



**HAL**  
open science

# Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh-Nagumo model

Nils Berglund, Damien Landon

► **To cite this version:**

Nils Berglund, Damien Landon. Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh-Nagumo model. 2011. hal-00591089v1

**HAL Id: hal-00591089**

**<https://hal.science/hal-00591089v1>**

Preprint submitted on 6 May 2011 (v1), last revised 5 Apr 2012 (v4)

**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.

# Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh–Nagumo model

Nils Berglund<sup>\*†</sup> and Damien Landon<sup>\*†</sup>

May 6, 2011

## Abstract

We study the stochastic FitzHugh–Nagumo equations, modelling the dynamics of neuronal action potentials, in parameter regimes characterised by mixed-mode oscillations. The interspike time interval is related to the random number of small-amplitude oscillations separating consecutive spikes. We prove that this number has an asymptotically geometric distribution, whose parameter is related to the principal eigenvalue of a substochastic Markov chain. We provide rigorous bounds on this eigenvalue in the small-noise regime, and derive an approximation of its dependence on the system’s parameters for a large range of noise intensities. This yields a precise description of the probability distribution of observed mixed-mode patterns and interspike intervals.

*Mathematical Subject Classification.* 60H10, 34C26 (primary) 60J20, 92C20 (secondary)

*Keywords and phrases.* FitzHugh–Nagumo equations, interspike interval distribution, mixed-mode oscillation, singular perturbation, fast–slow system, dynamic bifurcation, canard, substochastic Markov chain, principal eigenvalue, quasi-stationary distribution.

## 1 Introduction

Deterministic conduction-based models for action-potential generation in neuron axons have been much studied for over half a century. In particular, the four-dimensional Hodgkin–Huxley equations [HH52] have been extremely successful in reproducing the observed behaviour. Of particular interest is the so-called excitable regime, when the neuron is at rest, but reacts sensitively and reliably to small external perturbations, by emitting a so-called spike. Until recently, most research efforts have been concerned with the effect of deterministic perturbations. During the last decade, however, there has been growing interest in quantifying the effect of random perturbations as well. See for instance [TP01b, TTP02, Row07] for numerical studies of the effect of noise on the interspike interval distribution in the Hodgkin–Huxley equations.

Being four-dimensional, the Hodgkin–Huxley equations are notoriously difficult to study already in the deterministic case. For this reason, several simplified models have been introduced. In particular, the two-dimensional FitzHugh–Nagumo equations [Fit55,

---

<sup>\*</sup>MAPMO, CNRS – UMR 6628, Université d’Orléans, Fédération Denis Poisson – FR 2964, B.P. 6759, 45067 Orléans Cedex 2, France.

<sup>†</sup>Supported by ANR project MANDy, Mathematical Analysis of Neuronal Dynamics, ANR-09-BLAN-0008-01.

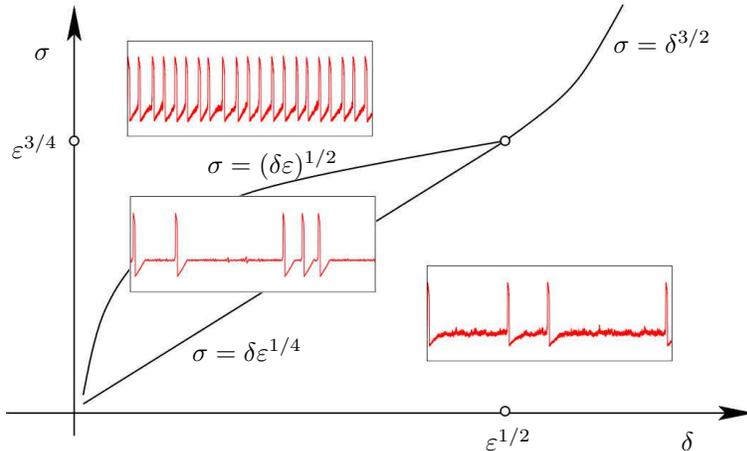


FIGURE 1. Schematic phase diagram of the stochastic FitzHugh–Nagumo equations. The parameter  $\sigma$  measures the noise intensity,  $\delta$  measures the distance to the singular Hopf bifurcation, and  $\varepsilon$  is the timescale separation. The three main regimes are characterised by rare isolated spikes, clusters of spikes, and repeated spikes.

[Fit61, NAY62], which generalise the Van der Pol equations, are able to reproduce one type of excitability, which is associated with a Hopf bifurcation (excitability of type II [Izh00]).

The effect of noise on the FitzHugh–Nagumo equations or similar excitable systems has been studied numerically [KP03, KP06, TGOS08, BKLLC11] and using moment methods [TP01a] and approximate solutions of the Fokker–Planck equations [LSG99, SK]. Rigorous results on the oscillatory (as opposed to excitable) regime have been obtained using the theory of large deviations [MVEE05, DT] and by a detailed description of sample paths near so-called canard solutions [Sow08].

An interesting connection between excitability and mixed-mode oscillations (MMOs) was observed by Kosmidis and Pakdaman [KP03, KP06], and further analysed by Muratov and Vanden-Eijnden [MVE08]. MMOs are patterns of alternating large- and small-amplitude oscillations (SAOs), which occur in a variety of chemical and biological systems. In the deterministic case, at least three variables are necessary to reproduce such a behaviour (see [DGK<sup>+</sup>11] for a recent review of deterministic mechanisms responsible for MMOs). As observed in [KP03, KP06, MVE08], in the presence of noise, already the two-dimensional FitzHugh–Nagumo equations can display MMOs. In fact, depending on the three parameters noise intensity  $\sigma$ , timescale separation  $\varepsilon$  and distance to the Hopf bifurcation  $\delta$ , a large variety of behaviours can be observed, including sporadic single spikes, clusters of spikes, bursting relaxation oscillations and coherence resonance. Figure 1 shows a simplified version of the phase diagram proposed in [MVE08].

In the present work, we build on ideas of [MVE08] to study in more detail the transition from rare individual spikes, through clusters of spikes and all the way to bursting relaxation oscillations. We begin by giving a precise mathematical definition of a random variable  $N$  counting the number of SAOs between successive spikes. It is related to a substochastic continuous-space Markov chain, keeping track of the amplitude of each SAO. We use this Markov process to prove that the distribution of  $N$  is asymptotically geometric, with a parameter directly related to the principal eigenvalue of the Markov chain (Theorem 2.2). A similar behaviour has been obtained for the length of bursting relaxation oscillations in a three-dimensional system [HM09]. In the weak noise regime, we derive rigorous bounds on the principal eigenvalue and on the expected number of SAOs (Theorem 2.4). Finally,

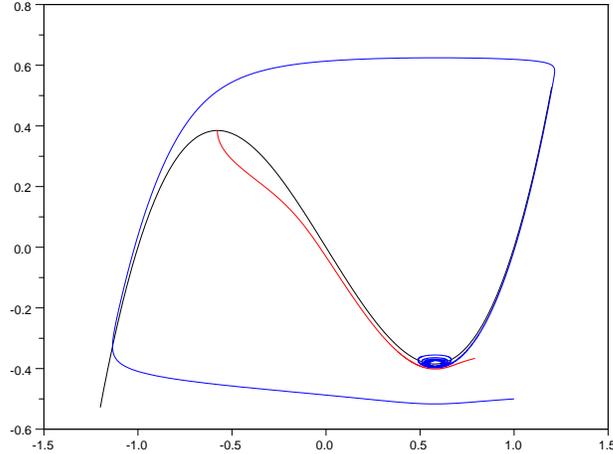


FIGURE 2. Two orbits of the deterministic FitzHugh–Nagumo equations (2.1) for  $\varepsilon = 0.05$ . The red orbit is the separatrix.

we derive an approximate expression for the distribution of  $N$  for all noise intensities up to the regime of repeated spiking.

The rest of this paper is organised as follows. Section 2 contains the precise definition of the model, the construction of the Markov chain, and the statement of all results. The following sections contain the proofs. Section 3 reviews a number of general results on continuous-space Markov chains, and contains the proof of Theorem 2.2. Section 4 studies the dynamics of the system near the separatrix, using a particular set of coordinates. In Section 5, we construct another system of coordinates, allowing us to describe the dynamics of each SAO, and to complete the proof Theorem 2.4.

## Acknowledgements

It’s a pleasure to thank Barbara Gentz, Simona Mancini and Khashayar Pakdaman for numerous inspiring discussions, Athanasios Batakis for advice on harmonic measures, and Christian Kuehn for sharing his deep knowledge on mixed-mode oscillations. NB was partly supported by the International Graduate College “Stochastics and real world models” at University of Bielefeld. NB and DL thank the CRC 701 at University of Bielefeld for hospitality.

## 2 Model and results

### 2.1 Definition of the model

We will consider random perturbations of the deterministic FitzHugh–Nagumo equations given by

$$\begin{aligned} \varepsilon \dot{x} &= x - x^3 + y \\ \dot{y} &= a - x, \end{aligned} \tag{2.1}$$

where  $a \in \mathbb{R}$  and  $\varepsilon > 0$  is a small parameter. The smallness of  $\varepsilon$  implies that  $x$  changes rapidly, unless the state  $(x, y)$  is close to the nullcline  $\{y = x^3 - x\}$ . Thus System (2.1) is called a fast-slow system,  $x$  being the fast variable and  $y$  the slow one.

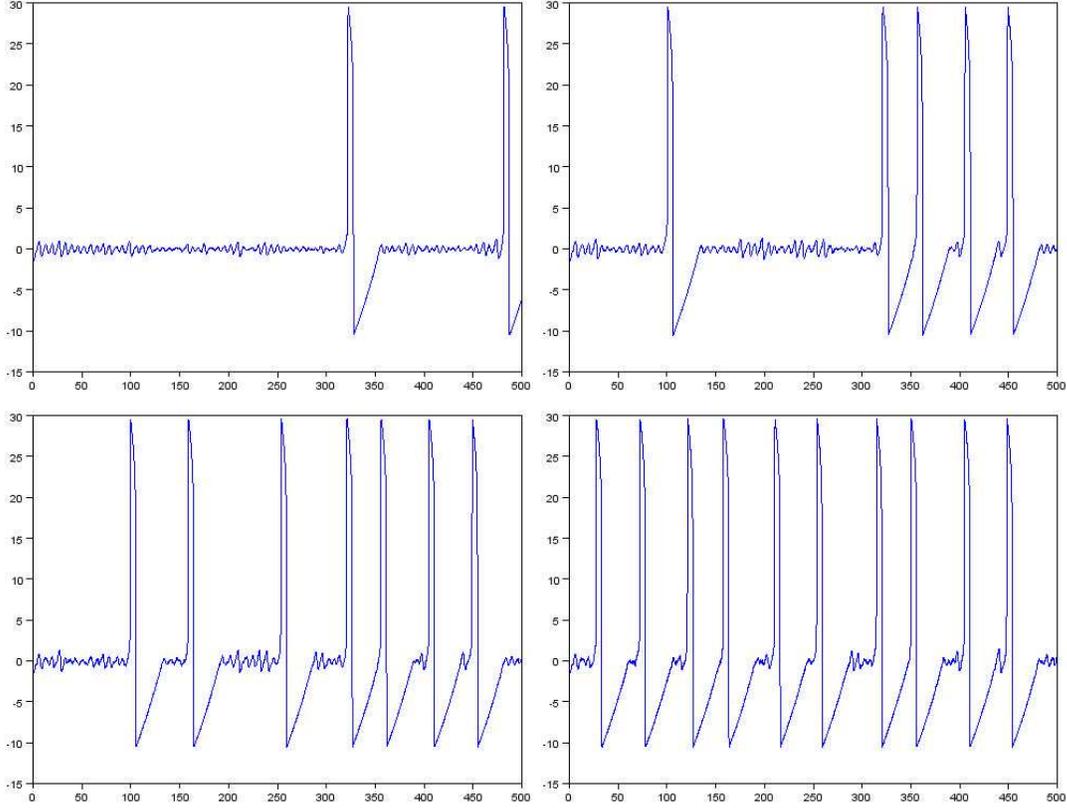


FIGURE 3. Examples of time series of the stochastic FitzHugh–Nagumo equations (2.2). The plots show the functions  $t \mapsto \xi_t$ , where the variable  $\xi$  is defined in Section 2.4. Parameter values are  $\varepsilon = 0.01$  and  $\delta = 3 \cdot 10^{-3}$  for the top row,  $\delta = 5 \cdot 10^{-3}$  for the bottom row. The noise intensities are given by  $\sigma_1 = \sigma_2 = 1.46 \cdot 10^{-4}$ ,  $1.82 \cdot 10^{-4}$ ,  $2.73 \cdot 10^{-4}$  and  $3.65 \cdot 10^{-4}$ .

System (2.1) admits a unique stationary point  $P = (a, a^3 - a)$ . The linearisation of the vector field at  $P$  has eigenvalues  $(-\delta \pm \sqrt{\delta^2 - \varepsilon})/\varepsilon$ , where  $\delta = (3a^2 - 1)/2$ , implying that  $P$  is a stable node when  $\delta > \sqrt{\varepsilon}$ , a stable focus when  $0 < \delta < \sqrt{\varepsilon}$ , an unstable focus when  $-\sqrt{\varepsilon} < \delta < 0$ , and an unstable node when  $\delta < -\sqrt{\varepsilon}$ . The system undergoes a so-called singular Hopf bifurcation at  $\delta = 0$  [BE86, BE92, Bra98].

We are interested in the excitable regime, when  $\delta$  is small and positive. In this situation,  $P$  is a stable stationary point, corresponding to a quiescent neuron. However, a small perturbation of the initial condition, e.g. a slight decrease of the  $y$ -coordinate, causes the system to make a large excursion to the region of negative  $x$ , before returning to  $P$  (Figure 2). This behaviour corresponds to a spike in the neuron’s membrane potential, followed by a return to the quiescent state. If  $\delta < \sqrt{\varepsilon}$ , the stationary point being a focus, the return to rest involves small-amplitude oscillations (SAOs), of exponentially decaying amplitude.

For later use, let us fix a particular orbit delimiting the spiking and quiescent regimes, called *separatrix*. An arbitrary but convenient choice for the separatrix is the negative-time orbit of the local maximum  $(-1/\sqrt{3}, 2/(3\sqrt{3}))$  of the nullcline (Figure 2).

In this work we consider random perturbations of the deterministic system (2.1) by Gaussian white noise. They are described by the system of Itô stochastic differential

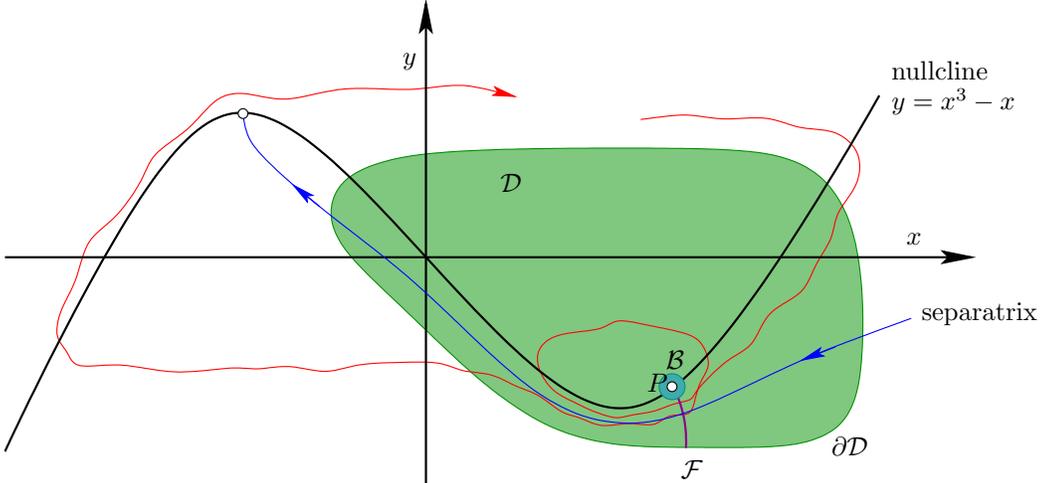


FIGURE 4. Definition of the number  $N$  of SAOs. The sample path (red) enters the region  $\mathcal{D}$ , and intersects twice the line  $\mathcal{F}$  before leaving  $\mathcal{D}$ , making another spike. Thus  $N = 2$  in this example. The separatrix is represented in blue.

equations (SDEs)

$$\begin{aligned} dx_t &= \frac{1}{\varepsilon}(x_t - x_t^3 + y_t) dt + \frac{\sigma_1}{\sqrt{\varepsilon}} dW_t^{(1)} \\ dy_t &= (a - x_t) dt + \sigma_2 dW_t^{(2)}, \end{aligned} \quad (2.2)$$

where  $W_t^{(1)}$  and  $W_t^{(2)}$  are independent, standard Wiener processes, and  $\sigma_1, \sigma_2 > 0$ . The scaling in  $1/\sqrt{\varepsilon}$  of the noise intensity in the first equation is chosen because the variance of the noise term then grows like  $\sigma_1^2 t/\varepsilon$ , so that  $\sigma_1^2$  measures the ratio of diffusion and drift for the  $x$ -variable, while  $\sigma_2^2$  plays the same rôle for the  $y$ -variable.

Figure 3 shows a selection of time series for the stochastic FitzHugh–Nagumo equations (2.2). For the chosen parameter values, one can clearly see large-amplitude spikes, separated by a random number of SAOs.

## 2.2 The random variable “Number of SAOs”

Let us now define an integer-valued random variable  $N$ , counting the number of small-amplitude oscillations the stochastic system performs between two consecutive spikes. The definition is going to be topological, making our results robust to changes in details of the definition. We start by fixing a bounded set  $\mathcal{D} \subset \mathbb{R}^2$ , with smooth boundary  $\partial\mathcal{D}$ , containing the stationary point  $P$  and a piece of the separatrix (Figure 4). Any excursion of the sample path  $(x_t, y_t)_t$  outside  $\mathcal{D}$  will be considered as a spike.  $N$  is defined as the number of times the sample path winds around  $P$  before leaving  $\mathcal{D}$ , and thus displaying a spike.

To define  $N$  precisely, we let  $\mathcal{B}$  be a small ball of radius  $\rho > 0$  centred in  $P$ . Then we draw a smooth curve  $\mathcal{F}$  from  $\mathcal{B}$  to the boundary  $\partial\mathcal{D}$ , which we parametrise by a variable  $r \in [0, 1]$  proportional to arclength (the results will be independent, however, of the choice of  $\mathcal{F}$  and of  $r$ ). We extend the parametrisation of  $\mathcal{F}$  to a polar-like parametrisation of all  $\mathcal{D} \setminus \mathcal{B}$ , i.e. we choose a diffeomorphism  $T : [0, 1] \times \mathbb{S}^1 \rightarrow \mathcal{D}$ ,  $(r, \varphi) \mapsto (x, y)$ , where

$T^{-1}(\mathcal{F}) = \{\varphi = 0\}$ ,  $T^{-1}(\partial\mathcal{D}) = \{r = 0\}$  and  $T^{-1}(\partial\mathcal{B}) = \{r = 1\}$ . We also arrange that  $\dot{\varphi} > 0$  near  $P$  for the deterministic flow.

Consider the process  $(r_t, \varphi_t)_t$  (where the angle  $\varphi$  has been lifted from  $\mathbb{S}^1$  to  $\mathbb{R}_+$ ). Given an initial condition  $(r_0, 0) \in T^{-1}(\mathcal{F})$  and an integer  $M \geq 1$ , we define the stopping time

$$\tau = \inf\{t > 0: \varphi_t \in \{2\pi, -2M\pi\} \text{ or } r_t \in \{0, 1\}\}. \quad (2.3)$$

There are four cases to consider:

- The case  $r_\tau = 0$  corresponds to the sample path  $(x_t, y_t)$  leaving  $\mathcal{D}$ , and thus to a spike. This happens with strictly positive probability, by ellipticity of the diffusion process (2.2). In this situation, we set by convention  $N = 1$ .
- In the case  $\varphi_\tau = 2\pi$  and  $r_\tau \in (0, 1)$ , the sample path has returned to  $\mathcal{F}$  after performing a complete revolution around  $P$ , staying all the while in  $\mathcal{D} \setminus \mathcal{B}$ . This corresponds to an SAO, and thus  $N \geq 2$ .
- The case  $r_\tau = 1$  corresponds to the sample path entering  $\mathcal{B}$ , which we consider as the neuron reaching the quiescent state. In that case we simply wait until the state leaves  $\mathcal{B}$  again and either hits  $\mathcal{F}$  or leaves  $\mathcal{D}$ .
- The case  $\varphi_\tau = -2M\pi$  and  $r_\tau \in (0, 1)$  represents the (unlikely) event that the sample path winds  $M$  time around  $P$  in the wrong direction. We introduce this case for technical reasons only, as we will need  $\tau$  to be the first-exit time of a bounded set. For simplicity, we also consider this situation as one SAO.

As long as  $\varphi_\tau \in (0, 1)$ , we repeat the above procedure, incrementing  $N$  at each iteration. This yields a sequence  $(R_0, R_1, \dots, R_{N-1})$  of random variables, describing the position of the successive intersections of the path with  $\mathcal{F}$ , separated by rotations around  $P$ , and up to the first exit from  $\mathcal{D}$ .

**Remark 2.1.** The above definition of  $N$  is the simplest one to analyse mathematically. There are several possible alternatives. One can, for instance, introduce a quiescent state  $(x, y) \in \mathcal{B}$ , and define  $N$  as the number of SAOs until the path either leaves  $\mathcal{D}$  or enters  $\mathcal{B}$ . This would allow to keep track of the number of SAOs between successive spikes and/or quiescent phases. Another possibility would be to count rotations in both the positive and negative directions. For simplicity, we stick here to the above simplest definition of  $N$ , but we plan to make a more refined study in a future work.

The sequence  $(R_n)_n$  forms a substochastic Markov chain on  $E = (0, 1)$ , with kernel

$$K(R, A) = \mathbb{P}\{\varphi_\tau = 2\pi, r_\tau \in A \mid \varphi_0 = 0, r_0 = R\}, \quad R \in E, A \subset E \text{ a Borel set}. \quad (2.4)$$

The Markov chain is substochastic because  $K(R, E) < 1$ , due to the positive probability of sample paths leaving  $\mathcal{D}$ . We can make it stochastic in the usual way by adding a cemetery state  $\Delta$  to  $E$  (the spike), and setting  $K(R, \Delta) = 1 - K(R, E)$ ,  $K(\Delta, \Delta) = 1$ . The number of SAOs is then given by

$$N = \inf\{n \geq 1: R_n = \Delta\} \in \mathbb{N} \cup \{\infty\} \quad (2.5)$$

(we set  $\inf \emptyset = \infty$ ). A suitable extension of the well-known Perron–Frobenius theorem (see [Jen12, KR50, Bir57] and Section 3) shows that  $K$  admits a maximal eigenvalue  $\lambda_0$ , which is real and simple. It is called the *principal eigenvalue* of  $K$ . If there exists a probability measure  $\pi_0$  such that  $\pi_0 K = \lambda_0 K$ , it is called the *quasi-stationary distribution (QSD)* of the kernel  $K$  [SVJ66].

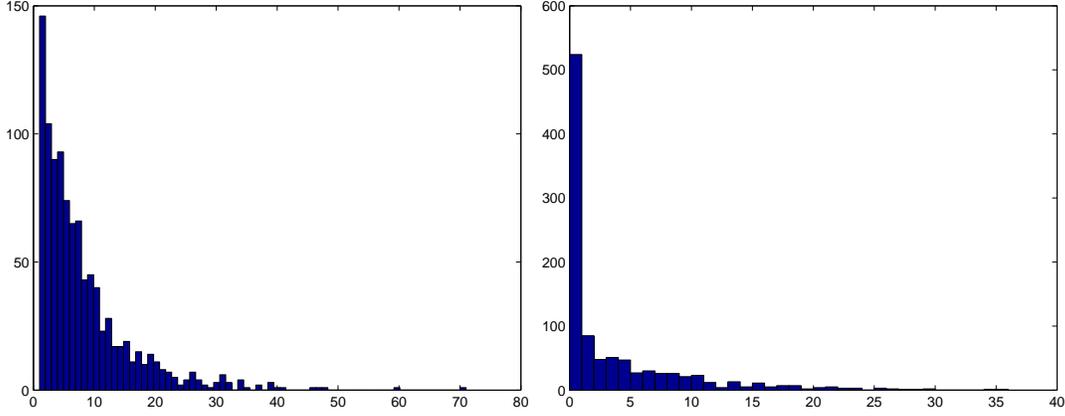


FIGURE 5. Histograms of numerically simulated distributions of the SAO number  $N$ . Each histogram has been obtained from a time series containing 1000 spikes. Parameter values are  $\tilde{\mu} = 0.05$ ,  $\tilde{\sigma} = 0.06$  and  $\tilde{\mu} = 0$ ,  $\tilde{\sigma} = 0.05$  (cf. (2.14) for their definition).

### 2.3 General results on the distribution of SAOs

Our first main result gives qualitative properties of the distribution of  $N$  valid in all parameter regimes.

**Theorem 2.2** (General properties of  $N$ ). *Assume that  $\sigma_1, \sigma_2 > 0$ . Then for any initial distribution  $\mu_0$  of  $R_0$  on the curve  $\mathcal{F}$ ,*

- *the kernel  $K$  admits a quasi-stationary distribution  $\pi_0$ ;*
- *the associated principal eigenvalue  $\lambda_0 = \lambda_0(\varepsilon, \delta, \sigma_1, \sigma_2)$  is strictly smaller than 1;*
- *the random variable  $N$  is almost surely finite;*
- *the distribution of  $N$  is “asymptotically geometric”, that is,*

$$\lim_{n \rightarrow \infty} \mathbb{P}^{\mu_0} \{N = n + 1 \mid N > n\} = 1 - \lambda_0 ; \quad (2.6)$$

- $\mathbb{E}^{\mu_0} \{r^N\} < \infty$  for  $r < 1/\lambda_0$  and thus all moments  $\mathbb{E}^{\mu_0} \{N^k\}$  of  $N$  are finite.

The proof is given in Section 3.3. It relies on the fact that we are able to show uniform positivity of the kernel  $K$ , which implies both the existence of the QSD  $\pi_0$  and the exponentially fast convergence to it.

Note that in the particular case where the initial distribution  $\mu_0$  is equal to the QSD  $\pi_0$ , the random variable  $R_n$  has the law  $\mu_n = \lambda_0^n \pi_0$ , and  $N$  follows an exponential law of parameter  $1 - \lambda_0$  :

$$\mathbb{P}^{\pi_0} \{N = n\} = \lambda_0^{n-1} (1 - \lambda_0) \quad \text{and} \quad \mathbb{E}^{\pi_0} \{N\} = \frac{1}{1 - \lambda_0} . \quad (2.7)$$

In general, however, the initial distribution  $\mu_0$  after a spike will be far from the QSD  $\pi_0$ , and thus the distribution of  $N$  will only be asymptotically geometric.

Theorem 2.2 allows to quantify the clusters of spikes observed in [MVE08]. To this end, we have to agree on a definition of clusters of spikes. One may decide that a cluster is a sequence of successive spikes between which there is no complete SAO, i.e.  $N = 1$  between consecutive spikes. If the time resolution is not very good, however, one may also fix a threshold SAO number  $n_0 \geq 1$ , and consider as a cluster a succession of spikes

separated by at most  $n_0$  SAOs. Let  $\mu_0^{(n)}$  be the arrival distribution on  $\mathcal{F}$  of sample paths after the  $n$ th spike. Then the probability to observe a cluster of length  $k$  is given by

$$\mathbb{P}^{\mu_0^{(0)}} \{N \leq n_0\} \mathbb{P}^{\mu_0^{(1)}} \{N \leq n_0\} \dots \mathbb{P}^{\mu_0^{(k-1)}} \{N \leq n_0\} \mathbb{P}^{\mu_0^{(k)}} \{N > n_0\} . \quad (2.8)$$

In general, the consecutive spikes will not be independent, and thus the distributions  $\mu_0^{(n)}$  will be different. For small noise, however, after a spike sample paths strongly concentrate near the stable branch of the nullcline (see the discussion in [BG09, Section 3.5.2]), and thus we expect all  $\mu_0^{(n)}$  to be very close to some constant distribution  $\mu_0$ . This implies that the lengths of clusters of spikes also follow an approximately geometric distribution :

$$\mathbb{P}\{\text{cluster of length } k\} \simeq p^k(1-p) \quad \text{where } p = \mathbb{P}^{\mu_0} \{N \leq n_0\} . \quad (2.9)$$

Figure 5 shows numerically simulated distributions of the SAO number. The geometric decay is clearly visible. Of course, the asymptotic relation (2.6) cannot be verified due to the finite sample size.

## 2.4 The weak-noise regime

In order to obtain more quantitative results, we start by transforming the FitzHugh-Nagumo equations to a more suitable form. The important part of dynamics occurs near the singular Hopf bifurcation point. We carry out the transformation in three steps, the first two of which have already been used in [BE86, BE92] :

1. An affine transformation  $x = (1-u)/\sqrt{3}$ ,  $y = (v-2/3)/\sqrt{3}$  translates the origin to the bifurcation point, and yields, in the deterministic case (2.1), the system

$$\begin{aligned} \varepsilon \dot{u} &= -v + u^2 - \frac{1}{3}u^3 , \\ \dot{v} &= \tilde{\delta} + u , \end{aligned} \quad (2.10)$$

where  $\tilde{\delta} = \sqrt{3}a - 1$  is small and positive. Note that the parameter  $\delta$  determining the stability of the stationary point  $P$  satisfies  $\delta = \tilde{\delta}(1 + \sqrt{3}a)/2 = \tilde{\delta} + \mathcal{O}(\tilde{\delta}^2)$ , so that we may use  $\delta$  and  $\tilde{\delta}$  interchangeably.

2. The scaling of space and time given by  $u = \sqrt{\varepsilon}\xi$ ,  $v = \varepsilon\eta$  and  $t = \sqrt{\varepsilon}t'$  yields

$$\begin{aligned} \dot{\xi} &= -\eta + \xi^2 - \frac{\sqrt{\varepsilon}}{3}\xi^3 , \\ \dot{\eta} &= \mu + \xi , \end{aligned} \quad (2.11)$$

where dots now indicate derivation with respect to  $t'$ , and  $\mu = \tilde{\delta}/\sqrt{\varepsilon}$ . On this scale, the nullcline  $\dot{\xi} = 0$  is close to the parabola  $\eta = \xi^2$ .

3. The nonlinear transformation  $\eta = \xi^2 + z - 1/2$  has the effect of straightening out the nullcline, and transforms (2.11) into

$$\begin{aligned} \dot{\xi} &= \frac{1}{2} - z - \frac{\sqrt{\varepsilon}}{3}\xi^3 , \\ \dot{z} &= \mu + 2\xi z + \frac{2\sqrt{\varepsilon}}{3}\xi^4 . \end{aligned} \quad (2.12)$$

Note that the stationary point now has coordinates  $(-\mu, 1/2)$ . If  $\mu = \varepsilon = 0$ , the line  $z = 0$  is invariant. Orbits starting in  $\{z < 0\}$  go to  $-\infty$ , while orbits starting in  $\{z > 0\}$  rotate around the stationary point. Thus  $\{z = 0\}$  is indeed the separatrix in that limit.

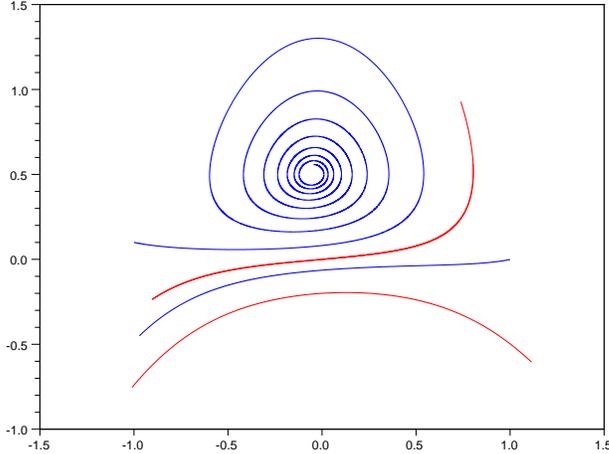


FIGURE 6. Some orbits of the deterministic equations (2.12) in  $(\xi, z)$ -coordinates.

Carrying out the same transformations for the stochastic system (2.2) yields the following result (we omit the proof, which is a straightforward application of Itô's formula).

**Proposition 2.3.** *In the new variables  $(\xi, z)$ , and on the new timescale  $t/\sqrt{\varepsilon}$ , the stochastic FitzHugh–Nagumo equations (2.2) take the form*

$$\begin{aligned} d\xi_t &= \left( \frac{1}{2} - z_t - \frac{\sqrt{\varepsilon}}{3} \xi_t^3 \right) dt + \tilde{\sigma}_1 dW_t^{(1)}, \\ dz_t &= \left( \tilde{\mu} + 2\xi_t z_t + \frac{2\sqrt{\varepsilon}}{3} \xi_t^4 \right) dt - 2\tilde{\sigma}_1 \xi_t dW_t^{(1)} + \tilde{\sigma}_2 dW_t^{(2)}, \end{aligned} \quad (2.13)$$

where

$$\begin{aligned} \tilde{\sigma}_1 &= -\sqrt{3}\varepsilon^{-3/4}\sigma_1, \\ \tilde{\sigma}_2 &= \sqrt{3}\varepsilon^{-3/4}\sigma_2, \\ \tilde{\mu} &= \mu - \tilde{\sigma}_1^2 = \frac{\tilde{\delta}}{\sqrt{\varepsilon}} - \tilde{\sigma}_1^2. \end{aligned} \quad (2.14)$$

When  $z$  is close to 0, the dynamics of  $z_t$  is dominated by two terms : the term  $\tilde{\mu} dt$ , which pushes sample paths upwards to the region of SAOs, and the noise terms. We can thus expect that if  $\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2 \ll \tilde{\mu}^2$ , then the upwards drift dominates, and the system will make many SAOs between two consecutive spikes. Going back to original parameters, the condition translates into  $\sigma_1^2 + \sigma_2^2 \ll (\varepsilon^{1/4}\delta)^2$ .

Making these ideas rigorous requires a precise control of the dynamics for small  $z$ . We carry out this analysis in Section 4. In order to analyse the SAO itself, we need to introduce another set of coordinates, based on the fact that the deterministic system (2.12) is integrable for  $\mu = \varepsilon = 0$ . We describe the dynamics in these coordinates in Section 5. As a result, we obtain the following theorem, which is proved in Section 5.4.

**Theorem 2.4** (Weak-noise regime). *Assume that  $\varepsilon$  and  $\delta/\sqrt{\varepsilon}$  are sufficiently small. Then there exists a constant  $\kappa > 0$  such that for  $\sigma_1^2 + \sigma_2^2 \leq (\varepsilon^{1/4}\delta)^2 / \log(\sqrt{\varepsilon}/\delta)$ , the principal eigenvalue  $\lambda_0$  satisfies*

$$1 - \lambda_0 \leq \exp \left\{ -\kappa \frac{(\varepsilon^{1/4}\delta)^2}{\sigma_1^2 + \sigma_2^2} \right\}. \quad (2.15)$$

Furthermore, for any initial distribution  $\mu_0$  of incoming sample paths, the expected number of SAOs satisfies

$$\mathbb{E}^{\mu_0}\{N\} \geq C(\mu_0) \exp\left\{\kappa \frac{(\varepsilon^{1/4}\delta)^2}{\sigma_1^2 + \sigma_2^2}\right\}. \quad (2.16)$$

Here  $C(\mu_0)$  is the probability that the incoming path hits  $\mathcal{F}$  above the separatrix.

Relation (2.16) shows that the average number of SAOs between two consecutive spikes is exponentially large in this regime. Note that each SAO requires a rescaled time of order 1 (see Section 5.1), and thus a time of order  $\sqrt{\varepsilon}$  in original units. It follows that the average interspike interval length is obtained by multiplying (2.16) by a constant times  $\sqrt{\varepsilon}$ .

Relation (2.6) shows that the distribution of  $N$  is asymptotically geometric with parameter given by (2.15). Hence the interspike interval distribution will be close to an exponential one, but with a periodic modulation due to the SAOs.

## 2.5 The transition from weak to strong noise

We now give an approximate description of how the dynamics changes with increasing noise intensity. Assume that  $\xi_0 = -L$  for some  $L > 0$  and that  $z_0$  is small. As long as  $z_t$  remains small, we may approximate  $\xi_t$  by  $\xi_0 + t/2$ , and thus  $z_t$  will be close to the solution of

$$dz_t^1 = \left(\tilde{\mu} + tz_t^1\right) dt - 2\tilde{\sigma}_1 t dW_t^{(1)} + \tilde{\sigma}_2 dW_t^{(2)}. \quad (2.17)$$

This linear equation can be solved explicitly. In particular, at time  $T = 4L$ ,  $\xi_t$  is close to  $L$  and we have the following result (see Section 4.1 for the proof).

**Proposition 2.5.** *Let  $2L^2 = \gamma|\log(c_- \tilde{\mu})|$  for some  $\gamma, c_- > 0$ . Then for any  $H$ ,*

$$\mathbb{P}\{z_T^1 \leq -H\} = \Phi\left(-\pi^{1/4} \frac{\tilde{\mu}}{\tilde{\sigma}} \left[1 + \mathcal{O}\left((H + z_0)\tilde{\mu}^{\gamma-1}\right)\right]\right), \quad (2.18)$$

where  $\tilde{\sigma}^2 = \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2$  and  $\Phi(x) = \int_{-\infty}^x e^{-u^2/2} du / \sqrt{2\pi}$  is the distribution function of the standard normal law.

Choosing  $\gamma$  large enough, the right-hand side is approximately constant for a large range of values of  $z_0$  and  $H$ . Since  $1 - \lambda_0$  is equal to the probability of leaving  $\mathcal{D}$  before completing the first SAO, when starting in the QSD  $\pi_0$ , we expect that

$$1 - \lambda_0 \simeq \Phi\left(-\pi^{1/4} \frac{\tilde{\mu}}{\tilde{\sigma}}\right) = \Phi\left(-\frac{(\pi\varepsilon)^{1/4}(\delta - \sigma_1^2/\varepsilon)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right), \quad (2.19)$$

provided  $\pi_0$  is concentrated near  $z = 0$ . We can identify three regimes, depending on the value of  $\tilde{\mu}/\tilde{\sigma}$ :

1. **Weak noise :**  $\tilde{\mu} \gg \tilde{\sigma}$ , which in original variables translates into  $\sqrt{\sigma_1^2 + \sigma_2^2} \ll \varepsilon^{1/4}\delta$ . This is the weak-noise regime already studied in the previous subsection, in which  $\lambda_0$  is exponentially close to 1, and thus spikes are separated by long sequences of SAOs.
2. **Strong noise :**  $\tilde{\mu} \ll -\tilde{\sigma}$ , which implies  $\mu \ll \tilde{\sigma}^2$ , and in original variables translates into  $\sqrt{\sigma_1^2 + \sigma_2^2} \gg \varepsilon^{3/4}$ . Then  $\lambda_0$  is exponentially small, of order  $e^{-(\sigma_1^2 + \sigma_2^2)/\varepsilon^{3/2}}$ . Thus with high probability, there will be no complete SAO between consecutive spikes, i.e., the neuron is spiking repeatedly.

3. **Intermediate noise** :  $|\tilde{\mu}| = \mathcal{O}(\tilde{\sigma})$ , which translates into  $\varepsilon^{1/4}\delta \leq \sqrt{\sigma_1^2 + \sigma_2^2} \leq \varepsilon^{3/4}$ . Then the mean number of SAOs is of order 1. In particular, when  $\sigma_1 = \sqrt{\varepsilon\delta}$ ,  $\tilde{\mu} = 0$  and thus  $\lambda_0$  is close to  $1/2$ .

An interesting point is that the transition from weak to strong noise is gradual, being characterised by a smooth change of the distribution of  $N$  as a function of the parameters. There is no clear-cut transition at  $\sigma_1 = \sqrt{\varepsilon\delta}$ , the only particularity of this parameter value being that  $\lambda_0$  is close to  $1/2$ .

### 3 Substochastic Markov kernels

In this section, we discuss properties of substochastic Markov chains with continuous state space, which will allow us to characterise the random number of SAOs. For the general theory of such processes, see [Ore71, Num84].

#### 3.1 General results

Let  $E$  be a bounded interval in  $\mathbb{R}$  and let  $\mathcal{E}$  be the Borel  $\sigma$ -algebra on  $E$ . Let  $K : E \times \mathcal{E} \rightarrow \mathbb{R}_+$  be a substochastic Markov kernel from  $(E, \mathcal{E})$  to itself, that is,

- for all  $x \in E$ ,  $K(x, \cdot)$  is a measure on  $(E, \mathcal{E})$ ;
- for all  $A \in \mathcal{E}$ ,  $K(\cdot, A)$  is an  $\mathcal{E}$ -measurable map;
- $K$  is *substochastic*, that is,  $K(x, E) \leq 1$  for all  $x \in E$ .

The kernel can be made stochastic by adding a cemetery state  $\Delta$  to  $E$ , and setting  $K(x, \Delta) = 1 - K(x, E)$ ,  $K(\Delta, E) = 0$ . This defines a Markov chain  $(X_n)_{n \geq 0}$  on  $\bar{E} = E \cup \{\Delta\}$  via the relation

$$\mathbb{P}\{X_{n+1} \in A | X_n = x\} = K(x, A) \quad \forall n \geq 0. \quad (3.1)$$

We consider  $K$  as a bounded linear operator on  $L^\infty(E)$ , acting on bounded measurable functions by

$$f(x) \mapsto (Kf)(x) = \int_E K(x, dy)f(y) = \mathbb{E}^x \{f(X_1)\}, \quad (3.2)$$

and as a bounded linear operator on  $L^1(E)$ , acting on finite measures  $\mu$  by

$$\mu(A) \mapsto (\mu K)(A) = \int_E \mu(dx)K(x, A) = \mathbb{P}^\mu \{X_1 \in A\}. \quad (3.3)$$

As the notations suggest, we think of measures as row vectors, and of functions of column vectors. We are interested in properties of the absorption time

$$N = \inf\{n \geq 1 : X_n = \Delta\} \in \mathbb{N} \cup \{\infty\}. \quad (3.4)$$

We begin by deriving an expression of the moment generating function of  $N$ . Let  $\mathbb{1}$  be the identity kernel defined by  $\mathbb{1}(x, A) = 1_{\{x \in A\}}$ . Define  $n$ -fold iterated kernels  $K^n$  by  $K^0 = \mathbb{1}$ ,  $K^1 = K$  and

$$K^{n+1}(x, A) = \int_E K^n(x, dy)K(y, A) \quad \forall n \geq 1. \quad (3.5)$$

Then  $K^n(x, \cdot)$  is the law of  $X_n$ , given that  $X_0 = x$ , and thus

$$\mathbb{P}^x \{N > n\} = \mathbb{P}^x \{X_n \in E\} = K^n(x, E). \quad (3.6)$$

More generally, if  $X_0$  is distributed according to some probability measure  $\mu_0$ , then

$$\mathbb{P}^{\mu_0}\{N > n\} = \mathbb{P}^{\mu_0}\{X_n \in E\} = \mu_0 K^n \mathbf{1} := \int_E \int_E \mu_0(dx) K^n(x, dy). \quad (3.7)$$

Here  $\mathbf{1}$  denotes the function which is identically equal to 1.

For  $r \in \mathbb{C}$ , the *potential kernel* associated with  $K$  is defined by

$$G^{(r)} = \sum_{n=0}^{\infty} r^n K^n = \mathbf{1} + rK + r^2 K^2 + \dots \quad (3.8)$$

It is known [Num84, Theorem 3.2] that there exists a *convergence parameter*  $R \in [0, \infty)$  such that  $G^{(r)}f = +\infty$  for all positive  $f$  if  $r > R$ , and  $G^{(r)}f < \infty$  on some set and for a class of  $f$  if  $r < R$ . Note that

$$G^{(r)} = \mathbf{1} + rKG^{(r)} = \mathbf{1} + G^{(r)}rK. \quad (3.9)$$

**Lemma 3.1.** *Assume  $N$  is almost surely finite. Then the moment generating function of  $N$  is given by*

$$\mathbb{E}^{\mu_0}\{r^N\} = 1 + (r-1)\mu_0 G^{(r)}\mathbf{1}. \quad (3.10)$$

PROOF: We have

$$\begin{aligned} \sum_{n \geq 0} r^n \mathbb{P}^{\mu_0}\{N > n\} &= \sum_{n \geq 0} r^n \sum_{k \geq n+1} \mathbb{P}^{\mu_0}\{N = k\} = \sum_{k \geq 1} \sum_{n=0}^{k-1} r^n \mathbb{P}^{\mu_0}\{N = k\} \\ &= \sum_{k \geq 1} \frac{1-r^k}{1-r} \mathbb{P}^{\mu_0}\{N = k\} = \frac{1}{1-r} [1 - \mathbb{E}^{\mu_0}\{r^N\}]. \end{aligned} \quad (3.11)$$

The result then follows from the fact that  $r^n \mathbb{P}^{\mu_0}\{N > n\} = \mu_0(r^n K^n)\mathbf{1}$ .  $\square$

If the convergence parameter  $R$  is strictly larger than 1, we can differentiate (3.10) with respect to  $r$  and evaluate the result at  $r = 1$ , yielding

$$\mathbb{E}^{\mu_0}\{N\} = \mu_0 G^{(1)}\mathbf{1}. \quad (3.12)$$

Higher moments of  $N$  can be determined by computing higher derivatives of the potential kernel in  $r = 1$ .

Next we want to show that the distribution of  $N$  has an asymptotically geometric decay. This is true if  $K$  has an isolated simple maximal eigenvalue  $\lambda_0$ , separated from the rest of its spectrum by a positive gap. Such a result was first proved by Frobenius and Perron for stochastic matrices, and later by Jentzsch for positive integral operators [Jen12], and by Krein and Rutman for more general linear operators [KR50].

In our case, we can work in the framework of the theory of positive operators on vector lattices [Bir57]. Both  $L^\infty(E)$  and  $L^1(E)$  are particular cases of Banach lattices (Banach spaces in which any two elements  $f, g$  admit a minimum  $f \wedge g$  and a maximum  $f \vee g$ ).

If  $s : E \rightarrow \mathbb{R}$  is a measurable function and  $\nu$  is a measure on  $(E, \mathcal{E})$  we denote by  $s \otimes \nu$  the product kernel defined by

$$(s \otimes \nu)(x, A) = s(x)\nu(A). \quad (3.13)$$

The kernel  $K$  is called *uniformly positive* if there exists a positive  $s$ , a finite measure  $\nu$  and a constant  $L$  such that

$$s \otimes \nu \leq K \leq L s \otimes \nu. \quad (3.14)$$

**Theorem 3.2** ([Bir57, Theorem 3 and Lemma 3]). *Any uniformly positive linear operator  $K$  on a Banach lattice  $L$  admits a unique positive unit eigenvector  $h$ . The corresponding eigenvalue  $\lambda_0$  is real and positive. Furthermore, for any  $f \in L$ , there exist constants  $M_1(f)$  and  $M_2(f)$  such that*

$$|K^n f - M_1(f)\lambda_0^n h| \leq M_2(f)(\lambda_0 \rho)^n h, \quad (3.15)$$

where  $\rho < 1$ .

Applying this theorem to  $f \mapsto Kf$  and to  $\mu \mapsto \mu K$ , we obtain the existence of an eigenvalue  $\lambda_0$ , and of left and right eigenvectors  $\pi_0 > 0$  and  $h_0 > 0$ , i.e., such that  $Kh_0 = \lambda_0 h_0$  and  $\pi_0 K = \lambda_0 \pi_0$ . We normalise the eigenvectors in such a way that

$$\pi_0(E) = \int_E \pi_0(dx) = 1, \quad \pi_0 h_0 = \int_E \pi_0(dx) h_0(x) = 1. \quad (3.16)$$

We call  $\pi_0$  the *quasistationary distribution* of the Markov chain (see e.g. [SVJ66]). Notice that

$$\lambda_0 = \lambda_0 \pi_0(E) = \int_E \pi_0(dx) K(x, E) \leq \pi_0(E) = 1, \quad (3.17)$$

with equality holding if and only if  $K(x, E) = 1$  for  $\pi_0$ -almost all  $x \in E$ .

**Remark 3.3.**

- To check that both eigenvectors correspond to the same eigenvalue, assume  $\pi_0 K = \lambda'_0 \pi_0$ . Then  $\lambda'_0 \pi_0 h_0 = \pi_0 K h_0 = \lambda_0 \pi_0 h_0$ , and thus  $\lambda'_0 = \lambda_0$  since  $\pi_0 h_0 > 0$ .
- In the case of the map  $f \mapsto Kf$ ,  $M_1(f) = \pi_0 f$ . Indeed,  $M_1(f)$  is a positive linear functional [Bir57, p. 225] and thus of the form  $M_1(f) = \mu f$  for some measure  $\mu$ . Relation (3.15) implies that  $M_1(Kf) = \lambda_0 M_1(f)$ , and thus  $\mu$  has to be a multiple of  $\pi_0$ . Since  $M_1(h_0) = h_0$ , in fact  $\mu = \pi_0$ . For the map  $\mu \mapsto \mu K$ , a similar argument yields  $M_1(\mu) = \mu h_0$ .
- We have  $G^{(r)} h_0 = (1 - r\lambda_0)^{-1} h_0$  for  $r < 1/\lambda_0$ , and thus the convergence parameter  $R$  satisfies  $R \geq 1/\lambda_0$ .

**Proposition 3.4.** *Assume that  $K$  is uniformly positive and that the principal eigenvalue  $\lambda_0$  satisfies  $\lambda_0 < 1$ . Then*

1.  $N$  is almost surely finite, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{P}^{\mu_0} \{N > n\} = 0. \quad (3.18)$$

2. The absorption time  $N$  is asymptotically geometric, that is,

$$\lim_{n \rightarrow \infty} \mathbb{P}^{\mu_0} \{N = n + 1 | N > n\} = 1 - \lambda_0. \quad (3.19)$$

3. The moment-generating function  $\mathbb{E}^{\mu_0} \{r^N\}$  is finite for any  $r < 1/\lambda_0$ .

PROOF: To prove (3.19), we start by writing

$$\mathbb{P}^{\mu_0} \{N = n + 1 | N > n\} = \frac{\mathbb{P}^{\mu_0} \{N = n + 1\}}{\mathbb{P}^{\mu_0} \{N > n\}} = \frac{\mu_0 K^n k_\Delta}{\mu_0 K^n \mathbf{1}}, \quad (3.20)$$

where  $k_\Delta(x) = K(x, \Delta) = 1 - K(x, E)$ , i.e.,  $k_\Delta = \mathbf{1} - K\mathbf{1}$ . Now (3.15) yields

$$M_1(\mathbf{1})\lambda_0^n h_0 - M_2(\mathbf{1})(\lambda_0 \rho)^n h_0 \leq K^n \mathbf{1} \leq M_1(\mathbf{1})\lambda_0^n h_0 + M_2(\mathbf{1})(\lambda_0 \rho)^n h_0. \quad (3.21)$$

Since  $M_1(\mathbf{1}) = \pi_0 \mathbf{1} = 1$  by the normalisation (3.16), dividing by  $\lambda_0^n$  and multiplying on the left by  $\mu_0$  we get

$$\mu_0 h_0 [1 - M_2(\mathbf{1})\rho^n] \leq \frac{\mu_0 K^n \mathbf{1}}{\lambda_0^n} \leq \mu_0 h_0 [1 + M_2(\mathbf{1})\rho^n]. \quad (3.22)$$

Similarly, we have

$$\mu_0 h_0 [M_1(k_\Delta) - M_2(k_\Delta)\rho^n] \leq \frac{\mu_0 K^n k_\Delta}{\lambda_0^n} \leq \mu_0 h_0 [M_1(k_\Delta) + M_2(k_\Delta)\rho^n]. \quad (3.23)$$

Since  $M_1(k_\Delta) = \pi_0 k_\Delta = \pi_0(\mathbf{1} - K\mathbf{1}) = 1 - \lambda_0 \pi_0 \mathbf{1} = 1 - \lambda_0$ , (3.19) follows by dividing (3.23) by (3.22) and taking the limit  $n \rightarrow \infty$ .

To prove (3.18), simply recall that  $\mathbb{P}^{\mu_0}\{N > n\} = \mu_0 K^n \mathbf{1}$ , which converges to 0 as  $n \rightarrow \infty$  by (3.22). The same equation implies that  $\mathbb{E}^{\mu_0}\{r^N\} = \mu_0 G^{(r)} \mathbf{1}$  can be bounded by a geometric series which converges for  $r < 1/\lambda_0$ .  $\square$

### 3.2 Bounds

In this section we provide some simple bounds on  $\lambda_0$  and  $\mathbb{E}^{\mu_0}\{N\}$ , in the case where there exists a subset  $A \subset E$  that the Markov chain is unlikely to leave. We assume in this section that  $K$  satisfies the uniform positivity condition (3.14). Let

$$\varepsilon_A = \sup_{x \in A} [1 - K(x, A)] \quad (3.24)$$

be the maximal probability to leave  $A$  when starting in  $A$ .

**Lemma 3.5.** *For any  $A \subset E$  with  $\nu(A) > 0$ ,*

$$\lambda_0 \geq 1 - \varepsilon_A. \quad (3.25)$$

PROOF: The relation  $\lambda_0 \pi_0 = \pi_0 K$  yields

$$\begin{aligned} \lambda_0 \pi_0(A) &= \int_A \pi_0(dx) K(x, A) + \int_{E \setminus A} \pi_0(dx) K(x, A) \\ &\geq \pi_0(A)(1 - \varepsilon_A) + \int_{E \setminus A} \pi_0(dx) s(x) \nu(A). \end{aligned} \quad (3.26)$$

Either  $\pi_0(A) = 1$ , and the result follows immediately. Or  $\pi_0(A) < 1$ , and thus  $\pi_0(E \setminus A) > 0$ , so that the second term on the right-hand side is strictly positive. It follows that  $\lambda_0 \pi_0(A) > 0$ , and we obtain (3.25) upon dividing by  $\pi_0(A)$ .  $\square$

**Lemma 3.6.** *For any  $A \subset E$  such that  $\varepsilon_A > 0$ ,*

$$\mathbb{E}^{\mu_0}\{N\} \geq \frac{\mu_0(A)}{\varepsilon_A}. \quad (3.27)$$

PROOF: For  $x \in A$ , let  $\theta(x) = \mathbb{E}^x\{N\} = G^{(1)}(x, E)$ . Then  $\theta(x) = \lim_{n \rightarrow \infty} \theta_n(x)$  where

$$\theta_n(x) = \sum_{m=0}^n K^m(x, E). \quad (3.28)$$

We have

$$\theta_{n+1}(x) = 1 + (K\theta_n)(x) \geq 1 + \int_A K(x, dy)\theta_n(y) . \quad (3.29)$$

Now let  $m_n = \inf_{x \in A} \theta_n(x)$ . Then  $m_0 = 1$  and

$$m_{n+1} \geq 1 + (1 - \varepsilon_A)m_n . \quad (3.30)$$

By induction on  $n$  we get

$$m_n \geq \frac{1}{\varepsilon_A} - \frac{(1 - \varepsilon_A)^{n+1}}{\varepsilon_A} , \quad (3.31)$$

so that  $\mathbb{E}^x \{N\} = \theta(x) \geq 1/\varepsilon_A$  for all  $x \in A$ . The result follows upon integrating against  $\mu_0$  over  $A$ .  $\square$

### 3.3 Proof of Theorem 2.2

Let us denote by  $K(x, dy)$  the kernel defined in (2.4). Note that  $y$  represents the first-exit location from the domain  $\mathcal{G} = (0, 1) \times (-2M\pi, 2\pi)$ , for an initial condition  $(x, 0)$ , in case the exit occurs through one of the lines  $\varphi = -2M\pi$  or  $\varphi = 2\pi$ . In harmonic analysis,  $K(x, dy)$  is called the harmonic measure for the generator of the diffusion in  $\mathcal{G}$  based at  $(x, 0)$ . In the case of Brownian motion, it has been proved in [Dah77] that sets of positive Hausdorff measure have positive harmonic measure. This result has been substantially extended in [BAKS84], where the authors prove that for a general class of hypoelliptic diffusions, the harmonic measure admits a smooth density  $k(x, y)$  with respect to Lebesgue measure  $dy$ . Our diffusion process being uniformly elliptic for  $\sigma_1, \sigma_2 > 0$ , it enters into the class of processes studied in that work.

Specifically, [BAKS84, Corollary 2.11] shows that  $k(x, y)$  is smooth, and its derivatives are bounded by a function of the distance from  $x$  to  $y$ . This distance being uniformly bounded below by a positive constant in our setting, there exists a constant  $L \in \mathbb{R}_+$  such that

$$\frac{\sup_{y \in E} k(x, y)}{\inf_{y \in E} k(x, y)} \leq L \quad \forall x \in E . \quad (3.32)$$

We set

$$s(x) = \inf_{y \in E} k(x, y) . \quad (3.33)$$

Then it follows that

$$s(x) \leq k(x, y) \leq Ls(x) \quad \forall x, y \in E . \quad (3.34)$$

Thus the kernel  $K$  fulfils the uniform positivity Condition (3.16) for  $\nu$  given by the Lebesgue measure.

With this property established, the remainder of the proof follows from results stated in Section 3.1. The existence of a QSD follows from Theorem 3.2. The fact that  $\lambda_0 < 1$  is a consequence of (3.17) and the fact that  $K(x, E) < 1$  due to ellipticity. The remaining claims follow from Proposition 3.4.  $\square$

## 4 Dynamics near the separatrix

In this section, we use the equations in  $(\xi, z)$ -variables given by

$$\begin{aligned} d\xi_t &= \left( \frac{1}{2} - z_t - \frac{\sqrt{\varepsilon}}{3} \xi_t^3 \right) dt + \tilde{\sigma}_1 dW_t^{(1)}, \\ dz_t &= \left( \tilde{\mu} + 2\xi_t z_t + \frac{2\sqrt{\varepsilon}}{3} \xi_t^4 \right) dt - 2\tilde{\sigma}_1 \xi_t dW_t^{(1)} + \tilde{\sigma}_2 dW_t^{(2)}, \end{aligned} \quad (4.1)$$

to describe the dynamics in a neighbourhood of the separatrix. To be more specific, we will assume that  $z$  is small, of the order of some power of  $\mu$ , and that  $\xi$  varies in an interval  $[-L, L]$ . The parameter  $L$  is given by

$$L^2 = \frac{\gamma}{2} (-\log(c_- \tilde{\mu})), \quad (4.2)$$

where  $c_- > 0$  and  $\gamma > 0$  are constants that will be chosen below. We define two broken lines

$$F_- = \left\{ \xi = -L \text{ and } z \leq \frac{1}{2} \right\} \cup \left\{ -L \leq \xi \leq 0 \text{ and } z = \frac{1}{2} \right\} \quad (4.3)$$

and

$$F_+ = \left\{ \xi = L \text{ and } z \leq \frac{1}{2} \right\} \cup \left\{ 0 \leq \xi \leq L \text{ and } z = \frac{1}{2} \right\}. \quad (4.4)$$

Given an initial condition  $(-L, z_0) \in F_-$ , our goal is to estimate where the sample path starting in  $(-L, z_0)$  hits  $F_+$  for the first time. This will yield the first part of the Markov kernel  $K$ .

### 4.1 The linearised process

Before analysing the full dynamics of (4.1) we consider some approximations of the system. The fact that  $\xi_t \simeq \xi_0 + t/2$  for small  $z$  motivates the change of variable

$$\xi = \frac{t}{2} + u, \quad (4.5)$$

which transforms the system (4.1) into

$$\begin{aligned} du_t &= (-z_t + \mathcal{O}(\tilde{\varepsilon})) dt + \tilde{\sigma}_1 dW_t^{(1)}, \\ dz_t &= (\tilde{\mu} + tz_t + 2u_t z_t + \mathcal{O}(\tilde{\varepsilon})) dt - 2\tilde{\sigma}_1 t dW_t^{(1)} - \tilde{\sigma}_1 u_t dW_t^{(1)} + \tilde{\sigma}_2 dW_t^{(2)}, \end{aligned} \quad (4.6)$$

where we write  $\tilde{\varepsilon} = \sqrt{\varepsilon} L^4$ . We choose an initial condition  $(0, z_0)$  at time  $t_0 = -2L$ . As a first approximation, consider the deterministic system

$$\begin{aligned} du_t^0 &= -z_t^0 dt, \\ dz_t^0 &= (\tilde{\mu} + tz_t^0) dt. \end{aligned} \quad (4.7)$$

The solution of the second equation is given by

$$z_t^0 = e^{t^2/2} \left[ z_0 e^{-t_0^2/2} + \tilde{\mu} \int_{t_0}^t e^{-s^2/2} ds \right]. \quad (4.8)$$

In particular, at time  $T = 2L$ , we have  $\xi_T = L + u_t \simeq L$  and the location of the first-hitting point of  $F_+$  is approximated by

$$z_T^0 = z_0 + \tilde{\mu} e^{T^2/2} \int_{t_0}^T e^{-s^2/2} ds = z_0 + \mathcal{O}(\tilde{\mu}^{1-\gamma}). \quad (4.9)$$

As a second approximation, we incorporate the noise terms and consider the linear SDE

$$\begin{aligned} du_t^1 &= -z_t^1 dt + \tilde{\sigma}_1 dW_t^{(1)}, \\ dz_t^1 &= (\tilde{\mu} + tz_t^1) dt - \tilde{\sigma}_1 t dW_t^{(1)} + \tilde{\sigma}_2 dW_t^{(2)}. \end{aligned} \quad (4.10)$$

**Proposition 4.1** (Proposition 2.5). *For any  $H$ ,*

$$\mathbb{P}\{z_T^1 \leq -H\} = \Phi\left(-\pi^{1/4} \frac{\tilde{\mu}}{\tilde{\sigma}} \left[1 + \mathcal{O}\left((H + z_0)\tilde{\mu}^{\gamma-1}\right)\right]\right), \quad (4.11)$$

where  $\tilde{\sigma}^2 = \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2$  and  $\Phi(x) = \int_{-\infty}^x e^{-u^2/2} du / \sqrt{2\pi}$  is the distribution function of the standard normal law.

PROOF: Solving the second equation in (4.10) by variation of the constant yields

$$z_T^1 = z_0 + e^{T^2/2} \left[ \tilde{\mu} \int_{t_0}^T e^{-s^2/2} ds - \tilde{\sigma}_1 \int_{t_0}^T s e^{-s^2/2} dW_s^{(1)} + \tilde{\sigma}_2 \int_{t_0}^T e^{-s^2/2} dW_s^{(2)} \right]. \quad (4.12)$$

Note that by the choice (4.2) of  $L$ , we have  $e^{T^2/2} = e^{2L^2} = (c_- \tilde{\mu})^{-\gamma}$ . The random variable  $z_T^1$  is Gaussian, with expectation

$$\mathbb{E}\{z_T^1\} = z_0 + \tilde{\mu} e^{2L^2} \int_{-2L}^{2L} e^{-s^2/2} ds \quad (4.13)$$

and variance

$$\text{Var}(z_T^1) = \tilde{\sigma}_1^2 e^{4L^2} \int_{-2L}^{2L} s^2 e^{-s^2} ds + \tilde{\sigma}_2^2 e^{4L^2} \int_{-2L}^{2L} e^{-s^2} ds. \quad (4.14)$$

Using this in the relation

$$\mathbb{P}\{z_T^1 \leq -H\} = \Phi\left(-\frac{H + \mathbb{E}\{z_T^1\}}{\sqrt{\text{Var}(z_T^1)}}\right) \quad (4.15)$$

yields the result. □

Let us now quantify the deviation between  $(u_t^1, z_t^1)$  and  $(u_t^0, z_t^0)$ .

**Proposition 4.2.** *Let*

$$\zeta(s) = e^{s^2} \left[ e^{-t_0^2} + \int_{t_0}^s e^{-u^2} du \right]. \quad (4.16)$$

*Then there exists a constant  $M > 0$  such that for all  $t \geq t_0$ , all  $h, h_1, h_2 > 0$  and all  $\rho \in (0, \tilde{\mu}^{2\gamma}/M)$ ,*

$$\mathbb{P}\left\{ \sup_{t_0 \leq s \leq t} \frac{|z_s^1 - z_s^0|}{\sqrt{\zeta(s)}} \geq h \right\} \leq \frac{2(t - t_0)}{\rho} \exp\left\{ -\frac{1}{8} \frac{h^2}{\tilde{\sigma}^2} (1 - M\rho \tilde{\mu}^{-2\gamma}) \right\} \quad (4.17)$$

and

$$\begin{aligned} & \mathbb{P} \left\{ \sup_{t_0 \leq s \leq t} |u_s^1 - u_s^0| \geq h_1 + h_2 \int_{t_0}^t \sqrt{\zeta(s)} ds \right\} \\ & \leq 2 \exp \left\{ -\frac{h_1^2}{2(t-t_0)\tilde{\sigma}_1^2} \right\} + \frac{2(t-t_0)}{\rho} \exp \left\{ -\frac{1}{8} \frac{h_2^2}{\tilde{\sigma}_2^2} (1 - M\rho\tilde{\mu}^{-2\gamma}) \right\}. \end{aligned} \quad (4.18)$$

PROOF: The difference  $(x^1, y^1) = (u^1 - u^0, z^1 - z^0)$  satisfies the system

$$\begin{aligned} dx_t^1 &= -y_t^1 dt + \tilde{\sigma}_1 dW_t^{(1)}, \\ dy_t^1 &= t y_t^1 dt - \tilde{\sigma}_1 t dW_t^{(1)} + \tilde{\sigma}_2 dW_t^{(2)}. \end{aligned} \quad (4.19)$$

The second equation admits the solution

$$y_t^1 = \tilde{\sigma}_2 e^{t^2/2} \int_{t_0}^t e^{-s^2/2} dW_s^{(2)} - \tilde{\sigma}_1 e^{t^2/2} \int_{t_0}^t s e^{-s^2/2} dW_s^{(1)} =: y_t^{1,1} + y_t^{1,2}. \quad (4.20)$$

We first estimate  $y_t^{1,1}$ . Let  $u_0 = t_0 < u_1 < \dots < u_K = t$  be a partition of  $[t_0, t]$ . The Bernstein-like estimate [BG02, Lemma 3.2] yields the bound

$$\mathbb{P} \left\{ \sup_{t_0 \leq s \leq t} \frac{1}{\sqrt{\zeta(s)}} \tilde{\sigma}_2 \left| \int_{t_0}^s e^{(s^2-u^2)/2} dW_u \right| \geq H_0 \right\} \leq 2 \sum_{k=1}^K P_k \quad (4.21)$$

for any  $H_0 > 0$ , where

$$P_k \leq \exp \left\{ -\frac{1}{2} \frac{H_0^2}{\tilde{\sigma}_2^2} \inf_{u_{k-1} \leq s \leq u_k} \frac{\zeta(s)}{\zeta(u_k)} e^{u_k^2 - s^2} \right\}. \quad (4.22)$$

The definition of  $\zeta(s)$  implies

$$\frac{\zeta(s)}{\zeta(u_k)} e^{u_k^2 - s^2} = 1 - \frac{1}{\zeta(u_k)} \int_s^{u_k} e^{u_k^2 - u^2} du \geq 1 - \int_s^{u_k} e^{t_0^2 - u^2} du. \quad (4.23)$$

Note that  $e^{t_0^2} = e^{4L^2} = \mathcal{O}(\tilde{\mu}^{-2\gamma})$ . For a uniform partition given by  $u_k - u_{k-1} = \rho$  with  $\rho \ll \tilde{\mu}^{2\gamma}$ , we can bound this last expression below by

$$1 - M\rho\tilde{\mu}^{-2\gamma} \quad (4.24)$$

for some constant  $M$ . This yields

$$\mathbb{P} \left\{ \sup_{0 \leq s \leq t} \frac{|y_s^{1,1}|}{\sqrt{\zeta(s)}} \geq H_0 \right\} \leq \frac{2(t-t_0)}{\rho} \exp \left\{ -\frac{1}{2} \frac{H_0^2}{\tilde{\sigma}_2^2} (1 - M\rho\tilde{\mu}^{-2\gamma}) \right\}. \quad (4.25)$$

Doing the same for  $y_s^{1,2}$  we obtain

$$\mathbb{P} \left\{ \sup_{0 \leq s \leq t} \frac{|y_s^{1,2}|}{\sqrt{\zeta(s)}} \geq H_1 \right\} \leq \frac{2(t-t_0)}{\rho} \exp \left\{ -\frac{1}{2} \frac{H_1^2}{\tilde{\sigma}_1^2} (1 - M\rho\tilde{\mu}^{-2\gamma}) \right\} \quad (4.26)$$

for any  $H_1 > 0$ . Letting  $h = H_0 + H_1$  with  $H_0 = H_1 = h/2$ , we obtain (4.17). Now we can express  $x_t^1$  in terms of  $y_t^1$  by

$$x_t^1 = - \int_{t_0}^t y_s^1 ds + \tilde{\sigma}_1 \int_{t_0}^t dW_s^{(1)}. \quad (4.27)$$

Then the Bernstein inequality

$$\mathbb{P} \left\{ \sup_{0 \leq s \leq t} \left| \tilde{\sigma}_1 \int_{t_0}^s dW_s^{(1)} \right| \geq h_1 \right\} \leq 2 \exp \left\{ -\frac{h_1^2}{2(t-t_0)\tilde{\sigma}_1^2} \right\} \quad (4.28)$$

yields (4.18).  $\square$

## 4.2 The nonlinear equation

We now turn to the analysis of the full system (4.1), or, equivalently, (4.6). Before that, we state a generalised Bernstein inequality that we will need several times in the sequel. Let  $W_t$  be an  $n$ -dimensional standard Brownian motion, and consider the martingale

$$M_t = \int_{t_0}^t g(X_s, s) dW_s = \sum_{i=1}^n \int_{t_0}^t g_i(X_t, t) dW_t^{(i)}, \quad (4.29)$$

where  $g = (g_1, \dots, g_n)$  takes values in  $\mathbb{R}^n$  and the process  $X_t$  is assumed to be adapted to the filtration generated by  $W_t$ . Then we have the following result (for the proof, see [BGK10, Lemma D.8]):

**Lemma 4.3.** *Assume that the integrand satisfies*

$$g(X_t, t)g(X_t, t)^T \leq G(t)^2 \quad (4.30)$$

*almost surely, for a deterministic function  $G(t)$ , and that the integral*

$$V(t) = \int_{t_0}^t G(s)^2 ds \quad (4.31)$$

*is finite. Then*

$$\mathbb{P}\left\{ \sup_{t_0 \leq s \leq t} M_s > x \right\} \leq e^{-x^2/2V(t)} \quad (4.32)$$

*for any  $x > 0$ .*

**Proposition 4.4.** *Assume  $z_0 = \mathcal{O}(\tilde{\mu}^{1-\gamma})$ . There exist constants  $C, \kappa, M > 0$  such that for  $t_0 \leq t \leq T + \mathcal{O}(|\log \tilde{\mu}|^{-1/2})$ , all  $\tilde{\sigma} \leq \tilde{\mu}$  and  $H > 0$ ,*

$$\mathbb{P}\left\{ \sup_{t_0 \leq s \leq t} \frac{|z_s - z_s^0|}{\sqrt{\zeta(s)}} \geq H \right\} \leq \frac{CT}{\tilde{\mu}^{2\gamma}} \left( \exp\left\{ -\kappa \frac{[H - M(T^2 \tilde{\mu}^{2-4\gamma} + T\tilde{\varepsilon} \tilde{\mu}^{-2\gamma})]^2}{\tilde{\sigma}^2} \right\} + e^{-\kappa \tilde{\mu}^2 / \tilde{\sigma}^2} \right) \quad (4.33)$$

*and for all  $H' > 0$ ,*

$$\mathbb{P}\left\{ \sup_{t_0 \leq s \leq t} |u_s - u_s^0| \geq H' \right\} \leq \frac{CT}{\tilde{\mu}^{2\gamma}} \left( \exp\left\{ -\kappa \frac{[H' - M(T^2 \tilde{\mu}^{2-4\gamma} + T\tilde{\varepsilon} \tilde{\mu}^{-2\gamma})]^2}{\tilde{\sigma}^2 \tilde{\mu}^{-2\gamma}} \right\} + e^{-\kappa \tilde{\mu}^2 / \tilde{\sigma}^2} \right). \quad (4.34)$$

PROOF: The upper bound on  $t$  implies that  $e^{t/2} = \mathcal{O}(\tilde{\mu}^{-\gamma})$ . Thus it follows from (4.7) and (4.8) that

$$z_s^0 = \mathcal{O}(\tilde{\mu}^{1-\gamma}) \quad \text{and} \quad u_s^0 = \mathcal{O}(T\tilde{\mu}^{1-\gamma}) \quad (4.35)$$

for  $t_0 \leq s \leq t$ . Given  $h, h_1, h_2 > 0$ , we introduce the stopping times

$$\begin{aligned} \tau_1 &= \inf\left\{ s \geq t_0 : |z_s^1 - z_s^0| \geq h\sqrt{\zeta(s)} \right\}, \\ \tau_2 &= \inf\left\{ s \geq t_0 : |u_s^1 - u_s^0| \geq h_1 + h_2 \int_{t_0}^s \sqrt{\zeta(s)} ds \right\}. \end{aligned} \quad (4.36)$$

The integral of  $\sqrt{\zeta(s)}$  is of order  $T\tilde{\mu}^{-\gamma}$  at most. Thus choosing  $h = h_1 = h_2 = \tilde{\mu}$  guarantees that

$$z_s^1 = \mathcal{O}(\tilde{\mu}^{1-\gamma}) \quad \text{and} \quad u_s^1 = \mathcal{O}(T\tilde{\mu}^{1-\gamma}) \quad (4.37)$$

for  $t_0 \leq s \leq t \wedge \tau_1 \wedge \tau_2$ . For these values of  $h$ ,  $h_1$  and  $h_2$ , Proposition 4.2 implies that

$$\begin{aligned}\mathbb{P}\{\tau_1 < t\} &\leq cT\tilde{\mu}^{-2\gamma} e^{-\kappa\tilde{\mu}^2/\tilde{\sigma}^2}, \\ \mathbb{P}\{\tau_2 < t\} &\leq cT\tilde{\mu}^{-2\gamma} e^{-\kappa\tilde{\mu}^2/\tilde{\sigma}^2}\end{aligned}\quad (4.38)$$

for some constants  $\kappa, c > 0$ . We consider the difference  $(x_t^2, y_t^2) = (u_t, z_t) - (u_t^1, z_t^1)$ , which satisfies the system of SDEs

$$\begin{aligned}dx_t^2 &= (-y_t^2 + \mathcal{O}(\tilde{\varepsilon})) dt, \\ dy_t^2 &= [ty_t^2 + 2(u_t^1 + x_t^2)(z_t^1 + y_t^2) + \mathcal{O}(\tilde{\varepsilon})] dt - 2\tilde{\sigma}_1(u_t^1 + x_t^2) dW_t^{(1)}.\end{aligned}\quad (4.39)$$

We introduce a Lyapunov function  $U_t > 0$  defined by

$$(U_t - C_0)^2 = \frac{(x_t^2)^2 + (y_t^2)^2}{2}.\quad (4.40)$$

The constant  $C_0$  will be chosen in order to kill the second-order terms arising from Itô's formula. Let

$$\tau^* = \inf\{t \geq t_0 : U_t = 1\}.\quad (4.41)$$

Applying Itô's formula and choosing  $C_0$  of order  $\tilde{\sigma}_1^2\tilde{\mu}^{-(1-\gamma)}$  yields

$$dU_t \leq [C_1 + C_2(t)U_t] dt + \tilde{\sigma}_1 g(t) dW_t^1,\quad (4.42)$$

where (using the fact that  $\tilde{\sigma} \leq \tilde{\mu}$ )

$$\begin{aligned}C_1 &= \mathcal{O}(T\tilde{\mu}^{2-2\gamma}) + \mathcal{O}(\tilde{\varepsilon}), \\ C_2(t) &= t \vee 0 + \mathcal{O}(T\tilde{\mu}^{1-\gamma}),\end{aligned}\quad (4.43)$$

and  $g(t)$  is at most of order 1 for  $t \leq \tau_1 \wedge \tau_2 \wedge \tau^*$ . Hence

$$U_{t \wedge \tau_1 \wedge \tau_2 \wedge \tau^*} \leq U_{t_0} + C_1(t - t_0) + \int_{t_0}^{t \wedge \tau_1 \wedge \tau_2 \wedge \tau^*} C_2(s)U_s ds + \tilde{\sigma}_1 \int_{t_0}^{t \wedge \tau_1 \wedge \tau_2 \wedge \tau^*} g(s) dW_s^1.\quad (4.44)$$

We introduce a last stopping time

$$\tau_3 = \inf\left\{t \geq t_0 : \left| \tilde{\sigma}_1 \int_0^{t \wedge \tau_1 \wedge \tau_2 \wedge \tau^*} g(s) dW_s^1 \right| \geq h_3\right\}.\quad (4.45)$$

Then Lemma 4.3 implies

$$\mathbb{P}\{\tau_3 < t\} \leq e^{-\kappa_3 h_3^2 / \tilde{\sigma}_1^2}\quad (4.46)$$

for a  $\kappa_3 > 0$ . Applying Gronwall's lemma to (4.44) we get

$$\begin{aligned}U_{t \wedge \tau_1 \wedge \tau_2 \wedge \tau_3 \wedge \tau^*} &\leq [U_{t_0} + C_1(t - t_0) + h_3] \exp\left\{\int_{t_0}^{t \wedge \tau_1 \wedge \tau_2 \wedge \tau_3 \wedge \tau^*} C_2(u) du\right\} \\ &= \mathcal{O}(T^2\tilde{\mu}^{2-3\gamma}) + \mathcal{O}(\tilde{\sigma}^2 T^2 \tilde{\mu}^{1-2\gamma}) + \mathcal{O}(\tilde{\varepsilon} T \tilde{\mu}^{-\gamma}).\end{aligned}\quad (4.47)$$

This shows in particular that  $\tau^* > t$ , provided we take  $\gamma$  small enough. Now (4.33) follows from the decomposition

$$\begin{aligned}\mathbb{P}\left\{\sup_{t_0 \leq s \leq t} \frac{|z_s - z_s^0|}{\sqrt{\zeta(s)}} \geq H\right\} &\leq \mathbb{P}\left\{\sup_{t_0 \leq s \leq t \wedge \tau_1 \wedge \tau_2 \wedge \tau_3} \frac{|z_s^1 - z_s^0|}{\sqrt{\zeta(s)}} \geq H - \sup_{t_0 \leq s \leq t \wedge \tau_1 \wedge \tau_2 \wedge \tau_3} \frac{U_s}{\sqrt{\zeta(s)}}\right\} \\ &\quad + \mathbb{P}\{\tau_1 > t\} + \mathbb{P}\{\tau_2 > t\} + \mathbb{P}\{\tau_3 > t\},\end{aligned}\quad (4.48)$$

and (4.34) is obtained in a similar way.  $\square$

### 4.3 Computation of the kernel

We can now proceed to estimating the contribution of the dynamics along the separatrix to the kernel of the Markov chain.

**Proposition 4.5.** *Fix some  $\gamma \in (0, 1/3)$  and an initial condition  $(\xi_0, z_0) = (-L, z_0) \in F_-$  with  $|z_0| = \mathcal{O}(\tilde{\mu}^{1-\gamma})$ .*

1. *Assume  $\sqrt{\varepsilon} \leq \tilde{\mu}^{2-2\gamma}$ . Then there exist constants  $C, \kappa_1, h_0 > 0$  such that the sample path starting in  $(\xi_0, z_0)$  will hit  $F_+$  for the first time at a point  $(\xi_1, z_1)$  such that*

$$\mathbb{P}\left\{|z_1 - z_T^0| \geq h \text{ or } \xi_1 \neq L\right\} \leq \frac{C}{\tilde{\mu}^{2\gamma}} \left( \exp\left\{-\kappa_1 \frac{h^2}{\tilde{\sigma}_1^2 \tilde{\mu}^{-2\gamma}}\right\} + \exp\left\{-\kappa_1 \frac{\tilde{\mu}^2}{\tilde{\sigma}^2}\right\} \right) \quad (4.49)$$

for all  $h \leq h_0 \tilde{\sigma}_1^2 \tilde{\mu}^{-2\gamma}$ .

2. *If  $\sqrt{\varepsilon} > \tilde{\mu}^{2-2\gamma}$ , then the first-hitting point of  $F_+$  satisfies*

$$\mathbb{P}\left\{z_1 \leq z_T^0 - h\right\} \leq \frac{C}{\tilde{\mu}^{2\gamma}} \left( \exp\left\{-\kappa_1 \frac{h^2}{\tilde{\sigma}_1^2 \tilde{\mu}^{-2\gamma}}\right\} + \exp\left\{-\kappa_1 \frac{\tilde{\mu}^2}{\tilde{\sigma}^2}\right\} \right) \quad (4.50)$$

for all  $h \leq h_0 \tilde{\sigma}_1^2 \tilde{\mu}^{-2\gamma}$ .

PROOF: Consider first the case  $\sqrt{\varepsilon} \leq \tilde{\mu}^{2-2\gamma}$ . We apply (4.34) with  $H' = \tilde{\mu}^{1-\gamma}$ . The condition  $\gamma < 1/3$  ensures that the error terms in the exponent are negligible. Thus with probability greater or equal than

$$1 - \frac{C}{\tilde{\mu}^{2\gamma}} e^{-\kappa \tilde{\mu}^2 / (2\tilde{\sigma}^2)}, \quad (4.51)$$

$|u_t - u_t^0|$  remains bounded by  $\tilde{\mu}^{1-\gamma}$  up to time  $T + \mathcal{O}(|\log \tilde{\mu}|^{-1/2})$ . Since  $u_T^0$  is of order  $\tilde{\mu}^{1-\gamma}$ , this implies that the first-hitting time  $\tau$  of  $F_+$  satisfies  $\tau = T + \mathcal{O}(\tilde{\mu}^{1-\gamma})$  with probability bounded below by (4.51).

For any decomposition  $h = h_0 + h_1$ , we can write

$$\mathbb{P}\{|z_\tau - z_T^0| \geq h\} \leq \mathbb{P}\{|z_\tau - z_\tau^0| \geq h_0\} + \mathbb{P}\{|z_\tau^0 - z_T^0| \geq h_1\}. \quad (4.52)$$

The first term on the right-hand side can be bounded by using (4.33), while the second one can be expressed in terms of the probability of  $\tau - T$  being large. The optimal bound is obtained by choosing  $h_0$  and  $h_1$  in such a way that  $h_0 = h_1 \tilde{\mu}^{-\gamma}$ , and yields (4.49).

In the case  $\sqrt{\varepsilon} \leq \tilde{\mu}^{2-2\gamma}$ , we can conclude in the same way by observing that  $z_t$  is bounded below by its value for  $\varepsilon = 0$ , the  $\varepsilon$ -dependent term of  $dz_t$  in (4.1) being positive.  $\square$

## 5 Dynamics during an SAO

### 5.1 Action–angle-type variables

In this section, we construct another set of coordinates allowing to describe the dynamics during a small-amplitude oscillation. In the limit  $\varepsilon \rightarrow 0$  and  $\tilde{\mu} \rightarrow 0$ , the deterministic system (2.11) becomes

$$\begin{aligned} \dot{\xi} &= \frac{1}{2} - z \\ \dot{z} &= 2\xi z. \end{aligned} \quad (5.1)$$

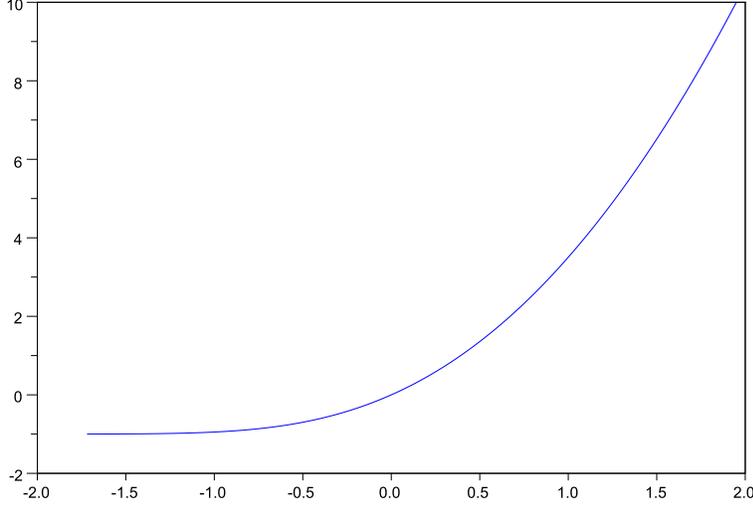


FIGURE 7. Graph of  $u \mapsto f(u)$ .

This system admits a first integral

$$K = 2z e^{-2z-2\xi^2+1} , \quad (5.2)$$

which is equivalent to the first integral found in [BE92]. The normalisation is chosen in such a way that  $K \in [0, 1]$  for  $z \geq 0$ . The separatrix is given in this limit by  $K = 0$ , while  $K = 1$  corresponds to the stationary point  $P$ . When  $\varepsilon$  and  $\tilde{\mu}$  are positive, we obtain

$$\dot{K} = \left[ 2\tilde{\mu}(1 - 2z) + \frac{4}{3}\sqrt{\varepsilon}\xi^4 \right] e^{-2z-2\xi^2+1} . \quad (5.3)$$

Observe that the term of order  $\sqrt{\varepsilon}$  is strictly positive.

In order to analyse the dynamics in more detail, it is useful to introduce an angle variable  $\phi$ . We define a coordinate transformation from  $(0, 1] \times \mathbb{S}^1$  to  $\mathbb{R} \times \mathbb{R}_+$  by

$$\begin{aligned} \xi &= -\sqrt{\frac{-\log K}{2}} \sin \phi \\ z &= \frac{1}{2} \left( 1 + f \left( \sqrt{\frac{-\log K}{2}} \cos \phi \right) \right) . \end{aligned} \quad (5.4)$$

Here  $f : \mathbb{R} \rightarrow (-1, +\infty)$  is defined as the solution of

$$\log(1 + f(u)) - f(u) = -2u^2 \quad (5.5)$$

such that

$$\text{sign } f(u) = \text{sign } u . \quad (5.6)$$

The graph of  $f$  is plotted in Figure 7.

**Lemma 5.1.** *The function  $f$  has the following properties:*

- *Lower bounds:*

$$f(u) > -1 \quad \text{and} \quad f(u) \geq 2u \quad \forall u \in \mathbb{R} . \quad (5.7)$$

- *Upper bounds:* There exist constants  $C_1, C_2 > 0$  and a function  $r : \mathbb{R}_- \rightarrow \mathbb{R}$ , with  $0 \leq r(u) \leq C_1 e^{-1-2u^2}$ , such that

$$f(u) \leq C_2 u + 2u^2 \quad \forall u \geq 0, \quad (5.8)$$

$$f(u) = -1 + e^{-1-2u^2} [1 + r(u)] \quad \forall u \leq 0. \quad (5.9)$$

- *Derivatives:*  $f \in \mathcal{C}^\infty$  and

$$f'(u) = 4u \frac{1 + f(u)}{f(u)}, \quad (5.10)$$

$$f''(u) = 4 \frac{1 + f(u)}{f(u)} \left( 1 - 4 \frac{u^2}{f(u)^2} \right). \quad (5.11)$$

- There exists a constant  $M > 0$  such that

$$0 < f''(u) \leq M \quad \forall u \in \mathbb{R}. \quad (5.12)$$

PROOF: The results follow directly from the implicit function theorem and elementary calculus.  $\square$

We can now derive an expression for the SDE in coordinates  $(K, \phi)$ . To ease notation, we introduce the function

$$X = X(K, \phi) = \sqrt{\frac{-\log K}{2}} \cos \phi, \quad (5.13)$$

a parameter  $\tilde{\sigma} > 0$  defined by

$$\tilde{\sigma}^2 = \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2, \quad (5.14)$$

and the two-dimensional Brownian motion  $dW_t = (d\tilde{W}_t^{(1)}, d\tilde{W}_t^{(2)})^T$ .

**Proposition 5.2.** *For  $z > 0$ , the system of SDEs (2.11) is equivalent to the system*

$$\begin{aligned} dK_t &= \tilde{\mu} f_1(K_t, \phi_t) dt + \tilde{\sigma} \psi_1(K_t, \phi_t) dW_t \\ d\phi_t &= f_2(K_t, \phi_t) dt + \tilde{\sigma} \psi_2(K_t, \phi_t) dW_t, \end{aligned} \quad (5.15)$$

where we introduced the following notations.

- The new drift terms are of the form

$$f_1(K, \phi) = -2K \frac{f(X)}{1 + f(X)} \left[ 1 + \frac{\sqrt{\varepsilon}}{\tilde{\mu}} R_{K,\varepsilon}(K, \phi) + \frac{\tilde{\sigma}^2}{\tilde{\mu}} R_{K,\sigma}(K, \phi) \right], \quad (5.16)$$

$$f_2(K, \phi) = \frac{f(X)}{2X} \left[ 1 + \frac{2\tilde{\mu} \tan \phi}{\log K (1 + f(X))} + \sqrt{\varepsilon} R_{\phi,\varepsilon}(K, \phi) + \tilde{\sigma}^2 R_{\phi,\sigma}(K, \phi) \right]. \quad (5.17)$$

- The remainders in the drift terms are bounded as follows. Let

$$\rho(K, \phi) = \begin{cases} \sqrt{|\log K|} & \text{if } \cos \phi \geq 0, \\ K^{-\cos^2 \phi} & \text{if } \cos \phi < 0. \end{cases} \quad (5.18)$$

Then there exists a constant  $M_1 > 0$  such that for all  $K \in (0, 1)$  and all  $\phi \in \mathbb{S}^1$ ,

$$\begin{aligned} |R_{K,\varepsilon}(K, \phi)| &\leq M_1 |\log K|^2, & |R_{K,\sigma}(K, \phi)| &\leq M_1 \rho(K, \phi), \\ |R_{\phi,\varepsilon}(K, \phi)| &\leq M_1 |\log K|^{3/2} \rho(K, \phi), & |R_{\phi,\sigma}(K, \phi)| &\leq M_1 \rho(K, \phi)^2 / |\log K|. \end{aligned} \quad (5.19)$$

Furthermore,  $-f(X)R_{K,\varepsilon}(K, \phi) \geq 0$ .

- The diffusion coefficients are given by

$$\begin{aligned}\psi_1(K, \phi) &= \left( 2\sqrt{2}\frac{\tilde{\sigma}_1}{\tilde{\sigma}}K \left[ \sqrt{-\log K} - \frac{f(X)}{1+f(X)} \right] \sin \phi, -2\frac{\tilde{\sigma}_2}{\tilde{\sigma}}K \frac{f(X)}{1+f(X)} \right), \\ \psi_2(K, \phi) &= \left( -\frac{\tilde{\sigma}_1}{\tilde{\sigma}} \sqrt{\frac{2}{-\log K}} \frac{1+f(X)\cos\phi}{[1+f(X)]\cos\phi}, \frac{\tilde{\sigma}_2}{\tilde{\sigma}} \frac{1}{\log K} \frac{f(X)}{1+f(X)} \tan \phi \right).\end{aligned}\quad (5.20)$$

- There exists a constant  $M_2 > 0$  such that for all  $K \in (0, 1)$  and all  $\phi \in \mathbb{S}^1$ ,

$$\|\psi_1(K, \phi)\|^2 \leq M_2 K^2 \rho(K, \phi)^2, \quad \|\psi_2(K, \phi)\|^2 \leq M_2 \frac{\rho(K, \phi)^2}{|\log K|^2}. \quad (5.21)$$

PROOF: The result follows from Itô's formula, by a straightforward though lengthy computation. The difference between the bounds obtained for  $\cos \phi \geq 0$  and  $\cos \phi < 0$  is due to the fact that terms such as  $f(X) \tan(\phi)/(1+f(X))$  can be bounded by a constant times  $\sqrt{-\log K}$  in the first case, and by a constant times  $K^{-\cos^2 \phi}$  in the second one, as a consequence of Lemma 5.1. The fact that  $-f(X)R_{K,\varepsilon}$  is positive follows from the positivity of the term of order  $\sqrt{\varepsilon}$  in (5.3).  $\square$

## 5.2 Averaging

In System (5.15), the variable  $K$  changes more slowly than the variable  $\phi$ , which is a consequence of the fact that  $K$  is a first integral when  $\tilde{\mu} = \varepsilon = \tilde{\sigma} = 0$ . This suggests to use an averaging approach to analyse the dynamics. However, since the behaviour near  $\phi = \pi$  has already been considered in the previous section, using  $(\xi, z)$ -coordinates, we only need to consider  $\phi \in [\phi_0, \phi_1]$ , where  $-\pi < \phi_0 < 0 < \phi_1 < \pi$ .

We look for a change of variables of the form

$$\bar{K} = K + \tilde{\mu}w(K, \phi) \quad (5.22)$$

which eliminates the term of order  $\tilde{\mu}$  in  $dK_t$ . Itô's formula yields

$$d\bar{K}_t = dK_t + \tilde{\mu} \frac{\partial w}{\partial \phi} d\phi_t + \tilde{\mu} \frac{\partial w}{\partial K} dK_t + \frac{1}{2} \tilde{\mu} \left( \frac{\partial^2 w}{\partial K^2} dK_t^2 + 2 \frac{\partial^2 w}{\partial K \partial \phi} dK_t d\phi_t + \frac{\partial^2 w}{\partial \phi^2} d\phi_t^2 \right). \quad (5.23)$$

Replacing  $dK_t$  et  $d\phi_t$  by their expressions in (5.15), we get

$$d\bar{K}_t = \tilde{\mu} \left( f_1 + \frac{\partial w}{\partial \phi} f_2 + \mathcal{O}(\tilde{\mu}) + \mathcal{O}(\tilde{\sigma}^2) \right) dt + \tilde{\sigma} \left( \psi_1 + \tilde{\mu} \left( \frac{\partial w}{\partial \phi} \psi_2 + \frac{\partial w}{\partial K} \psi_1 \right) \right) dW_t. \quad (5.24)$$

Thus choosing the function  $w$  in such a way that

$$f_1 + \frac{\partial w}{\partial \phi} f_2 = 0 \quad (5.25)$$

will decrease the order of the drift term in (5.24). We thus define the function  $w$  by the integral

$$w(K, \phi) = - \int_{\phi_0}^{\phi} \frac{f_1(K, \theta)}{f_2(K, \theta)} d\theta, \quad (5.26)$$

which is well-defined, since (5.7) shows that  $f_2(K, \phi)$  is bounded below by a positive constant, for sufficiently small  $\tilde{\mu}$ ,  $\varepsilon$  and  $\tilde{\sigma}$ .

**Lemma 5.3.** *Let  $\phi_0 \in (-\pi, -\pi/2)$  and  $\phi_1 \in (\pi/2, \pi)$  be such that  $\cos^2(\phi_0), \cos^2(\phi_1) \leq b$  for some  $b \in (0, 1)$ . Then*

$$w(K, \phi_1) = -\frac{\sqrt{2}e}{\sqrt{-\log K}} \left[ \frac{K^{\sin^2 \phi_0}}{-\sin \phi_0} + \frac{K^{\sin^2 \phi_1}}{\sin \phi_1} + r_1(K) \right] \left( 1 + r_2(K) + r_\varepsilon(K) \right), \quad (5.27)$$

where the remainder terms satisfy

$$\begin{aligned} r_1(K) &= \mathcal{O}(K \log(|\log K|)), \\ r_2(K) &= \mathcal{O}\left(\frac{1}{|\log K|} + \tilde{\mu}K^{-b} + \tilde{\sigma}^2\left(\frac{K^{-b}}{\tilde{\mu}} + \frac{K^{-2b}}{|\log K|}\right)\right), \\ 0 \leq r_\varepsilon(K) &\leq \sqrt{\varepsilon} \mathcal{O}\left(\frac{|\log K|^2}{\tilde{\mu}} + K^{-b}|\log K|^{3/2}\right). \end{aligned} \quad (5.28)$$

Furthermore, the derivatives of  $w$  satisfy the bounds

$$\frac{\partial w}{\partial K}(K, \phi) = \mathcal{O}\left(\frac{K^{-b}}{\sqrt{|\log K|}}\right), \quad \frac{\partial w}{\partial \phi}(K, \phi) = \mathcal{O}(K^{1-b}\sqrt{|\log K|}), \quad (5.29)$$

and

$$\frac{\partial^2 w}{\partial K^2} = \mathcal{O}\left(\frac{K^{-1-b}}{\sqrt{|\log K|}}\right), \quad \frac{\partial^2 w}{\partial K \partial \phi} = \mathcal{O}\left(\frac{K^{-b}}{\sqrt{|\log K|}}\right), \quad \frac{\partial^2 w}{\partial \phi^2} = \mathcal{O}(K^{1-b}\sqrt{|\log K|}). \quad (5.30)$$

PROOF: We split the integral into three parts. Using the change of variables  $t = \sin \phi$  and a partial fraction decomposition, we find that the leading part of the integral on  $[-\pi/2, \pi/2]$  satisfies

$$\int_{-\pi/2}^{\pi/2} \frac{4KX(K, \phi)}{1 + f(X(K, \phi))} d\phi = \mathcal{O}\left(K \frac{\log(|\log K|)}{\sqrt{|\log K|}}\right). \quad (5.31)$$

Next we consider the integral on  $[\phi_0, -\pi/2]$ . The change of variables  $u = \sqrt{-2 \log K} \sin \phi$ , (5.9) and asymptotic properties of the error function imply

$$\begin{aligned} \int_{\phi_0}^{-\pi/2} \frac{4KX}{1 + f(X)} d\phi &= 2K \int_{\sin \phi_0 \sqrt{-2 \log K}}^{-\sqrt{-2 \log K}} \frac{du}{1 + f\left(-\sqrt{\frac{-\log K}{2} - \frac{u^2}{4}}\right)} \\ &= 2e \int_{\sin \phi_0 \sqrt{-2 \log K}}^{-\sqrt{-2 \log K}} e^{-u^2/2} [1 + \mathcal{O}(K e^{-u^2/2})] du \\ &= -2e \frac{K^{\sin^2 \phi_0}}{\sqrt{-2 \log K} (-\sin \phi_0)} \left[ 1 + \mathcal{O}\left(\frac{1}{|\log K|}\right) + \mathcal{O}(K^{\cos^2 \phi_0}) \right]. \end{aligned} \quad (5.32)$$

The integral on  $[\pi/2, \phi_1]$  can be computed in a similar way. This yields the leading term in (5.27), and the form of the remainders follows from (5.19) with  $\rho = K^{-b}$ . The bound on  $\partial w / \partial \phi$  follows directly from (5.25), while the bound on  $\partial w / \partial K$  is obtained by computing the derivative of  $f_1/f_2$ . The bounds on second derivatives follow by similar computations.  $\square$

Notice that for the remainder  $r_2(K)$  to be small, we need that  $K^b \gg \tilde{\mu}$  and  $K^b \gg \tilde{\sigma}^2/\tilde{\mu}$ . Then the term  $r_\varepsilon(K)$  is of order  $\sqrt{\varepsilon}|\log \tilde{\mu}|^2/\tilde{\mu}$ , which is small for  $\tilde{\mu}/|\log \tilde{\mu}|^2 \gg \sqrt{\varepsilon}$ . If that is the case, then  $w(K, \phi_1)$  has order  $K^{1-b}/\sqrt{|\log K|}$ . Otherwise,  $w(K, \phi_1)$  has order  $\sqrt{\varepsilon}K^{1-b}|\log K|^{3/2}/\tilde{\mu}$ . In the sequel, we will sometimes bound  $1/\sqrt{|\log K|}$  by 1 to get simpler expressions.

### 5.3 Computation of the kernel

We can now proceed to the computation of the rotational part of the kernel of the Markov chain. Recall the broken lines  $F_{\pm}$  introduced in Section 4. For an initial condition  $(L, z_0) \in F_+$ , we want to compute the coordinates of the point  $(\xi_{\tau}, z_{\tau})$  at the first time

$$\tau = \inf\{t > 0: (\xi_t, z_t) \in F_-\} \quad (5.33)$$

that the path starting in  $(L, z_0)$  hits  $F_-$ .

We will assume that there is a  $\beta \in (0, 1]$  such that

$$(c_- \tilde{\mu})^{\beta} \leq z_0 \leq z_{\max} < \frac{1}{2}. \quad (5.34)$$

The  $(K, \phi)$ -coordinates of the initial condition are given by

$$\begin{aligned} K_0 &= 2z_0 e^{1-2z_0} e^{-2L^2} \geq 2(c_- \tilde{\mu})^{\beta+\gamma}, \\ \sin^2 \phi_0 &= \frac{2L^2}{-\log K_0} \geq \frac{\gamma}{\beta + \gamma}, \end{aligned} \quad (5.35)$$

with  $\phi_0 \in (-\pi, -\pi/2)$ . Thus Lemma 5.3 applies with  $b = \beta/(\beta + \gamma) < 1$ . Notice that

$$K_0^{\cos^2 \phi_0} \geq 2^b (c_- \tilde{\mu})^{\beta}. \quad (5.36)$$

**Proposition 5.4.** *Assume  $z_0$  satisfies (5.34) for a  $\beta < 1$ . Then there exists a constant  $\kappa > 0$  such that the following holds for sufficiently small  $\tilde{\mu}$  and  $\tilde{\sigma}$ .*

1. *If  $\sqrt{\varepsilon} \leq \tilde{\mu}/|\log \tilde{\mu}|^2$ , then with probability greater or equal than*

$$1 - e^{-\kappa \tilde{\mu}^2 / \tilde{\sigma}^2}, \quad (5.37)$$

*$(\xi_t, z_t)$  hits  $F_-$  for the first time at a point  $(-L, z_1)$  such that*

$$z_1 = z_0 + \tilde{\mu} A(z_0) + \frac{z_0}{1-2z_0} \left[ \tilde{\sigma} V(z_0) + \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{-2\beta}) \right]. \quad (5.38)$$

*The function  $A(z_0)$  is given by*

$$A(z_0) = \frac{e^{2z_0}}{L(1-2z_0)} \left[ 1 + \mathcal{O}(z_0 \log |\log \tilde{\mu}|) + \mathcal{O}\left(\frac{1}{|\log z_0|}\right) \right], \quad (5.39)$$

*and  $V(z_0)$  is a random variable satisfying*

$$\mathbb{P}\{\tilde{\sigma}|V(z_0)| \geq h\} \leq 2 \exp\left\{-\frac{\kappa h^2 \tilde{\mu}^{2\beta}}{\tilde{\sigma}^2}\right\} \quad \forall h > 0. \quad (5.40)$$

2. *If  $\sqrt{\varepsilon} > \tilde{\mu}/|\log \tilde{\mu}|^2$ , then  $(\xi_t, z_t)$  hits  $F_-$  for the first time either at a point  $(-L, z_1)$  such that  $z_1$  is greater or equal than the right-hand side of (5.38), or at a point  $(\xi_1, 1/2)$  with  $-L \leq \xi_1 \leq 0$ , again with a probability bounded below by (5.37).*

**PROOF:** We first consider the case  $\sqrt{\varepsilon} \leq \tilde{\mu}/|\log \tilde{\mu}|^2$ .

- **Step 1 :** To be able to bound various error terms, we need to assume that  $K_t$  stays bounded below. We thus introduce a second stopping time

$$\tau_1 = \inf\{t > 0: \cos \phi_t < 0, K_t^{\cos^2 \phi_t} < (c_- \tilde{\mu})^\beta\}. \quad (5.41)$$

We start by showing that  $\tau \wedge \tau_1$  is bounded with high probability. Proposition 5.2 implies the existence of a constant  $C > 0$  such that

$$d\phi_t \geq C dt + \tilde{\sigma} \psi_2(K_t, \phi_t) dW_t. \quad (5.42)$$

Integrating this relation between 0 and  $t$ , we get

$$\phi_t \geq \phi_0 + Ct + \tilde{\sigma} \int_0^t \psi_2(K_s, \phi_s) dW_s. \quad (5.43)$$

Lemma 4.3 and (5.21) provide the bound

$$\mathbb{P}\left\{\left|\tilde{\sigma} \int_0^{t \wedge \tau_1} \psi_2(K_s, \phi_s) dW_s\right| \geq h\right\} \leq \exp\left\{-\frac{\kappa h^2 \tilde{\mu}^{2\beta}}{\tilde{\sigma}^2}\right\} \quad (5.44)$$

for some  $\kappa > 0$ . Since by definition,  $\phi_{\tau \wedge \tau_1} - \phi_0 < 2\pi$ , we get

$$\mathbb{P}\left\{\tau \wedge \tau_1 > \frac{2\pi + h}{C}\right\} \leq \exp\left\{-\frac{\kappa h^2 \tilde{\mu}^{2\beta}}{\tilde{\sigma}^2}\right\}. \quad (5.45)$$

From now on, we work on the set  $\Omega_1 = \{\tau \wedge \tau_1 \leq (2\pi + 1)/C\}$ , which has probability greater or equal  $1 - e^{-\kappa \tilde{\mu}^{2\beta}/\tilde{\sigma}^2}$ .

- **Step 2 :** The SDE (5.23) for  $\bar{K}_t$  can be written

$$d\bar{K}_t = \bar{K}_t \bar{f}(\bar{K}_t, \phi_t) dt + \tilde{\sigma} \bar{K}_t \bar{\psi}(\bar{K}_t, \phi_t) dW_t, \quad (5.46)$$

where the bounds in Proposition 5.2 and Lemma 5.3 yield

$$\begin{aligned} \bar{f}(\bar{K}, \phi) &= \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{1-3\beta}), \\ \|\bar{\psi}(\bar{K}, \phi)\|^2 &= \mathcal{O}(\tilde{\mu}^{-2\beta}). \end{aligned} \quad (5.47)$$

By Itô's formula, the variable  $Q_t = \log \bar{K}_t$  satisfies

$$dQ_t = \tilde{f}(Q_t, \phi_t) dt + \tilde{\sigma} \tilde{\psi}(Q_t, \phi_t) dW_t, \quad (5.48)$$

where  $\tilde{f}(Q, \phi) = \bar{f}(e^Q, \phi) + \mathcal{O}(\tilde{\sigma}^2 \tilde{\mu}^{-2\beta})$  and  $\tilde{\psi}(Q, \phi) = \bar{\psi}(e^Q, \phi)$ . Setting

$$V_t = \int_0^t \tilde{\psi}(Q_s, \phi_s) dW_s, \quad (5.49)$$

we obtain, integrating (5.48) and using the fact that  $\tilde{\mu}^{1-3\beta} \leq \tilde{\mu}^{-2\beta}$ ,

$$Q_t = Q_0 + \tilde{\sigma} V_t + \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{-2\beta}). \quad (5.50)$$

Another application of Lemma 4.3 yields

$$\mathbb{P}\{\tilde{\sigma} |V_{t \wedge \tau_1}| \geq h_1\} \leq 2 \exp\left\{-\frac{\kappa_1 h_1^2 \tilde{\mu}^{2\beta}}{\tilde{\sigma}^2}\right\} \quad (5.51)$$

for some  $\kappa_1 > 0$ . A convenient choice is  $h_1 = \tilde{\mu}^{1-\beta}$ . From now on, we work on the set  $\Omega_1 \cap \Omega_2$ , where  $\Omega_2 = \{\tilde{\sigma} V_{t \wedge \tau_1} < \tilde{\mu}^{1-\beta}\}$  satisfies  $\mathbb{P}(\Omega_2) \geq 1 - e^{-\kappa_1 \tilde{\mu}^{2\beta}/\tilde{\sigma}^2}$ .

- **Step 3** : Returning to the variable  $\bar{K}$ , we get

$$\bar{K}_t = K_0 e^{\tilde{\sigma} V_t} \left[ 1 + \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{-2\beta}) \right], \quad (5.52)$$

and thus

$$K_t = K_0 e^{\tilde{\sigma} V_t} \left[ 1 + \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{-2\beta}) \right] - \tilde{\mu} w(K_t, \phi_t). \quad (5.53)$$

Using the implicit function theorem and the upper bound on  $w$ , we get the a priori bound

$$\frac{|K_t - K_0|}{K_0} = \mathcal{O}(\tilde{\mu}^{1-\beta} + \tilde{\sigma}^2 \tilde{\mu}^{-2\beta}). \quad (5.54)$$

- **Step 4** : The a priori estimate (5.54) implies that on  $\Omega_1 \cap \Omega_2$ , the sample path cannot hit  $F_-$  on the part  $\{-L \leq \xi \leq 0, z = 1/2\}$ . Indeed, this would imply that  $K_\tau \geq (c_- \tilde{\mu})^\gamma$ , while  $K_0 \leq a(c_- \tilde{\mu})^\gamma$  with  $a = 2z_{\max} e^{1-2z_{\max}} < 1$ . As a consequence, we would have  $(K_\tau - K_0)/K_0 > (1-a)/a$ , contradicting (5.54).

Let us now show that we also have  $\tau_1 \geq \tau$  on  $\Omega_1 \cap \Omega_2$ . Assume by contradiction that  $\tau_1 < \tau$ . Then we have  $K_{\tau_1} = (c_- \tilde{\mu})^\beta$  and  $\cos^2 \phi_{\tau_1} < \beta/(\beta+\gamma)$ , so that  $K_{\tau_1} = \mathcal{O}(\tilde{\mu}^{\beta+\gamma})$ . Thus  $\tilde{\mu} w(K_{\tau_1}, \phi_{\tau_1}) = \mathcal{O}(K_{\tau_1} \tilde{\mu}^{1-\beta}) = \mathcal{O}(\tilde{\mu}^{1+\gamma})$ . Together with the lower bound (5.35) on  $K_0$ , this implies that the right-hand side of (5.53) is larger than a constant times  $\tilde{\mu}^{\beta+\gamma}$  at time  $t = \tau_1$ . But this contradicts the fact that  $K_{\tau_1} = \mathcal{O}(\tilde{\mu}^{\beta+\gamma})$ .

- **Step 5** : The previous step implies that  $\xi_\tau = -L$  on  $\Omega_1 \cap \Omega_2$ . We can thus write

$$\phi_\tau = g(K_\tau) \quad \text{where} \quad \sin(g(K)) = \sqrt{\frac{2}{-\log K}} L, \quad (5.55)$$

with  $g(K) \in (\pi/2, \pi)$ . Notice that  $g(K_0) = -\phi_0$ . Furthermore, we have

$$g'(K) = \frac{L}{\sqrt{2} K (-\log K)^{3/2} \cos(g(K))}, \quad (5.56)$$

and thus  $K_\tau g'(K_\tau) = \mathcal{O}(1/|\log \tilde{\mu}|)$ . Using this in the Taylor expansion

$$w(K_\tau, \phi_\tau) = w(K_0, -\phi_0) + (K_\tau - K_0) \left[ \frac{\partial w}{\partial K}(K_\theta, g(K_\theta)) + \frac{\partial w}{\partial \phi}(K_\theta, g(K_\theta)) g'(K_\theta) \right], \quad (5.57)$$

which holds for some  $K_\theta \in (K_0, K_\tau)$ , yields the estimate

$$\frac{w(K_\tau, \phi_\tau)}{K_0} = \frac{w(K_0, -\phi_0)}{K_0} + \mathcal{O}(\tilde{\mu}^{1-2\beta} + \tilde{\sigma}^2 \tilde{\mu}^{-3\beta}). \quad (5.58)$$

Substitution in (5.53) yields the more precise estimate

$$K_1 = K_0 \left[ e^{\tilde{\sigma} V_\tau} - \tilde{\mu} \frac{w(K_0, -\phi_0)}{K_0} + \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{1-3\beta}) \right]. \quad (5.59)$$

- **Step 6** : Finally, we return to the variable  $z_1 = z_\tau$ . Eliminating  $\phi$  from the equations (5.4), it can be expressed in terms of  $K_\tau$  as

$$z_1 = G(K_\tau) := \frac{1}{2} \left[ 1 + f \left( -\sqrt{\frac{-\log K_\tau}{2} - L^2} \right) \right]. \quad (5.60)$$

Note that  $G(K_0) = z_0$ , while

$$G'(K_0) = -\frac{1}{2K_0} \frac{1 + f\left(-\sqrt{\frac{-\log K_0}{2} - L^2}\right)}{f\left(-\sqrt{\frac{-\log K_0}{2} - L^2}\right)} = \frac{z_0}{K_0(1 - 2z_0)}, \quad (5.61)$$

and

$$G''(K) = \frac{1}{2K^2} \frac{1 + f\left(-\sqrt{\frac{-\log K}{2} - L^2}\right)}{f\left(-\sqrt{\frac{-\log K}{2} - L^2}\right)} \left[ 1 - \frac{1}{f\left(-\sqrt{\frac{-\log K}{2} - L^2}\right)^2} \right], \quad (5.62)$$

which has order  $z_0^2/K_0^2$ . The Taylor expansion

$$z_1 = G(K_0) + (K_1 - K_0)G'(K_0) + \frac{(K_1 - K_0)^2}{2}G''(K_0) \quad (5.63)$$

thus becomes

$$z_1 = z_0 + \frac{K_1 - K_0}{K_0} \frac{z_0}{1 - 2z_0} + \mathcal{O}\left(\left[\frac{K_1 - K_0}{K_0} z_0\right]^2\right). \quad (5.64)$$

By (5.59), we have

$$\frac{K_1 - K_0}{K_0} = \tilde{\sigma}V_\tau - \tilde{\mu} \frac{w(K_0, -\phi_0)}{K_0} + \mathcal{O}(\tilde{\mu}^{2(1-\beta)} + \tilde{\sigma}^2 \tilde{\mu}^{-2\beta}), \quad (5.65)$$

and Lemma 5.3 yields

$$-\mu \frac{w(K_0, -\phi_0)}{K_0} = \frac{2e}{L} \left[ K_0^{-\cos^2 \phi_0} + \mathcal{O}(\log|\log K_0|) \right] \left[ 1 + \mathcal{O}\left(\frac{1}{|\log K_0|}\right) \right]. \quad (5.66)$$

Now (5.35) implies  $K_0^{-\cos^2 \phi_0} = e^{2z_0}(2e z_0)^{-1}$  and  $c_1|\log z_0| \leq |\log K_0| \leq c_2|\log \tilde{\mu}|$ . This completes the proof of the case the case  $\sqrt{\varepsilon} \leq \tilde{\mu}/|\log \tilde{\mu}|^2$ .

In the case  $\sqrt{\varepsilon} > \tilde{\mu}/|\log \tilde{\mu}|^2$ , we just use the fact that  $K_t$  is bounded below by its value in the previous case, as a consequence of (5.29).  $\square$

## 5.4 Proof of Theorem 2.4

To complete the proof of Theorem 2.4, it only remains to construct a subset  $A$  of the interval  $E$  on which the Markov chain is defined, such that if  $X_0 \in A$ , then  $X_1 \in A$  with high probability. Then the result will follow from Lemmas 3.5 and 3.6.

We choose the curve  $\mathcal{F}$  defining the Markov chain as being  $F_-$  in  $(\xi, z)$ -coordinates. Let  $(-L, z_0) \in F_-$  be an initial condition with  $|z_0| = \mathcal{O}(\tilde{\mu}^{1-\gamma})$ . The deterministic approximation of the trajectory starting at that point hits  $F_+$  at a point  $(L, z_T^0)$  such that  $z_T^0 \geq c_0 \tilde{\mu}^{1-\gamma}$  for some  $c_0 > 0$  (cf. (4.9)). Applying Proposition 4.5 with  $h$  of order  $\tilde{\mu}^{1-\gamma}$ , we obtain that the probability of the stochastic sample path starting in  $(-L, z_0)$  hitting  $F_+$  at a point  $(L, z_1)$  such that  $z_1 \geq c_1 \tilde{\mu}^{1-\gamma}$  for some  $c_1 \in (0, c_0)$  is bounded below by

$$1 - \frac{C}{\tilde{\mu}^{2\gamma}} \exp\left\{-\kappa_2 \frac{\tilde{\mu}^2}{\tilde{\sigma}^2}\right\} \quad (5.67)$$

for some  $\kappa_2 > 0$ .

We can then apply Proposition 5.4 with  $\beta = 1 - \gamma$  to show that the sample path starting in  $(L, z_1) \in F_+$  hits  $F_-$  in a point  $(-L, z_2)$ , where (5.40) with  $h = \tilde{\mu}^{1-\beta}$  shows that  $z_2 \geq z_1$  with probability greater or equal than

$$1 - 2 \exp\left\{-\kappa_3 \frac{\tilde{\mu}^2}{\tilde{\sigma}^2}\right\} \quad (5.68)$$

for some  $\kappa_3 > 0$ . This shows that the set  $A = \{(\xi, z) \in F_- : z \geq z_0\}$  is invariant with high probability. Indeed, we have shown that this is true for initial  $z$  of order  $\tilde{\mu}^{1-\gamma}$ , and sample paths starting at some higher  $z$  cannot arrive below without crossing the domain where we control the dynamics.

Finally, the prefactor  $C\tilde{\mu}^{-2\gamma}$  in (5.67) can be absorbed in the exponent, provided  $\tilde{\sigma}^2 \leq \tilde{\mu}^2/|\log \tilde{\mu}|$ .  $\square$

## References

- [BAKS84] Gérard Ben Arous, Shigeo Kusuoka, and Daniel W. Stroock, *The Poisson kernel for certain degenerate elliptic operators*, J. Funct. Anal. **56** (1984), no. 2, 171–209.
- [BE86] S.M. Baer and T. Erneux, *Singular Hopf bifurcation to relaxation oscillations I*, SIAM J. Appl. Math. **46** (1986), no. 5, 721–739.
- [BE92] ———, *Singular Hopf bifurcation to relaxation oscillations II*, SIAM J. Appl. Math. **52** (1992), no. 6, 1651–1664.
- [BG02] Nils Berglund and Barbara Gentz, *Pathwise description of dynamic pitchfork bifurcations with additive noise*, Probab. Theory Related Fields **122** (2002), no. 3, 341–388.
- [BG09] ———, *Stochastic dynamic bifurcations and excitability*, Stochastic Methods in Neuroscience (Carlo Laing and Gabriel Lord, eds.), Oxford University Press, 2009, pp. 64–93.
- [BGK10] Nils Berglund, Barbara Gentz, and Christian Kuehn, *Hunting french ducks in a noisy environment*, arXiv:1011.3193, 2010.
- [Bir57] Garrett Birkhoff, *Extensions of Jentzsch’s theorem*, Trans. Amer. Math. Soc. **85** (1957), 219–227.
- [BKLLC11] Peter Borowski, Rachel Kuske, Yue-Xian Li, and Juan Luis Cabrera, *Characterizing mixed mode oscillations shaped by noise and bifurcation structure*, Chaos **20** (2011), no. 4, 043117.
- [Bra98] B. Braaksma, *Singular Hopf bifurcation in systems with fast and slow variables*, Journal of Nonlinear Science **8** (1998), no. 5, 457–490.
- [Dah77] Björn E. J. Dahlberg, *Estimates of harmonic measure*, Arch. Rational Mech. Anal. **65** (1977), no. 3, 275–288.
- [DGK<sup>+</sup>11] M. Desroches, J. Guckenheimer, C. Kuehn, B. Krauskopf, H. Osinga, and M. Wechselberger, *Mixed-mode oscillations with multiple time scales*, SIAM Review, in press (2011).
- [DT] Catherine Doss and Michèle Thieullen, *Oscillations and random perturbations of a FitzHugh-Nagumo system*, Preprint hal-00395284 (2009).
- [Fit55] R. FitzHugh, *Mathematical models of threshold phenomena in the nerve membrane*, Bull. Math. Biophysics **17** (1955), 257–269.

- [Fit61] R. FitzHugh, *Impulses and physiological states in models of nerve membrane*, Biophys. J. **1** (1961), 445–466.
- [HH52] A. L. Hodgkin and A. F. Huxley, *A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes*, J. Physiol. (Lond.) **117** (1952), 500–544.
- [HM09] Pawel Hitczenko and Georgi S. Medvedev, *Bursting oscillations induced by small noise*, SIAM J. Appl. Math. **69** (2009), no. 5, 1359–1392.
- [Izh00] Eugene M. Izhikevich, *Neural excitability, spiking and bursting*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **10** (2000), no. 6, 1171–1266.
- [Jen12] Robert Jentzsch, *Über Integralgleichungen mit positivem Kern*, J. f. d. reine und angew. Math. **141** (1912), 235–244.
- [KP03] Efstratios K. Kosmidis and K. Pakdaman, *An analysis of the reliability phenomenon in the FitzHugh–Nagumo model*, J. Comput. Neuroscience **14** (2003), 5–22.
- [KP06] ———, *Stochastic chaos in a neuronal model*, Internat. J. Bifur. Chaos **16** (2006), no. 2, 395–410.
- [KR50] M. G. Kreĭn and M. A. Rutman, *Linear operators leaving invariant a cone in a Banach space*, Amer. Math. Soc. Translation **1950** (1950), no. 26, 128.
- [LSG99] Benjamin Lindner and Lutz Schimansky-Geier, *Analytical approach to the stochastic FitzHugh–Nagumo system and coherence resonance*, Physical Review E **60** (1999), no. 6, 7270–7276.
- [MVE08] Cyrill B. Muratov and Eric Vanden-Eijnden, *Noise-induced mixed-mode oscillations in a relaxation oscillator near the onset of a limit cycle*, Chaos **18** (2008), 015111.
- [MVVE05] C.B. Muratov, E. Vanden-Eijnden, and W. E. Skaggs, *Self-induced stochastic resonance in excitable systems*, Physica D **210** (2005), 227–240.
- [NAY62] J. S. Nagumo, S. Arimoto, and S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proc. IRE **50** (1962), 2061–2070.
- [Num84] Esa Nummelin, *General irreducible Markov chains and nonnegative operators*, Cambridge Tracts in Mathematics, vol. 83, Cambridge University Press, Cambridge, 1984.
- [Ore71] Steven Orey, *Lecture notes on limit theorems for Markov chain transition probabilities*, Van Nostrand Reinhold Co., London, 1971, Van Nostrand Reinhold Mathematical Studies, No. 34.
- [Row07] Peter Rowat, *Interspike interval statistics in the stochastic Hodgkin–Huxley model: Coexistence of gamma frequency bursts and highly irregular firing*, Neural Computation **19** (2007), 12151250.
- [SK] David W. J. Simpson and Rachel Kuske, *Mixed-mode oscillations in a stochastic, piecewise-linear system*, arXiv:1010.1504 (2010).
- [Sow08] Richard B. Sowers, *Random perturbations of canards*, J. Theoret. Probab. **21** (2008), no. 4, 824–889.
- [SVJ66] E. Seneta and D. Vere-Jones, *On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states*, J. Appl. Probability **3** (1966), 403–434.
- [TGOS08] Marc Turcotte, Jordi Garcia-Ojalvo, and Gurol M. Suel, *A genetic timer through noise-induced stabilization of an unstable state*, PNAS **105** (2008), no. 41, 1573215737.
- [TP01a] Seiji Tanabe and K. Pakdaman, *Dynamics of moments of FitzHugh–Nagumo neuronal models and stochastic bifurcations*, Phys. Rev. E **63** (2001), 031911.

- [TP01b] Seiji Tanabe and K. Pakdaman, *Noise-induced transition in excitable neuron models*, Biol. Cybern. **85** (2001), 269–280.
- [TTP02] Takayuki Takahata, Seiji Tanabe, and K. Pakdaman, *White-noise stimulation of the Hodgkin-Huxley model*, Biol. Cybern. **86** (2002), 403417.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Model and results</b>	<b>3</b>
2.1	Definition of the model . . . . .	3
2.2	The random variable “Number of SAOs” . . . . .	5
2.3	General results on the distribution of SAOs . . . . .	7
2.4	The weak-noise regime . . . . .	8
2.5	The transition from weak to strong noise . . . . .	10
<b>3</b>	<b>Substochastic Markov kernels</b>	<b>11</b>
3.1	General results . . . . .	11
3.2	Bounds . . . . .	14
3.3	Proof of Theorem 2.2 . . . . .	15
<b>4</b>	<b>Dynamics near the separatrix</b>	<b>16</b>
4.1	The linearised process . . . . .	16
4.2	The nonlinear equation . . . . .	19
4.3	Computation of the kernel . . . . .	21
<b>5</b>	<b>Dynamics during an SAO</b>	<b>21</b>
5.1	Action–angle-type variables . . . . .	21
5.2	Averaging . . . . .	24
5.3	Computation of the kernel . . . . .	26
5.4	Proof of Theorem 2.4 . . . . .	29

Nils Berglund and Damien Landon  
 Université d’Orléans, Laboratoire MAPMO  
 CNRS, UMR 6628  
 Fédération Denis Poisson, FR 2964  
 Bâtiment de Mathématiques, B.P. 6759  
 45067 Orléans Cedex 2, France  
*E-mail address:* [nils.berglund@univ-orleans.fr](mailto:nils.berglund@univ-orleans.fr), [damien.landon@univ-orleans.fr](mailto:damien.landon@univ-orleans.fr)