Automatica 99 (2019) 390–402

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Stability of Kalman filtering with a random measurement equation: Application to sensor scheduling with intermittent observations☆

Damián Edgardo Marelli [a,b,*], Tianju Sui [c], Eduardo Rath Rohr [d], Minyue Fu [e,a]

[a] School of Automation, Guandong University of Technology, Guangzhou, China
[b] French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Rosario 2000, Argentina
[c] School of Control Science and Engineering, Dalian University of Technology, Dalian, China
[d] ABB Corporate Research Center, Dättwil, 5405, Switzerland
[e] School of Electrical Engineering and Computer Science, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia

ARTICLE INFO

Article history:
Received 13 March 2017
Received in revised form 18 October 2018
Accepted 26 October 2018
Available online 22 November 2018

Keywords:
Kalman filters
Networked control systems
Sensor networks
Estimation theory
Statistical analysis
Stability analysis

ABSTRACT

Studying the stability of the Kalman filter whose measurements are randomly lost has been an active research topic for over a decade. In this paper we extend the existing results to a far more general setting in which the measurement equation, i.e., the measurement matrix and the measurement error covariance, are random. Our result also generalizes existing ones in the sense that it does not require the system matrix to be diagonalizable. For this general setting, we state a necessary and a sufficient condition for stability, and address their numerical computation. An important application of our generalization is a networking setting with multiple sensors which transmit their measurements to the estimator using a sensor scheduling protocol over a lossy network. We demonstrate how our result is used for assessing the stability of a Kalman filter in this multi-sensor setting.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

With the fast development of communications technologies, distributed control and monitoring are becoming increasingly popular. Packet losses resulting from communication links cause that the estimation accuracy of a Kalman filter deteriorates. Motivated by this, the stability condition of a Kalman filter when measurements are intermittently available has attracted a great attention in the recent years. In Sinopoli et al. (2004), the authors established the mathematical foundations for the estimation stability with measurement losses and pointed out that the covariance of the estimation error may not reach a steady state. Inspired by this, several authors have studied different aspects of the problem, using different assumptions on network models and protocols.

When a Kalman filter is subject to randomly intermittent observations (KFIO), its error covariance (EC) matrix becomes also random. Its asymptotic expected value, denoted by AEEC (for asymptotic expected error covariance), is typically used as a performance measure for designing the components of the system, the communication channel, and the estimator. There exists rich literature dedicated to finding the stability conditions of the KFIO (Dana, Gupta, Hespanha, Hassibi, & Murray, 2007; Huang & Dey, 2007; Liu & Goldsmith, 2004; Mo & Sinopoli, 2008, 2012; Park & Sahai, 2011; Plarre & Bullo, 2009; Quevedo, Ahlen, & Johansson, 2013; Rohr, Marelli, & Fu, 2010a, 2010b, 2011; Schenato, 2008; Schenato, Sinopoli, Franceschetti, Poolla, & Sastry, 2007; Sinopoli et al., 2004; Xie, 2008, 2012; Xie & Xie, 2007; You, Fu, & Xie, 2011). Some authors adopt the stability criterion used in Sinopoli et al. (2004), namely, a KFIO is said to be stable if its AEEC is finite (Sinopoli et al., 2004), and unstable otherwise. Other authors adopt the concept of peak error covariance introduced in Huang and Dey (2007). More recently, the equivalence between the two notions of stability has been studied in Xie (2012) and You et al. (2011).

In spite of the fact that there are many papers studying stability conditions of the KFIO, a necessary and sufficient condition for a general system is still not available. Most answers are only partial, in the sense that they depend on a particular structure of the system, or offer only a sufficient condition which is not necessary. In these papers, there are mainly two kinds of packet loss models:

☆ This work was supported by National Natural Science Foundation of China (61803068) and China Postdoctoral Science Foundation (2017M621134). The material in this paper was partially presented at the 13th IEEE International Conference on Control and Automation (IEEE ICCA 2017), Ohrid, Macedonia. This paper was recommended for publication in revised form by Associate Editor Hideaki Ishii under the direction of Editor Christos G. Cassandras.

* Corresponding author at: School of Automation, Guandong University of Technology, Guangzhou, China.
E-mail addresses: damian.marelli@newcastle.edu.au (D.E. Marelli), suitj@mail.dlut.edu.cn (T. Sui), eduardo.rohr@ch.abb.com (E.R. Rohr), minyue.fu@newcastle.edu.au (M. Fu).

https://doi.org/10.1016/j.automatica.2018.11.003
0005-1098/© 2018 Elsevier Ltd. All rights reserved.

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 (Elliott, 1963; Gilbert, 1960), and models the dropouts using a first-order Markov process. A generalization of these two models is the stationary high order Markov model, also known as finite state Markov channel (FSMC) (Sadeghi, Kennedy, Rapajic, & Shams, 2008). It permits modeling more complex channels with memory and fading. In the context of KFIO, this network model has been considered in Censi (2011), although not for the purposes of assessing stability.

For the i.i.d. packet loss model, the authors of Sinopoli et al. (2004) showed that there exists a critical value, such that the AEEC is bounded if the packet arrival probability is strictly greater than this value, and unbounded if the packet arrival probability falls below the critical value. They also provided lower and upper bounds on the critical measurement packet arrival probability. These bounds are only tight for systems whose observation matrix $\mathbf{C}$ is invertible, leading to a necessary a sufficient condition for stability of this kind of systems. This assumption was relaxed in Plarre and Bullo (2009), so that only the part of the matrix $\mathbf{C}$ corresponding to the observable subspace is required to be invertible. The assumption was further relaxed by Mo et al. in Mo and Sinopoli (2008), where they studied the case where the unstable eigenvalues of $\mathbf{A}$ have different magnitudes.

For the Gilbert–Elliott network model, the first work studying the stability of a KFIO is Huang & Dey (2007). In that work, a sufficient condition for stability was derived, considering the peak covariance criterion. For a scalar system, the authors showed that this sufficient condition is also necessary. In Xie (2008), 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 is also necessary. In the case of second order systems, You et al. (2011) gave a necessary and sufficient condition for stability. In Mo and Sinopoli (2012), the authors derived a necessary and sufficient stability condition for a kind of systems which they call *non-degenerate*. This result generalizes most necessary and sufficient stability conditions of systems using the Gilbert–Elliott network model.

For the general FSMC model, to the best of the authors knowledge, the only available work is Rohr, Marelli, & Fu (2014). The authors provided a necessary and a sufficient condition for the stability of the KFIO, which is valid under the assumption that the state transition matrix $\mathbf{A}$ is diagonalizable. This is the most general result known so far, since, as the authors show, every other available result follows as a particular case.

The goal of this paper is to generalize the result in Rohr et al. (2014), so that the resulting necessary and sufficient condition for stability can be applied to distributed Kalman filtering problems under a much more general setting. An important example is the state estimation on discrete-time systems, whose measurements are acquired by multiple sensors, and transmitted to the estimator using sensor scheduling over a lossy network. More precisely, in some wireless networked applications, only limited energy and bandwidth is available for data collection and transmission. Consequently, it is not feasible that all sensors transmit their measurements at every sampling time. Then, a method is required to decide which sensor sends its measurement to the estimator at each time. This decision-making process is referred to as *sensor scheduling*.

Sensor scheduling has been an active research problem for several years. For example, Walsh and Ye (2002) have studied the stability for the close-loop control problem with sensor scheduling. Also, Gupta, Chung, Hassibi, and Murray (2006) proposed a stochastic scheduling strategy for the networked state estimation problem, and derived the optimal probability distribution for selecting sensors at each sample time. Shi, Cheng, and Chen (2011b) considered a system with a single sensor, and studied the problem of whether or not to send its data to a remote estimator, with the aim of saving communications. They studied two scheduling schemes, according to the computational power of the sensor. If the sensor has sufficient power, and under a given communication constraint, they provide an optimal scheduling scheme to minimize the mean squared error (MSE). On the other hand, if the sensor has limited computation power, they gave a scheduling scheme that guarantees that the MSE remains within certain prescribed level. Also, an optimal periodic schedule, satisfying given communication and power constrains, is derived in Shi, Cheng, and Chen (2011a). Sui, You, and Fu (2015) studied the optimization of certain sensor scheduling frameworks for the CMSA/CA protocol. Other relevant works on sensor scheduling include Hovareshti, Gupta, & Baras (2007), Savage & Scala (2009), Shi, Epstein, Sinopoli, & Murray (2007), You & Xie (2013) and You, Xie, & Song (2013), to name a few.

When a sensor scheduling scheme is used together with a randomly lossy data transmission, both scheduling and packet loss directly affect the stability of estimation process. Our stability results are general enough so as to be applicable to assess stability in this setup. We show how this is done using two scheduling schemes, namely, time-based scheduling and random scheduling.

In order to achieve the above, we generalize the result in Rohr et al. (2014) in the following senses:

Model: we drop the diagonalizable assumption on the state transition matrix $\mathbf{A}$, making the result valid for any arbitrary matrix.

Measurements: we generalized the way in which measurements are produced in the following three directions:

(1) Instead of considering a random channel model in which a measurement can be either perfectly transmitted or totally lost, we consider a far more general random measurement equation, in which, at each sample time t, both the measurement matrix $\mathbf{C}_t$ and the measurement noise covariance $\mathbf{R}_t$, are randomly drawn from some known distribution.

(2) While in current works the most general statistical dependence condition for the packet loss process is given by the FSMC model, we assume a more general condition for the pair $(\mathbf{C}_t, \mathbf{R}_t)$. This condition is stated in Eq. (15).

(3) Also, while current works assume that the model for the packet loss process has stationary statistics, we generalize this assumption to the case where $(\mathbf{C}_t, \mathbf{R}_t)$ has cyclostationary statistics. This generalization is essential to the application of our results to a time-based scheduling setting (Section 7.1).

The paper is organized as follows. Section 2 introduces some mathematical background. Section 3 states the research problem. The main result (Theorem 14) is presented in Section 4, together with the general random model of the measurement equation, for which this result is valid. In Section 5, we provide some insight into this general random model. The stability condition stated in our main result is expressed in terms of certain quantity, whose computation is non trivial. In Section 6 we describe how to compute this quantity. In Section 7, we show how to apply our stability results for sensor scheduling. We draw our conclusions in Section 9. For the ease of reading, the formal proof of our main result is presented in Section 8.

2. Preliminaries

Throughout the paper we use the following notation.

Notation 1. *We use $\mathbb{N}$ to denote the set of natural numbers, $\mathbb{Z}$ to denote set of integers, $\mathbb{R}$ for the real numbers and $\mathbb{C}$ for the complex numbers. For a real or complex scalar, vector or matrix, we use * to*

denote its transpose conjugate. For an arbitrary set $\mathcal{S}$, we use $\mathcal{S}^N$ to denote the set of N-tuples with values in $\mathcal{S}$, and $\mathcal{S}^{\mathcal{I}}$ for the set of sequences with the same values indexed by $\mathcal{I}$. For $s \in \mathcal{S}^N$ or $s \in \mathcal{S}^{\mathcal{I}}$, we use $s(i)$ to denote the ith element in s.

We use $\mathcal{N}(\mu, \Sigma)$ to denote a normal distribution with mean μ and covariance matrix Σ and $\mathcal{CN}(\mu, \Sigma)$ to denote a circularly-symmetric complex normal distribution with the same mean and covariance. For an event A, we use $\mathbb{P}(A)$ to denote its probability. For a random variable x, $\mathbb{E}(x)$ denotes its expectation and $\mathbb{P}(x = a)$ denotes the probability of the event $\{x = a\}$. Following a standard convention, in order to simplify the notation, we use x to denote both, the random variable and the value defining the event. We therefore write $\mathbb{P}(x)$ as a shorthand notation for $\mathbb{P}(x = x)$.

We now introduce some required mathematical background on measure theory.

Let $(\mathcal{E}, \mathcal{B})$ be a measurable space. We use $\mathcal{M}(\mathcal{E})$ to denote the Banach space of signed measures on $\mathcal{E}$, which are bounded in the total variation norm (which we denote by $\|\cdot\|$). We also use $\mathcal{L}(\mathcal{E})$ to denote the set of bounded linear operators $\kappa : \mathcal{M}(\mathcal{E}) \to \mathcal{M}(\mathcal{E})$. We use $\sigma(\kappa)$ to denote the spectrum of $\kappa \in \mathcal{L}(\mathcal{E})$ and $\rho(\kappa)$ to denote its spectral radius. We also define $\prod_{n=1}^{N} \kappa_n = \kappa_N \cdots \kappa_1$ and $\kappa^N = \prod_{n=1}^{N} \kappa$.

An important subset of $\mathcal{M}(\mathcal{E})$ is that of probability measures, which we denote by $\mathcal{P}(\mathcal{E})$. We use $\mathcal{K}(\mathcal{E}) \subset \mathcal{L}(\mathcal{E})$ to denote the set of stochastic transition maps $\kappa : \mathcal{P}(\mathcal{E}) \to \mathcal{P}(\mathcal{E})$. A stochastic kernel is a map $\boldsymbol{\kappa} : \mathcal{E} \to \mathcal{P}(\mathcal{E})$ such that, for every $B \in \mathcal{B}$, the map $e \mapsto \boldsymbol{\kappa}(e)(B)$ is measurable. We use $\kappa \in \mathcal{K}(\mathcal{E})$ to denote the stochastic transition map induced by $\boldsymbol{\kappa}$ as follows

$$\kappa\mu(B) = \int_{\mathcal{E}} \kappa(\varepsilon)(B)\mu(d\varepsilon).$$

We finish this section by defining certain elements from the above spaces, which we will use in the rest of the paper. We define $\delta_e \in \mathcal{P}(\mathcal{E})$ by $\delta_e(A) = 1$ if $e \in A$ and 0 otherwise. For $D \in \mathcal{B}$ we use $\chi_D \in \mathcal{L}(\mathcal{E})$ to denote the map assigning each measure to its restriction to D, i.e.,

$$(\chi_D\mu)(B) = \mu(D \cap B), \text{ for all } B \in \mathcal{B}, \mu \in \mathcal{M}(\mathcal{E}).$$

For $\mu \in \mathcal{M}(\mathcal{E})$ and $\kappa \in \mathcal{L}(\mathcal{E})$, we define $\mathcal{U}(\kappa, \mu) \subset \mathcal{M}(\mathcal{E})$ to be the set of accumulation points of the sequence $\kappa^n\mu$, i.e., the set of all $\nu \in \mathcal{M}(\mathcal{E})$ such that, for every $\epsilon > 0$, there exist infinitely many $n \in \mathbb{N}$ such that $\|\kappa^n\mu - \nu\| < \epsilon$. We define

$$\mathcal{U}(\kappa) = \overline{\text{span}}\Big(\bigcup_{\mu \in \mathcal{M}(\mathcal{E})} \mathcal{U}(\kappa, \mu)\Big), \tag{1}$$

where $\overline{\text{span}}(\mathcal{U})$ denotes the closed linear span of the set $\mathcal{U}$. We also use

$$\breve{\kappa} = \kappa|_{\mathcal{U}(\kappa)} : \mathcal{U}(\kappa) \to \mathcal{U}(\kappa) \tag{2}$$

to denote the restriction of κ to $\mathcal{U}(\kappa)$. Finally, for a collection $\mathcal{U} \subset \mathcal{M}(\mathcal{E})$ of measures, we define the collection of sets $\mathcal{F}(\mathcal{U}) \subset \mathcal{B}$, as those which are not null with respect to some measure in $\mathcal{U}$, i.e.,

$$\mathcal{F}(\mathcal{U}) = \{A \in \mathcal{B} : \mu(A) > 0 \text{ for some } \mu \in \mathcal{U}\}. \tag{3}$$

3. Problem formulation

Consider the discrete-time linear system

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{w}_t, \tag{4}$$

$$\mathbf{y}_t = \mathbf{C}_t\mathbf{x}_t + \mathbf{v}_t, \tag{5}$$

where $\mathbf{x}_t \in \mathbb{C}^n$ is the vector of states, $\mathbf{y}_t \in \mathbb{C}^p$ is the vector of measurements, $\mathbf{w}_t \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$, with $\mathbf{Q} \geq 0$, is the process noise, $\mathbf{v}_t \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_t)$ with $\mathbf{R}_t \geq 0$, is the measurement noise, $\mathbf{A} \in \mathbb{C}^{n\times n}$ is the state matrix and $\mathbf{C}_t \in \mathbb{C}^{p\times n}$ is the measurement matrix at time t. It is assumed, without loss of generality, that $\mathbf{A}$ is in Jordan normal form. The initial state is $\mathbf{x}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{P}_0)$, with $\mathbf{P}_0 \geq 0$. Also, the set of random vectors $\{\mathbf{x}_0, \mathbf{w}_t, \mathbf{v}_t : t \geq 0\}$ is jointly statistically independent. At time t, the pair $\gamma_t = (\mathbf{C}_t, \mathbf{R}_t)$ is randomly drawn from the finite set $\mathcal{A} = \mathcal{C} \times \mathcal{R}$, where $\mathcal{C} = \left\{\mathbf{C}^{(1)}, \ldots, \mathbf{C}^{(D)}\right\}$ and $\mathcal{R} = \left\{\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(E)}\right\}$. For $T \in \mathbb{N}$, let $\Gamma_{t,T} = (\gamma_t, \ldots, \gamma_{t+T-1}) \in \mathcal{A}^T$ denote the random sequence of measurement matrices and noise covariances from time t up to time $t + T - 1$.

Our next step is to introduce the model describing the statistics of γ_t. We assume that γ_t is generated by a hidden Markov model whose state is an element of $\mathcal{E}$. More specifically, let $h : \mathcal{E} \to \mathcal{A}$ be a measurable function, $\mu_0 \in \mathcal{P}(\mathcal{E})$ and $\boldsymbol{\kappa}_t : \mathcal{E} \to \mathcal{P}(\mathcal{E})$, $t \in \mathbb{Z}$, be a sequence of stochastic kernels. The sequence γ_t is generated as follows: $\varrho_0 \sim \mu_0$, and, for each $t > 0$,

$$\varrho_t \sim \boldsymbol{\kappa}_t(\varrho_{t-1}), \tag{6}$$

$$\gamma_t = h(\varrho_t), \tag{7}$$

where $\varrho \sim \mu$ denotes that ϱ is independently drawn from the probability distribution μ. We assume that, for each $s \geq 0$, the $\{\mathbf{x}_0, \mathbf{w}_t, \mathbf{v}_t, \varrho_s : t \geq 0\}$ is jointly statistically independent.

Remark 2. We assume the above model for γ_t without loss of generality, as it is equivalent to the general model characterized by specifying $\mathbb{P}(\gamma_t|\gamma_s, s < t)$, for all $t \in \mathbb{Z}$ and all possible values of γ_s, $s \leq t$. To see this, notice that the latter can be written in the form (6)–(7) by taking $\mathcal{E} = \mathcal{A}^{\mathbb{N}}$, $\varrho_t = (\gamma_s : s \leq t)$, and defining the σ-algebra $\mathcal{B}$ to be the one generated by the sets $C_{t,A}$, $t \in \mathbb{N}$, $A \subset \mathcal{A}$, where

$$C_{t,A} = \left\{\Gamma \in \mathcal{A}^{\mathbb{N}} : \Gamma(t) \in A\right\}.$$

A Kalman filter is used to obtain an estimate $\hat{\mathbf{x}}_{t|t-1}$ of the state $\mathbf{x}_t$ given the knowledge of $\mathbf{y}_0, \ldots, \mathbf{y}_{t-1}$ and $\Gamma_{0,t}$. The update equation of the expected covariance (EC) $\mathbf{P}_t = \mathbb{E}\left(\tilde{\mathbf{x}}_t\tilde{\mathbf{x}}_t^*\right)$, with $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}$, is

$$\mathbf{P}_{t+1} = \psi_{\gamma_t}(\mathbf{P}_t), \tag{8}$$

with

$$\psi_{\gamma_t}(\mathbf{P}_t) = \mathbf{A}\mathbf{P}_t\mathbf{A}^* + \mathbf{Q} - \mathbf{A}\mathbf{P}_t\mathbf{C}_t^*\left(\mathbf{C}_t\mathbf{P}_t\mathbf{C}_t^* + \mathbf{R}_t\right)^{-1}\mathbf{C}_t\mathbf{P}_t\mathbf{A}^*.$$

In this work we derive a necessary condition and a sufficient condition, with a trivial gap between them (Remark 15 explains what this means), for the stability of the Kalman filter with a random measurement equation. This is done by studying the asymptotic norm of the expected error covariance (ANEEC). In order to define the ANEEC, we introduce the following notation

$$\boldsymbol{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right) = \psi_{\gamma_{t+T-1}} \cdots \psi_{\gamma_{t+1}}\psi_{\gamma_t}(\mathbf{P}_t),$$

i.e., $\boldsymbol{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)$ denotes the covariance matrix resulting at time $t+T$, after starting with covariance $\mathbf{P}_t$ at time t, and then applying the sequence of random measurement equations defined by $\Gamma_{t,T}$. This matrix depends on the random sequence $\Gamma_{t,T}$ and the initial covariance $\mathbf{P}_t$. In order to work with a quantity independent of these values, in defining the ANEEC, we take expectation with respect to $\Gamma_{t,T}$ and the supremum with respect to $\mathbf{P}_t$. This leads to the following definition.

Definition 3. The ANEEC is defined as

$$G = \sup_{\substack{t\in\mathbb{Z} \\ \mathbf{P}_t \geq 0}} \limsup_{T\to\infty} \left\|\mathbb{E}\left(\boldsymbol{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\|.$$

Remark 4. In (5) we assume that the measurements $\mathbf{y}_t \in \mathbb{R}^p$ have time-invariant dimension p. This assumption is done to simplify the presentation, and without loss of generality. This is because the case $\mathbf{y}_t \in \mathbb{R}^{p_t}$ with time-varying dimension p_t can be handled by defining p as the maximum number of rows among the matrices $\mathbf{C}^{(d)}$, $d = 1, \ldots, D$, and zero padding the matrices $\mathbf{C}^{(d)}$ and $\mathbf{R}^{(d)}$ so that all of them have p rows.

4. Main result

Our main result is stated in terms of certain partition of the system (4)–(5) into subsystems which we call finite multiplicative order (FMO) blocks. This partition is introduced next.

Definition 5. A set of complex numbers $x_i \in \mathbb{C}$, $i = 1, \ldots, I$, is said to have a common finite multiplicative order $N \in \mathbb{N}$ up to a constant $\alpha \in \mathbb{C}$, if $x_i^N = \alpha^N$, for all $i = 1, \ldots, I$. If there do not exist N and α satisfying the above, the set is said not to have common finite multiplicative order.

Example 6. The set of numbers $\{2, 2i, -2, -2i\}$ has common finite multiplicative order 4 up to 2. The set $\left\{1, e^{\sqrt{2}i}\right\}$ does not have common finite multiplicative order.

Notation 7. *Consider the following partition of* $\mathbf{A}$,

$$\mathbf{A} = \operatorname{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_K), \tag{9}$$

where the sub-matrices $\mathbf{A}_k$ *are chosen such that, for any* k, *the diagonal entries of* $\mathbf{A}_k$ *have a common finite multiplicative order* N_k *up to* α_k, *and for any* k *and* l *with* $k \neq l$, *the diagonal entries of the matrix* $\operatorname{diag}(\mathbf{A}_k, \mathbf{A}_l)$ *do not have common finite multiplicative order.*

Let $\bar{J} \in \mathbb{N}$ *be the largest size among the Jordan blocks of* $\mathbf{A}$, *and* $\bar{J}_k \in \mathbb{N}$ *be the largest size among the Jordan blocks of* $\mathbf{A}_k$. *For convenience, we assume that the sub-matrices* $\mathbf{A}_k$ *are ordered such that* $|\alpha_1| \geq |\alpha_2| \geq \cdots \geq |\alpha_K|$. *Also, when* $|\alpha_k| = |\alpha_{k+1}|$, *then* $\bar{J}_k \geq \bar{J}_{k+1}$.

For each $d = 1, \ldots, D$, *consider the partition*

$$\mathbf{C}^{(d)} = \left[\mathbf{C}_1^{(d)}, \ldots, \mathbf{C}_K^{(d)}\right],$$

such that, for each k, $\mathbf{C}_k^{(d)}$ *has the same number of columns as* $\mathbf{A}_k$. *Let* $\mathcal{C}_k = \left\{\mathbf{C}_k^{(d)} : d = 1, \ldots, D\right\}$.

Definition 8. In the above partition, each pair $(\mathbf{A}_k, \mathcal{C}_k)$ is called an FMO block of the system (4)–(5).

Remark 9. Notice that if $(\mathbf{A}_k, \mathcal{C}_k)$ is an FMO block, then each sub-matrix $\mathbf{A}_k$ can be written as

$$\mathbf{A}_k = \alpha_k \tilde{\mathbf{A}}_k, \tag{10}$$

$$\tilde{\mathbf{A}}_k = \operatorname{diag}\left\{\exp(i2\pi\theta_{k,1}), \ldots, \exp(i2\pi\theta_{k,K_k})\right\} + \mathbf{U}_k, \tag{11}$$

where $\mathbf{U}_k$ is strictly upper triangular, i.e., its non-zero entries lie above its main diagonal. Also, $\alpha_k \in \mathbb{C}$ and $\theta_{k,j} \in \mathbb{Q}$, for $j = 1, \ldots, K_k$. 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 all $m \in \mathbb{N}$.

In stating our main result, we use certain observability matrix $\mathbf{O}_k$ associated to each FMO block k of the system. Our next step is to introduce this matrix. The measurements $\mathbf{z}_{t,T}^* = \left[\mathbf{y}_t^*, \ldots, \mathbf{y}_{t+T-1}^*\right]$ available from time t up to $T-1$ can be written as

$$\mathbf{z}_{t,T} = \mathbf{O}\left(\Gamma_{t,T}\right)\mathbf{x}_t + \mathbf{f}_{t,T}\left(\Gamma_{t,T}\right), \tag{12}$$

where the observability matrix $\mathbf{O}\left(\Gamma_{t,T}\right)$ is given by

$$\mathbf{O}\left(\Gamma_{t,T}\right) = \left[\mathbf{O}_1\left(\Gamma_{t,T}\right) \quad \mathbf{O}_2\left(\Gamma_{t,T}\right) \quad \ldots \quad \mathbf{O}_K\left(\Gamma_{t,T}\right)\right], \tag{13}$$

$$\mathbf{f}_t\left(\Gamma_{t,T}\right) = \begin{bmatrix} \mathbf{v}_t \\ \mathbf{C}_{t+1}\mathbf{w}_t + \mathbf{v}_{t+1} \\ \vdots \\ \mathbf{C}_{T+t-1}\displaystyle\sum_{j=t}^{T+t-2}\mathbf{A}^{T+t-2-j}\mathbf{w}_j + \mathbf{v}_{t+T-1} \end{bmatrix},$$

with

$$\mathbf{O}_k\left(\Gamma_{t,T}\right) = \begin{bmatrix} \mathbf{C}_{t,k} \\ \mathbf{C}_{t+1,k}\mathbf{A}_k \\ \vdots \\ \mathbf{C}_{T+t-1,k}\mathbf{A}_k^{T-1} \end{bmatrix}, \quad \mathbf{C}_t = \left[\mathbf{C}_{t,1}, \ldots, \mathbf{C}_{t,K}\right],$$

such that, for each k, $\mathbf{C}_{t,k}$ have the same number of columns as $\mathbf{A}_k$.

Our main result is stated in terms of the probability that each matrix $\mathbf{O}_k\left(\Gamma_{t,T}\right)$ does not have full-column rank (FCR). The following definition identifies the event associated to sequences leading to this property.

Definition 10. For $k = 1, \ldots, K$, let

$$\mathcal{N}_k^{t,T} \triangleq \{\Gamma_{t,T} : \mathbf{O}_k\left(\Gamma_{t,T}\right) \text{ does not have FCR}\}. \tag{14}$$

We now state our main result. This requires Assumptions 12 and 13. These conditions are rather general. For this reason, their statement is somewhat technical. In Section 5 we give interpretations of these assumptions, as well as more practical conditions guaranteeing them. However, these assumptions also hold under conditions more general than those given in Section 5. An example of this appears in the proof of Corollary 26. This shows the value of the generality of Assumptions 12 and 13.

Definition 11. We say that the sequence γ_t, $t \in \mathbb{Z}$, is cyclostationary with period $\tau \in \mathbb{N}$, if, in (6)–(7), we have

$$\kappa_t = \kappa_{t+\tau}, \text{ for all } t \in \mathbb{N},$$

$$\mu_0 = \prod_{t=1}^{\tau}\kappa_t, \mu_0.$$

We say that it is stationary if it is cyclostationary with $\tau = 1$.

Assumption 12. The sequence γ_t, $t \in \mathbb{Z}$, in (6)–(7), is cyclostationary with period $\tau \in \mathbb{N}$ and

$$\zeta \triangleq \sup_{\substack{T\in\mathbb{N} \\ 0\leq t<\tau}} \sup_{\substack{\mathbb{P}(\varrho_{t-1})\neq 0 \\ \mathbb{P}(\Gamma_{t,T})\neq 0}} \frac{\mathbb{P}\left(\Gamma_{t,T}|\varrho_{t-1}\right)}{\mathbb{P}\left(\Gamma_{t,T}\right)} < \infty. \tag{15}$$

Assumption 13. For any $0 \leq t < \tau$, any multiple M of τ and any finite collection $\mathcal{D} = (D_m \in \mathcal{B} : m = 1, \ldots, M)$, let

$$\eta_t = \prod_{m=1}^{M}\chi_{D_m}\kappa_{t+m}.$$

Then,

$$\rho\left(\eta_t\right) = \rho\left(\check{\eta}_t\right). \tag{16}$$

Also, for any non-zero non-negative $\mu \in \mathcal{U}\left(\eta_t\right)$ and $A \in \mathcal{F}\left(\mathcal{U}\left(\eta_t\right)\right)$, there exists N such that

$$\check{\eta}_t^n \mu\left(A\right) > 0, \text{ for all } n \geq N. \tag{17}$$

We now state our main result. Its proof is deferred to Section 8.

Theorem 14. *Consider the system (4)–(5) satisfying Assumptions 12 and 13. For $k \in \{1, \ldots, K\}$, let*

$$\Phi_k = \max_{0 \le t < \tau} \limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}_k^{t,T}\right)^{1/T},$$

with $\mathcal{N}_k^{t,T}$ defined by (14). If

$$|\alpha_k|^2 \Phi_k < 1, \text{ for all } k \in \{1, \ldots, K\}, \tag{18}$$

then $G < \infty$, and if

$$|\alpha_k|^2 \Phi_k > 1, \text{ for some } k \in \{1, \ldots, K\}, \tag{19}$$

then $G = \infty$.

Remark 15. Notice that Theorem 14 is inconclusive in the case when $|\alpha_k|^2 \Phi_k = 1$. Trivial gaps of this kind are common in the literature (Mo & Sinopoli, 2012; Sinopoli et al., 2004).

5. About Assumptions 12 and 13

In this section we give an interpretation of the technical condition stated in Assumptions 12 and 13. We also show that these assumptions hold under certain conditions which are easier to interpret. This result is given in Proposition 18 stated below.

Definition 16. A random process γ_t is Markov of order $L \in \mathbb{N}$ if, for all $m \ge 1$,

$$\mathbb{P}\left(\gamma_t | \gamma_{t-L-m}, \ldots, \gamma_{t-1}\right) = \mathbb{P}\left(\gamma_t | \gamma_{t-L}, \ldots, \gamma_{t-1}\right).$$

Furthermore, it is called proper if all the above probabilities are strictly bigger than zero. Finally, the process is independent if $L = 0$.

Definition 17. A random process γ_t is called Gaussian hidden Markov if it is generated by a hidden Markov model like (6)–(7), but with (6) replaced by

$$\varrho_t = \mathbf{K}\varrho_{t-1} + \varepsilon_t,$$

where $\mathbf{K}$ is a stable matrix (i.e., $\rho(\mathbf{K}) < 1$) and $\varepsilon_t \sim \mathcal{N}(0, \Sigma)$, with $\{\mathbf{x}_0, \mathbf{w}_t, \mathbf{v}_t, \varepsilon_t : t \ge 0\}$ being a jointly independent set of random vectors.

Proposition 18. *Suppose that γ_t is cyclostationary with period τ and is either finite-order proper Markov, or Gaussian hidden Markov. Then Assumptions 12 and 13 hold.*

The reminder of this section is devoted to show Proposition 18.

If γ_t is an independent sequence, we could simply take $\mathcal{E} = \mathcal{A}$, $\varrho_t = \gamma_t$ and h to be the identity map. Then, (15) would hold trivially. Hence, (15) can be interpreted as a generalization of the independence property. More generally, the following two lemmas provide conditions under which (15) holds without the independence property.

Lemma 19. *If γ_t is cyclostationary with period $\tau \in \mathbb{N}$ and finite-order Markov, then (15) holds.*

Proof. Since γ_t is finite order Markov, we can take $\varrho_t = \Gamma_{t-L+1,L}$. We then have

$$\begin{aligned}
\zeta &= \sup_{\substack{T \in \mathbb{N} \\ 0 \le t < \tau}} \sup_{\substack{\mathbb{P}(\varrho_{t-1}) \neq 0 \\ \mathbb{P}(\Gamma_{t,T}) \neq 0}} \frac{\mathbb{P}\left(\Gamma_{t,T} | \varrho_{t-1}\right)}{\mathbb{P}\left(\Gamma_{t,T}\right)} \\
&= \sup_{\substack{T \in \mathbb{N} \\ 0 \le t < \tau}} \sup_{\substack{\mathbb{P}(\varrho_{t-1}) \neq 0 \\ \mathbb{P}(\Gamma_{t,T}) \neq 0}} \frac{\mathbb{P}\left(\varrho_{t-1} | \Gamma_{t,T}\right)}{\mathbb{P}\left(\varrho_{t-1}\right)} \\
&\le \sup_{0 \le t < \tau} \sup_{\mathbb{P}(\varrho_{t-1}) \neq 0} \frac{1}{\mathbb{P}\left(\varrho_{t-1}\right)} \\
&< \infty,
\end{aligned}$$

where the last inequality follows since the supremum operations are taken over finite sets.

Lemma 20. *If γ_t is cyclostationary with period τ and Gaussian hidden Markov, then (15) holds.*

Proof. Let $\mathcal{E} = \mathbb{R}^d$ and $\varrho_t \in \mathcal{E}$. Then there exists a partition $\mathcal{S} = \left\{\mathcal{S}^{(d,e)} : d = 1, \ldots, D, e = 1, \ldots, E\right\}$ of $\mathbb{R}^d$ (i.e., $\mathcal{S}^{(d,e)} \cap \mathcal{S}^{(d',e')} = \emptyset$, whenever $(d, e) \neq (d', e')$, and $\bigcup_{(d,e)=(1,1)}^{(D,E)} \mathcal{S}^{(d,e)} = \mathbb{R}^d$), such that

$$h(\varrho) = \left(\mathbf{C}^{(d)}, \mathbf{E}^{(e)}\right), \text{ for all } \varrho \in \mathcal{S}^{(d,e)}.$$

Let $\mathcal{S}_t = h^{-1}(\gamma_t)$, i.e., the unique element from the partition $\mathcal{S}$ such that $h(\varrho) = \gamma_t$, for all $\varrho \in \mathcal{S}_t$, and let $\mathcal{S}_{t,T} = \mathcal{S}_t \times \cdots \times \mathcal{S}_{t+T-1}$. We have that

$$\mathbb{P}\left(\varrho_{t-1} | \Gamma_{t,T}\right) = \mathbb{P}\left(\varrho_{t-1} | \varrho_{t,T} \in \mathcal{S}_{t,T}\right). \tag{20}$$

The right-hand side of (20) can be considered as the probability of ϱ_{t-1} conditioned on the future output of a stationary quantizer. Since the process ϱ_t is Gaussian, it is easy but tedious to show that there exists a constant $c > 0$ such that

$$\frac{\mathbb{P}\left(\varrho_{t-1} | \varrho_{t,T} \in \mathcal{S}_{t,T}\right)}{\mathbb{P}\left(\varrho_{t-1}\right)} < c, \text{ for all } t, T, \varrho_{t-1}, \text{ and } \mathcal{S}_{t,T}.$$

We then have

$$\begin{aligned}
\zeta &= \sup_{\substack{T \in \mathbb{N} \\ 0 \le t < \tau}} \sup_{\substack{\mathbb{P}(\varrho_{t-1}) \neq 0 \\ \mathbb{P}(\Gamma_{t,T}) \neq 0}} \frac{\mathbb{P}\left(\Gamma_{t,T} | \varrho_{t-1}\right)}{\mathbb{P}\left(\Gamma_{t,T}\right)} \\
&= \sup_{\substack{T \in \mathbb{N} \\ 0 \le t < \tau}} \sup_{\substack{\mathbb{P}(\varrho_{t-1}) \neq 0 \\ \mathbb{P}(\Gamma_{t,T}) \neq 0}} \frac{\mathbb{P}\left(\varrho_{t-1} | \Gamma_{t,T}\right)}{\mathbb{P}\left(\varrho_{t-1}\right)} \\
&= \sup_{\substack{T \in \mathbb{N} \\ 0 \le t < \tau}} \sup_{\mathcal{S}_{t,T}, \varrho_{t-1}} \frac{\mathbb{P}\left(\varrho_{t-1} | \varrho_{t,T} \in \mathcal{S}_{t,T}\right)}{\mathbb{P}\left(\varrho_{t-1}\right)} \\
&\le c < \infty.
\end{aligned}$$

We now turn our attention to condition (17). In this condition, the output of the probability transition mapping κ_t is restricted to a set D_m (via the restriction operator χ_{D_m}), which is taken from a finite family $\mathcal{D}$ of sets. The resulting operators are then composed by sequentially taking all sets within $\mathcal{D}$. This yields the map η_t. The set $\mathcal{U}(\eta_t)$ of measures is invariant under η_t. Condition (16) requires that the spectrum of η_t equals that of its restriction to its invariant subspace $\mathcal{U}(\eta_t)$. Also, $\mathcal{F}(\mathcal{U}(\eta_t))$ contains all sets which are non-null for some measure in $\mathcal{U}(\eta_t)$. Condition (17) requires that the measure of all sets in $\mathcal{F}(\mathcal{U}(\eta_t))$ becomes eventually and persistently strictly positive, when, starting from any measure in $\mathcal{U}(\eta_t)$, we sequentially apply the map η_t.

The above condition seems in principle difficult to verify. However, the two lemmas below show that it holds under conditions similar to those in Lemmas 19 and 20.

Lemma 21. *If γ_t is finite-order proper Markov, then Assumption 13 holds.*

Proof. Recall the definition of η_t given in Assumption 13. Let L be the Markov order of γ_t. Let $\mathcal{E} = \mathcal{A}^L$ and $\varrho_t = \Gamma_{t-L+1,L}$. It is easy to see that $\mathcal{F}(\mathcal{U}(\eta_t)) \subseteq \mathcal{G}$ where

$$\mathcal{G} \triangleq \left\{\varrho \in \mathcal{E} : \varrho(l) \subset D_{M-\mathrm{mod}(L-l,M)}, \lambda(A) > 0\right\}.$$

Now, for any μ and $n > L/M$, the measure $\eta_t^n\mu$ is strictly positive on any $A \in \mathcal{G}$. Hence, (17) holds. Also, $\eta_t^n\mu(\varrho) = 0$, for all $\varrho \notin \mathcal{G}$. Hence, (16) also holds and the result follows.

Lemma 22. *If γ_t is Gaussian hidden Markov, then Assumption 13 holds.*

Proof. Recall the definition of η_t given in Assumption 13. It is easy to see that

$$\mathcal{F}(\mathcal{U}(\eta_t)) = \{A \in \mathcal{B} : A \subset D_M, \lambda(A) > 0\}.$$

Also, for any μ, $\eta_t^n\mu$ has density g_n with respect to the Lebesgue's measure λ, and this density is λ-almost everywhere strictly positive on D_M. Hence, for all $A \in \mathcal{F}(\mathcal{U}(\eta_t))$,

$$\eta_t^n\mu(A) = \int_A g_n d\lambda > 0, \text{ for all } n \geq 1,$$

and (17) holds. Also, $\eta_t\mu(A) = 0$ for all $A \in \mathcal{B}$ satisfying $\lambda(B \cap D_M) = 0$. So (16) holds and the result follows.

Proof of Proposition 18. It follows by combining Lemmas 19–22.

6. Computing Φ_k

Our result in Theorem 14 is stated in terms of the quantities Φ_k, $k = 1, \ldots, K$. We introduce below results on how to compute this quantity. Due to space limitations, their proofs appear in the full version (Marelli, Sui, Rohr, & Fu, 2018, Appendix A). Since in our study the choice of $k = 1, \ldots, K$ is fixed, to remove k from the notation, we consider a generic FMO block $(\mathbf{A}, \mathcal{C})$. We start by introducing some necessary notation.

Notation 23. *Let $N \in \mathbb{N}$ be the smallest positive integer such that $\mathbf{A}^N = \alpha^N\mathbf{I}$. Let $\mathbb{K} = \left\{\ker(\mathbf{O}(\Gamma)) : \Gamma \in \mathcal{A}^N\right\} \cup \{\mathbb{C}^n, \emptyset\}$ be the set of all possible kernels of $\mathbf{O}(\Gamma)$, for sequences Γ of length N, including, the whole space $\mathbb{C}^n$ and the empty set $\emptyset$. Notice that $\mathbb{K}$ includes all possible kernels of $\mathbf{O}(\Gamma)$ for sequences Γ of length nN, for any $n \in \mathbb{N}$. For any $n \in \mathbb{N}$, define the map $\psi : \mathcal{A}^{nN} \to \mathbb{K}$ by*

$$\psi(\Gamma) = \ker(\mathbf{O}(\Gamma)).$$

Let $\mathcal{I} = \{0, \ldots, I\}$ and $\mathcal{K}_i$, $i \in \mathcal{I}$ denote all the elements in $\mathbb{K}$. The elements $\mathcal{K}_i$ are numerated such that, if $\mathcal{K}_i \subset \mathcal{K}_j$, then $i > j$ (notice that, in particular, $\mathcal{K}_0 = \mathbb{C}^n$ and $\mathcal{K}_I = \emptyset$). Let $M \in \mathbb{N}$ be any common multiple of N and τ. For each $t \in \mathbb{N}_0$, $i, j \in \mathcal{I}$, $e \in \mathcal{E}$, $A \in \mathcal{B}$, and $\pi \in \mathcal{L}(\mathcal{E})$, let $T_t : \mathcal{I} \times \mathcal{I} \times \mathcal{E} \times \mathcal{B} \to [0, 1]$ be defined by

$$T_t(i, j, e, A) = \mathbb{P}\left(\varrho_{t+M} \in A, \psi\left(\Gamma_{t,M}\right) \cap \mathcal{K}_j = \mathcal{K}_i | \varrho_t = e\right),$$

and $\varsigma_t : \mathcal{I} \times \mathcal{I} \to \mathcal{L}(\mathcal{E})$ by

$$\varsigma_t(i, j)\pi(A) = \int T_t(i, j, e, A)\,\pi(de).$$

The next result provides a method for evaluating Φ.

Proposition 24. *Let $(\mathbf{A}, \mathcal{C})$ be an FMO block. If Assumptions 12 and 13 hold, then*

$$\Phi = \max_{0 \leq t < \tau} \max_{0 \leq i < I} \rho(\breve{\varsigma}_t(i, i))^{1/M}. \tag{21}$$

The above result requires computing the spectral radius of the map $\breve{\varsigma}_t(i, i)$. If $\mathcal{E}$ is a discrete finite space, then $\breve{\varsigma}_t(i, i)$ becomes a matrix and $\rho(\breve{\varsigma}_t(i, i))$ can be easily computed. Otherwise, the following result can be used.

Proposition 25. *For every non-zero non-negative $\mu \in \mathcal{U}(\varsigma_t(i, i))$,*

$$\rho(\breve{\varsigma}_t(i, i)) = \lim_{n \to \infty} \left\|\breve{\varsigma}_t^n(i, i)\mu\right\|^{1/n}.$$

The following corollary states how the expression (21) simplifies in the particular case when there exists a single measurement matrix $\mathbf{C}^{(\alpha)}$ producing measurements which would never make the observability matrix have FCR (e.g., when measurements are lost), but any single measurement produced by any other matrix $\mathbf{C}^{(d)}$, $d \neq \alpha$, would.

Corollary 26. *If Assumption 12 holds, there exists $1 \leq \alpha \leq D$ such that pair $\left(\mathbf{A}, \mathbf{C}^{(\alpha)}\right)$ is non-observable, for each $d \neq \alpha$, $\mathbf{C}^{(d)}$ has FCR, and there exists $1 \leq \beta \leq E$ such that $\mathbb{P}\left(\mathbf{R}_t | \mathbf{C}_t = \mathbf{C}^{(e)}\right) = \delta_{\mathbf{R}^{(\beta)}}$ (i.e., there is only one possible value of $\mathbf{R}_t$ compatible with $\mathbf{C}_t = \mathbf{C}^{(\alpha)}$), then*

$$\Phi = \prod_{t=0}^{\tau-1} \mathbb{P}\left(\mathbf{C}_t = \mathbf{C}^{(\alpha)} | \mathbf{C}_s = \mathbf{C}^{(\alpha)}, s < t\right)^{1/\tau}. \tag{22}$$

7. Application: Sensor scheduling with packet loss

We have a linear system whose dynamics is given by

$$\mathbf{p}_{t+1} = \mathbf{F}\mathbf{p}_t + \mathbf{n}_t, \tag{23}$$

with $\mathbf{n}_t^{(s)} \sim \mathcal{N}(0, \mathbf{N})$. There are S sensors. For each $s = 1, \ldots, S$, sensor s measures

$$\mathbf{u}_t^{(s)} = \mathbf{H}^{(s)}\mathbf{p}_t + \mathbf{e}_t^{(s)}, \tag{24}$$

with $\mathbf{e}_t^{(s)} \sim \mathcal{N}\left(0, \mathbf{E}^{(s)}\right)$. We assume that measurements from only $R < S$ sensors can be transmitted at each time instant t. Then, at each time we transmit

$$\mathbf{r}_t = (\mathbf{M}_t \otimes \mathbf{I})\,\mathbf{u}_t,$$

where $\mathbf{u}_t^T = \left[\left(\mathbf{u}_t^{(1)}\right)^T, \ldots, \left(\mathbf{u}_t^{(S)}\right)^T\right]$ and $\mathbf{M}_t$ is the row-selection matrix determining the schedule at time t. Since there are packet losses, the actual transmission is given by

$$\mathbf{y}_t = (\mathbf{L}_t \otimes \mathbf{I})\,\mathbf{r}_t, \tag{25}$$

where $\mathbf{L}_t = \operatorname{diag}\left\{l_t^{(1)}, \ldots, l_t^{(R)}\right\}$ and $l_t^{(r)}$ is a binary random variable determining whether the packet associated with the rth scheduled measurement was lost ($l_t^{(r)} = 0$) or not ($l_t^{(r)} = 1$). Let $\mathbf{A} = \mathbf{V}\mathbf{F}\mathbf{V}^{-1}$ be the Jordan normal form of $\mathbf{F}$. Then, the system equations are given by (4)–(5), with

$$\mathbf{C}_t = \mathbf{B}_t\mathbf{H}\mathbf{V}^{-1}, \quad \mathbf{Q}_t = \mathbf{V}\mathbf{N}\mathbf{V}^T, \quad \mathbf{R}_t = \mathbf{B}_t\mathbf{E}\mathbf{B}_t^T,$$
$$\mathbf{x}_t = \mathbf{V}\mathbf{p}_t, \quad \mathbf{w}_t = \mathbf{V}\mathbf{n}_t, \quad \mathbf{v}_t = \mathbf{B}_t\mathbf{e}_t,$$

and

$$\mathbf{e}_t^T = \left[\left(\mathbf{e}_t^{(1)}\right)^T, \ldots, \left(\mathbf{e}_t^{(S)}\right)^T\right],$$
$$\mathbf{H}^T = \left[\left(\mathbf{H}^{(1)}\right)^T, \ldots, \left(\mathbf{H}^{(S)}\right)^T\right],$$
$$\mathbf{E} = \operatorname{diag}\left\{\mathbf{E}^{(1)}, \ldots, \mathbf{E}^{(S)}\right\},$$
$$\mathbf{B}_t = (\mathbf{L}_t \otimes \mathbf{I})(\mathbf{M}_t \otimes \mathbf{I}).$$

We consider below two scheduling strategies, namely, time-based schedule and random schedule.

7.1. Time-based schedule

In this case, the packet loss model $\mathbf{L}_t$ is stationary, independent of $\mathbf{w}_t$, $\mathbf{v}_t$ and $\mathbf{x}_0$, and either, finite-order proper Markov or Gaussian hidden Markov. The sequence of matrices $\mathbf{M}_t$, $t \in \mathbb{N}$, follows a periodic deterministic pattern, i.e.,

$$\mathbf{M}_t = \mathbf{M}_{t+\tau},$$

for all $t \in \mathbb{N}$ and some period $\tau \in \mathbb{N}$. Clearly, this leads to γ_t being cyclostationary with period τ, and satisfying the conditions of Proposition 18. Theorem 14 then holds.

7.2. Random schedule

In this case, both sequence of matrices $\mathbf{L}_t$ and $\mathbf{M}_t$ are randomly drawn at each $t \in \mathbb{N}$. This is done such that the sequences $(\mathbf{M}_t)$, $(\mathbf{L}_t)$, $(\mathbf{w}_t)$, $(\mathbf{v}_t)$ and $\mathbf{x}_0$ are mutually independent. The models describing the statistics of $\mathbf{L}_t$ and $\mathbf{M}_t$ are stationary and either finite-order proper Markov or Gaussian hidden Markov. This clearly leads to $\gamma_t = (\mathbf{C}_t, \mathbf{E}_t)$ satisfying the conditions of Proposition 18. Theorem 14 thus holds.

7.3. Example

In this section we use Theorem 14 and Corollary 26 to assess the stability of an example system. We consider a system whose state-transition matrix $\mathbf{A}$ is non-diagonalizable and whose measurement equation has statistics which are not finite-order Markov. Notice that, as mentioned in points 1 and 3 in the introduction, none of the results available in the literature could be used to assess the stability of a system with any of these two properties.

Consider a system whose dynamics is given by (4), with

$$\mathbf{A} = \operatorname{diag}\{\mathbf{A}_1, \mathbf{A}_2\}, \ \mathbf{A}_1 = \begin{bmatrix} \alpha_1 & 1 \\ 0 & \alpha_1 \end{bmatrix}, \ \mathbf{A}_2 = [\alpha_2],$$

for some $\alpha_1 > \alpha_2 > 0$. There are two sensors. For $i \in \{1, 2\}$, the measurement equation of sensor i is given by (24), with

$$\mathbf{H}^{(1)} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ and } \mathbf{H}^{(2)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix},$$

Due to communication constraints, the measurements from both sensors are alternatively transmitted, i.e.,

$$\mathbf{M}_t = \begin{cases} \begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} & t \text{ even}, \\ \begin{bmatrix} 0 & \mathbf{I} \end{bmatrix} & t \text{ odd}. \end{cases}$$

We assume that the communication channel has a packet loss model given by $\mathbf{L}_t = l_t \in \{0, 1\}$, where l_t is Gaussian hidden Markov. Hence, we have a time-based schedule, as described in Section 7.1. Thus, we can use the result of Theorem 14 to determine the stability of the Kalman filter.

We have

$$\mathbf{B}_t = \begin{cases} \begin{bmatrix} l_t\mathbf{I} & 0 \end{bmatrix} & t \text{ even}, \\ \begin{bmatrix} 0 & l_t\mathbf{I} \end{bmatrix} & t \text{ odd}. \end{cases}$$

Hence, the measurement equation of the aggregated system is given by (5), with $\mathbf{C}_t$ given by

$$\mathbf{C}_t = \begin{cases} l_t \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} & t \text{ even}, \\ l_t \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} & t \text{ odd}. \end{cases} \tag{26}$$

From Definition 8 the FMO blocks of the above system are $(\mathbf{A}_1, \mathcal{C}_1)$ and $(\mathbf{A}_2, \mathcal{C}_2)$ where

$$\mathcal{C}_k = \left\{\mathbf{C}_k^{(1)}, \mathbf{C}_k^{(2)}\right\}, \ k = 1, 2,$$

with

$$\mathbf{C}_1^{(1)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{C}_1^{(2)} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix},$$
$$\mathbf{C}_2^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \mathbf{C}_2^{(2)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Clearly, both FMO blocks satisfy the conditions of Corollary 26. Hence, we can apply this result to each block. Let

$$\lambda = \mathbb{P}\left(l_t = 0 | l_{2s} = 0, s < t\right).$$

We have

$$\mathbb{P}\left(\mathbf{C}_{t,1} = \mathbf{C}_1^{(0)} | \mathbf{C}_{s,1} = \mathbf{C}_1^{(0)}, s < t\right) = \begin{cases} \lambda & t \text{ even}, \\ 1 & t \text{ odd}, \end{cases}$$

$$\mathbb{P}\left(\mathbf{C}_{t,2} = \mathbf{C}_2^{(0)} | \mathbf{C}_{s,2} = \mathbf{C}_2^{(0)}, s < t\right) = \begin{cases} 1 & t \text{ even}, \\ \lambda & t \text{ odd}, \end{cases}$$

Then, from (22), since the cyclostationary period is $\tau = 2$, we obtain

$$\Phi_1 = \Phi_2 = \lambda^{1/2}.$$

It then follows from Theorem 14 that

$$\alpha_1^4\lambda < 1 \Rightarrow G < \infty,$$
$$\alpha_1^4\lambda > 1 \Rightarrow G = \infty.$$

8. Proof of the main result

This section presents a formal proof of the necessary and the sufficient conditions stated in Theorem 14. In Section 8.1 we derive certain preliminary results. More precisely, in Section 8.1.1 we provide lower and upper bounds on the growth rate of $\|\Psi\left(\mathbf{P}_t, \Gamma_{t,T}\right)\|$, and in Section 8.1.2 we state a technical condition to guarantee that the kernel of $\mathbf{O}\left(\Gamma_{t,T}\right)$ has certain desired orientation. In Section 8.2 we show the necessary condition. In Section 8.3.1 we derive a first sufficient condition, which differs from the desired one. This result is used in Section 8.3.2 to provide a second sufficient condition, seemingly stronger than the one in Theorem 14. We then show in Section 8.3.3 that the latter condition is indeed equivalent to the desired one.

8.1. Preliminary results

8.1.1. Bounds on the growth rate of $\|\Psi\left(\mathbf{P}_t, \Gamma_{t,T}\right)\|$

It turns out that the growth rate of $\|\Psi\left(\mathbf{P}_t, \Gamma_{t,T}\right)\|$ is determined by the location of the kernel of $\mathbf{O}\left(\Gamma_{t,T}\right)$. Recall from (12) that

$$\mathbf{z}_T = \mathbf{O}\left(\Gamma_{t,T}\right)\mathbf{x}_t + \mathbf{f}_t\left(\Gamma_{t,T}\right) \tag{27}$$
$$\mathbf{x}_T = \mathbf{A}^{T-t}\mathbf{x}_t + \mathbf{q}_{t,T} \tag{28}$$
$$\mathbf{q}_{t,T} = \sum_{j=t}^{t+T-1} \mathbf{A}^{t+T-1-j}\mathbf{w}_j. \tag{29}$$

From Anderson and Moore (1979, Ch. 5, Theorem 2.1), we have

$$\Psi\left(\mathbf{P}_t, \Gamma_{t,T}\right) = \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x},\mathbf{z}}\Sigma_{\mathbf{z}}^{\dagger}\Sigma_{\mathbf{x},\mathbf{z}}^{*}, \tag{30}$$

where † is the Moore–Penrose pseudo-inverse (Ben-Israel & Greville, 2003) and

$$\Sigma_{\mathbf{x}} = \mathbf{A}^T\mathbf{P}_t\mathbf{A}^{*T} + \mathbb{E}\left(\mathbf{q}_{t,T}\mathbf{q}_{t,T}^{*}\right),$$
$$\Sigma_{\mathbf{z}} = \mathbf{O}\left(\Gamma_{t,T}\right)\mathbf{P}_t\mathbf{O}\left(\Gamma_{t,T}\right)^{*} + \mathbb{E}\left(\mathbf{f}_t\left(\Gamma_{t,T}\right)\mathbf{f}\left(\Gamma_{t,T}\right)^{*}\right),$$
$$\Sigma_{\mathbf{x},\mathbf{z}} = \mathbf{A}^T\mathbf{P}_t\mathbf{O}\left(\Gamma_{t,T}\right)^{*} + \mathbb{E}\left(\mathbf{q}_{t,T}\mathbf{f}_t\left(\Gamma_{t,T}\right)^{*}\right).$$

Lemma 27. *Let* $\mathbf{A}$ *be a Jordan block of order* J *with eigenvalue* α*. Then, there exist* $c_1, c_2 \in \mathbb{R}$ *such that*

$$\|\mathbf{A}^t\| \leq |\alpha|^t c_1 t^{J-1} \tag{31}$$

$$\left\|\mathbf{A}^{-t}\right\|^{-1} \geq |\alpha|^t c_2 t^{1-J} \tag{32}$$

for all $t \in \mathbb{N}$.

Proof. The proof is divided in two steps.

Proof of (31): Notice that

$$\mathbf{A}^t = \alpha^t \mathbf{M}(t)$$

where

$$\mathbf{M}(t) = \begin{bmatrix} 1 & p_1(t) & \cdots & p_{J-1}(t) \\ 0 & 1 & \ddots & p_{J-2}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

and $p_j(t)$ is a polynomial in t of order j given by

$$p_j(t) = \binom{t}{j} \alpha^{-j}.$$

We then have

$$\begin{aligned} \left\|\mathbf{A}^t\right\| &= |\alpha|^t \|\mathbf{M}(t)\| \\ &\overset{(a)}{\leq} |\alpha|^t \left(1 + \sum_{j=1}^{J-1} |p_j(t)|\right) \\ &\leq |\alpha|^t c_1 t^{J-1}, \end{aligned}$$

for some $c_1 \in \mathbb{R}$, where (a) follows from Young's inequality (Katznelson, 2004, p. 115).

Proof of (32): Consider the matrix

$$\tilde{\mathbf{M}}(t) = \begin{bmatrix} 1 & \tilde{p}_1(t) & \cdots & \tilde{p}_{J-1}(t) \\ 0 & 1 & \ddots & \tilde{p}_{J-2}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

where $\tilde{p}_j(t)$ are polynomials in t of order j such that $\tilde{\mathbf{M}}(t) = \mathbf{M}^{-1}(t)$. This is always possible, since

$$\mathbf{M}(t)\tilde{\mathbf{M}}(t) = \begin{bmatrix} 1 & \hat{p}_1(t) & \cdots & \hat{p}_{J-1}(t) \\ 0 & 1 & \ddots & \hat{p}_{J-2}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

with

$$\hat{p}_j(t) = p_j(t) + \tilde{p}_j(t) + \sum_{i=1}^{j-1} p_{j-i}(t)\tilde{p}_i(t).$$

From the above, by making

$$\tilde{p}_j(t) = -p_j(t) - \sum_{i=1}^{j-1} p_{j-i}(t)\tilde{p}_i(t),$$

we have $\mathbf{M}(t)\tilde{\mathbf{M}}(t) = \mathbf{I}$. Then,

$$\left\|\mathbf{A}^{-t}\right\| = \|\alpha^{-t}\mathbf{M}^{-1}(t)\| \tag{33}$$

$$= |\alpha|^{-t}\|\tilde{\mathbf{M}}(t)\| \tag{34}$$

$$\leq |\alpha|^{-t}\left(1 + \sum_{j=1}^{J-1} |\tilde{p}_j(t)|\right) \tag{35}$$

$$\leq |\alpha|^{-t} c_3 t^{J-1}, \tag{36}$$

for some $c_3 \in \mathbb{R}$. Hence,

$$\left\|\mathbf{A}^{-t}\right\|^{-1} \geq |\alpha|^t c_3^{-1} t^{1-J}.$$

The result follows by making $c_2 = c_3^{-1}$.

The following two lemmas state bounds on the growth rate of $\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\|$. Firstly, some notation is introduced. Let $\mathbf{e}_{k,j}$ be the column vector with a 1 in one entry and zeros otherwise, such that $\mathbf{e}_{k,j}^\top \mathbf{A}\mathbf{e}_{k,j}$ equals the jth diagonal entry of the kth block $\mathbf{A}_k$ of $\mathbf{A}$. Let also $\mathcal{E}_k = \{\mathbf{e}_{k,1}, \ldots, \mathbf{e}_{k,K_k}\}$. The following lemma states an upper bound on the growth rate of $\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\|$.

Lemma 28. *Consider the system (4)–(5). If* $\ker\left\{\mathbf{O}\left(\Gamma_{t,T}\right)\right\} \subseteq \operatorname{span}\{\mathcal{E}_k, \ldots, \mathcal{E}_K\}$ *for some* $1 \leq k \leq K$*, then, there exist* $l_T > 0$ *and* $c_1 > 0$*, such that, for any* $\mathbf{P}_t$*,*

$$\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\| \leq |\alpha_k|^{2t} c_1 T^{2(\bar{J}-1)}\|\mathbf{P}_t\| + l_T. \tag{37}$$

Also, if $\ker\left\{\mathbf{O}\left(\Gamma_{t,T}\right)\right\} = \{\mathbf{0}\}$, then

$$\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\| \leq l_T. \tag{38}$$

Proof. Following the steps of the proof of Rohr et al. (2014, Lemma 20), we obtain (38) and

$$\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right) \leq \mathbf{M}_{t,T} + l_T\mathbf{I}, \tag{39}$$

where $\mathbf{M}_{t,T} = \mathbf{\Pi}\mathbf{A}^T\mathbf{P}_t\mathbf{A}^{*T}\mathbf{\Pi}$ and

$$\mathbf{\Pi} = \operatorname{diag}(\mathbf{0}_1, \ldots, \mathbf{0}_{k-1}, \mathbf{X}),$$

with $\mathbf{0}_j$ being a square matrix of zeros with the same dimension of $\mathbf{A}_j$ and $\mathbf{X}$ is a non-zero matrix with appropriate dimensions. Also,

$$l_T = \sup_{t\in\mathbb{Z}} \max_{\Gamma_{t,T}\in\mathcal{A}^T} \vartheta_t\left(\Gamma_{t,T}\right),$$

with

$$\vartheta_t\left(\Gamma_{t,T}\right) = \left\|\mathbb{E}\left(\mathbf{U}_t\left(\Gamma_{t,T}\right)\mathbf{U}_t^*\left(\Gamma_{t,T}\right)\middle|\Gamma_{t,T}\right)\right\|,$$

$$\mathbf{U}_t\left(\Gamma_{t,T}\right) = \mathbf{q}_{t,T} - \left(\mathbf{O}\left(\Gamma_{t,T}\right)\mathbf{A}^{-T}\right)^\dagger \mathbf{f}_t\left(\Gamma_{t,T}\right).$$

Notice that the map $\vartheta_t : \mathcal{A}^T \to \mathbb{R}$ is independent of $t \in \mathbb{Z}$. Hence, l_T is the result of a maximization over the finite set $\mathcal{A}^T$. This guarantees that l_T is finite. Hence, (38) clearly holds.

Let $\mathbf{M}_{t,T} = \mathbf{N}_{t,T}^*\mathbf{N}_{t,T}$, with $\mathbf{N}_{t,T} = \mathbf{P}_t^{1/2}\mathbf{A}^{*T}\mathbf{\Pi}$ and $\tilde{\mathbf{A}} = \operatorname{diag}\{\mathbf{0}, \ldots, \mathbf{0}, \mathbf{A}_k, \ldots, \mathbf{A}_K\}$. Then,

$$\begin{aligned} \|\mathbf{M}_{t,T}\| &= \|\mathbf{N}_{t,T}\|^2 \leq \|\mathbf{P}_t\|\|\mathbf{A}^{*T}\mathbf{\Pi}\|^2 \\ &= \|\mathbf{P}_t\|\|\tilde{\mathbf{A}}^{*T}\mathbf{\Pi}\|^2 \leq \|\mathbf{P}_t\|\|\tilde{\mathbf{A}}^{*T}\|^2\|\mathbf{\Pi}\|^2. \end{aligned}$$

Notice that $\|\tilde{\mathbf{A}}^{*T}\| = \max_{j\geq k}\|\mathbf{A}_j^T\|$. Using Lemma 27, we obtain

$$\|\mathbf{M}_{t,T}\| \leq \|\mathbf{P}_t\|\,|\alpha_k|^{2T} c_1 T^{2(\bar{J}-1)},$$

for some $c_1 \in \mathbb{R}$ and the result follows from (39).

The next lemma states a lower bound on the growth rate of $\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\|$.

Lemma 29. *Consider the system (4)–(5). If* $\ker\left\{\mathbf{O}\left(\Gamma_{t,T}\right)\right\} \cap \operatorname{span}\{\mathcal{E}_k\} \neq \{0\}$*, then, there exists* $c_2 > 0$ *such that for all* $\mathbf{P}_t$*,*

$$\left\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right\| \geq |\alpha_k|^{2T} c_2 T^{2(1-\bar{J})}\left\|\mathbf{P}_t^{-1}\right\|^{-1}.$$

Proof. Following the steps of the proof of Rohr et al. (2014, Lemma 21), we obtain

$$\left\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right\| \geq \left\|\mathbf{P}_t^{-1}\right\|^{-1}\left\|\mathbf{A}^T\mathbf{\Pi}\right\|^2, \tag{40}$$

where $\mathbf{\Pi} = \left[\mathbf{I} - \mathbf{O}\left(\Gamma_{t,T}\right)^{\dagger}\mathbf{O}\left(\Gamma_{t,T}\right)\right]$. Now, let $\mathbf{x} \in \ker\left\{\mathbf{O}\left(\Gamma_{t,T}\right)\right\} \cap \operatorname{span}\{\mathcal{E}_k\}$. Since $\mathbf{\Pi}$ is the projection onto the kernel of $\mathbf{O}\left(\Gamma_{t,T}\right)$, we have

$$\left\|\mathbf{A}^T\mathbf{\Pi}\mathbf{x}\right\| = \left\|\mathbf{A}^T\mathbf{x}\right\|.$$

Since $\mathbf{x} \in \operatorname{span}(\mathcal{E}_k)$, we have that

$$\|\mathbf{A}^T\mathbf{x}\| = \|\mathbf{A}_k^T\mathbf{v}\| \geq \|\mathbf{A}_k^{-T}\|^{-1}\|\mathbf{v}\| = \|\mathbf{A}_k^{-T}\|^{-1}\|\mathbf{x}\|,$$

for some $\mathbf{v} \in \mathbb{C}^{N_k}$. From Lemma 27, and the above, it follows that

$$\|\mathbf{A}^T\mathbf{\Pi}\mathbf{x}\| \geq c_2|\alpha_k|^T T^{1-\bar{J}}\|\mathbf{x}\|,$$

hence

$$\|\mathbf{A}^T\mathbf{\Pi}\| \geq |\alpha_k|^T T^{1-\bar{J}}. \tag{41}$$

The result then follows by substituting (41) into (40).

8.1.2. A condition to guarantee that $\ker\{\mathbf{O}(\Gamma_{t,T})\}$ *is orthogonal to* $\operatorname{span}\{\mathcal{E}_k\}$

Definition 30. A matrix $\mathbf{M}$ is said to have full column rank with strength $q \in \mathbb{N}_0$ ($\mathbb{N}_0 = \mathbb{N} \cup \{0\}$), denoted by FCR($q$), if $\mathbf{M}$ has more than q rows and the matrix obtained after removing any q rows from $\mathbf{M}$ still has FCR.

The main goal of this section is to show the following lemma.

Lemma 31. *There exists $Q \in \mathbb{N}_0$ such that, for any $1 \leq k \leq K$, if $\mathbf{O}_k(\Gamma_{t,T})$ has* FCR(Q)*, then* $\ker\{\mathbf{O}(\Gamma_{t,T})\} \perp \operatorname{span}\{\mathcal{E}_k\}$.

Proof. See the full version (Marelli et al., 2018, Proof of Lemma 31).

8.2. Proof of the necessary condition in Theorem 14

Following the steps of the proof in Rohr et al. (2014, Section V-B) we obtain

$$\left\|\mathbb{E}\left(\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\| \geq \max_{1\leq k\leq K}\frac{1}{n}\sum_{\Gamma_{t,T}\in\mathcal{N}_k^{t,T}}\mathbb{P}\left(\Gamma_{t,T}\right)\left\|\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right\|.$$

From Lemma 29, we have that, for all $k = 1, \ldots, K$ and $t \in \mathbb{N}$,

$$\begin{aligned}
&\left\|\mathbb{E}\left(\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\| \\
\geq&\frac{1}{n}\sum_{\Gamma_{t,T}\in\mathcal{N}_k^{t,T}}\mathbb{P}\left(\Gamma_{t,T}\right)|\alpha_k|^{2T}c_2T^{2(1-\bar{J})}\left\|\mathbf{P}_t^{-1}\right\|^{-1} \\
=&\frac{1}{n}|\alpha_k|^{2T}c_2T^{2(1-\bar{J})}\left\|\mathbf{P}_t^{-1}\right\|^{-1}\mathbb{P}\left(\mathcal{N}_k^{t,T}\right) \\
=&\left(|\alpha_k|^2\mathbb{P}\left(\mathcal{N}_k^{t,T}\right)^{1/T}T^{2(1-\bar{J})/T}\right)^T c_2\frac{\left\|\mathbf{P}_t^{-1}\right\|^{-1}}{n}.
\end{aligned}$$

For any $t \in \mathbb{N}$, put $\mathbf{P}_t = \mathbf{P}_0$. Then,

$$\begin{aligned}
&\max_{0\leq t<\tau}\limsup_{T\to\infty}\left\|\mathbb{E}\left(\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\| \\
\geq&\max_{0\leq t<\tau}\limsup_{T\to\infty}\left(|\alpha_k|^2\mathbb{P}\left(\mathcal{N}_k^{t,T}\right)^{1/T}T^{2(1-\bar{J})/T}\right)^T \\
&\times c_2\frac{\left\|\mathbf{P}_0^{-1}\right\|^{-1}}{n} \\
=&c_2\frac{\left\|\mathbf{P}_0^{-1}\right\|^{-1}}{n}\max_{0\leq t<\tau}\limsup_{T\to\infty}a_{t,T}^T,
\end{aligned}$$

where

$$a_{t,T} = |\alpha_k|^2\mathbb{P}\left(\mathcal{N}_k^{t,T}\right)^{1/T}T^{2(1-\bar{J})/T}.$$

Choose $k = 1, \ldots, K$ satisfying (19). Then

$$\begin{aligned}
&\max_{0\leq t<\tau}\limsup_{T\to\infty}a_{t,T} \\
=&\max_{0\leq t<\tau}\limsup_{T\to\infty}|\alpha_k|^2\mathbb{P}\left(\mathcal{N}_k^{t,T}\right)^{1/T}T^{2(1-\bar{J})/T} \\
=&|\alpha_k|^2\max_{0\leq t<\tau}\limsup_{T\to\infty}\mathbb{P}\left(\mathcal{N}_k^{t,T}\right)^{1/T}\lim_{T\to\infty}T^{2(1-\bar{J})/T} \\
=&|\alpha_k|^2\Phi_k\left(\lim_{T\to\infty}T^{1/T}\right)^{2(\bar{J}-1)} = |\alpha_k|^2\Phi_k > 1.
\end{aligned}$$

Hence, if we choose $\left\|\mathbf{P}_0^{-1}\right\|^{-1} > 0$, then

$$G = \max_{0\leq t<\tau}\limsup_{T\to\infty}\left\|\mathbb{E}\left(\mathbf{\Psi}\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\| = \infty,$$

and the result follows.

8.3. Proof of the sufficient condition in Theorem 14

8.3.1. First step

Define the map $\xi_{t,T} : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ by

$$\xi_{t,T}(x) = \sup_{\varrho_{t-1}\in\mathcal{E}}\sum_{\Gamma_{t,T}\in\mathcal{A}^T}\mathbb{P}\left(\Gamma_{t,T}|\varrho_{t-1}\right) \times\operatorname{Tr}\left\{\mathbf{\Psi}(x\mathbf{I}, \Gamma_{t,T})\right\}. \tag{42}$$

Lemma 32. *Let $T, S \in \mathbb{N}$ and $x, y > 0$. Then*

(1) if $x \geq y$, then $\xi_{t,T}(x) \geq \xi_{t,T}(y)$;
(2) $\xi_{t,T+S}(x) \leq \xi_{t+T,S} \circ \xi_{t,T}(x)$.

Proof. Proof of (1) From Sinopoli et al. (2004, Lemma 1c),

$$x \geq y \Rightarrow \mathbf{\Psi}(x\mathbf{I}, \Gamma_{t,T}) \geq \mathbf{\Psi}(y\mathbf{I}, \Gamma_{t,T}). \tag{43}$$

Then

$$\begin{aligned}
\xi_{t,T}(y) &= \sup_{\varrho_{t-1}}\sum_{\Gamma_{t,T}}\mathbb{P}\left(\Gamma_{t,T}|\varrho_{t-1}\right)\operatorname{Tr}\left\{\mathbf{\Psi}(y\mathbf{I}, \Gamma_{t,T})\right\} \\
&\leq \sup_{\varrho_{t-1}}\sum_{\Gamma_{t,T}}\mathbb{P}\left(\Gamma_{t,T}|\varrho_{t-1}\right)\operatorname{Tr}\left\{\mathbf{\Psi}(x\mathbf{I}, \Gamma_{t,T})\right\} \\
&= \xi_{t,T}(x).
\end{aligned}$$

Proof of (2) We have

$$\begin{aligned}
&\xi_{t,T+S}(x) \\
=&\sup_{\varrho_{t-1}}\sum_{\Gamma_{t,T+S}}\mathbb{P}\left(\Gamma_{t,T+S}|\varrho_{t-1}\right) \\
&\times\operatorname{Tr}\left\{\mathbf{\Psi}\left(x\mathbf{I}, \Gamma_{t,T+S}\right)\right\} \\
=&\sup_{\varrho_{t-1}}\sum_{\Gamma_{t,T}}\sum_{\Gamma_{t+T,S}}\mathbb{P}\left(\Gamma_{t+T,S}|\Gamma_{t,T}, \varrho_{t-1}\right)\times \\
&\mathbb{P}\left(\Gamma_{t,T}|\varrho_{t-1}\right)\operatorname{Tr}\left\{\mathbf{\Psi}\left(\mathbf{\Psi}\left(x\mathbf{I}, \Gamma_{t,T}\right), \Gamma_{t+T,S}\right)\right\} \\
\leq&\sup_{\varrho_{t+T-1}\atop\varrho_{t-1}}\sum_{\Gamma_{t,T}}\sum_{\Gamma_{t+T,S}}\mathbb{P}\left(\Gamma_{t+T,S}|\varrho_{t+T-1}\right)\times \\
&\mathbb{P}\left(\Gamma_{t,T}|\varrho_t\right)\operatorname{Tr}\left\{\mathbf{\Psi}\left(\mathbf{\Psi}\left(x\mathbf{I}, \Gamma_{t,T}\right), \Gamma_{t+T,S}\right)\right\}.
\end{aligned}$$

Using (43) and the concavity of $\Psi(\cdot, \Gamma_{t+T,S})$ (Sinopoli et al., 2004, Lemma 1e), we have

$$\begin{aligned}&\xi_{t,T+S}(x)\\ &\le \sup_{\varrho_{t+T-1}} \sum_{\Gamma_{t+T,S}} \mathbb{P}\left(\Gamma_{t+T,S}|\varrho_{t+T-1}\right)\\ &\times \mathrm{Tr}\left\{\Psi\left(\Xi(x), \Gamma_{t+T,S}\right)\right\},\end{aligned}$$

with

$$\Xi(x) = \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T}} \mathbb{P}\left(\Gamma_{t,T}\,|\varrho_{t-1}\right) \Psi\left(x\mathbf{I}, \Gamma_{t,T}\right).$$

Now, since $\Psi(x\mathbf{I}, \Gamma_{t,T}) \ge 0$,

$$\begin{aligned}\Xi(x) &\le \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T}} \mathbb{P}\left(\Gamma_{t,T}\,|\varrho_{t-1}\right) \mathrm{Tr}\left\{\Psi\left(x\mathbf{I}, \Gamma_{t,T}\right)\right\}\mathbf{I}\\ &= \xi_{t,T}(x)\mathbf{I}.\end{aligned}$$

Then, using (43), it follows that

$$\begin{aligned}\xi_{t,T+S}(x) &\le \sup_{\varrho_{t+T-1}} \sum_{\Gamma_{t+T,S}} \mathbb{P}\left(\Gamma_{t+T,S}|\varrho_{t+T-1}\right)\\ &\quad\times \mathrm{Tr}\left\{\Psi\left(\xi_{t,T}(x)\mathbf{I}, \Gamma_{t+T,S}\right)\right\}\\ &= \xi_{t+T,S}\left(\xi_{t,T}(x)\right) = \xi_{t+T,S} \circ \xi_{t,T}(x).\end{aligned}$$

The reason why the map $\xi_{t,T}(x)$ is particularly important for our analysis is because of the following result, which states a sufficient condition for G to be finite.

Lemma 33. *If there exist $T \in \mathbb{N}$ and $\bar{x} > 0$ such that $\max_{0\le t<\tau} \limsup_{k\to\infty} \xi_{t,kT}(x) \le \bar{x} < \infty$, for all $x > 0$, then*

$G < \infty$.

Proof. Following the steps in Rohr et al. (2014, Lemma 30), we can show that

$$\left\|\mathbb{E}\left(\Psi\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\| \le \xi_{t,T}(\|\mathbf{P}_t\|). \tag{44}$$

Now, from (44) Lemmas 28 and 32

$$\begin{aligned}&\left\|\mathbb{E}\left(\Psi\left(\mathbf{P}_t, \Gamma_{t,kT+S}\right)\right)\right\|\\ &\le \xi_{t,kT+S}(\|\mathbf{P}_t\|)\\ &\le \xi_{t+kT,S} \circ \xi_{t,kT}(\|\mathbf{P}_t\|)\\ &= \sup_{\varrho_{s-1}} \sum_{\Gamma_{s,S}} \mathbb{P}\left(\Gamma_{s,S}\,|\varrho_{s-1}\right) \mathrm{Tr}\left\{\Psi\left(\xi_{t,kT}(\|\mathbf{P}_t\|)\mathbf{I}, \Gamma_{s,S}\right)\right\}\\ &\le n \sup_{\varrho_{s-1}} \sum_{\Gamma_{s,S}} \mathbb{P}\left(\Gamma_{s,S}\,|\varrho_{s-1}\right) \left\|\Psi\left(\xi_{t,kT}(\|\mathbf{P}_t\|)\mathbf{I}, \Gamma_{s,S}\right)\right\|\\ &\le n \sup_{\varrho_{s-1}} \sum_{\Gamma_{s,S}} \mathbb{P}\left(\Gamma_{s,S}\,|\varrho_{s-1}\right)\\ &\times \left(\xi_{t,kT}(\|\mathbf{P}_t\|)|\alpha_1|^{2S} c_1 s^{2(\bar{J}-1)} + l_S\right)\\ &= n\xi_{t,kT}(\|\mathbf{P}_t\|)|\alpha_1|^{2S} c_1 s^{2(\bar{J}-1)} + n l_S.\end{aligned}$$

Hence

$$\begin{aligned}&\max_{0\le t<\tau} \limsup_{T\to\infty} \left\|\mathbb{E}\left(\Psi\left(\mathbf{P}_t, \Gamma_{t,T}\right)\right)\right\|\\ &\le \sup_{0<S<T} n|\alpha_1|^{2S} c_1 s^{2(\bar{J}-1)}\\ &\times \max_{0\le t<\tau} \limsup_{k\to\infty} \xi_{t,kT}(\|\mathbf{P}_t\|) + n l_S\\ &\le \sup_{0<S<T} n|\alpha_1|^{2S} c_1 s^{2(\bar{J}-1)}\bar{x} + n l_S.\end{aligned}$$

Finally, since the above bound is independent of t and $\mathbf{P}_t$,

$$G \le \sup_{0<S<T} n|\alpha_1|^{2S} c_1 s^{2(\bar{J}-1)}\bar{x} + n l_S < \infty.$$

8.3.2. Second step

An alternative sufficient condition for the ANEEC G to be bounded is now presented. It will be shown in Section 8.3.3 that this condition is equivalent to the sufficient condition in Theorem 14.

Notation 34. *For $k = 1, \ldots, K$, let*

$\mathcal{N}_{k,Q}^{t,T} = \{\Gamma_{t,T} : \mathbf{O}_k(\Gamma_{t,T})$ *does not have FCR*$(Q)\}$,

Notice that $\mathcal{N}_{k,0}^{t,T} = \mathcal{N}_k^{t,T}$.

The following lemma presents a sufficient condition for the ANEEC to be bounded.

Lemma 35. *Under Assumption 12, there exists $Q \in \mathbb{N}_0$ such that, if*

$$|\alpha_k|^2 \max_{0\le t<\tau} \limsup_{T\to\infty} \mathbb{P}\left(\mathcal{N}_{k,Q}^{t,T}\right)^{1/T} < 1 \text{ for all } k = 1, \ldots, K, \tag{45}$$

then $G < \infty$.

Proof. The proof is divided into 5 steps.

(1) In view of Lemma 31, there exists Q such that if $\mathbf{O}_k\left(\Gamma_{t,T}\right)$ has FCR(Q), for $k = 1, \ldots, K$, then $\mathbf{O}\left(\Gamma_{t,T}\right)$ has FCR. Recall from Notation 34 that $\mathcal{N}_{k,Q}^{t,T}$ is the set of sequences $\Gamma_{t,T} \in \mathcal{A}^T$ such that $\mathbf{O}_k(\Gamma_{t,T})$ does not have FCR with strength Q. Define the set

$$\mathcal{G}_{k,Q}^{t,T} \triangleq \begin{cases}\mathcal{N}_{1,Q}^{t,T}, & k = 1,\\ \bigcap_{j=1}^{k-1} \overline{\mathcal{N}_{j,Q}^{t,T}} \cap \mathcal{N}_{k,Q}^{t,T}, & k = 2, \ldots, K,\end{cases} \tag{46}$$

where $\overline{\mathcal{X}}$ denotes the complement of the set $\mathcal{X}$. Notice that $\mathcal{G}_{k,Q}^{t,T}$ is the set of sequences $\Gamma_{t,T} \in \mathcal{A}^T$ such that $\mathbf{O}_j(\Gamma_{t,T})$ has FCR(Q) for $1 \le j \le k-1$ and $\mathbf{O}_k(\Gamma_{t,T})$ does not have FCR(Q). Hence

$$\mathcal{G}_{k,Q}^{t,T} \subseteq \mathcal{N}_{k,Q}^{t,T}. \tag{47}$$

Let $\mathcal{N}^{t,T}$ be the set of sequences $\Gamma_{t,T} \in \mathcal{A}^T$ such that $\mathbf{O}(\Gamma_{t,T})$ does not have FCR. From Lemma 31, we have

$$\mathcal{A}^T = \overline{\mathcal{N}^{t,T}} \cup \bigcup_{k=1}^{K} \mathcal{G}_{k,Q}^{t,T},\ \forall t \in \mathbb{Z}.$$

From Lemma 28, there exists $c_1 > 0$ such that

$$\begin{aligned}&\Gamma_{t,T} \in \overline{\mathcal{N}^{t,T}} \implies \|\Psi(x\mathbf{I}, \Gamma_{t,T})\| \le l_T\\ &\Gamma_{t,T} \in \mathcal{G}_k^{t,T} \implies \|\Psi(x\mathbf{I}, \Gamma_{t,T})\| \le |\alpha_k|^{2T} c_1 T^{2(\bar{J}-1)} x + l_T.\end{aligned}$$

(2) Recall that $\Psi(x\mathbf{I}, \Gamma_{t,T})$ is a symmetric, positive-definite matrix with dimension n. Hence

$$\mathrm{Tr}\left(\Psi(x\mathbf{I}, \Gamma_{t,T})\right) \le n \left\|\Psi(x\mathbf{I}, \Gamma_{t,T})\right\|.$$

From (42) and the above, we have

$$\begin{aligned}&\xi_{t,T}(x)\\ &\le n \sup_{\varrho_{t-1}} \left\|\sum_{\Gamma_{t,T}} \mathbb{P}\left(\Gamma_{t,T}\,|\varrho_{t-1}\right) \Psi(x\mathbf{I}, \Gamma_{t,T})\right\|\\ &\le n \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T}\in\overline{\mathcal{N}^{t,T}}} \mathbb{P}\left(\Gamma_{t,T}\,|\varrho_{t-1}\right) \left\|\Psi(x\mathbf{I}, \Gamma_{t,T})\right\|\\ &+ n \sum_{k=1}^{K} \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T}\in\mathcal{G}_k^{t,T}} \mathbb{P}\left(\Gamma_{t,T}\,|\varrho_{t-1}\right) \left\|\Psi(x\mathbf{I}, \Gamma_{t,T})\right\|.\end{aligned} \tag{48}$$

Let

$$A_{t,T} = \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T} \in \overline{\mathcal{N}^{t,T}}} \mathbb{P}\left(\Gamma_{t,T} \,|\varrho_{t-1}\right),$$

$$B_{t,T,k} = \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T} \in \mathcal{G}_k^{t,T}} \mathbb{P}\left(\Gamma_{t,T} \,|\varrho_{t-1}\right),$$

$$C_{t,T,k} = \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T} \in \mathcal{N}_{k,Q}^{t,T}} \mathbb{P}\left(\Gamma_{t,T} \,|\varrho_{t-1}\right).$$

From (47), $B_{t,T,k} \le C_{t,T,k}$. Then, from (48), Lemma 28 and the above, we have

$$\begin{aligned}
&\xi_{t,T}(x)\\
&\le n l_T A_{t,T} + n \sum_{k=1}^{K} \left(|\alpha_k|^{2T} c_1 T^{2(\bar{J}-1)} x + l_T\right) B_{t,T,k}\\
&\le n l_T A_{t,T} + n \sum_{k=1}^{K} \left(|\alpha_k|^{2T} c_1 T^{2(\bar{J}-1)} x + l_T\right) C_{t,T,k}\\
&= n \sum_{k=1}^{K} \left(|\alpha_k|^2 C_{t,T,k}^{1/T}\right)^T c_1 T^{2(\bar{J}-1)} x\\
&+ n l_T \left(A_{t,T} + \sum_{k=1}^{K} C_{t,T,k}\right)\\
&= \beta_{t,T} x + \varphi_{t,T},
\end{aligned} \tag{49}$$

with

$$\beta_{t,T} = n c_1 T^{2(\bar{J}-1)} \sum_{k=1}^{K} \left(|\alpha_k|^2 C_{t,T,k}^{1/T}\right)^T,$$

$$\varphi_{t,T} = n l_T \left(A_{t,T} + \sum_{k=1}^{K} C_{t,T,k}\right).$$

(3) Recall the definition of ζ from (15). We have

$$\begin{aligned}
C_{t,T,k} &= \sup_{\varrho_{t-1}} \sum_{\Gamma_{t,T} \in \mathcal{N}_{k,Q}^{t,T}} \mathbb{P}\left(\Gamma_{t,T} \,|\varrho_{t-1}\right)\\
&= \sup_{\varrho_{t-1}} \sum_{\substack{\Gamma_{t,T} \in \mathcal{N}_{k,Q}^{t,T}\\ \mathbb{P}(\Gamma_{t,T}) \neq 0}} \frac{\mathbb{P}\left(\Gamma_{t,T} | \varrho_{t-1}\right)}{\mathbb{P}\left(\Gamma_{t,T}\right)} \mathbb{P}\left(\Gamma_{t,T}\right)\\
&\le \zeta \sum_{\substack{\Gamma_{t,T} \in \mathcal{N}_{k,Q}^{t,T}\\ \mathbb{P}(\Gamma_{t,T}) \neq 0}} \mathbb{P}\left(\Gamma_{t,T}\right) = \zeta \mathbb{P}\left(\mathcal{N}_{k,Q}^{t,T}\right).
\end{aligned}$$

Hence,

$$\beta_{t,T} \le \zeta n c_1 T^{2(\bar{J}-1)} \sum_{k=1}^{K} \left(|\alpha_k|^2 \mathbb{P}\left(\mathcal{N}_{k,Q}^{t,T}\right)^{1/T}\right)^T. \tag{50}$$

(4) From (45),

$$\max_{0 \le t < \tau} \limsup_{T \to \infty} |\alpha_k|^2 \mathbb{P}\left(\mathcal{N}_{k,Q}^{t,T}\right)^{1/T} < 1. \tag{51}$$

In view of (51), there exists $\bar{T} \in \mathbb{N}$ such $\bar{\beta} \triangleq \max_{0 \le t < \tau} \beta_{t,\bar{T}} < 1$. We also have that, for all $0 \le t < \tau$,

$$\varphi_{t,\bar{T}} \le \bar{\varphi} \triangleq n l_{\bar{T}} (1 + K). \tag{52}$$

Let $x_{k+1} = \bar{\beta} x_k + \bar{\varphi}$. We have that, for any x_0,

$$\lim_{k \to \infty} x_k \le \frac{\bar{\varphi}}{1 - \bar{\beta}}.$$

Also, from (49), $\xi_{t,k\bar{T}}(x_0) \le x_k$, for all $0 \le t < \tau$. Hence

$$\max_{0 \le t < \tau} \limsup_{k \to \infty} \xi_{t,k\bar{T}}(x_0) \le \lim_{k \to \infty} x_k \le \frac{\bar{\varphi}}{1 - \bar{\beta}} < \infty,$$

and the result follows from Lemma 33.

8.3.3. Third step

The main goal of this section is to show that, for any $q \in \mathbb{N}_0$ and $k \in \{1, \ldots, K\}$,

$$\max_{0 \le t < \tau} \limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}_{k,q}^{t,T}\right)^{1/T} = \max_{0 \le t < \tau} \limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}_{k}^{t,T}\right)^{1/T}.$$

Since in our analysis the value of k is fixed, we remove it from the notation.

Lemma 36. *For any $Q \in \mathbb{N}_0$ and $t \in \mathbb{Z}$*

$$\limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}_{Q}^{t,T}\right)^{1/T} \le \max_{0 \le i < l} \rho(\check{\zeta}_t(i, i))^{1/M},$$

with equality holding when $Q = 0$.

Proof. See the full version (Marelli et al., 2018, Proof of Lemma 46).

Lemma 37. *For any $q \in \mathbb{N}_0$ and $t \in \mathbb{Z}$,*

$$\limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}_{q}^{t,T}\right)^{1/T} = \limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}^{t,T}\right)^{1/T}.$$

Proof. We have $\mathcal{N}^{t,T} \subseteq \mathcal{N}_q^{t,T}$. Hence, using Lemma 36 we obtain,

$$\begin{aligned}
&\limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}^{t,T}\right)^{1/T} \le \limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}_q^{t,T}\right)^{1/T}\\
&\le \rho_t^{1/M} = \limsup_{T \to \infty} \mathbb{P}\left(\mathcal{N}^{t,T}\right)^{1/T},
\end{aligned}$$

and the result follows.

We are now ready to prove the sufficient condition in Theorem 14.

Proof of the sufficient condition in Theorem 14. The sufficient condition in Theorem 14, i.e., (19), follows immediately from Lemmas 35 and 37.

9. Conclusion

We stated a necessary and sufficient condition for stability of a Kalman filter under general assumptions on the linear system and its random measurement equation. We also studied how to numerically compute this condition for a given system. Furthermore, we used our result to assess the stability in a networked setting involving sensor scheduling and packet dropouts. This shows how our stability condition is a rather general one that could be applied in a widely range of applications, including those found in networked control settings.

References

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering, Vol. 11*. Prentice-Hall Englewood Cliffs, NJ.

Ben-Israel, A., & Greville, T. N. E. (2003). *Generalized inverses: Theory and applications, Vol. 15*. Springer Verlag.

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

Dana, A. F., Gupta, V., Hespanha, J. P., Hassibi, B., & Murray, R. M. (2007). Estimation over communication networks: Performance bounds and achievability results. In *Proc. amer. control conf.*, New York, NY (pp. 3450–3455), 9-13.

Elliott, E. O. (1963). Estimates of error rates for codes on burst-noise channels. *Bell System Technical Journal*, *42*(9), 1977–1997.

Gilbert, E. N. (1960). Capacity of a burst-noise channel. *Bell System Technical Journal*, *39*(9), 1253–1265.

Gupta, V., Chung, T. H., Hassibi, B., & Murray, R. M. (2006). On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage. *Automatica*, *42*(2), 251–260.

Hovareshti, P., Gupta, V., & Baras, J. (2007). Sensor scheduling using smart sensors. In *Decision and control, 46th IEEE conference on* (pp. 494–499).

Huang, M., & Dey, S. (2007). Stability of kalman filtering with markovian packet losses. *Automatica*, *43*(4), 598–607.

Katznelson, Yitzhak (2004). *An introduction to harmonic analysis (cambridge mathematical library)* (3rd ed.). Cambridge University Press.

Liu, X., & Goldsmith, A. (2004). Kalman filtering with partial observation losses. In *Proc. IEEE conf. decision control*, Paradise, The Bahamas (pp. 4180–4186).

Marelli, Damián, Sui, Tianju, Rohr, Eduardo, & Fu, Minyue (2018). Stability of Kalman filtering with a random measurement equation: Application to sensor scheduling with intermittent observations. arXiv:1806.08098v1 [cs.SY].

Mo, Y., & Sinopoli, B. (2008). A characterization of the critical value for kalman filtering with intermittent observations. In *IEEE conf. decision control*, Cancun, Mexico (pp. 2692–2697).

Mo, Y., & Sinopoli, B. (2012). Kalman filtering with intermittent observations: Tail distribution and critical value. *IEEE Transactions on Automatic Control*, *57*(3), 677–689.

Park, S. Y., & Sahai, A. (2011). Intermittent kalman filtering: Eigenvalue cycles and nonuniform sampling. In *Proc. amer. control conf.*, San Francisco, CA, (pp. 3692–3697), 06.

Plarre, K., & Bullo, F. (2009). On kalman filtering for detectable systems with intermittent observations. *IEEE Transactions on Automatic Control*, *54*(2), 386–390.

Quevedo, D. E., Ahlen, A., & Johansson, K. H. (2013). State estimation over sensor networks with correlated wireless fading channels. *IEEE Transactions on Automatic Control*, *58*(3), 581–593.

Rohr, E., Marelli, D., & Fu, M. (2010a). On the error covariance distribution for kalman filters with packet dropouts. In Mario Alberto Jordán (Ed.), *Discrete time systems* (pp. 71–92). InTech.

Rohr, E., Marelli, D., & Fu, M. (2010b). 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, (pp. 5881–5886), 15-17.

Rohr, E., Marelli, D., & Fu, M. (2011). Kalman filtering for a class of degenerate systems with intermittent observations. In *Proc. IEEE conf. decision control and european control conf.*, Orlando, FL (pp. 2422–2427).

Rohr, Eduardo Rath, Marelli, Damin, & Fu, Minyue (2014). Kalman filtering with intermittent observations: On the boundedness of the expected error covariance. *IEEE Transactions on Automatic Control*, *59*(10), 2724–2738.

Sadeghi, P., Kennedy, R., Rapajic, P., & Shams, R. (2008). Finite-state markov modeling of fading channels: a survey of principles and applications. *IEEE Signal Processing Magazine*, *25*(5), 57–80.

Savage, C., & Scala, B. (2009). Optimal scheduling of scalar gauss-markov systems with a terminal cost function. *IEEE Transactions on Automatic Control*, *54*(5), 1100–1105.

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

Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., & Sastry, S. S. (2007). Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, *95*(1), 163–187.

Shi, L., Cheng, P., & Chen, J. (2011a). Optimal periodic sensor scheduling with limited resources. *IEEE Transactions on Automatic Control*, *56*(9), 2190–2195.

Shi, L., Cheng, P., & Chen, J. (2011b). Sensor data scheduling for optimal state estimation with communication energy constraint. *Automatica*, *47*(8), 1693–1698.

Shi, L., Epstein, M., Sinopoli, B., & Murray, M. (2007). Effective sensor scheduling schemes in a sensor network by employing feedback in the communication loop. In *Control applications, 2007. CCA 2007. IEEE international conference on* (pp. 1006–1011).

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

Sui, Tianju, You, Keyou, & Fu, Minyue (2015). Optimal sensor scheduling for state estimation over lossy channel. *IET Control Theory & Applications*, *9*(16), 2458–2465.

Walsh, G. C., & Ye, H. (2002). Scheduling of networked control systems. *IEEE Control Systems Magazine*, *21*(1), 57–65.

Xie, L. (2008). Stability of a random riccati equation with markovian binary switching. *IEEE Transactions on Automatic Control*, *53*(7), 1759–1764.

Xie, L. (2012). Stochastic comparison, boundedness, weak convergence, and ergodicity of a random riccati equation with markovian binary switching. *SIAM Journal on Control and Optimization*, *50*(1), 532–558.

Xie, Li, & Xie, Lihua (2007). Peak covariance stability of a random riccati equation arising from kalman filtering with observation losses. *Journal of Systems Science and Complexity*, *20*(2), 262–272.

You, K., Fu, M., & Xie, L. (2011). Mean square stability for kalman filtering with markovian packet losses. *Automatica*, *47*(12), 2647–2657.

You, K., & Xie, L. (2013). Kalman filtering with scheduled measurements. *IEEE Transactions on Signal Processing*, *61*(6), 1520–1530.

You, K., Xie, L., & Song, S. (2013). Asymptotically optimal parameter estimation with scheduled measurements. *IEEE Transactions on Signal Processing*, *61*(14), 3521–3531.

Damián Marelli received his Bachelors Degree in Electronics Engineering from the Universidad Nacional de Rosario, Argentina in 1995. He also received his Bachelor (Honous) degree in Mathematics and Ph.D. degree in Electrical Engineering, both from the University of Newcastle, Australia in 2003. From 2004 to 2005 he was postdoc at the Laboratoire d'Analyse Topologie et Probabilités, CNRS / Université de Provence, France. From 2005 to 2015 he was Research Academic at the Centre for Complex Dynamic Systems and Control, the University of Newcastle, Australia. In 2007 he received a Marie Curie Postdoctoral Fellowship, hosted at the University of Vienna, and in 2011 he received a Lise Meitner Senior Fellowship, hosted at the Austrian Academy of Sciences. Since 2016, he is Professor at the School of Automation, Guangdong University of Technology, China and Independent Researcher appointment at the French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Argentina. His main research interests include system theory, statistical signal processing and distributed processing.

Tianju Sui received B.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China. in 2012 and 2017, respectively. He is currently working in Dalian University of Technology. His main research area includes networked estimation, distributed estimation and the security of Cyber–Physical systems.

Eduardo Rath 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 at ABB in Switzerland.

Minyue Fu received his Bachelor's Degree in Electrical Engineering from the University of Science and Technology of China, Hefei, China, in 1982, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Wisconsin-Madison in 1983 and 1987, respectively. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, Michigan. For the summer of 1989, he was employed by the Universite Catholoque de Louvain, Belgium, as a Maitre de Conferences Invited. He joined the Department of Electrical and Computer Engineering, the University of Newcastle, Australia, in 1989 and was promoted to a Chair Professor in Electrical Engineering in 2002. He has served as the Head of Department for Electrical and Computer Engineering and Head of School of Electrical Engineering and Computer Science over a period of 7 years. In addition, he was a Visiting Associate Professor at University of Iowa in 1995–1996, a Visiting Professor

at Nanyang Technological University, Singapore, 2002, and Visiting Professor at Tokyo University in 2003. He has held a ChangJiang Visiting Professorship at Shandong University, a visiting Professorship at South China University of Technology, and a Qian-ren Professorship at Zhejiang University in China. He was elected to Fellow of IEEE in late 2003. His main research interests include control systems, signal processing and communications. His current research projects include networked control systems, multi-agent systems, smart electricity networks and super-precision positioning control systems. He has been an Associate Editor for the IEEE Transactions on Automatic Control, IEEE Transactions on Signal Processing, Automatica, and Journal of Optimization and Engineering.