Spike detection for calcium activity

Hermine Biermé, Camille Constant, Anne Duittoz, Christine Georgelin

SPIKE DETECTION FOR CALCIUM ACTIVITY

HERMINE BIERMÉ, CAMILLE CONSTANT, ANNE DUITTOZ, AND CHRISTINE GEORGELIN

Abstract. We present in this paper a global methodology for the spike detection in a biological context of fluorescence recording of GnRH-neurons calcium activity. For this purpose we first propose a simple stochastic model that could mimic experimental time series by considering an autoregressive AR(1) process with a linear trend and specific innovations involving spiking times. Estimators of parameters with asymptotic normality are established and used to set up a statistical test on estimated innovations in order to detect spikes. We compare several procedures and illustrate on biological data the performance of our procedure.

Keywords: Autoregressive process, time series, peaks detection, neuronal activity.
2010 Mathematics Subject Classification. Primary: 62M10, 62F12, 62F03, ; Secondary: 92B25

1. Introduction

The neurohormone gonadotropin-releasing hormone (GnRH) controls the reproductive function in males and females mammals. GnRH controls the secretion of two pituitary hormones: luteinizing hormone (LH) and follicle stimulating hormone (FSH) that control the secretion of sexual hormones by gonads (estrogens, progesterone from ovary and androgens from testis) and gametogenesis (follicle development and spermatogenesis). GnRH is synthesized and secreted by a small number of neuroendocrine neurons, the so-called GnRH neurons. Studies on large animals such as sheep, allowed serial blood sampling from portal vessels in conscious animals and revealed that GnRH secretion was pulsatile and that every LH pulse was preceded by a GnRH pulse (Levine et al. 1982 [24], Clarke et al. 1982 [9], Caraty and al. in 1982 [8]).

Understanding the mechanisms underlying this pulsatility of secretion is one key-clue for the comprehension of infertility linked to hypothalamic dysfunctions. GnRH-neurons are dispersed in the anterior hypothalamus and in situ and simultaneous measurement of numerous neurons cannot be achieved. In vitro approaches have allowed to decipher part of some electrophysiological mechanisms [27, 34], showing the existence of periodic synchronization of electrical activity (see Moenter 2010 [27] for a review), intracellular calcium activity correlated with secretion (Constantin et al. 2009 [10], see Wray 2010 [34] for a review). In the present paper, data was obtained from in vitro primary cultures of GnRH neurons derived from mouse embryonic nasal placodes (see Constantin et al. 2009 [10] for biological methodology). In a previous study, Georgelin et al ([16]) showed the existence of a paracrine/ autocrine regulation by endogenous released GnRH on intracellular calcium activity in the neuronal population. The application of a GnRH antagonist blocking the paracrine/autocrine loop abolished high synchronization events between GnRH neurons which can be defined by a great percentage of neurons simultaneously exhibiting a peak of intracellular calcium (calcium event at a time t). This peaks have a duration of several seconds (3-8) and reflect a sustained increase in intracellular calcium. One crucial step in analyzing data is to correctly detect calcium peaks.

Fluctuations on ($[Ca^{2+}]_i$) are recorded as fluorescence signals captured at a given frequency (here 1 acquisition per second). This fluorescence is directly proportional to the amount of intracellular calcium. However, the intensity of fluorescence may present baseline variations due to methodological causes (fading) or environmental causes (pH, temperature...) but also to high frequency and low amplitude variation in intracellular calcium due synaptic input from neurons in the network. It is then important to be able to isolate the real signal from the noise. Fluorescence was quantified at each unit of time as the mean fluorescence intensity measured in each neuron (cell body) identified as a GnRH neuron. This leads to trajectories like the one showed in Figure 1.

The aim of this paper is to improve the detection of so-called calcium events. A $[Ca^{2+}]_i$ peak was detected at time t if the fluorescence value at t was greater or equal to the average value of the 5 previous points plus c_f times the standard deviation of five previous point. This coefficient c_f is empirically fixed by the biologists to separate "true calcium events" of so-called "noisy calcium events" regardless the observed neuron. Because the decision of being a calcium event is strongly linked to this coefficient c_f and will deeply influence the

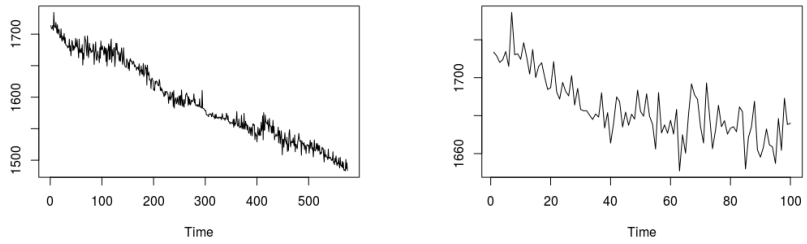


FIGURE 1. Neuronal trajectory from biological data: complete trajectory at left, first 100 values at right.

synchronization analysis, we want to have a robust method to decide whether a peak is an event or not and that could take in account the variability in the fluorescence response. Therefore we proposed a model for the calcium signal and two tests to determine the pics and we will compare them in Section 4.

Inspired by continuous Integrate-And-Fire models (see [7] or [17]), we propose a discrete stochastic model for single recording trajectories. Such a discretization procedure allows us to identify the model with an autoregressive process of order 1 (AR(1)) with a linear trend. Such a model can also be viewed as a linear trend with autoregressive errors and we can use main results developed for time series [5] and especially AR(1) processes. Spiking times are then described through the innovation process.

Several results were proposed to estimate the linear or autoregressive coefficients [30, 32]. We compare here the two main strategies: the first one based on a global contrast function and the second one (called two step estimation) based on a first estimation of the trend followed by the regression coefficient estimate on the detrended series in the spirit of [31]. Then this allows us to get access to estimated innovations, corresponding to a Gaussian mixture with two components with identical standard deviation. Therefore we can use classical mixture estimation [26] and devoted package (`fitgmdist` in Matlab and `Mclust` in R [14]) to estimate parameters and the two classes of times: those that are spiking times and those that are not. Hence we propose several tests for spiking times and compare their sensitivity/specificity on simulated trajectories and show their relevance compared to the initial strategy.

The paper is organized as follows. In Section 2 we describe the parametric model and explain the role of each parameters in link with our biological context. We set and study estimators in Section 3. We obtain explicit asymptotic variances and compare two strategies on some numerical simulations. Section 4 is devoted to our different test strategies, whose efficiency is compared using ROC curves on simulated data. We conclude this paper by several examples from a biological experiment. Some numerical illustrations and technical proofs are postponed to an Appendix section.

2. Model

2.1. **Discrete model.** In our biological context, the fluctuation of the calcium concentration is due on the one hand to natural clearance and on the other hand to sudden supplies released from the Reticulum. But ionic channels near the membrane of neurons can provide puffs of calcium and one has to consider also that indirect measurement can provide experimental noise. Moreover, a lot of experiments exhibit a more or less linear trend (see Section 5) probably due to the response fatigue of the fluorescence.
In a continuous setting, several models for spiking neurons like stochastic leaky integrate-and-fire models [7] have been proposed. In their simplest form, they are solutions of the diffusion equation (see [20]):

$$dX_t = \frac{1}{\tau}(\mu(t) - X_t)dt + \lambda dN_t + \sigma dW_t,$$

where μ is a time inhomogeneous input, W a Brownian motion and N an homogeneous Poisson process indepedent from W. We refer to [12] for a complete review on parameter estimation in the case $\lambda = 0$. Note that in general this equation is used to model the voltage potential for which one can extract electrophysiological events and restrict the study between consecutive spikes. In our context of calcium signals we are precisely

interested in detecting spiking activity (see [1]). Since our observed data are recorded in a discrete sampling with a constant step time and in order to take every phenomenon into account, we consider a kind of discrete version obtained from an infill scheme for this model (see Section 6.1 for more details). Hence we assume to observe the $n \geq 1$ points of the following autoregressive process with linear trend:

$$\forall k \in \{0, ..., n-1\}, \quad X_{n,k+1} = \phi X_{n,k} + a + bk/n + \lambda U_{k+1} + \sigma \varepsilon_{k+1} \tag{1}$$

where

- $\phi \in [0, 1)$ is the coefficient of the autoregressive process, linked to the clearance of the calcium;
- $a \in \mathbb{R}$ and $b \in \mathbb{R}$ correspond to the coefficients of the linear drift.
- $(U_k)_{k\in\{1,...,n\}}$ represents the jump times (instants of release from the Reticulum) : it is a family of iid random variables with common distribution $\mathcal{B}(\nu)$ for some $\nu \in [0, 1)$, the jump rate; the jump size is given by $\lambda \in (0, +\infty)$.
- $(\varepsilon_k)_{k\in\{1,...,n\}}$ represents the experimental noise or stochastic flux from the channels at the membrane with ratio given by $\sigma \in (0, +\infty)$: it is a family of iid random variables with common distribution $\mathcal{N}(0, 1)$.

Let us remark that we have simplified the Poisson process giving spiking times by a simple Bernoulli sequence. This approximation may be justify when the lapse time of acquisition is very short. Actually, considering an homogeneous Poisson point process of intensity $\mu > 0$, the interspiking duration is given by iid exponential random variables of parameter μ such that their entire upper part is given by iid geometrical random variables of parameter $\nu = 1 - e^{-\mu}$.

We give an illustration of a typical realization of this sequence in Figure 2. In order to understand the effect of the parameters in the the final trajectory we have also plotted the purely jumps one (case without noise nor drift) and the drifted jumps ones (only noise is removed). The initial value X_0 has been chosen as the experimental one observed in Figure 1. Let us emphasize that due to both regression coefficient and noise contributions jumps are not necessarily local maxima of the sample paths.

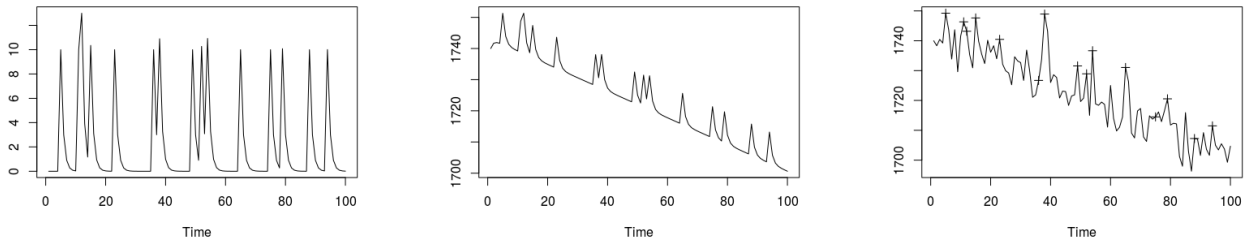


FIGURE 2. One simulation of a trajectory for $\nu = 0.15$, $\lambda = 10$, $\phi = 0.5$, $n = 100$. From left to right: only the jumps innovations ($\sigma = a = b = X_0 = 0$), the same jumps innovations with a drift ($a = 1200$, $b = -30$, $X_0 = 520$, $\sigma = 0$) the same jumps innovations with a drift and a gaussian noise ($a = 1200$, $b = -30$, $X_0 = 520$, $\sigma = 5$): the jumps times are marked with a cross.

In the sequel we will consider the centered innovations $(Z_k)_k$ corresponding to

$$Z_k = \lambda (U_k - \nu) + \sigma \varepsilon_k,$$

in such a way that $(Z_k)_k$ is an iid sequence of centered random variables with common variance $\sigma_Z^2 = \lambda^2 \nu(1 - \nu) + \sigma^2$. As a consequence, denoting $m = a + \lambda\nu$ we may simply rewrite (1) as

$$\forall k \in \{0, ..., n-1\}, \quad X_{n,k+1} = \phi X_{n,k} + m + bk/n + Z_{k+1}. \tag{2}$$

The introduction of a and b in the regression equation contributes to a linear trend of the sequence. It will therefore be convenient to assume furthermore that

$$\forall k \in \{0, ..., n\}, X_{n,k} = Y_k + c_n + dk/n \tag{3}$$

with $(Y_k)_k$ a stationary centered solution of the AR(1) equation

$$Y_{k+1} = \phi Y_k + Z_{k+1}, \tag{4}$$

It follows that the relation between a, b, c_n and d is given by

$$c_n(1-\phi)+d/n = a+\lambda\nu,\ d(1-\phi)=b \text{ iff } c_n = c - \frac{b}{n(1-\phi)^2},\ \text{ with } c=\frac{a+\lambda\nu}{1-\phi} \text{ and } d=\frac{b}{1-\phi}. \tag{5}$$

Actually, under (3) we have

$$\begin{aligned} X_{n,k+1} &= c_n + d\frac{k+1}{n} + Y_{k+1} \\ &= c_n + d\frac{k+1}{n} + \phi Y_k + Z_{k+1} \text{ by (4)} \\ &= \phi X_{n,k} + (1-\phi)c_n + (1-\phi)d\frac{k}{n} + d\frac{1}{n} + Z_{k+1} \text{ by (3)} \\ &= \phi X_{n,k} + (a+\lambda\nu) + b\frac{k}{n} + Z_{k+1}, \end{aligned}$$

by (5), so that we get (2). Let us remark that when $(\tilde{X}_{n,k})_{0\le k\le n}$ is any solution of (2), one will have

$$\tilde{X}_{n,k} = X_{n,k} + \phi^k(\tilde{X}_{n,0} - X_{n,0}), \tag{6}$$

that could be used to extend results in this more general setting using the fact that $|\phi|<1$.

The unknown parameters are therefore given by $(a,b,\phi,\lambda,\nu,\sigma)$ or equivalently by $(c,d,\phi,\lambda,\nu,\sigma)$. Our strategy will consist in first estimating the innovation process $(Z_k)_k$ through the estimation of (m,b,ϕ) for $m=a+\lambda\nu$ and then working on the estimated innovations in order to complete the estimation of the whole parameters and detect times of jumps simply defined as $J_n=\{k\in\{1,\dots,n\}\,;\,U_k=1\}$.

3. Parameter Estimation

3.1. **Preliminary useful results concerning AR(1) processes.** In this section we first give some useful theoretical results when considering $(Y_k)_k$ a stationary centered solution of the AR(1) equation (4) (see [5] for instance). Assuming that $(Z_k)_k$ is an iid sequence with $\mathbb{E}(Z_k)=0$ and $\mathrm{Var}(Z_k)=\sigma_Z^2<+\infty$, one can represent $(Y_k)_k$ as a causal MA(∞) process:

$$Y_k = \sum_{n=0}^{+\infty} \phi^n Z_{k-n}, \tag{7}$$

where the convergence of the series holds in $L^2(\Omega)$. In particular, it follows that $(Y_k)_k$ is a stationary centered second order process with covariance function given by

$$\rho_Y(k) = \mathrm{Cov}(Y_k,Y_0) = \frac{\phi^{|k|}}{1-\phi^2}\sigma_Z^2.$$

In order to compute asymptotic variances for the proposed estimators we will need the following results, that we give under general assumptions on innovations since it may be usefull for other settings.

Theorem 1. *Let $(Z_k)_k$ be a sequence of iid centered random variables with $\mathbb{E}(Z_0^4)<+\infty$ and (Y_k) its associated AR(1) process with autoregression coefficient $\phi\in(-1,1)$ given by (7). Then we have the following asymptotic normality*

$$\frac{1}{\sqrt{n}}\left(\sum_{k=0}^{n-1}Y_k, \sum_{k=0}^{n-1}\frac{k}{n}Y_k, \sum_{k=0}^{n-1}(Y_k^2-\rho_Y(0)), \sum_{k=0}^{n-1}Y_kZ_{k+1}\right) \xrightarrow[n\to+\infty]{d} \mathcal{N}(0,\Sigma_1),$$

with

$$\Sigma_1 = \frac{\sigma_Z^2}{(1-\phi)^2}\begin{pmatrix} 1 & \frac{1}{2} & \frac{\mathbb{E}(Z_0^3)}{(1+\phi)\sigma_Z^2} & 0 \\ \frac{1}{2} & \frac{1}{3} & \frac{\mathbb{E}(Z_0^3)}{2(1+\phi)\sigma_Z^2} & 0 \\ \frac{\mathbb{E}(Z_0^3)}{(1+\phi)\sigma_Z^2} & \frac{\mathbb{E}(Z_0^3)}{2(1+\phi)\sigma_Z^2} & \frac{(\mathbb{E}(Z_0^4)-3\sigma_Z^4)}{(1+\phi)^2\sigma_Z^2}+2\sigma_Z^2\frac{(1+\phi^2)}{(1-\phi)(1+\phi)^3} & \frac{2\sigma_Z^2\phi}{(1+\phi)^2} \\ 0 & 0 & \frac{2\sigma_Z^2\phi}{(1+\phi)^2} & \sigma_Z^2\frac{1-\phi}{1+\phi} \end{pmatrix}.$$

This result implies consistency of the corresponding empirical means but actually we have even stronger convergence as stated in the following corollary.

Corollary 1. *Let $(Z_k)_k$ be a sequence of iid centered rv with $\mathbb{E}(Z_0^4)<+\infty$ and (Y_k) its associated AR(1) process with autoregression coefficient $\phi\in(-1,1)$ given by (7). Then we have the following convergence almost surely and in L^2:*

i) $\frac{1}{n}\sum_{k=0}^{n-1} Y_k \to 0$ *and* $\frac{1}{n}\sum_{k=0}^{n-1} \frac{k}{n} Y_k \to 0$*;*

ii) $\frac{1}{n}\sum_{k=0}^{n-1} Y_k^2 \to \rho_Y(0) = \frac{1}{1-\phi^2}\sigma_Z^2$*;*

iii) $\frac{1}{n}\sum_{k=0}^{n-1} Y_k Z_{k+1} \to 0$ *and* $\frac{1}{n}\sum_{k=0}^{n-1} Y_k Y_{k+1} \to \rho_Y(1) = \frac{\phi}{1-\phi^2}\sigma_Z^2$.

Moreover, denoting $\overline{Y}_n = \frac{1}{n}\sum_{k=0}^{n-1} Y_k$*, one has*

$$\frac{1}{\sqrt{n}}\left(\sum_{k=0}^{n-1}\left[\left(Y_k-\overline{Y}_n\right)^2-\rho_Y(0)\right], \sum_{k=0}^{n-1}\left[\left(Y_k-\overline{Y_n}\right)\left(Y_{k+1}-\overline{Y}_n\right)\right)-\rho_Y(1)\right] \xrightarrow[n\to+\infty]{d} \mathcal{N}\left(0,\Sigma_\rho\right),$$

with

$$\Sigma_\rho = \rho_Y(0)^2\begin{pmatrix} (\eta-3)+2\frac{1+\phi^2}{1-\phi^2} & \phi(\eta-3)+\frac{4\phi}{1-\phi^2} \\ \phi(\eta-3)+\frac{4\phi}{1-\phi^2} & (\eta-3)\phi^2+4\frac{\phi^2}{1-\phi^2}+1+\phi^2 \end{pmatrix},$$

where $\eta = \frac{\mathbb{E}(Z_0^4)}{\sigma_Z^4}$*. Hence, for* $\hat{\phi}_n := \frac{\sum_{k=0}^{n-1}\left(Y_k-\overline{Y_n}\right)\left(Y_{k+1}-\overline{Y}_n\right)}{\sum_{k=0}^{n-1}\left(Y_k-\overline{Y}_n\right)^2}$ *we have*

$$\sqrt{n}\left(\hat{\phi}_n-\phi\right) \xrightarrow[n\to+\infty]{d} \mathcal{N}\left(0,1-\phi^2\right).$$

Proof. It is sufficient to remark that when $(U_k)_{k\in\mathbb{N}}$ is a sequence of L^2 centered random variables with $|\mathrm{Cov}(U_k,U_l)| \le r(|k-l|)$, for all $l,k\in\mathbb{N}$ and r satisfying $\sum_{k\in\mathbb{N}} r(k) < +\infty$ then

$$\frac{1}{n}\sum_{k=0}^{n-1} U_k \underset{n\to+\infty}{\longrightarrow} 0 \text{ and } \frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}U_k \underset{n\to+\infty}{\longrightarrow} 0 \text{ a.s. and in } L^2.$$

Actually, with a slight abuse of notation, let $\overline{U_n} := \frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}U_k$ and note that

$$\begin{aligned}\mathrm{Var}\left(\overline{U_n}\right) &= \frac{1}{n^2}\sum_{k=0}^{n-1}\sum_{l=0}^{n-1}\frac{kl}{n^2}\mathrm{Cov}(U_k,U_l)\\ &\le \frac{C}{n},\end{aligned}$$

with $C = \sum_{k\in\mathbb{Z}} r(|k|)$. This proves that $(\overline{U_n})_n$ converges to 0 in L^2 since $\mathbb{E}(\overline{U_n}) = 0$ by assumption. Moreover, by Markov's inequality one has $\mathbb{P}(|\overline{U_n}| \ge \varepsilon) \le \frac{C}{n\varepsilon^2}$ for $\varepsilon > 0$ ensuring that $(\overline{U_{n^2}})_n$ converges a.s. to 0 according to Borel Cantelli lemma. Now, let $n\in\mathbb{N}^*$ and note that, for $n^2+1 \le l \le (n+1)^2$, we can write

$$\overline{U_l} = \left(\frac{n^2}{l}\right)^2 \overline{U_{n^2}} + \frac{1}{l}\sum_{k=n^2}^{l-1}\frac{k}{l}U_k.$$

It follows that

$$V_n := \sup_{n^2+1\le l\le(n+1)^2} |\overline{U_l}| \le |\overline{U_{n^2}}| + \frac{1}{n^2}\sum_{k=n^2}^{(n+1)^2-1}|U_k|,$$

and then satisfies $\|V_n\|_{L^2} \le \frac{\sqrt{C}}{n} + \sqrt{r(0)}\frac{2n+1}{n^2} \le \frac{4\sqrt{C}}{n}$. Markov's inequality and Borel Cantelli lemma allow again to conclude that $(V_n)_n$ converges a.s. to 0 and therefore $(\overline{U_n})_n$ also. The result for $\frac{1}{n}\sum_{k=0}^{n-1} U_k$ can be even more easily obtained following the same sketch of proof. Now it is enough to remark that assumptions are fulfilled with $U_k = Y_k$ for i) and $U_k = Y_k^2 - \rho_Y(0)$ for ii) according to Lemma 1 Equation (5) (see Appendix). For iii) we can use $U_k = Y_k Z_{k+1}$ with Lemma 1 Equation (3) and using the AR(1) equation (4) we can write

$$\frac{1}{n}\sum_{k=0}^{n-1} Y_k Y_{k+1} = \phi\frac{1}{n}\sum_{k=0}^{n-1} Y_k^2 + \frac{1}{n}\sum_{k=0}^{n-1} Y_k Z_{k+1} \underset{n\to+\infty}{\longrightarrow} \phi\rho_Y(0) = \rho_Y(1) \text{ a.s. and in } L^2.$$

The central limit theorem and ϕ-estimator that follow are classical results (see [5] for instance) that we can recover by Theorem 1. Actually, since $\overline{Y}_n = O_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right)$ the asymptotic normality follows from the fact that

$$\left(\sum_{k=0}^{n-1}\left[Y_k^2 - \rho_Y(0)\right], \sum_{k=0}^{n-1}\left[Y_k Y_{k+1} - \rho_Y(1)\right]\right)$$

is just a linear transformation of $\left(\sum_{k=0}^{n-1}\left[Y_k^2 - \rho_Y(0)\right], \sum_{k=0}^{n-1} Y_k Z_{k+1}\right)$ with corresponding matrix given by $A = \begin{pmatrix} 1 & 0 \\ \phi & 1 \end{pmatrix}$ and we find $\Sigma_\rho = A\Sigma_{3,4}{}^t A$, where $\Sigma_{3,4}$ denotes the extracted matrix of Σ_1 corresponding to lines and rows number 3 and 4, that corresponds to the asymptotic covariance matrix obtained in Proposition 7.3.4. of [5]. Then using delta-method for $g(x,y) = \frac{y}{x}$ we may prove the asymptotic normality of $\hat{\phi}_n$ with asymptotic variance given by $Dg(\rho_Y(0),\rho_Y(1))\Sigma_\rho{}^t Dg(\rho_Y(0),\rho_Y(1))$ where $Dg(\rho_Y(0),\rho_Y(1)) = \frac{1}{\rho_Y(0)}\begin{pmatrix} -\phi & 1 \end{pmatrix}$. □

Remark 1. *Let us mention that these convergences may be strengthen using ergodic properties of $(Y_k)_k$. Actually, according to* [2]*, since the innovation Z_0 has a non trivial absolutely continuous component and $\mathbb{E}(\log(Z_0)^+) < +\infty$, the $AR(1)$ process $(Y_k)_k$ is strong mixing and hence ergodic. Therefore, by the ergodic theorem (see Corollary 9.1.3 of* [13] *for instance), for any measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}(g(Y_0)^2) < +\infty$ we have*

$$\frac{1}{n}\sum_{k=0}^{n-1} g(Y_k) \to \mathbb{E}(g(Y_0)), \text{ almost surely and in } L^2.$$

Let us also note that these results may be extended to non stationary solutions of the $AR(1)$ equation whatever the initial solution is in view of (6)*.*

3.2. Estimation of the innovations $(Z_k)_{k\in\{1,\dots,n\}}$.

3.2.1. *Estimation of $m = a + \lambda\nu$, b and ϕ.* In view of (2), we first derive an estimator for $\theta := (m, b, \phi) = (a - \lambda\nu, b, \phi)$ as a contrast based estimator. To this end let us consider the contrast function, defined for $\tilde{\theta} = (\tilde{m}, \tilde{b}, \tilde{\phi}) \in \mathbb{R}^3$ as

$$M_n(\tilde{\theta}) = \sum_{k=0}^{n-1}\left(X_{n,k+1} - \tilde{\phi}X_{n,k} - \tilde{m} - \tilde{b}\frac{k}{n}\right)^2.$$

The minimizer of this quadratic functional is given by

$$\widehat{\theta}_n^{(1)} := ({}^t A_n(X)A_n(X))^{-1}\,{}^t A_n(X)X = (\widehat{m}_n^{(1)}, \widehat{b}_n^{(1)}, \widehat{\phi}_n^{(1)}) \tag{8}$$

where $A_n(X) = \begin{pmatrix} 1 & 0 & X_{n,0} \\ & \dots & \\ 1 & \frac{n-1}{n} & X_{n,n-1} \end{pmatrix}$ is the random matrix of coefficients. Adapting classical proofs for M-estimators (see [33] for instance), we prove the following results (see Appendix).

Theorem 2. *Let us assume that $(X_{n,k})_{0\le k\le n}$ satisfies* (2) *for some $\theta = (m, b, \phi)$, following from* (3)*, with (Z_k) an iid sequence of centered random variables with $\mathbb{E}(Z_k^4) < +\infty$. Then,*

$$\widehat{\theta}_n^{(1)} := ({}^t A_n(X)A_n(X))^{-1}\,{}^t A_n(X)X \underset{n\to+\infty}{\longrightarrow} \theta \quad a.s. \tag{9}$$

Moreover,

$$\sqrt{n}\left(\widehat{\theta}_n^{(1)} - \theta\right) \underset{n\to+\infty}{\xrightarrow{d}} \mathcal{N}(0, \Sigma_2),$$

where, for $\sigma_Z^2 := Var(Z_0)$, the matrix Σ_2 is given by

$$\Sigma_2 = \begin{pmatrix} m^2\frac{1+\phi}{1-\phi} + 4\sigma_Z^2 & mb\frac{1+\phi}{1-\phi} - 6\sigma_Z^2 & -m(1+\phi) \\ mb\frac{1+\phi}{1-\phi} - 6\sigma_Z^2 & b^2\frac{1+\phi}{1-\phi} + 12\sigma_Z^2 & -b(1+\phi) \\ -m(1+\phi) & -b(1+\phi) & 1-\phi^2 \end{pmatrix},$$

or equivalently, writing $c = \frac{m}{1-\phi}$ and $d = \frac{b}{1-\phi}$, by

$$\Sigma_2 = \begin{pmatrix} c^2\left(1-\phi^2\right) + 4\sigma_Z^2 & cd\left(1-\phi^2\right) - 6\sigma_Z^2 & -c\left(1-\phi^2\right) \\ cd\left(1-\phi^2\right) - 6\sigma_Z^2 & d^2\left(1-\phi^2\right) + 12\sigma_Z^2 & -d\left(1-\phi^2\right) \\ -c\left(1-\phi^2\right) & -d\left(1-\phi^2\right) & 1-\phi^2 \end{pmatrix}.$$

Remark 2. *Let us note that for m and b be fixed, asymptotic variances of $(\hat{m}_n^{(1)})_n$ and $(\hat{b}_n^{(1)})_n$ are increasing as ϕ increases to 1, while it is the opposite for c and d be fixed since in this case $m = c(1-\phi)$ and $b = d(1-\phi)$ become smaller with ϕ.*

We check empirically this result for different values of ϕ in Figure 21, postponed to Section 6.4. As expected, in view of the asymptotic variance, the performances are better for ϕ not too close from 1.

3.2.2. *Estimation of the trend.* Theorem 2 allows us to build an estimator of the parameters $c = \frac{m}{1-\phi}$ and $d = \frac{b}{1-\phi}$ linked with the observed linear trend driven by $c_n = c - \frac{d}{n(1-\phi)}$ and d by considering

$$\hat{c}_n^{(1)} = \frac{\hat{m}_n^{(1)}}{1-\hat{\phi}_n^{(1)}} - \frac{\hat{b}_n^{(1)}}{n(1-\hat{\phi}_n^{(1)})^2} \text{ and } \hat{d}_n^{(1)} = \frac{\hat{b}_n}{1-\hat{\phi}_n}.$$

Since $c_n \to c$, as $n \to +\infty$, we can deduce from Theorem 2 both strong consistency and asymptotic normality.

Proposition 1. *Under the assumptions of Theorem 2*

$$\sqrt{n}\left((\hat{c}_n^{(1)}, \hat{d}_n^{(1)}, \hat{\phi}_n^{(1)}) - (c, d, \phi)\right) \xrightarrow[n\to+\infty]{d} \mathcal{N}(0, \Sigma_3),$$

with

$$\Sigma_3 = \begin{pmatrix} \frac{4\sigma_Z^2}{(1-\phi)^2} & -\frac{6\sigma_Z^2}{(1-\phi)^2} & 0 \\ -\frac{6\sigma_Z^2}{(1-\phi)^2} & \frac{12\sigma_Z^2}{(1-\phi)^2} & 0 \\ 0 & 0 & 1-\phi^2 \end{pmatrix}. \tag{10}$$

Proof. We consider the function from $\mathbb{R}^3$ to $\mathbb{R}^3$ defined $\forall (x,y,z) \in \mathbb{R}^3$ by:

$$g(x,y,z) = \left(\frac{x}{1-z}, \frac{y}{1-z}, z\right),$$

such that

$$\sqrt{n}\left(\left(\widehat{c}_n^{(1)}, \widehat{d}_n^{(1)}, \widehat{\phi}_n^{(1)}\right) - (c,d,\phi)\right) = \sqrt{n}\left(g(\hat{\theta}_n^{(1)}) - g(\theta)\right) \xrightarrow[n\to+\infty]{d} \mathcal{N}\left(0, Dg(\theta)\Sigma_2 {}^t Dg(\theta)\right).$$

by Theorem 2 and the delta-method ([33]), with

$$Dg(\theta) = \frac{1}{1-\phi}\begin{pmatrix} 1 & 0 & c \\ 0 & 1 & d \\ 0 & 0 & 1-\phi \end{pmatrix}.$$

□

Remark 3. *Note that the asymptotic covariance does not depend on c, d and that we get an asymptotic independence between $\left(\widehat{c}_n^{(1)}, \widehat{d}_n^{(1)}\right)$ and $\widehat{\phi}_n^{(1)}$ in contrast with previous results.*

We illustrate this result in Figure 22 (see Section 6.4) and compare for several values of ϕ and σ in Figure 3 below.

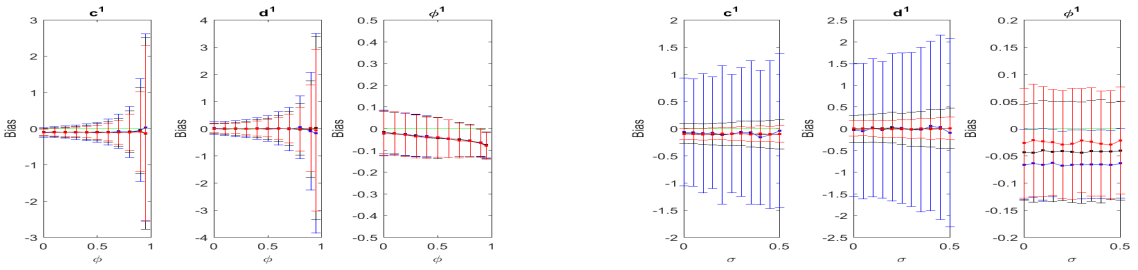


FIGURE 3. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c_n$. On the left: with respect to $\phi \in \{0, 0.1, \ldots, 0.9, 0.95\}$, in red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. On the right with respect to $\sigma \in \{0, 0.05, \ldots, 0.5\}$, in red $\phi = 0.1$, in black $\phi = 0.5$ and in blue $\phi = 0.9$

We can notice that, as expected, $\widehat{\phi}^{(1)}$ does not depend on the noise σ of the signal, but all the estimators $\widehat{c}_n^{(1)}$, $\widehat{d}_n^{(1)}$, $\widehat{\phi}^{(1)}$ depend of ϕ.

Now, a classical way to deal with trend consists in first estimating it and then removing it. When doing so, we have a 2-step estimation procedure by first getting estimators $\hat{c}_n^{(2)}$ and $\hat{d}_n^{(2)}$ given by linear regression of $(X_{n,k})_{0\leq k\leq n}$ and then computing the autoregression coefficient $\hat{\phi}_n^{(2)}$ by considering the autocorrelation function at lag 1 of the linear regression residuals given by

$$\hat{Y}_{n,k} := X_{n,k} - \hat{c}_n^{(2)} - \hat{d}_n^{(2)}k/n = Y_k + (c_n - \hat{c}_n^{(2)}) + (d - \hat{d}_n^{(2)})k/n. \tag{11}$$

More precisely, by (3) together with the fact that $\frac{1}{n+1}\sum_{k=0}^n \frac{k}{n} = \frac{1}{2}$ and $\frac{1}{n+1}\sum_{k=0}^n \left(\frac{k}{n} - \frac{1}{2}\right)^2 = \frac{n+2}{12n}$, we have for $n > 1$,

$$\begin{aligned}\hat{d}_n^{(2)} &= \frac{12n}{n+2} \times \frac{1}{n+1}\sum_{k=0}^n (\frac{k}{n} - \frac{1}{2})\left(X_{n,k} - \overline{X_n}\right) \\ &= \frac{12n}{n+2} \times \frac{1}{n+1}\sum_{k=0}^n (\frac{k}{n} - \frac{1}{2})\left(Y_k - \overline{Y}_n + d\left(\frac{k}{n} - \frac{1}{2}\right)\right)\end{aligned}$$

where $\overline{X_n} = \frac{1}{n+1}\sum_{k=0}^n X_{n,k}$ and $\overline{Y}_n = \frac{1}{n+1}\sum_{k=0}^n Y_k$, while

$$\begin{aligned}\hat{c}_n^{(2)} &= \overline{X_n} - \frac{1}{2}\hat{d}_n^{(2)} \\ &= c_n + \overline{Y}_n - \frac{1}{2}(\hat{d}_n^{(2)} - d).\end{aligned}$$

It follows that

$$\hat{\phi}_n^{(2)} = \sum_{k=0}^{n-1}(\hat{Y}_{n,k+1} - \overline{\hat{Y}}_n)(\hat{Y}_{n,k} - \overline{\hat{Y}}_n) \times \left(\sum_{k=0}^{n}(\hat{Y}_{n,k} - \overline{\hat{Y}}_n)^2\right)^{-1},$$

where $\overline{\hat{Y}}_n = \frac{1}{n+1}\sum_{k=0}^n \hat{Y}_{n,k}$.

In order to compare with our first estimators we can check that they have the same asymptotic covariance matrix.

Proposition 2. *Let us assume that $(X_{n,k})_{0\leq k\leq n}$ satisfies (2) for some $\theta = (m, b, \phi)$, following from (3) with $d = \frac{b}{1-\phi}$, $c = \frac{m}{1-\phi}$, $c_n = c - \frac{d}{n(1-\phi)}$, and $(Z_k)_k$ an iid sequence of centered random variables.*
If $\mathbb{E}(Z_k^4) < +\infty$, then $\mathbb{E}\left(\widehat{c}_n^{(2)}\right) = c_n$ and $\mathbb{E}\left(\widehat{d}_n^{(2)}\right) = d$ with

$$\left(\widehat{c}_n^{(2)}, \widehat{d}_n^{(2)}, \widehat{\phi}_n^{(2)}\right) \underset{n\to+\infty}{\longrightarrow} (c, d, \phi) \text{ a.s. and } \sqrt{n}\left((\hat{c}_n^{(2)}, \hat{d}_n^{(2)}, \hat{\phi}_n^{(2)}) - (c, d, \phi)\right) \underset{n\to+\infty}{\overset{d}{\longrightarrow}} \mathcal{N}\left(0, \Sigma_3\right),$$

with Σ_3 given by (10) of Proposition 1.

Proof. We first remark that $\hat{d}_n^{(2)} - d = A_n \cdot S_{n+1}(Y)$, where

$$S_n(Y) = \frac{1}{n}\left(\sum_{k=0}^{n-1} Y_k, \sum_{k=0}^{n-1}\frac{k}{n}Y_k, \sum_{k=0}^{n-1}(Y_k^2 - \rho_Y(0)), \sum_{k=0}^{n-1} Y_k Z_{k+1}\right),$$

have been introduced in Theorem 1 and $A_n = \left(-6\frac{n}{n+2}, 12\frac{n+1}{n+2}, 0, 0\right) \to A = (-6, 12, 0, 0)$. Similarly, $\hat{c}_n^{(2)} - c_n = (e_1 - \frac{1}{2}A_n) \cdot S_{n+1}(Y)$, with $e_1 = (1, 0, 0, 0)$.
If $\mathbb{E}(Z_k^4) < +\infty$, from Corollary 1 we have $S_{n+1}(Y) \underset{n\to+\infty}{\longrightarrow} 0$ a.s. and in L^2 with $\mathbb{E}\left(S_{n+1}(Y)\right) = 0$, proving the first statement. Moreover, by Theorem 1 we have $\sqrt{n}S_{n+1}(Y) \underset{n\to+\infty}{\overset{d}{\longrightarrow}} \mathcal{N}(0, \Sigma_1)$. Then, one can remark that for $h \in \{0, 1\}$ we have by (11)

$$\frac{1}{n}\sum_{k=0}^{n-1}\left[(\hat{Y}_{n,k+h} - \overline{\hat{Y}}_n)(\hat{Y}_{n,k} - \overline{\hat{Y}}_n) - \rho(h)\right] = \frac{1}{n}\sum_{k=0}^{n-1} Y_{k+h}Y_k - \rho(h) + R_n,$$

where the remaining term R_n tends to 0 a.s. with $\sqrt{n}R_n \to 0$ in probability using Slutsky Theorem and asymptotic normality of $S_{n+1}(Y)$. These lines ensure strong consistency of $\left(\widehat{c}_n^{(2)}, \widehat{d}_n^{(2)}, \widehat{\phi}_n^{(2)}\right)$. For asymptotic

normality, since

$$\sqrt{n}\left(\hat{c}_n^{(2)}, \hat{d}_n^{(2)}, \frac{1}{n+1}\sum_{k=0}^{n} Y_k^2, \frac{1}{n+1}\sum_{k=0}^{n} Y_{k+1}Y_k - (c_n, d, \rho(0), \rho(1))\right) = M_n\sqrt{n}S_{n+1}(Y),$$

for $M_n = \begin{pmatrix} 1+3\frac{n}{n+2} & -6\frac{n+1}{n+2} & 0 & 0 \\ -6\frac{n}{n+2} & 12\frac{n+1}{n+2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \phi & 1 \end{pmatrix} \underset{n\to+\infty}{\longrightarrow} M := \begin{pmatrix} 4 & -6 & 0 & 0 \\ -6 & 12 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \phi & 1 \end{pmatrix}$ and $c_n = c + O(1/n)$, we obtain

$$\sqrt{n}\left(\left(\hat{c}_n^{(2)}, \hat{d}_n^{(2)}, \frac{1}{n}\sum_{k=0}^{n-1}(\hat{Y}_{n,k+1} - \overline{\hat{Y}}_n)(\hat{Y}_{n,k} - \overline{\hat{Y}}_n), \frac{1}{n}\sum_{k=0}^{n-1}(\hat{Y}_{n,k} - \overline{\hat{Y}}_n)^2\right) - (c, d, \rho(0), \rho(1))\right) \underset{n\to+\infty}{\overset{d}{\longrightarrow}} \mathcal{N}(0, M\Sigma_1{}^t M),$$

and the stated result follows by the delta-method. □

We compare these two procedures in Figure 4. It seems that even for $n = 100$ both behave relatively similarly. This is also the case for larger values of n as illustrated in Figure 25. We only remark that the global estimation seems less biased for ϕ as ϕ is increasing than the 2-Step estimation but its standard deviation is a little bit higher. Let us emphasize that we chose initial conditions in order to be close from the chosen line and we saw that the global estimation was more robust with the choice of initial condition. Let us also remark that while the standard deviation for the estimate of ϕ is decreasing according to $(1-\phi^2)^{1/2}$ as ϕ tends to 1, absolute value of bias is increasing. Note also that the estimate of c is naturally biased (asymptotically not biased) but the estimate for d seems not too biased.

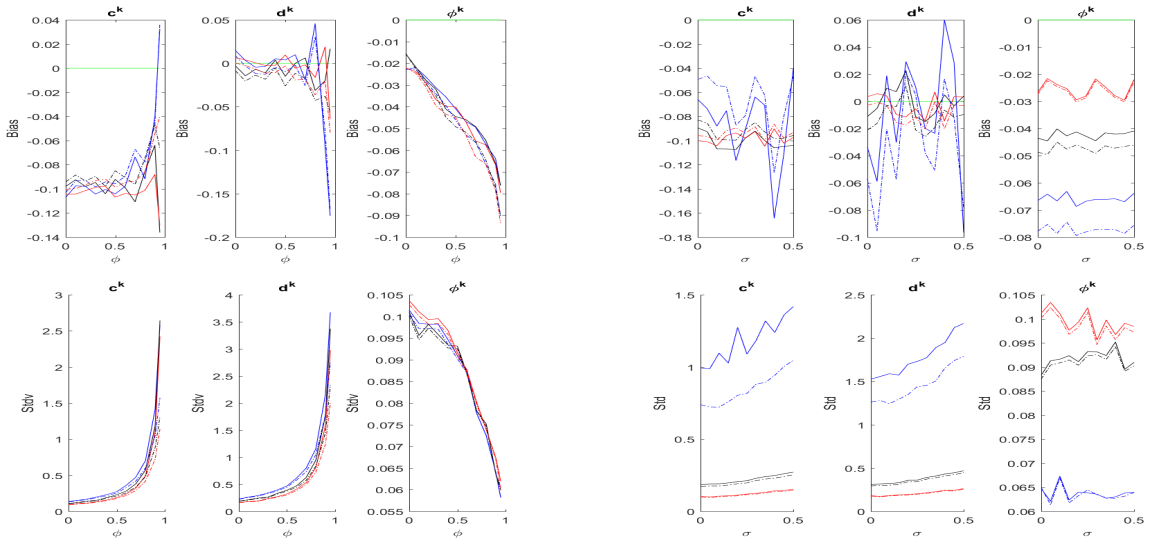


FIGURE 4. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c_n$ with respect to $\phi \in \{0, 0.1, \ldots, 0.9, 0.95\}$ on the left and to $\sigma \in \{0, 0.05, \ldots, 0.5\}$ on the right. Straight line for the global estimation and dashed line for the 2-step estimation. On the left in red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. On the right, in red $\phi = 0.1$, in black $\phi = 0.5$ and in blue $\phi = 0.9$.

For comparison we plot the theoretical values of the standard deviation in Figure 5.

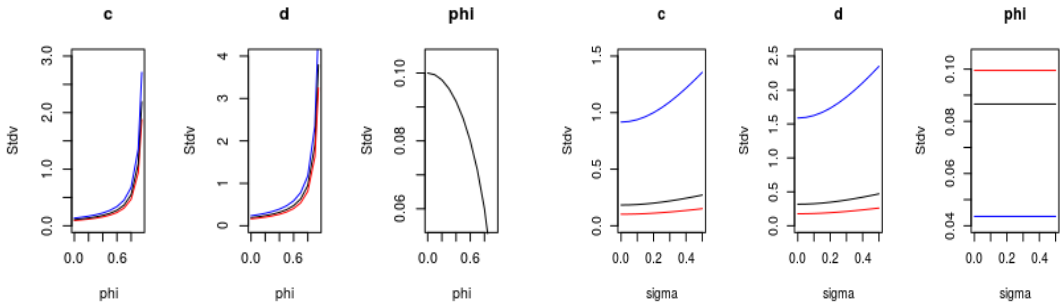


FIGURE 5. Theoretical asymptotic standard deviation for $\nu = 0.3$, $\lambda = 1$, $n = 100$, with respect to $\phi \in \{0, 0.1, \dots, 0.9, 0.95\}$ on the left and to $\sigma \in \{0, 0.05, \dots, 0.5\}$ on the right. On the left in red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. On the right, in red $\phi = 0.1$, in black $\phi = 0.5$ and in blue $\phi = 0.9$.

3.2.3. *Estimation of the innovations.* Theorem 2 and the 2-step estimation of Proposition 2 provide us two ways for the estimations of innovations $(Z_k)_{k\in\{1,\dots,n\}}$. Namely, we can consider for all $k \in \{1, ..., n\}$:

$$\widehat{Z}_k^{(1)} = X_{n,k} - \widehat{\phi}_n^{(1)} X_{n,k-1} - \widehat{m}_n^{(1)} - \widehat{b}_n^{(1)} \frac{k-1}{n} \tag{12}$$

or

$$\widehat{Z}_k^{(2)} = \hat{Y}_{n,k} - \widehat{\phi}_n^{(2)} \hat{Y}_{n,k-1} \text{ with } \hat{Y}_{n,k} \text{ given in (11).} \tag{13}$$

Proposition 3. *Under the assumptions of Theorem 1, we have for $i = 1, 2$ and $\alpha \in (0, 1/4)$,*

$$\forall k \geq 1, \widehat{Z}_k^{(i)} \underset{n\to+\infty}{\longrightarrow} Z_k \text{ a.s.} \quad \text{and } n^\alpha \sup_{1\leq k\leq n} |\widehat{Z}_k^{(i)} - Z_k| \underset{n\to+\infty}{\xrightarrow{\mathbb{P}}} 0.$$

Proof. Let $k \geq 1$, according to (12), choosing n large enough such that $|c_n| \leq 2|c|$ and $1 \leq k \leq n$,

$$\begin{aligned} |\widehat{Z}_k^{(1)} - Z_k| &\leq \left|\phi - \widehat{\phi}_n^{(1)}\right| |X_{n,k-1}| + \left|m - \widehat{m}_n^{(1)}\right| + \left|b - \widehat{b}_n^{(1)}\right| \\ &\leq \left|\phi - \widehat{\phi}_n^{(1)}\right| (|Y_{k-1}| + 2|c| + |d|) + \left|m - \widehat{m}_n^{(1)}\right| + \left|b - \widehat{b}_n^{(1)}\right| \end{aligned}$$

Note that by Theorem 1, this implies that $\widehat{Z}_k^{(1)} \longrightarrow Z_k$ a.s. since $(\widehat{m}_n^{(1)}, \widehat{b}_n^{(1)}, \widehat{\phi}_n^{(1)})$ is a strongly consistent estimator of (m, b, ϕ). Moreover for $\alpha \in (0, 1/4)$, let us choose $\beta \in (1/4, 1/2 - \alpha)$. For $\varepsilon > 0$ we write

$$\begin{aligned} \mathbb{P}(\sup_{0\leq k\leq n-1} |Y_k| > \varepsilon n^\beta) &\leq \mathbb{P}\left(\{|Y_0| > \varepsilon n^\beta\} \cup \left(\bigcup_{k=1}^{n-1} \{|Y_k| > \varepsilon n^\beta, |Y_{k-1}| \leq \varepsilon n^\beta\}\right)\right) \\ &\leq \mathbb{P}\left(\{|Y_0| > \varepsilon n^\beta\} \cup \left(\bigcup_{k=1}^{n-1} \{|Z_k| > (1-|\phi|)\varepsilon n^\beta, |Y_{k-1}| \leq \varepsilon n^\beta\}\right)\right) \\ &\leq \mathbb{P}(|Y_0| > \varepsilon n^\beta) + (n-1)\mathbb{P}(|Z_0| > (1-|\phi|)\varepsilon n^\beta) \\ &\leq \frac{\mathbb{E}(Z_0^4)}{(1-|\phi|)^4 \varepsilon^4} n^{1-4\beta}. \end{aligned}$$

Hence, $\frac{1}{n^\beta} \sup_{0\leq k\leq n-1} |Y_k|$ tends to 0 in probability and

$$n^\alpha \sup_{1\leq k\leq n} |\widehat{Z}_k^{(1)} - Z_k| \leq n^{\alpha+\beta} \left|\phi - \widehat{\phi}_n^{(1)}\right| \times \left[\frac{1}{n^\beta} \sup_{0\leq k\leq n-1} |Y_k|\right] + n^\alpha \left[(2|c| + |d|) \left|\phi - \widehat{\phi}_n^{(1)}\right| + \left|m - \widehat{m}_n^{(1)}\right| + \left|b - \widehat{b}_n^{(1)}\right|\right],$$

that tends to 0 in probability in view of Theorem 2 as $0 < \alpha < \alpha + \beta < 1/2$.
Since $\hat{Y}_{n,k} - Y_k = (c_n - \hat{c}_n^{(2)}) + (d - \hat{d}_n^{(2)})k/n$ we can also write for $1 \leq k \leq n$

$$\widehat{Z}_k^{(2)} - Z_k = [\hat{Y}_{n,k} - Y_k] - \widehat{\phi}_n^{(2)} [\hat{Y}_{n,k-1} - Y_{k-1}] + (\phi - \widehat{\phi}_n^{(2)}) Y_{k-1},$$

in such a way that

$$|\widehat{Z}_k^{(2)} - Z_k| \leq |c_n - \hat{c}_n^{(2)}| + |d - \hat{d}_n^{(2)}| \left(1 + |\widehat{\phi}_n^{(2)}|\right) + |\phi - \widehat{\phi}_n^{(2)}||Y_{k-1}|.$$

This ensures that $\widehat{Z}_k^{(2)} \underset{n\to+\infty}{\longrightarrow} Z_k$ a.s. and that $n^\alpha \sup_{1\le k\le n} |\widehat{Z}_k^{(2)} - Z_k|$ converges in probability for all $\alpha \in (0, 1/4)$ following the same lines as previously. □

Then, in the sequel, we will explicitly use our assumptions on the distribution of the innovations in order to complete estimation of all the parameters. Actually, this may be justified as in [23], when the common distribution of innovations has a density $f(., \theta)$ for some parameter $\theta \in \Theta$, with Θ an open set of $\mathbb{R}^k$, $k \ge 1$. Then, when considering an estimator $\hat{\theta}$ that maximizes the log-likelihood function

$$L_n(\theta) := \sum_{k=1}^{n} \log(f(Z_k, \theta)),$$

one can use an estimator $\hat{\hat{\theta}}$ that maximizes the estimated log-likelihood function

$$\hat{L}_n(\theta) := \sum_{k=1}^{n} \log(f(\widehat{Z}_k, \theta)),$$

with $(\widehat{Z}_k)$ the estimated innovations and obtain consistency and asymptotic normality under regularity assumptions on f (see Corollary of Theorem 4.1 [23]).

3.3. **Estimation of λ, ν and σ^2.** Let us recall that we assume that $Z_k = \lambda(U_k - \nu) + \sigma\varepsilon_k$, for $U \sim \mathcal{B}(\nu)$ and $\varepsilon_k \sim \mathcal{N}(0, \sigma^2)$. This implies that Z_k is a mixture between two Gaussian. More precisely, the density of Z_0 is given by

$$f_{Z_0}(z; \psi) = (1-\nu)\gamma_{-\lambda\nu,\sigma^2}(z) + \nu\gamma_{\lambda(1-\nu),\sigma^2}(z), \text{ for } z \in \mathbb{R}$$

where $\psi := (\nu, \lambda, \sigma)$ and γ_{m,ρ^2} denotes the density of a normal variable of mean m and of variance ρ^2.

Note that Maximum likelihood estimation even in this simple context of two Gaussian components suffers from several computational drawbacks [11, 22] and Fisher matrix are not explicitly tractable [3]. Then, it is classical to use the EM-algorithm to estimate the mixture parameters. We impose that the two components share the same variance and we consider the estimated innovations for all $k \in \{1, ..., n\}$, $\widehat{Z}_k^{(i)}$, for $i = 1$ or 2. We can use the R function `Mclust` on these estimated innovations [14] or the Matlab function `fitgmdist`, and we impose it two components with the same variance: one for the peaks, the other one for the innovations without peaks. This algorithm gives estimators of the model by iterations of two steps: expectation of the completed log-likelihood of the estimated innovations and choice of parameters which maximize this expectation. At each iteration, the log-likelihood of the estimated innovations increases, which enables us to have a better estimation at a given iteration than at the previous one. We can refer to [29] for more details. The algorithm stops when the log-likelihood increases by increments smaller than a tolerance. It gives us the parameters of the clustering:

- the estimated mean of each gaussian $(\widehat{\mu}_1, \widehat{\mu}_2)$: once sorted such that $\widehat{\mu}_1 \le \widehat{\mu}_2$ we should have $\widehat{\mu}_1$ that estimates $-\lambda\nu$ and $\widehat{\mu}_2$ that estimates $\lambda(1-\nu)$ so we can use $\widehat{\mu}_2 - \widehat{\mu}_1$ as an estimator of λ;
- the estimated proportion of each gaussian $(\widehat{p}_1, \widehat{p}_2)$: we keep $\widehat{p}_2$ as an estimator of ν;
- the common estimated variance $\widehat{v}$ which estimates σ^2.

Then we denote by $\hat{\psi}_n^{(i)} := (\hat{\nu}_n^{(i)}, \hat{\lambda}_n^{(i)}, \hat{\sigma}_n^{(i)})$ the corresponding estimators obtained by using $(\widehat{Z}_k^{(i)})_{1\le k\le n}$. The main problems of the EM-algorithms are that it not ensures to have the estimator of maximum likelihood if there exist local maxima (importance of the initialization) and that the convergence can be slow. We can refer to [26] to choose the initialization and to [21] to study different types of convergence. Here we use five replicates for each estimation and apply this algorithm with the previous estimated innovation so that we obtain two estimators $\hat{\psi}_n^{(i)} := (\hat{\nu}_n^{(i)}, \hat{\lambda}_n^{(i)}, \hat{\sigma}_n^{(i)})$ for $i = 1, 2$. We present bias as well as empirical 95% confidence intervals in Figure 6, while a specific comparison between global estimation and 2-step estimation for bias and standard deviation is presented in Figure 7.

Surprisingly, we notice that the estimators of ν, λ and σ seem not to depend on ϕ, whatever the estimation of the innovations (global or 2-step) and both of them performs rather similarly. We can also remark that here is a stronger bias for the λ estimation and as expected standard deviations increase with respect to σ.

Let us finally remark that we are therefore able to estimate the whole set of parameters since one can then set

$$\hat{a}_n^{(1)} = \hat{m}_n^{(1)} - \hat{\lambda}_n^{(1)}\hat{\nu}_n^{(1)}.$$

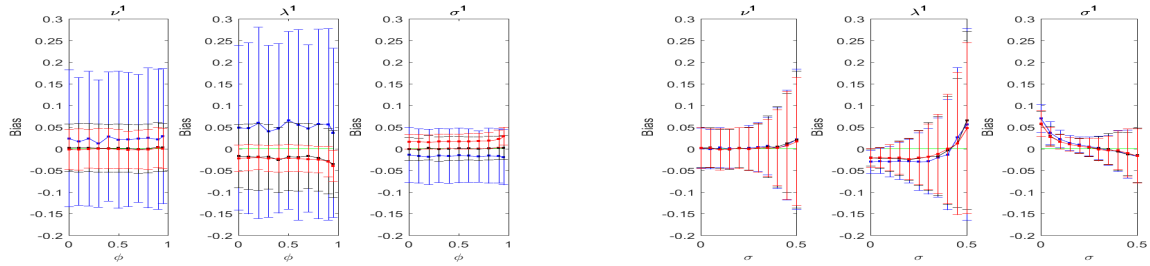


FIGURE 6. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c_n$ for EM estimators using estimated innovations from Theorem 2. On the left: with repect to $\phi \in \{0, 0.1, \ldots, 0.9, 0.95\}$, in red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. On the right with repect to $\sigma \in \{0, 0.05, \ldots, 0.5\}$, in red $\phi = 0.1$, in black $\phi = 0.5$ and in blue $\phi = 0.9$

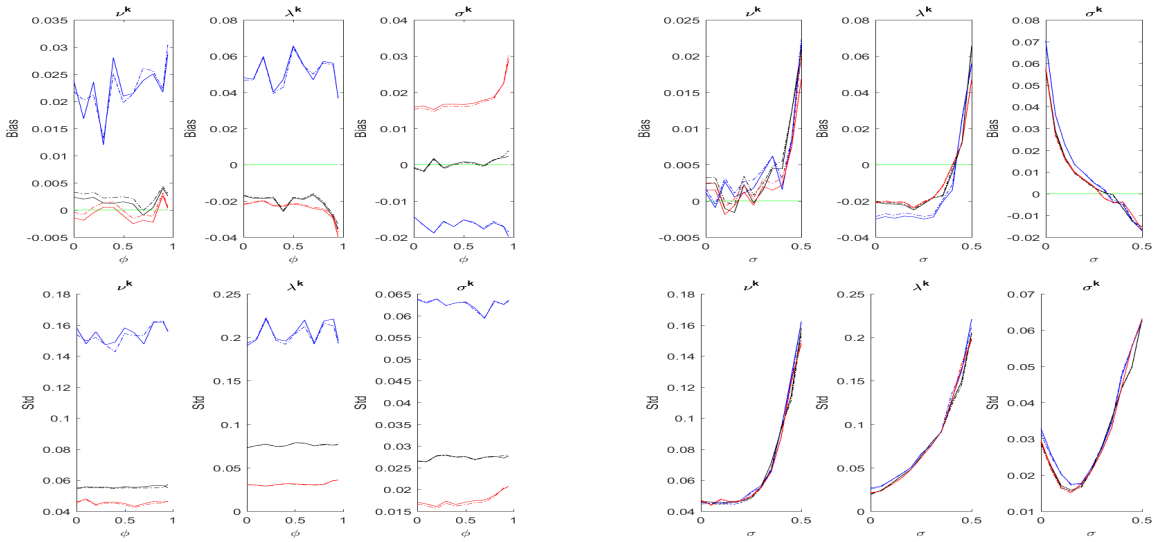


FIGURE 7. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c_n$ with respect to $\phi \in \{0, 0.1, \ldots, 0.9, 0.95\}$. Straight line for the global estimation and dashed line for the 2-Step estimation. On the left in red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. On the right $\phi = 0.1$, in black $\phi = 0.5$ and in blue $\phi = 0.9$.

4. Spike detection

4.1. **Tests on estimated innovations.** Let us recall that the innovations of the process X are given by

$$\forall k \in \{1, ..., n\}, \quad Z_k = \lambda (U_k - \nu) + \sigma \epsilon_k = X_k - \phi X_{k-1} - a - \lambda\nu - b\frac{k-1}{n},$$

which can be estimated through $\widehat{Z}_k^{(i)}$ where $i = 1$ or 2.

We have already remarked that Z_k is a mixture between two gaussians (see Section 3.3). As we want to detect the spikes, we have to find the $k \in \{1, ..., n\}$ such that $U_{k+1} = 1$. Therefore we can define two tests.

- T: we consider for a threshold $s \in \mathbb{R}$ the probabilities:

$$P_F := P_F(s, \nu, \lambda, \sigma) = \mathbb{P}(Z_k > s | U_{k+1} = 0) = 1 - \Phi\left(\frac{s + \lambda\nu}{\sigma}\right)$$

$$P_T := P_T(s, \nu, \lambda, \sigma) = \mathbb{P}(Z_k > s | U_{k+1} = 1) = 1 - \Phi\left(\frac{s - \lambda(1-\nu)}{\sigma}\right),$$

where Φ is the cumulative distribution function of a standard normal variable. We consider also for a tolerance $\alpha \in (0, 1)$:

$$s_\alpha = \min\{s \in \mathbb{R} | P_F \leq \alpha\} = \sigma q_{1-\alpha} - \lambda\nu,$$

where $q_{1-\alpha}$ denotes the quantile of order $1-\alpha$ of a standard Gaussian random variable. The test T consists in considering the instant k as a peak if $Z_k > s_\alpha$. Then P_F is the false detection probability at each time k and P_T is the true detection probability at each time k.

• T_c: EM algorithm can provide us the a posteriori probabilities $\hat{\nu}_k = \mathbb{P}\left(U_k = 1 | \, (Z_k)_{1\leq k\leq n}\right)$ which are the probabilities, given $(Z_k)_{1\leq k\leq n}$, that the instant k is a peak. Then T_c consists in considering the instant k as a peak if

$$1 - \hat{\nu}_k = \mathbb{P}\left(U_k = 0 | \, (Z_k)_{1\leq k\leq n}\right) \leq \alpha,$$

for a tolerance α. Let us note that the clustering by default is usually given by choosing $\alpha = \frac{1}{2}$.

Then, for $i \in \{1, 2\}$ we denote $T^{(i)}$ and $T_c^{(i)}$ the tests T and T_c, used respectively with the estimations $(\widehat{Z}_k^{(i)})_{1\leq k\leq n}$.

In order to enlight performances of both tests we consider a trajectory of size $n+1$ with $n = 100$. We recall that $(U_k)_{1\leq k\leq n}$ is the sequence of $\{0,1\}$-valued variables with true spiking times given by 1 and, for a tolerance α on false detection, $(T_k^{(i)}(\alpha))_{1\leq k\leq n}$, respectively $(T_{c,k}^{(i)}(\alpha))_{1\leq k\leq n}$, the sequence of $\{0,1\}$-valued variables with detected spiking times given by 1 using estimated innovations $\widehat{Z}_k^{(i)}$ and $\hat{\psi}_n^{(i)} = (\hat{\nu}_n^{(i)}, \hat{\lambda}_n^{(i)}, \hat{\sigma}_n^{(i)})$, described in the previous section. Then we can compute the true and false positive rates as

$$\begin{aligned}
\tau_T^{(i)}(\alpha) &:= \#\{1 \leq k \leq n; U_k = 1, T_k^{(i)}(\alpha) = 1\}/\#\{1 \leq k \leq n; U_k = 1\}, \\
\tau_F^{(i)}(\alpha) &:= \#\{1 \leq k \leq n; U_k = 0, T_k^{(i)}(\alpha) = 1\}/\#\{1 \leq k \leq n; U_k = 0\},
\end{aligned}$$

and similarly we write $\tau_{c,T}^{(i)}(\alpha)$ and $\tau_{c,F}^{(i)}(\alpha)$ when $T^{(i)}$ is replaced by $T_c^{(i)}$.

We estimate the true and false detection probabilities by considering $N = 1000$ independent simulations of sample paths and computing the empirical mean of the sample

$$\hat{P}_T^{(i)}(\alpha) = \frac{1}{N}\sum_{l=1}^{N}\left(\tau_T^{(i)}(\alpha)\right)^{(l)} \text{ and } \hat{P}_F^{(i)}(\alpha) = \frac{1}{N}\sum_{l=1}^{N}\left(\tau_F^{(i)}(\alpha)\right)^{(l)}, \tag{14}$$

as well as

$$\hat{P}_{c,T}^{(i)}(\alpha) = \frac{1}{N}\sum_{l=1}^{N}\left(\tau_{c,T}^{(i)}(\alpha)\right)^{(l)} \text{ and } \hat{P}_{c,F}^{(i)}(\alpha) = \frac{1}{N}\sum_{l=1}^{N}\left(\tau_{c,F}^{(i)}(\alpha)\right)^{(l)}. \tag{15}$$

We first present numerical results when ϕ is varying in Figure 8 with also 95% confidence intervals for $\alpha = 0.01$ as previously. Surprisingly, despite the values of ϕ near 1 affect the quality of the first estimated values, it seems to have no consequence on the results of the different tests since their performances do not vary too much with ϕ, as it was also the case for the second estimates using EM-algorithms. However the performances are highly depending on the level of noise σ as we can see in Figure 9 where we have plotted results according to σ varying between 0 and 0.5. As remarked previously, tests are performing similarly choosing the first or the second method of innovations. However as σ is increasing the level of T is higher than α but not exceeding 0.1. In contrast T_c seems to have a lower false detection rate but it has also a lower true detection rate.

A usual way to compare tests is to use ROC curves (for Receiver Operating Characteristic) that enable to measure a binary test's sensitivity. We can refer to [25] for more details. Since our tests depend on a level $\alpha \in (0,1)$ we can compute empirical ROC curves defined as the location of the pairs $(\hat{P}_F^{(i)}(\alpha), \hat{P}_T^{(i)}(\alpha))$ given by (14) for $T^{(i)}$ or $(\hat{P}_{c,F}^{(i)}(\alpha), \hat{P}_{c,T}^{(i)}(\alpha))$ given by (15) for $T_c^{(i)}$ when α is varying between 0 and 1. Let us quote that we can also compute the theoretical ROC curve associated to T and in this case the pairs (P_F, P_T) are related through the fact that the measure d' defined in [18] by

$$d' = \Phi^{-1}(P_T) - \Phi^{-1}(P_F),$$

where Φ is the cumulative distribution function of a standard normal variable, is constant. With the expressions of P_T and P_F and the symmetry of Φ we obtain $d' = \frac{\lambda}{\sigma}$, so each theoretical curve can be associated to a value of the ratio $\frac{\lambda}{\sigma}$ (and each point of a curve is related to a tolerance α and a threshold s_α). Note that d' increases when P_T increases and when P_F decreases, that is to say when the test is more sensitive and in our case it corresponds to a decrease of σ when λ is constant. This is confirmed by Figure 10 where the area under curve is decreasing with respect to σ increasing in $\{0.1, 0.3, 0.5\}$. The first ROC curve confirms that there is no significant difference between $T^{(1)}$ and $T^{(2)}$ for this kind of simulations. We omit a similar result for $T_c^{(1)}$ and

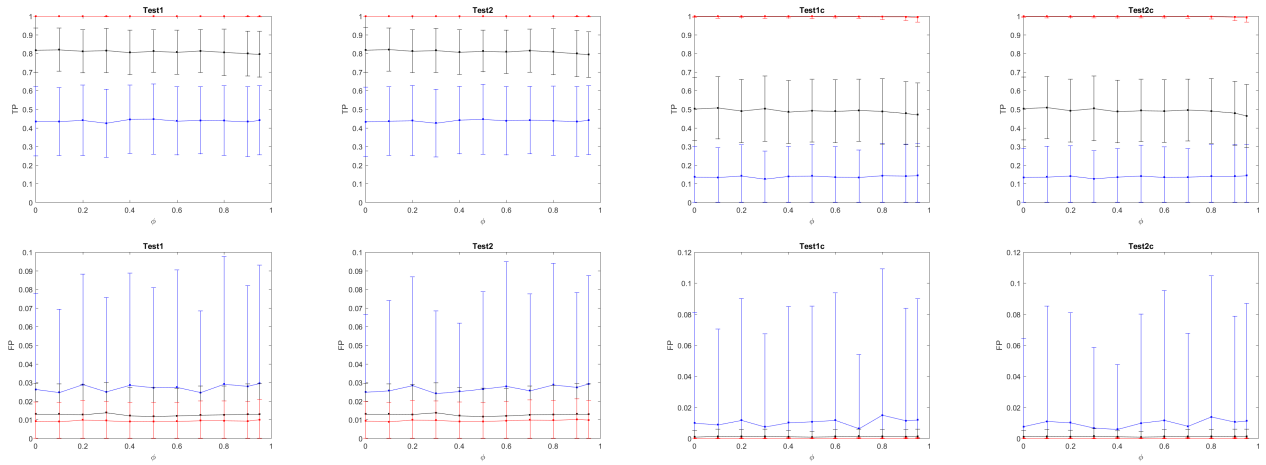
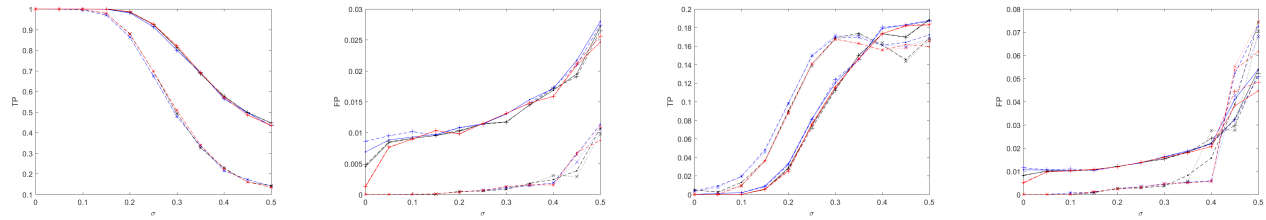


FIGURE 8. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c$ with respect to $\phi \in \{0, 0.1, \ldots, 0.9, 0.95\}$. Threshold given by $\alpha = 0.01$. In red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. First line: True positive rate according to the different Tests. Second line: false positive rate



Mean of TPR | Mean of FPR | Std of TPR | Std of FPR

FIGURE 9. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c$ with respect to $\sigma \in \{0, 0.05, \ldots, 0.5\}$. Threshold given by $\alpha = 0.01$. In red $\phi = 0.1$, in black $\phi = 0.5$ and in blue $\phi = 0.9$. Straigth line with dot for $T^{(1)}$, dashed line with plus for $T^{(2)}$, dashed line with dot for $T_c^{(1)}$, dotted line with cross for $T_c^{(2)}$.

$T_c^{(2)}$. The second ROC curve shows that $T^{(1)}$ seems to be better than $T_c^{(1)}$ and especially as σ is increasing for a fixed $\phi = 0.5$. It reveals that for a rate of false positive detection that we can tolerate, we can find a threshold which provides a bigger rate of true positive detection with $T^{(1)}$ than with $T_c^{(1)}$. This is also true for different values of ϕ since ROC curves completely coincide with ϕ for both tests as shown in the two last figures.

4.2. **Multiple test correction on false rate detection.** In order to correct our tests according to a level $\beta \in (0, 1)$ of false detection we propose to use the following corrections. For the first test, we use the correction proposed by Benjamini-Höchberg [4]. By this way we improve the power of a simple Bonferroni correction that would consist in choosing $\beta = \alpha/n$ meanwhile controling the false rate detection at the level β. More precisely, for all $1 \leq k \leq n$, let us recall that

$$Z_k = \lambda(U_k - \nu) + \sigma\epsilon_k,$$

such that under the assumption that $\{U_k = 0\}$ we have $\frac{Z_k + \lambda\nu}{\sigma} \sim \mathcal{N}(0, 1)$. Hence we consider $\frac{\hat{Z}_k + \hat{\lambda}\hat{\nu}}{\hat{\sigma}}$, where $\hat{Z}_k, \hat{\lambda}, \hat{\nu}, \hat{\sigma}$ are estimators of $Z_k, \lambda, \nu, \sigma$ and compute its associated p-value as

$$p_k = 1 - \Phi\left(\frac{\hat{Z}_k + \hat{\lambda}\hat{\nu}}{\hat{\sigma}}\right),$$

where Φ is the cumulative distribution function of the standard Gaussian distribution. We sort $(p_k)_k$ such that $p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(n)}$, and set $p_n^* = p_n$. Then we compute iteratively for k ranging from $n - 1$ to

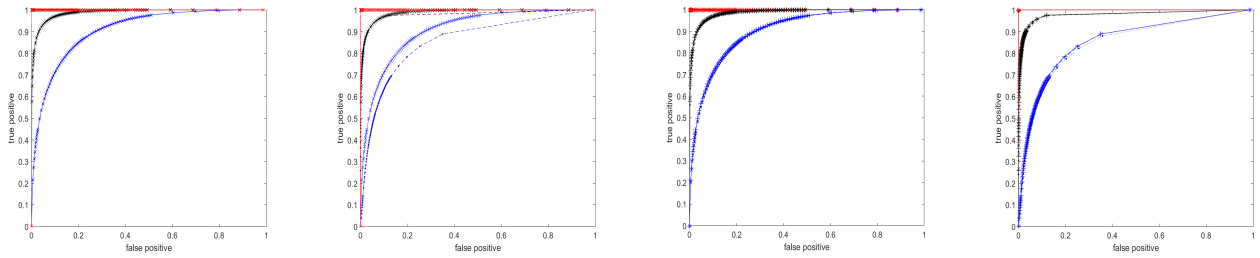


FIGURE 10. ROC curves obtained with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c_n$. In red $\sigma = 0.1$, in black $\sigma = 0.3$ and in blue $\sigma = 0.5$. First: $\phi = 0.5$, solid line with cross for $T^{(1)}$, dashed line with dot for $T^{(2)}$. Second: $\phi = 0.5$, solid line with cross for $T^{(1)}$, dashed line with dot for $T_c^{(1)}$. Third and Fourth: $T^{(1)}$ (left) and $T_c^{(1)}$ (right) with several values of ϕ, solid line with cross for $\phi = 0.1$, dotted line for $\phi = 0.9$, dashed with plus for $\phi = 0.5$.

1 the quantity $p_k^* = \min(p_{k+1}^*, \frac{n}{k} p_k)$. Then, using our two way of estimation, we can define for $i = 1, 2$ and $1 \leq k \leq n$,

$$\tilde{T}_k^{(i)}(\beta) = 1 \text{ if } p_k^{*(i)} \leq \beta \text{ and } \tilde{T}_k^{(i)}(\beta) = 0 \text{ otherwize.} \tag{16}$$

For the second test, we know from Theorem 1 of [28] that the optimal decision for false discovery rate is given by

$$\forall 1 \leq k \leq n, \;\; \hat{\nu}_k \geq t^*,$$

for some cutoff $t^* = \frac{c^*}{1+c^*} = \frac{1}{1+1/c^*}$. In order to control the number of false positive detection at a level β we first remark that the number of false positive, corresponding to the number of false discovery, is given by

$$\sum_{k=1}^{n} (1 - U_k) \mathbf{1}_{\hat{\nu}_k \geq t^*},$$

such that the posterior expected false positive number is given by

$$\sum_{k=1}^{n} (1 - \hat{\nu}_k) \mathbf{1}_{\hat{\nu}_k \geq t^*} \leq \frac{1 - t^*}{t^*} \sum_{k=1}^{n} t^* \mathbf{1}_{\hat{\nu}_k \geq t^*} \leq \frac{1 - t^*}{t^*} \sum_{k=1}^{n} \hat{\nu}_k \leq \frac{1 - t^*}{t^*} n\hat{\nu},$$

since $\sum_{k=1}^{n} \hat{\nu}_k = n\hat{\nu}$ by EM algorithm, which is the estimated expected number of jumps. Hence we propose to bound the false detection rate by $\frac{1-t^*}{t^*} \frac{\hat{\nu}}{1-\hat{\nu}} = \frac{1}{c^*} \frac{\hat{\nu}}{1-\hat{\nu}}$ and choose $c^*(\beta) = \frac{1}{\beta} \frac{\hat{\nu}}{1-\hat{\nu}}$ in order that $\frac{1}{c^*} \frac{\hat{\nu}}{1-\hat{\nu}} \leq \beta$. Then, using our two way of estimation, we can define for $i = 1, 2$ and $1 \leq k \leq n$,

$$\tilde{T}_{c,k}^{(i)}(\beta) = 1 \text{ if } \hat{\nu}_k^{(i)} \geq \frac{1}{1 + \beta \frac{1-\hat{\nu}}{\hat{\nu}}} \text{ and } \tilde{T}_{c,k}^{(i)}(\beta) = 0 \text{ otherwize.} \tag{17}$$

As for previous ones, on simulations, these new tests perform similarly for each method of estimation meaning that there is not significant difference between $i = 1$ and $i = 2$, and are not affected by the values of ϕ. We therefore focus on the comparison for the first estimation $i = 1$ and $\phi = 0.5$. We present these results in Figure 11 that emphasizes that these corrections achieve the required false detection rate in contrast with previous tests. The powers seem also comparable and are better than $T_c^{(1)}$ but lower than $T^{(1)}$. This is confirmed with ROC curves of Figure 12 where the corresponding values for $\alpha = \beta = 0.01$ are plotted in green. However in term of AUC comparison it seems that T^1 is better than $\tilde{T}^1$ (first plot) that behaves as T_c^1 (second plot). The two last plots are concerned with $\tilde{T}_c^1$. First let us mention that we make evolves $\alpha \in (0, 1)$ and $\beta \in (0, 1)$ for the corrected tests. Doing this we can check that the cutoff for $\tilde{T}_c^1$ is always larger than $\hat{\nu}$ that is around to $\nu = 0.3$ in our experiments. Hence letting $\beta \in (0, 1)$ does not allow to reach a rate 1 for false positive and we restrict the axis of curves between 0 and 0.3 consequently. Interestingly the third ROC curve shows a similar behavior between T^1 and $\tilde{T}_c^1$ and consequently we have improved the power of T_c^1 still controlling the rate of false positive. The last plot proves that the Bayesian correction of $\tilde{T}_c^1$ could do better than the frequentist one of $\tilde{T}^1$.

In Figure 23 several realizations are obtained considering fixed c, d also with the same jumps. Hence the linear slope is now identical for each realizations and the different values of ϕ only change a little the initial

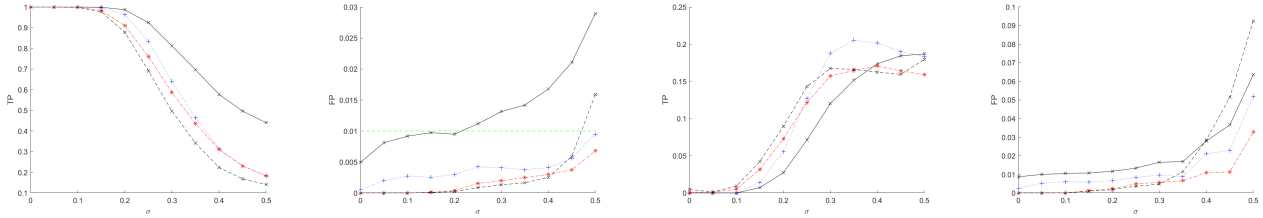


FIGURE 11. Comparison with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c$ and $\phi = 0.5$, with respect to $\sigma \in \{0, 0.05, \ldots, 0.5\}$. Threshold given by $\alpha = \beta = 0.01$. In black red straigth line for $T^{(1)}$ and black dashed line for ${T_c}^{(1)}$, in blue dotted line with plus for $\tilde{T}^{(1)}(\beta)$ and in red with cross for $\tilde{T}_c^{(1)}(\beta)$.

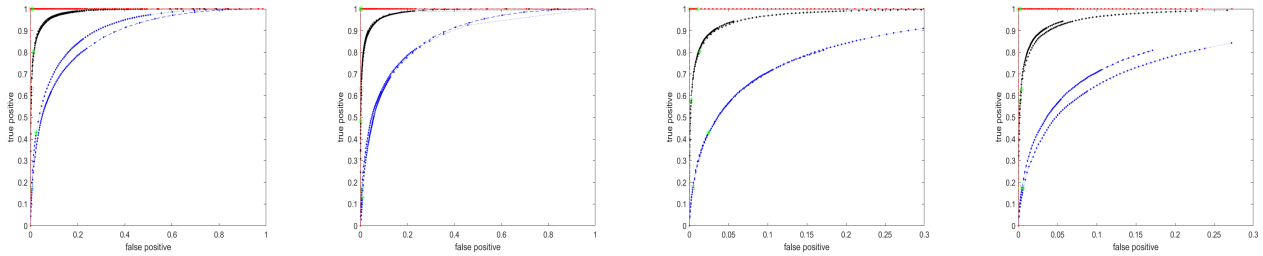


FIGURE 12. ROC curves obtained with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $n = 100$, $X_0 = c_n$ and $\phi = 0.5$. In red $\sigma = 0.1$, in black $\sigma = 0.3$, in blue $\sigma = 0.5$, in green values obtained for $\alpha = 0.01$ or $\beta = 0.01$. Dotted line with dot and a green star for $T^{(1)}$ (first and third), ${T_c}^{(1)}$ (second), $\tilde{T}^{(1)}$ (last); dashed line with dot and a green square for $\tilde{T}^{(1)}$ (first and second), for $\tilde{T}_c^{(1)}$ (third and last)

value. Let us emphasize that we chose as initial conditions c_n in order to have a point on the line. This choice corresponds to an ideal case for the 2-step estimations and it performs very similarly than the first global way of estimation. However we have remarked that the global estimation is more robust to changes with respect to initial conditions. We present in Figure 24 several realizations obtained considering fixed a, b and with the same jumps. The different values of ϕ contribute to changes in the linear trend and sample paths are very different. We give corresponding computed true positive rate in Table 2 and Table 4 and false positive rate in Table 3 and Table 5 for a fixed level chosen as $\alpha = \beta = 0.01$ for initial and corrected tests, using the two estimation procedures. Whatever the chosen test is, the performances are getting worser as σ is increasing. However, tests T and T_c are not performing similarly for a fixed level. The true and false positive rate for T are higher than for T_c. Despite we have chosen $\alpha = 0.01$ we can obtain up to 0.1 for false positive rate of T, in the worst case $\sigma = 0.5$, while false positive rate of T_c are of the accurate order. This is corrected with $\tilde{T}$ and $\tilde{T}_c$. In the worst case the false positive rate is 0.04 for $\tilde{T}$ (3 false detected jumps) and 0.01 for $\tilde{T}_c$ (1 false detected jump). We also improve the true positive rate compared to T_c but it remains quite far from the one of T.

5. Data analysis

We have considered a biological experiment of Anne Duittoz's team, composed with 14 trajectories of neurons observed during 600 seconds. More precisely, the observed values correspond to Calcium Green-1 mean fluorescence intensity of 14 selected regions corresponding to neurons in a Calcium imaging recording (see Figure 13). Inferring individual spikes from fluorescence signals is a high challenging problem but calcium imaging should allow to characterize neuronal events in vivo [1].

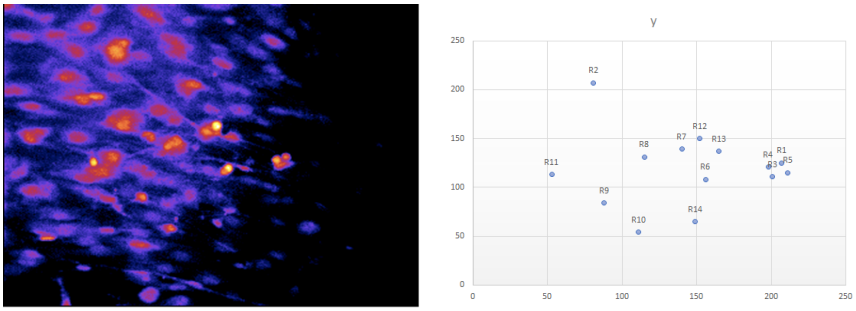


FIGURE 13. Image recording of 14 different Neurons. On the left: a typical image of the movie of 600 seconds (one frame per second). On the right: the locations of the 14 selected neurons with recorded paths under statistical analysis.

In order to apply our spike detection methodology we restrict the time domain to $\{301, \dots, 421\}$ of length $n+1$ with $n = 120$, where goodness of fit for the model seems reasonable. In this time period our two methods of estimation give similar results and we only present results obtained for the first one for sake of conciseness. Before presenting results on recorded neurons we present in Figure 14 results obtained on the movie where estimation is performed pixel by pixel with a size of image given by 247×325 coded with 8 Bytes meaning that the values are in $\{0, \dots, 255\}$. Note that, as expected, $\hat{c}_n^{(1)}$ captures the global image with a slope $\hat{d}_n^{(1)}$ around 0 except near neurons where we can also observe values of $\hat{\phi}_n^{(1)}$ greater than 0.5. Despite $\hat{\lambda}_n^{(1)}$ and $\hat{\sigma}_n^{(1)}$ seem coherent, $\hat{\nu}_n^{(1)}$ is very noisy and we will see that this estimation can perform badly, especially when outliers are present.

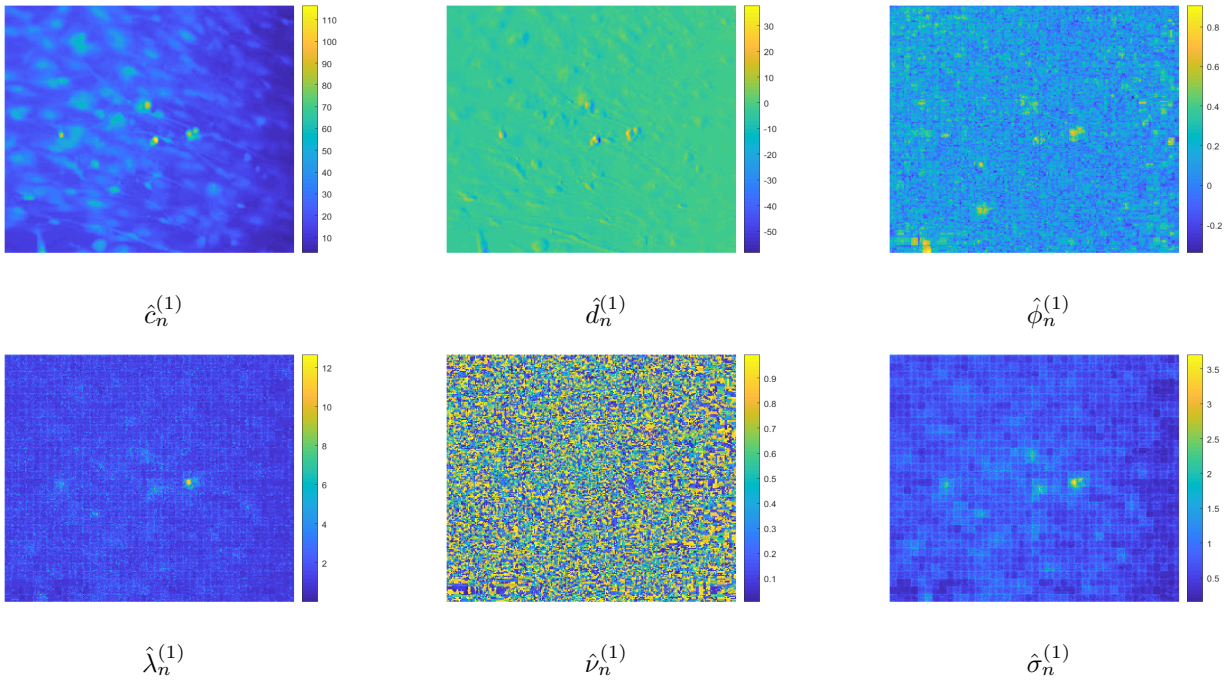
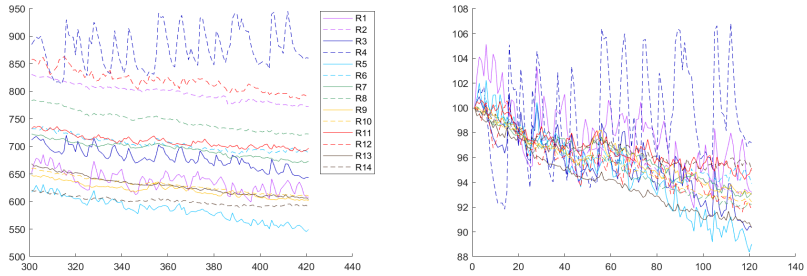


FIGURE 14. Estimation results on the recorded movie during the time lapse $\{301, \dots, 421\}$. Each pixel is analyzed as a realization of (3) with $n = 120$, whose parameters are estimated according to Section 3. Values of each pixels are given through the corresponding colorbar on the right of each image.

Then we restrict on the recorded neurons sample paths. In order to compare them, we normalized them in such a way that the time period corresponds to $\{0, \dots, n\}$ with initial value taken at 100 (see Figure 15), meaning that we multiply the signals by 100 divided by initial value.



Neurons Normalization

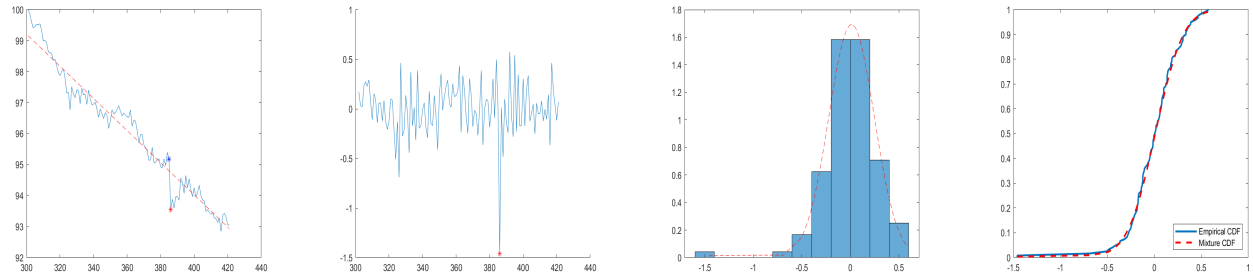
FIGURE 15. Recorded Neurons. On the left: the 14 selected trajectories $(X_k^j)_{301\leq k\leq 421}$ for $j = 1,\dots,14$. On the right: the 14 normalized trajectories for which the time is set to be $\{0,\dots,120\}$ and the initial value is 100, corresponding to $(X^j_{k+301} \times (100/X^j_{301}))_{0\leq k\leq 120}$ for $j = 1,\dots,14$.

In the whole experiment, we notice a high variability of the estimated parameters between the different neurons, especially for ϕ and λ, as it can be seen in Table 1. The last column correspond to the p-value obtained by performing a Kolmogorov-Smirnov goodness of fit test on the estimated innovation compared with the estimated Gaussian mixture. The Kolmogorov-Smirnov test is accepted for all neurons with a level given by the smallest p-value obtained for Neuron R4. This neuron is the only to exhibit an increasing trajectory.

Neuron	$\widehat{c}_n^{(1)}$	$\widehat{d}_n^{(1)}$	$\widehat{\phi}_n^{(1)}$	$\widehat{\lambda}_n^{(1)}$	$\widehat{\nu}_n^{(1)}$	$\widehat{\sigma}_n^{(1)}$	X_0	KS$p-$value
R1	101.2366	−6.9581	0.6732	2.5227	0.3718	0.9237	651.8600	0.6390
R2	99.1601	−6.2252	0.6929	1.4693	**0.9916**	0.2338	830.3800	0.8835
R3	99.2140	−7.6091	0.5887	2.5859	0.1469	0.6058	711.8500	0.2548
R4	97.6087	**2.8797**	0.6736	6.8259	0.1398	1.5634	885.1000	**0.0165**
R5	100.5374	−11.3993	0.4375	1.2457	0.2279	0.6656	617.1600	0.9908
R6	98.9209	−5.4705	0.7755	0.9122	0.0953	0.3720	732.3100	0.6739
R7	99.1936	−6.0570	0.8222	0.6450	**0.9754**	0.2547	722.6300	0.6982
R8	98.8731	−7.1818	0.8776	0.3688	0.6338	0.1376	783.6500	0.9776
R9	99.0390	−6.0332	0.8480	0.6253	0.0528	0.3078	647.0400	0.6350
R10	98.8467	−6.7943	0.7899	0.2901	**0.9277**	0.1967	660.1700	0.9049
R11	99.2561	−4.7800	0.7338	0.7595	0.3311	0.2590	732.3500	0.9390
R12	98.7342	−6.7195	0.7011	0.5676	0.7958	0.5834	858.1600	0.9022
R13	98.0884	−7.9994	0.8304	0.5031	0.1080	0.2048	668.5300	0.8813
R14	98.7415	−3.9130	0.7226	0.4726	0.3493	0.2322	620.6200	0.9853

TABLE 1. Estimated parameters for recorder Neurons. Estimation results on the normalized trajectories (see Figure 15) according to Section 3. In bold for R4 the only positive value estimated for d, corresponding to an increasing path for which the Kolmogorov Smirnov test between estimated innovation and estimated Gaussian mixture gives the smallest p-value (see column KS). In bold for R2, R7, R10, EM-estimations of ν are close from 1: these trajectories can not be analyzed directly by this model.

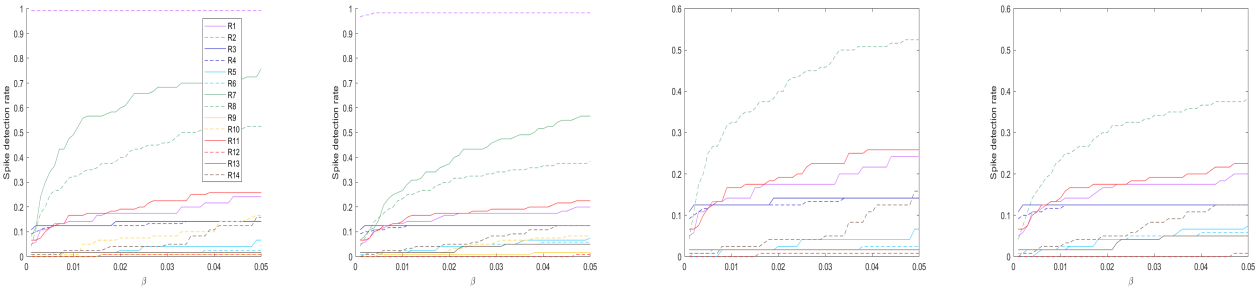
Note that Neurons R2, R7 and R10 exhibit a too high level for $\widehat{\nu}$. This can be explained by some few outliers in sample paths as explained in Figure 16. We restrict our attention on Neuron R2 and plot at left its sample path and the estimated trend (first) and the estimated innovations (second). Then we plot histogram and cumulant distribution function of innovations compared with the estimated Gaussian mixture. The red star corresponds to the minimal value of innovation that is an outlier for our Gaussian mixture model, corresponding to a big negative jump as we can see with the previous recorded value in blue star on R2. However EM algorithm isolates this value as a single Gaussian component that explains the high value of $\hat{\nu}_n^{(1)}$ (see Table 1).



R2 Innovations Histogram CDF

FIGURE 16. Outliers for Neuron R2. First figure: normalized trajectory of neuron R2 with a big negative jump in red star just after the blue star. The estimated trend is plotted in dashed red line. Second figure: plot of estimated innovations of R2 with a negative outlier in red star, corresponding to the negative jump of R2. Third figure: normalized histogram of estimated innovations of R2 and estimated Gaussian mixture density in dashed red line. Last figure: empirical cumulant distribution function of estimated innovations of R2 in blue and estimated Gaussian mixture cumulant distribution function in dashed red line.

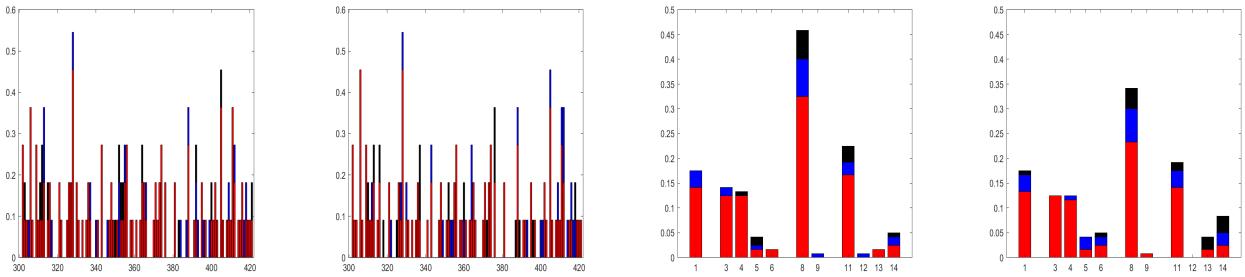
Then we proceed to our two different corrected tests $\tilde{T}$ and $\tilde{T}_c$ for level $\beta \in (0.001, 0.05)$ (see (16) and (17)). In Figure 17 we represent the rate of detected spikes by Neurons for the whole experiments. As expected we recover high detection rate for Neurons R2 and R7 but not for Neuron R10 (in dashed orange lines). Restricting to the other Neurons shows that Neuron R8 (in green dashed line) also exhibits a strange behavior with respect to detection. The comparison between the two tests again shows a better power for $\tilde{T}$.



$\tilde{T}$ $\tilde{T}_c$ $\tilde{T}$ $\tilde{T}_c$

FIGURE 17. Rate of spike detection according to test $\tilde{T}$ of (16) and test $\tilde{T}_c$ of (17) with respect to the level β of false positive detection. On the right: the whole experiment with the 14 recorded Neurons. On the right: Neurons R2, R7 and R10 are removed in view of ν estimates given by Table 1.

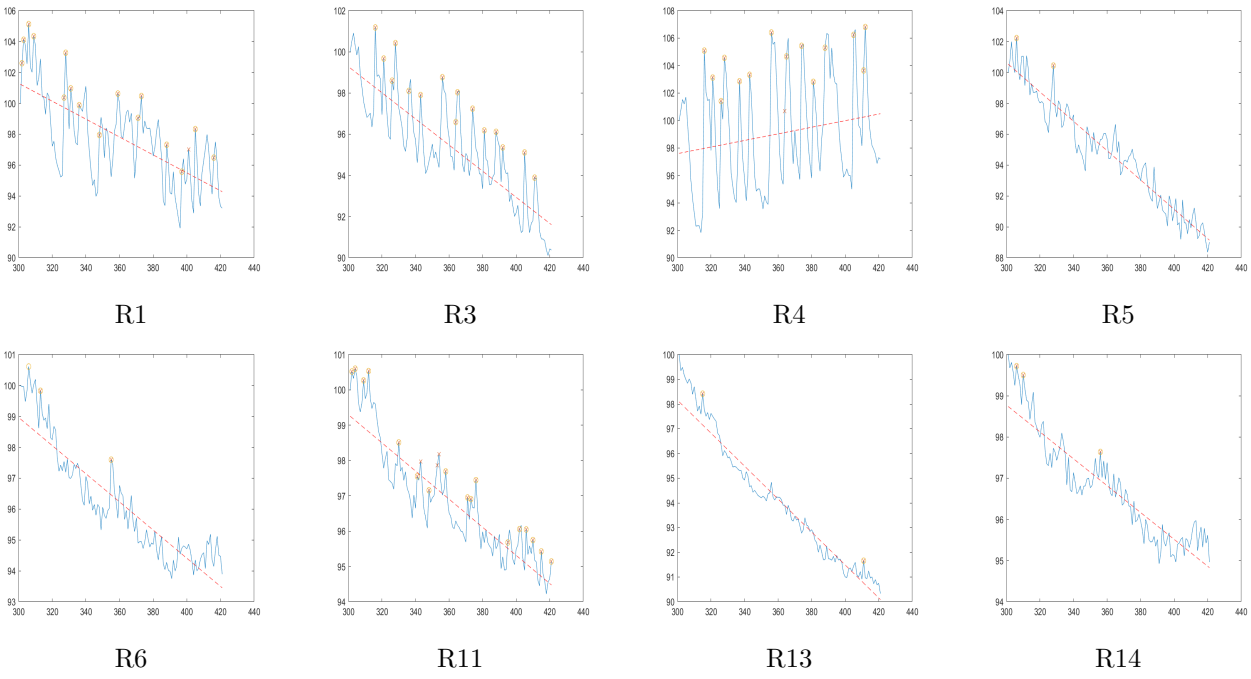
This is confirm in Figure 18 where we restrict the study to the eleven remaining neurons. Interestingly $\tilde{T}_c(\beta)$ seems to detect synchronisation event involving more neurons at level $\beta = 0.01$.



$\tilde{T}(\beta)$ $\tilde{T}_c(\beta)$ $\tilde{T}(\beta)$ $\tilde{T}_c(\beta)$

FIGURE 18. Test $\tilde{T}(\beta)$ of (16) and Test $\tilde{T}_c(\beta)$ of (17) for 11 Neurons (Neurons R2, R7 and R10 are removed), in red for $\beta = 0.01$, in blue for $\beta = 0.02$ and in black for $\beta = 0.03$. On the right: frequency of neurons detecting a spike per times. On the right: frequency of spike times per neurons.

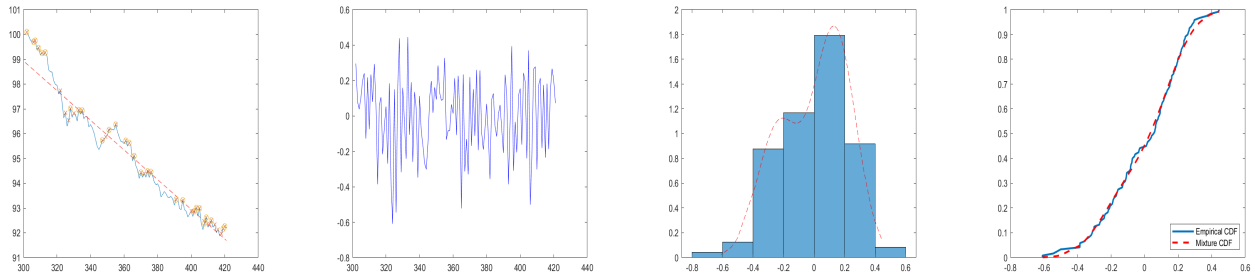
The time 306 is detected by more than 30% of the neurons and this percentage is not influenced by the level β. It is therefore probably a synchronisation event of the experiment. It involves Neurons R1, R5, R8 and R14 for $\tilde{T}$, while $\tilde{T}_c$ detects also Neuron R6. At time 328 both Neurons R1, R3, R4, R5, R8 are detected to have a spike by $\tilde{T}$ and $\tilde{T}_c$ at level $\beta = 0.01$. It also involves Neuron R11 at level $\beta = 0.02$. Finally, time 405 is detected for Neurons R1, R3, R4, R8 both for $\tilde{T}$ and $\tilde{T}_c$ at level $\beta = 0.01$ and involves Neuron R5 for $\tilde{T}$ at level $\beta = 0.03$ and $\tilde{T}_c$ at level $\beta = 0.02$. The sample paths with estimated spikes are presented in Figure 19. Note that Neuron R8 is involved in all these particular times. It is actually the Neuron having the higher rate of spiking times but it also has the greater value of ϕ that makes difficult to visually distinguish spiking times (see Figure 20).



R1 R3 R4 R5

R6 R11 R13 R14

FIGURE 19. Several neurons with normalized paths and detecting spike times with cross for $\tilde{T}(\beta)$ of (16) and circle for $\tilde{T}_c(\beta)$ of (17) with $\beta = 0.01$. The estimated trend is plotted in red dashed line.



R8 Innovations Histogram CDF

FIGURE 20. Innovations for Neuron R8. First figure: normalized trajectory of neuron R8. The estimated trend is plotted in dashed red line. Second figure: plot of estimated innovations of R8. Third figure: normalized histogram of estimated innovations of R8 and estimated Gaussian mixture density in dashed red line. Last figure: empirical cumulant distribution function of estimated innovations of R8 in blue and estimated Gaussian mixture cumulant distribution function in dashed red line.

In this biological problematic, there are still several issues to work on. In particular, we had to restrict arbitrarily time interval under study to avoid misleading estimates due to heterogeneous parts of the trajectories. This comes from the fact that our simple model does not fit a whole trajectory but can fit some parts with different parameters. It should therefore be necessary to adapt the procedure or the model with a piecewise one in order to mimic data trajectories. Then estimation and detection methods would have to be adapted to this more complex setting, using change-point analysis or sliding window for instance. The algorithm EM is also very sensitive with the presence of outliers. This could be manually corrected in practice but this procedure will be very time consuming.

Acknowledgments: We would like to sincerely thank the reviewers for their valuable comments and challenging remarks that helped us to improve substantially this paper.

6. Appendix

6.1. **From continuous time to discrete time.** We assume to observe a process X over a time interval $[0,T]$, $T \in \mathbb{R}^+$, satisfying

$$dX_t = \frac{1}{\tau}(\mu(t) - X_t)dt + \lambda dN_t + \sigma dW_t,$$

for a linear trend $\mu(t) = c + \frac{d}{T}t$, $c \in \mathbb{R}^+$, $d \in \mathbb{R}$, $\tau, \sigma, \lambda \in (0, +\infty)$, and W a Brownian motion independent from N a Poisson process of parameter $\nu \in (0,1]$. It follows, writing $\gamma = 1/\tau$ that

$$\mathrm{d}\left(e^{\gamma t}X_t\right) = \gamma e^{\gamma t}\left(c + \frac{d}{T}t\right)\mathrm{d}t + \sigma e^{\gamma t}\mathrm{d}W_t + \lambda e^{\gamma t}\mathrm{d}N_t. \tag{18}$$

Now, we assume to have $(n+1)$ observations uniformly distributed on $[0,T]$: X_0, $X_{\frac{T}{n}}$, ..., $X_{k\frac{T}{n}}$, ..., X_T. For $k \in \{0,...,n\}$, we will note $X_{n,k}$ for $X_{k\frac{T}{n}}$. Moreover, we introduce $\phi_n := e^{-\gamma\frac{T}{n}}$. By integrating (18) over $[k\frac{T}{n}, (k+1)\frac{T}{n}]$, we have:

$$\begin{aligned}\phi_n^{-(k+1)}X_{n,k+1} - \phi_n^{-k}X_{n,k} = c\left(\phi_n^{-(k+1)} - \phi_n^{-k}\right) + \frac{d}{n}\left((k+1)\phi_n^{-(k+1)} - k\phi_n^{-k}\right) - \frac{d}{\gamma T}\left(\phi_n^{-(k+1)} - \phi_n^{-k}\right) \\ + \sigma\int_{k\frac{T}{n}}^{(k+1)\frac{T}{n}} e^{\gamma t}\mathrm{d}W_t + \lambda\int_{k\frac{T}{n}}^{(k+1)\frac{T}{n}} e^{\gamma t}\mathrm{d}N_t.\end{aligned}$$

Or equivalently:

$$X_{n,k+1} = \phi_n X_{n,k} + c\,(1-\phi_n) + d\frac{k}{n}\,(1-\phi_n) + \frac{d}{n} - \frac{d}{\gamma T}\,(1-\phi_n) + \sigma \int_{k\frac{T}{n}}^{(k+1)\frac{T}{n}} e^{\gamma\left(t-(k+1)\frac{T}{n}\right)} \mathrm{d}W_t + \lambda \sum_{j=N_{k\frac{T}{n}}}^{N_{(k+1)\frac{T}{N}}} e^{\gamma\left(s_j-(k+1)\frac{T}{n}\right)},$$

where $(s_j)_j$ are the points of the Poisson process N. Assuming n is large enough:

$$\begin{aligned}\frac{d}{n} - \frac{d}{\gamma T}\,(1-\phi_n) &= \frac{d}{n} - \frac{d}{\gamma T}\left(1-e^{-\gamma\frac{T}{n}}\right)\\ &= \frac{d}{n} - \frac{d}{\gamma T}\left(-\gamma\frac{T}{n} + o\left(\frac{1}{n}\right)\right)\\ &= o\left(\frac{1}{n}\right);\end{aligned}$$

$$\int_{k\frac{T}{n}}^{(k+1)\frac{T}{n}} e^{\gamma\left(t-(k+1)\frac{T}{n}\right)} \mathrm{d}W_t = \left(W_{(k+1)\frac{T}{n}} - W_{k\frac{T}{n}}\right) + o\left(\frac{1}{n}\right),$$

with $W_{(k+1)\frac{T}{n}} - W_{k\frac{T}{n}} \sim \mathcal{N}\left(0, \frac{T}{n}\right)$ and $N_{(k+1)\frac{T}{n}} - N_{k\frac{T}{n}} \sim \mathcal{P}\left(\nu\frac{T}{n}\right)$. Then

$$\mathbb{P}\left(N_{(k+1)\frac{T}{n}} - N_{k\frac{T}{n}} = 1\right) = \nu\frac{T}{n} + o\left(\frac{1}{n}\right) \text{ and } \mathbb{P}\left(N_{(k+1)\frac{T}{n}} - N_{k\frac{T}{n}} = 0\right) = 1 - \nu\frac{T}{n} + o\left(\frac{1}{n}\right),$$

so we can consider the following approximation

$$X_{n,k+1} = \phi_n X_{n,k} + c\,(1-\phi_n) + d\frac{k}{n}\,(1-\phi_n) + \sigma\epsilon_{k+1} + \lambda U_{k+1},$$

where for all $k \in \{0,...,n\}$, $\epsilon_k \sim \mathcal{N}\left(0, \frac{T}{n}\right)$, $U_k \sim \mathcal{B}\left(\nu\frac{T}{n}\right)$ are independant variables. In this paper we first consider this approximation for a fix time interlaps $\frac{T}{n} = 1$.

6.2. **Proof of Theorem 1.** By Cramer-Wold device, it is sufficient to prove that for all u, v, w and x in $\mathbb{R}$:

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \Delta_{n,k} \xrightarrow[n\to+\infty]{d} \mathcal{N}\left(0, l\right),$$

where we denote $\Delta_{n,k} = \left(u + v\frac{k}{n} + wY_k + xZ_{k+1}\right) Y_k - w\rho_Y(0)$ for all $n \geq 1$ and $k \in \{0,...n\}$ and

$$l = \begin{pmatrix} u & v & w & x \end{pmatrix} \Sigma_1 \, {}^t\begin{pmatrix} u & v & w & x \end{pmatrix}.$$

As usual for proving asymptotic normality in time series we consider and (m_n)-dependent approximation of (Y_k) defined through $Y_k^{(m_n)} = \sum_{j=0}^{m_n} \phi^j Z_{k-j}$ and denote $\Delta_{n,k}^{(m_n)} = \left(u + v\frac{k}{n} + wY_k^{(m_n)} + xZ_{k+1}\right) Y_k^{(m_n)} - w\rho_{Y^{(m_n)}}(0)$, where $\rho_{Y^{(m_n)}}(0) = \mathrm{Var}(Y_k^{(m_n)})$. Let us remark that we may write

$$\Delta_{n,k} = \Delta_{n,k}^{(m_n)} + \left(u + v\frac{k}{n} + xZ_{k+1}\right) R_{m_n,k} + wR_{m_n,k}^2 + 2wY_k^{(m_n)} R_{m_n,k} + w\left(\rho_{Y^{(m_n)}}(0) - \rho_Y(0)\right)$$

where $R_{m_n,k} = \sum_{j=m_n+1}^{+\infty} \phi^j Z_{k-j} = \phi^{m_n+1} Y_{k-m_n-1}$. We will show that we can find $c > 0$ such that

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} (\Delta_{n,k} - \Delta_{n,k}^{(m_n)}) \right\|_{L^2} \leq c|\phi|^{m_n}. \tag{19}$$

Then, since $|\phi| < 1$ as soon as $m_n \to +\infty$, by Slutsky theorem, it is sufficient to prove that

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \Delta_{n,k}^{(m_n)} \xrightarrow[n\to+\infty]{d} \mathcal{N}\left(0, l\right).$$

For that, since now $(\Delta_{n,k}^{(m_n)})$ is a sequence of m_n-dependent random variables, it is enough to check the three points of Theorem 2 from [19] concerning central limit theorem for m_n dependent centered random variables. Namely, in our setting, considering $U_{n,k} = \frac{\Delta_{n,k}^{(m_n)}}{\sqrt{n}}$, it follows from

i) $\mathrm{Var}\left(\sum_{k=0}^{n-1} U_{n,k}\right) = \frac{1}{n}\,\mathrm{Var}\left(\sum_{k=0}^{n} \Delta_{n,k}^{(m_n)}\right) \underset{n\to+\infty}{\longrightarrow} l$;

ii) there exists $C>0$ such that $\sum_{k=0}^{n-1}\mathbb{E}(U_{n,k}^2) \le C$, for all $n \ge 1$;

iii) for all $\varepsilon > 0$, $L_n(\varepsilon) = m_n^2 \sum_{k=0}^{n-1}\mathbb{E}(U_{n,k}^2 \mathbf{1}_{|U_{n,k}|\ge \varepsilon m_n^{-2}}) \underset{n\to+\infty}{\longrightarrow} 0$.

In order to prove that and to compute explicitly the covariance matrix we need the following result.

Lemma 1. *Let $(Z_k)_k$ be a sequence of iid centered rv with $\mathbb{E}(Z_0^4) < +\infty$ and (Y_k) its associated AR(1) process with autoregression coefficient $\phi \in (-1,1)$. Then, for all $n \ge 1$ and $k,l \in \{0,...,n-1\}$ we have*

(1) $Cov\left(Y_k, Y_l^2\right) = \dfrac{\mathbb{E}\left(Z_0^3\right)}{1-\phi^3}\phi^{\frac{1}{2}(l-k)+\frac{3}{2}|l-k|}$,

(2) $Cov\left(Y_k, Y_l Z_{l+1}\right) = 0$,

(3) $Cov\left(Y_k Z_{k+1}, Y_l Z_{l+1}\right) = \dfrac{\sigma_Z^2}{1-\phi^2}\mathbb{E}\left(Z_0^2\right)\mathbf{1}_{\{k=l\}}$,

(4) $Cov\left(Y_k^2, Y_l Z_{l+1}\right) = 2\dfrac{\sigma_Z^4}{1-\phi^2}\phi^{2(k-l)-1}\mathbf{1}_{\{k\ge l+1\}}$,

(5) $Cov\left(Y_k^2, Y_l^2\right) = Cov\left(Y_k^2, Y_l^2\right) = \dfrac{\sigma_Z^4}{1-\phi^2}\left(\dfrac{\phi^{2|l-k|}(\mathbb{E}(Z_0^4)-3\sigma_Z^4)}{\sigma_Z^4(1+\phi^2)} + \dfrac{2\phi^{2|l-k|}}{1-\phi^2}\right)$.

It follows that one can find a constant $c>0$ such that

$$|Cov(\Delta_{n,k}, \Delta_{n,l})| \le c|\phi|^{|k-l|}.$$

Proof. We first consider k and l such that $0 \le k \le l \le n-1$:

$$(1)\quad \mathrm{Cov}\left(Y_k, Y_l^2\right) = \sum_{m,n,p=0}^{+\infty}\phi^{m+n+p}\mathbb{E}\left(Z_{k-m}Z_{l-n}Z_{l-p}\right).$$

Since the centered variables Z_k are i.i.d., the terms of this sum are equal to zero except if the indices $k-m$, $l-n$ and $l-p$ are equal then if $p=n=l-k+m$. So we have:

$$\begin{aligned}\mathrm{Cov}\left(Y_k, Y_l^2\right) &= \sum_{m=0}^{+\infty}\phi^{3m+2(l-k)}\mathbb{E}\left(Z_{k-l}^3\right)\\ &= \frac{\mathbb{E}\left(Z_0^3\right)}{1-\phi^3}\phi^{2(l-k)}.\end{aligned}$$

(2) Since Z_{l+1} is independant of $Y_k Y_l$, we get $\mathrm{Cov}(Y_k, Y_l Z_{l+1}) = 0$.

(3) Since Z_{l+1} is independant of $Y_k Y_l Z_{k+1}$ if $k<l$, we have

$$\begin{aligned}\mathrm{Cov}\left(Y_k Z_{k+1}, Y_l Z_{l+1}\right) &= \mathbb{E}\left(Y_k Z_{k+1} Y_l Z_{l+1}\right)\mathbf{1}_{\{k=l\}}\\ &= \mathbb{E}\left(Y_k^2 Z_{k+1}^2\right)\mathbf{1}_{\{k=l\}}\\ &= \mathbb{E}\left(Y_k^2\right)\mathbb{E}\left(Z_{k+1}^2\right)\mathbf{1}_{\{k=l\}}\\ &= \frac{\sigma_Z^4}{1-\phi^2}\mathbf{1}_{\{k=l\}}.\end{aligned}$$

(4) Since Z_{l+1} is independant of $Y_k^2 Y_l$, we obtain $\mathrm{Cov}\left(Y_k^2, Y_l Z_{l+1}\right) = 0$.

(5) We have

$$\mathrm{Cov}\left(Y_k^2, Y_l^2\right) = \sum_{m,n,p,q=0}^{+\infty}\phi^{m+n+p+q}\mathbb{E}\left(Z_{k-m}Z_{k-n}Z_{l-p}Z_{l-q}\right) - \rho_Y(0)^2.$$

The terms of this sum are equal to zero except if the indices $k-m$, $k-n$, $l-p$, $l-q$ are all equal or equal in twos, then we deduce that

$$\mathrm{Cov}\left(Y_k^2, Y_l^2\right) = \frac{\sigma_Z^4}{1-\phi^2}\left(\frac{\phi^{2(l-k)}(\mathbb{E}(Z_0^4)-3\sigma_Z^4)}{\sigma_Z^4(1+\phi^2)} + \frac{2\phi^{2(l-k)}}{1-\phi^2}\right).$$

Now if we have $0 \leq l < k \leq n-1$, we can check the equalities (3) and (5) using symetric property. For the other ones:
(1)

$$\mathrm{Cov}\left(Y_k, Y_l^2\right) = \sum_{m,n,p=0}^{+\infty} \phi^{m+n+p} \mathbb{E}\left(Z_{k-m} Z_{l-n} Z_{l-p}\right).$$

The terms of this sum are equal to zero except if the indices $k-m$, $l-n$ and $l-p$ are equal then if $p=n$ and $m = n+k-l$.

$$\mathrm{Cov}\left(Y_k, Y_l^2\right) = \sum_{n=0}^{+\infty} \phi^{3n+k-l} \mathbb{E}\left(Z_{k-l}^3\right) = \frac{\mathbb{E}\left(Z_0^3\right)}{1-\phi^3} \phi^{k-l}.$$

(2)

$$\begin{aligned} \mathrm{Cov}\left(Y_k, Y_l Z_{l+1}\right) &= \sum_{m,n=0}^{+\infty} \phi^{m+n} \mathbb{E}\left(Z_{k-m} Z_{l-n} Z_{l+1}\right) \\ &= 0 \end{aligned}$$

since the centered variables Z_{k-m}, Z_{l-n} and Z_{l+1} cannot be all equal.

(4) We can prove by induction that

$$\mathrm{Cov}\left(Y_k^2, Y_l Z_{l+1}\right) = 2\frac{\sigma_Z^4}{1-\phi^2} \phi^{2(k-l)-1}.$$

Actually, for $k = l+1$

$$\begin{aligned} \mathrm{Cov}\left(Y_{l+1}^2, Y_l Z_{l+1}\right) &= \mathrm{Cov}\left(\left(\phi Y_l + Z_{l+1}\right)^2, Y_l Z_{l+1}\right) \\ &= \phi^2 \mathrm{Cov}\left(Y_l^2, Y_l Z_{l+1}\right) + \mathrm{Cov}(Z_{l+1}^2, Y_l Z_{l+1}) + 2\phi \mathrm{Var}\left(Y_l Z_{l+1}\right) \\ &= 0 + 0 + 2\phi \mathrm{Var}(Y_l)\mathrm{Var}(Z_{l+1}) \\ &= 2\frac{\phi}{1-\phi^2}\sigma_Z^4 \end{aligned}$$

which proves the property for $k = l+1$. Now assuming that the result holds for $k > l$ we get

$$\begin{aligned} \mathrm{Cov}\left(Y_{k+1}^2, Y_l Z_{l+1}\right) &= \mathrm{Cov}\left(\left(\phi Y_k + Z_{k+1}\right)^2, Y_l Z_{l+1}\right) \\ &= \phi^2 \mathrm{Cov}\left(Y_k^2, Y_l Z_{l+1}\right) + \mathrm{Cov}(Z_{k+1}^2, Y_l Z_{l+1}) + 2\phi \mathrm{Cov}\left(Y_k Z_{k+1}, Y_l Z_{l+1}\right). \end{aligned}$$

As $k > l$, the two last terms are equal to zero and

$$\begin{aligned} \mathrm{Cov}\left(Y_{k+1}^2, Y_l Z_{l+1}\right) &= \phi^2 \mathrm{Cov}\left(Y_k^2, Y_l Z_{l+1}\right) \\ &= 2\frac{\sigma_Z^4}{1-\phi^2} \phi^{2(k+1-l)-1}, \end{aligned}$$

which proves using induction that (4) holds for any $k \geq l+1$.

Hence, since $|\phi| < 1$, one can find a constant $c > 0$ such that

$$\left|\mathrm{Cov}\left(\Delta_{n,k}, \Delta_{n,l}\right)\right| \leq c|\phi|^{|k-l|}.$$

□

Since $Y_k^{m_n} = Y_k - \phi^{m_n+1} Y_{k-m_n-1}$, we can check that there exists $c > 0$ such that

$$\left|\mathrm{Cov}\left(\Delta_{n,k} - \Delta_{n,k}^{m_n}, \Delta_{n,l} - \Delta_{n,l}^{m_n}\right)\right| \leq c|\phi|^{2m_n}|\phi|^{|k-l|} \text{ and } \left|\mathrm{Cov}\left(\Delta_{n,k}^{m_n}, \Delta_{n,l}^{m_n}\right)\right| \leq c|\phi|^{|k-l|}.$$

Hence

$$\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \left(\Delta_{n,k} - \Delta_{n,k}^{m_n}\right)\right) \leq \frac{c|\phi|^{2m_n}}{1-|\phi|},$$

that proves (19) and i) will follow from the fact that $\frac{1}{n}\mathrm{Var}\left(\sum_{k=0}^{n-1}\Delta_{n,k}\right) \underset{n\to+\infty}{\to} l$. Hence, we compute

$$\begin{aligned}\mathrm{Var}\left(\sum_{k=0}^{n-1}\Delta_{n,k}\right) &= \mathrm{Var}\left(\sum_{k=0}^{n-1}\left(u+v\frac{k}{n}+wY_k+xZ_{k+1}\right)Y_k\right)\\ &= \sum_{k,l=0}^{n-1}\mathrm{Cov}\left(\left(u+v\frac{k}{n}+wY_k+xZ_{k+1}\right)Y_k,\left(u+v\frac{l}{n}+wY_l+xZ_{l+1}\right)Y_l\right)\end{aligned}$$

$$\begin{aligned}&= \frac{1}{n}\sum_{k,l=0}^{n-1}\left(u^2n+uv(k+l)+v^2\frac{kl}{n}\right)\mathrm{Cov}(Y_k,Y_l)+(uwn+vwk)\,\mathrm{Cov}\left(Y_k,Y_l^2\right)+(uwn+vwl)\mathrm{Cov}\left(Y_k^2,Y_l\right)\\ &+ \frac{1}{n}\sum_{k,l=0}^{n-1}(uxn+vkx)\mathrm{Cov}\left(Y_k,Y_lZ_{l+1}\right)+(uxn+vlx)\mathrm{Cov}\left(Y_kZ_{k+1},Y_l\right)+nx^2\mathrm{Cov}\left(Y_kZ_{k+1},Y_lZ_{l+1}\right)\\ &+ \frac{1}{n}\sum_{k,l=0}^{n-1}xwn\mathrm{Cov}\left(Y_k^2,Y_lZ_{l+1}\right)+xwn\mathrm{Cov}\left(Y_kZ_{k+1},Y_l^2\right)+nw^2\mathrm{Cov}\left(Y_k^2,Y_l^2\right).\end{aligned}$$

Now in order to get an explicit asymptotic variance we need the following computations.

Lemma 2. *For $|\phi|<1$, we have the following behaviors when n tends to $+\infty$:*

(1) $\sum_{i=0}^{n}\sum_{j=0}^{n}\phi^{|j-i|}=n\frac{1+\phi}{1-\phi}+o(n)$;
(2) $\sum_{i=0}^{n}\sum_{j=0}^{n}i\phi^{|j-i|}=\frac{n^2}{2}\frac{1+\phi}{1-\phi}+o(n^2)$;
(3) $\sum_{i=0}^{n}\sum_{j=0}^{n}ij\phi^{|j-i|}=\frac{n^3}{3}\frac{1+\phi}{1-\phi}+o(n^3)$;
(4) $\sum_{i=0}^{n}\sum_{j=0}^{n}\phi^{\frac{1}{2}(j-i)+\frac{3}{2}|j-i|}=n\frac{1+\phi+\phi^2}{1-\phi^2}+o(n)$;
(5) $\sum_{i=0}^{n}\sum_{j=i+1}^{n}\phi^{2(j-i)-1}=n\frac{\phi}{1-\phi^2}+o(n)$;
(6) $\sum_{i=0}^{n}\sum_{j=0}^{n}i\phi^{\frac{1}{2}(j-i)+\frac{3}{2}|j-i|}=\frac{n^2}{2}\frac{1+\phi+\phi^2}{1-\phi^2}+o(n^2)$.

Then we obtain that $\frac{1}{n}\mathrm{Var}\left(\sum_{k=0}^{n-1}\Delta_{n,k}\right)\underset{n\to\infty}{\to} l$, where

$$\begin{aligned}l &= u^2\frac{\sigma_Z^2}{1-\phi^2}\frac{1-\phi}{1+\phi}+uv\frac{\sigma_Z^2}{1-\phi^2}\frac{1+\phi}{1-\phi}+v^2\frac{\sigma_Z^2}{1-\phi^2}\frac{1+\phi}{3(1-\phi)}+2uw\frac{\mathbb{E}\left(Z_0^3\right)}{\sigma_Z^2(1-\phi^3)}\frac{1+\phi+\phi^2}{1-\phi^2}\\ &+ vw\frac{\mathbb{E}\left(Z_0^3\right)}{1-\phi^3}\frac{1+\phi+\phi^2}{2(1-\phi^2)}+w^2\frac{\sigma_Z^4}{1-\phi^2}\frac{1+\phi^2}{1-\phi^2}\left(\frac{\mathbb{E}\left(Z_0^4\right)-3\sigma_Z^4}{\sigma_Z^4(1+\phi^2)}+\frac{2}{1-\phi^2}\right)\\ &+ 4wx\frac{\sigma_Z^4}{1-\phi^2}\frac{\phi}{1-\phi^2}+x^2\frac{\sigma_Z^4}{1-\phi^2}\\ &= \frac{\sigma_Z^2}{(1-\phi)^2}\left(u^2+uv+2uw\frac{\mathbb{E}(Z_0^3)}{\sigma_Z^2(1+\phi)}+\frac{v^2}{3}+\frac{\mathbb{E}\left(Z_0^3\right)}{\sigma_Z^2(1+\phi)}vw\right)\\ &+ \frac{\sigma_Z^2}{(1-\phi)^2}\left(w^2\left(\frac{\left(\mathbb{E}\left(Z_0^4\right)-3\sigma_Z^4\right)}{\sigma_Z^2(1+\phi)^2}+2\frac{\sigma_Z^2\left(1+\phi^2\right)}{(1-\phi)(1+\phi)^3}\right)+4wx\frac{\sigma_Z^2\phi}{(1+\phi)^2}+x^2\sigma_Z^2\frac{1-\phi}{1+\phi}\right)\\ &= \begin{pmatrix}u & v & w & x\end{pmatrix}{\Sigma_1}^t\begin{pmatrix}u & v & w & x\end{pmatrix}\neq 0.\end{aligned}$$

For $ii)$ note that for $k < n$

$$|U_{n,k}| = \left|\frac{1}{\sqrt{n}}\Delta_{n,k}^{m_n}\right| \leq \frac{1}{\sqrt{n}}V_k,$$

where $V_k = |u| + |v|W_k + |w|(W_k^2 + \rho_Y(0)) + |x||Z_{k+1}|W_k$ with $W_k = \sum_{j=0}^{+\infty}|\phi|^j|Z_{k-j}|$. Since $\mathbb{E}(Z_0^4) < +\infty$, we have for all $\alpha \in (0,4]$

$$\|W_k\|_{L^\alpha} \leq \|Z_0\|_{L^\alpha}\frac{1}{1-|\phi|}.$$

Hence, by Cauchy-Schwarz inequality,

$$\|V_k\|_{L^2} \leq |u| + |v|\|W_k\|_{L^2} + |w|\left(\|W_k\|_{L^4}^2 + \rho_Y(0)\right) + |x|\|Z_{k+1}\|_{L^4}\|W_k\|_{L^4} \leq C,$$

where

$$C = |u| + |v|\|Z_0\|_{L^2}\frac{1}{1-|\phi|} + |w|\left(\|Z_0\|_{L^4}^2\frac{1}{(1-|\phi|)^2} + \rho_Y(0)\right) + |x|\|Z_0\|_{L^4}^2\frac{1}{1-|\phi|}.$$

and $ii)$ holds since $\sum_{k=0}^{n-1}\mathbb{E}(U_{n,k}^2) \leq \frac{1}{n}\sum_{k=0}^{n-1}\|V_k\|_{L^2}^2 \leq C^2$.
For $iii)$ we remark that since $\mathbb{E}(V_0^2) < +\infty$ one has

$$\mathbb{E}(V_k^2\mathbf{1}_{V_k>n^{1/8}}) = \mathbb{E}(V_0^2\mathbf{1}_{V_0>n^{1/8}}) \underset{n\to+\infty}{\longrightarrow} 0,$$

and we can choose $m_n = \min\left(\mathbb{E}(V_0^2\mathbf{1}_{V_0>n^{1/8}})^{-1/4}, n^{1/8}\right) \underset{n\to+\infty}{\longrightarrow} +\infty$, in such a way that for $\varepsilon > 0$,

$$\begin{aligned} L_n(\varepsilon) &= m_n^2\sum_{k=0}^{n-1}\mathbb{E}\left(U_{n,k}^2\mathbf{1}_{m_n^2|U_{n,k}|>\varepsilon}\right) \\ &\leq m_n^2\sum_{k=0}^{n-1}\frac{1}{n}\mathbb{E}\left(V_k^2\mathbf{1}_{\frac{m_n^2}{\sqrt{n}}|V_k|>\varepsilon}\right) \\ &\leq m_n^2\mathbb{E}(V_0^2\mathbf{1}_{V_0>\varepsilon n^{1/4}}), \end{aligned}$$

where we have used the fact that $\frac{m_n^2}{\sqrt{n}} \leq n^{-1/4}$ in the last inequality. Since $\varepsilon n^{1/4} > n^{1/8}$, for n sufficiently large, we obtain

$$L_n(\varepsilon) \leq m_n^{-2},$$

such that $iii)$ holds.

6.3. **Proof of Theorem 2.** Let us write

$$H_n(X) = {}^tA_n(X)A_n(X) = \begin{pmatrix} n & \sum_{k=0}^{n-1}\frac{k}{n} & \sum_{k=0}^{n-1}X_k \\ \sum_{k=0}^{n-1}\frac{k}{n} & \sum_{k=0}^{n-1}\left(\frac{k}{n}\right)^2 & \sum_{k=0}^{n-1}\frac{k}{n}X_k \\ \sum_{k=0}^{n-1}X_k & \sum_{k=0}^{n-1}\frac{k}{n}X_k & \sum_{k=0}^{n-1}X_k^2 \end{pmatrix}.$$

In view of (3), if $(X_{n,k})$ satisfies (2) for some $\theta = (m, b, \phi) \in \mathbb{R}^2\times]-1,1[$, we can write $X_{n,k} = c_n + d\frac{k}{n} + Y_k$ with Y a stationary centered solution of the AR(1) equation (4), namely

$$Y_{k+1} = \phi Y_k + Z_{k+1},$$

and c_n, d are given by (5) ie $c_n = c - \frac{b}{n(1-\phi)^2}$, $d = \frac{b}{1-\phi}$ and $c = \frac{m}{1-\phi}$.

Hence, according to Corollary 1 and Lemma 1 above, the following convergences hold a.s and in L^2:

$$\begin{aligned} \frac{1}{n}\sum_{k=0}^{n-1}X_{n,k} &= c_n + d\frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n} + \frac{1}{n}\sum_{k=0}^{n-1}Y_k \underset{n\to+\infty}{\longrightarrow} c + d/2, \\ \frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}X_{n,k} &= c_n\frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n} + d\frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{k}{n}\right)^2 + \frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}Y_k \underset{n\to+\infty}{\longrightarrow} c/2 + d/3, \\ \frac{1}{n}\sum_{k=0}^{n-1}X_{n,k}^2 &= c_n^2 + 2c_nd\frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n} + d^2\frac{1}{n}\sum_{k=0}^{n-1}(\frac{k}{n})^2 + \frac{1}{n}\sum_{k=0}^{n-1}Y_k^2 + +2c_n\frac{1}{n}\sum_{k=0}^{n-1}Y_k + 2d\frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}Y_k \\ &\underset{n\to+\infty}{\longrightarrow} c^2 + d^2/3 + cd + \rho_Y(0), \end{aligned}$$

with $\rho_Y(0) = \sigma_Z^2/(1-\phi^2)$ as the stationary solution of the AR(1) equation. Therefore

$$\frac{1}{n}H_n(X) \underset{n\to+\infty}{\longrightarrow} H := \begin{pmatrix} 1 & 1/2 & c+d/2 \\ 1/2 & 1/3 & c/2+d/3 \\ c+d/2 & c/2+d/3 & c^2+d^2/3+cd+\rho_Y(0) \end{pmatrix} \text{ a.s. and in } L^2.$$

Note that for all $(u,v,w) \in \mathbb{R}^3$,

$$\begin{aligned} \begin{pmatrix} u & v & w \end{pmatrix} H \, {}^t\begin{pmatrix} u & v & w \end{pmatrix} &= (u+cw)^2 + \frac{1}{3}(v+dw)^2 + w^2\rho_Y(0) + (u+cw)(v+dw) \\ &= \left((u+cw)+\frac{1}{2}(v+dw)\right)^2 + \frac{1}{12}(v+dw)^2 + w^2\rho_Y(0), \end{aligned}$$

and therefore the matrix H is positive definite.

Now let us consider for $\tilde{\theta} = (\tilde{m}, \tilde{b}, \tilde{\phi}) \in \mathbb{R}^3$, the contrast function

$$M_n(\tilde{\theta}) = {}^t(X_n - A_n(X)\tilde{\theta})(X_n - A_n(X)\tilde{\theta}) = \sum_{k=0}^{n-1}\left(X_{n,k+1} - \tilde{\phi}X_{n,k} - \tilde{m} - \tilde{b}\frac{k}{n}\right)^2,$$

where $X_n := (X_{n,k+1})_{0\le k\le n-1}$. Let us write $\theta = (m,b,\phi)$ the true parameter such that

$$X_{n,k+1} - \phi X_{n,k} - m - b\frac{k}{n} = Z_{k+1},$$

meaning that $X_n - A_n(X)\theta = Z$ for $Z := (Z_{k+1})_{0\le k\le n-1}$. Then $\hat{\theta}_n = \operatorname{argmin}_{\tilde{\theta}\in\mathbb{R}^3} M_n(\tilde{\theta})$ satisfies $J_n(\hat{\theta}_n) = 0$ for $J_n = \nabla M_n$. But $J_n(\tilde{\theta}) = -2{}^tA_n(X)(X_n - A_n(X)\tilde{\theta})$ and $J_n(\theta) = J_n(\theta) - J_n(\hat{\theta}_n) = -2H_n(X)(\theta - \hat{\theta}_n)$. On the other hand, since $X_n - A_n(X)\theta = Z$, we get

$$J_n(\theta) = -2\begin{pmatrix} \sum_{k=0}^{n-1} Z_{k+1} \\ \sum_{k=0}^{n-1} \frac{k}{n} Z_{k+1} \\ \sum_{k=0}^{n-1} X_{n,k}Z_{k+1} \end{pmatrix}.$$

We will prove that $\frac{-1}{2n}J_n(\theta) \underset{n\to+\infty}{\longrightarrow} 0$ a.s. and in L^2 with $\frac{-1}{2\sqrt{n}}J_n(\theta) \underset{n\to+\infty}{\overset{d}{\to}} \mathcal{N}(0,\Sigma)$, with $\Sigma = \sigma_Z^2 H$. Since $(X_{n,k})$ satisfies (3) for θ the true parameter, we can write

$$\frac{1}{n}\sum_{k=0}^{n-1} X_{n,k}Z_{k+1} = c_n \times \frac{1}{n}\sum_{k=0}^{n-1} Z_{k+1} + d \times \frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}Z_{k+1} + \frac{1}{n}\sum_{k=0}^{n-1} Y_k Z_{k+1}.$$

Each empirical mean will tend to 0 a.s. and in L^2 according to Corollary 1 (choosing $\phi = 0$ for i)). This allows to conclude for the first point using the fact that (c_n) is a bounded sequence. Using the fact that $(\theta - \hat{\theta}_n) = \left(\frac{1}{n}H_n\right)^{-1}\frac{-1}{2n}J_n(\theta)$ we can conclude for the a.s. convergence. Now to prove the convergence in distribution, we use again the Cramer-Wold device (see Proposition 6.3.1 of [5] for instance) and consider for $(u,v,w) \in \mathbb{R}^3 \setminus \{0\}$,

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\left(u + v\frac{k}{n} + wX_{n,k}\right)Z_{k+1} = \frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\left((u+c_nw) + (v+dw)\frac{k}{n} + wY_k\right)Z_{k+1}.$$

The convergence will follow from a Lindeberg condition for triangular array of martingales [6]. Actually, let us write $\Delta_{n,k+1} = \left((u+c_nw) + (v+dw)\frac{k}{n} + wY_k\right)Z_{k+1}$ and $S_{n,l} = \sum_{k=0}^{l}\Delta_{n,k+1}$. We may consider the filtration $\mathcal{F}_{n,l} = \sigma(Z_k, k\le l) = \mathcal{F}_l$. It follows that

$$\mathbb{E}(\Delta_{n,k+1}|\mathcal{F}_{n,k}) = \mathbb{E}(\Delta_{n,k+1}|\mathcal{F}_k) = \left((u+c_nw) + (v+dw)\frac{k}{n} + wY_k\right)\mathbb{E}(Z_{k+1}) = 0 \text{ a.s.}$$

since Z_{k+1} is centered and independent from $\mathcal{F}_k$ and $(S_{n,l})$ is a martingale triangular array. Then let us write $\frac{S_n}{s_n} := \frac{S_{n,n-1}}{s_n}$, where $s_n^2 = \mathrm{Var}(S_{n,n-1})$. Hence according to Theorem 2 of [6] or [15] if

$$\frac{1}{s_n^2}\sum_{k=1}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k) \quad \xrightarrow[n\to+\infty]{\mathbb{P}} \quad 1, \tag{20}$$

$$\text{and } \frac{1}{s_n^2}\sum_{k=1}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2\mathbf{1}_{|\Delta_{n,k+1}|>\varepsilon s_n}|\mathcal{F}_k) \quad \xrightarrow[n\to+\infty]{\mathbb{P}} \quad 0, \text{ for all } \varepsilon>0, \tag{21}$$

we have $\frac{S_n}{s_n} \xrightarrow[n\to+\infty]{d} \mathcal{N}(0,1)$.

So let us first compute the asymptotic variance. We write $\langle S\rangle_n = \sum_{k=1}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k)$ and note that $s_n^2 = \mathbb{E}(\langle S\rangle_n)$. In our setting, it is clear that

$$\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k) = \sigma_Z^2\left((u+c_nw)+(v+dw)\frac{k}{n}+wY_k\right)^2.$$

By Corollary 1 we have $\frac{1}{n}\sum_{k=0}^{n-1}Y_k\to 0$, $\frac{1}{n}\sum_{k=0}^{n-1}\frac{k}{n}Y_k\to 0$ and $\frac{1}{n}\sum_{k=0}^{n-1}Y_k^2\to\rho_Y(0)$ a.s. and therefore,

$$\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k) \underset{n\to+\infty}{\longrightarrow} s^2 := \sigma_Z^2\left((u+cw)^2+\frac{1}{3}(v+dw)^2+w^2\rho_Y(0)+(u+cw)(v+dw)\right) \text{a.s.}$$

Note that we may deduce from these lines that the asymptotic covariance matrix is given by

$$\Sigma = \sigma_Z^2\begin{pmatrix} 1 & \frac{1}{2} & c+\frac{d}{2} \\ \frac{1}{2} & \frac{1}{3} & \frac{c}{2}+\frac{d}{3} \\ c+\frac{d}{2} & \frac{c}{2}+\frac{d}{3} & c^2+\frac{d^2}{3}+\rho_Y(0)+cd \end{pmatrix} = \sigma_Z^2 H.$$

Moreover, since $(u,v,w)\neq 0$ and H is positive definite we check that $s^2>0$. Then,

$$\frac{s_n^2}{n} = \mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k)\right) \underset{n\to+\infty}{\longrightarrow} s^2,$$

and (20) follows by writing

$$\frac{1}{s_n^2}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k) = \frac{n}{s_n^2}\times\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2|\mathcal{F}_k).$$

Now it remains to prove the Lindenberg condition (21). So let us fix $N\in\mathbb{N}^*$ large enough such that for all $n\geq 2N$ we have $s_n > s^2\sqrt{N}$, choose $C>0$ such that $|\Delta_{n,k+1}|\leq C(1+|Y_k|)|Z_{k+1}|$ and remark that for $n\geq 2N$, we have

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2\mathbf{1}_{|\Delta_{n,k+1}|>\varepsilon s_n}|\mathcal{F}_k) &\leq \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2\mathbf{1}_{|\Delta_{n,k+1}|>\varepsilon s\sqrt{N}}|\mathcal{F}_k)\\ &\leq \frac{C^2}{n}\sum_{k=0}^{n-1}\mathbb{E}((1+|Y_k|)^2Z_{k+1}^2\mathbf{1}_{|C(1+|Y_k|)|Z_{k+1}|>\varepsilon s\sqrt{N}}|\mathcal{F}_k).\end{aligned}$$

Hence, by stationarity, we get

$$\mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}(\Delta_{n,k+1}^2\mathbf{1}_{|\Delta_{n,k+1}|>\varepsilon s_n}|\mathcal{F}_k)\right) = C^2\mathbb{E}((1+|Y_0|)^2Z_1^2\mathbf{1}_{|(1+|Y_0|)|Z_1|>\varepsilon s\sqrt{N}/C)}) \underset{N\to+\infty}{\longrightarrow} 0,$$

since $\mathbb{E}((1+|Y_0|)^2Z_1^2)<+\infty$, and allows to get (21). We have therefore $\frac{S_n}{s_n}\xrightarrow[n\to+\infty]{d}\mathcal{N}(0,1)$ and consequently, by Slutsky's theorem, $\frac{-1}{2\sqrt{n}}J_n(\theta) = \frac{s_n}{\sqrt{n}}\frac{S_n}{s_n}\xrightarrow[n\to+\infty]{d}\mathcal{N}(0,s^2)$. But $\frac{-1}{2\sqrt{n}}J_n(\theta) = \frac{1}{n}H_n(X)\sqrt{n}(\theta-\hat{\theta}_n)$, and we can write

$$\sqrt{n}(\theta-\hat{\theta}_n) = \left(\frac{1}{n}H_n(X)\right)^{-1}\frac{-1}{2\sqrt{n}}J_n(\theta).$$

Again, by Slutsky's theorem, we may deduce that

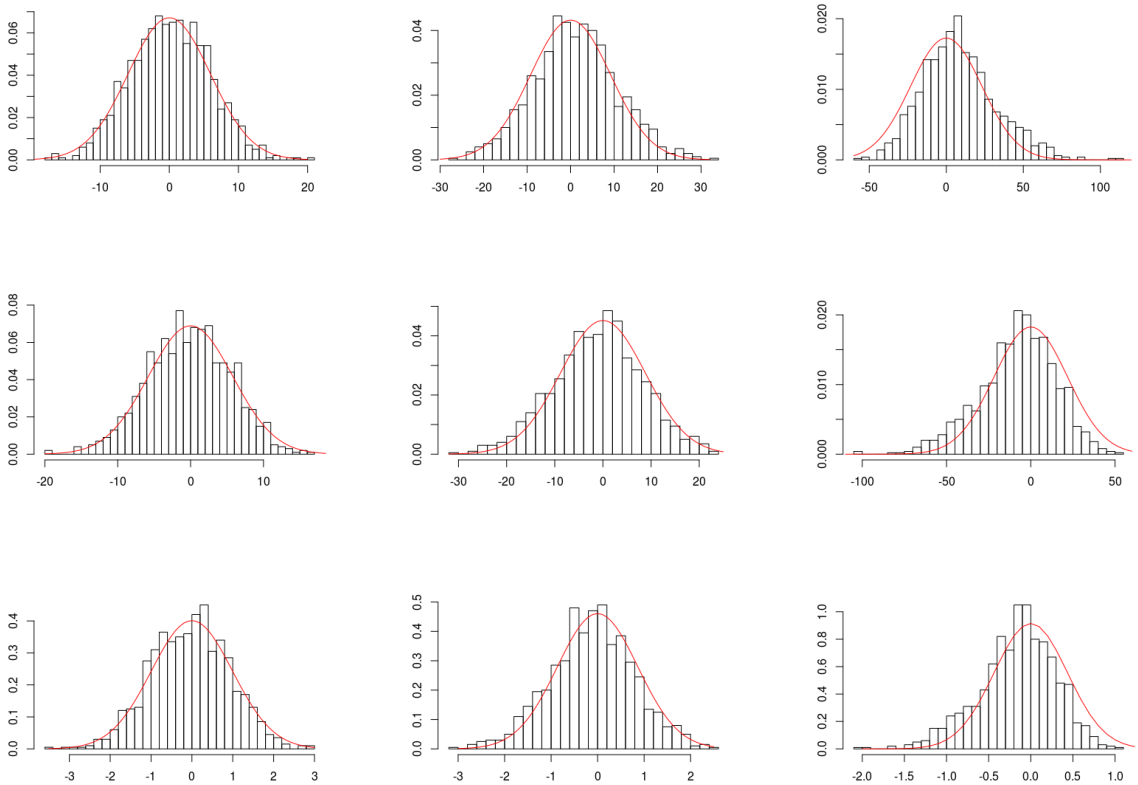
$$\sqrt{n}(\theta - \hat{\theta}_n) \xrightarrow[n\to+\infty]{d} H^{-1}\mathcal{N}(0,\Sigma) = \mathcal{N}\left(0, H^{-1}\Sigma^t H^{-1}\right) = \mathcal{N}\left(0, \sigma_Z^2 H^{-1}\right).$$

Note that $\det(H) = \frac{\rho_Y(0)}{12}$, so

$$\Sigma_2 = \sigma_Z^2 H^{-1} = (1-\phi^2)\begin{pmatrix} c^2 + 4\rho_Y(0) & cd - 6\rho_Y(0) & -c \\ cd - 6\rho_Y(0) & d^2 + 12\rho_Y(0) & -d \\ -c & -d & 1 \end{pmatrix},$$

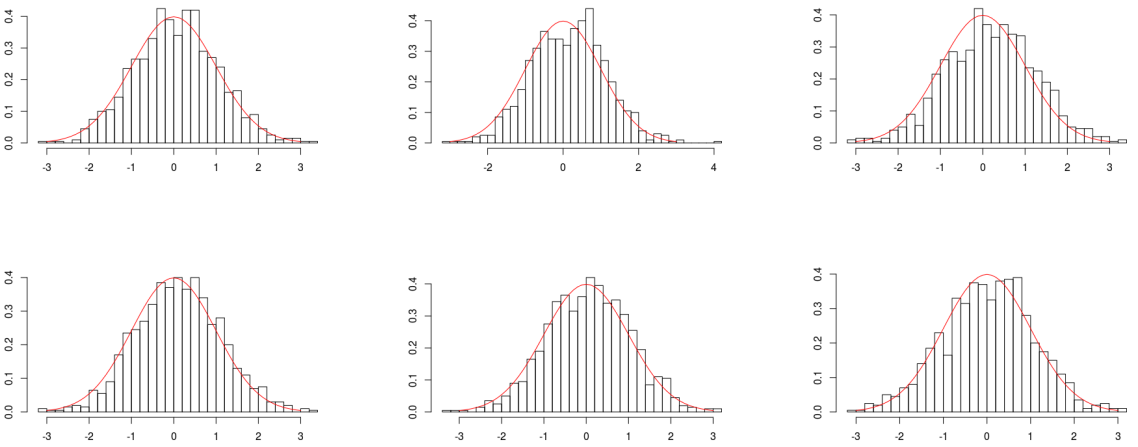
with $c = \frac{m}{1-\phi}$, $d = \frac{b}{1-\phi}$ and $\rho_Y(0) = \frac{1}{1-\phi^2}\sigma_Z^2$.

6.4. **Numerical results.** We first present some histograms obtained with 1000 simulations, $\nu = 0.3$, $\lambda = 1$, $\sigma = 0.2$, and $n = 1000$. The red lines correspond to the theoretical gaussian distribution with variance computed according to Theorem 2.



$\phi = 0.1$ $\phi = 0.5$ $\phi = 0.9$

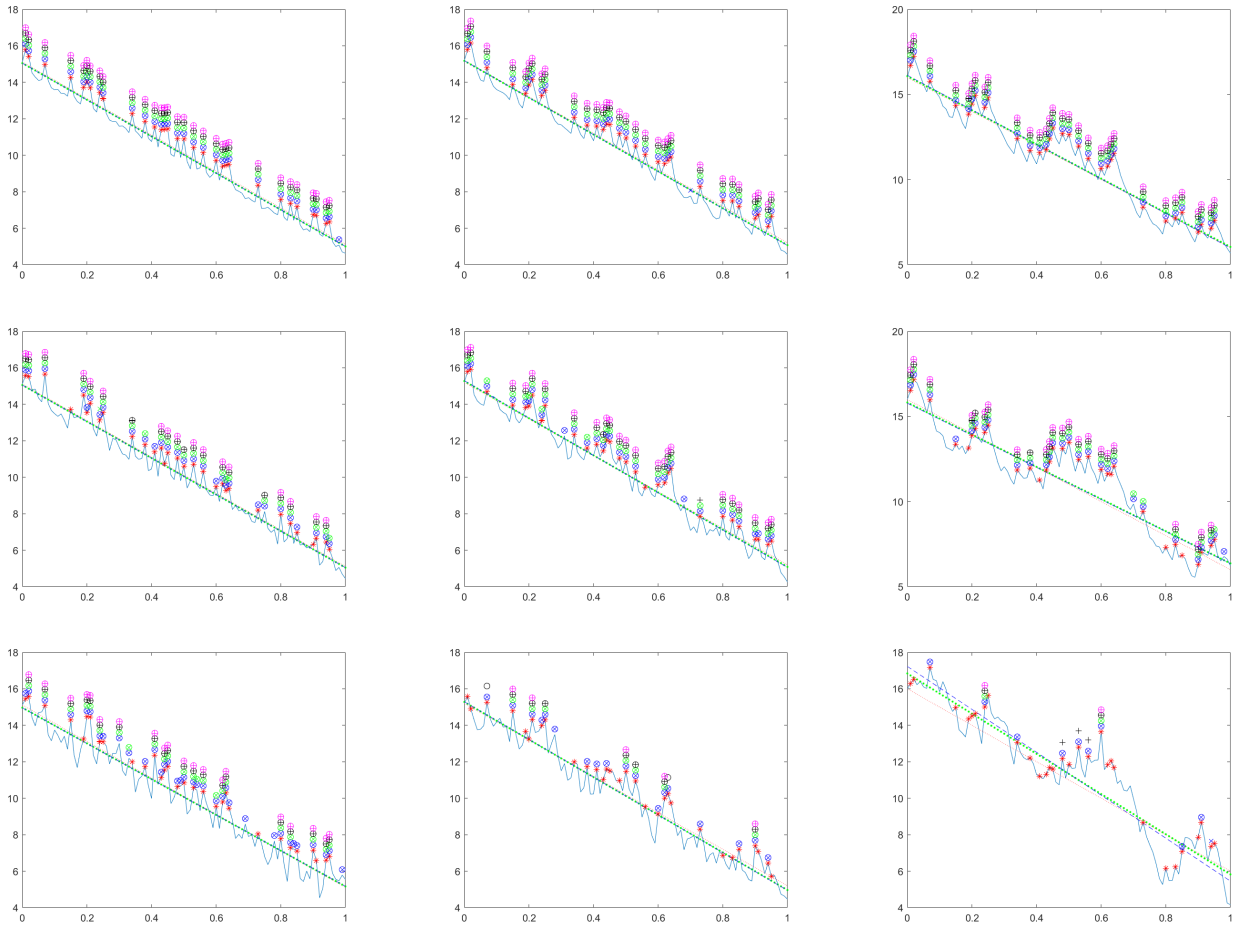
FIGURE 21. Histograms for $a = 5$, $b = -5$. First row: $(\widehat{m}_n^{(1)} - m)$. Second row: $(\widehat{b}_n^{(1)} - b)$. Third row: $(\widehat{\phi}_n^{(1)} - \phi)$.



$\phi = 0.1$ $\phi = 0.5$ $\phi = 0.9$

FIGURE 22. Histograms for $c = 5$, $d = -5$, $n = 1000$. First row: $(\widehat{c}_n^{(1)} - c)$. Second row: $(\widehat{d}_n^{(1)} - d)$.

Now we present typical realizations of trajectories for several choices of parameters but with the same jumps (in red star) in order to illustrate the parameters effect on shapes and tests. We set $\nu = 0.3$, $\lambda = 1$, $n = 100$ and consider tests: in blue cross detected jumps with $T^{(1)}$ for level $\alpha = 0.01$, in green cross detected jumps with the multiple test correction $\tilde{T}^{(1)}$ for level $\beta = 0.01$, in black plus detected jumps with the multiple test correction $\tilde{T}_c^{(1)}$ for level $\beta = 0.01$ and magenta plus indicate the detected jumps for $T_c^{(1)}$ for level $\alpha = 0.01$. We add circle for the corresponding detected jumps using $T^{(2)}$ in blue, $\tilde{T}^{(2)}$ in green, $\tilde{T}_c^{(2)}$ in black and $T_c^{(2)}$ in magenta. The blue line, respectively green line, is the estimated line with the first, respectively the second, estimators. The dot red line corresponds to the straight line with parameters (c_n, d) induced by our choices of values. The corresponding true and false positive rates are given in tables. We first present results for $c = 15$, $d = -10$ in Figure 23 and Tables 2 and 3.



$\phi = 0.1$ $\phi = 0.5$ $\phi = 0.9$

FIGURE 23. Test over one simulation for $c = 15$, $d = -10$, $b = d(1-\phi)$, $a = c(1-\phi) - \lambda\nu$, $X_0 = c_n$, with $c_n = c - d/(n(1-\phi))$. First line: $\sigma = 0.1$, second line: $\sigma = 0.3$, third line: $\sigma = 0.5$.

	$\sigma = 0.1$			$\sigma = 0.3$			$\sigma = 0.5$		
ϕ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$T^{(1)}$	1	1	1	0.9032	0.9677	0.8387	0.8710	0.5484	0.3226
$T^{(2)}$	1	1	1	0.9032	0.9677	0.8387	0.8710	0.5484	0.2903
$\tilde{T}^{(1)}$	1	1	1	0.7419	0.9355	0.8065	0.6452	0.2258	0.0645
$\tilde{T}^{(2)}$	1	1	1	0.7097	0.9355	0.8065	0.6452	0.2258	0.0645
$\tilde{T_c}^{(1)}$	1	1	1	0.6219	0.8065	0.7419	0.6129	0.2258	0.1613
$\tilde{T_c}^{(2)}$	1	1	1	0.6219	0.7742	0.7419	0.6129	0.2903	0.0645
$T_c^{(1)}$	1	1	1	0.5484	0.7419	0.6774	0.6129	0.1613	0.0645
$T_c^{(2)}$	1	1	1	0.5484	0.7419	0.6774	0.6129	0.1613	0.0645

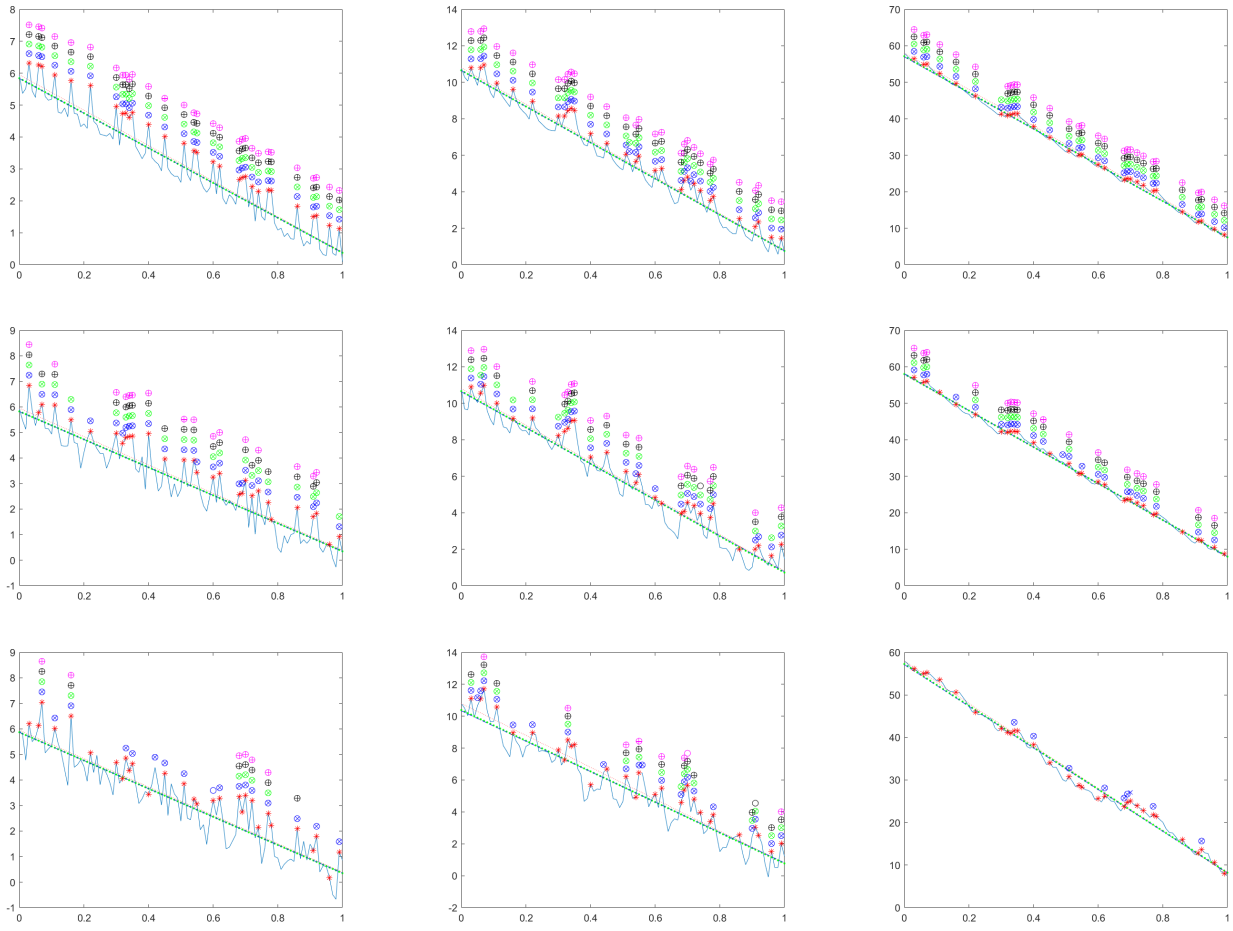
TABLE 2. TPR for tests over one simulation for $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $b = d(1-\phi)$, $a = c(1-\phi) - \lambda\nu$, $X_0 = c_n$, with $c_n = c - d/(n(1-\phi))$ and $n = 100$, as shown in Figure 23 with tolerance $\alpha = \beta = 0.01$. The number of jumps is 31.

	$\sigma = 0.1$			$\sigma = 0.3$			$\sigma = 0.5$		
ϕ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$T^{(1)}$	0.0145	0.0145	0	0.0145	0.0290	0.0290	0.1159	0.0145	0
$T^{(2)}$	0.0145	0	0	0.0145	0.0290	0.0290	0.1159	0.0145	0
$\tilde{T}^{(1)}$	0	0	0	0.0145	0	0.0145	0.0435	0	0
$\tilde{T}^{(2)}$	0	0	0	0.0145	0	0.0145	0.0435	0	0
$\tilde{T_c}^{(1)}$	0	0	0	0.0145	0	0	0.0145	0	0
$\tilde{T_c}^{(2)}$	0	0	0	0.0145	0	0	0.0145	0	0
$T_c^{(1)}$	0	0	0	0	0	0	0.0145	0	0
$T_c^{(2)}$	0	0	0	0	0	0	0.0145	0	0

TABLE 3. FPR for tests over one simulation for $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $b = d(1-\phi)$, $a = c(1-\phi) - \lambda\nu$, $X_0 = c_n$, with $c_n = c - d/(n(1-\phi))$ and $n = 100$, as shown in Figure 23 with tolerance $\alpha = \beta = 0.01$. The number of jumps is 31.

	$\sigma = 0.1$			$\sigma = 0.3$			$\sigma = 0.5$		
ϕ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$T^{(1)}$	1	1	1	0.9	0.9	0.7333	0.5	0.6	0.3
$T^{(2)}$	1	1	1	0.9	0.9	0.7333	0.5333	0.6	0.2667
$\tilde{T}^{(1)}$	1	1	1	0.7333	0.8	0.6333	0.2	0.4667	0
$\tilde{T}^{(2)}$	1	1	1	0.7333	0.8	0.6333	0.2	0.4667	0
$\tilde{T_c}^{(1)}$	1	1	0.9667	0.6667	0.6333	0.6333	0.2333	0.4	0
$\tilde{T_c}^{(2)}$	1	1	0.9667	0.6667	0.6667	0.6333	0.2333	0.4333	0
$T_c^{(1)}$	1	1	0.9667	0.5333	0.6333	0.5667	0.2	0.2333	0
$T_c^{(2)}$	1	1	0.9667	0.5333	0.6333	0.5667	0.2	0.6667	0

TABLE 4. TPR for tests over one simulation for $\nu = 0.3$, $\lambda = 1$, $a = 5$, $b = -5$, $b = d(1-\phi)$, $a = c(1-\phi) - \lambda\nu$, $X_0 = c_n$, with $c_n = c - d/(n(1-\phi))$ and $n = 100$, as shown in Figure 24 with tolerance $\alpha = \beta = 0.01$. The number of jumps is 30.



$\phi = 0.1$ $\phi = 0.5$ $\phi = 0.9$

FIGURE 24. Test over one simulation for $a = 5$, $b = -5$, $d = b/(1-\phi)$, $c_n = (a + \lambda\nu)/(1-\phi) - b/(n(1-\phi)^2)$, $X_0 = c_n$. First line: $\sigma = 0.1$, second line: $\sigma = 0.3$, third line: $\sigma = 0.5$.

	$\sigma = 0.1$			$\sigma = 0.3$			$\sigma = 0.5$		
ϕ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$T^{(1)}$	0	0.0286	0	0	0	0.0286	0.0143	0.0571	0
$T^{(2)}$	0	0.0286	0	0	0	0.0286	0.0143	0.0571	0
$\tilde{T}^{(1)}$	0	0.0143	0	0	0	0.0143	0	0.0143	0
$\tilde{T}^{(2)}$	0	0.0143	0	0	0	0.0143	0	0.0143	0
$\tilde{T_c}^{(1)}$	0	0	0	0	0	0.0143	0	0.0143	0
$\tilde{T_c}^{(2)}$	0	0	0	0	0	0.0143	0	0.0143	0
$T_c^{(1)}$	0	0	0	0	0	0.0143	0	0	0
$T_c^{(2)}$	0	0	0	0	0	0	0.0143	0	0

TABLE 5. FPR for tests over one simulation for $\nu = 0.3$, $\lambda = 1$, $a = 5$, $b = -5$, $b = d(1-\phi)$, $a = c(1-\phi) - \lambda\nu$, $X_0 = c_n$, with $c_n = c - d/(n(1-\phi))$ and $n = 100$, as shown in Figure 24 with tolerance $\alpha = \beta = 0.01$. The number of jumps is 30.

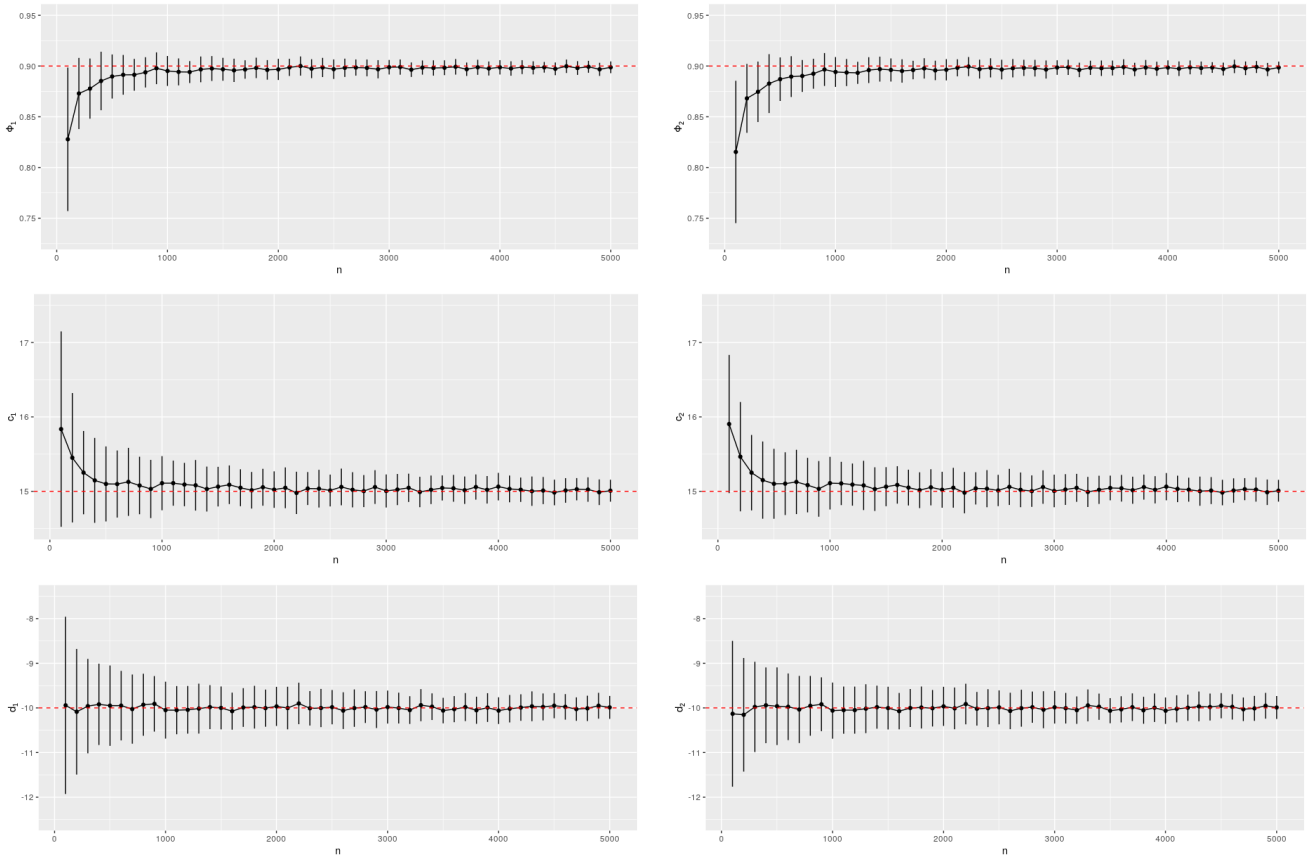


FIGURE 25. Means and standard deviation of the estimators of ϕ, c_n and d, according Theorem 2 (first column) or linear regression (second colum), over 100 simulations of $(X_{n,k})_{0\leq k\leq n}$, for $n \in \{100, 200, ..., 5000\}$ with $\phi = 0.9$, $\nu = 0.3$, $\lambda = 1$, $c = 15$, $d = -10$, $X_0 = c_n$, with $c_n = c - d/(n(1-\phi))$

REFERENCES

[1] F. Ali and A. C. Kwan. Interpreting in vivo calcium signals from neuronal cell bodies, axons, and dendrites: a review. *Neurophotonics*, 7(1):1 – 12, 2019.

[2] K. B. Athreya and S. G. Pantula. A note on strong mixing of ARMA processes. *Statist. Probab. Lett.*, 4(4):187–190, 1986.

[3] J. Behboodian. Information matrix for a mixture of two normal distributions. *Journal of Statistical Computation and Simulation*, 1(4):295–314, 1972.

[4] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57(1):289–300, 1995.

[5] P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer Series in Statistics. Springer, New York, 2006. Reprint of the second (1991) edition.

[6] B. M. Brown. Martingale central limit theorems. *Ann. Math. Statist.*, 42:59–66, 1971.

[7] A.N. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological Cybernetics*, 95:1–19, 2006.

[8] A. Caraty, P. Orgeur, and J.-C. Thiery. Demonstration of the pulsatile secretion of LH-RH into hypophysial portal blood of ewes using an original technic for multiple samples. *Comptes rendus des séances de l'Académie des sciences. Série III, Sciences de la vie*, 295:103–6, 10 1982.

[9] I. J. Clarke and J. T. Cummins. The temporal relationship between gonodotropin releasing hormone (GnRH) and luteinizing (LH) secretion in ovariectomized ewes. *Endocrinology*, 111(5):1737–1739, 1982.

[10] S. Constantin, A. Caraty, S. Wray, and A. Duittoz. Development of gonadotropin-releasing hormone-1 secretion in mouse nasal explants. *Endocrinology*, 150(7):3221–3227, July 2009.

[11] N. P. Dick and D. C. Bowden. Maximum likelihood estimation for mixtures of two normal distributions. *Biometrics*, 29(4):781–790, 1973.

[12] S. Ditlevsen and A. Samson. Parameter estimation in neuronal stochastic differential equation models from intracellular recordings of membrane potentials in single neurons: a review. *J. SFdS*, 157(1):6–21, 2016.

[13] P. Doukhan. *Stochastic models for time series*, volume 80 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Cham, 2018.

[14] C. Fraley, A. E. Raftery, T. B. Murphy, and L. Scrucca. MCLUST Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. *Technical Report No. 597*, 2012.
[15] P. Gaenssler, J. Strobel, and W. Stute. On central limit theorems for martingale triangular arrays. *Acta Math. Acad. Sci. Hungar.*, 31(3-4):205–216, 1978.
[16] C. Georgelin, C. Constant, H. Biermé, G. Chevrot, B. Piégu, R. Fleurot, G. Leng, and A. Duittoz. GnRH paracrine/autocrine action induced a non-stochastic behaviour andepisodic synchronisation of gnrh neurons activity: in vitro and in silico study. In preparation.
[17] W. Gerstner and W. Kistler. *Spiking Neuron Models: An Introduction*. Cambridge University Press, New York, NY, USA, 2002.
[18] D. M. Green and J. A. Swets. *Signal detection theory and psychophysics*. John Wiley & Sons Ltd, 1966.
[19] L. Heinrich. Asymptotic behaviour of an empirical nearest-neighbour distance function for stationary Poisson cluster processes. *Math. Nachr.*, 136:131–148, 1988.
[20] P. Jahn, R.W. Berg, J. Hounsgaard, and S. Ditlevsen. Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. *J. of Comput. Neurosci.*, 31:563–579, 2011.
[21] C. F. Jeff Wu. On the convergence properties of the EM algorithm. *Ann. Statist.*, 11(1):95–103, 03 1983.
[22] C. Kikawa, M. Shatalov, P. Kloppers, and A. Mkolesia. Parameter estimation for a mixture of two univariate gaussian distributions: A comparative analysis of the proposed and maximum likelihood methods. *Journal of Advances in Mathematics and Computer Science*, 12:1–8, 2015.
[23] R. J. Kulperger. On the residuals of autoregressive processes and polynomial regression. *Stochastic Process. Appl.*, 21(1):107–118, 1985.
[24] J. E. Levine, K. Y. Pau, V. D. Ramirez, and G. L. Jackson. Simultaneous measurement of luteinizing hormone-releasing hormone and luteinizing hormone release in unanesthetized, ovariectomized sheep. *Endocrinology*, 111(5):1449–1455, 1982.
[25] N. A. Macmillan and C. D. Creelman. *Detection Theory: A User's Guide*. Psychology Press. Taylor & Francis, 2004. Reprint of the second (1991) edition.
[26] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley Series in Probability and Statistics. Wiley, 2000.
[27] S. M. Moenter. GnRH neuron electrophysiology: A decade of study. *Brain Research*, 1364:10–24, October 2010.
[28] P. Müller, G. Parmigiani, C. Robert, and J. Rousseau. Optimal sample size for multiple testing: the case of gene expression microarrays. *J. Amer. Statist. Assoc.*, 99(468):990–1001, 2004.
[29] S. K. Ng, T. Krishnan, and G. J. McLachlan. *The EM algorithm*. Springer Handb. Comput. Stat. Springer, Heidelberg, 2012.
[30] P. Perron and T. Yabu. Testing for trend in the presence of autoregressive error: a comment. *J. Amer. Statist. Assoc.*, 107(498):844, 2012.
[31] D. Qiu, Q. Shao, and L. Yang. Efficient inference for autoregressive coefficients in the presence of trends. *J. Multivariate Anal.*, 114:40–53, 2013.
[32] A. Roy, B. Falk, and W. A. Fuller. Testing for trend in the presence of autoregressive error. *J. Amer. Statist. Assoc.*, 99(468):1082–1091, 2004.
[33] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.
[34] S. Wray. From nose to brain: development of gonadotrophin-releasing hormone-1 neurones. *Journal of neuroendocrinology*, 22(7):743–753, July 2010.

Hermine Biermé and Camille Constant, LMA UMR CNRS 7348, Université de Poitiers, Bât. H3 - Site du Futuroscope, TSA 61125, 11 bd Marie et Pierre Curie, 86073 Poitiers Cedex 9, France
Email address: hermine.bierme@math.univ-poitiers.fr, camille.constant@math.univ-poitiers.fr

Anne Duittoz, CNRS, IFCE, INRAE, Université de Tours, PRC, F-37380, Nouzilly, France
Email address: anne.duittoz@univ-tours.fr

Christine Georgelin, IDP UMR CNRS 7013, Université de Tours, Université d'Orléans, Parc de Grandmont 37200 Tours, France
Email address: christine.georgelin@univ-tours.fr