Error exponents for Neyman-Pearson detection of a continuous-time Gaussian Markov process from regular or irregular samples

W. Hachem, E. Moulines and F. Roueff

Abstract

This paper addresses the detection of a stochastic process in noise from a finite sample under various sampling schemes. We consider two hypotheses. The *noise only* hypothesis amounts to model the observations as a sample of a i.i.d. Gaussian random variables (noise only). The *signal plus noise* hypothesis models the observations as the samples of a continuous time stationary Gaussian process (the signal) taken at known but random time-instants and corrupted with an additive noise. Two binary tests are considered, depending on which assumptions is retained as the null hypothesis. Assuming that the signal is a linear combination of the solution of a multidimensional stochastic differential equation (SDE), it is shown that the minimum Type II error probability decreases exponentially in the number of samples when the False Alarm probability is fixed. This behavior is described by *error exponents* that are completely characterized. It turns out that they are related to the asymptotic behavior of the Kalman Filter in random stationary environment, which is studied in this paper. Finally, numerical illustrations of our claims are provided in the context of sensor networks.

Index Terms

Error exponents, Kalman filter, Gaussian Markov processes, Neyman-Pearson detection, Stein's Lemma, Stochastic Differential Equations.

The authors are with Institut Telecom / Telecom ParisTech / CNRS LTCI, France.
e-mails: walid.hachem, eric.moulines, francois.roueff@telecom-paristech.fr

I. Introduction

The detection of stochastic processes in noise has received a lot of attention in the past decades, see for instance the tutorial paper [1] and the references therein. More recently, in the context of sensor networks, there has been a rising interest in the analysis of detection performance when the stochastic process is sampled irregularly. An interesting approach in this direction has been initiated in [2] using *error exponents* for assessing the performance of the optimal detection procedure. In this paper, we follow this approach in the following general setting. Given two integers p and q, and A a positive stable square matrix, we consider the q-dimensional stochastic process defined as the stationary solution of the stochastic differential equation (SDE)

$$\mathrm{d}X(t) = -A\,X(t)\,\mathrm{d}t + B\,\mathrm{d}W(t), \quad t \geq 0 \tag{1}$$

where $(W(t),\, t \geq 0)$ is a p-dimensional Brownian motion and B is a $q \times p$ matrix. The SDE (1) is widely used to describe continuous time signals (see [3], [4] and the references therein). We are interested in the detection of the signal $(X(t),\, t \geq 0)$ from a finite sample obtained from a general spacing model. Let C be a $d \times q$ matrix, $(T_n, n \geq 1)$ be a renewal sampling process and $(V_n, n \geq 1)$ be a noise sequence. Based on the observed samples $Y_{1:N} = (Y_1, \ldots, Y_N)$ and $T_{1:N} = (T_1, \ldots, T_N)$, our goal is to decide whether for $n = 1, \ldots, N$, $Y_n = V_n$ or $Y_n = CX(T_n) + V_n$. The first situation will be referred to as the *noise hypothesis* and the second as the *signal plus noise hypothesis*.

The renewal hypothesis on $(T_n,\, n \geq 1)$ means that $T_n = \sum_{k=1}^{n} I_k$ where the $(I_k, k \geq 1)$ are nonnegative i.i.d. random variables called *holding times* with common distribution denoted by τ. This is a standard model for possibly irregular sampling, see [2], [5], [6]. The most usual examples are:

- The Poisson point process. In this case, the $(I_k,\, k \geq 1)$ are i.i.d. with exponential distribution $\tau(\mathrm{d}x) = \lambda \exp(-\lambda x)\,\mathrm{d}x$. This model has been considered in [2], [5], [6] to model the situation where the signal is measured in time by N identical asynchronous sensors.
- The Bernoulli process. This is the discrete time counterpart of the Poisson process. In this case, the $(I_k,\, k \geq 1)$ are i.i.d. and have geometric distribution up to a multiplicative time constant $S > 0$, *i.e.* $\tau(\{S\ell\}) = \tau(\{S\})(1 - \tau(\{S\}))^{\ell-1}$ for $\ell = 1, 2, \ldots$ In practice, this model corresponds to a regular sampling with period S for which observations are missing at random, with failure probability $1 - \tau(\{S\})$. The regular sampling process corresponds to $\tau(\{S\}) = 1$.

Two binary tests are considered in this work: either $\mathsf{H0}$ is the noise hypothesis and $\mathsf{H1}$ is the signal plus noise hypothesis, or the opposite. Constraining the False Alarm probability (probability for deciding $\mathsf{H1}$ under $\mathsf{H0}$) to lie beneath an $\varepsilon \in (0, 1)$, it is well known that the minimum Type II error probability is achieved by the Neyman-Pearson test. It will be shown in this paper that this minimum Type II error probability $\beta_N(\varepsilon)$ satisfies $\beta_N(\varepsilon) = \exp(-N(\xi + o(1)))$ as $N \to \infty$, where the error exponent ξ does not depend on ε. The error exponent ξ is an indicator of the performance of the detection test. Its value will be shown to depend on the distribution of signal (given by A, B and C) and on the distribution of the sampling process (given by τ). An important goal in sensor network design is to optimize the sampling process. Characterizing the error exponents offers useful guidelines in this direction. For instance, [2], [7]–[10] provide useful insights on such concrete problems as the choice of the optimum mean sensor spacing possibly subject to a cost or a power constraint. Other application examples are considered in [11], [12]. In these contributions, error exponents are used to propose optimum routing strategies for conveying the sensors data to the fusion center.

In the context of Neyman-Pearson detection, these error exponents are given by the limits of the likelihood ratios, provided that these limits exist. Let $Z_{1:N} = (Z_1, \ldots, Z_N)$ be a sequence of N observed random vectors. Assume a binary test is performed on this sequence, and assume that under hypothesis $\mathsf{H0}$, the distribution of $Z_{1:N}$ has the density $f_{0,N}$, while under $\mathsf{H1}$, this distribution has the density $f_{1,N}$. Fix $\varepsilon \in (0,1)$ and let $\beta_N(\varepsilon)$ be the minimum over all tests of the Type II error probability when the False Alarm probability α is constrained to satisfy $\alpha \leq \varepsilon$. Let

$$\mathcal{L}_N(Z_{1:N}) = \frac{1}{N} \log \left(\frac{f_{0,N}(Z_{1:N})}{f_{1,N}(Z_{1:N})} \right)$$

be the normalized Log Likelihood Ratio (LLR) associated with the received $Z_{1:N}$. Then the asymptotic behavior of $\beta_N(\varepsilon)$ can be obtained from the following theorem, found in [13].

Theorem 1 *Assume there is a real number ξ such that the random variable $\mathcal{L}_N(Z_{1:N})$ satisfies*

$$\mathcal{L}_N(Z_{1:N}) \xrightarrow[N\to\infty]{} \xi \quad \textit{in probability under } \mathsf{H0}. \tag{2}$$

Then for every $\varepsilon \in (0,1)$,

$$-\frac{1}{N} \log \beta_N(\varepsilon) \xrightarrow[N\to\infty]{} \xi \ .$$

In the case where the Z_i's are i.i.d. under both hypotheses, the analogue of Theorem 1 appeared in [14] and is known as Stein's lemma. The generalization to Theorem 1 can be found in [13], [15]. In our case, the observed process is $Z_{1:N} = (Z_1, \ldots, Z_N)$ with $Z_n = (Y_n, T_n)$, in other words, the measurements consist in the sampled received signal and the sampling moments. Let us consider that $Y_{1:N} = V_{1:N}$ under $\mathsf{H0}$ and $Y_{1:N} = (CX(T_n) + V_n)_{1 \leq n \leq N}$ under $\mathsf{H1}$. Recall that the probability distribution of $T_{1:N}$

does not depend on the hypothesis to be tested. In these conditions, the LLR is given by

$$\mathcal{L}_N(Z_{1:N}) = \frac{1}{N} \log \left(\frac{f_{0,N}(Y_{1:N} \,|\, T_{1:N})}{f_{1,N}(Y_{1:N} \,|\, T_{1:N})} \right) \tag{3}$$

where $f_{0,N}(.\,|\,T_{1:N})$ and $f_{1,N}(.\,|\,T_{1:N})$ are the densities of $Y_{1:N}$ conditionally to $T_{1:N}$ under $\mathsf{H0}$ and $\mathsf{H1}$ respectively. It is clear that $f_{0,N}(.\,|\,T_{1:N}) = \mathcal{N}(0,1)$. Being solution of the SDE (1), the process $(X(t),\, t \geq 0)$ is a Gaussian process. In consequence, $f_{1,N}(.\,|\,T_{1:N}) = \mathcal{N}(0, R(T_{1:N}))$ where matrix $R(T_{1:N})$ is a covariance matrix that depends on $T_{1:N}$. In the light of Theorem 1 we need to establish the convergence in probability of the Right Hand Side (RHS) of Eq. (3) towards a constant ξ, and to characterize this constant, under the assumption $Y_{1:N} = V_{1:N}$. Alternatively, if we consider that $\mathsf{H0}$ is the Signal plus Noise hypothesis $Y_{1:N} = (CX(T_n) + V_n)_{1 \leq n \leq N}$, then we study the convergence of $-\mathcal{L}_N$ under this assumption.

Theorem 1 has been used for detection performance analysis in [2], [7], [16], [17]. In the closely related Bayesian framework, error exponents have been obtained in [8]–[10], [18]. The closest contributions to this paper are [2], [7], [17] which consider different covariance structure for the process and different sensors locations models. In [7], Sung, Tong and Poor consider the scalar version of the SDE (1) and a regular sampling. In [17], the authors essentially generalize the results of [7] to situations where the sensor locations follow some deterministic periodic patterns. In [2], the sampling process (sensor locations) is a renewal process as in our paper, and the detector discriminates among two scalar diffusion processes described by Eq. (1). Moreover, the observations are noiseless. Here, due to the presence of additive noise, our technique for establishing the existence of the error exponents and for characterizing them differ substantially from [2]. We establish the convergence of the LLR $\mathcal{L}_N(Z_{1:N})$ by studying the stability (and ergodicity) of the Kalman filter, using Markov chains techniques.

The paper is organized as follows. In Section II, the main assumptions and notations are introduced and the main results of the paper are stated. Proofs of these results are presented in Section III. Some particular cases of interest are presented in Section IV. Section V is devoted to numerical illustrations and to a discussion about the impact of the sampling scheme on the detection performance. The proofs in Section III rely heavily on a theorem for Markov chains stability shown in appendix B. The other appendices contain technical results needed in the proofs.

II. THE ERROR EXPONENTS

We consider the following hypothesis test based on observations (Y_n, T_n), $n = 1, \ldots, N$, that we shall call the “H0-Noise” test:

$$\begin{aligned} \mathsf{H0} &: Y_n = V_n \quad \text{for } n = 1, \ldots, N \\ \mathsf{H1} &: Y_n = CX(T_n) + V_n \quad \text{for } n = 1, \ldots, N \end{aligned} \tag{4}$$

where C is the $d \times q$ observation matrix and the involved processes satisfy the following set of conditions.

Assumption 1 *The following assertions hold.*

(i) The process $(X(t),\, t \geq 0)$ *is a stationary solution of the stochastic differential equation* (1) *where* $(W(t),\, t \geq 0)$ *is a* p*-dimensional Brownian motion.*

(ii) $(T_n,\, n \geq 1)$ *is a renewal process, that is* $T_n = \sum_1^n I_k$ *where* $(I_n,\, n \geq 1)$ *is a sequence of i.i.d. non-negative r.v.'s with distribution* τ *and* $\tau(\{0\}) < 1$.

(iii) $(V_n,\, n \geq 1)$ *is a sequence of i.i.d. r.v.'s with* $V_1 \sim \mathcal{N}(0, 1_d)$.

(iv) The processes $(X(t),\, t \geq 0)$, $(T_n,\, n \geq 1)$ *and* $(V_n,\, n \geq 1)$ *are independent.*

Here 1_d denotes the $d \times d$ identity matrix.

In order to be able to apply Theorem 1, we now develop the expression of the LLR given by (3). To that end, we derive the expressions of the likelihood functions $f_{0,N}(Y_{1:N} \,|\, T_{1:N})$ and $f_{1,N}(Y_{1:N} \,|\, T_{1:N})$. The density $f_{0,N}(.\,|\, T_{1:N})$ is simply the density $\mathcal{N}(0, 1_{Nd})$ of $(V_1, \ldots, V_N)$, therefore

$$f_{0,N}(Y_{1:N} \,|\, T_{1:N}) = \frac{1}{(2\pi)^{Nd/2}} \exp\left(-\frac{1}{2} \sum_{n=1}^{N} Y_n^{\mathrm{T}} Y_n \right) . \tag{5}$$

We now develop $f_{1,N}(.\,|\, T_{1:N})$ by mimicking the approach developed in [19] and in [7]. Let $Q(x)$ be the $q \times q$ symmetric nonnegative matrix defined by

$$Q(x) = \int_0^x \mathrm{e}^{-uA} B B^{\mathrm{T}} \mathrm{e}^{-uA^{\mathrm{T}}} \mathrm{d}u . \tag{6}$$

As A is positive stable, the covariance matrix $Q(\infty)$ exists (by Lemma 3) and is the unique solution of the so called *Lyapunov's equation* $QA^{\mathrm{T}} + AQ = BB^{\mathrm{T}}$, see [20, Chap. 2]. Solving Eq. (1) between T_n and T_{n+1} (see [3, Chap. 5]), the conditional distribution of the process $(X_n,\, n \geq 1)$ given $(I_n,\, n \geq 1)$ is characterized by the recursion equation

$$X_{n+1} = \mathrm{e}^{-I_{n+1}A} X_n + U_{n+1}, \quad n \geq 0 , \tag{7}$$

where the conditional distribution of the sequence $(X_0, U_n,\, n \geq 1)$ is that of a sequence of independent r.v.'s, $X_0 \sim \mathcal{N}(0, Q(\infty))$ and $U_n \sim \mathcal{N}(0, Q_n)$ with $Q_n = Q(I_n)$ defined by (6).

Now we write

$$f_{1,N}(Y_{1:N}\,|\,T_{1:N}) = \prod_{n=1}^{N} f_{1,n,N}\left(Y_n\,|\,Y_{1:n-1},T_{1:N}\right) \tag{8}$$

where $f_{1,n,N}(.\,|\,Y_{1:n-1},T_{1:N})$ is the density of Y_n conditionally to $(Y_{1:n-1},T_{1:N})$. In view of Eq. (7), $Y_n = CX_n + V_n$ and the assumptions on (V_n), these conditional densities are Gaussian, in other words

$$f_{1,n,N}\left(Y_n\,|\,Y_{1:n-1},T_{1:N}\right) = \frac{1}{\det(2\pi\Delta_n)^{1/2}}\exp\left(-\frac{1}{2}(Y_n-\widehat{Y}_n)^{\mathrm{T}}\Delta_n^{-1}(Y_n-\widehat{Y}_n)\right) \tag{9}$$

where $\widehat{Y}_n = \mathbb{E}\left[Y_n\,|\,Y_{1:n-1},T_{1:N}\right]$ and $\Delta_n = \mathrm{Cov}\left(Y_n-\widehat{Y}_n\,|\,T_{1:N}\right)$ are respectively the conditional expectation of the current observation Y_n given the past observations and the so-called innovation covariance matrix under H1. From Equations (5), (8) and (9), the LLR $\mathcal{L}_N$ writes

$$\begin{aligned}\mathcal{L}_N(Y_{1:N},T_{1:N}) &= \frac{1}{N}\log f_{0,N}(Y_{1:N}\,|\,T_{1:N}) - \frac{1}{N}\log f_{1,N}(Y_{1:N}\,|\,T_{1:N})\\ &= \frac{1}{2N}\sum_{n=1}^{N}\log\det\Delta_n + \frac{1}{2N}\sum_{n=1}^{N}(Y_n-\widehat{Y}_n)^T\Delta_n^{-1}(Y_n-\widehat{Y}_n) - \frac{1}{2N}\sum_{n=1}^{N}Y_n^{\mathrm{T}}Y_n\end{aligned} \tag{10}$$

As (Y_n) is described under H1 by the state equations

$$\mathsf{H1}:\begin{cases}X_{n+1} = \mathrm{e}^{-I_{n+1}A}X_n + U_{n+1}\\ Y_n = CX_n + V_n\end{cases}\quad\text{for } n=1,\dots,N, \tag{11}$$

it is well known that $\widehat{Y}_n$ and Δ_n can be computed using the Kalman filter recursive equations. Define the $q\times 1$ vector $\widehat{X}_n$ and the $q\times q$ matrix P_n as

$$\widehat{X}_n = \mathbb{E}[X_n\,|\,Y_{1:n-1},T_{1:N}] \quad\text{and}\quad P_n = \mathrm{Cov}\left(X_n-\widehat{X}_n\,|\,T_{1:N}\right)\ .$$

The Kalman recursions which provide these quantities are [21, Prop. 12.2.2]:

$$\widehat{X}_{n+1} = \mathrm{e}^{-I_{n+1}A}\left(1_q - P_nC^{\mathrm{T}}\left(CP_nC^{\mathrm{T}}+1_d\right)^{-1}C\right)\widehat{X}_n + \mathrm{e}^{-I_{n+1}A}P_nC^{\mathrm{T}}\left(CP_nC^{\mathrm{T}}+1_d\right)^{-1}Y_n \tag{12}$$

$$P_{n+1} = \mathrm{e}^{-I_{n+1}A}\left(1_q - P_nC^{\mathrm{T}}\left(CP_nC^{\mathrm{T}}+1_d\right)^{-1}C\right)P_n\mathrm{e}^{-I_{n+1}A^{\mathrm{T}}} + Q_{n+1}\ . \tag{13}$$

The recursion is started with the initial conditions $\widehat{X}_1 = 0$ and $P_1 = Q(\infty)$. With these quantities at hand, $\widehat{Y}_n$ and Δ_n are given by

$$\widehat{Y}_n = C\widehat{X}_n \quad\text{and}\quad \Delta_n = CP_nC^{\mathrm{T}} + 1_d\ . \tag{14}$$

With these expressions at hand, our purpose is to study the asymptotic behavior of $\mathcal{L}_N$ given by Eq. (10) assuming that that (Y_n) is i.i.d. with $Y_1\sim\mathcal{N}(0,1_d)$ (under H0).

In our analysis, we shall require Model (1) to be *controllable*, *i.e.*, (A,B) satisfies $\sum_{\ell=0}^{q-1}A^\ell BB^{\mathrm{T}}A^{\mathrm{T}^\ell} > 0$. Recall that (A,B) is controllable if and only if the matrix $Q(x)$ defined by Equation (6) is nonsingular for any $x>0$ (see [22, Chap. 6] for a proof).

Since under H0, $\eta_n = (I_{n+1}, Y_n)$, $n \geq 1$, are i.i.d. r.v.'s, the Kalman equations (12)-(13) can be written as a random iteration scheme. To this end we introduce some notation. Let $\mathcal{P}_q$ denote the cone of $q \times q$ nonnegative symmetric matrices. For any $\eta = (I, Y) \in [0, \infty) \times \mathbb{R}^d$, we denote by F_η the $\mathbb{R}^q \times \mathcal{P}_q$-valued function defined for all $\mathbf{w} = (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^q \times \mathcal{P}_q$ by

$$F_\eta(\mathbf{w}) := \begin{bmatrix} \mathrm{e}^{-IA}\left(1_q - \mathbf{p}C^{\mathrm{T}}\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right)^{-1}C\right)\mathbf{x} + \mathrm{e}^{-IA}\mathbf{p}C^{\mathrm{T}}\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right)^{-1}Y \\ \mathrm{e}^{-IA}\left(1_q - \mathbf{p}C^{\mathrm{T}}\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right)^{-1}C\right)\mathbf{p}\mathrm{e}^{-IA^{\mathrm{T}}} + Q(I) \end{bmatrix} . \tag{15}$$

Using this notation the Kalman equations read

$$W_n = F_{\eta_n}(W_{n-1}), \quad n \geq 1 ,$$

where $W_n = (\widehat{X}_{n+1}, P_{n+1})$. Under H0, since (η_n) is a sequence of i.i.d. random variables, $(W_n)_{n\geq 0}$ is a Markov chain starting at $W_0 = (0, Q(\infty))$. Observe also that since the second component of F_η, denoted by $\tilde{F}_I(\mathbf{p})$ in the following, does not depend on $\mathbf{x}$, $(P_n)_{n\geq 1}$ also is a Markov chain starting at $P_1 = Q(\infty)$. Since it neither depends on Y, this is also true under H1 as well. Let $[0, Q(\infty)]$ denote the subset of all matrices $\mathbf{p} \in \mathcal{P}_q$ such that $\mathbf{p} \leq Q(\infty)$. It is easy to see that, for any $I \geq 0$, $[0, Q(\infty)]$ is a stable set for $\tilde{F}_I$. Indeed, suppose that $\mathbf{p} \in [0, Q(\infty)]$, then

$$\begin{aligned} \tilde{F}_I(\mathbf{p}) &= \mathrm{e}^{-IA}\left(1_q - \mathbf{p}C^{\mathrm{T}}\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right)^{-1}C\right)\mathbf{p}\mathrm{e}^{-IA^{\mathrm{T}}} + Q(I) \\ &\leq \mathrm{e}^{-IA}\mathbf{p}\mathrm{e}^{-IA^{\mathrm{T}}} + Q(I) \\ &\leq \mathrm{e}^{-IA}Q(\infty)\mathrm{e}^{-IA^{\mathrm{T}}} + Q(I) = Q(\infty) , \end{aligned}$$

by definition of Q in (6). Hence, in the following, we consider (W_n) and (P_n) as chains valued in $\mathbb{R}^q \times [0, Q(\infty)]$ and $[0, Q(\infty)]$, respectively. We will denote by Π and $\tilde{\Pi}$ the transition kernels associated to the chains (W_n) (under H0) and (P_n), respectively, that is, for test functions f and $\tilde{f}$ defined on $\mathbb{R}^q \times [0, Q(\infty)]$ and $[0, Q(\infty)]$,

$$\begin{aligned} \Pi f(\mathbf{w}) &= \mathbb{E}[f(F_\eta(\mathbf{w}))], \quad \mathbf{w} \in \mathbb{R}^q \times [0, Q(\infty)] \\ \tilde{\Pi}\tilde{f}(\mathbf{p}) &= \mathbb{E}[\tilde{f}(\tilde{F}_I(\mathbf{p}))], \quad \mathbf{p} \in [0, Q(\infty)] , \end{aligned}$$

where $\eta = (I, Y)$ is such that $I \sim \tau$, $Y \sim \mathcal{N}(0, 1_d)$ and I and Y are independent. In the following we will simply use the notation $\eta \sim \tau \otimes \mathcal{N}(0, 1_d)$. We now state our main results.

First we determine the limit of the LLR under the signal hypothesis H1.

Theorem 2 *Suppose that Assumption 1 holds with a state realization (A, B, C) such that A is positive stable and (A, B) is controllable. Then the transition kernel $\tilde{\Pi}$ has a unique invariant distribution μ.*

Moreover, if $Y_{1:N}$ *is defined as in* H1 *in (4), then as* $N \to \infty$*, we have*

$$-\mathcal{L}_N(Y_{1:N}, T_{1:N}) \to \xi_{\textit{H0:Signal}} \quad \textit{almost surely } (\text{a.s.}) ,$$

where $\mathcal{L}_N(Y_{1:N}, T_{1:N})$ *is defined in (10) and*

$$\xi_{\textit{H0:Signal}} = \frac{1}{2}\left(\operatorname{tr}\left(CQ(\infty)C^{\mathrm{T}}\right) - \int \log\det\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right) \mathrm{d}\mu(\mathbf{p})\right) \tag{16}$$

is positive and finite.

Remark 1 *This paper deals with the detection of a stationary signal. That is why the matrix* A *is assumed to be positive stable. An interesting problem, however, consists in searching for minimum conditions on* τ *that guarantee the existence of* μ *when* A *is not positive stable. This problem could also be refined by studying the existence of some moments of* μ*. Recently, this study has been undertaken in [23], [24] and [25] in the case where the sampling process is a Bernoulli process. The approach of [23] and [24] is based on the so called random dynamical systems theory.*

Let us now provide the limit of the LLR under the noise hypothesis H0.

Theorem 3 *Suppose that Assumption 1 holds with a state realization* (A, B, C) *such that* A *is positive stable and* (A, B) *is controllable. Then the transition kernel* Π *has a unique invariant distribution* ν*. Moreover, if* $Y_{1:N}$ *is defined as in* H0 *in (4), then as* $N \to \infty$*, we have*

$$\mathcal{L}_N(Y_{1:N}, T_{1:N}) \to \xi_{\textit{H0:Noise}} \quad \text{a.s.} ,$$

where $\mathcal{L}_N(Y_{1:N}, T_{1:N})$ *is defined in (10) and*

$$\xi_{\textit{H0:Noise}} = \frac{1}{2}\int \left\{\log\det\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right) + \operatorname{tr}\left[C(\mathbf{x}\mathbf{x}^{\mathrm{T}} - \mathbf{p})C^{\mathrm{T}}\left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right)^{-1}\right]\right\} \mathrm{d}\nu(\mathbf{x}, \mathbf{p}) \tag{17}$$

is positive and finite.

Remark 2 *From Theorems 2 and 3, and by definitions of* Π *and* $\tilde{\Pi}$*, we immediately see that* μ *is the second marginal distribution of* ν*,* $\mu(\cdot) = \nu(\mathbb{R}^q \times \cdot)$*.*

By Theorems 1, 2 and 3, we get that the error exponents associated to the hypotheses testing problem (4) are given by $\xi_{\text{H0:Noise}}$ and $\xi_{\text{H0:Signal}}$. More precisely, we have the following result.

Corollary 1 *Consider, under Assumption 1, the hypotheses test* (4)*. For* $N \geq 1$ *and* $\varepsilon \in (0, 1)$*, let* $\beta_N(\varepsilon)$ *be the minimum of error probabilities over all tests for which the false alarm probability is at most* ε*. Then, as* $N \to \infty$*,* $N^{-1}\log\beta_N(\varepsilon) \to \xi_{\textit{H0:Noise}}$*, as defined in (17).*

Now, if we interchange the roles of H0 *and* H1 *in* (4) *(call this test the "H0-Signal" test), we obtain* $N^{-1}\log\beta_N(\varepsilon) \to \xi_{\textit{H0:Signal}}$*, as defined in (16).*

III. Proofs of main results

This section is devoted to the proofs of Theorems 2 and 3. These results follow from an analysis of the Markov chains induced by the transition kernels Π and $\tilde{\Pi}$, or, equivalently, of the random iteration functions F_η and $\tilde{F}_I$ defined in (15). We provide a fairly general result in the appendix, Theorem 4, to deal with this general framework. Based on moment contraction conditions, this latter result establishes the existence and uniqueness of the invariant distribution and a law of large numbers for functions with precise polynomial growth conditions at infinity. In this section we establish some useful preliminary results related to moment contraction conditions for random iteration functions F_η and $\tilde{F}_I$, and then prove Theorems 2 and 3 by applying Theorem 4.

A. Preliminary results

We start with a series of preliminary results for which we recall the following notations and assumptions: τ is a distribution on $[0,\infty)$ such that $\tau(\{0\}) < 1$, $(I_n)_{n\geq 1}$ is a sequence of i.i.d. r.v.'s distributed according to the distribution τ and $(\eta_n)_{n\geq 1}$ is a sequence of i.i.d. r.v.'s distributed according to the distribution $\tau \otimes \mathcal{N}(0, 1_d)$.

We further denote by I and η two generic r.v.'s having same distribution as I_1 and η_1, respectively. For any $\mathbf{x} \in \mathbb{R}^q$ and $\mathbf{p} \in [0, Q(\infty)]$, we define two Markov chains induced by Π and $\tilde{\Pi}$ and starting at $\mathbf{w} = (\mathbf{x}, \mathbf{p})$ and $\mathbf{p}$, respectively

$$\begin{cases} Z_0^{\mathbf{w}} = \mathbf{w} \quad \text{and} \quad \tilde{Z}_0^{\mathbf{p}} = \mathbf{p} \\ Z_k^{\mathbf{w}} = F_{\eta_k}(Z_{k-1}^{\mathbf{w}}) \quad \text{and} \quad \tilde{Z}_k^{\mathbf{p}} = \tilde{F}_{I_k}(\tilde{Z}_{k-1}^{\mathbf{p}}), \quad k \geq 1 \ . \end{cases}$$

As noticed earlier, $\tilde{Z}_k^{\mathbf{p}}$ corresponds to the second component of $Z_k^{\mathbf{w}}$ for each k and is valued in $[0, Q(\infty)]$. We introduce the following notation for the Kalman gain matrix

$$G(\mathbf{p}) = \mathbf{p}C^{\mathrm{T}}(1_d + C\mathbf{p}C^{\mathrm{T}})^{-1} \ ,$$

and the short-hand notation for $G(\tilde{Z}_k^{\mathbf{p}})$ (the Kalman gain matrix at time k):

$$G_k^{\mathbf{p}} = \tilde{Z}_k^{\mathbf{p}}C^{\mathrm{T}}(1_d + C\tilde{Z}_k^{\mathbf{p}}C^{\mathrm{T}})^{-1}, \quad k \geq 0 \ , \tag{18}$$

As for the Kalman transition matrix, we set

$$\Theta(I, \mathbf{p}) = \mathrm{e}^{-IA}(1_q - G(\mathbf{p})C) \ ,$$

and the short-hand notation for $\Theta(I_k, \tilde{Z}^{\mathbf{p}}_{k-1})$ (the Kalman transition matrix at time k):

$$\Theta^{\mathbf{p}}_k = \mathrm{e}^{-I_k A}(1_q - G^{\mathbf{p}}_{k-1}C), \quad n \geq 1 \; . \tag{19}$$

Using this notation and $Q_k = Q(I_k)$, the Kalman covariance update equation $\tilde{Z}^{\mathbf{p}}_k = \tilde{F}_{I_k}(\tilde{Z}^{\mathbf{p}}_{k-1})$ can be expressed for all $k \geq 1$ as

$$\begin{aligned} \tilde{Z}^{\mathbf{p}}_k &= \Theta^{\mathbf{p}}_k \tilde{Z}^{\mathbf{p}}_{k-1} \mathrm{e}^{-I_k A^{\mathrm{T}}} + Q_k \\ &= \Theta^{\mathbf{p}}_k \tilde{Z}^{\mathbf{p}}_{k-1} \Theta^{\mathbf{p}\,\mathrm{T}}_k + \mathrm{e}^{-I_k A}(1_q - G^{\mathbf{p}}_{k-1}C)\tilde{Z}^{\mathbf{p}}_{k-1} C^{\mathrm{T}} G^{\mathbf{p}\,\mathrm{T}}_{k-1} \mathrm{e}^{-I_k A^{\mathrm{T}}} + Q_k && (20) \\ &= \Theta^{\mathbf{p}}_k \tilde{Z}^{\mathbf{p}}_{k-1} \Theta^{\mathbf{p}\,\mathrm{T}}_k + \overline{Q}_k \; , && (21) \end{aligned}$$

where

$$\overline{Q}_k = \mathrm{e}^{-I_k A} G^{\mathbf{p}}_{k-1} G^{\mathbf{p}\,\mathrm{T}}_{k-1} \mathrm{e}^{-I_k A^{\mathrm{T}}} + Q_k, \quad k \geq 1 \; . \tag{22}$$

We also denote a product of successive Kalman transition matrices by

$$\Theta^{\mathbf{p}}_{n,m} = \Theta^{\mathbf{p}}_n \Theta^{\mathbf{p}}_{n-1} \ldots \Theta^{\mathbf{p}}_{m+1} \; , \quad 0 \leq m < n \; . \tag{23}$$

Note that $\Theta^{\mathbf{p}}_{n,n-1} = \Theta^{\mathbf{p}}_n$. If $m = n$, we will use the convention $\Theta^{\mathbf{p}}_{n,n} = 1_q$.
Finally, we denote by $|x|$ the Euclidean norm of the vector x, $\lambda_{\min}(H)$ and $\lambda_{\max}(H)$ the minimum and maximum eigenvalues of the matrix H and by $\|H\|$ its operator norm, $\|H\| = \lambda_{\max}(H^{\mathrm{T}}H)^{1/2}$.

We first derive a deterministic bound for $\Theta^{\mathbf{p}}_{n,m}$ based on (21), which relies on a Lyapunov function argument similar to that in [26, Theorem 2.4] and [27, Sec. 4].

Lemma 1 *For any $0 \leq m < n$, we have*

$$\|\Theta^{\mathbf{p}}_{n,m}\|^2 \leq \|\tilde{Z}^{\mathbf{p}}_n\| \, \|(\tilde{Z}^{\mathbf{p}}_m)^{-1}\| \prod_{k=m+1}^{n} \left(1 - \frac{\lambda_{\min}(Q_k)}{\|\tilde{Z}^{\mathbf{p}}_k\|}\right) \; , \tag{24}$$

Proof: Obviously $\overline{Q}_k \geq Q_k$, hence $\lambda_{\min}(\overline{Q}_k) \geq \lambda_{\min}(Q_k)$. Now, for a given $\mathbf{x}_n \in \mathbb{R}^q$, define the backward recursion $\mathbf{x}_k = \Theta^{\mathbf{p}\,\mathrm{T}}_{k+1}\mathbf{x}_{k+1}$ for k decreasing from $n-1$ down to m, and set $V_k = \mathbf{x}^{\mathrm{T}}_k \tilde{Z}^{\mathbf{p}}_k \mathbf{x}_k$ for $k = m, \ldots, n$. We have

$$V_n - V_{n-1} = \mathbf{x}^{\mathrm{T}}_n \tilde{Z}^{\mathbf{p}}_n \mathbf{x}_n - \mathbf{x}^{\mathrm{T}}_n \Theta^{\mathbf{p}}_n \tilde{Z}^{\mathbf{p}}_{n-1} \Theta^{\mathbf{p}\,\mathrm{T}}_n \mathbf{x}_n = \mathbf{x}^{\mathrm{T}}_n \overline{Q}_n \mathbf{x}_n$$

by (21), and moreover, $\mathbf{x}^{\mathrm{T}}_n \overline{Q}_n \mathbf{x}_n \geq |\mathbf{x}_n|^2 \lambda_{\min}(\overline{Q}_n) \geq |\mathbf{x}_n|^2 \lambda_{\min}(Q_n) \geq V_n \lambda_{\min}(Q_n)/\|\tilde{Z}^{\mathbf{p}}_n\|$. Hence, $V_{n-1} \leq V_n \left(1 - \lambda_{\min}(Q_n)/\|\tilde{Z}^{\mathbf{p}}_n\|\right)$. Iterating, we obtain

$$V_m \leq V_n \prod_{k=m+1}^{n} \left(1 - \frac{\lambda_{\min}(Q_k)}{\|\tilde{Z}^{\mathbf{p}}_k\|}\right) \leq |\mathbf{x}_n|^2 \|\tilde{Z}^{\mathbf{p}}_n\| \prod_{k=m+1}^{n} \left(1 - \frac{\lambda_{\min}(Q_k)}{\|\tilde{Z}^{\mathbf{p}}_k\|}\right) \; . \tag{25}$$

On the other hand, by (23), $V_m = \mathbf{x}_m^{\mathrm{T}} \tilde{Z}_m^{\mathbf{p}} \mathbf{x}_m = \mathbf{x}_n^{\mathrm{T}} \Theta_{n,m}^{\mathbf{p}} \tilde{Z}_m^{\mathbf{p}} \Theta_{n,m}^{\mathbf{p}\,\mathrm{T}} \mathbf{x}_n$, hence $\lambda_{\min}(\tilde{Z}_m^{\mathbf{p}})\ \mathbf{x}_n^{\mathrm{T}} \Theta_{n,m}^{\mathbf{p}} \Theta_{n,m}^{\mathbf{p}\,\mathrm{T}} \mathbf{x}_n \leq V_m$. This, with Inequality (25), implies (24). ■

Next we deduce a uniform moment bound for $\Theta_{n,m}^{\mathbf{p}}$.

Lemma 2 *Assume that the matrix A is positive stable and that the pair (A, B) is controllable. For any $r > 0$, there exist $K > 0$ and $\rho \in (0, 1)$ such that*

$$\mathbb{E}\left[\sup_{\mathbf{p}\in[0,Q(\infty)]} \|\Theta_{n,m}^{\mathbf{p}}\|^{2r}\right] \leq K\,\rho^{n-m}, \quad 0 \leq m < n\ .$$

Proof: Recall that for $\mathbf{p} \in [0, Q(\infty)]$, we have $\tilde{Z}_k^{\mathbf{p}} \in [0, Q(\infty)]$ for all $k \geq 1$. Note that G is continuous and, by Lemma 3, $\sup_k \|\mathrm{e}^{-I_k A}\| < \infty$. Hence,

$$\tilde{Z}^* = \sup_{\mathbf{p}\in[0,Q(\infty)]} \sup_{k\geq 1} \|\tilde{Z}_k^{\mathbf{p}}\| < \infty \quad \text{and} \quad \Theta^* = \sup_{\mathbf{p}\in[0,Q(\infty)]} \sup_{k\geq 1} \|\Theta_k^{\mathbf{p}}\| < \infty\ .$$

Let $0 \leq m < n$. Let $\epsilon > 0$ that we will choose arbitrarily small later. Denote $T = \inf\{k \geq m \mid I_k \geq \epsilon\}$. Then we have, by (21) and (22),

$$\lambda_{\min}(\tilde{Z}_T^{\mathbf{p}}) \geq \lambda_{\min}(Q_T) \geq \lambda_{\min}(Q(\epsilon)) > 0\ , \tag{26}$$

by Lemma 4. We now write, denoting by $\mathbb{1}_A$ the indicator function of the event A,

$$\begin{aligned}\|\Theta_{n,m}^{\mathbf{p}}\|^{2r} &= \sum_{k=m}^{n-1} \|\Theta_{n,k}^{\mathbf{p}}\Theta_{k,m}^{\mathbf{p}}\|^{2r}\,\mathbb{1}_{\{T=k\}} + \|\Theta_{n,m}^{\mathbf{p}}\|^{2r}\,\mathbb{1}_{\{T\geq n\}} \\ &\leq \sum_{k=m}^{n-1} \|\Theta_{n,k}^{\mathbf{p}}\|^{2r}(\Theta^*)^{2r(k-m)}\,\mathbb{1}_{\{T=k\}} + (\Theta^*)^{2r(n-m)}\,\mathbb{1}_{\{T\geq n\}}\ .\end{aligned}$$

For any $k < n$, applying Lemma 1 and the bound (26), we have, on the event $T = k$,

$$\|\Theta_{n,k}^{\mathbf{p}}\|^{2r} \leq (\tilde{Z}^*\lambda_{\min}(Q(\epsilon))^{-1})^{2r} \prod_{j=T+1}^{n}\left(1 - \frac{\lambda_{\min}(Q_j)}{\tilde{Z}^*}\right)^{2r}\ ,$$

Observe that, for any $k \geq m$, $\{T = k\} = \{I_m < \epsilon, \ldots, I_{k-1} < \epsilon, I_k \geq \epsilon\}$. Hence $T - m + 1$ is a geometric r.v. with parameter $\tau_\epsilon = \tau([\epsilon, \infty))$ and $(Q_{T+i})_{i\geq 1}$ is i.i.d., independent of T, and follows the same distribution as $Q(I)$. Moreover, by Lemma 4, $\lambda_{\min}(Q(I)) > 0$ for $I > 0$, and since, $\tau(\{0\}) < 1$, we have $\gamma = \mathbb{E}[(1 - \lambda_{\min}(Q(I))/\tilde{Z}^*)^{2r}] < 1$. Thus we get

$$\begin{aligned}&\mathbb{E}\left[\sup_{\mathbf{p}\in[0,Q(\infty)]} \|\Theta_{n,m}^{\mathbf{p}}\|^{2r}\right] \\ &\leq (\tilde{Z}^*\lambda_{\min}(Q(\epsilon))^{-1})^{2r} \sum_{k=m}^{n-1} \gamma^{n-k}(\Theta^*)^{2r(k-m)}\tau_\epsilon(1-\tau_\epsilon)^{k-m} + (\Theta^*)^{2r(n-m)}\sum_{k\geq n}\tau_\epsilon(1-\tau_\epsilon)^{k-m} \\ &\leq \{(\tilde{Z}^*\lambda_{\min}(Q(\epsilon))^{-1})^{2r}\tau_\epsilon(n-m) + 1\}\tilde{\rho}^{n-m}\ .\end{aligned}$$

where we chose $\epsilon > 0$ small enough so that $(\Theta^*)^{2r}(1-\tau_\epsilon) < 1$ and set $\tilde{\rho} = \gamma \vee \{(\Theta^*)^{2r}(1-\tau_\epsilon)\} < 1$. This gives the result for any $\rho \in (\tilde{\rho}, 1)$ by conveniently choosing K. ■

We conclude this series of preliminary results with useful Lipschitz bounds for the mappings $\mathbf{p} \mapsto \tilde{Z}_n^{\mathbf{p}}$, $\mathbf{p} \mapsto \Theta_{n,m}^{\mathbf{p}}$ and $\mathbf{p} \mapsto G_n^{\mathbf{p}}$.

Proposition 1 *We have, for any nonnegative symmetric matrices* $\mathbf{p}$ *and* $\mathbf{q}$*,*

$$\tilde{Z}_n^{\mathbf{p}} - \tilde{Z}_n^{\mathbf{q}} = \Theta_{n,0}^{\mathbf{p}}(\mathbf{p}-\mathbf{q})\Theta_{n,0}^{\mathbf{q}\mathrm{T}}, \quad n \geq 1 \,. \tag{27}$$

Moreover, there exists a constant $C > 0$ *such that, for all* $\mathbf{p}, \mathbf{q} \in [0, Q(\infty)]$*,*

$$\|G_n^{\mathbf{p}} - G_n^{\mathbf{q}}\| \leq C\, \|\mathbf{p}-\mathbf{q}\|\, \|\Theta_{n,0}^{\mathbf{p}}\|\, \|\Theta_{n,0}^{\mathbf{q}}\| \,, \qquad n \geq 1 \,, \tag{28}$$

$$\|\Theta_{n,m}^{\mathbf{p}} - \Theta_{n,m}^{\mathbf{q}}\| \leq C\, \|\mathbf{p}-\mathbf{q}\| \sum_{j=m+1}^{n} \|\Theta_{n,j}^{\mathbf{q}}\|\, \|\Theta_{j-1,m}^{\mathbf{p}}\|\, \|\Theta_{j-1,1}^{\mathbf{q}}\|\, \|\Theta_{j-1,1}^{\mathbf{p}}\| \,, \qquad 0 \leq m < n \,. \tag{29}$$

Proof: Let us prove (27). By induction, it is sufficient to show that

$$\tilde{F}_I(\mathbf{p}) - \tilde{F}_I(\mathbf{q}) = \Theta(I, \mathbf{p})(\mathbf{p}-\mathbf{q})\Theta^{\mathrm{T}}(I, \mathbf{q}) \,. \tag{30}$$

By continuity of $\tilde{F}_I$ and $\Theta(I, \cdot)$, we may assume that $\mathbf{p}$ and $\mathbf{q}$ are invertible. In this case, the matrix inversion lemma gives that

$$(\mathbf{p} - \mathbf{p}C^{\mathrm{T}}(C\mathbf{p}C + 1_d)^{-1}C\mathbf{p}) = (\mathbf{p}^{-1} + C^{\mathrm{T}}C)^{-1} \,, \tag{31}$$

and the same is true with $\mathbf{q}$ replacing $\mathbf{p}$. Hence

$$\begin{aligned}\tilde{F}_I(\mathbf{p}) - \tilde{F}_I(\mathbf{q}) &= \mathrm{e}^{-IA}\left[(\mathbf{p}^{-1} + C^{\mathrm{T}}C)^{-1} - (\mathbf{q}^{-1} + C^{\mathrm{T}}C)^{-1}\right]\mathrm{e}^{-IA^{\mathrm{T}}} \\ &= \mathrm{e}^{-IA}(\mathbf{p}^{-1} + C^{\mathrm{T}}C)^{-1}\mathbf{p}^{-1}\left[\mathbf{p}-\mathbf{q}\right]\mathbf{q}^{-1}(\mathbf{q}^{-1} + C^{\mathrm{T}}C)^{-1}\mathrm{e}^{-IA^{\mathrm{T}}} \,.\end{aligned}$$

Using again (31) and the definition of Θ, we get (30), which achieves the proof of (27).

We now prove (28). Observe that G is continuously differentiable on the compact set $[0, Q(\infty)]$. Hence $\|G(\mathbf{p}) - G(\mathbf{q})\| \leq C\|\mathbf{p}-\mathbf{q}\|$ for some constant $C > 0$. Thus, since $G_n^{\mathbf{p}} = G(\tilde{Z}_n^{\mathbf{p}})$, the bound (28) follows from (27).

Finally we prove (29). We have, for all $0 \leq m < n$ (recall the convention $\Theta_{n,n}^{\mathbf{p}} = \Theta_{m,m}^{\mathbf{p}} = 1_q$),

$$\Theta_{n,m}^{\mathbf{p}} - \Theta_{n,m}^{\mathbf{q}} = \sum_{j=m+1}^{n} \Theta_{n,j}^{\mathbf{q}}(\Theta_j^{\mathbf{p}} - \Theta_j^{\mathbf{q}})\Theta_{j-1,m}^{\mathbf{p}} \,.$$

On the other hand, $\Theta_j^{\mathbf{p}} - \Theta_j^{\mathbf{q}} = \mathrm{e}^{-I_jA}(G_{j-1}^{\mathbf{q}} - G_{j-1}^{\mathbf{p}})C$, and (29) thus follows from (28). ■

B. *Proof of Theorem 2.*

Using (27) in Proposition 1, Lemma 2 and the Hölder inequality, we obtain that, for any $q > 0$ there exists $C > 0$ and $\alpha \in (0,1)$ such that

$$\mathbb{E}\left[\|\tilde{Z}_n^{\mathbf{p}} - \tilde{Z}_n^{\mathbf{q}}\|^q\right] \leq C\alpha^n, \quad \mathbf{p}, \mathbf{q} \in [0, Q(\infty)],\ n \geq 1 \ . \tag{32}$$

This corresponds to Condition (i) in Theorem 4. Condition (ii) is trivially satisfied for any s and $r = 1$ since here $\mathcal{X} = [0, Q(\infty)]$ is a compact state space. Hence, by Theorem 4(a) we obtain the existence and uniqueness of μ.

Next, we show that $-\mathcal{L}_N(Y_{1:N}, T_{1:N})$ defined in (10) converges to $\xi_{\text{H0:Signal}}$ in probability when $Y_n = CX_n + V_n$ for all $n \geq 1$. Since, for all $n \geq 1$, $P_n = \tilde{Z}_{n-1}^{Q(\infty)}$ and $\log\det\Delta_n$ is a Lipschitz function of P_n, we have by Theorem 4(b) that

$$\frac{1}{N}\sum_{n=1}^{N}\log\det\Delta_n \xrightarrow[N\to\infty]{a.s.} \int \log\det(C\mathbf{p}C^{\mathrm{T}} + 1_d)\ \mu(\mathrm{d}\mathbf{p}) \ . \tag{33}$$

This is true independently of the definition of (Y_n) and hence will also be used in the proof of Theorem 3. In contrast the specific definition of (Y_n) here implies that $\hat{Y}_n = \mathbb{E}[Y_n \mid Y_{1:n-1}, T_{1:N}]$ and $\Delta_n = \mathrm{Cov}(Y_n - \hat{Y}_n \mid T_{1:N})$. Hence $((Y_n - \hat{Y}_n)^{\mathrm{T}}\Delta_n^{-1}(Y_n - \hat{Y}_n))_{n\geq 1}$ is a sequence of i.i.d. $\mathcal{N}(0,1)$ r.v.'s, which yields

$$\frac{1}{N}\sum_{n=1}^{N}(Y_n - \hat{Y}_n)^{\mathrm{T}}\Delta_n^{-1}(Y_n - \hat{Y}_n) \xrightarrow[N\to\infty]{a.s.} d \ .$$

On the other hand, in (10) this limit cancels with

$$\frac{1}{N}\sum_{n=1}^{N}V_n^{\mathrm{T}}V_n \xrightarrow[N\to\infty]{a.s.} d \ , \tag{34}$$

which appears in the last term of (10) when developing $Y_n^{\mathrm{T}}Y_n = V_n^{\mathrm{T}}V_n + X_n^{\mathrm{T}}C^{\mathrm{T}}CX_n + 2X_n^{\mathrm{T}}C^{\mathrm{T}}V_n$. Hence it only remains to show that

$$\frac{1}{N}\sum_{n=1}^{N}X_n^{\mathrm{T}}C^{\mathrm{T}}CX_n \xrightarrow[N\to\infty]{a.s.} \mathrm{tr}(CQ(\infty)C^{\mathrm{T}}) \ , \tag{35}$$

$$\frac{1}{N}\sum_{n=1}^{N}X_n^{\mathrm{T}}C^{\mathrm{T}}V_n \xrightarrow[N\to\infty]{a.s.} 0 \ . \tag{36}$$

To this end, recall that (X_n) is a Markov chain, whose distribution is defined by the recurrence equation (7) and the initial condition $X_0 \sim \mathcal{N}(0, Q(\infty))$. We shall establish the ergodicity of this Markov chain by again applying Theorem 4. For any $\mathbf{x} \in \mathbb{R}^q$, we denote by $(X_n^{\mathbf{x}})$ the Markov chain defined with the same recurrence equation but with initial condition $X_0 = \mathbf{x}$. Then we have, by iterating,

$$X_n^{\mathbf{x}} = \mathrm{e}^{-\sum_{j=1}^n I_j A}\,\mathbf{x} + \sum_{k=1}^n \mathrm{e}^{-\sum_{j=k+1}^n I_j A}\,U_k\ , n \geq 1\ .$$

with the convention $\sum_{j=n+1}^n I_j = 0$. Recall that, given I_n, the conditional distribution of U_n is $\mathcal{N}(0, Q_n)$ and $Q_n = Q(I_n) \in [0, Q(\infty)]$. Hence $\mathbb{E}[|U_n|^s \mid I_n]$ is a bounded r.v. for any $s > 0$. By Lemma 3, we have $\mathbb{E}[\|\mathrm{e}^{-\sum_{j=k+1}^n I_j A}\|^s] \leq K\,(\mathbb{E}[\mathrm{e}^{-asI_1}])^{n-k}$ for same $K, s > 0$. Hence, we obtain, for any $s > 0$, for some constants $C > 0$ and $\alpha \in (0,1)$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^q$,

$$\mathbb{E}[|X_n^{\mathbf{x}} - X_n^{\mathbf{y}}|^s] \leq C\alpha^n\,(1 + |\mathbf{x}|^s + |\mathbf{y}|^s)\ ,$$
$$\mathbb{E}[|X_n^{\mathbf{x}}|^s] \leq C(1 + |\mathbf{x}|^s)\ .$$

These are conditions (i) and (ii) of Theorem 4 with $r = p = 1$. Moreover, (X_n) has a constant marginal distribution, namely $\mathcal{N}(0, Q(\infty))$, so that the invariant distribution μ of Theorem 4(a) is necessary $\mu = \mathcal{N}(0, Q(\infty))$. Now, applying Theorem 4(b) and Theorem 4(c), we get (35) and (36), with $a = 2$ and $a = 1$ respectively.

To achieve the proof of Theorem 2, it remains to prove that $\xi_{\text{H0:Signal}} > 0$. This results from $\log\det(C\mathbf{p}C^{\mathrm{T}} + 1_d) < \mathrm{tr}(CQ(\infty)C^{\mathrm{T}})$ for every $\mathbf{p} \in [0, Q(\infty)]$.

C. *Proof of Theorem 3.*

Let $\mathbf{w} = (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^q \times [0, Q(\infty)]$. We denote the first component of $Z_k^{\mathbf{w}}$ by $\underline{Z}_k^{\mathbf{w}}$ so that $Z_k^{\mathbf{w}} = (\underline{Z}_k^{\mathbf{w}}\ \tilde{Z}_k^{\mathbf{p}})$. Using the notation introduced above, we have, for all $k \geq 1$, $\underline{Z}_k^{\mathbf{w}} = \Theta_k^{\mathbf{p}} \underline{Z}_{k-1}^{\mathbf{w}} + \mathrm{e}^{-I_k A} G_{k-1}^{\mathbf{p}} Y_{k-1}$, and, by iterating,

$$\underline{Z}_n^{\mathbf{w}} = \Theta_{n,0}^{\mathbf{p}}\mathbf{x} + \sum_{k=1}^n \Theta_{n,k}^{\mathbf{p}} \mathrm{e}^{-I_k A} G_{k-1}^{\mathbf{p}} Y_{k-1}, \quad n \geq 1\ .$$

By continuity of G, it is bounded on the compact set $[0, Q(\infty)]$, hence $\sup_{\mathbf{p},n} \|G_k^{\mathbf{p}}\| < \infty$. Also by Lemma 3, $\sup_k \|\mathrm{e}^{-I_k A}\| < \infty$. Applying these bounds, Lemma 2, the Minkowski Inequality and the Hölder Inequality in the previous display, we obtain, for any $s > 0$ and some constant $C > 0$,

$$\mathbb{E}[|\underline{Z}_n^{\mathbf{w}}|^s] \leq C\,(1 + |\mathbf{x}|^s), \quad \mathbf{w} = (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^q \times [0, Q(\infty)],\ n \geq 1\ . \tag{37}$$

Since the second component of $Z_k^{\mathbf{w}}$ stays in the compact set $[0, Q(\infty)]$, this implies Condition (i) in Theorem 4 with $r = 1$ for the complete chain $(Z_k^{\mathbf{p}})_{k \geq 0}$.

Let now $\mathbf{v} = (\mathbf{y}, \mathbf{q}) \in \mathbb{R}^q \times [0, Q(\infty)]$. We have

$$\underline{Z}_n^{\mathbf{w}} - \underline{Z}_n^{\mathbf{v}} = \Theta_{n,0}^{\mathbf{p}}\mathbf{x} - \Theta_{n,0}^{\mathbf{q}}\mathbf{y} + \sum_{k=1}^n (\Theta_{n,k}^{\mathbf{p}} - \Theta_{n,k}^{\mathbf{q}}) \mathrm{e}^{-I_k A} G_{k-1}^{\mathbf{p}} Y_{k-1} + \sum_{k=1}^n \Theta_{n,k}^{\mathbf{q}} \mathrm{e}^{-I_k A} (G_{k-1}^{\mathbf{p}} - G_{k-1}^{\mathbf{q}}) Y_{k-1}\ .$$

Note that, using Lemma 2, the bounds (28) and (29) in Proposition 1, the Hölder Inequality and the Minkowski Inequality, we obtain, for any $r > 0$ and some constants $C > 0$, and $\rho \in (0,1)$ not depending on $\mathbf{p}, \mathbf{q}$,

$$\mathbb{E}[\|G_n^{\mathbf{p}} - G_n^{\mathbf{q}}\|^r] \leq C\rho^n \quad \text{and} \quad \mathbb{E}[\|\Theta_{n,m}^{\mathbf{p}} - \Theta_{n,m}^{\mathbf{q}}\|^r] \leq C\rho^n, \quad 0 \leq m < n \; .$$

Using these bounds, Lemma 2 and the previous two displays, we thus obtain, for any $q > 0$ and some constants $C > 0$, and $\alpha \in (0,1)$ not depending on $\mathbf{w}, \mathbf{v}$,

$$\mathbb{E}\left[|\underline{Z}_n^{\mathbf{w}} - \underline{Z}_n^{\mathbf{v}}|^q\right] \leq C\alpha^n \left(1 + |\mathbf{x}|^q + |\mathbf{y}|^q\right), \quad n \geq 1 \; .$$

This, with (32), implies Condition (i) in Theorem 4 with $p = 1$ for the chain $(Z_k^{\mathbf{p}})_{k\geq 0}$. Hence Theorem 4(a) applies, which yields the existence and uniqueness of the measure ν. Moreover we get that $\int |\mathbf{w}|^r \mathrm{d}\nu(\mathbf{w}) < \infty$ for any $r > 0$.

Let us now show that $\mathcal{L}_N(Y_{1:N}, T_{1:N})$ converges to $\xi_{\text{H0:Noise}}$ when $Y_n = V_n$ for all $n \geq 1$. Some of the terms appearing in (10) are identical to the case where $Y_n = CX_n + V_n$ for all $n \geq 1$ investigated for the proof of Theorem 2. Writing

$$(Y_n - \hat{Y}_n)^{\mathrm{T}} \Delta_n^{-1} (Y_n - \hat{Y}_n) = V_n^{\mathrm{T}} \Delta_n^{-1} V_n + 2V_n^{\mathrm{T}} \Delta_n^{-1} C\hat{X}_n + \hat{X}_n^{\mathrm{T}} C^{\mathrm{T}} \Delta_n^{-1} C\hat{X}_n \; ,$$

and using $\mu = \nu(\mathbb{R}^q, \cdot)$, (33), (34) and some algebra, it is in fact sufficient to prove that

$$\frac{1}{N}\sum_{n=1}^{N} \hat{X}_n^{\mathrm{T}} C^{\mathrm{T}} \Delta_n^{-1} C\hat{X}_n \xrightarrow[N\to\infty]{\text{a.s.}} \int \mathbf{x}^{\mathrm{T}} C^{\mathrm{T}} (C\mathbf{p}C^{\mathrm{T}} + 1_d)^{-1} C\mathbf{x} \, \mathrm{d}\nu(\mathbf{x}, \mathbf{p}) \; , \tag{38}$$

$$\frac{1}{N}\sum_{n=1}^{N} V_n^{\mathrm{T}} \Delta_n^{-1} V_n \xrightarrow[N\to\infty]{\text{a.s.}} \int \mathrm{tr}(C\mathbf{p}C^{\mathrm{T}} + 1_d)^{-1} \, \mathrm{d}\mu(\mathbf{p}) \; , \tag{39}$$

$$\frac{1}{N}\sum_{n=1}^{N} V_n^{\mathrm{T}} \Delta_n^{-1} C\hat{X}_n \xrightarrow[N\to\infty]{\text{a.s.}} 0 \; . \tag{40}$$

Now, these limits hold by observing that $(\hat{X}_n, P_n) = Z_{n-1}^{(0,Q(\infty))}$ and by apply Theorem 4(b) with $a = 1$ for (38), Theorem 4(c) with $a = 1$ for (40) and Theorem 4(c) with $a = 2$ for (40).

It remains to prove that $\xi_{\text{H0:Noise}} > 0$. From Equation (17), $\xi_{\text{H0:Noise}} \geq \int f(\mathbf{p}) \mathrm{d}\mu(\mathbf{p})$ where $f(\mathbf{p}) = 0.5\left(\log\det(C\mathbf{p}C^{\mathrm{T}} + 1_d) - C\mathbf{p}C^{\mathrm{T}}(C\mathbf{p}C^{\mathrm{T}} + 1_d)^{-1}\right)$. This function satisfies $f(\mathbf{p}) \geq 0$ and $f(\mathbf{p}) = 0$ if and only if $C\mathbf{p}C^{\mathrm{T}} = 0$. Let $\tilde{Z} \in [0, Q(\infty)]$ be a random variable with the invariant distribution μ, and assume that $C\tilde{Z}C^{\mathrm{T}} = 0$ with probability one. From Equation (15) we have with probability one

$$0 = C\tilde{F}_I(\tilde{Z})C^{\mathrm{T}} = C\mathrm{e}^{-IA}\tilde{Z}\mathrm{e}^{-IA^{\mathrm{T}}}C^{\mathrm{T}} - C\mathrm{e}^{-IA}\tilde{Z}C^{\mathrm{T}}\left(C\tilde{Z}C^{\mathrm{T}} + 1_d\right)^{-1} C\tilde{Z}\mathrm{e}^{-IA^{\mathrm{T}}}C^{\mathrm{T}} + CQ(I)C^{\mathrm{T}}$$
$$= C\mathrm{e}^{-IA}\tilde{Z}\mathrm{e}^{-IA^{\mathrm{T}}}C^{\mathrm{T}} + CQ(I)C^{\mathrm{T}} = CQ(I)C^{\mathrm{T}}$$

Due to the controllability of (A, B) and the fact that $\tau(\{0\}) < 1$, this is a contradiction. Therefore $f(\tilde{Z}) > 0$ with probability one, and hence $\xi_{\text{H0:Noise}} > 0$, which achieves the proof of Theorem 3.

IV. Some particular cases

Different particular cases and limit situations will be considered in this section. We begin with the case where the sampling is regular, *i.e.*, I_1 is equal to a constant that we take equal to one without loss of generality. In this case, we obtain compact expressions for the error exponents. We then consider the case where the holding times are large with high probability, *i.e.*, the sensors tend to be far apart. Finally, we consider the case where the SDE (1) is a scalar equation. In the scalar case, we will be able to analyze the impact of $\mathbb{E}[I_1]$, the Signal to Noise Ratio, and the distribution of I_1 on $\xi_{\text{H0:Signal}}$. All proofs are deferred to Appendix C.

Regular sampling

When the sampling is regular, the model for (Y_n) under $\mathsf{H1}$ (see Eqs. (11)) is a general model for stable Gaussian multidimensional ARMA processes corrupted with a Gaussian white noise. In this case we denote by $\Phi = \mathrm{e}^{-I_{n+1}A} = \exp(-A)$ and by $Q = Q(1) = \int_0^1 \exp(-uA)BB^{\mathrm{T}}\exp(-uA^{\mathrm{T}})\,\mathrm{d}u$ the state transition matrix and the excitation covariance matrix respectively.

Proposition 2 (Regular sampling) *Assume that $I_1 = 1$ and define $\xi_{\text{H0:Signal}}$ and $\xi_{\text{H0:Noise}}$ by (16) and (17), respectively. Then we have*

$$\xi_{\text{H0:Signal}} = \frac{1}{2}\left(\mathrm{tr}\left(CQ(\infty)C^{\mathrm{T}}\right) - \log\det\left(CP_RC^{\mathrm{T}} + 1_d\right)\right) \ , \tag{41}$$

$$\xi_{\text{H0:Noise}} = \frac{1}{2}\Big(\log\det\left(CP_RC^{\mathrm{T}} + 1_d\right) - \mathrm{tr}\left[CP_RC^{\mathrm{T}}\left(CP_RC^{\mathrm{T}} + 1_d\right)^{-1}\right] + \mathrm{tr}\left[C\Sigma C^{\mathrm{T}}\left(CP_RC^{\mathrm{T}} + 1_d\right)^{-1}\right]\Big) \ , \tag{42}$$

where P_R is the unique solution of the matrix equation

$$P = \Phi P \Phi^{\mathrm{T}} - \Phi P C^{\mathrm{T}}\left(CPC^{\mathrm{T}} + 1_d\right)^{-1} CP\Phi^{\mathrm{T}} + Q \tag{43}$$

and where the $q \times q$ symmetric matrix Σ is the unique solution of the matrix linear equation

$$\Sigma - \Phi(1_q - GC)\Sigma(1_q - GC)^{\mathrm{T}}\Phi^{\mathrm{T}} = \Phi G G^{\mathrm{T}} \Phi^{\mathrm{T}} \tag{44}$$

with $G = P_RC^{\mathrm{T}}\left(CP_RC^{\mathrm{T}} + 1_d\right)^{-1}$.

Equation (43) is the celebrated discrete algebraic Riccati equation. Its solution P_{R} is the asymptotic (steady state) error covariance matrix when the sampling is regular. The matrix $G = P_{\mathrm{R}} C^{\mathrm{T}} \left(C P_{\mathrm{R}} C^{\mathrm{T}} + I \right)^{-1}$ is the Kalman filter steady state gain matrix [28, Chap. 4].

Large Holding Times

We now study the behavior of the error exponents when the holding times are large with high probability. We shall say that a family (τ_s) of probability distributions on $[0, \infty)$ "escapes to infinity" if, as $s \to \infty$

$$\text{for all } K > 0,\ \tau_s([0, K]) \xrightarrow[s\to\infty]{} 0.$$

In order to study the large holding time behavior of the error exponents, we index the distribution of the holding times by s and assume that τ_s escapes to infinity. A typical particular case that illustrates this situation is when we assume that the I_n are equal in distribution to $s\bar{I}$ where $\bar{I}$ is some nonnegative random variable, and when we study the behavior of the error exponents for large values of s.

Proposition 3 (Large holding times) *Assume that (τ_s) escapes to infinity and define $\xi_{\text{H0:Signal}}$ and $\xi_{\text{H0:Noise}}$ by (16) and (17), respectively. Then, as $s \to \infty$,*

$$\xi_{\text{H0:Noise}} \to \frac{1}{2} \Big(\log\det \left(CQ(\infty)C^{\mathrm{T}} + 1_d \right) - \operatorname{tr} \Big[CQ(\infty)C^{\mathrm{T}} \left(CQ(\infty)C^{\mathrm{T}} + 1_d \right)^{-1} \Big] \Big), \tag{45}$$

$$\xi_{\text{H0:Signal}} \to \frac{1}{2} \left(\operatorname{tr} \left(CQ(\infty)C^{\mathrm{T}} \right) - \log\det \left(CQ(\infty)C^{\mathrm{T}} + 1_d \right) \right). \tag{46}$$

Given an $\mathbb{R}^d$-valued i.i.d. sequence (Y_n) such that $Y_1 \sim \mathcal{N}(0, 1_d)$ under H0 and $Y_1 \sim \mathcal{N}(0, CQ(\infty)C^{\mathrm{T}} + 1_d)$ under H1, it is well known that, under H0, the LLR converges to the Kullback-Leibler divergence $D\left(\mathcal{N}(0, 1_d) \,\|\, \mathcal{N}(0, CQ(\infty)C^{\mathrm{T}} + 1_d)\right)$, which equals the RHS of (45), while under H1, the negated LLR converges to $D\left(\mathcal{N}(0, CQ(\infty)C^{\mathrm{T}} + 1_d) \,\|\, \mathcal{N}(0, 1_d)\right)$, which equals the RHS of (46). This can be explained as follows. When τ_s escapes to infinity, two consecutive samples $X(T_n)$ and $X(T_{n+1})$ are asymptotically uncorrelated. Hence, under the signal hypothesis and as τ_s escapes to infinity, the process (Y_n) can be seen as a centered Gaussian i.i.d. sequence with covariance matrix $CQ(\infty)C^{\mathrm{T}} + 1_d$.

The Scalar Case

In the scalar case, the SDE (1) becomes

$$\mathrm{d}X(t) = -a\, X(t)\, \mathrm{d}t + b\, \mathrm{d}W(t), \quad t \geq 0, \tag{47}$$

where $W(t)$ is a scalar Brownian motion and (a, b) are known real non zero constants. The SDE (47) defines a so called Ornstein-Uhlenbeck (O-U) process. Under the stationary assumption, $a > 0$ and the initial value $X(0)$ is independent from $W(t)$ and follows the law $\mathcal{N}(0, Q(\infty))$ where the variance $Q(\infty)$ is given by $Q(\infty) = b^2/(2a)$. We observe $(Y_n, T_n)_{1 \le n \le N}$ where (Y_n) is a scalar process and we write the H0-Noise test as

$$\mathsf{H0} : Y_n = V_n \quad \text{for } n = 1, \dots, N \tag{48}$$

$$\mathsf{H1} : Y_n = X(T_n) + V_n \quad \text{for } n = 1, \dots, N \ , \tag{49}$$

where the observation noise process (V_n) is i.i.d. with $V_1 \sim \mathcal{N}(0, 1)$. Solving Equation (47) between T_n and T_{n+1} we obtain that $X_n = X(T_n)$ is given by

$$X_{n+1} = e^{-aI_{n+1}} X_n + U_{n+1}, \quad n \in \mathbb{N}, \tag{50}$$

where, given (I_n), (U_n) is a sequence of independent variables such that $U_n \sim \mathcal{N}(0, Q_n)$ with $Q_n = Q(\infty)(1 - e^{-2aI_n})$. The distribution of the process (Y_n) under $\mathsf{H1}$ is completely described by the scalars a and $Q(\infty)$ and by the distribution τ of I_1, and so are the error exponents. Recall that we assume here that I_1 is integrable. In this case, we can set $\mathbb{E}[I_1] = 1$ by including the mean holding time into a. The parameter $Q(\infty)$ determines the marginal distribution of X, since $X(t) \sim \mathcal{N}(0, Q(\infty))$ for every $t \ge 0$, and can thus be interpreted as the Signal to Noise Ratio SNR $= \mathbb{E}[X_n^2]/\mathbb{E}[V_n^2]$.

We now provide the error exponents when the sampling is regular. In the scalar case, it is easy to solve Equations (43) and (44) in the statement of Proposition 2 and to obtain the error exponents in closed forms.

Corollary 2 (Regular sampling in the scalar case) *In the scalar case, defining $\xi_{\mathit{H0:Signal}}$ and $\xi_{\mathit{H0:Noise}}$ as in (41) and (42), respectively, we have*

$$\xi_{\mathit{H0:Signal}} = \frac{1}{2} \left(\mathit{SNR} - \log\left(1 + P_R\right) \right) \ , \tag{51}$$

$$\xi_{\mathit{H0:Noise}} = \frac{1}{2} \left(\log\left(1 + P_R\right) + \frac{P_R}{P_R + 1} \left(\frac{\Phi^2 P_R}{P_R^2 + 2P_R + 1 - \Phi^2} - 1 \right) \right) \ , \tag{52}$$

where $\Phi = \exp(-a)$ and

$$P_R = \frac{(\mathit{SNR} - 1)(1 - \Phi^2) + \sqrt{(\mathit{SNR} - 1)^2 (1 - \Phi^2)^2 + 4\mathit{SNR}(1 - \Phi^2)}}{2} \ .$$

We note that (52) was first proved in [7, Theorem 1]. The proof of (51) is straightforward and thus omitted.

We now consider a general distribution for the holding times and consider the behavior of $\xi_{\text{H0:Signal}}$ with respect to a, the Signal to Noise Ratio SNR $= Q(\infty)$, and the distribution of I_1.

Proposition 4 *Define $\xi_{\textit{H0:Signal}}$ and $\xi_{\textit{H0:Noise}}$ by (16) and (17), respectively. In the scalar case, the following properties hold true.*

(i) The error exponent $\xi_{\textit{H0:Signal}}$ decreases as a increases. Moreover, as $a \to 0$, $\xi_{\textit{H0:Signal}} = Q(\infty)/2$ and, as $a \to \infty$, $\xi_{\textit{H0:Signal}} = (Q(\infty) - \log(Q(\infty)+1))/2$.

(ii) $\xi_{\textit{H0:Signal}}$ increases as $Q(\infty)$ increases.

(iii) The distribution τ of I_1 that minimizes $\xi_{\textit{H0:Signal}}$ under the constraint $\mathbb{E}[I_1] = 1$ is $\tau = \delta_1$.

The proof of Proposition 4 can be found in Appendix C.

Some practical design guidelines can be inferred from this proposition: from the stand point of the error exponent theory, when H0 refers to the presence of a noisy O-U signal, one has an interest in choosing close sensors if one wants to reduce the Type II error probability. This probability is reduced by exploiting the correlations between the X_n. As regards the sampling strategy, the worst sampling from the error exponent stand point is the regular sampling. We note that the problem of determining the best distribution τ, that is, the one that maximizes $\xi_{\text{H0:Signal}}$ with a given mean is an open question.

In the setting of Theorem 3, the behavior of $\xi_{\text{H0:Noise}}$ with respect to a has been analyzed in the regular sampling case only (Corollary 2) in [7]. The authors of [7] proved that when SNR ≥ 0 dB, $\xi_{\text{H0:Noise}}$ is an increasing function of a while when SNR < 0 dB, $\xi_{\text{H0:Noise}}$ admits a maximum with respect to a. By a numerical estimation of $\xi_{\text{H0:Noise}}$ (see below), we observe a similar behavior in the case of a Poisson sampling. However, a more formal characterization of the shape of $\xi_{\text{H0:Noise}}$ for a general distribution τ seems to be difficult. A more detailed discussion on the behavior of $\xi_{\text{H0:Noise}}$ is provided in the next section.

V. NUMERICAL ILLUSTRATION AND INTERPRETATION OF THE RESULTS

Let us first describe the simulation procedure. Let $(\widehat{X}_\infty, P_\infty)$ be a random element of $\mathbb{R}^q \times [0, Q(\infty)]$ distributed according to the invariant distribution ν of the Markov process $(\widehat{X}_n, P_n)$. Then the error exponent defined by (17) can be written as

$$\xi_{\text{H0:Noise}} = \frac{1}{2}\mathbb{E}\left[\log\det\left(CP_\infty C^{\mathrm{T}} + 1_d\right) + C\left(\widehat{X}_\infty \widehat{X}_\infty^{\mathrm{T}} - P_\infty\right)C^{\mathrm{T}}\left(CP_\infty C^{\mathrm{T}} + 1_d\right)^{-1}\right]$$

and the error exponent defined by (16) as

$$\xi_{\text{H0:Signal}} = \frac{1}{2}\left(\mathrm{tr}\left(CQ(\infty)C^{\mathrm{T}}\right) - \mathbb{E}\left[\log\det\left(CP_\infty C^{\mathrm{T}} + 1_d\right)\right]\right).$$

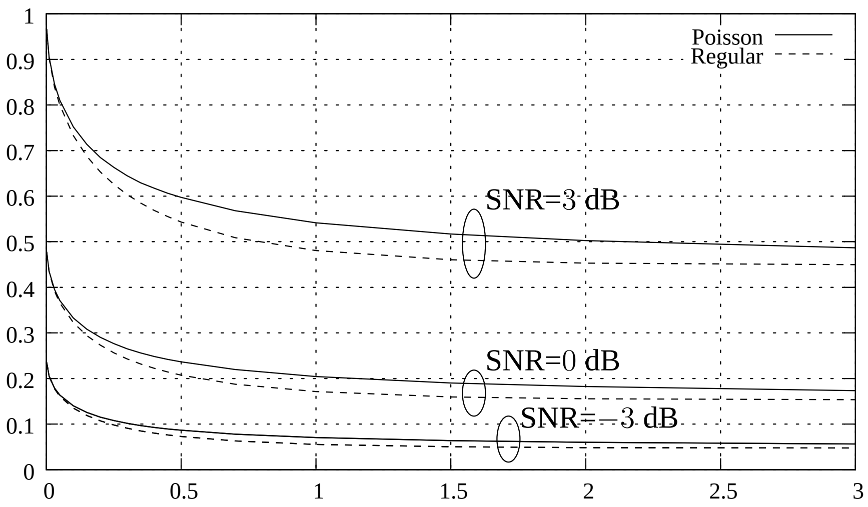

Fig. 1. Scalar case: $\xi_{\text{H0:Signal}}$ vs a for SNR $= -3, 0$ and 3 dB

By the stability of the Markov chain $(\widehat{X}_n, P_n)$ shown in Section III, we estimate the error exponents by simulating the Kalman Equations (12)-(13) with (Y_n) i.i.d. $\mathcal{N}(0, 1_d)$, and by estimating the expectation in the equations above by empirical means taken on $(\widehat{X}_n, P_n)_{n=1,\dots,N}$ for N large enough. A scalar case and a vector case are considered.

A. The scalar case

Figures 1 and 2 describe the behavior of the error exponents in the scalar case. Poisson sampling with $\mathbb{E}[I_1] = 1$ and regular sampling with $I_1 = 1$ are both displayed in the figures.

In Fig. 1, the error exponent $\xi_{\text{H0:Signal}}$ is plotted as a function of a for SNR $(= Q(\infty))\ = -3, 0$ and 3 dB. We note that the empirical results match the theoretical findings stated in Proposition 4.

In Fig. 2, $\xi_{\text{H0:Noise}}$ is plotted *vs* a for SNR $= -3, 0$ and 3 dB. We notice that $\xi_{\text{H0:Noise}} \to 0$ as $a \to 0$. Moreover, this error exponent is an increasing function of a for SNR $= 0$ and 3 dB while it has a maximum with respect to a for SNR $= -3$ dB. As said in Section IV, this behavior has been established in [7] in the case of a regular sampling. We also notice that Poisson sampling is worse than regular sampling for SNR $= 3$ dB and better than regular sampling for SNR $= -3$ dB. We will further discuss these findings in Section V-C below.

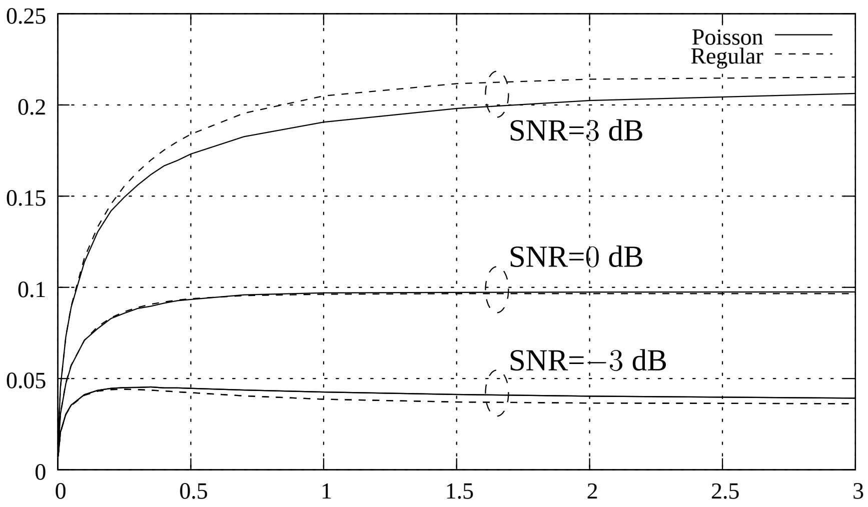

Fig. 2. Scalar case: $\xi_{\text{H0:Noise}}$ vs a for SNR $= -3, 0$ and 3 dB

B. The vector case

We now consider a vector case and investigate whether the qualitative findings in the scalar case are again observed. The following 2-dimensional process is considered.

$$\mathrm{d}X(t) = -\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} X(t)\,\mathrm{d}t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathrm{d}W(t)$$

where $W(t)$ is a scalar Brownian motion. We take $C = 1_2$ in (4). Figures 3 and 4 concern the behavior of $\xi_{\text{H0:Signal}}$ and $\xi_{\text{H0:Noise}}$ in the vector case. Both Poisson and regular sampling are considered. In the Poisson sampling case, we assume that the I_n are equal in distribution to sI where I is an exponential random variable with mean one, and we plot the error exponents in terms of the mean holding time s. In the regular sampling case, s is simply the sensor spacing. The last parameter is the SNR given by

$$\text{SNR} = \frac{\mathbb{E}[|CX_n + V_n|^2]}{\mathbb{E}[|V_n|^2]} = \frac{\operatorname{tr}\left(CQ(\infty)C^{\mathrm{T}}\right)}{d} \ .$$

In this experimental setting, a behavior comparable to the scalar case is observed for both tests. In the case of the H0-Signal test, we also observe that $\xi_{\text{H0:Signal}}$ decreases as s increases, and Poisson sampling enjoys a higher error exponent than the regular sampling for the three considered SNR. In the case of the H0-Noise test, the error exponent increases with s at high SNR, while it has a maximum with respect to s at low SNR, and the Poisson sampling is worse than the regular sampling at high SNR.

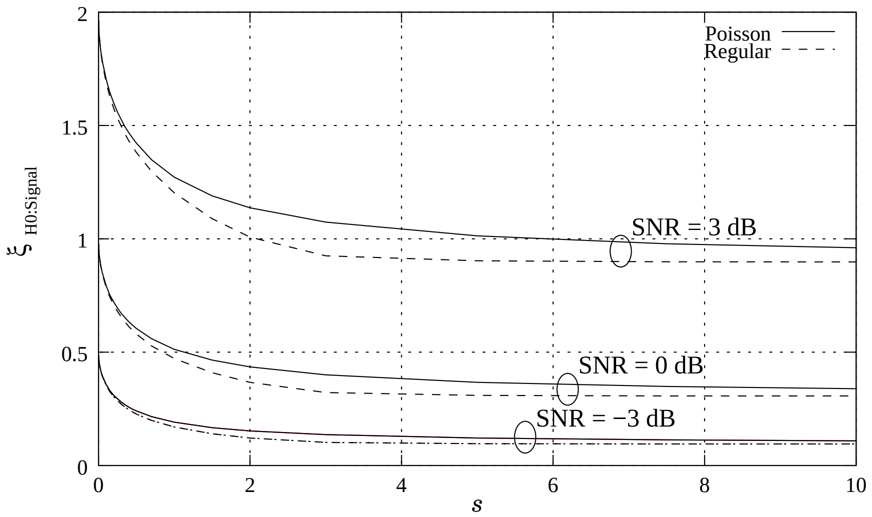

Fig. 3. Vector case: $\xi_{\text{H0:Signal}}$ vs s for SNR $= -3, 0$ and 3 dB

C. Discussion on the error exponents behavior.

Small sampling spacing: Figures 1 and 2 show that $\xi_{\text{H0:Signal}}$ increases to $Q(\infty)/2$ as $a \downarrow 0$ (as predicted by Proposition 4) while $\xi_{\text{H0:Noise}}$ decreases to zero as $a \downarrow 0$. This behavior has the following heuristic interpretation. At $a = 0$, Equations (50) boil down to $X_N = \cdots = X_n = \cdots = X_0 \sim \mathcal{N}(0, Q(\infty))$. Under H1, it is easy to show that the corresponding negated LLR converges to $X_0^2/2$, which has expectation $Q(\infty)/2$, the limit of $\xi_{\text{H0:Signal}}$ as $a \downarrow 0$. In contrast, as already noticed in [7], a direct derivation shows that the error exponent $\xi_{\text{H0:Noise}}$ of the limit model is zero, since the Neyman-Pearson Type II error probability decreases as $\mathcal{O}(1/\sqrt{N})$, that is much more slowly than the usual exponential decreasing. This is in accordance with the observed behavior on simulations, namely, $\xi_{\text{H0:Noise}} \to 0$ as $a \downarrow 0$.

Small versus large SNR: Here we denote the SNR by $q = Q(\infty)$ and let $\varphi_\tau(a) = \mathbb{E}\left[\mathrm{e}^{-2aI_1}\right]$, for I_1 with distribution τ. In Appendix D we provide a heuristic calculation that yields the following approximations. As $q \to 0$,

$$\xi_{\text{H0:Noise}} \sim q^2 \frac{1 + \varphi_\tau(a)}{4(1 - \varphi_\tau(a))} \tag{53}$$

and, as $q \to \infty$,

$$\xi_{\text{H0:Noise}} - \frac{1}{2}\log(q) \to \frac{1}{2}\left(\mathbb{E}\left[\log(1 - \mathrm{e}^{-2aI_1})\right] - 1\right) \ . \tag{54}$$

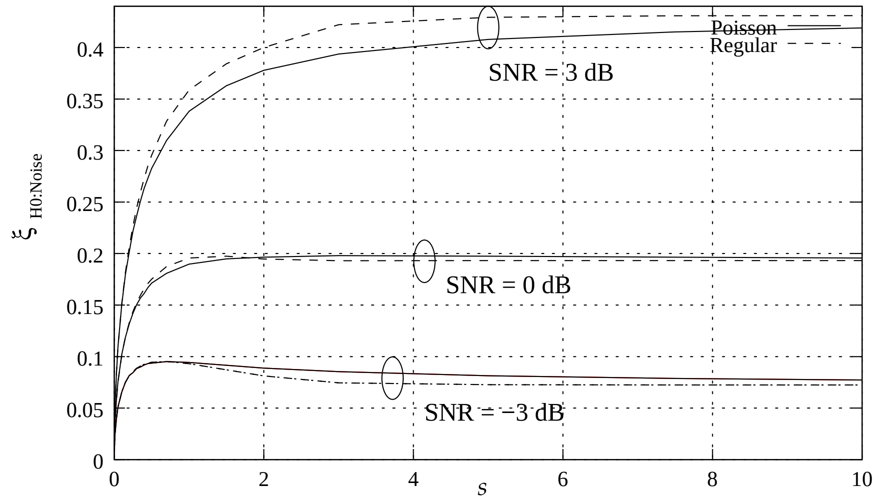

Fig. 4. Vector case: $\xi_{\text{H0:Noise}}$ vs s for SNR $= -3, 0$ and 3 dB

These results show that for small SNR, the regular sampling has the lowest error exponent, and for high SNR the regular sampling enjoys the highest error exponent in the set of distributions τ for which $\mathbb{E}[I_1] = 1$. Indeed, as $\exp(-x)$ is convex and $(1+x)/(1-x)$ is increasing on $[0, 1)$, the q^2 term at the right hand side of (53) is minimum when $I_1 = 1$ with probability one. At low SNR, the error exponent loss L due to the use of a regular sampling (that is, the SNR to pay to achieve an equivalent error exponent) is

$$L = 5 \log_{10} \left(\frac{(1+\varphi_\tau(a))(1-\exp(-2a))}{(1-\varphi_\tau(a))(1+\exp(-2a))} \right) \text{ dB} . \tag{55}$$

At large SNR, as $\log(1-\exp(-x))$ is concave, the right hand side of (54) is maximum when $I_1 = 1$ with probability one. Thus, at high SNR, the error exponent gain G due to the use of the regular sampling is

$$G = 10 \left(\log_{10}(1-\exp(-2a)) - \mathbb{E}\left[\log_{10}(1-\mathrm{e}^{-2aI_1}) \right] \right) \text{ dB} . \tag{56}$$

These results are illustrated in Figure 5 for the low SNR regime, and in Figure 6 for the high SNR regime, where $\xi_{\text{H0:Noise}}$ is plotted as a function of the SNR q for $a = 1$. In these figures, the curves termed "Asymp. Poisson" and "Asymp. Regular" represent the asymptotic values of $\xi_{\text{H0:Noise}}$ provided by Equations (53) for Fig. 5 and (54) for Fig. 6. At low SNR, the asymptotic loss incurred by the regular sampling with respect to a Poisson sampling as predicted by (55) is $L = 0.91$ dB, and this is confirmed numerically in Figure 5. At high SNR, (56) predicts for the regular sampling the asymptotic gain $G = 2.03$ dB with respect to the Poisson sampling. This is confirmed numerically by Figure 6.

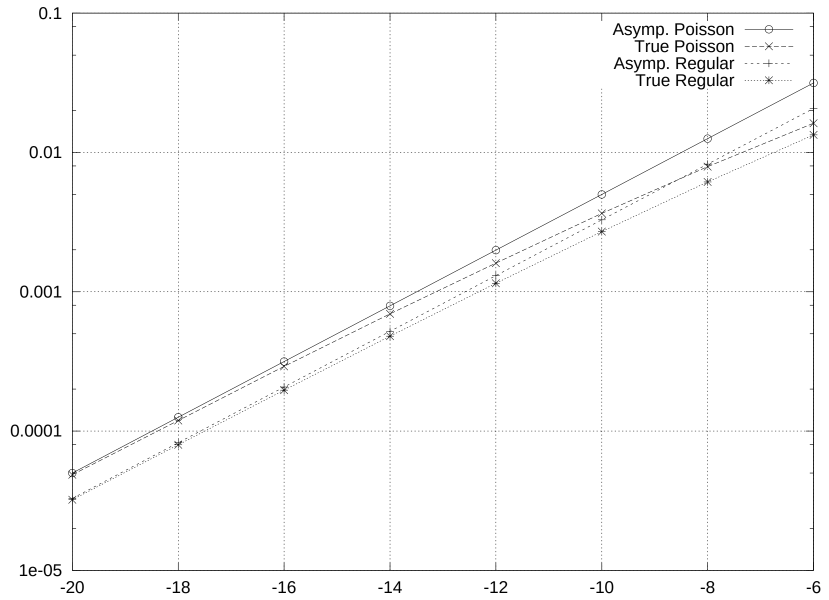

Fig. 5. $\xi_{\text{H0:Noise}}$ (log scale) vs SNR in dB in the low SNR regime, $a = 1$.

APPENDIX

A. Technical lemmas

In this section we provide some useful technical lemmas.

Lemma 3 *Assume that A is positive stable. Then there exists constants $a > 0$ and $K > 0$ such that $\|e^{-xA}\| \le K \exp(-xa)$ for $x \ge 0$.*

Proof: Let $a > 0$ be smaller than the real parts of all the eigenvalues of A and $\mathcal{C}$ be a rectangle in the complex half plane $\{z : \Re(z) \ge a\}$ whose interior contains all these eigenvalues (here $\Re(z)$ is the real part of z). Applying Theorem 6.2.28 of [20], we have $e^{-xA} = \frac{1}{2\pi\,i} \int_{\mathcal{C}} e^{-x\lambda} (\lambda\, I - A)^{-1} \, d\lambda$. Hence, we have

$$\|e^{-xA}\| \le e^{-xa} \int_{\mathcal{C}} \|(\lambda\, I - A)^{-1}\| \, d\lambda.$$

By continuity of $\lambda \mapsto (\lambda\, I - A)^{-1}$ on $\mathcal{C}$, the previous integral is finite, which gives the result. ■

Lemma 4 *Assume that the matrix A is positive stable and that the pair (A, B) is controllable. Then the matrix function $Q(x)$ defined by (6) is strictly increasing in the positive semidefinite ordering from 0 to $Q(\infty)$ as x increases from 0 to ∞.*

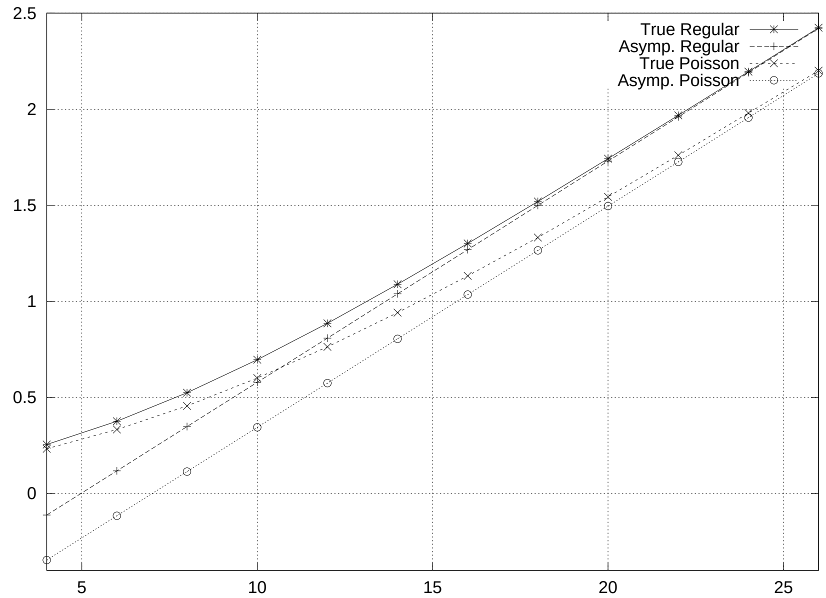

Fig. 6. $\xi_{\mathrm{H0:Noise}}$ vs SNR in dB in the high SNR regime, $a = 1$.

Proof: Since (A, B) is controllable, $Q(x) > 0$ for any $x > 0$. Assume that $x < y$. We have $Q(y) - Q(x) = \int_x^y \exp(-uA)BB^{\mathrm{T}} \exp(-uA^{\mathrm{T}})du = \exp(-xA)Q(y-x)\exp(-xA^{\mathrm{T}}) > 0$ which proves the lemma. ∎

B. A stability result on Markov chains

Here we present our Swiss knife result on Markov chains. We follow the approach in [29] for obtaining the geometric ergodicity of Markov chains using simple moment conditions, although we use a more direct proof inspired from [30]. For the *a.s.* convergence of the empirical mean, we will rely on the following standard result for martingales [31].

Lemma 5 *Let $(M_n)_{n\geq 0}$ be a martingale sequence and $X_n = M_n - M_{n-1}$ be its increments. If there exists $p \in [1, 2]$ such that $\sum_{k\geq 1} k^{-p}\mathbb{E}[|X_k|^p] < \infty$, then $M_n/n \xrightarrow[n\to\infty]{\text{a.s.}} 0$.*

We adopt the following setting for our generic Markov chain. Let $(\eta_k,\, k \geq 1)$ be an i.i.d. sequence of random variables valued in E and let $\mathcal{X}$ be a closed subset of $\mathbb{R}^d$. We further denote by η a random variable having the same distribution as the η_k's and independent of them. Let $F_y(x)$ be defined for all $y \in \mathrm{E}$ and $x \in \mathcal{X}$ with values in $\mathcal{X}$ and such that $(x, y) \mapsto F_y(x)$ is a measurable $\mathcal{X} \times \mathrm{E} \to \mathcal{X}$ function.

This allows to define a Markov chain $(Z_k^x,\, k \geq 0)$ by

$$\begin{cases} Z_0^x = x \;, \\ Z_k^x = F_{\eta_k}(Z_{k-1}^x), \quad k \geq 1 \;. \end{cases} \tag{57}$$

This Markov chain is valued in $\mathcal{X}$ and start at time 0 with the value x. We denote by P the corresponding kernel defined on any bounded continuous function $f : \mathcal{X} \to \mathbb{R}$ by

$$Pf(x) = \mathbb{E}(f(Z_1^x)) = \mathbb{E}(f \circ F_{\eta_1}(x)), \quad x \in \mathcal{X} \;,$$

where $\circ$ denotes the composition operator. Observe that (57) implies, for all $n \geq 1$

$$Z_n^x = F_{\eta_n} \circ \cdots \circ F_{\eta_1}(x) \;.$$

Recalling that $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$, we denote for any $p \geq 1$ and $f : \mathcal{X} \to \mathbb{R}$,

$$\|f\|_{\mathrm{Lip}_p} = \sup_{x,x' \in \mathcal{X}^2} \frac{|f(x) - f(x')|}{|x - x'|\,(1 + |x|^{p-1} + |x'|^{p-1})} \;,$$

which is the Lipschitz norm for $p = 1$. We now state the main result of this appendix.

Theorem 4 *Define $(Z_k^x,\, k \geq 0)$ as in (57). Assume that F_η is* a.s. *continuous, and that, for some $C > 0$, $\alpha \in (0,1)$, $p, r \geq 0$, $q \geq 1$ and $s \geq p$,*

(i) For all $(x, x') \in \mathcal{X}^2$ and $n \geq 1$,

$$\mathbb{E}\left[|F_{\eta_n} \circ \cdots \circ F_{\eta_1}(x) - F_{\eta_n} \circ \cdots \circ F_{\eta_1}(x')|^q\right] \leq C\alpha^n \left(1 + |x|^{pq} + |x'|^{pq}\right) \;.$$

(ii) For all $x \in \mathcal{X}$ and $n \geq 1$,

$$\mathbb{E}\left[|F_{\eta_n} \circ \cdots \circ F_{\eta_1}(x)|^s\right] \leq C(1 + |x|^{rs}) \;. \tag{58}$$

Then the following conclusions hold.

(a) There exists a unique probability measure μ on $\mathcal{X}$ such that

$$\xi \sim \mu \text{ and } \xi \text{ independent of } \eta \Rightarrow F_\eta(\xi) \sim \mu \;. \tag{59}$$

Moreover such a measure μ has a finite s-th moment.

(b) Let $a \in [1, s \wedge \{1 + s(q-1)/q\}]$ and $f : \mathcal{X} \to \mathbb{R}$ such that $\|f\|_{\mathrm{Lip}_a} < \infty$. Suppose in addition that $s > b = p + r(a-1)$. Then, for all $x \in \mathcal{X}$,

$$\frac{1}{n}\sum_{k=1}^{n} f(Z_k^x) \xrightarrow[n\to\infty]{\text{a.s.}} \mu(f) = \int f d\mu \;.$$

(c) Let $(U_n)_{n\geq 1}$ be a sequence of i.i.d. real-valued random variables such that $\mathbb{E}[|U_1|^{1+\epsilon}] < \infty$ for some $\epsilon > 0$ and, for all $n \geq 1$, U_n is independent of $\eta_1, \dots, \eta_n$. Then, under the same assumptions as in (b), if moreover $s > a$, then, for all $x \in \mathcal{X}$,

$$\frac{1}{n}\sum_{k=1}^{n} U_k f(Z_k^x) \xrightarrow[n\to\infty]{\text{a.s.}} m\ \mu(f)\ ,$$

where $m = \mathbb{E}[U_1]$.

Proof: Let us introduce the backward recurrence process starting at x defined by $Y_0 = x$ and

$$Y_n = F_{\eta_1} \circ \dots \circ F_{\eta_n}(x), \quad n \geq 1\ .$$

Note that for any n, $Y_n \stackrel{d}{=} Z_n^x$, that is, the processes (Y_n) and (Z_n^x) has the same marginal distributions. Moreover, using (i) and the Jensen Inequality, we have

$$\mathbb{E}\left[\sum_{n\geq 0} |Y_{n+1} - Y_n|\right] \leq \sum_{n\geq 0} C^{1/q}\alpha^{n/q}\ (1 + |x|^p + \mathbb{E}\left[|F_{\eta_n}(x)|^p\right])\ .$$

By (ii), since $s \geq p$, $\mathbb{E}\left[|F_{\eta_n}(x)|^p\right] < \infty$ and thus $\sum_{n\geq 0} |Y_{n+1} - Y_n| < \infty$ *a.s.* By completeness of the state space $\mathcal{X}$, Y_n converges in $\mathcal{X}$ *a.s.* We denote the limit by ξ and its probability distribution by μ. By *a.s.* continuity of F_η, we have $F_\eta(Y_n) \to F_\eta(\xi)$ *a.s.* On the other hand $F_\eta(Y_n) \sim Y_{n+1} \to \xi$ *a.s.* Hence μ satisfies (59), that is, μ is an invariant distribution of the induced Markov chain. Moreover by (ii), we have $\sup_n \mathbb{E}[|Y_n|^s] < \infty$ which by Fatou's Lemma implies that $\mathbb{E}[|\xi|^s] < \infty$. Let us show that μ is the unique invariant distribution. By (i), for any $x, y \in \mathcal{X}$, $Z_n^x - Z_n^y \xrightarrow[n\to\infty]{a.s.} 0$. Now draw x and y according to two invariant distributions, respectively, so that $(Z_n^x)_{n\geq 0}$ and $(Z_n^y)_{n\geq 0}$ are two sequences with constant marginal distributions. Then necessarily these two distributions are the same and thus μ is the unique invariant distribution, which achieves the proof of (a).

We now prove (b). First observe that f is continuous and $f(x) = \mathcal{O}(|x|^a)$ as $|x| \to \infty$. Hence by (a), since $a \leq s$, f is integrable with respect to μ. Also, by (ii), $\mathbb{E}[|f(Z_k^x)|] < \infty$ for all $k \geq 1$ and $x \in \mathcal{X}$. We use the classical Poisson equation for decomposing the empirical mean of the Markov chain as the empirical mean of martingale increments plus a negligible remainder. Using that $\|f\|_{\mathrm{Lip}_a} < \infty$, the Hölder inequality and (i), we have, for any $x, y \in \mathcal{X}$,

$$\mathbb{E}\sum_{k\geq 1} |f(Z_k^x) - f(Z_k^y)| \leq \sum_{k\geq 1} C^{1/q}\alpha^{k/q}\,(1 + \|Z_k^x\|_{q'(a-1)}^{a-1} + \|Z_k^y\|_{q'(a-1)}^{a-1})\ (1 + |x|^p + |y|^p)\ ,$$

where we used the notation $\|\cdot\|_p = (\mathbb{E}[|\cdot|^p])^{1/p}$ and $q' = q/(q-1)$. Since $q'(a-1) \leq s$, we can apply the Jensen Inequality and (ii) to bound $\|Z_k^x\|_{q'(a-1)}$ and $\|Z_k^y\|_{q'(a-1)}$. We obtain, for some constant $c > 0$

$$\mathbb{E}\sum_{k\geq 1} |f(Z_k^x) - f(Z_k^y)| \leq c\ (1 + |x|^b + |y|^b)\ .$$

with $b = p + r(a-1)$. Since we assumed $s > b$, using (a), the right-hand side of the previous display is integrable in y with respect to μ and we get

$$\sum_{k\geq 1} |\mathbb{E}[f(Z_k^x)] - \mu(f)| \leq \int \mathbb{E} \sum_{k\geq 1} |f(Z_k^x) - f(Z_k^y)|\, \mu(\mathrm{d}y) \leq c'\,(1 + |x|^b)\ .$$

Hence we may define the real-valued function

$$\hat{f}(x) = \sum_{k\geq 1} \{\mathbb{E}[f(Z_k^x)] - \mu(f)\}\ ,$$

which is the solution of the Poisson equation $f(x) - \mu(f) = \hat{f}(x) - P\hat{f}(x)$ and satisfies

$$\sup_{x\in\mathcal{X}} (1 + |x|^b)^{-1} |\hat{f}(x)| < \infty\ . \tag{60}$$

Hence the decomposition

$$\frac{1}{n}\sum_{k=1}^{n} \{f(Z_k^x) - \mu(f)\} = \frac{1}{n}\sum_{k=1}^{n} \{\hat{f}(Z_k^x) - P\hat{f}(Z_k^x)\} = \frac{1}{n}\sum_{k=1}^{n} X_k + \frac{1}{n}\{P\hat{f}(x) - P\hat{f}(Z_n^x)\}\ ,$$

where $X_k = \hat{f}(Z_k^x) - P\hat{f}(Z_{k-1}^x)$, $k \geq 1$. Observe that $(X_k)_{k\geq 1}$ is a sequence of martingale increments. By the Jensen Inequality, we have $\mathbb{E}[|P\hat{f}(Z_n^x)|^{s/b}] \leq \mathbb{E}[|\hat{f}(Z_{n+1}^x)|^{s/b}]$ and by (60) and (ii), $\sup_{n\geq 1} \mathbb{E}[|\hat{f}(Z_{n+1}^x)|^{s/b}] < \infty$. Since $s/b > 1$, by the Markov Inequality and Borel-Cantelli's lemma, this implies that $P\hat{f}(Z_n^x)/n \to 0$ *a.s.* We also get that $\sup_{k\geq 1} \mathbb{E}[|X_k|^{s/b}] < \infty$ and, by Lemma 5 $\sum_{k=1}^{n} X_k/n \to 0$ *a.s.* This proves (b).

We conclude with the proof of (c). Using (b) we may replace U_k by $U_k - m$, that is, we assume $m = 0$ without loss of generality. Then $(U_k f(Z_k^x))_{k\geq 1}$ is a sequence of martingale increments. Let $u = (1+\epsilon) \wedge s/a > 1$. We have $\sup_{k\geq 1} \mathbb{E}[|U_k f(Z_k^x)|^u] = \mathbb{E}[|U_1|^u] \sup_{k\geq 1} \mathbb{E}[|f(Z_k^x)|^u] < \infty$ by (ii) and the result follows from Lemma 5. ∎

C. Proofs for Section IV.

1) Proof of Proposition 2: Given any deterministic nonnegative matrix $\mathbf{p} \in [0, Q(\infty)]$, the sequence of covariance matrices $\tilde{Z}_k^{\mathbf{p}} = \tilde{F}_1(\tilde{Z}_{k-1}^{\mathbf{p}})$ where $\tilde{F}_1$ is the second component of (15) with $I = 1$ is a deterministic sequence. From Lemma 1 and Proposition 1–Eq. (27), one get that $\|\tilde{Z}_K^{\mathbf{p}} - \tilde{Z}_K^{\mathbf{q}}\| \leq \alpha \|\mathbf{p} - \mathbf{q}\|$ for $\alpha \in (0,1)$ and K large enough. Hence, by the fixed point theorem, $\tilde{Z}_k^{\mathbf{p}}$ converges to a limit P_{R} defined as the unique solution in $[0, Q(\infty)]$ of the equation $P = \tilde{F}_1(P)$ which is the discrete algebraic Riccati equation (43). This amounts to say that the invariant distribution μ defined in Theorem 2 coincides with $\delta_{P_{\mathrm{R}}}$. It remains to show that Equation (43) has no solutions outside $[0, Q(\infty)]$. Indeed, assume that $\mathbf{p}$ is a solution of (43). Consider the state equations (11) where it is assumed that $X(0) \sim \mathcal{N}(0, \mathbf{p})$. By the very nature of the Kalman filter, the covariance matrix P_k satisfies

$$P_k \le \mathbb{E}\left[X_k X_k^{\mathrm{T}}\right] = \mathrm{e}^{-T_k A}\mathbf{p}\mathrm{e}^{-T_k A^{\mathrm{T}}} + Q(T_k) < \mathrm{e}^{-T_k A}\mathbf{p}\mathrm{e}^{-T_k A^{\mathrm{T}}} + Q(\infty)$$

As $P_k = \mathbf{p}$ for any k, we have $\mathbf{p} \le Q(\infty)$ by taking the limit as $k \to \infty$.

We now consider the invariant distribution ν defined in Theorem 3. This distribution writes $\nu = \nu_X \otimes \delta_{P_{\mathrm{R}}}$, and we shall show that $\nu_X = \mathcal{N}(0, \Sigma)$ where Σ is the unique solution of Equation (44). To that end, we begin by showing that the steady state Kalman filter transition matrix $\Theta = \Phi(1_q - GC)$ with $G = P_{\mathrm{R}} C^{\mathrm{T}}(C P_{\mathrm{R}} C^{\mathrm{T}} + 1_d)^{-1}$ has all its eigenvalues $\{\lambda_i\}$ in the open unit disk. Indeed, taking the limit in (20), we get $P_{\mathrm{R}} = \Theta P_{\mathrm{R}} \Theta^{\mathrm{T}} + \Phi G G^{\mathrm{T}} \Phi^{\mathrm{T}} + Q$. Assuming t_i is an eigenvector of Θ with eigenvalue λ_i, we obtain from this last equation that $(1 - |\lambda_i|^2) t_i^{\mathrm{T}} P_{\mathrm{R}} t_i = t_i^{\mathrm{T}} \Phi G G^{\mathrm{T}} \Phi^{\mathrm{T}} t_i + t_i^{\mathrm{T}} Q t_i > 0$ due to $Q = Q(1) > 0$, hence $|\lambda_i| < 1$. Consequently, the matrix equation (44) has a unique solution $\Sigma = \sum_{n=0}^{\infty} \Theta^n \Phi G G^{\mathrm{T}} \Phi^{\mathrm{T}} (\Theta^{\mathrm{T}})^n$ [28, Chap. 4.2]. When $Z_k = (\underline{Z}_k, \tilde{Z}_k) \in \mathbb{R}^q \times [0, Q(\infty)]$ follows the distribution ν, we have (see (15)) $\underline{Z}_k = \Theta \underline{Z}_{k-1} + \Phi G Y_k$. Recall that $Y_k \sim \mathcal{N}(0, 1_d)$ and is independent with $\underline{Z}_{k-1}$. In these conditions, it is clear that $\underline{Z}_k \sim \mathcal{N}(0, \Sigma)$ when $\underline{Z}_{k-1} \sim \mathcal{N}(0, \Sigma)$. Therefore, $\nu = \mathcal{N}(0, \Sigma) \otimes \delta_{P_{\mathrm{R}}}$ is invariant, and it is the unique invariant distribution. Replacing ν and μ with their values at the right hand sides of (17) and (16), we obtain (42) and (41) respectively. Proposition 2 is proven.

2) Proof of Proposition 3: We assume that the holding times I_n are equal in distribution to I^s (distributed as τ_s) to point out the dependence on s. We also denote the invariant distribution of the Markov chain $(\tilde{Z}_k)$ defined in Section III as μ_s. We begin by proving that μ_s converges weakly to $\delta_{Q(\infty)}$ as $s \to \infty$ (we will use the notation $\mu_s \Rightarrow \delta_{Q(\infty)}$). By Lemma 3 we have $\mathbb{E}[\| \exp(-I^s A)\|^2] \le K\mathbb{E}[\exp(-2aI^s)] = \int \exp(-2ax)\tau_s(\mathrm{d}x)$ with $a > 0$. Given a $K > 0$, we have $\int \exp(-2ax)\tau_s(\mathrm{d}x) = \int_0^K \exp(-2ax)\tau_s(\mathrm{d}x) + \int_K^\infty \exp(-2ax)\tau_s(\mathrm{d}x) \le \tau_s([0, K]) + \exp(-2aK)$. Since τ_s escapes to infinity, $\mathbb{E}[\|\mathrm{e}^{-I^s A}\|^2] \to 0$ as $s \to \infty$, which implies that $\mathrm{e}^{-I^s A} \to 0$ in probability as $s \to \infty$. Moreover, we have

$$\begin{aligned}\|Q(I^s) - Q(\infty)\| &= \left\| \int_{I^s}^{\infty} \exp(-uA) B B^{\mathrm{T}} \exp(-uA^{\mathrm{T}})\, \mathrm{d}u \right\| \\ &\le \|B\|^2 \int_{I^s}^{\infty} \|\exp(-uA)\|^2\, \mathrm{d}u \le K \int_{I^s}^{\infty} \exp(-2ua)\, \mathrm{d}u = (K/2a)\exp(-2aI^s)\end{aligned}$$

hence $Q(I^s) \to Q(\infty)$ in probability as $s \to \infty$. Now, assume that the random variable $\tilde{Z} \in [0, Q(\infty)]$ is distributed as μ_s. Recalling that $\tilde{F}$ is the random iteration function defined as the second component of Equation (15), we have $\|\tilde{F}_{I^s}(\tilde{Z}) - Q(\infty)\| \le K\|\mathrm{e}^{-I^s A}\|^2 + \|Q(I^s) - Q(\infty)\|$, hence $\tilde{F}_{I^s}(\tilde{Z}) \to Q(\infty)$ in probability as $s \to \infty$. As $\tilde{F}_{I^s}(\tilde{Z}) \sim \mu_s$, $\mu_s \Rightarrow \delta_{Q(\infty)}$. Due to the continuity of the $\log\det$ on the compact set $[0, Q(\infty)]$, we have, as $s \to \infty$, $\int \log(1+\mathbf{p})\mathrm{d}\mu_s(\mathbf{p}) \to \int \log(CQ(\infty)C^{\mathrm{T}} + 1_d)$, and (46) results from (16).

Now assume that $Z = (\underline{Z}, \tilde{Z}) \in \mathbb{R}^q \times [0, Q(\infty)]$ follows the invariant distribution ν, and let $(\underline{Z}_1, \tilde{Z}_1) = F_{(I^s,V)}(Z)$, where $F_{(I^s,V)}$ is defined by Equation (15). In particular, we have $\underline{Z}_1 = \Theta(I^s, \tilde{Z})\underline{Z} + \mathrm{e}^{-I^s A}G(\tilde{Z})V$. As $\mathbb{E}[\|\mathrm{e}^{-I^s A}\|^2] \to 0$ and $\tilde{Z} \leq Q(\infty)$, we have $\mathbb{E}[\|\Theta(I^s, \tilde{Z})\|^2] = \mathbb{E}[\|\mathrm{e}^{-I^s A}(I - G(\tilde{Z})C)\|^2] \to 0$ as $s \to \infty$ and $\mathbb{E}[|\mathrm{e}^{-I^s A}G(\tilde{Z})V|^2] \to 0$, hence $\mathbb{E}[|\underline{Z}_1|^2] \to 0$. The third term in the RHS of the Expression (17) of $\xi_{\text{H0:Noise}}$ satisfies

$$\int \mathbf{x}^{\mathrm{T}} C^{\mathrm{T}} \left(C\mathbf{p}C^{\mathrm{T}} + 1_d\right)^{-1} C\mathbf{x}\ \mathrm{d}\nu(\mathbf{x}, \mathbf{p}) \leq \|C\|^2 \int |\mathbf{x}|^2 \mathrm{d}\nu(\mathbf{x}, \mathbf{p}) = \|C\|^2 \mathbb{E}[|\underline{Z}_1|^2] \to 0 \ ,$$

as $s \to \infty$. As $\mu_s \Rightarrow \delta_{Q(\infty)}$, the second term in the RHS of (17) converges to $-\mathrm{tr}[CQ(\infty)C^{\mathrm{T}}(CQ(\infty)C^{\mathrm{T}} + 1_d)^{-1}]$, which achieves the proof of Proposition 3.

3) Proof of Proposition 4: We begin with *(i)*. In the scalar case, the covariance update equation (13) writes

$$P_{k+1} = \tilde{F}^a_{I_{k+1}}(P_k) = \mathrm{e}^{-2aI_{k+1}} \left(\frac{P_k}{P_k + 1} - Q(\infty)\right) + Q(\infty) \ . \tag{61}$$

Given a sequence of holding times $(I_k)_{k\geq 1}$ and two positive numbers $a_1 \geq a_2$, consider the two Markov chains $\tilde{Z}^{\mathbf{p}}_{a_i,k} = \tilde{F}^{a_i}_{I_k}(\tilde{Z}^{\mathbf{p}}_{a_i,k-1})$ for $i = 1, 2$, both starting at the same value $\mathbf{p} = Q(\infty)$. Let $f(\mathbf{p}) = \mathbf{p}/(\mathbf{p}+1) - Q(\infty)$. As $f(Q(\infty)) < 0$ and $0 < \exp(-2a_1 I_1) \leq \exp(-2a_2 I_1)$, it is clear that $\tilde{Z}^{\mathbf{p}}_{a_1,1} \geq \tilde{Z}^{\mathbf{p}}_{a_2,1}$. Assume that $\tilde{Z}^{\mathbf{p}}_{a_1,k-1} \geq \tilde{Z}^{\mathbf{p}}_{a_2,k-1}$. As $f(\mathbf{p})$ is negative and increasing for $\mathbf{p} \in [0, Q(\infty)]$ and $0 < \exp(-2a_1 I_k) \leq \exp(-2a_2 I_k)$, we have $\tilde{Z}^{\mathbf{p}}_{a_1,k} = \exp(-2a_1 I_k) f(\tilde{Z}^{\mathbf{p}}_{a_1,k-1}) + Q(\infty) \geq \exp(-2a_2 I_k) f(\tilde{Z}^{\mathbf{p}}_{a_2,k-1}) + Q(\infty) = \tilde{Z}^{\mathbf{p}}_{a_2,k}$.

From Theorem 2, both the chains $\tilde{Z}^{\mathbf{p}}_{a_1,k}$ and $\tilde{Z}^{\mathbf{p}}_{a_2,k}$ have unique invariant distributions μ_1 and μ_2 respectively, and by repeating the arguments of the proof of Theorem 2,

$$\frac{1}{N}\sum_{k=0}^{N-1} \log\left(1 + \tilde{Z}^{\mathbf{p}}_{a_i,k}\right) \xrightarrow[N\to\infty]{} \int \log(1+\mathbf{p})\,\mathrm{d}\mu_i(\mathbf{p}) \quad \text{for } i = 1, 2, \quad \textit{a.s.}$$

As $\tilde{Z}^{\mathbf{p}}_{a_1,k} \geq \tilde{Z}^{\mathbf{p}}_{a_2,k}$ for all k, by passing to the limit we have $\int \log(1+\mathbf{p})\mathrm{d}\mu_1(\mathbf{p}) \geq \int \log(1+\mathbf{p})\mathrm{d}\mu_2(\mathbf{p})$. As $\xi_{\text{H0:Signal}} = 0.5\left(Q(\infty) - \int \log(1+\mathbf{p})\mathrm{d}\mu(\mathbf{p})\right)$ in the scalar case (see Expression (16)), this error exponent decreases with a.

We now show that $\lim_{a\to 0} \xi_{\text{H0:Signal}} = Q(\infty)/2$. Assume that $\tilde{Z} \in [0, Q(\infty)]$ has the invariant distribution that we denote μ_a. From Eq. (61), we have $\mathbb{E}[\tilde{Z}] = \mathbb{E}[\tilde{F}^a_I(\tilde{Z})] = \mathbb{E}[\mathrm{e}^{-2aI}]\left(\mathbb{E}\left[\frac{\tilde{Z}}{\tilde{Z}+1} - Q(\infty)\right]\right) + Q(\infty)$ which yields

$$\mathbb{E}\left[\frac{\tilde{Z}^2 + (1 - \mathbb{E}[\mathrm{e}^{-2aI}])\tilde{Z}}{\tilde{Z}+1}\right] = Q(\infty)(1 - \mathbb{E}[\mathrm{e}^{-2aI}]) \ .$$

As $\tilde{Z} \leq Q(\infty)$, we have $\mathbb{E}\left[\frac{\tilde{Z}^2}{Q(\infty)+1}\right] \leq \mathbb{E}\left[\frac{\tilde{Z}^2}{\tilde{Z}+1}\right] \leq Q(\infty)(1-\mathbb{E}[\mathrm{e}^{-2aI}])$. By the dominated convergence theorem, $\mathbb{E}[\exp(-2aI)] \to_{a\to 0} 1$, therefore $\mathbb{E}[\tilde{Z}^2] \to 0$ as $a \to 0$. It follows that μ_a converges weakly to δ_0 as $a \to 0$, therefore $\int \log(1+p)\mathrm{d}\mu_a(p) \to 0$. Hence $\lim_{a\to 0} \xi_{\text{H0:Signal}} = Q(\infty)/2$.

The limit as $a \to \infty$ is obtained by applying Proposition 3.

To show *(ii)*, the argument is similar to the one used above to show that $\xi_{\text{H0:Signal}}$ decreases as a increases.

We now prove *(iii)*. Consider the Markov chain $\tilde{Z}_k = \tilde{F}^a_{I_k}(\tilde{Z}_{k-1})$ where $\tilde{F}^a_{I_k}$ is given by (61). Assuming that $\tilde{Z}_k$ has the invariant distribution μ, we have

$$\begin{aligned}
\mathbb{E}\left[\tilde{Z}_k\right] &= \mathbb{E}\left[\mathrm{e}^{-2aI_k}\right]\left(\mathbb{E}\left[\frac{\tilde{Z}_k}{\tilde{Z}_k+1}\right] - Q(\infty)\right) + Q(\infty) \\
&\leq \mathrm{e}^{-2a}\left(\mathbb{E}\left[\frac{\tilde{Z}_k}{\tilde{Z}_k+1}\right] - Q(\infty)\right) + Q(\infty) \\
&\leq \mathrm{e}^{-2a}\frac{\mathbb{E}[\tilde{Z}_k]}{\mathbb{E}[\tilde{Z}_k]+1} + Q(\infty)\left(1-\mathrm{e}^{-2a}\right) \\
&= h\left(\mathbb{E}[\tilde{Z}_k]\right)
\end{aligned}$$

where the first inequality is due to the convexity of e^{-2ax} in conjunction with $\tilde{Z}_k \leq Q(\infty)$ with probability one, and the second inequality is due to the concavity of $x/(1+x)$. If we choose $\tau = \delta_1$, then the corresponding invariant distribution is δ_{P_R} where P_R is the unique solution of the equation $h(p) = p$ (see Proposition 2). As $h(p) - p$ is decreasing, $\int p\,\mathrm{d}\mu(p) = \mathbb{E}[\tilde{Z}_k] \leq P_\text{R}$. As $\log$ is an increasing concave function, the error exponent satisfies

$$\begin{aligned}
\xi_{\text{H0:Signal}} &= \frac{1}{2}\left(Q(\infty) - \int \log(p+1)\,\mathrm{d}\mu(p)\right) \\
&\geq \frac{1}{2}\left(Q(\infty) - \log\left(\int p\,\mathrm{d}\mu + 1\right)\right) \geq \frac{1}{2}\left(Q(\infty) - \log(P_\text{R}+1)\right)
\end{aligned}$$

which achieves the proof of Proposition 4.

D. Heuristic calculations for small VS large SNR discussion

In the scalar case, the Kalman recursions (12)-(13) write

$$\widehat{X}_{n+1} = \frac{e^{-aI_{n+1}}}{P_n+1}\widehat{X}_n + e^{-aI_{n+1}}\frac{P_n}{P_n+1}Y_n \tag{62}$$

$$P_{n+1} = e^{-2aI_{n+1}}\frac{P_n}{P_n+1} + q(1-e^{-2aI_{n+1}}) \tag{63}$$

We assume that vector $(\widehat{X}_n, P_n)$ has the invariant distribution ν. In order to obtain (53) and (54), we study the asymptotic behavior of

$$\xi_{\text{H0:Noise}} = \frac{1}{2}\mathbb{E}\left[\log(P_n+1) + \frac{\widehat{X}_n^2 - P_n}{P_n+1}\right] \tag{64}$$

for small and large values of q.

We start with $q \to 0$. Let us expand the RHS of (63) to the order $\mathcal{O}(q)$, by recalling that $P_n \le q$ and by taking the expectations, we obtain $\mathbb{E}[P_n] = \mathbb{E}[P_{n+1}] = \varphi_\tau(a)\mathbb{E}[P_n] + q(1-\varphi_\tau(a)) + o(q)$. Hence $\mathbb{E}[P_n] = q + o(q)$. As $P_n \le q$, we also have $P_n = q + o(q)$. Inserting this back into (63), we obtain $P_n = q + q^2 B_n + o(q^2)$ with

$$q + q^2 B_{n+1} = \mathrm{e}^{-2aI_{n+1}}\left(q + q^2 B_n\right)(1-q) + q(1-\mathrm{e}^{-2aI_{n+1}}) + o(q^2)\ .$$

By identifying the coefficient of q^2 in the two members we get $B_{n+1} = \mathrm{e}^{-2aI_{n+1}}(B_n - 1)$. Taking the expectations and recalling that we are under an invariant distribution, we obtain

$$\mathbb{E}[P_n] = q + q^2\mathbb{E}[B_n] + o(q^2) = q - q^2\frac{\varphi_\tau(a)}{1-\varphi_\tau(a)} + o(q^2)\ .$$

Turning to Equation (62) and developing as above, we have

$$\begin{aligned}\mathbb{E}\left[\widehat{X}_n^2\right] = \mathbb{E}\left[\widehat{X}_{n+1}^2\right] &= \varphi_\tau(a)\mathbb{E}\left[\left(\frac{\widehat{X}_n}{P_n+1}\right)^2\right] + \varphi_\tau(a)\mathbb{E}\left[\left(\frac{P_n}{P_n+1}\right)^2\right] \\ &= \varphi_\tau(a)\mathbb{E}\left[\widehat{X}_n^2\right] + \varphi_\tau(a)q^2 + o(q^2)\end{aligned}$$

hence

$$\mathbb{E}\left[\frac{\widehat{X}_n^2}{P_n+1}\right] = \mathbb{E}\left[\widehat{X}_n^2\right] + o(q^2) = q^2\frac{\varphi_\tau(a)}{1-\varphi_\tau(a)}\ .$$

Similarly, we have

$$\mathbb{E}\left[\frac{P_n}{P_n+1}\right] = \mathbb{E}\left[\left(q+q^2B_n\right)(1-q)\right] + o(q^2) = q - \frac{q^2}{1-\varphi_\tau(a)} + o(q^2)\ .$$

Plugging these expressions into (64) and recalling that $\log(1+x) = x - x^2/2 + o(x^2)$ we obtain (53).

Next we consider $q \to \infty$. It is easily seen from (63) that

$$\frac{q+1}{q} \ge \frac{P_{n+1}+1}{q} \downarrow 1 - \mathrm{e}^{-2aI_{n+1}} \quad \text{as } q \to \infty$$

therefore, by the monotone convergence theorem, $\mathbb{E}[\log(P_n+1)] - \log q \to \mathbb{E}\left[\log(1-\mathrm{e}^{-2aI_1})\right]$. Moreover, we readily have $\mathbb{E}[(P_n+1)/P_n] \to 1$. Using (62), as $P_n \sim q(1-\mathrm{e}^{-2aI_{n-1}})$, we have $\mathbb{E}[\widehat{X}_n^2] \to \varphi(a)$ and $\mathbb{E}[\widehat{X}_n^2/(P_n+1)] \to 0$. Replacing into (64), we obtain (54).

ACKNOWLEDGMENTS

The authors thank the associate editor and the reviewers for their remarks which substantially improved the paper.

REFERENCES

[1] T. Kailath and H.V. Poor, "Detection of stochastic processes," *IEEE Trans. on Information Theory*, vol. 44, no. 6, pp. 2230–2259, Oct. 1998.

[2] S. Misra and L. Tong, "Error exponents for the detection of Gauss-Markov signals using randomly spaced sensors," *IEEE Trans. on Signal Processing*, vol. 56, no. 8, pp. 3385–3396, Aug. 2008.

[3] B. Øksendal, *Stochastic Differential Equations: An Introduction With Applications*, Springer Verlag, 6th edition, 2003.

[4] E. Allen, *Modeling with Itô Stochastic Differential Equations*, vol. 22 of *Mathematical Modelling: Theory and Applications*, Springer, 2007.

[5] M. Micheli, "Random Sampling of a Continuous-Time Stochastic Dynamical System: analysis, State Estimation, and Applications," M.S. thesis, UC Berkeley, 2001.

[6] M. Micheli and M.I. Jordan, "Random sampling of a continuous-time stochastic dynamical system," in *Proc. of the International Symposium on the Mathematical Theory of Networks and Systems*, Univ. of Notre Dame, South Bend, Indiana, USA, Aug. 2002.

[7] Y. Sung, L. Tong, and H.V. Poor, "Neyman-Pearson detection of Gauss-Markov signals in noise: Closed-form error exponents and properties," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1354–1365, Apr. 2006.

[8] J.-F. Chamberland and V.V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. on Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[9] J.-F. Chamberland and V.V. Veeravalli, "How dense should a sensor network be for detection with correlated observations ?," *IEEE Trans. on Information Theory*, vol. 52, no. 11, pp. 5099–5106, Nov. 2006.

[10] S.K. Jayaweera, "Bayesian fusion performance and system optimization for distributed stochastic Gaussian signal detection under communication constraints," *IEEE Trans. on Signal Processing*, vol. 55, no. 4, pp. 1238–1250, Apr. 2007.

[11] Y. Sung, S. Misra, L. Tong, and A. Ephremides, "Signal processing for application-specific ad hoc networks," *IEEE Signal Processing Magazine*, vol. 23, no. 5, pp. 74–83, Sept. 2006.

[12] Y. Sung, S. Misra, L. Tong, and A. Ephremides, "Cooperative routing for distributed detection in large sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 471–483, Feb. 2007.

[13] Po-Ning Chen, "General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds," *IEEE Trans. on Information Theory*, vol. 42, no. 1, pp. 316–323, Jan 1996.

[14] H. Chernoff, "A measure for asymptotic efficiency for tests of a hypothesis based on a sum of observations," *Ann. Math. Statist.*, vol. 23, no. 4, pp. 493–507, 1952.

[15] H. Luschgy, A. Rukhin, and I. Vajda, "Adaptive tests for stochastic processes in the ergodic case," *Stoch. Process. Appl.*, vol. 45, no. 1, pp. 45–59, 1993.

[16] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov random field on nearest-neighbor graph," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hawaii, USA, 2007.

[17] Y. Sung, X. Zhang, L. Tong, and H.V. Poor, "Sensor configuration and activation for field detection in large sensor arrays," *IEEE Trans. on Signal Processing*, vol. 56, no. 2, pp. 447–463, Feb. 2008.

[18] R.K. Bahr, "Asymptotic analysis of error probabilities for the nonzero-mean Gaussian hypothesis testing problem," *IEEE Trans. on Information Theory*, vol. 36, no. 3, pp. 597–607, May 1990.

[19] F. Schweppe, "Evaluation of likelihood functions for Gaussian signals," *IEEE Trans. on Information Theory*, vol. 11, no. 1, pp. 61–70, Jan. 1965.

[20] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, 1991.

[21] P. Brockwell and R. Davis, *Time Series: Theory and Methods*, Springer Series in Statistics, 1991.

[22] H.H. Rosenbrock, *State Space and Multivariable Theory*, Nelson, 1970.

[23] S. Kar, B. Sinopoli, and J. Moura, "Kalman Filtering with Intermittent Observations: Weak Convergence to a Stationary Distribution," submitted to *IEEE Trans. Automatic Control*, [online] `arXiv:0903.2890v1`, Mar. 2009.

[24] S. Kar and J. Moura, "Kalman Filtering with Intermittent Observations: Weak Convergence and Moderate Deviations," [online] `arXiv:0910.4686`, Oct. 2009.

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

[26] L. Guo, "Stability of recursive stochastic tracking algorithms," *SIAM J. on Control and Optimization*, vol. 32, pp. 1195–1125, 1994.

[27] B.D.O. Anderson and J.B. Moore, "Detectability and stabilizability of time-varying discrete-time linear systems," *SIAM J. Control and Optimization*, vol. 19, no. 1, pp. 20–32, 1981.

[28] B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, 1979.

[29] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive algorithms and stochastic approximations*, vol. 22 of *Applications of Mathematics (New York)*, Springer-Verlag, Berlin, 1990, Translated from the French by Stephen S. Wilson.

[30] P. Diaconis and D. Freedman, "Iterated random functions," *SIAM Review*, vol. 41, no. 1, pp. 45–76, Mar. 1999.

[31] Y.S. Chow, "Local convergence of martingales and the law of large numbers," *Ann. Math. Statist.*, vol. 36, no. 2, pp. 552–558, 1965.