HAL open science

Hypoelliptic stochastic FitzHugh-Nagumo neuronal model: mixing, up-crossing and estimation of the spike rate

Jose R. Leon, Adeline Samson

To cite this version:

Jose R. Leon, Adeline Samson. Hypoelliptic stochastic FitzHugh-Nagumo neuronal model: mixing, up-crossing and estimation of the spike rate. The Annals of Applied Probability, 2018, 28 (4), pp.2243-2274. 10.1214/17-AAP1355 . hal-01492590

HAL Id: hal-01492590

https://hal.science/hal-01492590v1

Submitted on 20 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hypoelliptic stochastic FitzHugh-Nagumo neuronal model: mixing, up-crossing and estimation of the spike rate

José R. León[*,†,‡] **and Adeline Samson**[†]

Abstract: The FitzHugh-Nagumo is a well-known neuronal model that describes the generation of spikes at the intracellular level. We study a stochastic version of the model from a probabilistic point of view. The hypoellipticity is proved, as well as the existence and uniqueness of the stationary distribution. The bi-dimensional stochastic process is β-mixing. The stationary density can be estimated with an adaptive non-parametric estimator. Then, we focus on the distribution of the length between successive spikes. Spikes are difficult to define directly from the continuous stochastic process. We study the distribution of the number of up-crossings. We link it to the stationary distribution and propose an estimator of its expectation. We finally prove mathematically that the mean length of inter-up-crossings interval in the FitzHugh-Nagumo model is equal to its up-crossings rate. We illustrate the proposed estimators on a simulation study. Different regimes are explored, with no, few or high generation of spikes.

MSC 2010 subject classifications: Primary 60J70; secondary 62M05, 60H10, 60J60, 62P10, 35Q62, 37A50.
Keywords and phrases: Hypoelliptic diffusion, FitzHugh-Nagumo model, invariant density, non-parametric estimation, up-crossings, pulse rate, spike rate estimation.

1. Introduction

Neurons are excitable cells that are linked thanks to synapses into a huge network. If the electric membrane potential, the voltage, of a neuron is sufficiently high, the neuron is able to produce an action potential, also called a spike, which is a stereotype all-or-non fast and large electric signal. Spikes allow the neuron to activate its synaptic contacts and to modulate their voltage. Spikes can be viewed as the basic element of information traveling from one neuron to another in the network. It is therefore of tremendous importance to understand and describe the individual voltage and the generation of spikes.

Neuronal spiking (also called firing) is a complex process that involves interactions between numerous cells. Modeling this mechanism mathematically is therefore difficult. Several neuronal models have been developed, the most

[*]Escuela de Matemática, Facultad de Ciencias, Universidad Central de Venezuela, Caracas
[†]Université Grenoble Alpes, Laboratoire Jean Kuntzmann, UMR CNRS 5224
[‡]INRIA project/team AIRSEA

famous is the 4 equations dynamical system of [12] that accurately describes the electrical mechanism of a single neuron. The model has an oscillatory behavior to reproduce the alternance of spiking phases and non spiking phases. To produce such behavior, the differential equations driving the potential (denoted X_t in the following) is coupled to differential equations related to the fraction of open ion channels of different kinds (conductances). However this chaotic system is difficult to study from a mathematical point of view. Several relaxed models have then been proposed, most of them reducing the dimension of the system. We can cite the Morris-Lecar model that simplifies the three channel equations of the Hodgkin-Huxley model into only one non-linear equation modeling the membrane conductance evolution [19]. Another model is the FitzHugh-Nagumo (FHN) model, which has a polynomial drift. FitzHugh-Nagumo and Morris-Lecar models share the properties of sub-threshold and supra-threshold response, that is they intrinsically model the regenerative firing (spiking) mechanism in an excitable neuron. FitzHugh-Nagumo model is less plausible than conductance-based models: it has been built as an oscillatory system, not from physical assumptions. It has however the advantage of being more directly amenable to a mathematical analysis than Hodgkin-Huxley or Morris-Lecar thanks to its polynomial drift. Note that another class of neuronal models is the class of Leaky-Integrate-Fire (LIF) models [see 10, for a review], where voltage is modeled by a one-dimensional process. The main drawback is that spikes are not generated automatically and a (fixed) threshold has to be introduced, which is unrealistic.

Stochastic versions of neuronal models have been proposed to describe various sources of randomness [10, 17, 2] . Stochastic noise can be introduced in the first equation mimicking noisy presynaptic currents [see among others 23, 18, 10]. In the second class of stochastic models, noise affects the other differential equations, describing the randomness of the conductance dynamics, like random opening and closing of ion channels [16, 17]. This second class of models can be viewed as a diffusion approximation of ion channels modeled by point processes [see among others 20].

In this paper, we focus on the second class of stochastic FitzHugh-Nagumo (FHN) model. It is defined as follows. Let X_t denote the membrane potential of the neuron at time t and C_t a recovery variable that models the channel kinetics. We assume that $((X_t, C_t),\ t \geq 0)$ is governed by the following Itô stochastic differential equation (SDE):

$$\left\{ \begin{array}{rcl} dX_t &=& \frac{1}{\varepsilon}(X_t - X_t^3 - C_t - s)dt, \\ dC_t &=& (\gamma X_t - C_t + \beta)\, dt + \tilde{\sigma} dW_t, \end{array} \right. \tag{1.1}$$

where W_t is a standard Brownian motion, ε is the time scale separation usually very small (X has a much faster time scale than C), s is the magnitude of the stimulus current, $\tilde{\sigma}$ the diffusion coefficient, β, γ are positive constants that determine the position of the fixed point and the duration of an excitation. FHN has already been studied extensively in physical papers [see among others 17, 2]. Our objective is to revisite theoretically some of these results and to propose

non-parametric estimators.

The three objectives of the paper are the followings. In Section 2, we study some probabilistic properties of the FitzHugh-Nagumo model: hypoellipticity, Feller, invariant probability, mixing property. To prove these results, we take advantage of the fact that the stochastic FitzHugh-Nagumo model with noise on the second equation is a generalization of van der Pol equations and belongs to the class of stochastic Damping Hamiltonian systems. One of the main reference is [24] for an overview of the theoretical properties of these models, [see also 13, 5].

In Section 3, we consider the questions related to the neuronal modeling. As said previously, spikes are the essential element of information exchange in the neural network. It is thus very important to understand their distribution. The distribution of spikes is difficult to study from scratch. Attempts have been proposed using point processes, but describe only the spike trains and not the neuronal voltage [21, 22]. When using voltage data, the first difficulty is the definition of the spike itself. In the pioneer work of [16], a spike is described as a "long" excursion on the phase space before returning back into the neighborhood of the fixed point. The pulse rate is measured by time averaging the number of pulses during time interval $[0, T]$ and the mean time between two pulses can also be estimated. They show that the pulse rate is the inverse of the mean length. But to our best knowledge, this has not be proven theoretically for the hypoelliptic FHN. Attempts have been based on Gaussian approximation of the voltage process [23, 8] leading to a Gaussian stochastic modeling. This Gaussian approximation does however not fit with real data. In this paper, we propose to study spike generation through the modeling of the voltage by the FitzHugh-Nagumo model, avoiding any normal approximation or a vague definition of a spike. The idea is the following. A spike occurs when X_t crosses a certain threshold, the spike shape being then almost deterministic. Note that it is known from voltage data that the threshold is not fixed: the voltage X_t has not always the same value when entering the spiking phases. We thus focus on the distribution of the process of up-crossings of X_t at a large level u. If a spike occurs, the distribution of up-crossings should remain the same for any level value u in a given interval. Finally the distribution of the length of the interval between two successive spikes is studied and we prove that its expectation is the inverse of the up-crossing rate. Note that [3] also study the generation of spikes for a bi-dimensional FitzHugh-Nagumo model. Their model, although based on the same deterministic system as ours, has noise in both components. This fact implies that the solution (X_t, C_t) is a classical diffusion (elliptic). Hence each coordinate is a continuous but non-differentiable function. The number of crossings of such a function is infinity in every bounded interval (allowing for instance the existence of local time). This prevents defining the spikes via the up-crossings. The authors define the spikes as large excursions in the space of phases and study their distributions. Their method is therefore very different than ours.

In Section 4, we propose a non-parametric estimation of the stationary density

based on a kernel estimator. The 2-dimensional bandwidth is selected automatically from the data with a Goldenshluger and Lepski's approach. We deduce an estimator of the up-crossing rate. These estimators are illustrated on simulations in different excitation regimes of the neuron.

2. Properties of the FitzHugh-Nagumo model

Model (1.1) follows a non-linear drift with singular diffusion coefficient (no noise on the first coordinate). It is not easy to study directly this kind of models because standard probabilistic tools assume a non degenerate diffusion coefficient. To take advantage of probabilistic tools that have already been developed for some hypoelliptic systems, we introduce a change of variable of the second coordinate. This allows us to enter the class of stochastic Damping Hamiltonian systems, that have been widely studied [see among others 24, 13, 5]. We can then prove some useful properties of the FitzHugh-Nagumo model (hypoellipticity, Feller, existence of a stationary distribution, β-mixing). Let us first introduce the change of variable.

2.1. A stochastic Damping Hamiltonian system

The change of variable is the following. Let $Y_t = \frac{1}{\varepsilon}(X_t - X_t^3 - C_t - s)$. Applying Itô's formula, the FitzHugh-Nagumo system (1.1) can be rewritten:

$$\left\{ \begin{array}{lcl} dX_t & = & Y_t dt, \\ dY_t & = & \frac{1}{\varepsilon}\left(Y_t(1-\varepsilon-3X_t^2) - X_t(\gamma-1) - X_t^3 - (s+\beta)\right) dt - \frac{\tilde{\sigma}}{\varepsilon} dW_t, \end{array} \right. \tag{2.1}$$

Thanks to this transformation, we can notice that system (2.1) is a stochastic Damping Hamiltonian system. These systems have been introduced to describe the dynamics of a particle with X_t referring to its position and Y_t to its velocity. The movement of the particle is guided by a potential $V(x)$ and by a damping force $c(x)$:

$$\left\{ \begin{array}{lcl} dX_t & = & Y_t dt, \\ dY_t & = & -(c(X_t)\, Y_t + \partial_x V(X_t))dt + \sigma\, dW_t \end{array} \right. \tag{2.2}$$

Its infinitesimal generator $\mathcal{L}$ is

$$\mathcal{L} = \frac{\sigma^2}{2}\,\partial_{yy} + y\,\partial_x - (c(x)\,y + \partial_x V(x))\partial_y\,.$$

These models have been studied by [24] under some conditions for $V(\cdot)$, $c(\cdot)$ and the diffusion coefficient, that we recall here:

(H1) The potential $V(x)$ is lower bounded, smooth over $\mathbb{R}$, V and ∇V have polynomial growth at infinity.
(H2) The damping coefficient $c(x)$ is continuous, for all $N > 0$, $\sup_{|x|\leq N} |c(x)| < +\infty$ and for all $x \in \mathbb{R}$, $c(x) \geq c$.
(H3) There exists $\sigma_1 > 0$ such that $0 < \sigma < \sigma_1$.

In our case, the damping force is $c(x) = \frac{1}{\varepsilon}(3x^2 - 1 + \varepsilon)$, the potential is $V(x) = \frac{1}{\varepsilon}(\frac{x^4}{4} + \frac{\gamma-1}{2}x^2 + (s+\beta)x)$ and the diffusion coefficient is $\sigma = \frac{\tilde{\sigma}}{\varepsilon} > 0$. One can prove easily that conditions (H1), (H2) and (H3) are fulfilled under weak assumptions. Indeed

- The potential $V(x) = \frac{1}{\varepsilon}(\frac{x^4}{4} + \frac{\gamma-1}{2}x^2 + (s+\beta)x)$ is continuous, goes to ∞ when $x \to \pm\infty$ and is thus lower bounded, smooth over $\mathbb{R}$, V and $\nabla V = \frac{1}{\varepsilon}(x^3 + (\gamma-1)x + (s+\beta))$ have polynomial growth at infinity. This implies (H1).
- The damping coefficient $c(x) = \frac{1}{\varepsilon}(3x^2 - 1 + \varepsilon)$ is continuous, upper bounded on sets $\{x| \leq N\}$ and for all $x \in \mathbb{R}$, $c := 1 - \frac{1}{\varepsilon}$ implies (H2).
- The diffusion coefficient is $\sigma = \frac{\tilde{\sigma}}{\varepsilon}$. We assume that the (unknown) parameters ε and σ are such that $\varepsilon \in [\varepsilon_0, \varepsilon_1]$ and $\tilde{\sigma} \in [\tilde{\sigma}_0, \tilde{\sigma}_1]$. This implies (H3) with $\sigma_1 := \frac{\tilde{\sigma}_1}{\varepsilon_0}$.

We can also notice that

$$\frac{x.\nabla V(x)}{|x|} \to +\infty\,, \quad \text{as } |x| \to +\infty.$$

This is condition (0.5) of [24]. It can be interpreted as follows: the force $-\nabla V(x)$ is "strong enough" for $|x|$ large to ensure a quick return of the system to compact subsets of $\mathbb{R}^2$.

2.2. Hypoellipticity and β-mixing

In this section, we prove some theoretical properties for process $(Z_t) = (X_t, Y_t)$: (Z_t) is strong Feller, hypoelliptic, the existence of a unique invariant probability and a β-mixing property. The main reference used in all the proofs is [24].

Hypoellipticity and stationary distribution We first focus on the hypoellipticity of (Z_t).

Proposition 2.1. ***(Hypoellipticity)*** *Let $Z_t = (X_t, Y_t)$ be the solution of system (2.1). The stochastic process (Z_t) is hypoelliptic and strong Feller.*

Hypoellipticity can be interpreted as the fact that the one-dimensional noise entering the second coordinate propagates to the two-dimensional space. It ensures that the distribution $P_t(z, \cdot)$ of the process Z_t starting from $Z_0 = z$ has a smooth density, denoted $p_t(z, \cdot)$ in the following.

Proof. of Proposition 2.1. We start by proving the hypoellipticity. Let us denote A_0, A_1 the differential operators

$$\begin{aligned} A_0 &= y\,\partial_x - (\nabla V(x) + c(x)y)\,\partial_y \\ A_1 &= \sigma\partial_y \end{aligned}$$

Let $[A, B]$ denote the Lie bracket between operators A and B. We have

$$[A_0, A_1] = -\sigma\, \partial_x + \sigma c(x)\, \partial_y$$

Thus

$$\mathrm{Span}(A_0, A_1, [A_0, A_1]) = \mathrm{Span}(\partial_x, \partial_y).$$

This implies that the system is hypoelliptic.

Second, we prove that the stochastic process (Z_t) is strong Feller. Let us denote $P_t f(z) = \mathbb{E}_z(f(Z_t)) = \int f(u) p_t(z, u) du$ where $p_t(z, \cdot)$ is the transition density of the system. We want to prove that if f is $L^\infty(\mathbb{R}^2)$ then $P_t f(x)$ is continuous. The coefficients of the infinitesimal generator $\mathcal{L}$ are $\mathcal{C}^\infty$. By Hrmander's theorem, this implies that $p_t(z, u)$ is $\mathcal{C}^\infty$. Thus as f is bounded, $\int f(u) p_t(z, u) du = P_t f(z)$ is continuous. So finally (Z_t) is strong Feller. □ □

We now prove the existence and uniqueness of an invariant probability.

Proposition 2.2. ***(Stationary distribution)*** *Let (X_t, Y_t) be solution of system (2.1).*

1. *The process (X_t, Y_t) is positive recurrent with a unique invariant probability measure μ.*
2. *Moments of any order of μ exist: for all $k_1, k_2 \in \mathbb{N}$,*

$$\mathbb{E}(X_t^{k_1} Y_t^{k_2}) = \int x^{k_1} y^{k_2} d\mu(x, y) < +\infty$$

Proof is given in Appendix.

Proposition 2.2 ensures that the invariant measure has a smooth density $\mu(dz) = p(z)dz$ [4]. In the following, we will denote p^x and p^y the marginal of p with respect to x and y.

Mixing We now study the mixing property of the stationary distribution. We first recall the definition.

Definition 2.1. *Let $\{Z_t\}_{t \in \mathbb{R}^+}$ a stationary stochastic process. Introducing the σ-algebra $\mathcal{F}_t = \sigma(\{Z_s : s \leq t\})$ and $\mathcal{F}^t = \sigma\{Z_s : s \geq t\}$. We say that Z is β-mixing, with mixing coefficient β_t, if*

$$\beta_t = \mathbb{E}[\sup\{|\mathbb{P}(U|\mathcal{F}_u) - \mathbb{P}(U)| : U \in \mathcal{F}^{u+t}\}]$$

and $\beta_t \to 0$ when $t \to \infty$.

We can prove that the process Z is β-mixing.

Proposition 2.3. ***(Mixing)*** *Let (X_t, Y_t) be solution of system (2.1).*

1. *There exist constants $D > 0$ and $0 < \rho < 1$ such that for all z,*

$$\left| P_t f(z) - \int f d\mu \right| \leq D \sup_a \left(\frac{\left| f(a) - \int f d\mu \right|}{\Psi(a)} \right) \Psi(z)\, \rho^t \,. \tag{2.3}$$

   where Ψ is a Lyapounov function defined by (A.2).
2. *The skeleton chain $\tilde{Z}_k = Z_{kh}$ for $h > 0$ is exponentially β-mixing with β-mixing coefficient β_{kh} such that*

$$\beta_{kh} \leq D'' ||\Psi||_1 \rho^{kh}$$

Note that the large deviation principle also applies for the occupation empirical measure $L_t = \frac{1}{t} \int_0^t \delta_{Z_s} ds$, this is a direct consequence of Theorem 3.1 of [24].

Proof of Proposition 2.3. Let us start with the proof of 1. Theorem 2.4 of [24] shows the existence of constants $D > 0$ and $0 < \rho < 1$ such that for all z, inequality (2.3) holds. Then, since Ψ is μ integrable and larger than 1, inequality (2.3) implies that the Markov chain $(Z_i)_{i \in \mathbb{N}}$, $Z_0 \sim p_s(z)dz$ is exponentially β-mixing.

Now we prove 2. First, we remark that the Liapunov function Ψ is integrable with respect to the invariant measure [see 24]. Property 1 of Proposition 2.3 implies that

$$||P_t(f) - \int f d\mu||_1 \leq \rho^t\, D \sup_a \left(\frac{\left| f(a) - \int f d\mu \right|}{\Psi(a)} \right) ||\Psi||_1,$$

where $||\cdot||_1$ denotes the $\mathbb{L}^1$ norm with respect to the invariant measure. We can deduce the following inequality in norm of total variation

$$||P_t(z, \cdot) - \mu||_{TV} \leq D'' \Psi(z) \rho^t.$$

One can apply this inequality to the skeleton chain $\tilde{Z}_k = Z_{kh}$ for a certain h. Let us denote $\tilde{P}_k$ the discrete semi-group associated to $\tilde{Z}$. Then we get

$$||\tilde{P}_k(z, \cdot) - \mu||_{TV} \leq D'' \Psi(z) \rho^{kh}.$$

We can deduce from [9] Chapter 2, Section 2.4, that the β-mixing coefficient β_{kh} is equal to

$$\beta_{kh} = \int ||\tilde{P}_k(z, \cdot) - \mu||_{VT}\, d\mu(z) \leq D'' ||\Psi||_1 \rho^{kh}.$$

So we have $\beta_{kh} \leq D'' ||\Psi||_1 \rho^{kh}$. Hence the skeleton chain is exponentially β-mixing. □

3. Number of up-crossings and spike rate

When a spike is emitted, it has a deterministic shape. The emission of a spike occurs when the voltage crosses a certain level, this level being not fixed. This means that when the voltage is high enough, there is no way back and a spike occurs with probability 1. We thus focus on describing the spike generation process.

3.1. Spikes and previous results

Spiking regime In the literature, spikes of FitzHugh-Nagumo have been defined as a long excursion in the phase space. One of the main references is [16]. An example of a phase space is given in Figure 1 for three different sets of parameters: left: $\varepsilon = 0.1$, middle: $\varepsilon = 0.4$, right: $\varepsilon = 0.5$. Let us comment first the left plots ($\varepsilon = 0.1$) that generates automatically spikes. As explained in [16], the fixed point is on the left bottom of the phase space. It corresponds to the dynamic of the potential between two pulses. Then the trajectory reaches the right branch, that belongs to the excited state: X increases while C remains almost constant. Then it moves along this branch upwards until it reaches its top, with C that increases. Then it switches to the left branch, that belongs to a refractory phase of the neuron: X decreases and C stays high. Finally, the trajectory relaxes into the fixed point with X back to the resting potential and C which decreases. When such a long excursion occurs, a "spike" or a "pulse" is observed in the voltage variable.

Now, let us comment the right plots of Figure 1 with $\varepsilon = 0.5$. In that case, the potential stays in the vicinity of the fixed point and no excursion in the excited state occurs. Finally, let us comment the middle plots of Figure 1 with $\varepsilon = 0.4$. We can observe excursions on the right branch, as when $\varepsilon = 0.1$, but these excursions are less large. It is less clear if one should consider these excursions as spikes or pulses. The definition of a spike is therefore not clear.

Spike rate Nevertheless, given this definition of a spike, spike rate has been studied. We recall some results provided by [16]. Let N_t denote the number of pulses during time interval $[0, t]$. The spike rate is defined as

$$\rho := \lim_{t\to\infty} \frac{N_t}{t}. \tag{3.1}$$

The process N_t is random and the limit above is to be understood almost surely. It is the expected number of spikes by unit of time. The spike rate can also be defined as follows. Let us denote T_i the ith interspike interval, i.e. the time between the i and $i+1$ spikes (or pulses, or long excursions). The mean time $< T >$ between two spikes, i.e. the mean length of interspike intervals, can be defined using the ergodic theorem as

$$< T >= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} T_i. \tag{3.2}$$

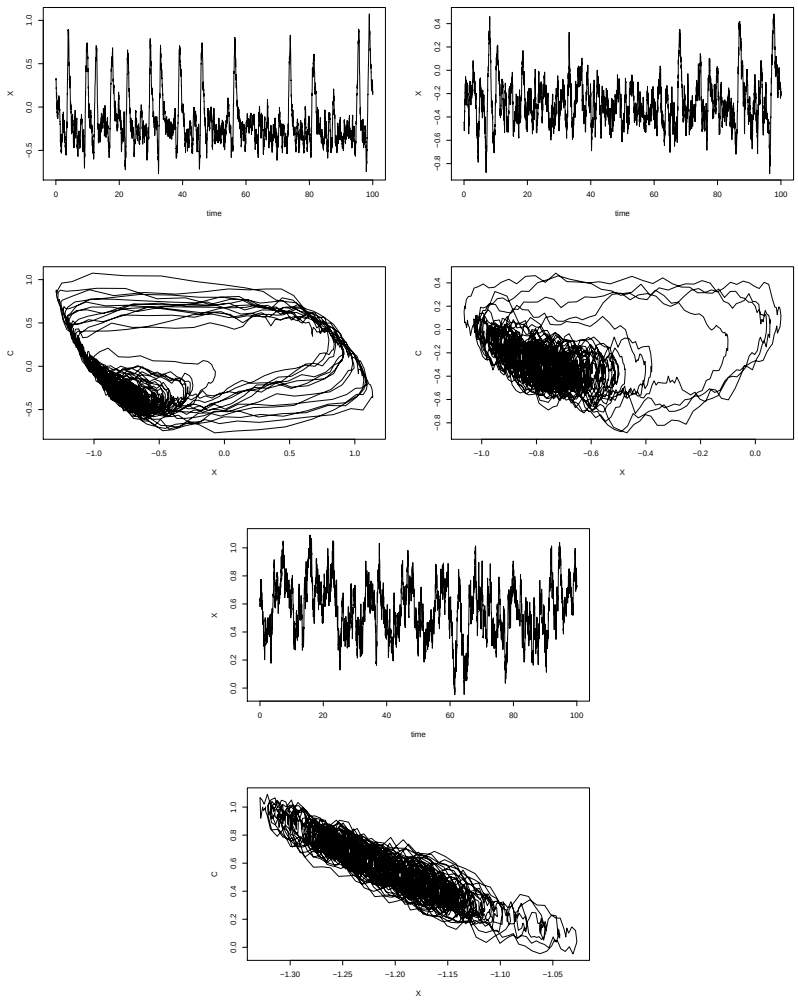

FIG 1. *Simulations of the FitzHugh-Nagumo. Voltage variable X_t versus time (top line) and the corresponding trajectory in phase space C_t versus X_t (bottom line) for $s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, left: $\varepsilon = 0.1$ and $\gamma = 1.5$, middle: $\varepsilon = 0.4$ and $\gamma = 1.5$, right: $\varepsilon = 0.5$ and $\gamma = 0.2$.*

[16] state the following relationship between these two quantities:

$$\rho = \lim_{N\to\infty} \left(\frac{T_0}{N} + \frac{1}{N}\sum_{i=1}^{N} T_i + \frac{T_{N+1}}{N} \right)^{-1} = \frac{1}{< T >}, \tag{3.3}$$

where T_0 and T_{N+1} are the time intervals until the first or after the last spike. The spiking times (T_i) are also random and the above limit is in the almost sure sense. Relation (3.3) means that the limit in t is equal to the limit in N of the random processes $(N_t, t \geq 0)$ and $(T_i, i \geq 1)$. This relation is intuitive as explained by [16]. It is true for a Poisson process. However, it is not straightforward for any point process. Moreover, the two processes $(N_t, t \geq 0)$ and $(T_i, i \geq 1)$ are difficult to define from the stochastic process (X_t, C_t) or (X_t, Y_t), and consequently to prove (3.3).

The objective of the next two subsections is to give a formal mathematical framework to formula (3.1), (3.2) and (3.3). The starting point is a precise definition of the objects of study: spikes are difficult to define from (X_t, C_t) or (X_t, Y_t) (should one take the beginning of the spike ? the maximum ?).

An alternative is to study the number of spikes occurence through the number of up-crossing of process X_t at a certain (large) level u. This has the advantage of defining the occurence of a spike through a precisely defined random variable (see below). Hence, theoretical results can be derived. This is the methodology we consider in this paper.

3.2. Number of up-crossings

Process $Z_t = (X_t, Y_t)$ defines a mesure $\mathbb{P}^z$ in the space $\Omega := C(\mathbb{R}^+, \mathbb{R}^2)$. This means that X_t is an a.s. continuously differentiable process and $\dot{X}_t = Y_t$. Let us define the number of up-crossings of process $X_\cdot$ at level u in $[0, t]$ as

$$U_t^{X_\cdot}(u) = \#\{s \leq t : \, X_s = u, \, Y_s > 0\}.$$

Note that the natural model to study the up-crossing is model (2.1) rather than model (2.2) because it only requires the second coordinate to be positive.

Heuristic of up-crossing process Let us first give an intuition why $U_t^{X_\cdot}(u)$ is a process linked to the generation of spikes. If we forget the boundary effects, the random variable N_t will be equal to the number of up-crossings $U_t^{X_\cdot}(u)$ at level u, for a set of large values u. Indeed, when a spike occurs, we expect $U_t^{X_\cdot}(u)$ to be constant for all values u that correspond to the right branch of the phase space (Figure 1), ie to the (deterministic) increase phase of the potential. Note that for too large values of u (larger than the maximum of the spikes), $U_t^{X_\cdot}(u)$ is expected to be zero. We therefore expect the distribution of $U_t^{X_\cdot}$ to be stable for an interval of values of u and then to decrease abruptly to 0. On the contrary, if a neuron is not in a spiking regime (when ε is large for example, see middle plots

of Figure 1), (small) excursions do not correspond to spikes and up-crossings will vary with u. These two very different behaviors of $U_t^{X_\cdot}$ imply that knowing how process $U_t^{X_\cdot}(u)$ varies with u gives automatically a definition of a spike and the fact that the neuron is in a spiking regime or not. Process $U_t^{X_\cdot}(u)$ can therefore be seen as a properly defined stochastic process that describes generation of spikes. Note that a spike occurs the first time process (X_t, Y_t) hits the plan $\{X_s = u,\, Y_s > 0\}$ of $\mathbb{R}^2$.

It is therefore important to study the number of up-crossings. In this section, we first prove the Rice's formula that links the expectation of $U_t^{X_\cdot}(u)$ with the stationary density of process $Z_t = (X_t, Y_t)$. Then, we prove an ergodic theorem for $U_t^{X_\cdot}(u)$: the expected value of $U_t^{X_\cdot}(u)$ by unit of time converges to an integral with respect to the stationary density. This limit, that depends on u, will be used in section 3.3 to estimate the spike rate.

Rice's formula Let us denote B_Ψ the space of measurable functions

$$B_\Psi = \{f : \mathbb{R}^2 \to \mathbb{R} : \sup_{(x,y)} \frac{|f(x,y)|}{\Psi(x,y)} < \infty\}.$$

We consider the norm $||f||_\Psi = \sup_{(x,y)} \frac{|f(x,y)|}{\Psi(x,y)}$. We first prove the Rice's formula on the expectation of the number of up-crossings.

Proposition 3.1. ***(Rice's formula)*** *Let $Z_t = (X_t, Y_t)$ be the stationary solution of the FitzHugh-Nagumo system. The Rice's formula holds true*

$$\mathbb{E}U_t^{X_\cdot}(u) = t \int_0^\infty y p(u,y) dy. \tag{3.4}$$

Proof. of Proposition 3.1. The first step is the proof of a technical lemma.

Lemma 3.1. *Let us define the function $G_{y_1}(x,y) := |y - y_1|$. Set $z = (x,y)$. The semigroup $\{P_t\}_{t\geq 0}$ satisfies*

$$P_t G_{y_1}(z) \to G_{y_1}(x,y), \quad \textit{if} \quad t \to 0$$

uniformly in x.

The proof of Lemman 3.1 is given in Appendix. Then, Rice's formula is proved using the following result from [1]'s book, that we recall now.
Let $\{X_t\}_{t\in\mathbb{R}^+}$ be a stochastic process and Y its derivative. Let us first introduce the Rice's formula. For any $u \in \mathbb{R}$ and for I an interval of time, where the crossings are counting, the Rice's formula is

$$E(U_t^X(u)) = \int_I dt \int_0^\infty y p_{X_t,Y_t}(u,y) dy,$$

It holds true under the following conditions:

(C1) Function $(t,x) \to p_{X_t}(x)$ is continuous for $t \in I$, x in a neighborhood of u.
(C2) $(t,x,y) \to p_{X_t,Y_t}(x,y)$ is continuous for $t \in I$ and x in a neighborhood of u and $y \in \mathbb{R}$.
(C3) Let define $A_1(t_1,t_2,x) = \int_{\mathbb{R}^2} |y_1 - y| p_{X_{t_1},Y_{t_1},Y_{t_2}}(x,y,y_1)dydy_1$. The process satisfies condition A41 in page 76 of [1]'s book if $A_1(t_1,t_2,x)$ tends to zero as $t_2 - t_1 \to 0$ uniformly, for t_1 and t_2 in a compact and x in a neighborhood of u.

We must now verify if conditions (C1), (C2), (C3) hold. Under the stationary regime, the two marginal densities, $p^x(x)$ and $p^y(y)$, of the invariant density $p(x,y)$ correspond to the density of X_s and the density of Y_s respectively. They are both C^∞ functions. By stationarity, the density of (X_s, Y_s) is $p(x,y)$. Then (C1)-(C2) hold true.

Consider the vector (X_0, Y_0, X_s, Y_s). It has a density $p_s(x,y,x_1,y_1)p(x,y)$. To prove (C3), let us observe that for any $s > 0$, thanks to the stationary regime,

$$A_1(s,x) := A_1(0,s,x) = \int_{\mathbb{R}^2} |y_1 - y| \left(\int_{\mathbb{R}} p_s(x,y,x_1,y_1)p(x,y)dx_1 \right) dydy_1.$$

Then we can write

$$A_1(s,x) := \int_{\mathbb{R}} \left(P_s G_y\right)(x,y)p(x,y)dy.$$

Lemma 3.1 implies that $P_t(G_{\tilde{y}})(x,y) \to G_{\tilde{y}}(x,y)$ uniformly in x, where (x,y) is the point of depart of our process. In the particular case $\tilde{y} = y$ we get $P_t(G_y)(x,y) \to G_y(x,y) = 0$. This yields, by using the Bounded Convergence Theorem, that $A_1(s,x) \to 0$ uniformly in x and the Rice's formula for the first moment of the number of up-crossings holds true. □

Ergodic theorem We now prove that the Ergodic theorem can be applied.

Theorem 3.1. *Let $Z_t = (X_t, Y_t)$ be the stationary solution of the FitzHugh-Nagumo system. For any $u \in \mathbb{R}$:*

$$\frac{U_t^{X_\cdot}(u)}{t} \to \int_0^\infty yp(u,y)dy \ \ a.s. \tag{3.5}$$

This result gives the limit of the expected number of up-crossings by unit of time. The fact that it is an integral with respect to the invariant density allows us to estimate this quantity (see Section 4). This limit will also be used in the estimation of the mean length between two spikes, as explained below.

Proof. of Theorem 3.1. We follow [7]'s book section 11.5. Proposition 2.2 states that process (X_t, Y_t) is exponentially ergodic. Let us define the two processes

of the number of up-crossings in interval $(s, s+1]$ and in interval $(s-1, s]$ respectively:

$$\begin{aligned}\zeta_1(s) &= U^{X_\cdot}_{s+1}(u) - U^{X_\cdot}_{s}(u) := U^{X_\cdot}_{(s,s+1]}(u),\ s \geq 0\\ \zeta_2(s) &= U^{X_\cdot}_{s}(u) - U^{X_\cdot}_{s-1}(u) := U^{X_\cdot}_{(s-1,s]}(u),\ s \geq 1.\end{aligned}$$

These two processes are stationary. Let $\theta_\cdot$ denote the shift operator for the stationary Markov process (X_t, Y_t). It holds $\zeta_1(s) = \theta_s \circ \zeta_1(0)$ and $\zeta_2(s) = \theta_s \circ \zeta_2(1)$. The Ergodic Theorem and Rice's formula (3.4) assure that

$$\frac{1}{t}\int_0^t \zeta_1(s)ds \to \int_0^\infty yp(u,y)dy \text{ a.s.} \tag{3.6}$$

and similarly for ζ_2. Let us prove the following chain of inequalities from [7] (Cramer & Leadbetter (1967) pag. 238).

$$\int_0^{t-1} \zeta_1(s)ds \leq U^{X_\cdot}_t(u) \leq \int_0^{t+1} \zeta_2(s)ds \tag{3.7}$$

holds. Let us show the left inequality

$$\begin{aligned}\int_0^{t-1} U^{X_\cdot}_{(s,s+1]}(u)ds &= \int_0^{t-1} \big(U^{X_\cdot}_{s+1}(u) - U^{X_\cdot}_{s}(u)\big)ds\\ &= \int_1^t U^{X_\cdot}_s(u)ds - \int_0^{t-1} U^{X_\cdot}_s(u)ds\\ &= \int_{t-1}^t U^{X_\cdot}_s(u)ds - \int_0^1 U^{X_\cdot}_s(u)ds \leq U^{X_\cdot}_t(u),\end{aligned}$$

where we have used that function $U^{X_\cdot}_t(u)$ is nondecreasing. The right inequality can be proved similarly. Gathering (3.6) and (3.7) implies the theorem. □

3.3. Spike rate

Now we want to link the spike rate ρ with the up-crossing process. We start by the intuition and some heuristic and then, we formalize this link.

Heuristic As explained earlier, the up-crossing process gives a definition of a spike: if the up-crossing process is constant for any level u in a (large) interval, a spike occurs. Then, for a given level u in this interval, the number of up-crossing $U^{X_\cdot}_t(u)$ is equal to the number of spikes. This naturally gives an approximation of the variable N_t introduced by [16]. Let us define $\lambda(u)$ the limit of the number of up-crossings at level u:

$$\lambda(u) = \lim_{t\to\infty} \frac{U^{X_\cdot}_t(u)}{t}.$$

The ergodic theorem (Theorem 3.1) gives an explicit formula of this limit:

$$\lambda(u) = \int_0^\infty y\, p(u,y)\, dy. \tag{3.8}$$

Finally, we can naturally link the spike rate obtained via the random variable N_t (excursions rate) and the rate of up-crossings:

$$\lambda(u) \approx \rho. \tag{3.9}$$

Formal link This heuristic reasoning may be led by a more formal way. We want to describe formally the time between successive up-crossings of level u. Time between successive up-crossings has been studied by [7] (Chapter 11).

The idea is the following. Let $C^{X.}_{[0,t]}(u)$ denote the number of all crossings at level u on the interval $[0,t]$. Let us define for all $k \in \mathbb{N}^*$ the set:

$$H_k(\tau,t) = \mathbb{P}\{U^{X.}_{(-\tau,0)}(u) \geq 1,\, C^{X.}_{(0,t)}(u) \leq k\}.$$

The function of interest for us is when $k = 1$. Assume an up-crossing occurs at time $\tau = 0$. Then the trajectory is over level u just after τ. If $C^{X.}_{(0,t)}(u) = 1$, the crossing is a down-crossing and no up-crossing occurs in interval $[0,t]$. Thus $H_1(\tau,t)$ is the probability to have a up-crossing in the interval $(-\tau,0)$ and no up-crossing in interval $[0,t]$.

To study the interval between two spikes or up-crossings, we are interested in defining a conditional probability. For that purpose, let us introduce the following probability

$$\omega(t) := \mathbb{P}\{\text{there exists at least one upcrossing in time}\, t\}.$$

We know by using the ergodic theorem that

$$U^{X.}_{[0,t]}(u) = t\lambda(u) + o(t) \;\text{ when } t \to \infty. \tag{3.10}$$

As process X is continuous and differentiable, the stream of upcrossings, i.e. the times when the process crosses level u with positive derivative, is stationary and regular. This implies that $\mathbb{P}\{U^{X.}_{[0,t]}(u) > 1\} = o(t)$. From [7] (page 54), we can prove that

$$\omega(t) = t\ell(u) + o(t) \;\text{ when } t \to 0,$$

for a certain constant $\ell(u)$. The property of a regular up-crossing process implies that $\ell(u) = \lambda(u)$. This important fact links the ergodic limit (3.10) with the behavior of $\omega(t)$ in a neighborhood of zero. This result will be very useful for us.

Then, [7] prove that for all k, there exists a finite function Φ_k defined by the following limit:

$$\Phi_k(t) := \lim_{\tau \to 0} \frac{H_k(\tau,t)}{\omega(\tau)}. \tag{3.11}$$

This last function represents *the conditional probability of no more than k crossings in the interval $(0,t)$, given that an up-crossing occurred "at" time zero.* For $k=1$, $\Phi_1(t)$ is the conditional probability that no up-crossing occurs in interval $[0,t]$, given that an up-crossing occurred "at" time zero.

It is then natural to introduce the following function:

$$F_2(t) := 1 - \Phi_1(t). \tag{3.12}$$

The first remarkable fact is that $F_2(\cdot)$ is a real distribution function.

Proposition 3.2. *The function F_2 defined by (3.12) is a distribution function.*

Moreover, it may be regarded as the *distribution function of the length of the interval of an arbitrarily chosen up-crossing and the next up-crossing'*, following [7].

Proof. of Proposition 3.2. It holds easily that $0 \le F_2(t) \le 1$.

Then to prove that it is a distribution function, set $u_0(t) = \mathbb{P}\{U^{X.}_{(0,t)} = 0\}$. We have

$$\begin{aligned} u_0(t) - u_0(t+\tau) &= \mathbb{P}\{U^{X.}_{(0,t)}(u) = 0\} - \mathbb{P}\{U^{X.}_{(-\tau,t)}(u) = 0\} \\ &= \mathbb{P}\{U^{X.}_{(-\tau,t)}(u) \ge 1,\, U^{X.}_{(0,t)}(u) = 0\} \\ &= H_1(\tau,t) + o(\tau). \end{aligned}$$

where we use [7] (page 225) for the last equality: "the probability of more than one crossing in $(-\tau,0)$ is $o(\tau)$, whereas if the only crossing in $(-\tau,0)$ is an up-crossing, then $U^{X.}(0,t) = 0$ when (and only when) $C^{X.}_{(0,t)}(u) = 0$".
Thus

$$\lim_{\tau^+\to 0} \frac{u_0(t+\tau) - u_0(t)}{\tau} = -\lim_{\tau\to 0} \frac{H_1(\tau,t)}{\omega(\tau)} \frac{\omega(\tau)}{\tau} = -\lambda(u)\Phi_1(t). \tag{3.13}$$

Thus the function u_0 has right-hand side derivative: $D^+u_0(t) = -\lambda(u)\Phi_1(t)$. Moreover the Lebesgue theorem gives

$$u_0(T) - u_0(0) = \int_0^T D^+u_0(t)dt.$$

As u_0 is bounded, the derivative $D^+u_0(t)$ is integrable over $(0,\infty)$. This in particular implies that $D^+u_0(t) \to 0$ whenever $t \to \infty$, thus $\lim_{t\to\infty} \Phi_1(t) = 0$ and therefore $F_2(t) \to 1$ when $t \to \infty$. Finally as F_2 is non-decreasing it is a real distribution function. □

As F_2 can be interpreted as the distribution function of the length of an inter-up-crossings interval (interval between two successive up-crossings), we are interested in computing its first two moments. Always following [7], page 227, we get

Proposition 3.3. *The expectation of distribution F_2, that is the mean length of the interval between two successive up-crossings, is given by:*

$$\int_0^\infty t dF_2(t) = \frac{1}{\lambda(u)}.$$

The second moment of F_2 is:

$$\int_0^\infty t^2 dF_2(t) = \frac{2}{\lambda(u)} \int_0^\infty u_0(t) dt, \tag{3.14}$$

where $u_0(t) = \mathbb{P}\{U^{X.}_{(0,t)} = 0\}$.

The expectation of the length of an inter-up-crossings interval is equal to the inverse of the rate of up-crossings at level u (formula (4.6)) and thus to the spike rate. This the principal advantage of our approach. We give a theoretical justification to the link between length of the excursions and the spike rate.

Proof. of Proposition 3.3. Start with the mean of F_2. Given formula (3.13), one easily obtain:

$$\int_0^\infty t dF_2(t) = \int_0^\infty [1 - F_2(t)] dt = \frac{1}{\lambda(u)} [u_0(0) - u_0(\infty)].$$

Then we use that $u_0(0) = 1$ and $u_0(\infty) = 0$ to obtain

$$\int_0^\infty t dF_2(t) = \frac{1}{\lambda(u)}.$$

For the second moment, we simply apply twice an integration by parts to obtain

$$\int_0^\infty t^2 dF_2(t) = \frac{2}{\lambda(u)} \int_0^\infty u_0(t) dt.$$

□

Finally, estimating $\lambda(u)$ gives a direct estimation of the mean length of inter-spike intervals, as well as its variance. This is detailed in Section 4.

4. Estimation of invariant density and spike rate

The objective is to estimate the invariant density p and to deduce estimators of the spike rate and $\lambda(u)$. We start with the invariant density.

4.1. Invariant density estimation

In the neuronal context, the ion channel coordinate Y_t can not be measured and only discrete observations of X at discrete times $i\delta$, $i = 1, \dots, n$ with discretization step δ are available. The density p has no explicit formula. We therefore use the non-parametric adaptive estimation of p from observations $(X_{1\delta}, \dots X_{n\delta})$, proposed by [6].

Let us detail their estimator. Let K be some kernel $\mathcal{C}^2$ function with compact support A such that its partial derivatives functions $\frac{\partial K}{\partial x}$ and $\frac{\partial K}{\partial y}$ are in $\mathbb{L}^2(\mathbb{R})$, $\int K(x,y)dxdy = 1$ and $\int K^2(x,y)dxdy < \infty$. For all bandwidth $b = (b_1, b_2)$ with $b_1 > 0, b_2 > 0$, for all $(x,y) \in \mathbb{R}^2$, we denote

$$K_b(x,y) = \frac{1}{b_1 b_2} K\left(\frac{x}{b_1}, \frac{y}{b_2}\right).$$

When both coordinates are observed, the natural estimator of p for all $z = (x,y) \in \mathbb{R}^2$, is:

$$\tilde{p}_b(z) = \tilde{p}_b(x,y) := \frac{1}{n}\sum_{i=1}^{n} K_b\left(x - X_{i\delta}, y - Y_{i\delta}\right) = \frac{1}{n}\sum_{i=1}^{n} K_b(z - Z_{i\delta}). \tag{4.1}$$

When only X is observed, we replace Y by increments of X. Indeed, for any $i = 1, \dots, n$, when δ is small enough, we have

$$X_{(i+1)\delta} - X_{i\delta} = \int_{i\delta}^{(i+1)\delta} Y_t dt \approx \delta Y_{i\delta} \tag{4.2}$$

Let us thus define the following approximation of $Y_{i\delta}$:

$$\bar{Y}_{i\delta} = \frac{X_{(i+1)\delta} - X_{i\delta}}{\delta}$$

and define the 2-dimensional kernel estimator by

$$\hat{p}_b(x,y) := \frac{1}{n}\sum_{i=1}^{n} K_b\left(x - X_{i\delta}, y - \bar{Y}_{i\delta}\right). \tag{4.3}$$

The bandwidth $b = (b_1, b_2)$ has to be chosen to realize a trade-off between the bias of $\hat{p}_b$ and its variance. This is automatically achieved using the adaptive estimation procedure proposed by [6]. We can apply their procedure because we have already proved that the invariant density p decreases exponentially and is β-mixing (Section 2). Their procedure, inspired by [11], is the following. Let $\mathcal{B}_n = \{(b_{1,k}, b_{2,\ell}), k, \ell = 1/\sqrt{n}, \dots, c/\sqrt{n}\}$ be the set of possible bandwidths. Set for all $z = (x,y)$ and all $b, b' \in \mathcal{B}_n$

$$\hat{p}_{b,b'}(z) = K_{b'} \star \hat{p}_b(z) = \frac{1}{n}\sum_{i=1}^{n} K_{b'} \star K_b(x - X_{i\delta}, y - \bar{Y}_{i\delta}).$$

Now let

$$A(b) = \sup_{b' \in \mathcal{B}_n} \left(\|\hat{p}_{b,b'} - \hat{p}_{b'}\|^2 - V(b') \right)_+$$

with

$$V(b) = \kappa_1 \frac{1}{n b_1 b_2} \sum_{i=0}^{n-1} \beta(i\delta) + \kappa_2 \frac{\delta}{b_1 b_2^3},$$

where κ_1, κ_2 are numerical constants and $\beta(i\delta)$ are the β-mixing coefficients. The selection is then made by setting

$$\hat{b} = \arg\min_{b \in \mathcal{B}_n} (A(b) + V(b)) \tag{4.4}$$

[6] prove an oracle inequality for the final estimator $\hat{p}_{\hat{b}}$.

Theorem 4.1 (Comte et al's result)**.** *Set $p_b(z) = K_b \star p$ the function that is estimated without bias by $\hat{p}_n$. We have*

$$\mathbb{E}(\|\hat{p}_{\hat{b}} - p\|^2) \leq C \inf_{b \in \mathcal{B}_n} \left(\|p_b - p\|^2 + V(b) \right) + C \frac{\log(n)}{n\delta}$$

As explained by [6], the Goldenshluger and Lepski's procedure (4.4) is numerically demanding due to the double convolutions $\hat{p}_{b,b'}$, especially in the multidimensional case. They propose a simplified procedure based on [15], that we also implement in this paper. The selection of the bandwidth is the following:

$$\hat{\hat{b}} = \arg\min_{b \in \mathcal{B}_n} \left(\|\hat{p}_b - \hat{p}_{b_{min}}\|^2 + V(b) \right) \tag{4.5}$$

with $\kappa_1 = 0.1$ and $\kappa_2 = 0.001$ and $b_{min} = (1/\sqrt{n}, 1/\sqrt{n})$, as given in [6]. By plugging $\hat{\hat{b}}$ into 4.3 we obtain $\hat{p} := \hat{p}_{\hat{\hat{b}}}$ which is the final estimator of p.

4.2. Spike rate estimation

Equation (3.8) provides a good start to estimate the spike rate. The quantity that we estimate is $\lambda(u)$ for a large level u. By plugging the kernel estimator $\hat{p}$ of the invariant density, we define the following estimator of $\lambda(u)$:

$$\hat{\lambda}(u) = \int_0^\infty y \hat{p}(u, y)\, dy.$$

For some specific choices of kernel K, estimator $\hat{\lambda}(u)$ has an explicit expression. More precisely, let us consider a multiplicative two-dimensional kernel $K(x, y) = k(x)k(y)$, where k is a continuous and bounded kernel, such that $\int k(v)dv = 1$. Then we have

$$\hat{p}(u, y) = \frac{1}{n \hat{b}_1 \hat{b}_2} \sum_{i=1}^{n} k\left(\frac{u - X_{i\delta}}{\hat{b}_1} \right) k\left(\frac{y - \bar{Y}_{i\delta}}{\hat{b}_2} \right),$$

with $\hat{b}_1, \hat{b}_2$ the bandwidth estimated adaptively by (4.5). We get

$$\begin{aligned}\hat{\lambda}(u) &= \frac{1}{n\hat{b}_1\hat{b}_2}\sum_{i=1}^{n} k\left(\frac{u-X_{i\delta}}{\hat{b}_1}\right)\int_0^\infty yk\left(\frac{y-\bar{Y}_{i\delta}}{\hat{b}_2}\right)dy \\ &= \frac{1}{n\hat{b}_1}\sum_{i=1}^{n} k\left(\frac{u-X_{i\delta}}{\hat{b}_1}\right)\left(\hat{b}_2\int_{-\frac{\bar{Y}_{i\delta}}{\hat{b}_2}}^{\infty} yk(y)dy + \bar{Y}_{i\delta}\int_{-\frac{\bar{Y}_{i\delta}}{\hat{b}_2}}^{\infty} k(y)dy\right)\end{aligned}$$

For a Gaussian centered kernel k, we obtain:

$$\hat{\lambda}(u) = \frac{1}{n\hat{b}_1}\sum_{i=1}^{n} k\left(\frac{u-X_{i\delta}}{\hat{b}_1}\right)\left(\frac{\hat{b}_2}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\bar{Y}_{i\delta}}{\hat{b}_2}\right)^2} + \bar{Y}_{i\delta}\left(1-\Phi\left(-\frac{\bar{Y}_{i\delta}}{\hat{b}_2}\right)\right)\right), \tag{4.6}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the centered and reduced normal distribution.

Note that a CLT can be concluded for $\hat{\lambda}(u)$.

The next step is the estimation of the variance, and more precisely of the second moment of F_2, given by (3.14). First, we need to estimate $u_0(t) = \mathbb{P}\{U^{X.}_{(0,t)} = 0\}$. The idea is to link this function with the survival function of inter up-crossings interval.

For a fixed level u, let us assume that one up-crossing occurs at time 0 and let denote $\{T_i^u,\, i \geq 0\}$ the successive times of up-crossings after time 0 with $T_0^u = 0$.
Thanks to the stationarity of the process, the $(T_{i+1}^u - T_i^u,\, i \geq 0)$ are identically distributed. For any $i \geq 0$, we can rewrite $u_0(t)$ as follows

$$u_0(t) = \mathbb{P}(U^X_{(0,t)} = 0) = \mathbb{P}(T_1^u > t) = \mathbb{P}((T_{i+1}^u - T_i^u) > t).$$

A natural estimator of $u_0(t)$ from observations on interval $[0,T]$ is

$$\hat{u}_0^T(t) = \frac{1}{U^X_{(0,T)}}\sum_{i=0}^{U^X_{(0,T)}} \mathbf{1}_{(T_{i+1}^u - T_i^u)>t}, \tag{4.7}$$

where $U^X_{(0,T)}$ is the number of up-crossings in the interval $[0,T]$.

Lemma 4.1. *Estimator $\hat{u}_0^T(t)$ (4.7) based on observations on interval $[0,T]$ is a consistent estimator of $u_0(t)$ when T goes to infinity.*

Proof. of Lemma 4.1. Set $[\cdot]$ for the integer part. Let us rewrite $\hat{u}_0^T(t)$ as

$$\hat{u}_0^T(t) = \frac{1}{\frac{U^X_{(0,T)}}{T}}\frac{1}{T}\sum_{i=0}^{U^X_{(0,T)}} \mathbf{1}_{(T_{i+1}^u - T_i^u)>t}.$$

By the ergodic theorem, we have $U^X_{(0,T)} \sim T\int_0^\infty yp(u,y)dy$ a.s. We can thus deduce that

$$\left|\frac{1}{T}\left(\sum_{i=0}^{U^X_{(0,T)}} \mathbf{1}_{(T^u_{i+1}-T^u_i)>t} - \sum_{i=0}^{[T\int_0^\infty yp(u,y)dy]} \mathbf{1}_{(T^u_{i+1}-T^u_i)>t}\right)\right| \leq \left|\frac{U^X_{(0,T)}}{T} - \frac{[T\int_0^\infty yp(u,y)dy]}{T}\right| \to 0.$$

Thus we have $\lim_{T\to\infty} \hat{u}^T_0(t) = \lim_{T\to\infty} \frac{1}{\frac{U^X_{(0,T)}}{T}} \frac{1}{T} \sum_{i=0}^{[T\int_0^\infty yp(u,y)dy]} \mathbf{1}_{(T^u_{i+1}-T^u_i)>t}$.

We now focus on studying the RHS. The stationarity and the ergodic theorem imply that

$$\frac{1}{T}\sum_{i=0}^{[T\int_0^\infty yp(u,y)dy]} \mathbf{1}_{(T^u_{i+1}-T^u_i)>t} \to \int_0^\infty yp(u,y)dy\,\mathbb{P}(T^u_1 > t) \text{ a.s.}$$

So finally, we obtain:

$$\lim_{T\to\infty} \hat{u}^T_0(t) \to \frac{1}{\int_0^\infty yp(u,y)dy}\int_0^\infty yp(u,y)dy\,\mathbb{P}(T^u_1 > t) = \mathbb{P}(T^u_1 > t) = u_0(t) \text{ a.s.} \tag{4.8}$$

□

To estimate the second moment of F_2, we plug $\hat{u}^T_0(t)$ into formula (3.14):

$$\frac{2}{\hat{\lambda}(u)}\int_0^\infty \hat{u}^T_0(t)dt = \frac{2}{\hat{\lambda}(u)}\frac{1}{U^X_{(0,T)}}\sum_{i=0}^{U^X_{(0,T)}}(T^u_{i+1}-T^u_i) = \frac{2}{\hat{\lambda}(u)}\frac{T^u_{U^X_{(0,T)}}}{U^X_{(0,T)}}.$$

The final estimator of the variance of the length between two successive upcrossings at level u based on observations on the interval $[0,T]$ is thus

$$\hat{V}^u = \frac{2}{\hat{\lambda}(u)}\frac{T^u_{U^X_{(0,T)}}}{U^X_{(0,T)}} - \frac{1}{\hat{\lambda}(u)^2}. \tag{4.9}$$

5. Simulation

Three sets of parameter values of the FitzHugh-Nagumo are used in the simulations. A set that allows spike generation: $\varepsilon = 0.1$, $s = 0$, $\gamma = 1.5$, $\beta = 0.8$, $\sigma = 0.3$; a set that generates small excursions $s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.4$ and $\gamma = 1.5$ and a set that does not generate spikes $s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.5$ and $\gamma = 0.2$. Trajectories are simulated with time step $\delta = 0.02$ ms, $n = 20000$ and a Itô-Taylor scheme of order 2 of system (1.1). Figure 2 shows an example of such a simulation with trajectories of (X_t, U_t) for the first set of parameters that generates spikes (($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.1$ and $\gamma = 1.5$), as well as the transformed coordinate $Y_t = \frac{1}{\varepsilon}(X_t - X_t^3 - C_t - s)$ between 0 and 200ms ($n = 1000$).

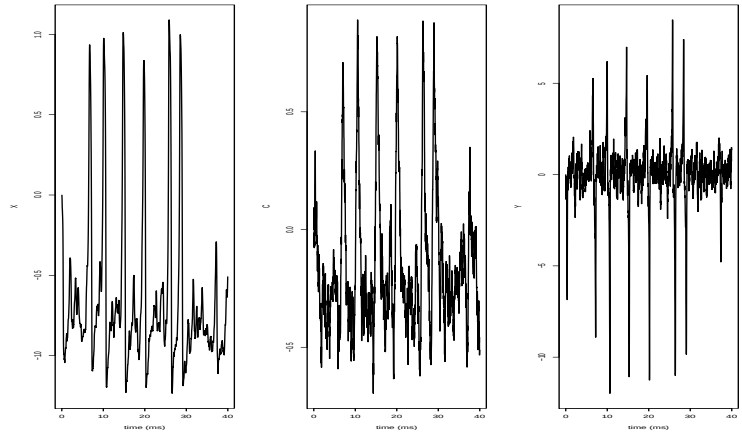

FIG 2. *FitzHugh-Nagumo simulation. Left: coordinate X along time. Middle: coordinate C. Right: transformed coordinate Y. Simulations are performed with $\delta = 0.02$, $n = 2000$. Parameters are $\varepsilon = 0.1$, $s = 0$, $\gamma = 1.5$, $\beta = 0.8$, $\sigma = 0.3$.*

We apply the adaptive estimation procedure (4.5) to estimate the invariant density p. The true density p has no closed form. To compare the estimator with the truth, we approximate p by numerically solving the associated hypoelliptic Fokker-Planck equation. A finite difference method is used to solve the Fokker-Planck equation [14]. The solver is very stable for the set of parameters that does not generate spikes ($\varepsilon = 0.5$). This is illustrated in the bottom plots of Figure 3. The density estimator $\hat{p}$ (red dotted line) is very closed to the 'true' stationary density p (black plain line). However, the finite difference approximation of stationary density appears to be unstable for the set of parameters that generates spikes or even small excursions (see black lines of top and middle plots of Figure 3). We tried to decrease the step of the discretization grid but the approximation remains unstable (recall that the PDE is hypoelliptic). We then compare the estimator with a Monte Carlo approximation of the stationary density. More precisely, 10 000 trajectories have been simulated on an interval $[0, T = n\delta]$. The last point of each trajectory is stored in a sample of iid simulation under the stationary regime. Then a kernel estimation procedure for iid data has been applied. The comparison with our estimation based on only one trajectory of dependent data is shown on Figure 4. The estimation is very close to the approximation of the stationary regime. Note also that in the spiking regime, as expected, the estimator of the marginal density in x (red dotted line) has two bounds, one corresponding to the subthreshold activity and the other to the spiking activity. To conclude, our estimation procedure is stable whatever the value of ε and fast to compute compared to the PDE solver of the Monte Carlo approximation.

Then the spike rate is estimated by formula (4.6) for the three sets of pa-

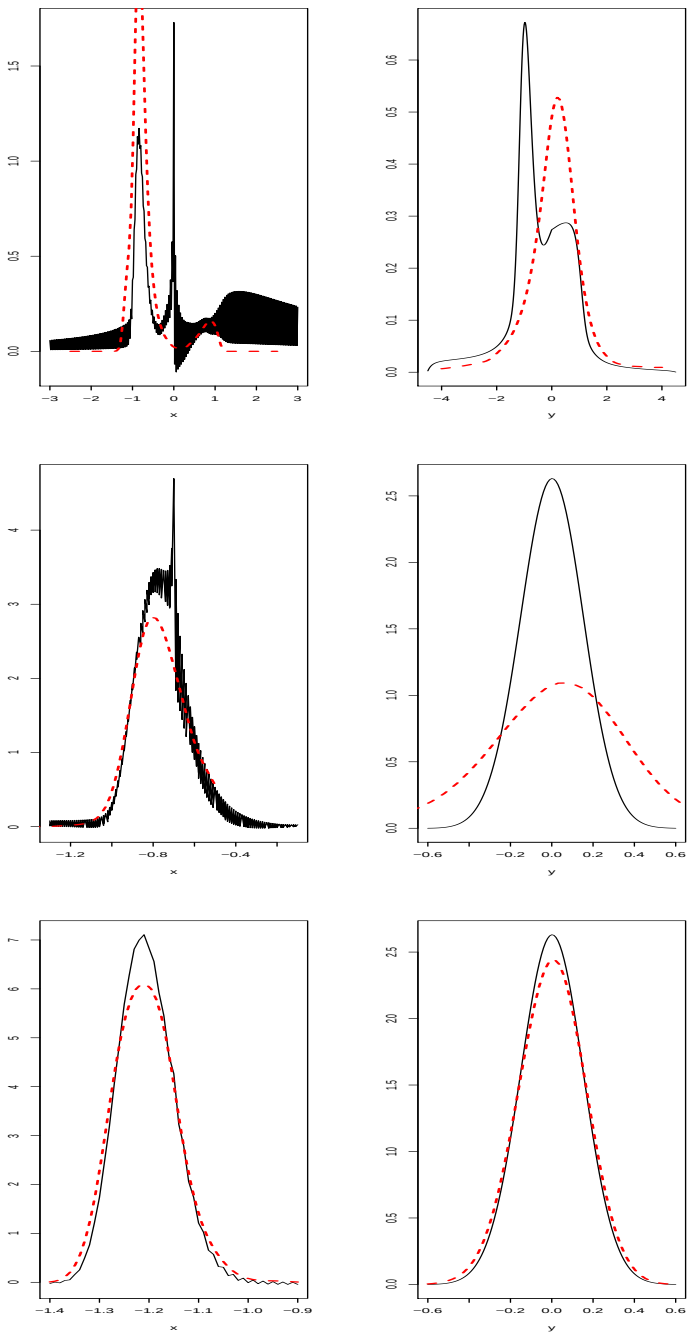

FIG 3. *Invariant density estimation. Left: marginal in x of the estimation $\hat{p}_b$ (red line) and true density approximated by a finite difference scheme (black line). Right: marginal in y of the estimation $\hat{p}_b$ (red line) and true density approximated by a finite difference scheme (black line). Top line: set of parameters that generate spikes ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.1$ and $\gamma = 1.5$). Middle line: set of parameters that generates few and small excursions ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.4$ and $\gamma = 1.5$) . Bottom line: set of parameters that does not generate spikes ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.5$ and $\gamma = 0.2$). The finite difference scheme is unstable in the two first cases (black curve is very noisy).*

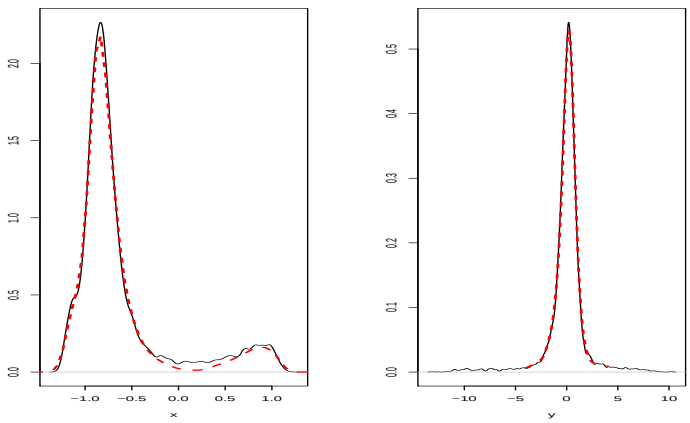

FIG 4. *Invariant density estimation. Left: marginal in x of the estimation $\hat{p}_b$ (red line) and true density approximated by a Monte Carlo scheme (black line). Right: marginal in y of the estimation $\hat{p}_b$ (red line) and true density approximated by a Monte Carlo scheme (black line). Simulations are performed with parameters that generate spikes ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.1$ and $\gamma = 1.5$).*

rameters that generate spikes, few spikes or no spikes. The expected number of up-crossings is estimated for level u between -0.5 and 1.5. The three curves (as functions of u) are plotted in Figure 5: black plain line for the set of parameters that generate spikes ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.1$ and $\gamma = 1.5$); red dotted line for the set of parameters that generates few and small excursions ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.4$ and $\gamma = 1.5$); green dashed line for the set of parameters that does not generate spikes ($s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.5$ and $\gamma = 0.2$). As expected, when spikes occur, the estimator is stable for $u \in [0; 0.8]$, interval that corresponds to the amplitude of large excursions and then suddenly decreases to 0. When no spike occur, the estimator is null. When only small sub-threshold excursions occur, the estimator decreases slowly to 0.

We now focus on the two first cases (large or small excursions with $\varepsilon = 0.1$ and $\varepsilon = 0.4$, respectively). For both cases, we estimate the spike rate with the two approaches presented in the paper. First, we compute the mean of $\hat{\lambda}(u)$ for $u \in [0.1, 0.6]$, this value is denoted $\bar{\lambda}$. Second, we compute ρ defined as the number of spikes divided by the length of the time interval. As said before, spikes are defined as excursions in the phase space. Different thresholds $v \in [0.1, 0.7]$ are used to define excursions $N_t(v)$. The mean of the corresponding spike rates is denoted $\bar{\rho}$ and superimposed on Figure 5 (horizontal lines). Table 1 displays the estimations $\bar{\rho}$, $\bar{\lambda}$, as well as the estimated mean and standard deviation (estimator (4.9)) of the length of intervals defined by two successive up-crossings. One can see that the two estimators $\bar{\lambda}$ and $\bar{\rho}$ are of the same order in both regimes. Recalling that $\bar{\lambda}$ is based on the estimation of the stationary density, this implies that its estimation is good. Therefore, the non-parametric

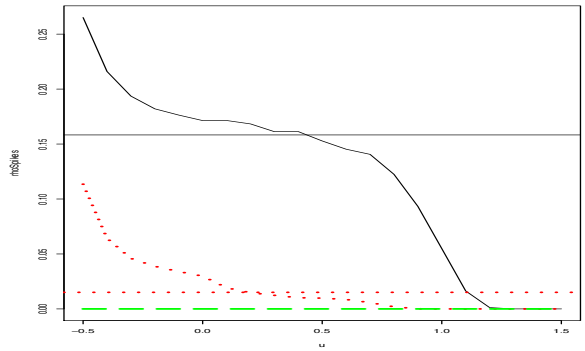

FIG 5. *Spike rate estimators* $\hat{\lambda}(u)$ *and* $\hat{\rho}$ *computed as the mean of* $\rho(0.1), \ldots, \rho(0.6)$. *Black plain curve: evolution of* $\hat{\lambda}(u)$ *with* u *and black plain line* $\hat{\rho}$ *for a set of parameters that generate spikes (*$s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.1$ *and* $\gamma = 1.5$*). Red dotted curve: evolution of* $\hat{\lambda}(u)$ *with* u *and red dotted line* $\hat{\rho}$ *for a set of parameters that generates few and small excursions (*$s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.4$ *and* $\gamma = 1.5$*). Green dashed curve: evolution of* $\hat{\lambda}(u)$ *with* u *for a set of parameters that does not generate spikes (*$s = 0$, $\beta = 0.8$, $\tilde{\sigma} = 0.3$, $\varepsilon = 0.5$ *and* $\gamma = 0.2$*).*

regime	$\bar{\rho}$	$\bar{\lambda}$	length of intervals mean (sd)
$\varepsilon = 0.1$	0.1568	0.1609	6.35 (6.32)
$\varepsilon = 0.4$	0.0115	0.0111	93.13 (82.70)

TABLE 1

FitzHugh-Nagumo simulation for two regimes ($\varepsilon = 0.1$ *and* $\varepsilon = 0.4$*). Estimation of the spike rate by up-crossings approach (*$\bar{\lambda}$*) and by number of excursions (*$\bar{\rho}$ *) and estimation of the mean and standard deviation of the length of intervals defined by two successive up-crossings.*

estimation is a good alternative to the numerical approximation of the solution of the hypoelliptic Fokker-Planck equation, especially when the numerical scheme is not stable for ε small. The up-crossing approach allows to estimated some characteristics of the length of inter up-crossings intervals (mean and standard deviation). In the spiking regime, mean and standard deviation are close. This is not the case when ε increases.

6. Conclusion

The FitzHugh-Nagumo is a neuronal model that describes the generation of spikes at the intracellular level. In this paper, we study a stochastic version of the model from a probabilistic point of view. The hypoellipticity is proved, as well as the existence and uniqueness of the stationary distribution. The bi-dimensional stochastic process is β-mixing. The stationary distribution can be estimated with an adaptive non-parametric estimator. Then, we focus on the distribution of the length between successive spikes. We propose to study this distribution through the distribution of the number of up-crossings. The distribution function of the length of the interval between two successive up-crossings is defined through the stationary distribution. This allows to propose an estimator of the expectation of this distribution. We also derive the second moment of this distribution, that allows to estimate the variance.

We illustrate the proposed estimators on a simulation study. Different regimes are explored, for different values of ε: regime with no, few or high generation of spikes. The true stationary density has no explicit distribution. It can be approximated numerically by solving the Fokker-Planck equation. We consider a finite difference scheme, which is however unstable in spiking regime. At the other hand, the non-parametric estimation of the stationary distribution reveals to be stable even in spiking regime. We also implement the estimator of the mean length of the interval between two successive up-crossings. This estimator is based on the estimator of the stationary distribution. It reveals to be close to mean spiking rate in the spiking regime.

It would be of interest in the future to apply the same approach to other stochastic intra-cellular neuronal models and to estimate the characteristics of their spiking process.

Appendix A: Proofs

A.1. Proof of Proposition 2.2

The proof follows the proof of Theorem 3.1 from [24] for the specific FHN model. Steps are the followings: 1/ find a Lyapounov function that is lower bounded, 2/ show that this function is greater than the exponential of the Hamiltonian of the system, 3/ apply Theorem 2.4 of [24] to prove the existence and unicity of the invariant probability, the existence of the moments.

1. The first step is finding a Lyapounov function $\Psi(x,y)$ such that there exist a compact $K \in \mathbb{R}^2$ and constants $C, \xi > 0$, such that

$$-\frac{\mathcal{L}\Psi}{\Psi} \geq \xi 1_{K^c} - C 1_K. \tag{A.1}$$

The choice of the Lyapounov function is not trivial. Following [24], we choose

$$\Psi(x,y) = e^{F(x,y) - \inf_{\mathbb{R}^2} F} \tag{A.2}$$

with

$$\begin{aligned} F(x,y) &= aH(x,y) + by\,G(x) + y\,W'(x) + b\,U(x) \\ &= a\left(\frac{1}{2}y^2 + V(x)\right) + by\,G(x) + y\,W'(x) + b\,U(x) \end{aligned}$$

where $H(x,y) = \frac{1}{2}y^2 + V(x)$ is the Hamiltonian and the functions $G(x), W(x), U(x)$ will now be defined, such as the two constants a and b. With this form of $\Psi(x,y)$, we have

$$-\frac{\mathcal{L}\Psi}{\Psi} = -\mathcal{L}F - \frac{1}{2}\sigma^2 |\partial_y F|^2.$$

We have to compute $-\frac{\mathcal{L}\Psi}{\Psi}$. We have:

$$\begin{aligned} \mathcal{L}F &= \frac{a}{2}\sigma^2 - ay^2 c(x) + y^2(bG'(x) + W''(x)) \\ &\quad -y(bc(x)G(x) - bU'(x) + c(x)W'(x)) \\ &\quad -V'(x)(bG(x) + W'(x)) \\ \frac{1}{2}\sigma^2(\partial_y F(x,y))^2 &= \frac{1}{2}\sigma^2\left(ay + bG(x) + W'(x)\right)^2 \end{aligned}$$

We will now detail our choice of functions $G(x), U(x), W(x)$ and constants a, b. We can bound $\frac{1}{2}\sigma^2(\partial_y F(x,y))^2$ by

$$\frac{1}{2}\sigma^2(\partial_y F(x,y))^2 \leq 2\sigma^2\left(a^2y^2 + (bG(x) + W'(x))^2\right)$$

Thus

$$\begin{aligned} -\frac{\mathcal{L}\Psi}{\Psi} &\geq -\frac{a}{2}\sigma^2 + y^2(ac(x) - 2\sigma^2 a^2 - 2(bG'(x) + W''(x))) \\ &\quad +y\left(b\left(c(x)G(x) - U'(x)\right) + c(x)W'(x)\right) \\ &\quad +V'(x)(bG(x) + W'(x)) - 2\sigma^2(bG(x) + W'(x))^2 \end{aligned}$$

We consider a function $G(x)$ such that

$$G(x) = \Phi(|x|)\frac{x}{|x|}, \qquad \text{for } x \neq 0$$

with $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ is a non decreasing smooth function equal to zero on a small neighborhood of 0 and equal to 1 for $|x| \geq 1$.

We choose a such that $0 < a < \frac{c}{2\sigma^2}$ where $c = 1 - \frac{1}{\varepsilon}$ is the lower bound of $c(x)$. Then we choose b such that

$$bG'(x) < \frac{1}{8}(ac - 2\sigma^2 a^2)$$

We choose the function W, of compact support, concave such that for all x

$$-W''(x) \geq -\frac{1}{8}(ac - 2\sigma^2 a^2)$$

This allows to control the term in y^2:

$$(ac(x) - 2\sigma^2 a^2 - 2(bG'(x) + W''(x))) \geq (ac - 2\sigma^2 a^2 - \frac{1}{2}(ac - 2\sigma^2 a^2)) > 0$$

Now, we bound the term in y. First note that by definition of W, there exists a constant M_1 such that

$$yc(x)W'(x) \geq -M_1|y|$$

Then, we define the function U such that the term $c(x)G(x) - U'(x)$ can be controlled. We choose $U(x)$ verifying

$$\sup_{x\in\mathbb{R}} |c(x)G(x) - U'(x)| < +\infty$$

Because $c(x) = \frac{1}{\varepsilon}(3x^2 - 1 + \varepsilon)$, we can take $U(x)$ such that

$$U'(x) = \begin{cases} \frac{3}{\varepsilon}x^2\Phi(|x|) & \text{if} \quad x \geq 0 \\ -\frac{3}{\varepsilon}x^2\Phi(|x|) & \text{if} \quad x \leq 0 \end{cases}$$

Thus, we obtain that there exists a constant M_2 such that

$$y\left(b\left(c(x)G(x) - U'(x)\right) + c(x)W'(x)\right) \geq -M2|y|.$$

Now, we want to control the constant term $(bG(x) + W'(x))(V'(x) - 2\sigma^2(bG(x) + W'(x)))$. Given the form of function G and $V'(x) = x^3 + x(\gamma - 1) + (s + \beta)$, we have that $V'(x)G(x) \to +\infty$ as $|x| \to +\infty$ and that $G^2(x) \to +\infty$ as $|x| \to +\infty$. The function W has compact support thus $W'(x)V'(x) \to 0$, $W'(x)G(x) \to 0$ and $W'^2(x) \to 0$ as $|x| \to +\infty$. So, the constant term is lower bounded.
Finally,

$$\lim_{|x|+|y|\to\infty} (-\frac{\mathcal{L}\Psi}{\Psi}) = +\infty.$$

We thus have the existence of a compact such that (A.1) holds.

2. We want to show that Ψ is lower bounded by 1. We can choose for any fixed $\delta > 0$ the constant a in $]\frac{c}{2\sigma^2} - \frac{\delta}{2}, \frac{c}{2\sigma^2}[$. The Lyapounov function is

defined as $\Psi(x,y) = e^{aH(x,y)+by\,G(x)+y\,W'(x)+b\,U(x)-\inf_{\mathbb{R}^2} F}$, we can prove that there exists a constant B such that

$$\Psi \geq B \exp\left((\frac{c}{2\sigma^2} - \delta) H(x,y) \right).$$

Therefore $\Psi \geq 1$.

3. Applying Theorem 2.4 of [24] leads to the fact that
   - There is a unique invariant probability measure μ.
   - μ satisfies $\int \Psi d\mu < +\infty$. Given the exponential form of Ψ, this means that μ integrates any polynom in (x,y) and thus the existence of any moments.

□.

Proof. of Lemma 3.1. Let $F \in B_\Psi$ and $\{F_n\}$ be a sequence of bounded functions such that $F_n \uparrow F$. Inequality (2.3) gives $||P_t F||_\Psi \leq (D\rho^t + 1)||F||_\Psi < \infty$. Moreover, as in [24], let us introduce the exponential local martingale

$$M_t = \exp\left[-\frac{1}{\sigma} \int_0^t (c(X_s,Y_s)Y_s + \nabla_x V(X_s)) dW_s + \frac{1}{2\sigma^2} \int_0^t (c(X_s,Y_s)Y_s + \nabla_x V(X_s))^2 ds \right].$$

By using monotone convergence theorem first and then the probabilistic representation as in [24], we get

$$\mathbb{E}^{\mathbb{P}^0_z}[M_t F(\sigma \int_0^t W_s ds, \sigma W_t)] = \lim_{n\to\infty} \mathbb{E}^{\mathbb{P}^0_z}[M_t F_n(\sigma \int_0^t W_s ds, \sigma W_t)]$$

$$= \lim_{n\to\infty} \int p_t(z,z') F_n(z') dz' = \int p_t(z,z') F(z') dz' = P_t F(z) < \infty.$$

Then we recover the representation $P_t F(z) = \mathbb{E}^{\mathbb{P}^0_z}[M_t F(\sigma \int_0^t W_s ds, \sigma W_t)]$. Now function G_{y_1} trivially belongs to B_Ψ. Hence $P_t G_{y_1}(z) = \mathbb{E}^{\mathbb{P}^0_z}[M_t |\sigma W_t - y_1|]$. Thus by using the Levy modulus of continuity of Brownian motion we get

$$|P_t G_{y_1}(z) - |y - y_1|| \leq \mathbb{E}^{\mathbb{P}^0_z}[M_t |\sigma W_t - y|] \leq C \sqrt{t \log \frac{1}{t}} \mathbb{E}^z[M_t] = O(t^{1/2-\varepsilon})$$

for $\varepsilon > 0$. This proves the lemma. Remark that the convergence is uniform in z. □

Acknowledgments

Adeline Samson has been partially supported by the LabExPERSYVAL-Lab (ANR-11-LABX-0025-01). José R. Leòn has been partially supported by the INRIA International Chairs Program.

References

[1] Azais, J. M. and Wschebor, M. (2009). *Level sets and extrema of random processes and fiels*. Wiley.

[2] Berglund, N. and Gentz, B. (2006). *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*. Springer.

[3] Berglund, N. and Landon, D. (2012). Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh-Nagumo model. *Nonlinearity* **25**.

[4] Cattiaux, P., Leon, J. R. and Prieur, C. (2014). Estimation for stochastic damping Hamiltonian systems under partial observation. I. invariant density. *Stochastic Processes and their Applications* **124** 1236–1260.

[5] Cattiaux, P., Leon, J., Pineda Centeno, A. and Prieur, C. (2017). An overlook on statistical inference issues for stochastic damping hamiltonian systems under the fluctuation-dissipation condition. *Statistics* **51** 11-29.

[6] Comte, F., Prieur, C. and Samson, A. (2016). Adaptive estimation for stochastic damping Hamiltonian systems under partial observation. *submitted*.

[7] Cramer, H. and Leadbetter, M. R. (1967). *Stationary and related stochastic processes*. J. Wiley.

[8] Di Bernardino, E., Léon, J. R. and Tchumatchenko, T. (2014). Cross-Correlations and Joint Gaussianity in Multivariate Level Crossing Models. *Journal of Mathematical Neuroscience* **4** 22.

[9] Doukhan, P. (1994). *Mixing Properties and examples*. Springer.

[10] Gerstner, W. and Kistler, W. (2002). *Spiking neuron models*. Cambridge University Press.

[11] Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality. *Annals of Statistics* **39** 1608-1632.

[12] Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *J. Physiol.* **117** 500-544.

[13] Konakov, V., Menozzi, S. and Molchanov, S. (2010). Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Ann. Institut Henri Poincare,* **46** 908-923.

[14] Kumar, P. and Narayanan, S. (2006). Solution of Fokker-Planck equation by finite element and finite difference methods for nonlinear systems. *Sadhana* **31** 445-461.

[15] Lacour, C., Massart, P. and Rivoirard, V. (2016). Estimator selection: a new method with applications to kernel density estimation. *submitted*.

[16] Lindner, B. and Schimansky-Geier, L. (1999). Analytical approach to the stochastic FitzHugh-Nagumo system and coherence resonance. *Physical review E* **60** 7270-7276.

[17] Lindner, B., Garcia-Ojalvo, J., Neiman, A. and Schimansky-

Geier, L. (2004). Effects of noise in excitable systems. *Physics Reports* **392** 321-424.

[18] Longtin, A. (2000). Effect of noise on the tuning properties of excitable systems. *chaos solitons and fractals* **11** 1835-1848.

[19] Morris, C. and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **35** 193-213.

[20] Pakdaman, K., Thieullen, M. and Wainrib, G. (2010). Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Adv. in Appl. Probab.* **42** 761-794.

[21] Reynaud-Bouret, P., Rivoirard, V. and Tuleau-Malot, C. (2013). Inference of functional connectivity in Neurosciences via Hawkes processes. *1st IEEE Global Conference on Signal and Information Processing, Austin, Texas.*

[22] Reynaud-Bouret, P., Rivoirard, V., Grammont, F. and Tuleau-Malot, C. (2014). Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *Journal of Mathematical Neuroscience* **4** 3.

[23] Tuckwell, H. and Rodriguez, R. (1998). Analytical and simulation results for stochastic Fitzhugh-Nagumo neurons and neural networks. *Journal of Computational Neuroscience* **5** 91-113.

[24] Wu, L. (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Processes and their Applications* **91** 205-238.