Kalman Filtering with Intermittent Observations: On the Boundedness of the Expected Error Covariance

Eduardo Rath Rohr, Damián Marelli, and Minyue Fu

***Abstract*—This paper addresses the stability of a Kalman filter when measurements are intermittently available due to constraints in the communication channel between the sensor and the estimator. We give a necessary condition and a sufficient condition, with a trivial gap between them, for the boundedness of the expected value of the estimation error covariance. These conditions are more general than the existing ones in the sense that they only require the state matrix of the system to be diagonalizable and the sequence of packet losses to be a stationary finite order Markov process. Hence, we extend the class of systems for which these conditions are known in two directions, namely, by including degenerate systems, and by considering network models more general than i.i.d. and Gilbert-Elliott. We show that these conditions recover known results from the literature when evaluated for non-degenerate systems under the assumption of i.i.d. or Gilbert-Elliott packet loss models.**

I. Introduction

Characterizing the behavior of a Kalman filter when measurements are intermittently available has attracted a great interest in the recent years. This is partly due to the development of communications technologies, which today permit distributed control and monitoring in a broad range of applications. When measurements sent through a communication channel are subject to random losses, the estimation accuracy of a Kalman filter deteriorates. In [1], the authors established the mathematical foundations for the basic problem and pointed out that the covariance of the estimation error does not reach a steady state. Since then, several authors have studied different aspects of the problem, using different assumptions on network models and protocols.

In a Kalman filter with intermittent observations (KFIO), the error covariance (EC) matrix becomes random. A central issue in the study of KFIO is the study the stochastic properties of its EC. One of these properties, which has been recently studied, is the asymptotic probability distribution of the EC (APDEC). The APDEC provides a comprehensive description of the performance of the KFIO, as it allows the system designer to know the probability that the EC will lie inside an acceptable range. In [2] the authors provide lower and upper bounds on the APDEC under the assumption that the sensor is capable of sending either the state estimation (instead of the raw measurements), or a packet containing a group of measurements having enough information to produce a bounded EC. This setting was relaxed in [3], where the authors assumed that a sensor only sends one measurement per sampling interval. Moreover, the bounds on the APDEC presented in [3] can be made monotonically tighter at the expense of increased computations.

E. R. Rohr was with the School of Electrical Engineering and Computer Science, The University of Newcastle, Australia. He is now with ABB Corporate Research Center, Switzerland. Email: eduardo.rohr@ch.abb.com. D. Marelli is with the School of Electrical Engineering and Computer Science, The University of Newcastle, Australia. He is also with the Acoustics Research Institute, Austrian Academy of Sciences, Austria. Email: damian.marelli@newcastle.edu.au. M. Fu is with the School of Electrical Engineering and Computer Science, The University of Newcastle, Australia. He also holds a Qian-ren Professorship at the State Key Laboratory of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, China. Email: minyue.fu@newcastle.edu.au

This work was partly supported by the Austrian Science Fund (FWF, project M1230-N13).

Although previous works were concerned with finding bounds on the APDEC, it was not until [4], [5] that the existence of a unique and invariant APDEC was shown. In [4] the authors consider the case where the measurements are dropped according to a Bernoulli process, while [5] adopts the Gilbert-Elliott network model. Other properties of the APDEC are also studied in [6], [7].

Another stochastic property of the EC that has been subject of study is the asymptotic expected EC (AEEC). There exists a rich literature dedicated to finding the stability conditions of the KFIO [1], [5], [8]–[23]. Most authors adopt the stability criterion used in [1], namely, a KFIO is said to be stable if its AEEC is finite [1], and unstable otherwise. Other authors, adopt the concept of peak error covariance, introduced in [8]. More recently, the equivalence between the two notions of stability has been studied in [5], [9]. The AEEC can in principle be determined from the knowledge of its APDEC. To this end, notice that the APDEC can still be evaluated even if the KFIO is unstable [4]. However, determining the stability of the KFIO from its APDEC is not the preferred approach, as it is technically cumbersome and often leads to separate necessary and sufficient conditions with a non-trivial gap between them. In general, the study of stability uses different approaches from those used for studying the APDEC, and is important for its own sake, as it provides a hard limit for the design of the components of the system, the communication channel, and the estimator.

Despite the considerable effort in obtaining conditions for the stability of the KFIO, a complete answer is not available yet. Partial answers depend on the structure of the system under consideration and on the given network model. The latter can be classified according to the stochastic properties of the packet dropouts. Two network models dominate the literature on the topic. The first one considers the dropouts as a sequence of independent and identically distributed (i.i.d.) binary random variables. The second one is known as the Gilbert-Elliott model [24], [25], and models the dropouts using a first order Markov process. This model has been adopted by many authors in an attempt to account for some communication channel phenomena, such as fading and congestion [8]–

[11]. In many applications, particularly when the network conditions change slowly in comparison with the sampling time, the use of a stationary higher order Markov process (also known as finite state Markov channel (FSMC) [26]) produces a more accurate description of the packet dropouts. In the context of KFIO, this network model has been studied in [6], where the existence of a stationary APDEC was investigated. Notice that the i.i.d. and the Gilbert-Elliott network models are particular cases of the FSMC model.

Even when the simplest i.i.d. network model is used, determining if a KFIO is stable is still an open problem. In [1], the authors showed that there exists a critical value, such that the AEEC is bounded if the arrival probability is strictly greater than this value, and unbounded if the arrival probability falls below the critical value. They also provided lower and upper bounds on the critical measurement arrival probability. The bounds are only tight for systems whose observation matrix $\mathbf{C}$ is invertible. This condition was relaxed in [12] to only requiring that the part of the matrix $\mathbf{C}$ corresponding to the observable subspace is invertible. The set of systems for which the necessary conditions for stability are also sufficient was further extended in [13], where the authors studied the case where the unstable eigenvalues of $\mathbf{A}$ have different magnitudes.

The introduction of the Gilbert-Elliott network model in the context of KFIO was done in [8], where sufficient conditions for a slightly different notion of stability, namely, the peak covariance stability, was derived. For scalar plants, the authors showed that this sufficient condition is also necessary. In [10], a new sufficient condition for the stability of the peak covariance was established. In the particular case where the observation matrix $\mathbf{C}$ has full column rank (FCR), the sufficient condition for stability of the peak covariance matches the necessary one presented in [14]. In [9], a necessary and sufficient condition for the stability of the peak covariance for second order systems was derived, while for higher order systems, only a necessary condition was presented.

In [11] the authors derive a necessary and sufficient condition for the stability of the KFIO for a class of systems called non-degenerate. Although this paper presents several novel ideas, we found a technical problem in the argument of the proof of the main result. In Section II, we provide an example that contradicts the claim in [11, Theorem 4].

Some interesting connections between the i.i.d. and the Gilbert-Elliott channel models are presented in [5]. The author showed that, under mild conditions, the notions of peak covariance stability and the usual AEEC stability are equivalent. The author also showed that when the Gilbert-Elliott model is adopted for systems in which the observation matrix $\mathbf{C}$ has FCR, there exists a critical value on the recovery probability (i.e., the probability to receive a measurement given that the previous one was lost), for the stability of the AEEC. Moreover, this critical value matches the critical probability of the i.i.d. case. For general systems, this connection is not straightforward and the necessary and sufficient conditions presented have a non-trivial gap.

In this paper we study the stability conditions of the KFIO. We extend in two directions the class of systems for which the conditions for stability are known. Firstly, we consider that the measurements are dropped according to a FSMC model. This model is substantially more general than the i.i.d. and the Gilbert-Elliott models studied so far, and permits modeling more complex channels with memory and fading [26]. Secondly, we also consider degenerate systems in our analysis, making the only requirement on the system to be that the dynamics matrix $\mathbf{A}$ is diagonalizable. Degenerate systems represent an important class of systems, as it includes, among others, all systems with scalar measurements whose matrix $\mathbf{A}$ has repeated or complex-pairs eigenvalues.

We derive a necessary condition and a sufficient condition for the stability of the KFIO that are tight up to a trivial gap. The main result is cast in its most general form, i.e., in terms of the eigenvalues of $\mathbf{A}$ and the normalized asymptotic probability that the measurement loss pattern belongs to certain classes (more details below). For non-degenerate systems, this probability is simply the recovery rate of the FSMC or the Gilbert-Elliott models and the packet dropping rate when the i.i.d. channel model is considered. When the system is degenerate, this asymptotic probability does not have a straightforward interpretation. Nevertheless, we provide a method to compute it. Since diagonalizable matrices form a dense subset within the set of square matrices, we conjecture that similar conditions also hold for systems that only admit a Jordan form, but this is yet to be proved.

Necessary and sufficient conditions for the stability of the KFIO are often derived by analyzing the system as a whole [1], [8]–[12], [14]–[16], [20]. In contrast, we decompose the system into a number of subsystems, which we call finite multiplicative order (FMO) blocks. A key property of this partition is that we are able to derive a quantity, which is computed for each block, such that, if it is smaller than one on each FMO block, then the KFIO is stable, and if it is greater than one in at least one FMO block, then the KFIO is unstable (see Theorem 7). This leads straightforwardly to the desired necessary condition and sufficient condition (having a trivial gap when some of these quantities equals to one). Thanks to this property, we can carry out the stability analysis for each FMO block separately.

The paper is organized as follows. Section II states the problem and introduce our formal assumptions. The main result is presented in Section III, as well as a discussion on how other results from the literature follow from it. In Section IV, we highlight a byproduct of our analysis, namely, a necessary condition and a sufficient condition for the measurement arrival patterns leading to a received observability matrix (ROM) having FCR. The formal proof of our main result is presented in Section V. We draw our conclusions in Section VI. For ease of reading, some proofs are included in the appendix.

We adopt the following notational conventions. The sets of natural, integer, rational, real and complex numbers are denoted by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. Also denote $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For given $T \in \mathbb{N}$, $\mathbb{B}^T$ denotes the set of binary sequences of length T. We denote the i-th element of the sequence $S \in \mathbb{B}^T$ by $S[i]$, $i = 1, \cdots, T$. A set with the elements a, b and c is denoted by $\{a, b, c\}$, while a sequence with the same elements by (a, b, c). Also, $(0)^T$ denotes a sequence of T zeros, and for $P \in \mathbb{B}^{T_1}$ and $S \in \mathbb{B}^{T_2}$, (P, S) denotes sequence concatenation. The length of the sequence $S \in \mathbb{B}^T$ is denoted by $|S| = T$. For symmetric matrices $\mathbf{X}, \mathbf{Y}$ of same dimensions, we use $\mathbf{X} > \mathbf{Y}$ ($\mathbf{X} \geq \mathbf{Y}$) to indicate that $\mathbf{X} - \mathbf{Y}$ is positive definite (resp. positive semi-definite). We say that $\mathbf{X} \succ \mathbf{Y}$ ($\mathbf{X} \succeq \mathbf{Y}$) if all the entries of the matrix $\mathbf{X} - \mathbf{Y}$ are positive (resp. non-negative). For a matrix $\mathbf{X}$, $\mathbf{X}^*$ denotes the transpose conjugate and $\mathbf{X}'$ denotes transpose. Also, $\|\mathbf{X}\|$ denotes the operator norm of $\mathbf{X}$ (i.e., the largest singular value), $\rho(\mathbf{X})$ denotes the spectral radius (i.e., the absolute value of the largest eigenvalue), and $\mathrm{Tr}(\mathbf{X})$ denotes the trace. We use $\circ$ to denotes the composition of functions (i.e. $f \circ g(x) = f(g(x))$). Also $f^{(j)}(x)$ denotes the composition of $f(\cdot)$ j times. We use $\mathbf{A}$ for matrices, $\mathbf{a}$ for vectors, and $\mathcal{A}$ for sets. We use $\overline{\mathcal{A}}$ to denote the complement of the set $\mathcal{A}$.

II. Problem Formulation

Consider the discrete-time linear system

$$\begin{cases} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{C}\mathbf{x}_t + \mathbf{v}_t \end{cases} \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^n$ is the vector of states, $\mathbf{y} \in \mathbb{R}^p$ is the vector of measurements, $\mathbf{w} \sim N(\mathbf{0}, \mathbf{Q})$ with $\mathbf{Q} \geq 0$ is the process noise, $\mathbf{v} \sim N(\mathbf{0}, \mathbf{R})$ with $\mathbf{R} \geq 0$ is the measurement noise, $\mathbf{A} \in \mathbb{C}^{n \times n}$ is the dynamics matrix and $\mathbf{C} \in \mathbb{C}^{p \times n}$ is the measurement matrix. The initial state is $\mathbf{x}_0 \sim N(\mathbf{0}, \mathbf{P}_0)$, with $\mathbf{P}_0 \geq 0$. The measurements are sent to an estimator through a network subject to random packet losses, but without delays. We assume that an error correcting scheme is used such that if an error is introduced during transmission, it can be detected. If the transmission error cannot be corrected, then the corresponding measurement is discarded. Let g_t be a binary random variable describing the arrival of a valid measurement at time t. We denote $g_t = 1$ when $\mathbf{y}_t$ is available for the estimator and $g_t = 0$ otherwise.

We run a Kalman filter to obtain an estimate $\hat{\mathbf{x}}_t$ of the state $\mathbf{x}_t$. The update equation of the EC matrix $\mathbf{P}_t$ (i.e., the covariance of the error $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$ given the measurements received up to time $t-1$) can be written as follows [1]:

$$\mathbf{P}_t = \begin{cases} \Phi_0(\mathbf{P}_{t-1}) & , \ g_{t-1} = 0, \\ \Phi_1(\mathbf{P}_{t-1}) & , \ g_{t-1} = 1, \end{cases} \tag{2}$$

with

$$\Phi_0(\mathbf{X}) = \mathbf{A}\mathbf{X}\mathbf{A}^* + \mathbf{Q},$$
$$\Phi_1(\mathbf{X}) = \mathbf{A}\mathbf{X}\mathbf{A}^* + \mathbf{Q} - \mathbf{A}\mathbf{X}\mathbf{C}^* \left(\mathbf{C}\mathbf{X}\mathbf{C}^* + \mathbf{R}\right)^{-1} \mathbf{C}\mathbf{X}\mathbf{A}^*.$$

Let Γ_t be the binary sequence indicating whether the measurements $y_\tau, \tau = 0, \cdots, t-1$ are available, i.e.,

$$\Gamma_t \triangleq (g_0, \cdots, g_{t-1}). \tag{3}$$

For a given matrix $0 \leq \mathbf{X} \in \mathbb{R}^{n \times n}$ and sequence $S \in \mathbb{B}^T$, we define the map $\mathbf{\Psi} : \mathbb{R}^{n \times n} \times \mathbb{B}^T \to \mathbb{R}^{n \times n}$, by

$$\mathbf{\Psi}(\mathbf{X}, S) = \Phi_{S[T]} \circ \Phi_{S[T-1]} \circ \ldots \Phi_{S[1]}(\mathbf{X}). \tag{4}$$

Notice that the EC at time t only depends on the initial EC, $\mathbf{P}_0$, and the sequence of available measurements up to time $t-1$, i.e.,

$$\mathbf{P}_t = \mathbf{\Psi}(\mathbf{P}_0, \Gamma_t) = \Phi_{g_{t-1}} \circ \Phi_{g_{t-2}} \circ \ldots \Phi_{g_0}(\mathbf{P}_0). \tag{5}$$

In this paper we derive a necessary condition and a sufficient condition, with a trivial gap between them, for the boundedness of the asymptotic value of the norm of the expected error covariance, which we call the AEEC norm. The definition of this quantity is given below.

Definition 1. For a given initial EC $\mathbf{P}_0 \geq 0$, the AEEC norm is defined as

$$G(\mathbf{P}_0) \triangleq \limsup_{t \to \infty} G_t(\mathbf{P}_0), \tag{6}$$

where

$$G_t(\mathbf{P}_0) \triangleq \|\mathbb{E}(\mathbf{P}_t)\| = \left\| \sum_{S \in \mathbb{B}^t} \mathbb{P}(\Gamma_t = S) \mathbf{\Psi}(\mathbf{P}_0, S) \right\|. \tag{7}$$

Remark 2. Conditions to guarantee that

$$\limsup_{t \to \infty} G_t(\mathbf{P}_0) = \liminf_{t \to \infty} G_t(\mathbf{P}_0),$$

as well as the independence of this limit with the initial covariance $\mathbf{P}_0$, are given in [6, Proposition 6] and [27, Th. 2.4]. We state our results in terms of $\limsup_{t \to \infty} G_t(\mathbf{P}_0)$ to make them valid even when these conditions cannot be guaranteed.

Remark 3. Our criterion for stability of a KFIO is that the AEEC norm $G(\mathbf{P}_0) < \infty$. This choice can be roughly justified as follows. Notice that, conditioned on a given packet arrival sequence Γ_t, the estimation error $\tilde{\mathbf{x}}_t$ has normal distribution. For such a distribution, having an infinite covariance norm implies that any state component, on the eigenspace associated with the resulting infinite eigenvalues, is equally likely. This in turn means that the Kalman filter provides no information about the state on that eigenspace. Hence, the packet loss condition under which the AEEC norm is finite is the condition under which the Kalman filter, on average over the sequences Γ_t, provides useful information about each component of the state.

The results are obtained under the assumptions that the matrix $\mathbf{A}$ is diagonal and that the measurements are dropped according to a FSMC model. We formally introduce these assumptions below. Notice that if $\mathbf{A}$ is diagonalizable, the results in this paper still apply after a similarity transformation.

Assumption 1. The matrix $\mathbf{A}$ in (1) is diagonal.

Definition 4. Let g_t, $t \in \mathbb{Z}$, be a stationary random process. Its Markov order ν is defined as the smallest non-negative integer such that, for all $\mu \geq 1$, the following holds

$$\mathbb{P}(g_t = 1 | g_{t-\nu-\mu}, \cdots, g_{t-1}) \tag{8}$$
$$= \mathbb{P}(g_t = 1 | g_{t-\nu}, \cdots, g_{t-1}). \tag{9}$$

We say that the communication channel follows the FSMC model if the measurement drop process g_t is a stationary random process with a finite Markov order ν.

Assumption 2. The packet dropout process g_t is stationary, and its Markov order ν is finite. Also, $0 < \mathbb{P}(g_t = 1|g_{t-\nu}, \cdots, g_{t-1}) < 1$, for any $g_{t-\nu}, \cdots, g_{t-1}$.

Remark 5. Notice that the i.i.d. network model is a special case of the FSMC model with $\nu = 0$. It is fully characterized by the parameter $\mathbb{P}(g_t = 1) \triangleq \lambda$ [1]. Similarly, a Gilbert-Elliott model is obtained using the FSMC model with $\nu = 1$ and is fully characterized by two parameters: the recovery rate $q = \mathbb{P}(g_t = 1|g_{t-1} = 0)$ and the failure rate $p = \mathbb{P}(g_t = 0|g_{t-1} = 1)$ [8].

Remark 6. The general FSMC model has been widely used to model wireless channels in a variety of applications (see [26] for a survey of principles and applications). The problem of channel modeling, i.e., how to obtain a model of the channel based on its statistics, has been studied by several authors. For instance, [28] and [29] presented methods to obtain a model given a set of observations from a channel.

III. Main Results

In this section we present our main results. In Section III-A we state our stability conditions. These conditions require splitting the system into blocks (i.e., sub-systems), and are stated in terms of certain asymptotic probability associated to each block. We classify these block in two kinds, namely, degenerate and non-degenerate ones. In Section III-B we state an analytical expression for the aforementioned probability, valid for the case of non-degenerate blocks. In Section III-C we point out the extra difficulty introduced by degenerate blocks, which prevents us from stating an analytical expression for this probbility. To cope with this, we provide in Section III-D a numerical method to evaluate it. Finally, in Section III-E we summarize the resulting procedure for testing stability.

A. Stability conditions

We first consider the structure of matrix A. Recall that $\mathbf{A}$ is diagonal. We will partition the matrix $\mathbf{A}$ into sub-matrices whose eigenvalues have the same modulus and rational phase differences. Formally, consider the following partition:

$$\mathbf{A} = \mathrm{diag}(\mathbf{A}_1, \ \cdots, \ \mathbf{A}_K), \tag{10}$$

where the sub-matrices $\mathbf{A}_k \in \mathbb{C}^{J_k \times J_k}$ are chosen such that, for any k, $\mathbf{A}_k$ has FMO up to a constant (i.e., there exists $N_k \in \mathbb{N}$ and $\alpha_k \in \mathbb{C}$ such that $\mathbf{A}_k^{N_k} = \alpha_k^{N_k}\mathbf{I}$[1]), and for any k and l with $k \neq l$, the matrix $\mathrm{diag}(\mathbf{A}_k, \mathbf{A}_l)$ does not have FMO up to any constant. Then, each sub-matrix $\mathbf{A}_k$ can be written as

$$\mathbf{A}_k = \alpha_k \tilde{\mathbf{A}}_k \tag{11}$$

$$\tilde{\mathbf{A}}_k = \mathrm{diag}(\exp(i2\pi\theta_{k,1}), \ \cdots, \exp(i2\pi\theta_{k,J_k})), \tag{12}$$

with $\alpha_k \in \mathbb{C}$ and $\theta_{k,j} \in \mathbb{Q}$ for $j = 1, \cdots, J_k$. Notice that, for any k and l with $k \neq l$, α_k/α_l is not a root of unity, i.e., $(\alpha_k/\alpha_l)^m \neq 1$ for any $m \in \mathbb{N}$. We assume that the sub-matrices $\mathbf{A}_k$ are ordered such that $|\alpha_1| \geq |\alpha_2| \geq \cdots \geq |\alpha_K|$. We also do the partition

$$\mathbf{C} = [\mathbf{C}_1 \ \cdots \ \mathbf{C}_K], \tag{13}$$

such that each $\mathbf{C}_k$ has the same number of columns as $\mathbf{A}_k$. We then call the pair $(\mathbf{A}_k, \mathbf{C}_k)$ the k-th FMO block of the system (1).

Consider a time interval $[0, t-1]$ and its associated measurement arrival sequence Γ_t. Let t_i, $i = 1, \cdots, s$, be all the time instants such that $\Gamma_t[t_i + 1] = 1$ (notice that $\Gamma_t[t_i + 1] = g_t$). Then, the available measurements can be written in a vector $\mathbf{z}_t$ as follows

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{y}'_{t_1} & \mathbf{y}'_{t_2} & \cdots & \mathbf{y}'_{t_s} \end{bmatrix}' = \mathbf{O}(\Gamma_t)\mathbf{x}_0 + \mathbf{f}(\Gamma_t) \tag{14}$$

where

$$\mathbf{O}(\Gamma_t) \triangleq \begin{bmatrix} \mathbf{O}_1(\Gamma_t) & \mathbf{O}_2(\Gamma_t) & \dots & \mathbf{O}_K(\Gamma_t) \end{bmatrix}, \tag{15}$$

with

$$\mathbf{O}_k(\Gamma_t) \triangleq \left[\left(\mathbf{C}_k \mathbf{A}_k^{t_1}\right)' \quad \left(\mathbf{C}_k \mathbf{A}_k^{t_2}\right)' \quad \cdots \quad \left(\mathbf{C}_k \mathbf{A}_k^{t_s}\right)' \right]'. \tag{16}$$

Also, $\mathbf{f}(\Gamma_t) = [\mathbf{f}'_1, \cdots, \mathbf{f}'_s]'$, with $\mathbf{f}_i = \sum_{j=0}^{t_i-1} \mathbf{C}\mathbf{A}^{t_i-1-j}\mathbf{w}_j + \mathbf{v}_{t_i}$, for $i = 1, \cdots, s$. For $k = 1, \cdots, K$, let $\mathcal{N}_k^t$ denote the subset of sequences S in $\mathbb{B}^t$ such that $\mathbf{O}_k(S)$ does not have FCR, i.e.,

$$\mathcal{N}_k^t \triangleq \{S \in \mathbb{B}^t : \mathbf{O}_k(S) \text{ does not have FCR}\}. \tag{17}$$

To simplify the notation, in the rest of the paper we will use $\mathbb{P}(\mathcal{N}_k^t)$ to denote $\mathbb{P}(\Gamma_t \in \mathcal{N}_k^t)$. The main result is presented in terms of the quantity $\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t}$. We will present in Lemmas 11 and 15 a method to compute this quantity.

We now state the main result of the paper. The formal proof is deferred to Section V.

Theorem 7. *Consider the system* (1) *satisfying Assumptions 1 and 2. If*

$$|\alpha_k|^2 \limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} < 1, \ \textit{for all } k \in \{1, \cdots, K\}, \tag{18}$$

then the AEEC norm $G(\mathbf{P}_0)$ is finite for any $\mathbf{P}_0 \geq 0$, and if

$$|\alpha_k|^2 \limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} > 1, \ \textit{for some } k \in \{1, \cdots, K\}, \tag{19}$$

then $G(\mathbf{P}_0)$ is infinite for any $\mathbf{P}_0 \geq 0$.

Notice that the result is valid for any initial condition $\mathbf{P}_0 \geq 0$, so we will omit the argument of $G(\cdot)$ in the rest of the paper. Also, there is a trivial gap between (18) and (19), i.e., we do not state whether G is finite or not when $|\alpha_k|^2 \limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} = 1$. This type of gap is common in the literature, see e.g., [1], [11].

Remark 8. Notice that the result in Theorem 7 permits carrying out the stability analysis in each FMO block independently. This allows splitting the problem into smaller sub-problems that are easier to analyze. This is a key property of our decomposition of the system into FMO blocks.

Contrary to most results available in the literature, where the conditions for the AEEC to be finite are cast in terms of the parameters of the network model, we state our result in terms of the quantity $\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t}$. This permits stating the result for the general FSMC network model. We

[1] Notice that Assumption 1 is required for this property to hold.

now address the issue of how to evaluate this quantity. To this end, we use the following definition.

Definition 9. The k-th FMO block $(\mathbf{A}_k, \mathbf{C}_k)$ is said to be degenerate if $\mathbf{C}_k$ does not have FCR, and non-degenerate otherwise. Also, the system (1) is said to be degenerate if at least one of its FMO blocks is degenerate, and non-degenerate otherwise.

Remark 10. Our system decomposition into FMO blocks is somehow related to the partition used in [11]. Although for a different purpose, the authors of [11], decomposed the system into subsystems such that all eigenvalues of a subsystem have the same magnitude. In our decomposition, we further require that the phase differences of the complex eigenvalues in a subsystem are rational numbers. In both cases, the system is said to be degenerate if the corresponding observation matrix $\mathbf{C}_k$ of at least one subsystem fails to have FCR. Notice that according to these definitions, some systems classified as degenerate according to the definition in [11] are non-degenerate according to our definition. While in [11] the authors use this decomposition mainly for providing a sufficient condition for the ROM (of the whole system) to have FCR, we use the decomposition to split the stability analysis into sub-problems.

We treat degenerate and non-degenerate FMO blocks separately. In both cases, the results are first derived for the general FSMC network, and then used to state those for the i.i.d. and the Gilbert-Elliott networks. For non-degenerate FMO blocks, we show in Lemma 11 that $\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t}$ is simply the recovery rate of the network, i.e., the probability to receive a measurement after a long sequence of lost ones. This result, together with Theorem 7, extends the class of communication channels for which the stability conditions of the KFIO are known. It also recovers most known results in the literature.

B. Non-Degenerate FMO blocks

The lemma below shows that for non-degenerate FMO blocks, the quantity $\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t}$ is simply the recovery rate of the network, i.e., the probability to receive a measurement after a long sequence of dropped ones.

Lemma 11. *Let $(\mathbf{A}_k, \mathbf{C}_k)$ be a non-degenerate FMO block, and suppose that Assumption 2 holds. Then,*

$$\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} = \mathbb{P}\left(g_t = 0 | g_{t-\nu} = 0, \cdots, g_{t-1} = 0\right). \tag{20}$$

Proof: To simplify the notation, suppose $\mathbf{A} = \mathbf{A}_1$ so we can omit the subscript k. Since the FMO block is non-degenerate, the matrix $\mathbf{C}$ has FCR. Hence, if at least one measurement in the sequence Γ_t is available, then $\mathbf{O}(\Gamma_t)$ has FCR. That is, the only sequence that produces a matrix $\mathbf{O}(\Gamma_t)$ that does not have FCR, occurs when all t measurements are lost, i.e., $\Gamma_t[1] = \cdots = \Gamma_t[t] = 0$. Let $S_0 = (S_1, S_2)$ be a sequence composed of t zeros, with $S_1 = (0)^\nu$ and $S_2 = (0)^{t-\nu}$. We have

$$\begin{aligned}
\mathbb{P}\left(\mathcal{N}^t\right) =& \mathbb{P}\left(\Gamma_t = S_0\right) \\
=& \mathbb{P}\left(\Gamma_\nu = S_1\right) \mathbb{P}\left(\Gamma_t = S_0 | \Gamma_\nu = S_1\right) \\
=& \mathbb{P}\left(\Gamma_\nu = S_1\right) \prod_{\tau=\nu}^{t-1} \mathbb{P}\left(g_\tau = 0 | g_0, \cdots, g_{\tau-1} = 0\right) \\
=& \mathbb{P}\left(\Gamma_\nu = S_1\right) \mathbb{P}\left(g_\tau = 0 | g_{\tau-\nu} = 0, \cdots, g_{\tau-1} = 0\right)^{t-\nu} \\
=& \frac{\mathbb{P}\left(\Gamma_\nu = S_1\right) \mathbb{P}\left(g_\tau = 0 | g_{\tau-\nu} = 0, \cdots, g_{\tau-1} = 0\right)^t}{\mathbb{P}\left(g_\tau = 0 | g_{\tau-\nu} = 0, \cdots, g_{\tau-1} = 0\right)^\nu},
\end{aligned}$$

and the result follows immediately since

$$\limsup_{t\to\infty} \left(\frac{\mathbb{P}\left(\Gamma_\nu = S_1\right)}{\mathbb{P}\left(g_\tau = 0 | g_{\tau-\nu} = 0, \cdots, g_{\tau-1} = 0\right)^\nu} \right)^{1/t} = 1.$$

■

Combining Theorem 7 and Lemma 11 we obtain the following corollary.

Corollary 12. *For an i.i.d. network with $\mathbb{P}(g_t = 1) = \lambda$, the AEEC norm G is finite if*

$$|\alpha_1|^2 (1 - \lambda) < 1, \tag{21}$$

and is infinite if

$$|\alpha_1|^2 (1 - \lambda) > 1. \tag{22}$$

Also, for a Gilbert-Elliott network model with recovery rate $\mathbb{P}(g_t = 1 | g_{t-1} = 0) = q$, G is finite if

$$|\alpha_1|^2 (1 - q) < 1, \tag{23}$$

and is infinite if

$$|\alpha_1|^2 (1 - q) > 1. \tag{24}$$

Remark 13. To this date, the most general necessary and sufficient conditions for the boundedness of the AEEC are due to [11], where a non-degenerate system and a Gilbert-Elliott network model are considered. Although we show in Example 19 below, that there is a flaw in their proof, the above corollary states that their main result [11, Theorem 8] still holds.

C. Degenerate FMO blocks

When the matrix $\mathbf{C}_k$ does not have FCR for some k, then more than one measurement in the sequence Γ_t must be available in order for $\mathbf{O}_k(\Gamma_t)$ to have FCR. Moreover, the time when the measurement is received is important to determine how many measurements are needed. We show this with the following example:

Example 14. Consider the FMO block $(\mathbf{A}, \mathbf{C})$, with

$$\mathbf{A} = \alpha \ \mathrm{diag}(1, -1) \qquad \mathbf{C} = [1 \ 1]. \tag{25}$$

If all the available measurements are from even time instants, i.e., $\Gamma_t = (1, 0, 1, 0, 1, 0, \cdots)$, then the matrix $\mathbf{O}(\Gamma_t)$ does not have FCR:

$$\mathbf{O}\left((1, 0, 1, 0, 1, 0)\right) = \begin{bmatrix} 1 & \alpha^2 & \alpha^4 \\ 1 & \alpha^2 & \alpha^4 \end{bmatrix}'.$$

More generally, since there exists a positive integer N such that $\mathbf{A}^N = \alpha^N \mathbf{I}$, if the measurement y_t is available, the measurements y_{t+jN}, for $j \in \mathbb{N}$, will not increase the rank of $\mathbf{O}$.

It is clear from the previous example that the probability to observe a sequence of measurements Γ_t such that $\mathbf{O}(\Gamma_t)$ has FCR depends on the structure of the system under analysis as well as on the parameters of the communication channel. Lemma 15 below states a numerical method to evaluate $\limsup_{t\to\infty} \mathbb{P}(\mathcal{N}^t)^{1/t}$ for general degenerate systems using the FSMC channel model. Furthermore, we show in Corollary 17 that this method results in a closed-form expression when the packet loss model is i.i.d..

D. Computation of $\limsup_{t\to\infty} \mathbb{P}(\mathcal{N}_k^t)^{1/t}$

We now address the issue of computing the quantity $\limsup_{t\to\infty} \mathbb{P}(\mathcal{N}_k^t)^{1/t}$. Our strategy to do so is as follows. For a given block, whose index k is omitted from the notation, we first arrange the packet arrival (binary) sequences into certain groups (which we denote below by $\mathcal{L}_{i,m}^n$). We then build different probability transition matrices ($\mathbf{D}_i$) between different sets of groups. These transitions are those which occur when a new sub-sequence is concatenated to the current one. Then, we express the quantity $\limsup_{t\to\infty} \mathbb{P}(\mathcal{N}_k^t)^{1/t}$ in terms of the spectral radii of these transition matrices.

We start with some necessary definitions. Fix $k \in \{1, \cdots, K\}$. For convenience of notation we omit the subscript k, i.e., we assume that $\mathbf{A} = \mathbf{A}_1$. Let $N \in \mathbb{N}$ be the smallest positive integer greater than or equal to ν (the order of the packet drop model) such that $\mathbf{A}^N = \alpha^N \mathbf{I}$. Since, for all $n, l \in \mathbb{N}_0$, $\mathbf{CA}^n$ and $\mathbf{CA}^{n+lN}$ are linearly dependent, we can restrict our characterization to sequences of length N. To do so, for each $t \in \mathbb{N}$, we define the map $\psi : \mathbb{B}^t \to \mathbb{B}^N$ such that, for all $n = 1, \cdots, N$, the n-th element of $\psi(S)$ is given by

$$\psi(S)[n] \triangleq \begin{cases} 1, & S[n+lN] = 1 \text{ for some } l \in \mathbb{N}_0 \\ & \text{and } 1 \le n + lN \le t, \\ 0, & \text{otherwise,} \end{cases} \tag{26}$$

i.e., ψ maps a sequence $S \in \mathbb{B}^t$ of arbitrary length t, to a sequence $\psi(S) \in \mathbb{B}^N$ of length N, such that $\operatorname{rank}(\mathbf{O}(S)) = \operatorname{rank}(\mathbf{O}(\psi(S)))$. Recall the definition of $\mathcal{N}^N$ in (17) (with the subscript k omitted). Let I be the number of sequences in $\mathcal{N}^N$ and $F_i \in \mathbb{B}^N, i = 1, \cdots, I$ be the elements of $\mathcal{N}^N$ i.e., $\mathcal{N}^N = \{F_1, \cdots, F_I\}$. Let the elements F_i be enumerated such that, for all $t \in \mathbb{N}$ and $S \in \mathbb{B}^t$, if $\psi((F_i, S)) = F_j$, then $j \ge i$ (notice that such enumeration is not unique in general). A possible way to obtain this enumeration is as follows. Partition the set $\mathcal{N}^N$ into subsets $\mathcal{F}_j$, $j = 1, \cdots, J$, such that the sequences in $\mathcal{F}_j$ have j nonzero elements. The enumeration is then done such that, all the sequences in $\mathcal{F}_j$ have smaller indexes than those in $\mathcal{F}_{j+1}$. Notice that, after doing so, $F_1 = (0, \cdots, 0)$. Now partition

$$\mathcal{N}^{nN} = \mathcal{L}_1^n \cup \cdots \cup \mathcal{L}_I^n$$

with $\mathcal{L}_i^n$ being the set of sequences S of length nN such that its associated sequence $\psi(S)$ equals F_i, i.e.,

$$\mathcal{L}_i^n \triangleq \{S \in \mathbb{B}^{nN} : \psi(S) = F_i\}. \tag{27}$$

For any $S \in \mathbb{B}^t$ with $t \ge \nu$, denote its tail sequence

$$\delta(S) \triangleq (S[t-\nu+1], \cdots, S[t]). \tag{28}$$

For every $i = 1, \cdots, I$, it is clear that the set of all $\delta(S)$ with $S \in \mathcal{L}_i^n$ and $n \in \mathbb{N}$ is a finite set. Denote this set by $\{H_{i,1}, \cdots, H_{i,M_i}\}$. Then, for each $n \in \mathbb{N}$, further partition

$$\mathcal{L}_i^n = \mathcal{L}_{i,1}^n \cup \ldots \cup \mathcal{L}_{i,M_i}^n$$

with

$$\mathcal{L}_{i,m}^n \triangleq \{S \in \mathcal{L}_i^n : \delta(S) = H_{i,m}\}. \tag{29}$$

Notice that $\mathcal{L}_{i,m}^2$ is not an empty set, since $(F_i, (0)^{N-\nu}, H_{i,m}) \in \mathcal{L}_{i,m}^2$. Now construct the $M_i \times M_i$ probability transition matrix $\mathbf{D}_i$, whose (r,c)-th entry equals the probability that $\Gamma_{(n+1)N} \in \mathcal{L}_{i,r}^{(n+1)}$ given that $\Gamma_{nN} \in \mathcal{L}_{i,c}^n$ for $n \ge 2$, i.e.,

$$[\mathbf{D}_i]_{r,c} \triangleq \mathbb{P}\left(\Gamma_{3N} \in \mathcal{L}_{i,r}^3 | \Gamma_{2N} \in \mathcal{L}_{i,c}^2\right). \tag{30}$$

By Assumption 2, $\mathbf{D}_i$ has a finite dimension and is independent of n.

Recall that $\mathbf{D}_i$ is associated to the FMO block k. In general, for $k = 1, \cdots, K$ define

$$\sigma_k \triangleq \left(\max_{1 \le i \le I} \rho(\mathbf{D}_i)\right)^{1/N}. \tag{31}$$

We can now state our result for numerically evaluating $\limsup_{t\to\infty} \mathbb{P}(\mathcal{N}_k^t)^{1/t}$ (The proof is deferred to the Appendix).

Lemma 15. *Let σ_k be as defined in* (31). *We have*

$$\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} = \sigma_k. \tag{32}$$

Example 16. Consider the FMO block in Example 14 subject to random measurement losses according to the Gilbert-Elliott model with recovery rate q and failure rate p. We have that $N = 2$ and $\mathcal{N}^N = \{F_1, F_2, F_3\}$, with $F_1 = (0,0)$, $F_2 = (0,1)$, and $F_3 = (1,0)$. We have $\mathbf{D}_1 = \mathbb{P}\left(\Gamma_{3N} = (0)^6 | \Gamma_{2N} = (0)^4\right) = (1-q)^2$. For $i = 2$, we have $H_{2,1} = (0)$ and $H_{2,2} = (1)$. Then,

$$\begin{aligned}
[\mathbf{D}_2]_{1,1} &= \mathbb{P}(\Gamma_{3N} = (0,1,0,0,0,0) | \Gamma_{2N} = (0,1,0,0)) \\
[\mathbf{D}_2]_{1,2} &= \mathbb{P}(\Gamma_{3N} = (0,1,0,1,0,0) | \Gamma_{2N} = (0,1,0,1)) \\
[\mathbf{D}_2]_{2,1} &= \mathbb{P}(\Gamma_{3N} = (0,1,0,0,0,1) | \Gamma_{2N} = (0,1,0,0)) \\
[\mathbf{D}_2]_{2,2} &= \mathbb{P}(\Gamma_{3N} = (0,1,0,1,0,1) | \Gamma_{2N} = (0,1,0,1))
\end{aligned}$$

$$\mathbf{D}_2 = \begin{bmatrix} (1-q)^2 & p(1-q) \\ q(1-q) & pq \end{bmatrix},$$

and $\max \operatorname{eig}(\mathbf{D}_2) = (1-q)^2 + pq$. For $i = 3$, we have $H_{3,1} = (0)$. Then,

$$\begin{aligned}
\mathbf{D}_3 =& \mathbb{P}(\Gamma_{3N} = (1,0,0,0,x,0) | \Gamma_{2N} = (1,0,0,0)) \\
=& (1-q)^2 + q(1-q) = 1-q,
\end{aligned}$$

where x can be either 0 or 1. Hence,

$$\begin{aligned}
\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}^t\right)^{1/t} &= \left(\max\{(1-q), (1-q)^2 + pq\}\right)^{1/2} \\
&= \begin{cases} (1-q)^{1/2}, & p \le 1-q \\ \left((1-q)^2 + pq\right)^{1/2}, & p > 1-q. \end{cases}
\end{aligned}$$

For the i.i.d. packet loss model, the corollary below explicitly expresses $\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t}$.

Corollary 17. *To simplify the notation, suppose that $\mathbf{A} = \mathbf{A}_1$, so we can omit the subindex k. Consider an i.i.d. network model with packet receiving probability λ. Let $F_i \in \mathbb{B}^N, i = 1, \cdots, I$ such that $\bigcup_{i=1}^{I} F_i = \mathcal{N}^N$. Let ζ_i be the number of zeros in the sequence $F_i \in \mathcal{N}^N$. Then,*

$$\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}^t\right)^{1/t} = \max_i (1-\lambda)^{\zeta_i/N}. \tag{33}$$

Proof: For a given i, let z_j, $1 \le j \le \zeta_i$, be all the integers such that $F_i[z_j] = 0$. Using the i.i.d. assumption, we have

$$\begin{aligned}
\mathbf{D}_i &= \mathbb{P}(\Gamma_{3N} \in \mathcal{L}_i^3 | \Gamma_{2N} \in \mathcal{L}_i^2) \\
&= \mathbb{P}(g_{2N+z_1-1} = 0 \cap \cdots \cap g_{2N+z_{\zeta_i}-1} = 0) \\
&= (1-\lambda)^{\zeta_i}.
\end{aligned}$$

The result then follows from Lemma 15. ■

Example 18. Consider the system in Example 14 with an i.i.d. network model with packet receiving probability λ. We have that $N = 2$, $\zeta_1 = 2$, $\zeta_2 = \zeta_3 = 1$, $\mathbf{D}_1 = (1-\lambda)^2$, $\mathbf{D}_2 = \mathbf{D}_3 = (1-\lambda)$. Then, from Corollary 17, we have

$$\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}^t\right)^{1/t} = (1-\lambda)^{1/2}. \tag{34}$$

Using Theorem 7, the critical probability to receive a measurement λ_c can be obtained by solving $|\alpha|^2(1-\lambda_c)^{1/2} = 1$, hence, $\lambda_c = 1 - |\alpha|^{-4}$. That is, if $\lambda < \lambda_c$, then $G = \infty$ and if $\lambda > \lambda_c$, then the AEEC is bounded. Note that this critical value matches the result reported in [22, Theorem 4].

E. Procedure for stability test

We now present a procedure that summarizes how the main results of the paper can be used to determine if the KFIO corresponding to a given system and communication channel is stable.

1) *Obtain the FSMC model of the communication channel.* The FSMC model is determined by the probabilities

$$\mathbb{P}\left(g_\nu = 1 | \Gamma_\nu = S\right) \text{ for all } S \in \mathbb{B}^\nu, \tag{35}$$

where ν is the order of the FSMC model. There are several methods to obtain these probabilities from a set of channel observations, see e.g., [28] and [29].

2) *Partition the system into FMO blocks $(\mathbf{A}_k, \mathbf{C}_k)$, $k = 1, \cdots, K$, according to* (10)-(13). Recall that each FMO block has an associated scalar α_k which equals the magnitude of all the eigenvalues of $\mathbf{A}_k$.

3) *Determine whether the system is degenerate or non-degenerate using to Definition 9.*

1) If the system is non-degenerate: the associated KFIO is stable if

$$|\alpha_1|^2 \mathbb{P}\left(g_t = 0 | g_{t-\nu} = 0, \cdots, g_{t-1} = 0\right) < 1 \tag{36}$$

and unstable if

$$|\alpha_1|^2 \mathbb{P}\left(g_t = 0 | g_{t-\nu} = 0, \cdots, g_{t-1} = 0\right) > 1. \tag{37}$$

2) If the system is degenerate:
   a) *For each $k = 1, \cdots K$, compute σ_k. This in turn requires:*
      i) *Enumerate the sequences $F_i \in \mathbb{B}^N, i = 1, \cdots, I$, such that $\mathcal{N}_k^N = \{F_1, \cdots, F_I\}$.*
      ii) *For each $i = 1, \cdots, I$, enumerate all the possible tails $H_{i,m} \in \mathbb{B}^\nu$ $m = 1, \cdots, M_i$, such that, for any $n \in \mathbb{N}$ and $S \in \mathbb{B}^{nN}$, if $\psi(S) = F_i$, then $\delta(S) = H_{i,m}$, for some m.*
      iii) *For $1 \le r, c \le M_i$, compute each entry $[\mathbf{D}_i]_{r,c}$ using* (30) *and* (35).
      iv) *Compute σ_k using* (31).
   b) The associated KFIO is stable if

$$|\alpha_k|^2 \sigma_k < 1, \text{ for all } k \in \{1, \cdots, K\}, \tag{38}$$

   and unstable if

$$|\alpha_k|^2 \sigma_k > 1, \text{ for some } k \in \{1, \cdots, K\}. \tag{39}$$

IV. Characterization of measurement arrival patterns

In a standard Kalman filter, where all past measurements are available at the estimator, a sufficient condition for stability of the filter is that the observability matrix $\left[\mathbf{C}' \ (\mathbf{CA})' \ \cdots \ \left(\mathbf{CA}^{\iota-1}\right)'\right]'$ has FCR, where ι is the observability index [30, pp. 157]. When measurements are only intermittently available, the probability to receive measurements such that $\mathbf{O}\left(\Gamma_t\right)$ has FCR plays a major role in the study of the stability of KFIO. Identifying the sequences S such that $\mathbf{O}(S)$ has FCR has been a challenge and a source of the gap between necessary and sufficient conditions for the stability of the KFIO for general systems. In this section we summarize results which provide answers to this question in some particular cases, and point out how a byproduct of our analysis (i.e., Corollary 27) gives an answer to the general case considered in this work.

In the particular case when $\mathbf{C}$ has FCR, the question becomes trivial, since if $S\left[\tau\right] = 1$ for some $0 \le \tau \le t-1$, i.e., at least one measurement is received, then $\mathbf{O}(S)$ has FCR. This property was used in [1], [8], [10], [12]. For second-order systems which are non-degenerate, [9] showed that any two measurements are enough. Other than these cases, a precise characterization is not available, and only sufficient conditions on the loss patterns have been considered. This leads to conservative sufficient conditions on the stability of the KFIO, and is the source of the gap between necessary and sufficient conditions in some works. For example, a sufficient condition to obtain a ROM with FCR is to receive a sequence of at least ι (the observability index) consecutive measurements. This idea was exploited in [8], [10], [14]. In [13], the authors showed that for systems whose unstable eigenvalues have distinct magnitudes, if the time interval between received measurements is large enough, then the ROM has FCR when the number of measurements received matches the number of unstable eigenvalues. This idea was also used in [9]. In [11], the authors claimed that for non-degenerate systems, if the number of available measurements is the same as the number of subsystems with unstable eigenvalues, then the ROM has FCR. However, there is a flaw in their argument. We show below that this claim is not true in general.

Example 19. Let $\mathbf{A} = \operatorname{diag}(2, 3, -5)$, $\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, $\mathbf{Q} = \mathbf{I}$, and $\mathbf{R} = 1$. Notice that the system is observable. According to [11, Definitions 1-5], the system is non-degenerate and has 3 quasi-equiblocks. Let $\Gamma_t = \left((0)^{t-4}, 1, 1, 0, 1\right)$. Note that

$$\operatorname{rank}(\mathbf{O}(\Gamma_t)) = \operatorname{rank}\left(\begin{bmatrix} \mathbf{CA}^{t-4} \\ \mathbf{CA}^{t-3} \\ \mathbf{CA}^{t-1} \end{bmatrix}\right) = 2,$$

i.e., even if the number of measurements available matches the number of quasi-equiblocks, $\mathbf{O}(\Gamma_t)$ can fail to have FCR. In [11, eq. (35)], the matrix $\mathbf{U}$ plays a similar role as our ROM $\mathbf{O}(\Gamma_t)$. We believe that there is a misprint in the definition of $\mathbf{U}$ in [11]: the powers i_j should be replaced by $k_j - k$. In any case, using this example, the matrix $\mathbf{U}$ is singular, causing [11, Theorem 4] not to hold. Adopting the notation used in [11], $\Gamma_t = \left((0)^{t-4}, 1, 1, 0, 1\right)$ is equivalent to $i_1 = 1, i_2 = 2, i_3 = 1, i_4 = t - 4$. For some $\sigma > 0$, we have

$$\begin{aligned} h^{i_1} \circ g \circ h^{i_2-1} \circ g \circ \cdots \circ h^{i_l-1} \circ g \circ h^{i_{l+1}-1}(\Sigma) \\ \geq h \circ g \circ h \circ g \circ g \circ h^{t-5}(\sigma \mathbf{I}). \quad (40) \end{aligned}$$

Also,

$$h^{t-5}(\sigma \mathbf{I}) = \mathbf{A}^{2(t-5)}\sigma + \sum_{j=0}^{t-6} \mathbf{A}^{2j} \geq \overline{\sigma}\mathbf{I},$$

where $\overline{\sigma}$ can be made arbitrarily large by increasing t. Hence,

$$h \circ g \circ h \circ g \circ g \circ h^{t-5}(\sigma \mathbf{I}) \geq h \circ g \circ h \circ g \circ g(\overline{\sigma}\mathbf{I}).$$

After some algebra, we obtain

$$\begin{aligned} &\operatorname{trace}(h \circ g \circ h \circ g \circ g(\overline{\sigma}\mathbf{I})) \\ =&\frac{136 \times 10^9\overline{\sigma}^3 + 793 \times 10^9\overline{\sigma}^2 + 10.3 \times 10^9\overline{\sigma} + 300 \times 10^6}{4021888\overline{\sigma}^2 + 870238\overline{\sigma} + 29497} \\ \geq&\overline{\sigma}. \end{aligned}$$

We conclude that the EC cannot be bounded by a function of i_1, i_2, i_3, contradicting the claim of [11, Theorem 4].

Characterizing the sequences of received measurements that produce a bounded EC is important not only for stability analysis, but also for designing encoders, communication protocols, and efficient power transmission control strategies, that guarantee stability in a stochastic sense. As Example 19 above shows, [11, Theorem 4] does not hold in general, and therefore cannot be used to determine the number of measurements required to guarantee stability. On the other hand, this number can be obtained using Corollary 27 in Section V-C. As shown in Example 19, this number is not necessarily the number of quasi-equiblocks of the system.

V. Proof of the main result

In this section we prove the claims in Section III. Before presenting the formal proofs, we give a brief overview of the ideas in the proofs.

To prove the necessary condition in (19) we use the fact stated in Lemma 21, which says that if $\Gamma_t \in \mathcal{N}_k^t$, then the EC grows at rate at least $|\alpha_k|^{2t}$. When this is used in the definition of G in (7), the necessary condition thus depends on the convergence of the series

$$\sum_{S \in \mathcal{N}_k^t} \mathbf{\Psi}(\mathbf{P}_0, S)\, \mathbb{P}(\Gamma_t = S)$$

for each k when $t \to \infty$.

The proof of the sufficient condition (18) is technically more involved. There are two main steps in the proof. In the first step, we present an alternative sufficient condition in Lemma 23. To prove Lemma 23, we use the fact that the location of the kernel of $\mathbf{O}(\Gamma_t)$ can be used to bound the growth rate of the EC (Lemma 20). We then study a condition (Lemma 25) on the received measurements that guarantees the desired kernel location. In the second step, we show that this condition is equivalent to the one given in Lemma 23. We do so in Lemma 24, showing that the asymptotic probability of the condition in Lemma 23 matches the one in (18).

A. Bounds on the Growth Rate of $\|\mathbf{P}_t\|$

We will first provide lower and upper bounds on the growth rate of $\|\mathbf{P}_t\|$. It turns out that the growth rate of $\|\mathbf{P}_t\|$ is determined by the location of the kernel of $\mathbf{O}(\Gamma_t)$. Recall from (14) that

$$\mathbf{z}_t = \mathbf{O}(\Gamma_t)\mathbf{x}_0 + \mathbf{f}(\Gamma_t) \quad (41)$$
$$\mathbf{x}_t = \mathbf{A}^t\mathbf{x}_0 + \mathbf{q}_t, \quad (42)$$

where $\mathbf{q}_t = \sum_{j=1}^{t} \mathbf{A}^{t-j}\mathbf{w}_{j-1}$. To simplify the notation, in this section we will omit the argument (Γ_t) when it is clear from the context. From [31, Ch. 5, Theorem 2.1], we have

$$\mathbf{P}_t = \mathbf{\Sigma}_{\mathbf{x}} - \mathbf{\Sigma}_{\mathbf{x},\mathbf{z}}\mathbf{\Sigma}_{\mathbf{z}}^{\dagger}\mathbf{\Sigma}_{\mathbf{x},\mathbf{z}}^{*} \quad (43)$$

with

$$\begin{aligned} \mathbf{\Sigma}_{\mathbf{x}} &= \mathbf{A}^t\mathbf{P}_0\mathbf{A}^{*t} + \mathbb{E}(\mathbf{q}_t\mathbf{q}_t^*), \\ \mathbf{\Sigma}_{\mathbf{z}} &= \mathbf{O}_t\mathbf{P}_0\mathbf{O}_t^* + \mathbb{E}(\mathbf{f}_t\mathbf{f}_t^*), \\ \mathbf{\Sigma}_{\mathbf{x},\mathbf{z}} &= \mathbf{A}^t\mathbf{P}_0\mathbf{O}_t^* + \mathbb{E}(\mathbf{q}_t\mathbf{f}_t^*). \end{aligned}$$

Let $\mathbf{e}_{k,j}$ be the column vector with a 1 in one entry and zeros otherwise such that $\mathbf{e}_{k,j}'\mathbf{A}\mathbf{e}_{k,j}$ equals the j-th diagonal entry of the k-th block $\mathbf{A}_k$ of $\mathbf{A}$. Let also $\mathcal{E}_k = \{\mathbf{e}_{k,1}, \cdots, \mathbf{e}_{k,J_k}\}$. The following lemma states an upper bound on the growth rate of $\|\mathbf{P}_t\|$.

Lemma 20. *If* $\ker(\mathbf{O}) \subseteq \operatorname{span}(\{\mathcal{E}_k, \cdots, \mathcal{E}_K\})$ *for some* $1 \leq k \leq K$, *then there exists* $l_t > 0$, *independent of* $\mathbf{P}_0$, *such that, for all* $t \in \mathbb{N}$,

$$\|\mathbf{P}_t\| \leq |\alpha_k|^{2t}\|\mathbf{P}_0\| + l_t. \quad (44)$$

Also, if $\ker(\mathbf{O}) = \mathbf{0}$, *then*

$$\|\mathbf{P}_t\| \leq l_t. \quad (45)$$

Proof: From (41)-(42), we have that

$$\mathbf{z}_t = \mathbf{O}\mathbf{A}^{-t}\mathbf{x}_t + \mathbf{f} - \mathbf{O}\mathbf{A}^{-t}\mathbf{q}_t. \quad (46)$$

Consider the following sub-optimal estimator

$$\hat{\mathbf{x}}_t = \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{z}_t. \tag{47}$$

Using (46) and (42) in the above we have

$$\begin{aligned}\hat{\mathbf{x}}_t &= \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\mathbf{x}_t + \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{f} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\mathbf{q}_t \\ &= \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\mathbf{A}^t\mathbf{x}_0 + \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{f}.\end{aligned}$$

Then, the error $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$ is given by

$$\begin{aligned}\tilde{\mathbf{x}}_t &= \left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\mathbf{A}^t\mathbf{x}_0 + \mathbf{q}_t - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{f} \\ &= \left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\mathbf{A}^t\mathbf{x}_0 + \mathbf{U},\end{aligned}$$

with $\mathbf{U} = \mathbf{q}_t - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{f}$. Since $\mathbf{P}_t$ comes from the optimal estimator, we have

$$\begin{aligned}\mathbf{P}_t &\leq \mathbb{E}\left(\tilde{\mathbf{x}}_t\tilde{\mathbf{x}}_t^*\right) && (48)\\ &= \left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\mathbf{A}^t\mathbf{P}_0\mathbf{A}^{*t} && (49)\\ &\quad\times \left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right) + \mathbb{E}\left(\mathbf{U}\mathbf{U}^*\right). && (50)\end{aligned}$$

Since $\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)$ is the projection onto the kernel of $\mathbf{O}\mathbf{A}^{-t}$, (45) follows by making

$$l_t = \max_{S\in\mathbb{B}^t} \left\|\mathbb{E}\left(\mathbf{U}\mathbf{U}^*\right)\right\|. \tag{51}$$

Now, $\ker(\mathbf{O}) \subseteq \operatorname{span}\left(\{\mathcal{E}_k, \cdots, \mathcal{E}_K\}\right)$ for some $1 \leq k \leq K$ implies that $\ker\left(\mathbf{O}\mathbf{A}^{-t}\right) \subseteq \operatorname{span}\left(\{\mathcal{E}_k, \cdots, \mathcal{E}_K\}\right)$. So

$$\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t} = \operatorname{diag}(\mathbf{0}_1, \cdots, \mathbf{0}_{k-1}, \mathbf{X}), \tag{52}$$

where $\mathbf{X}$ is a non-zero matrix with appropriate dimensions, and $\mathbf{0}_j$ is a square matrix of zeros with the same dimension of $\mathbf{A}_j$. Let $\tilde{\mathbf{A}} = \operatorname{diag}\left(\mathbf{0}, \cdots, \mathbf{0}, \mathbf{A}_k, \cdots, \mathbf{A}_K\right)$. From (52),

$$\begin{aligned}&\mathbf{A}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right) = \\ &\qquad = \tilde{\mathbf{A}}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right). \qquad (53)\end{aligned}$$

From (50) and (51) we have

$$\begin{aligned}\|\mathbf{P}_t\| &\leq \left\|\left(\mathbf{P}_0^{1/2}\mathbf{A}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\right)^* \times\right. \\ &\qquad \left.\times \mathbf{P}_0^{1/2}\mathbf{A}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\right\| + l_t \\ &= \left\|\mathbf{P}_0^{1/2}\mathbf{A}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\right\|^2 + l_t \\ &\leq \|\mathbf{P}_0\| \left\|\mathbf{A}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\right\|^2 + l_t \\ &\overset{(a)}{=} \|\mathbf{P}_0\| \left\|\tilde{\mathbf{A}}^{*t}\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\right\|^2 + l_t \\ &\leq \|\mathbf{P}_0\| \left\|\tilde{\mathbf{A}}^{*t}\right\|^2 \left\|\left(\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}\right)\right\|^2 + l_t \\ &\overset{(b)}{=} |\alpha_k|^{2t} \|\mathbf{P}_0\| + l_t,\end{aligned}$$

where (a) follows from (53) and (b) follows since $\mathbf{I} - \left(\mathbf{O}\mathbf{A}^{-t}\right)^\dagger \mathbf{O}\mathbf{A}^{-t}$ is a projection (and therefore has unit norm). This concludes the proof. ■

Next, we give a lower bound on the growth rate of $\|\mathbf{P}_t\|$.

Lemma 21. *If $\ker(\mathbf{O}) \cap \operatorname{span}(\mathcal{E}_k) \neq \mathbf{0}$ for some k, then, for all $t \in \mathbb{N}$,*

$$\|\mathbf{P}_t\| \geq |\alpha_k|^{2t} \left\|\mathbf{P}_0^{-1}\right\|^{-1}.$$

Proof: A lower bound for $\mathbf{P}_t$ can be obtained by making $\mathbf{Q} = \mathbf{0}$, $\mathbf{R} = \mathbf{0}$, and $\mathbf{P}_0 = \underline{\lambda}\mathbf{I}$, where $\underline{\lambda} = \left\|\mathbf{P}_0^{-1}\right\|^{-1}$ is the smallest eigenvalue of $\mathbf{P}_0$. Then, from (43), we have

$$\begin{aligned}\mathbf{P}_t &\geq \underline{\lambda}\left(\mathbf{A}^t\mathbf{A}^{*t} - \mathbf{A}^t\mathbf{O}^*\left(\mathbf{O}\mathbf{O}^*\right)^\dagger \mathbf{O}\mathbf{A}^{*t}\right) \\ &= \underline{\lambda}\mathbf{A}^t\left(\mathbf{I} - \mathbf{O}^*\left(\mathbf{O}\mathbf{O}^*\right)^\dagger \mathbf{O}\right)\mathbf{A}^{*t} \\ &= \underline{\lambda}\mathbf{A}^t\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)\mathbf{A}^{*t}.\end{aligned}$$

Since $\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)$ is a projection, we have $\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)^2 = \mathbf{I} - \mathbf{O}^\dagger\mathbf{O}$, thus $\|\mathbf{P}_t\| \geq \underline{\lambda}\left\|\mathbf{A}^t\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)\right\|^2$. Now, let $\mathbf{x} \neq \mathbf{0} \in \ker(\mathbf{O}) \cap \operatorname{span}(\mathcal{E}_k)$. Notice that $\mathbf{O}\mathbf{x} = \mathbf{0}$ and $\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)\mathbf{x} = \mathbf{x}$. Also, $\mathbf{A}\mathbf{x} = \operatorname{diag}\left(\mathbf{0}_1, \cdots, \mathbf{0}_{k-1}, \mathbf{A}_k\tilde{\mathbf{x}}, \mathbf{0}_{k+1}, \cdots, \mathbf{0}_K\right)$, where $\mathbf{0}_j$ is a zero matrix with the same dimensions of $\mathbf{A}_j$ and $\tilde{\mathbf{x}}$ is a non-zero vector. We have

$$\left\|\mathbf{A}^t\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)\mathbf{x}\right\| = \left\|\mathbf{A}^t\mathbf{x}\right\| = |\alpha_k|^t \|\mathbf{x}\|.$$

Hence, $\left\|\mathbf{A}^t\left(\mathbf{I} - \mathbf{O}^\dagger\mathbf{O}\right)\right\| \geq |\alpha_k|^t$. Then, $\|\mathbf{P}_t\| \geq \underline{\lambda}|\alpha_k|^{2t} = |\alpha_k|^{2t}\left\|\mathbf{P}_0^{-1}\right\|^{-1}$. ■

B. Proof of the Necessary Condition in Theorem 7

Recall that $\mathbf{P}_0 \in \mathbb{R}^{n\times n}$. From (7) we have

$$\begin{aligned}G_t &= \left\|\sum_{S\in\mathbb{B}^t} \mathbb{P}\left(\Gamma_t = S\right)\mathbf{\Psi}(\mathbf{P}_0, S)\right\| \\ &\geq \frac{1}{n}\operatorname{Tr}\left(\sum_{S\in\mathbb{B}^t} \mathbb{P}\left(\Gamma_t = S\right)\mathbf{\Psi}(\mathbf{P}_0, S)\right) \\ &= \frac{1}{n}\sum_{S\in\mathbb{B}^t} \mathbb{P}\left(\Gamma_t = S\right)\operatorname{Tr}\left(\mathbf{\Psi}(\mathbf{P}_0, S)\right) \\ &\geq \frac{1}{n}\sum_{S\in\mathbb{B}^t} \mathbb{P}\left(\Gamma_t = S\right)\left\|\mathbf{\Psi}(\mathbf{P}_0, S)\right\| \\ &\geq \max_{1\leq k\leq K} \frac{1}{n}\sum_{S\in\mathcal{N}_k^t} \mathbb{P}\left(\Gamma_t = S\right)\left\|\mathbf{\Psi}(\mathbf{P}_0, S)\right\|,\end{aligned}$$

Notice from (17) that if $S \in \mathcal{N}_k^t$ then $\ker(\mathbf{O}) \cap \operatorname{span}(\mathcal{E}_k) \neq \mathbf{0}$. Using Lemma 21, we have

$$\begin{aligned}G_t &\geq \max_{1\leq k\leq K} \frac{1}{n}\sum_{S\in\mathcal{N}_k^t} \mathbb{P}\left(\Gamma_t = S\right)|\alpha_k|^{2t}\left\|\mathbf{P}_0^{-1}\right\|^{-1} \\ &= \max_{1\leq k\leq K} \frac{1}{n}|\alpha_k|^{2t}\left\|\mathbf{P}_0^{-1}\right\|^{-1}\mathbb{P}\left(\mathcal{N}_k^t\right) \\ &= \max_{1\leq k\leq K} \left(|\alpha_k|^2\mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t}\right)^t \frac{\left\|\mathbf{P}_0^{-1}\right\|^{-1}}{n}.\end{aligned}$$

Hence, if

$$\limsup_{t\to\infty} |\alpha_k|^2\mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} > 1$$

for some $1 \leq k \leq K$, then $\lim_{t\to\infty} G_t = \infty$.

C. Proof of the Sufficient Condition in Theorem 7

In this section we first introduce an alternative sufficient condition for the norm of the AEEC G to be bounded. We then show that this alternative condition is equivalent to the sufficient condition in Theorem 7. We start with the following definition.

Definition 22. A matrix $\mathbf{M}$ is said to have full column rank with strength $Q \in \mathbb{N}_0$, denoted by $\mathrm{FCR}(Q)$, if the matrix obtained after removing any Q rows from $\mathbf{M}$ still has FCR.

For $k = 1, \cdots, K$, let

$$\mathcal{N}_{k,Q}^t \triangleq \{S \in \mathbb{B}^t : \mathbf{O}_k(S) \text{ does not have } \mathrm{FCR}(Q)\}. \quad (54)$$

Notice that $\mathcal{N}_{k,0}^t = \mathcal{N}_k^t$. We now present an alternative sufficient condition for the asymptotic norm of the AEEC to be bounded.

Lemma 23. *There exists $Q \in \mathbb{N}_0$ such that, if*

$$|\alpha_k|^2 \limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_{k,Q}^t\right)^{1/t} < 1 \text{ for all } k = 1, \cdots, K, \quad (55)$$

then $G < \infty$.

The sufficient condition stated in Lemma 23 seems stronger than the one in Theorem 7. We show now that, under Assumption 2 they are indeed equivalent.

Lemma 24. *If Assumption 2 holds, then, for any $Q \in \mathbb{N}$,*

$$\limsup_{t\to\infty} \; \mathbb{P}\left(\mathcal{N}_k^t\right)^{1/t} = \limsup_{t\to\infty} \; \mathbb{P}\left(\mathcal{N}_{k,Q}^t\right)^{1/t}. \quad (56)$$

The proof of Lemma 24 is given in the appendix. Before proving Lemma 23, we introduce some technical results. The next lemma provides a sufficient condition on the sequences $S \in \mathbb{B}^t$ to guarantee a certain convenient location of the kernel of $\mathbf{O}(S)$. This result is required to show the decoupling property of our system decomposition, as mentioned in Remark 8.

Lemma 25. *Let $\mathbf{A}$ be a diagonal matrix and $\mathbf{O}_k$ be as defined in* (16)*. There exists $Q \in \mathbb{N}_0$ such that, for any $1 \le k \le K$, if $\mathbf{O}_k$ has $\mathrm{FCR}(Q)$, then $\ker(\mathbf{O}) \perp \mathrm{span}(\mathcal{E}_k)$.*

The proof of Lemma 25 uses the following result.

Lemma 26 (Immediate consequence of [32], Theorem 1.1)**.** *Let $\alpha_1, \cdots, \alpha_K \in \mathbb{C}$ with $\alpha_k \neq 0, k = 1, \cdots, K$ and $(\alpha_k/\alpha_j)^t \neq 1$ for all $k \neq j$ and all $t \in \mathbb{N}$. Let $a_1, \cdots, a_K \in \mathbb{C}$ with $a_k \neq 0$ for some $1 \le k \le K$. Then, there exists at most a finite number L of non-negative integers t such that*

$$a_1\alpha_1^t + a_2\alpha_2^t + \cdots + a_K\alpha_K^t = 0. \quad (57)$$

Proof of Lemma 25: Let $\mathbf{O}(S)\mathbf{v} = \mathbf{0}$, with $\mathbf{v} = [\mathbf{v}_1' \; \cdots \; \mathbf{v}_K']'$. To show the result is enough to show that, if $\mathbf{O}_k(S)$ has $\mathrm{FCR}(Q)$, then $\mathbf{v}_k = \mathbf{0}$. We do so by contradiction. Let $\mathbf{a} = [a_1 \, a_2 \, \ldots \, a_K]$. Suppose $\mathbf{v}_k \neq \mathbf{0}$. From Lemma 26, there exists at most L values of $t \in \mathbb{N}$ such that (57) holds for some $\mathbf{a} \neq 0$. Recall from (12) that for each $\tilde{\mathbf{A}}_k$ there exists integers N_k such that $\tilde{\mathbf{A}}_k^{N_k} = \mathbf{I}$. Let $\tilde{Q} = NL + 1$, where $N \in \mathbb{N}$ is such that $\mathrm{diag}(\tilde{\mathbf{A}}_1, \cdots, \tilde{\mathbf{A}}_K)^N = \mathbf{I}$. Let $Q = p\tilde{Q}$ and recall that p is the dimension of $\mathbf{y}_t$. Also, for $j = 1, \cdots p$, let $\mathbf{C}_{k,j}$ be the j-th row of $\mathbf{C}_k$. Then, $\mathbf{O}_k(S)$ having $\mathrm{FCR}(Q)$ implies that there exists j and t_m, $m = 1, \cdots, \tilde{Q}$, such that $S[t_m] = 1$ and $\mathbf{C}_{k,j}\tilde{\mathbf{A}}_k^{t_m}\mathbf{v}_k \neq \mathbf{0}$. This in turn implies that there exists some $r \in \{0, \cdots, N-1\}$ for which the set $\mathcal{T}_r(S) \triangleq \{t : S[t] = 1, \mathbf{C}_{k,j}\tilde{\mathbf{A}}_k^t\mathbf{v}_k \neq \mathbf{0} \text{ and } t \bmod N = r\}$ has at least $L + 1$ elements. Let $t \in \mathcal{T}_r(S)$ and define $a_k = \mathbf{C}_{k,j}\tilde{\mathbf{A}}_k^t\mathbf{v}_k$, for $k = 1, \cdots, K$. Then, $\mathbf{a} \neq \mathbf{0}$ since $a_k \neq 0$. Hence, from Lemma 26, (57) can hold for at most L values of t, and therefore cannot hold for all $t \in \mathcal{T}_r(S)$. Thus, $\mathbf{O}(S)\mathbf{v} = \mathbf{0}$ cannot hold, which implies that $\mathbf{v}_k$ must be $\mathbf{0}$. ∎

The following corollary is a byproduct of Lemmas 21, 20 and 25, which characterizes the growth rate of the EC according to the sequence of available measurements.

Corollary 27. *Let $\mathbf{A}$ and $\mathbf{C}$ be partitioned as in* (10)-(13)*. The following holds true.*

1. If $\mathbf{O}_k(\Gamma_t)$ does not have FCR, then

$$\|\mathbf{P}_t\| \ge |\alpha_k|^{2t} \left\|\mathbf{P}_0^{-1}\right\|^{-1}. \quad (58)$$

2. There exists $Q \in \mathbb{N}$ such that if $\mathbf{O}_j(\Gamma_t)$ has $FCR(Q)$, for $j = 1, \cdots, k-1$, then

$$\|\mathbf{P}_t\| \le |\alpha_k|^{2t} \|\mathbf{P}_0\| + l_t, \quad (59)$$

where $l_t > 0$ is a constant independent of $\mathbf{P}_0$.

Proof: 1. If $\mathbf{O}_k(\Gamma_t)$ does not have FCR, then $\ker(\mathbf{O}(\Gamma_t)) \cap \mathrm{span}(\mathcal{E}_k) \neq \mathbf{0}$ and (58) follows from Lemma 21.

2. If $\mathbf{O}_j(\Gamma_t)$ has $\mathrm{FCR}(Q)$ for $j = 1, \cdots, k-1$, then from Lemma 25, we have $\ker(\mathbf{O}(\Gamma_t)) \perp \mathrm{span}(\mathcal{E}_j), j = 1, \cdots, k-1$, or equivalently, $\ker(\mathbf{O}(\Gamma_t)) \subseteq \mathrm{span}(\xi_k, \cdots, \xi_K)$. Then (59) follows from Lemma 20. ∎

Remark 28. Corollary 27 states the relationship between the sequence of received measurements Γ_t and the growth rate of the norm $\|\mathbf{P}_t\|$ of the EC. This is an important byproduct of our result, as this relationship can be used in a number of problems, such as devising control strategies for the transmission power used in wireless sensor networks, while maintaining a desired performance [33], [34], sensor placement algorithms [35]–[37], and network communication protocols for state estimation and control [20], [38], [39], etc.

The next two lemmas are used to obtain an upper bound on the AEEC. Recall from Assumption 2 that the Markov chain that describes the packet loss process is irreducible and has a stationary distribution. Also, from the time reversal property [40, Theorem 1.9.1], we have that, for all $\mu \ge 1$, $\mathbb{P}(g_t = 1 | g_{t+1}, \cdots, g_{t+\nu+\mu}) = \mathbb{P}(g_t = 1 | g_{t+1}, \cdots, g_{t+\nu})$. Define $\Delta_t = (g_t, \cdots, g_{t+\nu-1})$, i.e., Δ_t is the set of ν samples, immediately following Γ_t. Also, for $S \in \mathbb{B}^t$, with $t \ge \nu$, define $\phi(S)$ to be the first ν samples from S, i.e.,

$$\phi(S) \triangleq (S[1], \cdots, S[\nu]).$$

Define the mapping

$$\xi_T(x) = \max_{Y \in \mathbb{B}^\nu} \mathrm{Tr}\left(\sum_{S \in \mathbb{B}^T} \mathbb{P}(\Gamma_T = S \,|\Delta_T = Y\,)\, \mathbf{\Psi}(x\mathbf{I}, S)\right). \quad (60)$$

Notice that for $T = 1$ with the i.i.d. packet loss model, $x \ge \xi_T(x)$ is a scalar version of the ARE studied in [1].

In Lemma 30, we show that if there exists $x > 0 \in \mathbb{R}$ such that $x \geq \xi_T(x)$ for some T, then the norm of the AEEC is bounded. The next lemma is required to show that result.

Lemma 29. *For $T, N \in \mathbb{N}$ and $x \geq 0$, we have*

$$\xi_{T+V}(x) \leq \xi_V \circ \xi_T(x). \tag{61}$$

Proof: Let $S \in \mathbb{B}^T$ and $U \in \mathbb{B}^V$. We have

$$\xi_{T+V}(x) = \max_{Y \in \mathbb{B}^\nu} \operatorname{Tr}\left[\sum_{(S,U)\in\mathbb{B}^{T+V}} \mathbf{\Psi}\left(x\mathbf{I}, (S,U)\right)\right. \tag{62}$$

$$\left.\times \mathbb{P}\left(\Gamma_{T+V} = (S,U) \,|\Delta_{T+V} = Y\right)\right]. \tag{63}$$

Notice that

$$\begin{aligned}&\mathbb{P}\left(\Gamma_{T+V} = (S,U) \,|\Delta_{T+V} = Y\right)\\ &\quad= \mathbb{P}\left(\Gamma_T = S \,|\Delta_T = \phi\left((U,Y)\right)\right) \times\\ &\quad\times \mathbb{P}\left(g_T, \cdots, g_{T+V-1} = U \,|\Delta_{T+V} = Y\right).\end{aligned} \tag{64}$$

Since the Markov process describing the measurement losses is stationary, we have

$$\begin{aligned}&\mathbb{P}\left(g_T, \cdots, g_{T+V-1} = U \,|\Delta_{T+V} = Y\right) =\\ &\qquad\mathbb{P}\left(\Gamma_V = U \,|\Delta_V = Y\right).\end{aligned} \tag{65}$$

From (63)-(65), we have

$$\begin{aligned}&\xi_{T+V}(x)\\ &= \max_{Y\in\mathbb{B}^\nu} \operatorname{Tr}\left(\sum_{S\in\mathbb{B}^T}\sum_{U\in\mathbb{B}^V} \mathbb{P}\left(\Gamma_T = S \,|\Delta_T = \phi\left((U,Y)\right)\right) \times\right.\\ &\quad\left.\times\ \mathbb{P}\left(\Gamma_V = U \,|\Delta_V = Y\right) \mathbf{\Psi}\left(\mathbf{\Psi}\left(x\mathbf{I}, S\right), U\right)\right).\end{aligned}$$

Using the concavity of $\mathbf{\Psi}(\cdot, U)$ [1, Lemma 1e], we have

$$\begin{aligned}\xi_{T+V}(x) &\leq \max_{Y\in\mathbb{B}^\nu} \operatorname{Tr}\left(\sum_{U\in\mathbb{B}^V} \mathbb{P}\left(\Gamma_V = U \,|\Delta_V = Y\right) \times\right.\\ &\quad\left.\times \mathbf{\Psi}\left(\Xi_T(x), U\right)\right)\end{aligned}$$

with $\Xi_T(x) = \sum_{S\in\mathbb{B}^T} \mathbb{P}\left(\Gamma_T = S \,|\Delta_T = \phi((U,Y))\right) \mathbf{\Psi}(x\mathbf{I}, S)$. Now, since $\mathbf{\Psi}(x\mathbf{I}, S) \geq 0$,

$$\begin{aligned}&\Xi_T(x)\\ &\leq \sum_{S\in\mathbb{B}^T} \mathbb{P}\left(\Gamma_T = S \,|\Delta_T = \phi((U,Y))\right) \operatorname{Tr}\left(\mathbf{\Psi}(x\mathbf{I}, S)\right)\mathbf{I}\\ &\leq \max_{W\in\mathbb{B}^\nu} \operatorname{Tr}\left(\sum_{S\in\mathbb{B}^T} \mathbb{P}\left(\Gamma_T = S \,|\Delta_T = W\right) \mathbf{\Psi}(x\mathbf{I}, S)\right)\mathbf{I}\\ &=\xi_T(x)\mathbf{I}.\end{aligned}$$

Then, using the monotonicity of $\mathbf{\Psi}(\cdot, U)$ [1, Lemma 1c],

$$\begin{aligned}&\xi_{T+V}(x)\\ &\leq \max_{Y\in\mathbb{B}^\nu} \operatorname{Tr}\left(\sum_{U\in\mathbb{B}^V} \mathbb{P}\left(\Gamma_V = U \,|\Delta_V = Y\right) \mathbf{\Psi}\left(\xi_T(x)\mathbf{I}, U\right)\right)\\ &=\xi_V\left(\xi_T\left(x\right)\right) = \xi_V \circ \xi_T\left(x\right).\end{aligned}$$

■

Lemma 30. *If there exists $T \in \mathbb{N}$ and $x > 0$ such that $x > \xi_T(x)$, then there exists $x_0 > 0$ and $l \in \mathbb{R}$ such that $G \leq x_0|\alpha_1|^{2(T-1)} + l$.*

Proof: From the concavity of $\mathbf{\Psi}(\cdot, S)$ [1, Lemma 1e], it follows that $\xi_T(\cdot)$ is concave. Hence, for all $y > x$, we have $y > \xi_T(y)$. Also, from the monotonicity of $\mathbf{\Psi}(\cdot, S)$ [1, Lemma 1c], $\xi_T(\cdot)$ is monotonic. Then, for $j \in \mathbb{N}$ we have $y > \xi_T^{(j)}(y) \geq \xi_T^{(j+1)}(y)$ where $\xi_T^{(j)}$ denotes the composition of ξ_T j times. Also, for $j \in \mathbb{N}$, we have $0 < \xi_T^{(j)}(0) \leq \xi_T^{(j+1)}(0) < y$. Since the sequence $\xi_T^{(j)}(0)$ is monotonic and bounded, it must converge to a fixed point x_0, i.e.

$$\lim_{j\to\infty} \xi_T^{(j)}(z) = x_0 = \xi_T(x_0). \tag{66}$$

Recall from (7) that $G_T(\mathbf{P}_0) = \left\|\sum_{S\in\mathbb{B}^T} \mathbb{P}(\Gamma_T = S)\mathbf{\Psi}\left(\mathbf{P}_0, S\right)\right\|$. Using the monotonicity of $\mathbf{\Psi}(\cdot, S)$, we have that for all $Z \in \mathbb{N}$

$$\begin{aligned}G_Z(\mathbf{P}_0) \leq& \left\|\sum_{S\in\mathbb{B}^Z} \mathbb{P}(\Gamma_Z = S)\mathbf{\Psi}\left(\|\mathbf{P}_0\|\mathbf{I}, S\right)\right\|\\ \leq&\operatorname{Tr}\left(\sum_{S\in\mathbb{B}^Z} \mathbb{P}(\Gamma_Z = S)\mathbf{\Psi}\left(\|\mathbf{P}_0\|\mathbf{I}, S\right)\right)\\ =& \sum_{Y\in\mathbb{B}^\nu} \mathbb{P}(\Delta_Z = Y)\times \qquad (67)\\ &\times \operatorname{Tr}\left(\sum_{S\in\mathbb{B}^Z} \mathbb{P}(\Gamma_Z = S|\Delta_Z = Y)\mathbf{\Psi}\left(\|\mathbf{P}_0\|\mathbf{I}, S\right)\right)\\ \leq&\xi_Z(\|\mathbf{P}_0\|).\end{aligned} \tag{68}$$

Using (68), (61), and (66) we obtain

$$\begin{aligned}\limsup_{j\to\infty} G_{Tj}(\mathbf{P}_0) &\leq \limsup_{j\to\infty}\ \xi_{Tj}(\|\mathbf{P}_0\|)\\ &\leq \limsup_{j\to\infty}\ \xi_T^{(j)}(\|\mathbf{P}_0\|) = x_0.\end{aligned} \tag{69}$$

Now, from Lemma 20, for each $i = 0, \cdots, T-1$ and $j = 1, 2, \cdots$, we have $G_{Tj+i} \leq x_0|\alpha_1|^{2i} + l_i$, where l_i is a constant independent of $\mathbf{P}_0$. Hence, $\limsup_{t\to\infty} G_t \leq \sup_{0\leq i<T} x_0|\alpha_1|^{2i} + l_i$, and the result follows by making $l = \sup_{0\leq i<T} l_i$. ■

We are now ready to prove Lemma 23.

Proof of Lemma 23: The proof is divided in five steps.

1) In view of Lemma 25, there exists Q such that if $\mathbf{O}_k$ has FCR(Q), for $k = 1, \cdots, K$, then $\mathbf{O}$ has FCR. Recall from (54) that $\mathcal{N}_{k,Q}^t$ is the set of binary sequences $S \in \mathbb{B}^t$ such that $\mathbf{O}_k(S)$ does not have FCR with strength Q. Define

$$\mathcal{G}_k^t \triangleq \begin{cases}\mathcal{N}_{1,Q}^t & , k = 1\\ \mathcal{N}_{k,Q}^t \cap \bigcap_{j=1}^{k-1} \overline{\mathcal{N}}_{j,Q}^t & , k = 2, \cdots, K,\end{cases} \tag{70}$$

i.e., $\mathcal{G}_k^t$ is the set of binary sequences $S \in \mathbb{B}^t$ such that $\mathbf{O}_j(S)$ has FCR(Q) for $1 \leq j \leq k-1$ and $\mathbf{O}_k(S)$ does not have FCR(Q).

Recall that $\mathcal{N}^t$ is the set of sequences $S \in \mathbb{B}^t$ such that $\mathbf{O}(S)$ does not have FCR. From Lemma 25, we have $\mathbb{B}^t = \overline{\mathcal{N}}^t \cup \bigcup_{k=1}^{K} \mathcal{G}_k^t$. That is, for any sequence $S \in \mathbb{B}^t$, either $\mathbf{O}\left(S\right)$

has FCR, or $\mathbf{O}_k(S)$ does not have FCR(Q) for some $1 \leq k \leq K$, or both conditions are satisfied. Recall that $\mathbf{\Psi}(x\mathbf{I}, S)$ is a symmetric, positive-definite matrix with dimension n. Hence $\mathrm{Tr}(\mathbf{\Psi}(x\mathbf{I}, S)) \leq n \|\mathbf{\Psi}(x\mathbf{I}, S)\|$. From (60) and the above, we have

$$\xi_t(x) \leq n \max_{Y\in\mathbb{B}^\nu} \left\| \sum_{S\in\mathbb{B}^t} \mathbb{P}(\Gamma_t = S \,|\Delta_t = Y\,)\, \mathbf{\Psi}(x\mathbf{I}, S) \right\| \quad (71)$$

$$\leq n \left\| \sum_{S\in\overline{\mathcal{N}}^t} \mathbf{\Psi}(x\mathbf{I}, S)\, \mathbb{P}(\Gamma_t = S|\Delta_t = Y^\star) \right\| +$$

$$+ n \sum_{k=1}^K \left\| \sum_{S\in\mathcal{G}_k^t} \mathbf{\Psi}(x\mathbf{I}, S)\, \mathbb{P}(\Gamma_t = S|\Delta_t = Y^\star) \right\|, \quad (72)$$

where $Y^\star$ is the argument that maximizes (71).

2) From Lemma 20, we have that

$$S \in \overline{\mathcal{N}}^t \implies \|\mathbf{\Psi}(x\mathbf{I}, S)\| \leq l_t \quad (73)$$

$$S \in \mathcal{G}_k^t \implies \|\mathbf{\Psi}(x\mathbf{I}, S)\| \leq |\alpha_k|^{2t} x + l_t. \quad (74)$$

Also, from (70), we have that $\mathbb{P}(\mathcal{G}_k) \leq \mathbb{P}(\mathcal{N}_{k,Q})$.

3) From (72) and the above, we have

$$\begin{aligned}
\xi_t(x) \leq & n \sum_{S\in\overline{\mathcal{N}}^t} l_t \mathbb{P}(\Gamma_t = S|\Delta_t = Y^\star) + \\
& + n \sum_{k=1}^K \sum_{S\in\mathcal{G}_k} \left(|\alpha_k|^{2t} x + l_t\right) \mathbb{P}(\Gamma_t = S|\Delta_t = Y^\star) \\
\leq & n l_t \mathbb{P}\left(\overline{\mathcal{N}}^t|\Delta_t = Y^\star\right) + \\
& + n \sum_{k=1}^K \left(|\alpha_k|^{2t} x + l_t\right) \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right) \\
= & n \sum_{k=1}^K \left(|\alpha_k|^2 \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right)^{1/t}\right)^t x + \\
& + n l_t \left(\mathbb{P}\left(\overline{\mathcal{N}}^t|\Delta_t = Y^\star\right) + \sum_{k=1}^K \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right)\right) \\
= & \beta_t x + \gamma_t,
\end{aligned}$$

with $\beta_t = n\sum_{k=1}^K \left(|\alpha_k|^2 \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right)^{1/t}\right)^t$ and

$$\gamma_t = n l_t \left(\mathbb{P}\left(\overline{\mathcal{N}}^t|\Delta_t = Y^\star\right) + \sum_{k=1}^K \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right)\right).$$

4) Let $t > \nu$. Then,

$$\begin{aligned}
& \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right) \\
= & \sum_{S\in\mathcal{N}_{k,Q}^t} \mathbb{P}(\Gamma_t = S|\Delta_t = Y^\star) \\
= & \sum_{S\in\mathcal{N}_{k,Q}^t} \mathbb{P}(\Gamma_t = S) \frac{\mathbb{P}(\Delta_t = Y^\star|\Gamma_t = S)}{\mathbb{P}(\Delta_t = Y^\star)} \\
\leq & \sum_{S\in\mathcal{N}_{k,Q}^t} \mathbb{P}(\Gamma_t = S) \max_{S\in\mathcal{N}_{k,Q}^t} \frac{\mathbb{P}(\Delta_t = Y^\star|\Gamma_t = S)}{\mathbb{P}(\Delta_t = Y^\star)}
\end{aligned}$$

Taking $\delta(\cdot)$ as defined in (28), we get

$$\begin{aligned}
& \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right) \\
= & \sum_{S\in\mathcal{N}_{k,Q}^t} \mathbb{P}(\Gamma_t = S) \max_{S\in\mathcal{N}_{k,Q}^t} \frac{\mathbb{P}(\Delta_t = Y^\star|\delta(\Gamma_t) = \delta(S))}{\mathbb{P}(\Delta_t = Y^\star)} \\
= & \zeta \mathbb{P}\left(\mathcal{N}_{k,Q}^t\right),
\end{aligned}$$

where

$$\zeta = \max_{S\in\mathcal{N}_{k,Q}^t} \frac{\mathbb{P}(\Delta_t = Y^\star|\delta(\Gamma_t) = \delta(S))}{\mathbb{P}(\Delta_t = Y^\star)} > 0.$$

Then,

$$\begin{aligned}
\limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_{k,Q}^t|\Delta_t = Y^\star\right)^{1/t} = & \limsup_{t\to\infty} \zeta^{1/t} \mathbb{P}\left(\mathcal{N}_{k,Q}^t\right)^{1/t} \\
= & \limsup_{t\to\infty} \mathbb{P}\left(\mathcal{N}_{k,Q}^t\right)^{1/t}. \quad (75)
\end{aligned}$$

5) From (55) and (75), there exists $t_0 \in \mathbb{N}$ such that, for all $t > t_0$, $\beta_t < 1$. Then, for all $t > t_0$, there exists $x > 0$ such that $x = \beta_t x + \gamma_t$, and therefore, $x > \xi_t(x)$. Then, the result follows from Lemma 30. ■

The sufficient condition in Theorem 7, i.e., (19) follows from Lemmas 23 and 24.

VI. Conclusion

We have derived a necessary condition and a sufficient condition, having only a trivial gap, for the boundedness of the expected value of the estimation error covariance of a Kalman filter subject to random measurement losses. The results were obtained for a general finite state Markov channel (FSMC) packet loss model and assuming that the system's state matrix is diagonalizable. The existing literature usually adopts either i.i.d. or Gilbert-Elliott packet loss model and assumes non-degenerate systems or special cases of degenerate systems. In these cases, our conditions retrieve the known results for the boundedness of the asymptotic expected error covariance. When the more general FSMC packet loss model and non-degenerate systems are adopted, we extend the known results by providing a closed-form expression to determine the critical parameter that determines the boundedness of the asymptotic expected error covariance. Finally, when degenerate systems and an FSMC packet loss model are considered, we provide a numerical method to assess whether the asymptotic expected error covariance is bounded or not.

A key byproduct of our work is a novel characterization of the received measurement patterns leading to a ROM having FCR. Ensuring a ROM with FCR leads to an estimation error with bounded covariance, in a number of estimation methods based on the MMSE criterion (including Kalman filtering). Our result in Lemma 25 states that if the ROM is decomposed in horizontally concatenated sub-matrices, and each of these sub-matrices has FCR of order Q, for a given integer Q, then the ROM has FCR. This result can find applications in, for example, the problem of distributed sensor scheduling [41]–[43]. This problem can be considered as dual to the problem studied in this paper, since it consists in selecting which sensors should send their measurements at each time instant in order to maximize estimation accuracy and/or minimize energy/bandwidth utilization. Hence, each sub-matrix component of the ROM is generated from the measurements transmitted from each sensor. It then follows from our result that, if the local scheduling strategy is such that each sensor sends enough measurements to ensure that its corresponding sub-matrix component has FCR of order Q, then the estimation error covariance is bounded.

APPENDIX

To simplify the notation we suppose $\mathbf{A} = \mathbf{A}_1$, so we can omit the subscript k.

Before proving Lemma 24 we present some technical lemmas. Recall the notation introduced before Lemma 15. In particular, $N \in \mathbb{N}$ is such that $\mathbf{A}^N = \alpha^N \mathbf{I}$. We split a sequence $S \in \mathbb{B}^t$ into L sub-sequences of length multiple of N with the property that the first $L-1$ sub-sequences U are such that $\mathbf{O}(U)$ has FCR. Formally, for $t \in \mathbb{N}$ and $S \in \mathbb{B}^t$, put $S_0 = S$, and consider the following recursions,

$$\begin{aligned} t_l &\triangleq \min\left\{t : (S_{l-1}[1], \cdots, S_{l-1}[tN]) \in \overline{\mathcal{N}}^{tN}\right\}, \\ S_l &\triangleq (S_{l-1}[t_l N + 1], \cdots, S_{l-1}[|S_{l-1}|]), \end{aligned}$$

where $|S|$ denotes the length of the sequence S. The recurrence is stopped at $l = L$, where L is such that $S_L \in \mathcal{N}^{|S_L|}$. Then, we define the maps η and τ by

$$\eta(S) \triangleq L - 1, \quad \tau(S) \triangleq S_L,$$

i.e., $\eta(S)$ is denotes the maximum number of contiguous sub-blocks of S which are in $\overline{\mathcal{N}}^{t_l N}$, $l = 1, \cdots, L$, and $\tau(S)$ denotes the remaining sub-block of S. We allow $\tau(S)$ to be an empty sequence, denoted by $()$. Recall the sequences $F_i, i = 1, \cdots, I$ and $H_{i,m}, i = 1, \cdots, I$ and $m = 1, \cdots, M_i$, defined in Section III-C. Let $\tilde{M}_1 = 2^\nu$, and for $i = 2, \cdots, I$, $\tilde{M}_i = M_i$. Also, define $\tilde{H}_{i,m}$, $i = 1, \cdots, I$, $m = 1, \cdots, \tilde{M}_i$ such that $\tilde{H}_{i,m} = H_{i,m}$, for $i > 1$, and $\tilde{H}_{1,m}$, $m = 1, \cdots, \tilde{M}_1$ are all sequences in $\mathbb{B}^\nu$. Then, for any $i = 1, \cdots, I$, $m = 1, \cdots, \tilde{M}_i$, and $Q \in \mathbb{N}_0$, we define

$$\begin{aligned} \mathcal{L}^n_{i,m,Q} \triangleq \Big\{S \in \mathbb{B}^{nN} : \psi(\tau(S)) = F_i, \\ \delta(S) = \tilde{H}_{i,m}, \eta(S) = Q\Big\}. \end{aligned} \quad (76)$$

In (76) we have used $\tilde{H}_{1,m}$ (i.e., all sequences in $\mathbb{B}^\nu$) instead of $H_{1,m}$. The reason for this is that, when $\tau(S) = ()$, then $\psi(\tau(S)) = F_1 = (0)^N$. In this case $M_1 = 1$ and the only possibility for $H_{1,m}$ is $H_{1,1} = (0, \cdots, 0)$. However, notice that in (76), when $\tau(S) = ()$, $\delta(S)$ can be any sequence of length ν. Hence, $\tilde{H}_{1,m}$ is used.

Lemma 31. *For any $Q, l \in \mathbb{N}_0$,*

$$\mathbb{P}\left(\mathcal{N}^{lN}_Q\right) \le \mathbf{u}_b' \mathbf{B}^l \mathbf{z}_b, \quad (77)$$

with equality holding when $Q = 0$. In the above, $\mathbf{B} \in \mathbb{R}^{b\times b}$, with $b = (Q+1)d$ and $d = \sum_{i=1}^I \tilde{M}_i$, is given by

$$\mathbf{B} = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{N} & \mathbf{D} & \ddots & \vdots \\ \mathbf{0} & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{N} & \mathbf{D} \end{bmatrix}, \quad (78)$$

with $\mathbf{D} \in \mathbb{R}^{d\times d}$ and $\mathbf{N} \in \mathbb{R}^{d\times d}$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{1,1} & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \mathbf{0} \\ \mathbf{D}_{I,1} & \cdots & \mathbf{D}_{I,I} \end{bmatrix}, \ \mathbf{N} = \begin{bmatrix} \mathbf{N}_1 & \cdots & \mathbf{N}_I \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}. \quad (79)$$

The sub-matrix $\mathbf{D}_{i_1,i_2} \in \mathbb{R}^{\tilde{M}_{i_1}\times \tilde{M}_{i_2}}$ is given by[2]

$$[\mathbf{D}_{i_1,i_2}]_{j_1,j_2} = \mathbb{P}\left(\mathcal{L}^3_{i_1,j_1,1} \middle| \mathcal{L}^2_{i_2,j_2,1}\right). \quad (80)$$

Also, $\mathbf{N}_i \in \mathbb{R}^{\tilde{M}_1\times\tilde{M}_i}$ is given by

$$[\mathbf{N}_i]_{j_1,j_2} = \mathbb{P}\left(\mathcal{L}^3_{1,j_1,2} \middle| \mathcal{L}^2_{i,j_2,1}\right), \quad (81)$$

Finally, $\mathbf{u}_b = [1 \ \cdots \ 1]' \in \mathbb{R}^b$ and $\mathbf{z}_b = [\mathbf{z}' \ 0 \ \cdots \ 0]' \in \mathbb{R}^b$, where $\mathbf{z} = [z_1 \ \cdots \ z_{2^\nu}]'$, with $z_m = \mathbb{P}(\Gamma_\nu = H_{1,m})$, $m = 1, \cdots, 2^\nu$, i.e., $\mathbf{z}$ is the probability distribution of Γ_ν.

Proof: For $l \in \mathbb{N}$ let $\mathbf{m}(l) = \begin{bmatrix}\mathbf{m}_0'(l) & \cdots & \mathbf{m}_Q'(l)\end{bmatrix}'$, with $\mathbf{m}_q(l) = \begin{bmatrix}\mathbf{m}_{q,1}'(l) & \cdots & \mathbf{m}_{q,I}'(l)\end{bmatrix}'$, and

$$\mathbf{m}_{q,i}(l) = \begin{bmatrix} \mathbb{P}\left(\mathcal{L}^l_{i,1,q}\right) & \cdots & \mathbb{P}\left(\mathcal{L}^l_{i,\tilde{M}_i,q}\right) \end{bmatrix}'. \quad (82)$$

Let also $\mathbf{m}(0) = \begin{bmatrix}\mathbf{m}_{0,1}'(0) & \mathbf{0}\cdots\mathbf{0}\end{bmatrix}' \in \mathbb{R}^b$, with

$$\mathbf{m}_{0,1}(0) = \begin{bmatrix} \mathbb{P}\left((g_{-\nu}, \cdots, g_{-1}) = \tilde{H}_{1,1}\right) \\ \vdots \\ \mathbb{P}\left((g_{-\nu}, \cdots, g_{-1}) = \tilde{H}_{1,\tilde{M}_1}\right) \end{bmatrix}. \quad (83)$$

It is easy to see that $\mathbf{m}(l) = \mathbf{B}\mathbf{m}(l-1)$, for all $l \in \mathbb{N}$, and $\{S : \eta(S) \le Q\} = \bigcup_{i,j,q} \mathcal{L}^l_{i,j,q}$. Hence,

$$\mathbb{P}(\eta(S) \le Q) = \mathbf{u}_b'\mathbf{m}(l) = \mathbf{u}_b'\mathbf{B}^l\mathbf{m}(0). \quad (84)$$

From the stationarity of the Markov process we have that $\mathbf{m}(0) = \mathbf{z}_b$. Substituting into (84), we have $\mathbb{P}(\eta(S) \le Q) = \mathbf{u}_b'\mathbf{B}^l\mathbf{z}_b$. Then, (77) follows since, for all $S \in \mathbb{B}^{lN}$, $\eta(S) > Q \implies S \in \overline{\mathcal{N}}^{lN}_Q$. Finally, the equality holds when $Q = 0$, since in this case, $\eta(S) > 0 \iff S \in \overline{\mathcal{N}}^{lN}$. ■

Lemma 32. *Let $\mathbf{u}_b$ and $\mathbf{z}_b$ be as defined in Lemma 31. Then*

$$\limsup_{t\to\infty}(\mathbf{u}_b'\mathbf{B}^t\mathbf{z}_b)^{1/t} = \max_i \rho(\mathbf{D}_{i,i}), \quad (85)$$

where $\rho(\mathbf{X})$ denotes the spectral radius of the matrix $\mathbf{X}$.

Proof: We split the proof in two steps:

1) Let $\beta = \max_i \rho(\mathbf{D}_{i,i})$. Since $\mathbf{B}$ is block triangular, its eigenvalues are those of $\mathbf{D}$. Since $\mathbf{D}$ is also block triangular, its set of eigenvalues is the union of those of each submatrix $\mathbf{D}_{i,i}$, $i = 1, \cdots, I$. Hence,

$$\limsup_{t\to\infty}(\mathbf{u}_b'\mathbf{B}^t\mathbf{z}_b)^{1/t} \le \beta. \quad (86)$$

2) Let $\beta_i = \rho(\mathbf{D}_{i,i})$. Since $\mathbf{N} \succeq \mathbf{0}$, we have that $\mathbf{u}_b'\mathbf{B}^t\mathbf{z}_b \ge \mathbf{u}_d'\mathbf{D}^t\mathbf{z}_d$. Also, since $\mathbf{D}$ is block triangular, we have

$$\mathbf{u}_d'\mathbf{D}^t\mathbf{z}_d \ge \max_{1\le i\le I} \mathbf{u}_{\tilde{M}_i}'\mathbf{D}_{i,1}\mathbf{D}_{i,i}^{t-1}\mathbf{z}_{\tilde{M}_1} = \max_{1\le i\le I} \tilde{\mathbf{u}}_i\mathbf{D}_{i,i}^{t-1}\mathbf{z}_{\tilde{M}_1},$$

[2]Notice that Assumption 2 is required so that the matrices $\mathbf{D}_{i_1,i_2}$ have finite dimension.

where $\tilde{\mathbf{u}}_i = \mathbf{u}'_{\tilde{M}_i}\mathbf{D}_{i,1}$. Hence,

$$\limsup_{t\to\infty}(\mathbf{u}'_b\mathbf{B}^t\mathbf{z}_b)^{1/t} \geq \limsup_{t\to\infty}\left(\max_{1\leq i\leq I}\tilde{\mathbf{u}}'_i\mathbf{D}^t_{i,i}\mathbf{z}_{\tilde{M}_1}\right)^{1/t}. \tag{87}$$

For $i = 1$, we have that $\limsup_{t\to\infty}(\tilde{\mathbf{u}}'_1\mathbf{D}^t_{1,1}\mathbf{z}_{\tilde{M}_1})^{1/t} = \beta_1$. Now, for $i > 1$ and from Assumption 2, we have that $\mathbf{D}_{i,i} \succ \mathbf{0}$ and from Perron's Theorem [44, Theorem 8.2.11], we have that β_i is an eigenvalue of $\mathbf{D}_{i,i}$ with multiplicity one, and its associated left $\mathbf{l}$ and right $\mathbf{r}$ eigenvectors are positive. Then we have

$$\begin{aligned}\lim_{t\to\infty}(\tilde{\mathbf{u}}'_i\mathbf{D}^t_{i,i}\mathbf{z}_{\tilde{M}_1})^{1/t} &= \lim_{t\to\infty}\left(\tilde{\mathbf{u}}'_i\left(\frac{\mathbf{D}_{i,i}}{\beta_i}\right)^t\mathbf{z}_{\tilde{M}_1}\right)^{1/t}\beta_i \\ &\overset{(a)}{=} \lim_{t\to\infty}\left(\frac{\tilde{\mathbf{u}}'_i\mathbf{r}\mathbf{l}'\mathbf{z}_{\tilde{M}_1}}{\mathbf{l}'\mathbf{r}}\right)^{1/t}\beta_i = \beta_i,\end{aligned} \tag{88}$$

where (a) follows from Perron's theorem. Putting (88) into (87) we have $\limsup_{t\to\infty}(\mathbf{u}'_b\mathbf{B}^t\mathbf{z}_b)^{1/t} \geq \beta$, and the result follows from (86). ■

Lemma 33. *For any $Q \in \mathbb{N}_0$,*

$$\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t_Q\right)^{1/t} \leq \max_i \rho\left(\mathbf{D}_{i,i}\right)^{1/N},$$

with equality holding when $Q = 0$.

Proof: Let $k(t) = \max\{k \in \mathbb{N} : kN \leq t\}$. Then,

$$\mathbb{P}\left(\mathcal{N}^t_Q\right)^{1/t} \leq \mathbb{P}\left(\mathcal{N}^{k(t)N}_Q\right)^{1/t} \leq \mathbb{P}\left(\mathcal{N}^{k(t)N}_Q\right)^{1/(k(t)+1)N}.$$

Then,

$$\begin{aligned}&\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t_Q\right)^{1/t} \\ &\leq \limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^{k(t)N}_Q\right)^{1/(k(t)+1)N} \\ &= \limsup_{k\to\infty}\mathbb{P}\left(\mathcal{N}^{kN}_Q\right)^{1/(k+1)N} \\ &= \left(\limsup_{k\to\infty}\mathbb{P}\left(\mathcal{N}^{kN}_Q\right)^{1/k}\left(\mathbb{P}\left(\mathcal{N}^{kN}_Q\right)^{1/k}\right)^{-1/(k+1)}\right)^{1/N}.\end{aligned}$$

Notice that $\limsup_{k\to\infty}\left(\mathbb{P}\left(\mathcal{N}^{kN}_Q\right)^{1/k}\right)^{-1/(k+1)} = 1$. Using Lemmas 31 and 32, we have

$$\begin{aligned}\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t_Q\right)^{1/t} &\leq \left(\limsup_{k\to\infty}\mathbb{P}\left(\mathcal{N}^{kN}_Q\right)^{1/k}\right)^{1/N} \\ &\leq \left(\limsup_{k\to\infty}\left(\mathbf{u}'_b\mathbf{B}^k\mathbf{z}_b\right)^{1/k}\right)^{1/N} \\ &= \max_i\rho\left(\mathbf{D}_{i,i}\right)^{1/N}.\end{aligned} \tag{89}$$

To show the result for $Q = 0$, notice that

$$\begin{aligned}&\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t\right)^{1/t} \\ &\geq \limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^{k(t)N}\right)^{\frac{1}{(k(t)-1)N}} \\ &= \limsup_{k\to\infty}\mathbb{P}\left(\mathcal{N}^{kN}\right)^{\frac{1}{(k-1)N}} \\ &= \left(\limsup_{k\to\infty}\mathbb{P}\left(\mathcal{N}^{kN}\right)^{1/k}\left(\mathbb{P}\left(\mathcal{N}^{kN}\right)^{1/k}\right)^{\frac{1}{k-1}}\right)^{1/N}.\end{aligned} \tag{90}$$

Similarly, we have that $\limsup_{k\to\infty}\left(\mathbb{P}\left(\mathcal{N}^{kN}\right)^{1/k}\right)^{1/(k-1)} = 1$. Hence,

$$\begin{aligned}\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t\right)^{1/t} &\geq \left(\limsup_{k\to\infty}\left(\mathbf{u}'_b\mathbf{B}^k\mathbf{z}_b\right)^{1/k}\right)^{1/N} \\ &= \max_i\rho(\mathbf{D}_{i,i})^{1/N}.\end{aligned} \tag{91}$$

Then, the result follows from (89) and (91). ■

We are now ready to prove Lemma 15 and Lemma 24.

Proof of Lemma 15: It follows from Lemma 33 since $\mathbf{D}_1 = [\mathbf{D}_{1,1}]_{1,1}$ and for $i > 1$, $\mathbf{D}_i = \mathbf{D}_{i,i}$. ■

Proof of Lemma 24: From Lemma 33 and Lemma 15, we have

$$\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t_Q\right)^{1/t} \leq \max_i\rho(\mathbf{D}_i)^{1/N} = \limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t\right)^{1/t}. \tag{92}$$

Now, notice that $\mathcal{N}^t \subseteq \mathcal{N}^t_Q$, with equality holding only for $Q = 0$. Hence, $\mathbb{P}\left(\mathcal{N}^t\right) \leq \mathbb{P}\left(\mathcal{N}^t_Q\right)$ and

$$\limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t\right)^{1/t} \leq \limsup_{t\to\infty}\mathbb{P}\left(\mathcal{N}^t_Q\right)^{1/t}. \tag{93}$$

The result then follows from (92) and (93). ■

REFERENCES

[1] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[2] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: a probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 594–604, Mar. 2010.

[3] E. Rohr, D. Marelli, and M. Fu, "Kalman filtering with intermittent observations: Bounds on the error covariance distribution," in *Proc. IEEE Conf. Decision Control and European Control Conf.*, Orlando, FL, 2011, pp. 2416–2421.

[4] S. Kar, B. Sinopoli, and J. M. F. Moura, "Kalman filtering with intermittent observations: Weak convergence to a stationary distribution," *IEEE Trans. Autom. Control*, vol. 57, no. 2, pp. 405–420, Feb. 2012.

[5] L. Xie, "Stochastic comparison, boundedness, weak convergence, and ergodicity of a random riccati equation with markovian binary switching," *SIAM J. on Control and Optimization*, vol. 50, no. 1, pp. 532–558, 2012.

[6] A. Censi, "Kalman filtering with intermittent observations: Convergence for semi-markov chains and an intrinsic performance measure," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 376–381, Feb. 2011.

[7] S. Kar and J. M. F. Moura, "Moderate deviations of a random riccati equation," *IEEE Trans. Autom. Control*, vol. 57, no. 9, pp. 2250–2265, Sep. 2012.

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

[9] K. You, M. Fu, and L. Xie, "Mean square stability for kalman filtering with markovian packet losses," *Automatica*, vol. 47, no. 12, pp. 2647–2657, 2011.

[10] L. Xie, "Stability of a random riccati equation with markovian binary switching," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1759–1764, Aug. 2008.

[11] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: Tail distribution and critical value," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 677–689, Mar. 2012.

[12] K. Plarre and F. Bullo, "On kalman filtering for detectable systems with intermittent observations," *IEEE Trans. Autom. Control*, vol. 54, no. 2, pp. 386–390, Feb. 2009.

[13] Y. Mo and B. Sinopoli, "A characterization of the critical value for kalman filtering with intermittent observations," in *IEEE Conf. Decision Control*, Cancun, Mexico, Dec. 2008, pp. 2692–2697.

[14] L. Xie and L. Xie, "Peak covariance stability of a random riccati equation arising from kalman filtering with observation losses," *J. of Systems Science and Complexity*, vol. 20, no. 2, pp. 262–272, 2007.

[15] E. Rohr, D. Marelli, and M. Fu, "Statistical properties of the error covariance in a kalman filter with random measurement losses," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Atlanta, USA, 15-17 2010, pp. 5881–5886.

[16] ——, "On the error covariance distribution for kalman filters with packet dropouts," in *Discrete Time Systems*, M. A. Jordán, Ed. InTech, 2010, pp. 71–92.

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

[18] A. F. Dana, V. Gupta, J. P. Hespanha, B. Hassibi, and R. M. Murray, "Estimation over communication networks: Performance bounds and achievability results," in *Proc. Amer. Control Conf.*, New York, NY, 9-13 2007, pp. 3450–3455.

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

[20] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, Jun. 2008.

[21] E. Rohr, D. Marelli, and M. Fu, "Kalman filtering for a class of degenerate systems with intermittent observations," in *Proc. IEEE Conf. Decision Control and European Control Conf.*, Orlando, FL, 2011, pp. 2422–2427.

[22] S. Y. Park and A. Sahai, "Intermittent kalman filtering: Eigenvalue cycles and nonuniform sampling," in *Proc. Amer. Control Conf.*, San Francisco, CA, 06 2011, pp. 3692–3697.

[23] D. Quevedo, A. Ahlen, and K. Johansson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. Autom. Control*, to be published.

[24] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J*, vol. 39, no. 9, pp. 1253–1265, 1960.

[25] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J*, vol. 42, no. 9, pp. 1977–1997, 1963.

[26] P. Sadeghi, R. Kennedy, P. Rapajic, and R. Shams, "Finite-state markov modeling of fading channels: a survey of principles and applications," *IEEE Signal Processing Mag.*, vol. 25, no. 5, pp. 57–80, 2008.

[27] P. Bougerol, "Kalman filtering with random coefficients and contractions," *SIAM Journal on Control and Optimization*, vol. 31, no. 4, pp. 942–959, 1993.

[28] M. J. Weinberger, J. J. Rissanen, and M. Feder, "A universal finite memory source," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 643–652, 1995.

[29] F. Babich, O. E. Kelly, and G. Lombardi, "A context-tree based model for quantized fading," *IEEE Commun. Lett.*, vol. 3, no. 2, pp. 46–48, 1999.

[30] C.-T. Chen, *Linear System Theory and Design*, 3rd ed. Oxford University Press, Inc., 1998.

[31] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall Englewood Cliffs, NJ, 1979, vol. 11.

[32] J. H. Evertse, H. P. Schlickewei, and W. M. Schmidt, "Linear equations in variables which lie in a multiplicative group," *Annals of mathematics*, pp. 807–836, 2002.

[33] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, "On kalman filtering over fading wireless channels with controlled transmission powers," *Automatica*, vol. 48, no. 7, pp. 1306–1316, Jul. 2012.

[34] Z. Kenan and T. M. Lok, "Power control for uplink transmission with mobile users," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 5, pp. 2117–2127, jun 2011.

[35] S. Chakrabarti, E. Kyriakides, and D. G. Eliades, "Placement of synchronized measurements for power system observability," *Power Delivery, IEEE Transactions on*, vol. 24, no. 1, pp. 12–19, 2009.

[36] X. Tai, D. Marelli, E. Rohr, and M. Fu, "Optimal pmu placement for power system state estimation with random communication packet losses," in *Control and Automation (ICCA), 2011 9th IEEE International Conference on*, Santiago, Chile, 19-21 2011, pp. 444–448.

[37] S. Azizi, A. S. Dobakhshari, S. A. Nezam Sarmadi, and A. M. Ranjbar, "Optimal pmu placement by an equivalent linear formulation for exhaustive search," *Smart Grid, IEEE Transactions on*, vol. 3, no. 1, pp. 174–182, 2012.

[38] R. Ambrosino, B. Sinopoli, K. Poolla, and S. Sastry, "Optimal sensor density for remote estimation over wireless sensor networks," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, sept. 2008, pp. 599–606.

[39] E. Garone, B. Sinopoli, and A. Casavola, "Lqg control over lossy tcp-like networks with probabilistic packet acknowledgements," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, dec. 2008, pp. 2686–2691.

[40] J. R. Norris, *Markov Chains*. Cambridge Univ Press, 1997.

[41] S. Arai, Y. Iwatani, and K. Hashimoto, "Fast sensor scheduling for spatially distributed sensors," *IEEE Trans. Autom. Control*, vol. 56, no. 8, pp. 1900–1905, Aug. 2011.

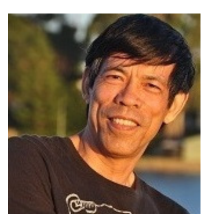
[42] W. Yang and H. Shi, "Sensor selection schemes for consensus based distributed estimation over energy constrained wireless sensor networks," *Neurocomputing*, vol. 87, no. 0, pp. 132–137, Jun. 2012.

[43] K. Cohen and A. Leshem, "A time-varying opportunistic approach to lifetime maximization of wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 58, no. 10, pp. 5307–5319, Oct. 2010.

[44] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge Univ Pr, 1990.

Eduardo Rohr received his B.Sc. and M.Sc. degrees in Electrical Engineering from the Pontifical Catholic University of Rio Grande do Sul, Brazil, in 2006 and 2008 respectively. In 2013 he received a Ph.D. in Electrical Engineering from the University of Newcastle, Australia. Since 2013 Dr. Rohr has been working as a Scientist at the ABB Corporate Research Center in Switzerland.

Damián Marelli received his Bachelors Degree in Electronics Engineering from the Universidad Nacional de Rosario, Argentina in 1995, a Ph.D. degree in Electrical Engineering and a Bachelor (Honous) degree in Mathematics from the University of Newcastle, Australia in 2003. From 2004 to 2005 he held a postdoctoral position at the Laboratoire d'Analyse Topologie et Probabilités, CNRS / Université de Provence, France. Since 2005 he is Research Academic at the School of Electrical Engineering and Computer Science at the University of Newcastle, Australia. In 2007 he received a Marie Curie Postdoctoral Fellowship, hosted at the Faculty of Mathematics, University of Vienna, Austria, and in 2011 he received a Lise Meitner Senior Fellowship, hosted at the Acoustics Research Institute of the Austrian Academy of Sciences. His main research interests include time-frequency analysis, system identification, statistical signal processing and sensor networks.



Minyue Fu (F'83) received the Bachelor's degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of Wisconsin-Madison, in 1983 and 1987, respectively. From 1983 to 1987, he held a Teaching Assistant- ship and a Research Assistantship at the University of Wisconsin-Madison. He worked as a Computer Engineering Consultant at Nicolet Instruments, Inc., Madison, during 1987. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI. He joined the Department of Electrical and Computer Engineering, University of Newcastle, Australia, in 1989, where he is currently Chair Professor in Electrical Engineering and Head of School of Electrical Engineering and Computer Science. In addition, he was a Visiting Associate Professor at the University of Iowa during 1995/1996, and a Senior Fellow/Visiting Professor at Nanyang Technological University, Singapore, 2002. He holds a Qian-ren Professorship at Zhejiang University, China. His main research interests include control systems and signal processing.

Dr. Fu has been an Associate Editor for the IEEE Transactions on Automatic Control, IEEE Transactions on Signal Processing, Automatica, and the Journal of Optimization and Engineering.