ON THE ERROR EXPONENTS FOR DETECTING RANDOMLY SAMPLED NOISY DIFFUSION PROCESSES

Walid Hachem, Eric Moulines, Jamal Najim and François Roueff

CNRS/ Telecom ParisTech (ENST), France.
hachem, moulines, najim, roueff@enst.fr

ABSTRACT

This paper deals with the detection of a continuous random process described by an Ornstein-Uhlenbeck (O-U) stochastic differential equation. Randomly spaced sensors or equivalently a random time sampler which deliver noisy samples of the process are used for this detection. Two types of tests are considered: either **H0** refers to the presence of the noisy O-U process or **H0** refers to the sole presence of noise. For any fixed false alarm probability, it is shown that the Type II error probability decreases to zero exponentially in the number of samples. The exponents, which do not depend on the false alarm probability, are characterized. This work completes former contributions that consider noiseless O-U process with a random sampling or noisy O-U processes with a regular sampling.

***Index Terms*—** Error Exponents, Neyman-Pearson Detection, Ornstein-Uhlenbeck Processes, Sensor Networks, Stability of Markov Processes.

1. INTRODUCTION

Problem Description

Let $(X(t), t \geq 0)$ be the continuous time process defined as the solution of the Ornstein-Uhlenbeck stochastic differential equation

$$dX(t) = -a\, X(t)\, dt + b\, dB(t) \tag{1}$$

where $B(t)$ is a Brownian motion and $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}$ are known[1]. It is assumed that the initial value $X(0)$ is independent of $(B(t), t \geq 0)$ and follows the law $\mathcal{N}(0, c)$ with $c = b^2/(2a)$, which ensures that the solution $(X(t), t \geq 0)$ of (1) is a strict sense stationary process on the positive real line. Let $(T_n)_{n\in\mathbb{N}}$ be a random point process (with $0 = T_0 < T_1 < T_2 < \cdots$) which represents the sampling moments of $X(t)$. Note that the parameter t might be a location parameter instead of being a time parameter, in which case the T_n represent random sensor locations. It will be assumed that the so called holding times $I_n = T_{n+1} - T_n$ are independent and identically distributed (iid) random variables[2]. In particular, when the distribution of the I_n is exponential, (T_n) is a Poisson process. It is furthermore assumed that (T_n) is independent of $(X(t), t \geq 0)$. Solving Equation (1) between T_n and T_{n+1}, it is well known that the process $(X_n) = (X(T_n))_{n\geq 1}$ is characterized by the difference equation

$$X_{n+1} = e^{-aI_n} X_n + U_n, \quad n \in \mathbb{N} \tag{2}$$

with the initial condition $X_0 \sim \mathcal{N}(0, c)$. The "input process" (U_n) is characterized statistically by the fact that the sequence (U_n, I_n) is an iid sequence independent of X_0, and the distribution of U_n conditionally to the holding time I_n is $\mathcal{N}\left(0, c(1 - e^{-2aI_n})\right)$. Note that if the holding times are all equal to a constant, in other words, if the sampling of $X(t)$ is regular, then (X_n) is a Gaussian autoregressive process of order one.

We assume that the sensor's output is corrupted by an iid noise (V_n) such that $V_n \sim \mathcal{N}(0, 1)$ and we denote by $(Y_n)_{n=1,\ldots,N}$ the signal received in a window of size N. We shall consider in turn the two following hypothesis tests:

$$\text{Test 1}: \begin{cases} \mathbf{H0} & : \quad Y_n = X_n + V_n \\ \mathbf{H1} & : \quad Y_n = V_n \end{cases} \qquad \text{for } n = 1, \ldots, N\,. \tag{3}$$

$$\text{Test 2}: \begin{cases} \mathbf{H0} & : \quad Y_n = V_n \\ \mathbf{H1} & : \quad Y_n = X_n + V_n \end{cases} \qquad \text{for } n = 1, \ldots, N \tag{4}$$

Our performance analysis of these tests will be based on the following result. Let $Y_{1:N} = (Y_1, \ldots, Y_N)$ and $T_{1:N} = (T_1, \ldots, T_N)$, and for $i = 0, 1$, let $f_{i,N}(\mathbf{y} \mid \mathbf{t})$ be the density of $Y_{1:N}$ conditionnally to $T_{1:N}$ according to hypothesis **Hi**. Denote by

$$\mathcal{L}_N(Y_{1:N} \mid T_{1:N}) = \frac{1}{N} \log\left(\frac{f_{0,N}(Y_{1:N} \mid T_{1:N})}{f_{1,N}(Y_{1:N} \mid T_{1:N})}\right) \tag{5}$$

the associated Log Likelihood Ratio (LLR). Fix $\varepsilon \in (0, 1)$, and denote by $\beta(\varepsilon)$ the minimum over all tests of the Type II error probability when the false alarm probability α is constrained to satisfy $\alpha \leq \varepsilon$. The minimum $\beta(\varepsilon)$ is attained by a Neyman-Pearson (N-P) test. If the sequence

[1] Eq. (1) is sometimes written $X'(t) = -aX(t) + bN(t)$ where $N(t)$ is a "white noise".

[2] When $\mathbb{E}[I_n] < \infty$, process (T_n) is called a renewal process.

$(\mathcal{L}_N(Y_{1:N} \mid T_{1:N}))$ converges in probability towards a constant ξ as $N \to \infty$ under **H0**, then (see for instance [1])

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

The constant ξ is called the *error exponent* of the N-P test. The study of the behavior of ξ with respect to parameters such as a, the Signal to Noise Ratio (SNR), or the probability distribution of the I_n leads to interesting guidelines to assess the detector performance or the sensor network dimensioning.

There is a number of papers devoted to the detection of correlated Gaussian signals by means of sensor networks, see *e.g.* the tutorial paper [2]. In this context, contributions [3, 4, 5, 6] study the error exponents of N-P or Bayesian tests. The closest contributions to this paper are [3] and [6]. Sung *et.al.* [3] consider Test 2 above with regularly spaced sensors. Following the approach of [7], they develop the LLR $\mathcal{L}_N$ in terms of an innovation process. Our approach starts from the same idea (see Section 2). In [6], sensor location is random and the detector discriminates among two noiseless O-U processes (Eq. (2)). Due to the noisy character of the received signal, our technique for establishing the existence of ξ and for characterizing this error exponent differs substantially from the one used in [6].

The main results of the paper will be provided in Section 2 along with the main ideas of the proofs. Some implications of these results will be discussed in Section 3. Some numerical illustrations will be also shown in Section 3.

2. MAIN RESULTS

The asymptotic behaviors of the minimum Type II error probabilities for Tests 1 and 2 are provided by the two following theorems:

Theorem 1 *Given two real numbers $a > 0$ and $c > 0$, consider the stochastic process (X_n) described by Equation* (2) *where*

- *The initial value X_0 is independent of the process (U_n, I_n) and follows the probability law $\mathcal{N}(0, c)$.*
- *The sequence (U_n, I_n) is iid with $\mathbb{P}[I_n = 0] < 1$ and the distribution of U_n conditional to I_n is $\mathcal{N}\left(0, c(1 - e^{-2aI_n})\right)$.*

Let (V_n) be an iid sequence independent of $(X_0, (U_n, I_n)_{n\in\mathbb{N}})$ such that $V_n \sim \mathcal{N}(0,1)$. Consider Test 1 described in (3) *where N samples of the sequence $(Y_n, I_n)_{n=1,\dots,N}$ are observed. Then the following assertions hold true:*

1. *For $p \in \mathbb{R}_+$, let $\pi(p, .)$ be the probability distribution of the random variable $\exp(-2aI_1)\left(\frac{p}{p+1} - c\right) + c$. There exists a unique probability measure μ that satisfies the equation*

$$\mu(du) = \int \pi(p, du)\, \mu(dp) \; .$$

Moreover, the support of μ is included in $[0, c]$.

2. *Let $\varepsilon \in (0, 1)$. For a given N, let $\beta_N(\varepsilon)$ be the minimum of the Type II error probabilities over all tests for which the false alarm probability α satisfies $\alpha \leq \varepsilon$. Then*

$$-\frac{1}{N}\log\beta_N(\varepsilon) \xrightarrow[N\to\infty]{} \xi_{\textit{H0:Signal}} = \frac{1}{2}\left(c - \int \log(1+p)\,\mu(dp)\right) \in (0, \infty) \; . \quad (6)$$

Theorem 2 *Assume the setting of Theorem 1 with the difference that the roles of* **H0** *and* **H1** *are interchanged (Test 2 described in* (4)*). Then the following hold true:*

1. *For $\mathbf{u} = (x, p) \in \mathbb{R}\times\mathbb{R}_+$, let $\Pi(\mathbf{u}, .)$ be the probability distribution of the random vector*

$$W^{\mathbf{u}} = \left[\exp(-aI_1)\left(\frac{x}{p+1} + \frac{p}{p+1}Y_1\right), \; \exp(-2aI_1)\left(\frac{p}{p+1} - c\right) + c\right]$$

where it is recalled that $Y_1 \sim \mathcal{N}(0,1)$ and I_1 are independent. There exists a unique probability measure ν on $\mathbb{R}\times\mathbb{R}_+$ that satisfies the equation

$$\nu(d\mathbf{w}) = \int \Pi(\mathbf{u}, d\mathbf{w})\, \nu(d\mathbf{u}) \; . \quad (7)$$

2. *The minimum of the Type II error probabilities satisfies*

$$-\frac{1}{N}\log\beta_N(\varepsilon) \xrightarrow[N\to\infty]{} \xi_{\textit{H0:Noise}} = \frac{1}{2}\left(\int \log(1+p)\,\mu(dp) - \int \frac{p}{p+1}\,\mu(dp) + \int \frac{x^2}{p+1}\,\nu(dx, dp)\right) \in (0, \infty) \quad (8)$$

where the law μ is the one described in the statement of Theorem 1. It coincides with the marginal law $\nu(\mathbb{R}, .)$.

Theorems 1 and 2: Sketch of the Proof

In the contexts of both Theorems 1 and 2, we have to prove that $\mathcal{L}_N$ converges in probability to constant values which will

be $\xi_{\text{H0:Signal}}$ and $\xi_{\text{H0:Noise}}$ respectively.
Denote by f_{V} (resp. f_{S}) the density of $Y_{1:N}$ conditionnally to $T_{1:N}$ for the model $Y_n = V_n$ (resp. for the model $Y_n = X_n + V_n$). Hence, Test 1 assumes $f_{0,N} = f_{\text{S}}$ and $f_{1,N} = f_{\text{V}}$ while Test 2 assumes the opposite. Let us derive the expressions of these two densities. Obviously f_{V} is the standard multivariate Gaussian density. Considering f_{S}, we have

$$f_{\text{S}}(Y_{1:N} \mid T_{1:N}) = \prod_{n=1}^{N} f_{\text{S}}\left(Y_n \mid (\vec{Y}_{n-1}, \vec{T}_n)\right)$$

where we recall that $\vec{Y}_{n-1} = (Y_1, \ldots, Y_{n-1})$ and $\vec{T}_n = (T_1, \ldots, T_n)$. The conditional densities at the right hand side of this equation are Gaussian, and write

$$f_{\text{S}}\left(Y_n \mid (\vec{Y}_{n-1}, \vec{T}_n)\right) = \frac{1}{\sqrt{2\pi\Delta_n^2}} \exp\left(-\frac{(Y_n - \widehat{Y}_n)^2}{2\Delta_n^2}\right)$$

where $\widehat{Y}_n = \mathbb{E}\left[Y_n \mid \vec{Y}_{n-1}, \vec{T}_n\right]$ is the mean of Y_n conditional to its "past" and $\Delta_n^2 = \mathbb{E}\left[(Y_n - \widehat{Y}_n)^2 \mid \vec{T}_n\right]$ is the innovation variance of the model $Y_n = X_n + V_n$.
As is well known, these two quantities can be calculated recursively with the help of the Kalman filter equations. Recall that the received signal Y_n is decribed by the state equations

$$\begin{aligned} X_{n+1} &= e^{-aI_n} X_n + U_n \\ Y_n &= X_n + V_n \ . \end{aligned}$$

Defining $\widehat{X}_n$ and P_n as $\widehat{X}_n = \mathbb{E}[X_n \mid (\vec{Y}_{n-1}, \vec{T}_n)]$ and $P_n = \mathbb{E}[(X_n - \widehat{X}_n)^2 \mid \vec{T}_n]$, the Kalman recursions that give $(\widehat{X}_{n+1}, P_{n+1})$ in terms of $(\widehat{X}_n, P_n)$ are provided by *e.g.* [8, Prop. 12.2.2]:

$$\widehat{X}_{n+1} = \frac{e^{-aI_n}}{P_n + 1}\widehat{X}_n + e^{-aI_n}\frac{P_n}{P_n + 1}Y_n \tag{9}$$

$$P_{n+1} = e^{-2aI_n}\frac{P_n}{P_n + 1} + Q_n \tag{10}$$

where we put $Q_n = \mathbb{E}[U_n^2 \mid I_n] = c(1 - e^{-2aI_n})$. The conditional mean and variance $\widehat{Y}_n$ and Δ_n^2 are then given by

$$\widehat{Y}_n = \mathbb{E}\left[X_n + V_n \mid \vec{Y}_{n-1}, \vec{T}_n\right] = \widehat{X}_n \tag{11}$$

$$\Delta_n^2 = \mathbb{E}\left[(X_n + V_n - \widehat{X}_n)^2 \mid \vec{T}_n\right] = P_n + 1 \ . \tag{12}$$

Using these results, the LLR (5) writes in the setting of Theorem 1 (where $f_{0,N} = f_{\text{S}}$ and $f_{1,N} = f_{\text{V}}$)

$$\mathcal{L}_N = -\frac{1}{2N}\sum_{n=1}^{N}\log\Delta_n^2 - \frac{1}{2N}\sum_{n=1}^{N}\frac{(Y_n - \widehat{Y}_n)^2}{\Delta_n^2} + \frac{1}{2N}\sum_{n=1}^{N}Y_n^2 \tag{13}$$

where the $(\widehat{Y}_n, \Delta_n^2)$ are given by Equations (9)-(12).
To prove Theorem 1, we have to study the asymptotic behavior of $\mathcal{L}_N$ assuming the conditional density of $Y_{1:N}$ is f_{S}. To that end, some results pertaining to the asymptotic behavior of Markov chains are used. Consider for instance the first term at the right hand side of Eq. (13), denoted as $\chi_{1,N}$. Recall that $\Delta_n^2 = P_n + 1$ where P_n is decribed by the recursion (10). By (10) the sequence (P_n) forms a homogeneous Markov chain. The asymptotic behavior of $\chi_{1,N}$ is intimately related with the *stability* (or ergodicity) of the chain (P_n). Similarly, to prove Theorem 2, we need to establish the convergence in probability of $-\mathcal{L}_N$ towards a constant and characterize this constant, the conditional density of $Y_{1:N}$ being this time f_{V}. In this case also, the Kalman recursion (9)-(10) generates a homogeneous Markov chain whose stability will be established.
The asymptotic behavior of these Markov chains is the core of our proof. Eventually, in the setting of Theorem 2 we show that the $\mathbb{R} \times \mathbb{R}_+$-valued Markov chain $(\widehat{X}_n, P_n)$ given by (9)-(10) with Y_n iid $\sim \mathcal{N}(0, 1)$ is stable and its stationary distribution ν is its unique invariant distribution (given as such by Eq. (7)). Let $(\widehat{X}_\infty, P_\infty)$ be a random vector with law ν. We show that the error exponent $\xi_{\text{H0:Noise}}$ writes as

$$\xi_{\text{H0:Noise}} = \frac{1}{2}\left(\mathbb{E}\left[\log\left(1 + P_\infty\right)\right] + \mathbb{E}\left[\frac{\widehat{X}_\infty^2 - P_\infty}{P_\infty + 1}\right]\right) \tag{14}$$

Similarly, the error exponent $\xi_{\text{H0:Signal}}$ provided by Th. 1 is

$$\xi_{\text{H0:Signal}} = \frac{1}{2}\left(c - \mathbb{E}\left[\log\left(1 + P_\infty\right)\right]\right) \tag{15}$$

These equations coincide with Eq. (8) and (6) respectively.

3. DISCUSSION AND NUMERICAL ILLUSTRATION

Comments and consequences of Theorems 1 and 2

We provide here some observations on the influence of the system design parameters on the error exponents. The first parameter we consider is the parameter a which captures the effects of both the "memory" of the O-U process and the mean sensor spacing (assuming w.l.o.g. $\mathbb{E}[I_n] = 1$). Another key parameter is the Signal to Noise Ratio SNR $= \mathbb{E}[X_n^2]$ (recall that $\mathbb{E}[V_n^2] = 1$). Recalling that $X_n = X(T_n)$ and that $X(t)$ is stationary and independent from (T_n) we simply have SNR $= \mathbb{E}[X(T_n)^2] = \mathbb{E}[X(t)^2] = c$. Notice that the error exponents for both Tests 1 and 2 are completely determined by the parameters a and SNR and by the probability law of I_n.
A few remarks are in order. These assertions will not be proven because of lack of space:

1. In the case of a regular sampling ($I_n = 1$), we obtain explicit expressions for $\xi_{\text{H0:Signal}}$ and $\xi_{\text{H0:Noise}}$. Note that the expression of $\xi_{\text{H0:Noise}}$ in this case has been found in [3, Th. 1].

2. If a is large, *i.e.*, the continuous O-U process (Eq. (1)) is weakly correlated and/or the sensors are far apart, we

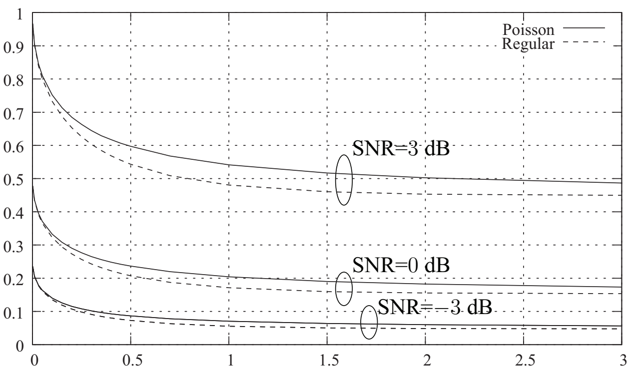

Fig. 1. Test 1: $\xi_{\text{H0:Signal}}$ vs a for SNR $= -3, 0$ and 3 dB

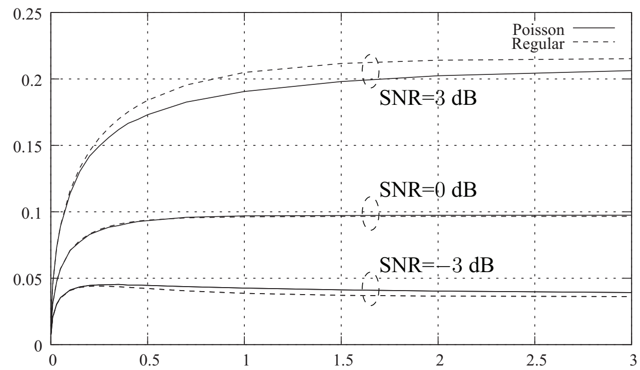

Fig. 2. Test 2: $\xi_{\text{H0:Noise}}$ vs a for SNR $= -3, 0$ and 3 dB

have

$$\xi_{\text{H0:Signal}} \underset{a\to\infty}{\longrightarrow} \frac{1}{2}\left(\text{SNR} - \log(1+\text{SNR})\right),$$
$$\xi_{\text{H0:Noise}} \underset{a\to\infty}{\longrightarrow} \frac{1}{2}\left(\log(1+\text{SNR}) - \frac{\text{SNR}}{\text{SNR}+1}\right).$$

3. In the setting of Theorem 1, the error exponent $\xi_{\text{H0:Signal}}$ decreases as a increases. Moreover, $\lim_{a\to 0}\xi_{\text{H0:Signal}} = \text{SNR}/2$.
   One practical implication of this assertion is the following: from the stand point of the error exponent theory, when **H0** stands for 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.

Numerical illustration

We begin this paragraph by describing the simulation technique. By ergodicity of the Markov process $(\widehat{X}_n, P_n)$, to estimate the error exponents, we simply replace the expectation operators in the equations (14)-(15) above with empirical means taken on $(\widehat{X}_n, P_n)_{n=1,\dots,N}$, for a large snapshot size N.

In Fig. 1, the error exponent $\xi_{\text{H0:Signal}}$ is plotted vs a for SNR $= -3, 0$ and 3 dB. Poisson sampling as well as regular sampling is considered in this figure. Remarks 2 (for $\xi_{\text{H0:Signal}}$) and 3 are confirmed. One interesting observation is that the error exponent with Poisson sampling is better than the error exponent with regular sampling in the context of Test 1.

In Fig. 2, $\xi_{\text{H0:Noise}}$ is plotted vs a also for SNR $= -3, 0$ and 3 dB. We notice that $\xi_{\text{H0:Noise}}$ increases for SNR $= 0$ and 3 dB while it has a maximum with respect to a for SNR $= -3$ dB. This behavior has been established in [3] in the case of a regular sampling. We also notice that Poisson sampling is worse than regular sampling for SNR > 0 dB and better than regular sampling for SNR < 0 dB from the viewpoint of the error exponent.

4. REFERENCES

[1] 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 IT*, vol. 42, no. 1, pp. 316–323, Jan 1996.

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

[3] 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 IT*, vol. 52, no. 4, pp. 1354–1365, Apr. 2006.

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

[5] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov random field on nearest-neighbor graph," in *Proc. ICASSP*, Hawaii, USA, 2007.

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

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

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