \qquad \forall\, \mathbb{E}[P_0] \geq 0 \tag{23}$$

*where $\bar{S} = (1-\lambda)A\bar{S}A' + Q$ and $\bar{V} = g_\lambda(\bar{V})$.*

The previous theorems give lower and upper bounds for both the critical probability $\lambda_c$ and for the mean error covariance $\mathbb{E}[P_t]$. The lower bound $\underline{\lambda}$ is expressed in closed form. We now resort to numerical algorithms for the computation of the remaining bounds $\overline{\lambda}, \bar{S}$ and $\bar{V}$.

The computation of the upper bound $\overline{\lambda}$ can be reformulated as the iteration of an LMI feasibility problem. To establish this we need the following theorem:

**Theorem 5.** *If $(A, Q^{\frac{1}{2}})$ is controllable and $(A, C)$ is detectable, then the following statements are equivalent:*

*(a) $\exists \bar{X}$   such that   $\bar{X} > g_\lambda(\bar{X})$*

*(b) $\exists \bar{K}, \bar{X} > 0$   such that   $\bar{X} > \phi(\bar{K}, \bar{X})$*

*(c) $\exists \bar{Z}$ and $0 < \bar{Y} \leq I$ such that*

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

*Proof:*   (a)$\Longrightarrow$(b) If $\bar{X} > g_\lambda(\bar{X})$ exists, then $\bar{X} > 0$ by Lemma 1(g). Let $\bar{K} = K_{\bar{X}}$. Then $\bar{X} > g_\lambda(\bar{X}) = \phi(\bar{K}, \bar{X})$ which proves the statement. (b)$\Longrightarrow$(a) Clearly $\bar{X} > \phi(\bar{K}, \bar{X}) \geq g_\lambda(\bar{X})$

which proves the statement. (b)$\Longleftrightarrow$(c) Let $F = A + KC$, then:

$$X > (1 - \lambda)AXA' + \lambda FXF' + Q + \lambda KRK'$$

is equivalent to

$$\begin{bmatrix} X - (1 - \lambda)AXA' & \sqrt{\lambda}F \\ \sqrt{\lambda}F' & X^{-1} \end{bmatrix} > 0,$$

where we used the Schur complement decomposition and the fact that $X - (1 - \lambda)AXA' \geq \lambda FXF' + Q + \lambda KRK' \geq Q > 0$. Using one more time the Schur complement decomposition on the first element of the matrix we obtain

$$\Theta = \begin{bmatrix} X & \sqrt{\lambda}F & \sqrt{1 - \lambda}A \\ \sqrt{\lambda}F' & X^{-1} & 0 \\ \sqrt{1 - \lambda}A' & 0 & X^{-1} \end{bmatrix} > 0.$$

This is equivalent to

$$\Lambda = \begin{bmatrix} X^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \Theta \begin{bmatrix} X^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} > 0$$

$$= \begin{bmatrix} X^{-1} & \sqrt{\lambda}X^{-1}F & \sqrt{1 - \lambda}X^{-1}A \\ \sqrt{\lambda}F'X^{-1} & X^{-1} & 0 \\ \sqrt{1 - \lambda}A'X^{-1} & 0 & X^{-1} \end{bmatrix} > 0.$$

Let us consider the change of variable $Y = X^{-1} > 0$ and $Z = X^{-1}K$, in which case the previous LMI is equivalent to:

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

Since $\Psi(\alpha Y, \alpha K) = \alpha \Psi(Y, K)$, then $Y$ can be restricted to $Y \leq I$, which completes the theorem.

∎

Combining theorems 3 and 5 we immediately have the following corollary

**Corollary 1.** *The upper bound $\bar{\lambda}$ is given by the solution of the following optimization problem,*

$$\bar{\lambda} = \text{argmin}_\lambda \Psi(Y, Z) > 0, \quad 0 \leq Y \leq I.$$

This is a quasi-convex optimization problem in the variables $(\lambda, Y, Z)$ and the solution can be obtained by iterating LMI feasibility problems and using bisection for the variable $\lambda$.

The lower bound $\bar{S}$ for the mean covariance matrix can be easily obtained via standard Lyapunov Equation solvers. The upper bound $\bar{V}$ can be found by iterating the MARE or by solving an semidefinite programming (SDP) problem as shown in the following.

**Theorem 6.** *If $\lambda > \bar{\lambda}$, then the matrix $\bar{V} = g_\lambda(V)$ is given by:*

*(a) $\bar{V} = \lim_{t \to \infty} V_t$; $V_{t+1} = g_\lambda V_t$ where $V_0 \geq 0$*

*(b)*

$$\text{argmax}_V \quad Trace(V)$$

$$\textit{subject to} \quad \begin{bmatrix} AVA' - V & \sqrt{\lambda}AVC' \\ \sqrt{\lambda}CVA' & CVC' + R \end{bmatrix} \geq 0, \quad V \geq 0$$

*Proof:* (a) It follows directly from Theorem 1.

(b) It can be obtained by using the Schur complement decomposition on the equation $V \leq g_\lambda(V)$. Clearly the solution $\bar{V} = g_\lambda(\bar{V})$ belongs to the feasible set of the optimization problem. We now show that the solution of the optimization problem is the fixed point of the MARE. Suppose it is not, i.e., $\hat{V}$ solves the optimization problem but $\hat{V} \neq g_\lambda(\hat{V})$. Since $\hat{V}$ is a feasible point of the optimization problem, then $\hat{V} < g_\lambda(\hat{V}) = \hat{\hat{V}}$. However, this implies that $Trace(\hat{V}) < Trace(\hat{\hat{V}})$, which contradicts the hypothesis of optimality of matrix $\hat{V}$. Therefore $\hat{V} = g_\lambda(\hat{V})$ and this concludes the theorem. ∎

## IV. SPECIAL CASES AND EXAMPLES

In this section we present some special cases in which upper and lower bounds on the critical value $\lambda_c$ coincide and give some examples. From Theorem 1, it follows that if there exists a $\tilde{K}$ such that $F$ is the zero matrix, then the convergence condition of the MARE is for $\lambda > \lambda_c = 1 - 1/\alpha^2$, where $\alpha = \max_i |\sigma_i|$, and $\sigma_i$ are the eigenvalues of $A$.

- **C is invertible**. In this case a choice of $K = -AC^{-1}$ makes $F = 0$. Note that the scalar case also falls under this category. Figure (1) shows a plot of the steady state of the upper
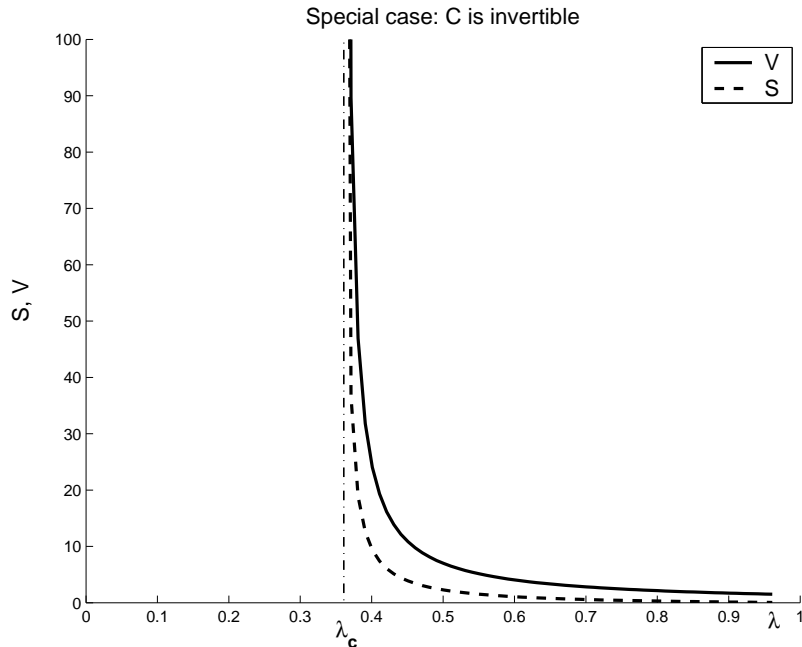
Fig. 2. Example of transition to instability in the scalar case. The dashed line shows the asymptotic value of the lower bound $(\bar{S})$, the solid line the asymptotic value of the upper bound $(\bar{V})$, and the dash-dot line shows the asymptote $(\lambda_c)$.

and lower bounds versus $\lambda$ in the scalar case. The discrete time LTI system used in this simulation has $A = -1.25$, $C = 1$, with $v_t$ and $w_t$ having zero mean and variance $R = 2.5$ and $Q = 1$, respectively. For this system we have $\lambda_c = 0.36$. The transition clearly appears in the figure, where we see that the steady state value of both upper and lower bound tends to infinity as $\lambda$ approaches $\lambda_c$. The dashed line shows the lower bound, the solid line the upper bound, and the dash-dot line shows the asymptote.

- **A has a single unstable eigenvalue**. In this case, regardless of the dimension of $C$ (and as long as the pair $(A, C)$ is detectable), we can use Kalman decomposition to bring to zero the unstable part of $F$ and thereby obtain tight bounds. Figure (2) shows a plot for the system $A = \begin{pmatrix} 1.25 & 1 & 0 \\ 0 & .9 & 7 \\ 0 & 0 & .60 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 2 \end{pmatrix}$

with $v_t$ and $w_t$ having zero mean and variance $R = 2.5$ and $Q = 20 * I_{3x3}$, respectively. This time, the asymptotic value for trace of upper and lower bound is plotted versus $\lambda$. Once again $\lambda_c = 0.36$.
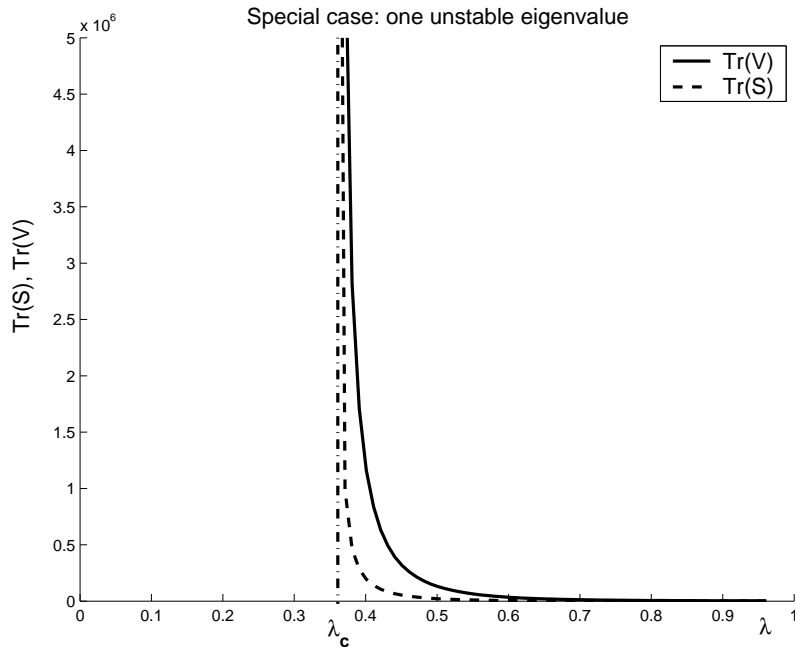
Fig. 3. Example of transition to instability with a single unstable eigenvalue in the MIMO case. The dashed line shows the asymptotic value of the trace of lower bound ($\bar{S}$), the solid line the asymptotic value of trace of the upper bound ($\bar{V}$), and the dash-dot line shows the asymptote ($\lambda_c$).

In general $F$ cannot always be made zero and we have shown that while a lower bound on $\lambda_c$ can be written in closed form, an upper bound on $\lambda_c$ is the result of a LMI. Figure (3) shows an example where upper and lower bounds have different convergence conditions. The system used for this simulation is $A = \begin{pmatrix} 1.25 & 0 \\ 1 & 1.1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \end{pmatrix}$ with $v_t$ and $w_t$ having zero mean and variance $R = 2.5$ and $Q = 20 * I_{2x2}$, respectively.

Finally, in Figure (4) we report results of another experiment, plotting the state estimation error of another system at two similar values of $\lambda$, one being below and one above the critical value. We note a dramatic change in the error at $\lambda_c \approx 0.125$. The figure on the left shows the estimation error with $\lambda = 0.1$. The figure on the right shows the estimation error for the same system evolution with $\lambda = 0.15$. In the first case the estimation error grows dramatically, making it practically useless for control purposes. In the second case, a small increase in $\lambda$ reduces the estimation error by approximately three orders of magnitude.
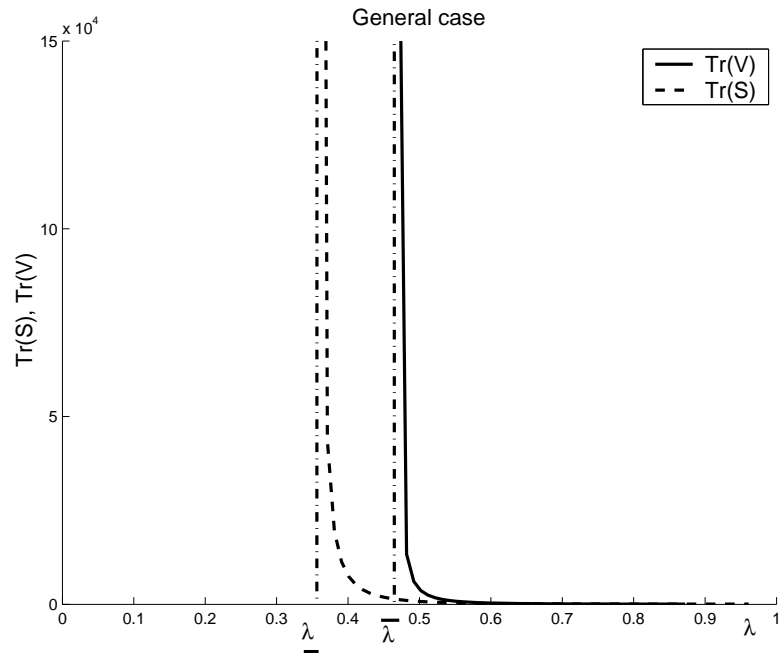
Fig. 4. Transition to instability in the general case, with arbitrary A and C. In this case lower and upper bounds do not have the same asymptote.
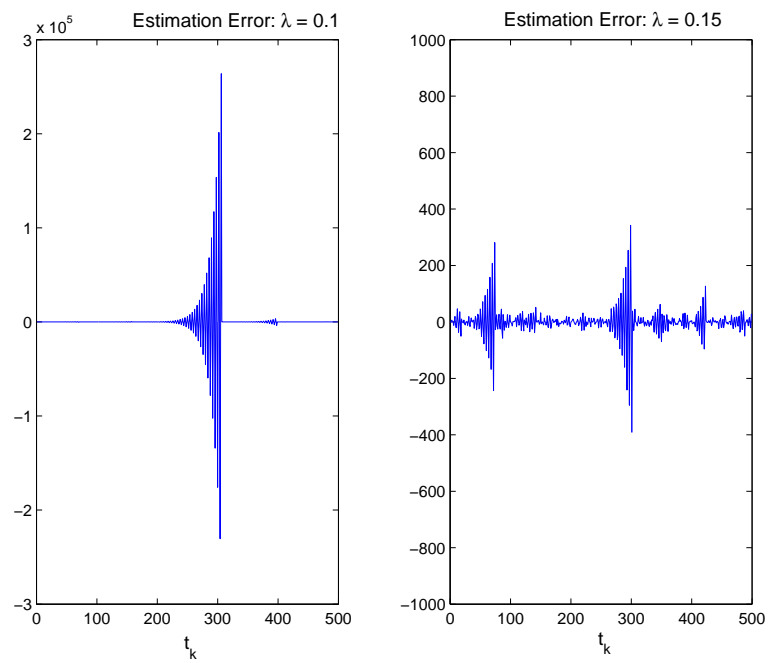


Fig. 5. Estimation error for $\lambda$ below (left) and above (right) the critical value

## V. STATIC VERSUS OPTIMAL DYNAMIC KALMAN GAIN

In this section we compare the performance of filtering with static and dynamic gains for a scalar discrete system. For the static estimator we follow the jump linear system approach of [14]. The scalar static estimator case has been also worked out in [21].

Consider the dynamic state estimator

$$
\begin{aligned}
\hat{x}_{t+1}^d &= A\hat{x}_t^d + \gamma_t K_t^d(y_t - \hat{y}_t) \\
K_t^d &= AP_tC'(CP_tC' + R)^{-1} \\
P_{t+1} &= AP_tA' + Q - \gamma_t AP_tC'(CP_tC' + R)^{-1}CP_tA'
\end{aligned}
\tag{24}
$$

where the Kalman gain $K_t^d$ is time-varying. Also consider the static state estimator

$$
\hat{x}_{t+1}^s = A\hat{x}_t^d + \gamma_t K_s(y_t - \hat{y}_t)
\tag{25}
$$

where the estimator gain $K_s$ is constant. If no data arrives, i.e. $\gamma_t = 0$, both estimators simply propagate the state estimate of the previous time-step.

The performance of the dynamic state estimator (24) has been analyzed in the previous sections. The performance of static state estimator (25) can be readily obtained using jump linear system theory [13] [14]. To do so, let us consider the estimator error $e_{t+1}^s \overset{\Delta}{=} x_{t+1} - \hat{x}_{t+1}^s$. Substituting Equations (1) for $x_{t+1}$ and (25) for $\hat{x}_{t+1}^s$, we obtain the dynamics of the estimation error:

$$
e_{t+1}^s = (A - \gamma_t K_s C)e_t^s + v_t + \gamma_t K_s w_t.
\tag{26}
$$

Using the same notation of Chapter 6 in Nilsson [14], where he considers the general system:

$$
z_{k+1} = \Phi(r_k)z_k + \Gamma(r_k)e_k
$$

the system (26) can be seen as jump linear system switching between two states $r_k \in \{1, 2\}$ given by:

$$
\begin{aligned}
\Phi(1) &= A - K_sC \qquad \Gamma(1) = [1 \; K_s] \\
\Phi(2) &= A \qquad\qquad\;\; \Gamma(2) = [1 \; 0]
\end{aligned}
$$

where the noise covariance $\mathbb{E}[e_k e_k'] = R_e$, the transition probability matrix $Q_\pi$ and the steady state probability distribution $\pi^\infty$ are given by:

$$R_e = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \qquad Q_\pi = \begin{bmatrix} \lambda & 1-\lambda \\ \lambda & 1-\lambda \end{bmatrix} \qquad \pi^\infty = \begin{bmatrix} \lambda & 1-\lambda \end{bmatrix}$$

Following the methodology proposed in Nilsson [14], it is possible to show that the system above is mean square stable, i.e., $\lim_{t\to\infty} \mathbb{E}[(e_t^s)' e_t^s] = 0$ if and only if the transition probability is

$$\lambda < \lambda_s = \frac{1}{1 - \left(1 - \frac{K_s C}{A}\right)^2} \left(1 - \frac{1}{A^2}\right). \tag{27}$$

If the system is mean square stable, the steady state error covariance $P_\infty^s = \lim_{t\to\infty} \mathbb{E}[e_t^s (e_t^s)']$ is given by:

$$P_\infty^s = \frac{Q + K_s^2 R}{1 - \lambda(A - K_s C)^2 - (1-\lambda)A^2}. \tag{28}$$

Calculations to obtain Equations (27) and (28) are tedious but straightforward, therefore they are omitted.

It is immediately evident that the transition probability $\lambda_s$ of the estimator (25) using a static gain is always greater than the transition probability $\lambda_c$ of the estimator (24) which uses a dynamic gain, in fact

$$\lambda_s = \lambda_c \frac{1}{1 - \left(1 - \frac{K_s C}{A}\right)^2}$$

with equality only in the case when $K_s = \frac{A}{C}$. A natural choice for the estimator gain $K_s$ is the steady state Kalman gain for the closed loop system ($r = 1$), which is always different from $\frac{A}{C}$. Figure 6 shows the steady state error covariance for the scalar system considered in the previous section, where $A = -1.5$, $C = 1$, $Q = 1$, $R = 2.5$. The steady state Kalman gain for this system is $K_{SSKalman} = -0.70$, and the gain for largest mean square stability range is $K_s = \frac{A}{C} = -1.25$. The figure also displays the upper bound of the state error covariance $\bar{V}$ for the dynamic estimator (24) that can be computed as indicated in Theorem 6. The steady state error covariance of the static predictor for the two different gains is always greater than our upper bound $\bar{V}$. This is not surprising, since the dynamic estimator is optimal over all possible estimators as shown in Section II. It is interesting to note that the static predictor with steady state Kalman gain is close to the upper bound of the optimal predictor for arrival probability
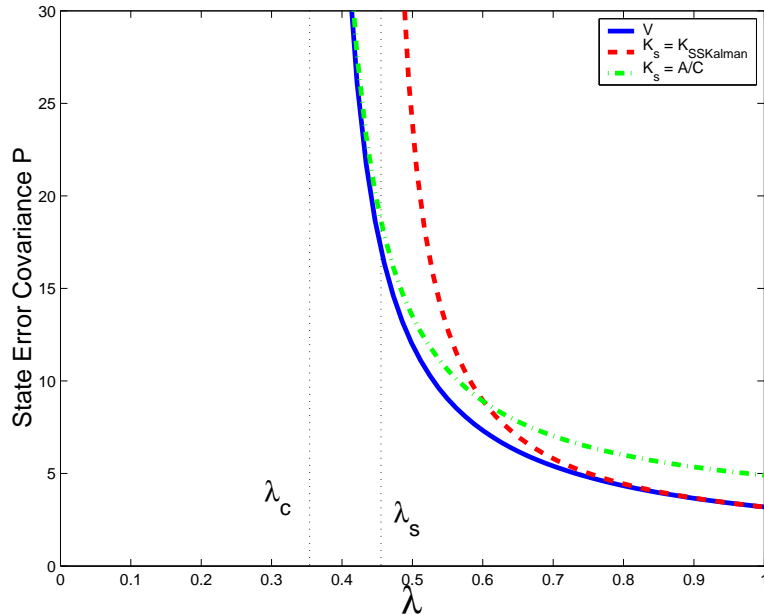
Fig. 6. Error covariance bound $\bar{V}$ for dynamic predictor obtained from our theory and steady state error covariance for two static predictors obtained from JLS theory using a steady state Kalman gain and maximum stability margin gain

close to unity, while the static predictor giving the largest stability margin approaches the optimal predictor asymptotically for arrival probability close to the critical arrival probability.

From this example, it seems that the upper bound for the dynamic estimator $\bar{V}$ gives an estimate of the minimum steady state covariance that can be achieved with a static estimator for any given arrival probability if the static gain $K_s$ is chosen optimally. Then the MARE could be used to find the minimum steady state covariance and then the corresponding steady state modified Kalman gain, thus providing an useful tool for optimal static estimator design. Future work will explore this possibility.

## VI. CONCLUSIONS

In this paper we have presented an analysis of Kalman filtering in the setting of intermittent observations. We have shown how the expected estimation error covariance depends on the tradeoff between loss probability and the system dynamics. Such a result is useful to the system designer who must assess the relationship between the dynamics of the system whose state is to be estimated and the reliability of the communication channel through which that system is measured.

Our motivating application is a distributed sensor network that collects observations and sends them to one or more central units that are responsible for estimation and control. For example, in a pursuit evasion game in which mobile pursuers perform their control actions based on the current estimate of the positions of both pursuers and evaders, the sensing capability of each pursuer is generally limited, and an embedded sensor network is essential for providing a larger overall view of the terrain. The results that we have presented here can aid the designer of the sensor network in the choice of the number and disposition of the sensors.

This application also suggests a number of interesting directions for further work. For example, although we have assumed independent Bernoulli probabilities for the observation events, in the sensor network there will generally be temporal and spatial sources of variability that lead to correlations among these events. While it is possible to compute posterior state estimates in such a setting, it would be of interest to see if a priori bounds of the kind that we have obtained here can be obtained in this case. Similarly, in many situations there may be correlations between the states and the observation events; for example, such correlations will arise in the pursuit evasion game when the evaders move near the boundaries of the sensor network. Finally, the sensor network setting also suggests the use of smoothing algorithms in addition to the filtering algorithms that have been our focus here. In particular, we may be willing to tolerate a small amount of additional delay to wait for the arrival of a sensor measurement, if that measurement is expected to provide a significant reduction in uncertainty. Thus we would expect that the tradeoff that we have studied here between loss probability and the system dynamics should also be modulated in interesting ways by the delay due to smoothing.

As another motoviation for considering more complex models that remove the assumption that the arrivals are independent Bernoulli, one can imagine situations where some of the sensing is done locally and therefore measurements are available at all sampling times, while measurements taken at distant locations are available at irregular intervals. This would imply different dropping rates for different channels.

## VII. APPENDIX A

In order to give complete proofs of our main theorems, we need to prove some preliminary lemmas. The first lemma displays some useful properties of the MARE.

**Lemma 1.** *Let the operator*

$$\phi(K, X) = (1 - \lambda)(AXA' + Q) + \lambda(FXF' + V) \tag{29}$$

*where $F = A + KC$, $\quad V = Q + KRK'$. Assume $X \in \mathbb{S} = \{S \in \mathbb{R}^{n \times n} | S \geq 0\}$, $R > 0$, $Q \geq 0$, and $(A, Q^{\frac{1}{2}})$ is controllable. Then the following facts are true:*

*(a) With $K_X = -AXC'(CXC' + R)^{-1}$, $g_\lambda(X) = \phi(K_X, X)$*

*(b) $g_\lambda(X) = \min_K \phi(K, X) \leq \phi(K, X) \forall K$*

*(c) If $X \leq Y$, then $g_\lambda(X) \leq g_\lambda(Y)$*

*(d) If $\lambda_1 \leq \lambda_2$ then $g_{\lambda_1}(X) \geq g_{\lambda_2}(X)$*

*(e) If $\alpha \in [0, 1]$, then $g_\lambda(\alpha X + (1 - \alpha)Y) \geq \alpha g_\lambda(X) + (1 - \alpha)g_\lambda(Y)$*

*(f) $g_\lambda(X) \geq (1 - \lambda)AXA' + Q$*

*(g) If $\bar{X} \geq g_\lambda(\bar{X})$, then $\bar{X} > 0$*

*(h) If $X$ is a random variable, then $(1 - \lambda)A\mathbb{E}[X]A' + Q \leq \mathbb{E}[g_\lambda(X)] \leq g_\lambda(\mathbb{E}[X])$*

*Proof:* (a) Define $F_X = A + K_X C$, and observe that

$$F_X XC' + K_X R = (A + K_X C)XC' + K_X R = AXC' + K_X(CXC' + R) = 0.$$

Next, we have

$$
\begin{aligned}
g_\lambda(X) &= (1 - \lambda)(AXA' + Q) + \lambda(AXA' + Q - AXC'(CXC' + R)^{-1}CXA') \\
&= (1 - \lambda)(AXA' + Q) + \lambda(AXA' + Q + K_X CXA') \\
&= (1 - \lambda)(AXA' + Q) + \lambda(F_X XA' + Q) \\
&= (1 - \lambda)(AXA' + Q) + \lambda(F_X XA' + Q) + (F_X XC' + K_X R)K' \\
&= \phi(K_X, X)
\end{aligned}
$$

(b) Let $\psi(K, X) = (A + KC)X(A + KC)' + KRK' + Q$. Note that

$$\text{argmin}_K \phi(K, X) = \text{argmin}_K FXF' + V = \text{argmin}_K \psi(X, K).$$

Since $X, R \geq 0$, $\phi(K, X)$ is quadratic and convex in the variable $K$, therefore the minimizer can be found by solving $\frac{\partial \psi(K, X)}{\partial K} = 0$, which gives:

$$2(A + KC)XC' + 2KR = 0 \implies K = -AXC'(CXC' + R)^{-1}.$$

Since the minimizer corresponds to $K_X$ defined above, the fact follows from fact (1).

(c) Note that $\phi(K, X)$ is affine in $X$. Suppose $X \leq Y$. Then

$$g_\lambda(X) = \phi(K_X, X) \leq \phi(K_Y, X) \leq \phi(K_Y, Y) = g_\lambda(Y).$$

This completes the proof.

(d) Note that $AXC'(CXC' + R)^{-1}CXA \geq 0$. Then

$$
\begin{aligned}
g_{\lambda_1}(X) &= AXA' + Q - \lambda_1 \, AXC'(CXC' + R)^{-1}CXA \\
&\geq AXA' + Q - \lambda_2 \, AXC'(CXC' + R)^{-1}CXA = g_{\lambda_2}(X)
\end{aligned}
$$

(e) Let $Z = \alpha X + (1 - \alpha)Y$ where $\alpha \in [0, 1]$. Then we have

$$
\begin{aligned}
g_\lambda(Z) &= \phi(K_Z, Z) \\
&= \alpha(A + K_Z \, C)X(A + K_Z \, C)' + (1 - \alpha)(A + K_Z \, C)Y(A + K_Z \, C)' + \\
&\quad + (\alpha + 1 - \alpha)(K_Z \, R \, K_Z' + Q) \\
&= \alpha\phi(K_Z, X) + (1 - \alpha)\phi(K_Z, Y) \\
&\geq \alpha\phi(K_X, X) + (1 - \alpha)\phi(K_Y, Y) \\
&= \alpha g_\lambda(X) + (1 - \alpha)g_\lambda(Y).
\end{aligned}
\tag{30}
$$

(f) Note that $F_X X F_X' \geq 0$ and $KRK' \geq 0$ for all $K$ and $X$. Then

$$
\begin{aligned}
g_{\lambda_1}(X) &= \phi(K_X, X) = (1 - \lambda)(AXA' + Q) + \lambda(F_X X F_X' + K_X R K_X' + Q) \\
&\geq (1 - \lambda)(AXA' + Q) + \lambda Q = (1 - \lambda)AXA' + Q.
\end{aligned}
$$

(g) From fact (f) follows that $\bar{X} \geq g_{\lambda_1}(\bar{X}) \geq (1 - \lambda)A\bar{X}A' + Q$. Let $\hat{X}$ such that $\hat{X} = (1 - \lambda)A\hat{X}A' + Q$. Such $\hat{X}$ must clearly exist. Therefore $\bar{X} - \hat{X} \geq (1 - \lambda)A(\bar{X} - \hat{X})A' \geq 0$. Moreover the matrix $\hat{X}$ solves the Lyapunov Equation $\hat{X} = \tilde{A}\hat{X}\tilde{A}' + Q$ where $\tilde{A} = \sqrt{1 - \lambda}A$. Since $(\tilde{A}, Q^{\frac{1}{2}})$ is detectable, it follows that $\hat{X} > 0$ and so $\bar{X} > 0$, which proves the fact.

(h) Using fact (f) and linearity of expectation we have

$$\mathbb{E}[g_\lambda(X)] \geq \mathbb{E}[(1 - \lambda)AXA' + Q] = (1 - \lambda)A\mathbb{E}[X]A' + Q,$$

fact (e) implies that the operator $g_\lambda()$ is concave, therefore by Jensen's Inequality we have $\mathbb{E}[g_\lambda(X)] \leq g_\lambda(\mathbb{E}[X])$. ∎

**Lemma 2.** *Let $X_{t+1} = h(X_t)$ and $Y_{t+1} = h(Y_t)$. If $h(X)$ is a monotonically increasing function then:*

$$X_1 \geq X_0 \quad \Rightarrow \quad X_{t+1} \geq X_t, \quad \forall t \geq 0$$
$$X_1 \leq X_0 \quad \Rightarrow \quad X_{t+1} \leq X_t, \quad \forall t \geq 0$$
$$X_0 \leq Y_0 \quad \Rightarrow \quad X_t \leq Y_t, \qquad \forall t \geq 0$$

*Proof:* This lemma can be readily proved by induction. It is true for $t = 0$, since $X_1 \geq X_0$ by definition. Now assume that $X_{t+1} \geq X_t$, then $X_{t+2} = h(X_{t+1}) \geq h(X_{t+1}) = X_{t+1}$ because of monotonicity of $h(\cdot)$. The proof for the other two cases is analogous. ∎

It is important to note that while in the scalar case $X \in \mathbb{R}$ either $h(X) \leq X$ or $h(X) \geq X$; in the matrix case $X \in \mathbb{R}^{n \times n}$, it is not generally true that either $h(X) \geq X$ or $h(X) \leq X$. This is the source of the major technical difficulty for the proof of convergence of sequences in higher dimensions. In this case convergence of a sequence $\{X_t\}_0^\infty$ is obtained by finding two other sequences, $\{Y_t\}_0^\infty, \{Z_t\}_0^\infty$ that bound $X_t$, i.e., $Y_t \leq X_t \leq Z_t, \forall t$, and then by showing that these two sequences converge to the same point.

The next two lemmas show that when the MARE has a solution $\bar{P}$, this solution is also stable, i.e., every sequence based on the difference Riccati equation $P_{t+1} = g_\lambda(P_t)$ converges to $\bar{P}$ for all initial positive semidefinite conditions $P \geq 0$.

**Lemma 3.** *Define the linear operator*

$$\mathcal{L}(Y) = (1 - \lambda)(AYA') + \lambda(FYF')$$

*Suppose there exists $\overline{Y} > 0$ such that $\overline{Y} > \mathcal{L}(\overline{Y})$.*

*(a) For all $W \geq 0$,*

$$\lim_{k \to \infty} \mathcal{L}^k(W) = 0$$

*(b) Let $U \geq 0$ and consider the linear system*

$$Y_{k+1} = \mathcal{L}(Y_k) + U \quad \text{initialized at} \quad Y_0.$$

*Then, the sequence $Y_k$ is bounded.*

*Proof:* (a) First observe that $0 \leq \mathcal{L}(Y)$ for all $0 \leq Y$. Also, $X \leq Y$ implies $\mathcal{L}(X) \leq \mathcal{L}(Y)$. Choose $0 \leq r < 1$ such that $\mathcal{L}(\overline{Y}) < r\overline{Y}$. Choose $0 \leq m$ such that $W \leq m\overline{Y}$. Then,

$$0 \leq \mathcal{L}^k(W) \leq m\mathcal{L}^k(\overline{Y}) < mr^k\overline{Y}$$

The assertion follows when we take the limit $r \to \infty$, on noticing that $0 \leq r < 1$.

(b) The solution of the linear iteration is

$$
\begin{aligned}
Y_k &= \mathcal{L}^k(Y_0) + \sum_{t=0}^{k-1} \mathcal{L}^t(U) \\
&\leq \left( m_{Y_0} r^k + \sum_{t=0}^{k-1} m_U r^t \right) \overline{Y} \\
&\leq \left( m_{Y_0} r^k + \frac{m_U}{1-r} \right) \overline{Y} \\
&\leq \left( m_{Y_0} + \frac{m_U}{1-r} \right) \overline{Y}
\end{aligned}
$$

proving the claim. ∎

**Lemma 4.** *Consider the operator $\phi(K, X)$ defined in Equation (29). Suppose there exists a matrix $\overline{K}$ and a positive definite matrix $\overline{P}$ such that*

$$ \overline{P} > 0 \quad and \quad \overline{P} > \phi(\overline{K}, \overline{P}). $$

*Then, for any $P_0$, the sequence $P_t = g_\lambda^t(P_0)$ is bounded, i.e. there exists $M_{P_0} \geq 0$ dependent of $P_0$ such that*

$$ P_t \leq M \quad for\ all \quad t. $$

*Proof:* First define the matrices $\overline{F} = A + \overline{K}C$ and consider the linear operator

$$ \mathcal{L}(Y) = (1 - \lambda)(AYA') + \lambda(\overline{F}Y\overline{F}') $$

Observe that

$$ \overline{P} > \phi(\overline{K}, \overline{P}) = \mathcal{L}(\overline{P}) + Q + \lambda \overline{K}R\overline{K}' \geq \mathcal{L}(\overline{P}). $$

Thus, $\mathcal{L}$ meets the condition of Lemma 3. Finally, using fact (b) in Lemma 1 we have

$$ P_{t+1} = g_\lambda(P_t) \leq \phi(\overline{K}, P_t) = \mathcal{L}P_t + Q + \lambda \overline{K}R\overline{K}' = \mathcal{L}(P_t) + U. $$

Since $U = \lambda \overline{K}R\overline{K}' + Q \geq 0$, using Lemma 3, we conclude that the sequence $P_t$ is bounded. ∎

We are now ready to give proofs for Theorems 1-4.

*A. Proof of Theorem 1*

(a) We first show that the modified Riccati difference equation initialized at $Q_0 = 0$ converges. Let $Q_k = g_\lambda^k(0)$. Note that $0 = Q_0 \leq Q_1$. It follows from Lemma 1(c) that

$$Q_1 = g_\lambda(Q_0) \leq g_\lambda(Q_1) = Q_2.$$

A simple inductive argument establishes that

$$0 = Q_0 \leq Q_1 \leq Q_2 \leq \cdots \leq M_{Q_0}.$$

Here, we have used Lemma 4 to bound the trajectory. We now have a monotone non-decreasing sequence of matrices bounded above. It is a simple matter to show that the sequence converges, i.e.

$$\lim_{k \to \infty} Q_k = \overline{P}.$$

Also, we see that $\overline{P}$ is a fixed point of the modified Riccati iteration:

$$\overline{P} = g_\lambda(\overline{P}),$$

which establishes that it is *a* positive semidefinite solution of the MARE.

Next, we show that the Riccati iteration initialized at $R_0 \geq \overline{P}$ also converges, and to the same limit $\overline{P}$. First define the matrices

$$\overline{K} = -A\overline{P}C' \left( C\overline{P}C' + R \right)^{-1}, \quad \overline{F} = A + \overline{K}C$$

and consider the linear operator

$$\hat{\mathcal{L}}(Y) = (1 - \lambda)(AYA') + \lambda(\overline{F}Y\overline{F}').$$

Observe that

$$\overline{P} = g_\lambda(\overline{P}) = \mathcal{L}(\overline{P}) + Q + \overline{K}R\overline{K}' > \hat{\mathcal{L}}(\overline{P}).$$

Thus, $\hat{\mathcal{L}}$ meets the condition of Lemma 3. Consequently, for all $Y \geq 0$,

$$\lim_{k \to \infty} \hat{\mathcal{L}}^k(Y) = 0.$$

Now suppose $R_0 \geq \overline{P}$. Then,

$$R_1 = g_\lambda(R_0) \geq g_\lambda(\overline{P}) = \overline{P}.$$

A simple inductive argument establishes that

$$R_k \geq \overline{P} \quad \text{for all } k.$$

Observe that

$$
\begin{aligned}
0 \leq (R_{k+1} - \overline{P}) &= g_\lambda(R_k) - g_\lambda(\overline{P}) \\
&= \phi(K_{R_k}, R_k) - \phi(K_{\overline{P}}, \overline{P}) \\
&\leq \phi(K_{\overline{P}}, R_k) - \phi(K_{\overline{P}}, \overline{P}) \\
&= (1 - \lambda)A(R_k - \overline{P})A' + \lambda F_{\overline{P}}(R_k - \overline{P})F'_{\overline{P}} \\
&= \hat{\mathcal{L}}(R_k - \overline{P}).
\end{aligned}
$$

Then, $0 \leq \lim_{k \to \infty}(R_{k+1} - \overline{P}) \leq 0$, proving the claim.

We now establish that the Riccati iteration converges to $\overline{P}$ for all initial conditions $P_0 \geq 0$. Define $Q_0 = 0$ and $R_0 = P_0 + \overline{P}$. Consider three Riccati iterations, initialized at $Q_0, P_0$, and $R_0$. Note that

$$Q_0 \leq P_0 \leq R_0.$$

It then follows from Lemma 2 that

$$Q_k \leq P_k \leq R_k \quad \text{for all } k.$$

We have already established that the Riccati equations $P_k$ and $R_k$ converge to $\overline{P}$. As a result, we have

$$\overline{P} = \lim_{k \to \infty} P_k \leq \lim_{k \to \infty} Q_k \leq \lim_{k \to \infty} R_k = \overline{P},$$

proving the claim.

(b) Finally, we establish that the MARE has a unique positive semidefinite solution. To this end, consider $\hat{P} = g_\lambda(\hat{P})$ and the Riccati iteration initialized at $P_0 = \hat{P}$. This yields the constant sequence

$$\hat{P}, \hat{P}, \cdots$$

However, we have shown that every Riccati iteration converges to $\overline{P}$. Thus $\overline{P} = \hat{P}$.

## B. Proof of Theorem 2

First we note that the two cases expressed by the theorem are indeed possible. If $\lambda = 1$ the modified Riccati difference equation reduces to the standard Riccati difference equation, which is known to converge to a fixed point, under the theorem's hypotheses. Hence, the covariance matrix is always bounded in this case, for any initial condition $P_0 \geq 0$. If $\lambda = 0$ then we reduce to open loop prediction, and if the matrix $A$ is unstable, then the covariance matrix diverges for some initial condition $P_0 \geq 0$. Next, we show the existence of a single point of transition between the two cases. Fix a $0 < \lambda_1 \leq 1$ such that $\mathbb{E}_{\lambda_1}[P_t]$ is bounded for any initial condition $P_0 \geq 0$. Then, for any $\lambda_2 \geq \lambda_1$ $\mathbb{E}_{\lambda_2}[P_t]$ is also bounded for all $P_0 \geq 0$. In fact we have

$$
\begin{aligned}
\mathbb{E}_{\lambda_1}[P_{t+1}] &= \mathbb{E}_{\lambda_1}[AP_tA' + Q - \gamma_{t+1}AP_tC'(CP_tC' + R)^{-1}CP_tA] \\
&= \mathbb{E}[AP_tA' + Q - \lambda_1 AP_tC'(CP_tC' + R)^{-1}CP_tA] \\
&= \mathbb{E}[g_{\lambda_1}(P_t)] \\
&\geq \mathbb{E}[g_{\lambda_2}(P_t)] \\
&= \mathbb{E}_{\lambda_2}[P_{t+1}],
\end{aligned}
$$

where we exploited fact (d) of Lemma 1 to write the above inequality . We can now choose

$$
\lambda_c = \{\inf \lambda^* : \lambda > \lambda^* \Rightarrow \mathbb{E}_\lambda[P_t] \text{is bounded, for all } P_0 \geq 0\},
$$

completing the proof.

## C. Proof of Theorem 3

Define the Lyapunov operator $m(X) = \tilde{A}X\tilde{A}' + Q$ where $\tilde{A} = \sqrt{1-\lambda}A$. If $(A, Q^{\frac{1}{2}})$ is controllable, also $(\tilde{A}, Q^{\frac{1}{2}})$ is controllable. Therefore, it is well known that $\hat{S} = m(\hat{S})$ has a unique strictly positive definite solution $\hat{S} > 0$ if and only if $\max_i |\sigma_i(\tilde{A})| < 1$, i.e. $\sqrt{1-\lambda} \max_i |\sigma_i(A)| < 1$, from which follows $\underline{\lambda} = 1 - \frac{1}{\alpha^2}$. If $\max_i |\sigma_i(\tilde{A})| \geq 1$ it is also a well known fact that there is no positive semidefinite fixed point to the Lyapunov equation $\hat{S} = m(\hat{S})$, since $(\tilde{A}, Q^{\frac{1}{2}})$ is controllable.

Let us consider the difference equation $S_{t+1} = m(S_t)$, $S_0 = 0$. It is clear that $S_0 = 0 \leq Q = S_1$. Since the operator $m()$ is monotonic increasing, by Lemma 2 it follows that the sequence $\{S_t\}_0^\infty$ is monotonically increasing, i.e. $S_{t+1} \geq S_t$ for all $t$. If $\lambda < \underline{\lambda}$ this sequence

does not converge to a finite matrix $\bar{S}$, otherwise by continuity of the operator $m$ we would have $\bar{S} = m(\bar{S})$, which is not possible. Since it is easy to show that a monotonically increasing sequence $S_t$ that does not converge is also unbounded, then we have

$$\lim_{t \to \infty} S_t = \infty.$$

Let us consider now the mean covariance matrix $\mathbb{E}[P_t]$ initialized at $\mathbb{E}[P_0] \geq 0$. Clearly $0 = S_0 \leq \mathbb{E}[P_0]$. Moreover it is also true

$$S_t \leq \mathbb{E}[P_t] \Longrightarrow S_{t+1} = (1 - \lambda)AS_tA' + Q \leq (1 - \lambda)A\mathbb{E}[P_t]A' + Q \leq \mathbb{E}[g_\lambda(P_t)] = \mathbb{E}[P_{t+1}],$$

where we used fact (h) from Lemma 1. By induction, it is easy to show that

$$S_t \leq \mathbb{E}[P_t] \ \forall t, \ \forall \mathbb{E}[P_0] \geq 0 \Longrightarrow \lim_{t \to \infty} \mathbb{E}[P_t] \geq \lim_{t \to \infty} S_t = \infty.$$

This implies that for any initial condition $\mathbb{E}[P_t]$ is unbounded for any $\lambda < \underline{\lambda}$, therefore $\underline{\lambda} \leq \lambda_c$, which proves the first part of the Theorem.

Now consider the sequence $V_{t+1} = g_\lambda(V_t)$, $V_0 = \mathbb{E}[P_0] \geq 0$. Clearly

$$\mathbb{E}[P_t] \leq V_t \implies \mathbb{E}[P_{t+1}] = \mathbb{E}[g_\lambda(P_t)] \leq g_\lambda(\mathbb{E}[P_t]) \leq [g_\lambda(V_t)] = V_{t+1},$$

where we used facts (c) and (h) from Lemma 1. Then a simple induction argument shows that $V_t \geq \mathbb{E}[P_t]$ for all $t$. Let us consider the case $\lambda > \bar{\lambda}$, therefore there exists $\hat{X}$ such that $\hat{X} \geq g_\lambda(\hat{X})$. By Lemma 1(g) $\bar{X} > 0$, therefore all hypotheses of Lemma 3 are satisfied, which implies that

$$\mathbb{E}[P_t] \leq V_t \leq M_{V_0} \ \forall t.$$

This shows that $\lambda_c \leq \bar{\lambda}$ and concludes the proof of the Theorem.

*D. Proof of Theorem 4*

Let consider the sequences $S_{t+1} = (1 - \lambda)AS_tA' + Q$, $S_0 = 0$ and $V_{t+1} = g_\lambda(V_t)$, $V_0 = \mathbb{E}[P_0] \geq 0$. Using the same induction arguments in Theorem 3 it is easy to show that

$$S_t \leq \mathbb{E}[P_t] \leq V_t \ \forall t.$$

From Theorem 1 follows that $\lim_{t \to \infty} V_t = \bar{V}$, where $\bar{V} = g_\lambda V$. As shown before the sequence $S_t$ is monotonically increasing. Also it is bounded since $S_t \leq V_t \leq M$. Therefore $\lim_{t \to \infty} S_t = \bar{S}$, and by continuity $\bar{S} = (1 - \lambda)A\bar{S}A' + Q$, which is a Lyapunov equation. Since $\sqrt{1 - \lambda}A$ is

stable and $(A, Q^{\frac{1}{2}})$ is controllable, then the solution of the Lyapunov equation is strictly positive definite, i.e. $\bar{S} > 0$. Adding all the results together we get

$$0 < \bar{S} = \lim_{t \to \infty} S_t \leq \lim_{t \to \infty} \mathbb{E}[P_t] \leq \lim_{t \to \infty} V_t = \bar{V},$$

which concludes the proof.

# REFERENCES

[1] Smart dust project home page. http://robotics.eecs.berkeley.edu/ pister/SmartDust/.

[2] NEST project at Berkeley home page. http://webs.cs.berkeley.edu/nest-index.html.

[3] Seismic sensor research at berkeley, home page. http://www.berkeley.edu/news/media/releases /2001/12/13_snsor.html.

[4] P. Varaiya, "Smart cars on smart roads: Problems of control," *IEEE Transactions on Automatic Control*, vol. 38(2), February 1993.

[5] J. Lygeros, D. N. Godbole, and S. S. Sastry, "Verified hybrid controllers for automated vehicles," *IEEE Transactions on Automatic Control*, vol. 43(4), 1998.

[6] B. Sinopoli, C. Sharp, S. Schaffert, L. Schenato, and S. Sastry, "Distributed control applications within sensor networks," *IEEE Proceedings Special Issue on Distributed Sensor Networks*, November 2003.

[7] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering on Automatic Control*, vol. 82(D), pp. 35–45, 1960.

[8] P. S. Maybeck, *Stochastic models, estimation, and control*, ser. Mathematics in Science and Engineering, 1979, vol. 141.

[9] M. Micheli and M. I. Jordan, "Random sampling of a continiuous-time stochastic dynamical system," in *Proceedings of 15th International Symposium on the Mathematical Theory of Networks and Systems (MTNS)*, University of Notre Dame, South Bend, Indiana, August 2002.

[10] M. Micheli, "Random sampling of a continuous-time stochastic dynamical system: Analysis, state estimation, applications," Master's Thesis, University of California at Berkeley, Deparment of Electrical Engineering, 2001.

[11] M. Athans, R. Ku, and S. B. Gershwin, "The uncertainty threshold principle, some fundamental limitations of optimal decision making under dynamic uncertainty," *IEEE Transactions on Automatic Control*, vol. 22(3), pp. 491–495, June 1977.

[12] R. Ku and M. Athans, "Further results on the uncertainty threshold principle," *IEEE Transactions on Automatic Control*, vol. 22(5), pp. 491–495, October 1977.

[13] M. Mariton, *Jump Linear Systems in Automatic Control*. Marcel Dekker, 1990.

[14] J. Nilsson, "Real-time control systems with delays," Ph.D. dissertation, Department of Automatic Control, Lund Institute of Technology, 1998.

[15] J. Nilsson, B. Bernhardsson, and B. Wittenmark, "Stochastic analysis and control of real-time systems with random time delays." [Online]. Available: citeseer.nj.nec.com/101333.html

[16] Q. Ling and M. Lemmon, "Soft real-time scheduling of networked control systems with dropouts governed by a markov chain," in *American Control Conference*, June 2003, denver, CO.

[17] T. Fortmann, Y. Bar-Shalom, M. Scheffe, and S. Gelfand, "Detection thresholds for tracking in clutter-a connection between estimation and signal processing," *IEEE Transactions on Automatic Control*, vol. AC-30, no. 3, pp. 221–228, March 1985.

[18] N. M. Dempster, A. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B*, vol. 39, pp. 185–197, 1977.

[19] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*. Springer Verlag, March 2000.

[20] S. Lauritzen, *Graphical Models*. Clarendon Press, 1996.

[21] C. N. Hadjicostis and R. Touri, "Feedback control utilizing packet dropping network links," in *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, NV, Dec 2002, invited.