FLUID LIMIT THEOREMS FOR STOCHASTIC HYBRID SYSTEMS WITH APPLICATION TO NEURON MODELS

K. PAKDAMAN,* *Université Paris VII*

M. THIEULLEN,** *Université Paris VI*

G. WAINRIB,*** *Ecole Polytechnique, Université Paris VII and Université Paris VI*

Abstract

In this paper we establish limit theorems for a class of stochastic hybrid systems (continuous deterministic dynamics coupled with jump Markov processes) in the fluid limit (small jumps at high frequency), thus extending known results for jump Markov processes. We prove a functional law of large numbers with exponential convergence speed, derive a diffusion approximation, and establish a functional central limit theorem. We apply these results to neuron models with stochastic ion channels, as the number of channels goes to infinity, estimating the convergence to the deterministic model. In terms of neural coding, we apply our central limit theorems to numerically estimate the impact of channel noise both on frequency and spike timing coding.

Keywords: Stochastic hybrid system; piecewise-deterministic Markov process; fluid limit; Langevin approximation; kinetic model; neuron model; Hodgkin–Huxley; Morris–Lecar; stochastic ion channels

2010 Mathematics Subject Classification: Primary 60F05; 60F17; 60J75
Secondary 92C20; 92C45

1. Introduction

In this paper we consider stochastic hybrid systems where a continuous deterministic dynamic is coupled with a jump Markov process. Such systems were introduced in [6] as piecewise-deterministic Markov processes. They have been subsequently generalized to cover a wide range of applications: communication networks, biochemistry, and, more recently, DNA replication modelling [2], [14], [20], [23]. We are interested in the fluid limit for these systems, considering the case of small jumps of size $1/N$ at high frequency N, with a view towards application to neural modelling.

The general class of models we consider is described in Section 2.1, and, for the sake of clarity, we describe here a simple example which retains the main features. Consider a population of N independent individuals, each of them being described by a jump Markov process $u_k(t)$ for $k = 1, \ldots, N$ with states 0 and 1, and with identical transition rates $\alpha > 0$

and $\beta > 0$ as follows:

$$0 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} 1.$$

As an empirical measure, we define the proportion of individuals in state 1 at time t by

$$e_N(t) = \frac{1}{N} \sum_{k=1}^{N} u_k(t).$$

The model becomes hybrid when we assume a global coupling through a variable $V_N \in \mathbb{R}$, in the sense that the rates $\alpha(V_N)$ and $\beta(V_N)$ are functions of V_N. The variable V_N is itself a solution of a differential equation, between the jumps of $e_N(t)$:

$$\frac{dV_N}{dt} = f(V_N, e_N),$$

where $f : \mathbb{R}^2 \to \mathbb{R}$. In the general case, this model is extended with more general non-autonomous jump Markov processes, the global variable can be vector valued, and the transition rates can be functions of the empirical measure (Section 2.1).

We prove convergence in probability on finite time intervals, with techniques inspired by [1], of the solution X_N of the stochastic hybrid system to a deterministic limit $X = (v, g)$. For the example above, X is the solution to

$$\frac{dv}{dt} = f(v, g), \qquad \frac{dg}{dt} = (1 - g)\alpha(v) - g\beta(v).$$

We derive a diffusion approximation and prove a functional central limit theorem that helps characterize the fluctuations of both the discrete and continuous variables around the deterministic solution. We find that these fluctuations are a Gaussian process which corresponds to the asymptotic law of the linearized diffusion approximation. We further obtain an exponential speed of convergence which relates the tail distribution of the error $E_N(T) = \sup_{[0,T]} |X_N - X|^2$ to the size parameter N and the time window T: for $\Delta > 0$ and large N,

$$P[E_N(T) > \Delta] \le e^{-\Delta N H(T)}. \tag{1.1}$$

Thus, the convergence result can be extended to large time intervals $[0, T(N)]$, provided that $T = T(N)$ is such that $NH(T(N)) \to \infty$. Inequality (1.1) is a new result which provides an estimate of the number, N, of individuals required to reach a given level of precision. This number increases with the time scale on which one wants this precision to be achieved. For systems subject to finite-size stochasticity, sometimes called demographic stochasticity, it provides a relation between the reliability time scale and the population size N. There are other ways of obtaining a law of large numbers, for example using the convergence of the master equation or of the generators [9], [38]. We want to highlight here that our proof is based on exponential inequalities for martingales. Other ways of obtaining a law of large numbers would not be likely to provide an estimate such as (1.1).

Our mathematical reference on the fluid limit is the seminal paper [22], which contains a law of large numbers and a central limit theorem for sequences of classical jump Markov processes. Recently, a spatially extended version of a stochastic hybrid system has been considered in [1],

for a standard neuron model. The author showed convergence in probability up to finite time windows to a deterministic fluid limit expressed in terms of a partial differential equation coupled with ordinary differential equations. In the present paper, we consider a class of nonspatial models which includes multi compartmental models, by increasing the dimension. We extend the results of [22] to stochastic hybrid models at the fluid limit.

Neurons are subject to various sources of fluctuations, intrinsic (membrane noise) and extrinsic (synaptic noise). Clarifying the impact of noise and variability in the nervous system is an active field of research [10], [31]. The intrinsic fluctuations in single neurons are mainly caused by ion channels, also called channel noise, whose impact and putative functions have been intensively investigated [30], [32], [40], mainly by numerical simulations. Our motivation is to study the intrinsic fluctuations in neuron models, and we think that stochastic hybrid systems are a natural tool for this purpose. Ion channels open and close, through voltage induced electromagnetic conformational change, thus enabling ion transfer and action potential generation. Because of thermal noise, one of the main features of these channels is their stochastic behaviour.

Neurons encode incoming signals into trains of stereotyped pulses referred to as action potentials (APs). It is the mean firing frequency, that is, the number of APs within a given time window, and the timing of the APs that are the main conveyors of information in nervous systems. Channel noise due to the seemingly random fluctuations in the opening and closing times of transmembrane ion channels induces jitter in the AP timing and, consequently, in the mean firing frequency as well. In terms of modelling, our starting point is the stochastic interpretation of the Hodgkin–Huxley formalism [15]. In this setting, ion channels are usually modelled with independent Markov jump processes, whose transition rates can be estimated experimentally [37]. These stochastic discrete models are coupled with a continuous dynamic for the membrane potential, leading to a piecewise-deterministic Markov process. Thus, the individuals are the ion channels and the global variable V_N the voltage potential (cf. Section 3).

Deterministic hybrid kinetic equations appear to be a common formalism suitable for each stage of nervous system modelling, as shown in [8]. This latter study provides us with a framework to introduce stochastic hybrid processes to model action potential generation and synaptic transmission, as stochastic versions of deterministic kinetic models coupled with differential equations through the transition rates.

On the side of neuron modelling applications, the limit behaviour of a similar but less general model is considered in [11], using an asymptotic development of the master equation as $N \to \infty$, which formally leads to a deterministic limit and a Fokker–Planck equation (Langevin approximation), providing the computation of the diffusion coefficients. The Langevin approximation is also studied in [36], but in a simplified case where the transition rates are constants (independent of V_N), which is actually the case studied in [22]. Our mathematical results extend these previous studies to a wider class of models (if we put aside the spatial aspects in [1]), providing a rigorous approach for the Langevin approximation, and establishing a central limit theorem which describes the effect of channel noise on the membrane potential [34]. The convergence speed provides a quantitative insight into the following question: if a neuron needs to be reliable during a given time scale, what would be a sufficient number of ion channels? We thus provide a mathematical foundation for the study of stochastic neuron models, and we apply our results to standard models, quantifying the effect of noise on neural coding. In particular, both frequency coding (Section 3.5.1) and spike timing coding (Section 3.5.2) are numerically studied using the Morris–Lecar neuron model with a large number of stochastic ion channels.

Generally speaking, stochastic hybrid models in the fluid limit would arise in multiscale systems with a large population of stochastic agents coupled through a global variable, leading to an emergent cooperative behaviour. Starting from a microscopic description (for instance, ion channels), the central limit theorem as stated in the present paper leads to a description of the fluctuations of the global variable (the membrane potential). So, in the perspective of applications, it would be interesting to investigate how our framework and results could be developed in fields other than neural modelling, for instance in chemical kinetics, population dynamics, tumor modelling, economics, or opinion dynamics theory. In a more mathematical perspective, it would be interesting to consider a wider class of models, for instance by including spatial aspects as in [1] or by weakening the independence assumption. Other questions could be investigated, for instance those concerning escape problems, first passage times, and large deviations, whenever N is large or not.

Our paper is organized as follows. In Section 2 we define our model and formulate the main results. In Section 3 we apply our results to neuron models. In Section 4 we give the proof of the law of large numbers and its convergence speed (Theorem 2.1), and in Section 5 we give the proof of the Langevin approximation and the central limit theorems (Theorems 2.2–2.5).

2. Model and main results

This section contains the definition of our general model and states the main theorems.

2.1. Model

Model 2.1. (Stochastic hybrid model.) *Let $p, q, N \in \mathbb{N}^+$ and $r_j \in \mathbb{N}^*$ for all $1 \leq j \leq q$. Let $d = \sum_{j=1}^q r_j$. We define the stochastic hybrid model (S_N) to be the model whose solution,*

$$X_N(t) = (V_N(t), \boldsymbol{e}_N(t)) \in \mathbb{R}^p \times \mathbb{R}^d, \qquad t \geq 0,$$

satisfies

$$\frac{\mathrm{d}V_N}{\mathrm{d}t} = f(X_N),$$

where $\boldsymbol{e}_N = (e_N^{(1)}, \ldots, e_N^{(q)})$, $e_N^{(j)} \in \mathbb{R}^{r_j}$, and the processes $e_N^{(j)}(t)$ are q independent jump Markov processes. Note that the differential equation for V_N holds only between the jump times of the process $\boldsymbol{e}_N$, with updated initial conditions.

Let us introduce some notation. Each process $e_N^{(j)}$ represents an empirical measure for a population j of jump processes: the state dependent transition rates for each of these processes in population j are denoted by $\alpha_{k,l}^{(j)} \colon \mathbb{R}^{p+d} \to \mathbb{R}_+^*$. Moreover, if $\boldsymbol{e} = (e^{(1)}, \ldots, e^{(q)}) \in \mathbb{R}^d$, we denote the kth element of $e^{(j)}$ by $\{e^{(j)}\}_k$.

Then, for $1 \leq j \leq q$, processes $e_N^{(j)}(t)$ are characterized by

- *their state space: $E_N^{(j)} = \{(x_1, \ldots, x_{r_j}) \in \{0, 1/N, \ldots, 1\}^{r_j} \mid \sum_{k=1}^{r_j} x_k = 1\}$;*

- *their intensity $\lambda_N^{(j)}$: for $X = (V, \boldsymbol{e}) \in \mathbb{R}^p \times \mathbb{R}^d$, $\lambda_N^{(j)}(X) = N\tilde{\lambda}^{(j)}(X)$ with*

$$\tilde{\lambda}^{(j)}(X) = \sum_{k=1}^{r_j} \{e^{(j)}\}_k \sum_{l=1, l \neq k}^{r_j} \alpha_{k,l}^{(j)}(X);$$

- *their jump law $\mu_N^{(j)}$: we define $u_a^{(j)} = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^{r_j}$ and $u_{a,b}^{(j)} = u_a^{(j)} - u_b^{(j)}$ for $1 \leq a, b \leq r_j$. The transition of an individual agent in subpopulation j from*

one state a to another state b corresponds to a jump of $z = u_{b,a}^{(j)}/N$ for the process $e_N^{(j)}$. Thus, we define

$$X + \frac{1}{N}\Delta X_{a,b}^j = (V, e^{(1)}, \ldots, \tilde{e}^{(j)}, \ldots, e^{(q)}), \qquad \tilde{e}^{(j)} = e^{(j)} + \frac{1}{N}u_{a,b}^{(j)}.$$

So that the jump law for a jump of z in subpopulation j is given by

$$\mu_N^{(j)}(X, z) = \frac{\{e^{(j)}\}_a \alpha_{a,b}^{(j)}(X)}{\tilde{\lambda}^{(j)}(X)} \quad \text{if } z = \frac{1}{N}u_{b,a}^{(j)}$$

for all $1 \le a, b \le r_j$ such that $\{e^{(j)}\}_a \neq 0$ and $\{e^{(j)}\}_b \neq 1$, and

$$\mu_N^{(j)}(X, z) = 0 \quad \text{otherwise.}$$

For a more formal definition, we refer the reader to [6].

For $1 \le k \le r_j$, the kth component $\{e_N^{(j)}\}_k$ of vector $e_N^{(j)}$ can be interpreted as the proportion of agents of type j which are in state k in a population of size N.

For the presentation of the results to be clearer, we assume in the stochastic model (S_N) that each population is of the same size N. However, the results can be extended to the case where all the population sizes are proportional to a large number $A \to \infty$, which is the case for the application to neuron models (see Section 3).

We show in Theorem 2.1, below, that this stochastic hybrid model (S_N) converges as $N \to \infty$ to the following deterministic model (D).

Model 2.2. (Deterministic model.) *We define the deterministic model to be the model whose solution $X = (v, \boldsymbol{g}) \in \mathbb{R}^p \times \mathbb{R}^d$ with $\boldsymbol{g} = (\boldsymbol{g}^{(1)}, \ldots, \boldsymbol{g}^{(q)})$ satisfies*

$$\dot{v} = f(v, \boldsymbol{g}), \qquad \dot{\boldsymbol{g}}_k^{(j)} = \sum_{1 \le i \le r_j, \, i \neq k} \alpha_{i,k}^{(j)}(X)\boldsymbol{g}_i^{(j)} - \alpha_{k,i}^{(j)}(X)\boldsymbol{g}_k^{(j)} \tag{D}$$

for all $1 \le j \le q$ and all $1 \le k \le r_j$. The first equation is the same as in the stochastic model (deterministic part) and the second equation corresponds to the usual rate equation, with a gain term and a loss term. We define $F \colon \mathbb{R}^p \times \mathbb{R}^d \to \mathbb{R}^p \times \mathbb{R}^d$ such that (D) might be written $\dot{X} = F(X)$.

Example 2.1. We now describe an example of the stochastic hybrid model in a simple setting motivated by applications. This is the setting we will use in Sections 4 and 5 in order to make the arguments clearer. We consider the case where $p = q = 1$ and $r_1 = 2$. For $N \in \mathbb{N}^*$, we can construct a stochastic hybrid process as follows: first let us introduce a collection of N independent jump Markov processes $u^{(k)}$ for $1 \le k \le N$ such that $u_t^{(k)} \colon 0 \to 1$ with rate $\alpha(V_N)$ and $u_t^{(k)} \colon 1 \to 0$ with rate $\beta(V_N)$:

$$\alpha(V_N)$$
$$0 \; \rightleftarrows \; 1,$$
$$\beta(V_N)$$

where V_N is defined below. In this case, the stochastic hybrid model (S_N) can be written as

$$\dot{V}_N(t) = f(V_N(t), e_N(t)), \qquad e_N(t) = \left(\frac{1}{N}\sum_{k=1}^N \delta_0(u_t^{(k)}), \frac{1}{N}\sum_{k=1}^N \delta_1(u_t^{(k)})\right),$$

$$V_N(0) = v_0, \qquad e_N(0) = (u_0, 1 - u_0).$$

Note that the components of $e_N(t)$ are the proportions of processes $u^{(k)}$ in states 0 and 1, and if we define $u_N(t) = (1/N) \sum_{k=1}^{N} \delta_1(u_t^{(k)})$ then we have $e_N(t) = (1 - u_N(t), u_N(t))$, so that the solution is determined by the pair $X_N(t) = (V_N(t), u_N(t))$. Thus, each element of the sequence of jump Markov processes $\{u_N\}_{N \geq 1}$ is characterized by

- its state space $E_N = \{0, 1/N, 2/N, \ldots, 1\}$;

- its intensity $\lambda_N(V_N(t), u) = N[u\beta(V_N(t)) + (1 - u)\alpha(V_N(t))]$; this intensity is time dependent through $V_N(t)$;

- its jump law

$$\mu_N(V_N(t), u, y) = \mu^+(V_N(t), u)\delta_{y, u+1/N} + \mu^-(V_N(t), u)\delta_{y, u-1/N},$$

where

$$\mu^+(V, u) = \frac{(1 - u)\alpha(V)}{u\beta(V) + (1 - u)\alpha(V)} \quad \text{and} \quad \mu^-(V, u) = \frac{u\beta(V)}{u\beta(V) + (1 - u)\alpha(V)}.$$

This jump law is also time dependent through $V_N(t)$.

The deterministic system (D) takes the form

$$\dot{v}(t) = f(v(t), g(t)), \qquad \dot{g}(t) = (1 - g(t))\alpha(v(t)) - g(t)\beta(v(t)),$$
$$v(0) = v_0, \qquad g(0) = u_0.$$

Fluid limit assumption. In the sequel we will be interested in the asymptotic behaviour of the stochastic hybrid models (S_N) under the fluid limit assumption. Let us now recall what this assumption means. Let (W_N) be a sequence of homogeneous Markov jump processes with state spaces $E_N \subset \mathbb{R}^k$, intensities $\lambda_N(w)$, and jump law $\mu_N(w, dy)$. Define the flow as $F_N(w) = \lambda_N(w) \int_{E_N} (z - w)\mu_N(w, dz)$. The fluid limit occurs if the flow admits a limit and if the second-order moment of the jump size converges to 0 when $N \to \infty$. Our stochastic hybrid models (S_N) are in the fluid limit since the jumps are of size $1/N$ and the intensity is proportional to N. As mentioned in Example 2.1, such a setting appears when considering proportions in a population of independent agents. However, this independence between agents is not necessary to satisfy the fluid limit assumption.

2.2. Law of large numbers for stochastic hybrid systems

We state here the first result concerning the convergence of the stochastic hybrid model (S_N), which is a functional law of large numbers on finite time windows and associated convergence rate.

Theorem 2.1. *Let $\varepsilon > 0$, $\delta > 0$, and $T > 0$. Assume that the functions $\alpha_{i,j}$ and f are C^1, and that they satisfy the following condition:*

(H1) *the solution v of (D) is bounded for any initial value on $[0, T]$ and, for all $N \geq 1$, the solution $V_N(t)$ of (S_N) is uniformly bounded in N on $[0, T]$.*

Let X^0 be a given initial condition for (D), and let $X = (v, \boldsymbol{g})$ be the corresponding solution. Then there exist initial conditions $(X_N^0)_{N \geq 1}$ for (S_N) and $N_0 \geq 0$ such that, for all $N \geq N_0$, the solution $X_N = (V_N, e_N^{(1)}, \ldots, e_N^{(q)})$ satisfies, for all $1 \leq j \leq q$,

$$\mathrm{P}\left[\sup_{0 \leq t \leq T} \|V_N(t) - v(t)\| \geq \delta \right] \leq \varepsilon,$$

$$\mathrm{P}\left[\sup_{0\leq t\leq T}\|e_N^{(j)}(t) - g^{(j)}(t)\| \geq \delta\right] \leq \varepsilon.$$

Moreover, if we define

$$P_N(T, \Delta) := \mathrm{P}\left[\sup_{0\leq t\leq T}\|V_N(t) - v(t)\|^2 + \sum_{j=1}^{q}\|e_N^{(j)}(t) - g^{(j)}(t)\|^2 > \Delta\right],$$

there exist two constants $B(T) > 0$ and $C > 0$ such that, for sufficiently small Δ,

$$\limsup_{N\to\infty}\frac{1}{N}\log P_N(T, \Delta) \leq -\frac{\Delta e^{-B(T)T}}{CT}. \tag{2.1}$$

Moreover, if

(H2) *assumption (H1) holds on $[0, +\infty)$*

then the constant $B(T) = BT$ is proportional to T.

Interpretation of the convergence speed. We have obtained in (2.1) an upper bound for the convergence speed which can help to answer the following issue. Given a number of channels N, and given an error Δ and a confidence probability $1 - p$ (e.g. $p = 0.01$), the time window $[0, T]$ for which we can be sure (up to probability $1 - p$) that the distance between the stochastic and the deterministic solutions (starting at the same point) is less than Δ is given by (2.1). In Section 3.3 we present numerical simulation results for the stochastic Hodgkin–Huxley model illustrating the obtained bound for the convergence speed.

Remark 2.1. Assumption (H2) and, thus, (H1) is satisfied for most neuron models, for instance for the Hodgkin–Huxley model [4].

2.3. Functional central limit theorem and the exit problem

Let $X_N = (V_N, e_N^{(1)}, \ldots, e_N^{(q)})$ be the solution to the stochastic model (S_N), and let $X = (v, g^{(1)}, \ldots, g^{(q)})$ be the solution to the deterministic system (D) with identical initial conditions $X(0) = X_N(0) = X_0 \in \mathbb{R}^{p+d}$. Before stating the theorem, we need to introduce some further notation. For $X = (v, e) \in \mathbb{R}^p \times \mathbb{R}^d$, and $1 \leq j \leq q$, $1 \leq i, k \leq r_j$, let

$$b_k^{(j)}(X) = \sum_{1\leq i\leq r_j, \, i\neq k}\alpha_{i,k}^{(j)}(X)e_i^{(j)} - \alpha_{k,i}^{(j)}(X)e_k^{(j)},$$

$$H_{i,k}^{(j)} = \alpha_{i,k}^{(j)}(X)e_i^{(j)} + \alpha_{k,i}^{(j)}(X)e_k^{(j)}, \qquad \lambda_k^{(j)}(X) = \sum_{1\leq i\leq r_j, \, i\neq k}H_{i,k}^{(j)}.$$

We then define a matrix $G^{(j)}(X)$ by, for $1 \leq k, l \leq r_j$,

$$G_{k,k}^{(j)}(X) = \lambda_k^{(j)}(X), \qquad G_{k,l}^{(j)}(X) = H_{k,l}^{(j)}(X) = G_{l,k}^{(j)}(X), \qquad l \neq k.$$

Consider the $(p + d)$-dimensional processes

$$Z_N = \begin{pmatrix} Y_N \\ P_N^{(1)} \\ \vdots \\ P_N^{(q)} \end{pmatrix} := \sqrt{N}\begin{pmatrix} V_N - v \\ e_N^{(1)} - g^{(1)} \\ \vdots \\ e_N^{(q)} - g^{(q)} \end{pmatrix}.$$

With this setting, we have the following result.

Theorem 2.2. *Under the same hypotheses as in Theorem 2.1, the process Z_N converges in law, as $N \to \infty$, to the process*

$$
Z = \begin{pmatrix} Y \\ \boldsymbol{P}^{(1)} \\ \vdots \\ \boldsymbol{P}^{(q)} \end{pmatrix},
$$

whose characteristic function $\Psi(t, \theta) = \mathrm{E}[\mathrm{e}^{\mathrm{i}\langle \theta, Z(t) \rangle}]$ satisfies

$$
\frac{\partial \Psi}{\partial t} = \sum_{j=1}^{q} \left\{ \sum_{l \in L} \sum_{k=1}^{r_j} \theta_k^{(j)} \frac{\partial b_k^{(j)}}{\partial x_l} \frac{\partial \Psi}{\partial \theta_l} - \frac{1}{2} \sum_{k,l=1}^{r_j} \theta_k^{(j)} \theta_l^{(j)} G_{k,l}^{(j)} \Psi \right\} + \sum_{m=1}^{p} \sum_{l \in L} \theta_m \frac{\partial f^m}{\partial x_l} \frac{\partial \Psi}{\partial \theta_l},
$$

where $G_{k,l}^{(j)}$, $\partial f^m / \partial x_l$ and $\partial b_k^{(j)} / \partial x_l$ are evaluated at $X(t)$ for $1 \leq m \leq p$, $1 \leq j \leq q$, $1 \leq k \leq r_j$, and $l \in L$, $\theta = ((\theta_m)_{1 \leq m \leq p}, (\theta_k^{(j)})_{1 \leq j \leq q}, 1 \leq k \leq r_j) = (\theta_l)_{l \in L}$, and $L = \{(m)_{1 \leq m \leq p}, (j, k)_{1 \leq j \leq q, 1 \leq k \leq r_j}\}$.

In the next theorem we provide a characterization of the fluctuations of the first exit time and location for the stochastic hybrid process X_N, using the central limit theorem, Theorem 2.2. Let $\phi \colon \mathbb{R}^{p+d} \to \mathbb{R}$ be continuously differentiable. Define

$$
\tau_N := \inf\{t \geq 0; \phi(X_N(t)) \leq 0\} \quad \text{and} \quad \tau := \inf\{t \geq 0; \phi(X(t)) \leq 0\}
$$

to be the first passage times through $\phi = 0$ for the stochastic hybrid process and its deterministic limit, respectively. Recall that F defines the deterministic model (D): $\dot{X} = F(X)$.

Theorem 2.3. *Assume that the initial condition $X(0)$ satisfies $\phi(X(0)) > 0$. Suppose that $\tau < \infty$ and $\nabla \phi(X(\tau)) F(X(\tau)) < 0$. Define the random variable*

$$
\pi(\tau) := -\frac{\nabla \phi(X(\tau)) Z(\tau)}{\nabla \phi(X(\tau)) F(X(\tau))},
$$

where Z is the limit of Z_N as obtained in Theorem 2.2. Then the following convergences in distribution hold when $N \to \infty$:

$$
\sqrt{N}(\tau_N - \tau) \to \pi(\tau), \qquad \sqrt{N}(X_N(\tau_N) - X(\tau)) \to Z(\tau) + \pi(\tau) F(X(\tau)).
$$

2.4. Langevin approximation

Our second result is a central limit theorem that provides a way to build a diffusion or Langevin approximation of the solution of the stochastic hybrid system (S_N).

As before, $X_N(t) = (V_N(t), \boldsymbol{e}_N(t)) \in \mathbb{R}^p \times \mathbb{R}^d$ is the solution to the stochastic hybrid model (S_N).

Let $R_N(t) = \{(R_N^{(j)})_k(t)\}_{1 \leq j \leq q, 1 \leq k \leq r_j}$ with $R_N^{(j)} \in \mathbb{R}^{r_j}$ be defined as

$$
R_N^{(j)}(t) = \sqrt{N} \left(\boldsymbol{e}_N^{(j)}(t) - \boldsymbol{e}_N^{(j)}(0) - \int_0^t b^{(j)}(X_N(s)) \, \mathrm{d}s \right),
$$

where $b^{(j)}$ denotes the vector $(b_k^{(j)})_{1 \leq k \leq r_j} \in \mathbb{R}^{r_j}$ for fixed j.

Theorem 2.4. *Under the same hypotheses of Theorem 2.1, the process R_N converges in law, as $N \to \infty$, to the process $R = \{(R^{(j)})_k(t)\}_{1 \le j \le q}$, $1 \le k \le r_j$, with*

$$R^{(j)}(t) = \int_0^t \sigma^{(j)}(X(s)) \, dW_s^j,$$

where

- *$X = (v, g)$ is the solution of the deterministic model (D) with initial conditions $X(0) = X_N(0) = X_0$;*
- *the W^j are independent standard r_j-dimensional Brownian motions;*
- *$\sigma^{(j)}(X)$ is the square root of matrix $G^{(j)}(X)$ (as defined in Section 2.3).*

This theorem leads to the following degenerate diffusion approximation $\tilde{X}_N = (\tilde{V}_N, \tilde{g}_N) \in \mathbb{R}^p \times \mathbb{R}^d$ for sufficiently large N:

$$d\tilde{V}_N = f(\tilde{X}_N(t)) \, dt, \qquad d\tilde{g}_N^{(j)} = b^{(j)}(\tilde{X}_N(t)) \, dt + \frac{1}{\sqrt{N}} \sigma^{(j)}(\tilde{X}_N(t)) \, dW_t^j. \qquad (2.2)$$

Note that this approximation may not have the same properties as the original process, even in the limit $N \to \infty$ when considering for instance large deviations (cf. [26]).

Linearization. Let us define $\tilde{Z}_N := (\tilde{Y}_N, \tilde{P}_N) := \sqrt{N}(\tilde{X}_N - X) \in \mathbb{R}^p \times \mathbb{R}^d$, where we recall that $\tilde{X}_N$ is the Langevin approximation defined in (2.2). Thus, for $1 \le j \le q$ and $1 \le k \le r_j$,

$$d\tilde{Y}_N = \sqrt{N}(f(\tilde{X}_N) - f(X)) \, dt,$$

$$d\tilde{P}_N^{(j)} = \sqrt{N}(b^{(j)}(\tilde{X}_N) - b^{(j)}(X)) \, dt + \sigma^{(j)}(\tilde{X}_N) \, dW_t^j.$$

As an asymptotic linearization of $\tilde{Z}_N$, we define the diffusion process $\Theta = (\gamma, \rho) \in \mathbb{R}^p \times \mathbb{R}^d$ by

$$d\gamma_m = \sum_{l \in L} \frac{\partial f^m}{\partial x_l} \Theta_l \, dt, \qquad d\rho^{(j)} = \sum_{l \in L} \frac{\partial b^{(j)}}{\partial x_l} \Theta_l \, dt + \sigma^{(j)}(X) \, dW_t^{(j)}.$$

The following theorem states that the central limit fluctuations are given by the linearization of the Langevin approximation around the deterministic solution.

Theorem 2.5. *The processes Z (cf. Theorem 2.2) and Θ have the same law.*

3. Application to neuron models

In this section we show how our previous theorems can be applied to standard neuron models, taking into account ion channel stochasticity.

3.1. Kinetic formalism in neuron modelling

Kinetic models can be used in many parts of nervous system modelling, such as in ion channel kinetics, synapse, and neurotransmitters release modelling [8]. As already mentioned in the introduction, compared to our general formalism, the stochastic individuals are the ion channels and the global variable V_N the voltage potential. Constituted of several subunits called gates, voltage-gated ion channels are metastable molecular devices that can open and close. There exist different types of channel according to the kind of ions, and they are distributed

within the neuron membrane (soma, axon hillock, nodes of Ranvier, dendritic spines) with heterogeneous densities.

In what follows, we consider the model of Hodgkin and Huxley, which has been extended in different ways to include stochastic ion channels. In numerical studies, different versions have been used, from a *two-state gating* interpretation (see, e.g. [33]) to a *multistate Markov* scheme [7], [30]. In [34], two of these models are compared, one with a complete multistate Markov scheme, and the other inspired from [24] with a multistate scheme for the sodium ion and a two-state gating for the potassium ion. Here we consider only single-compartment neurons, but in order to deal with spatial heterogeneities of axonal or ion channel properties for instance, multi-compartmental models can be introduced as a discretized description of the spatial neuron, with Ohm's law coupling between compartments.

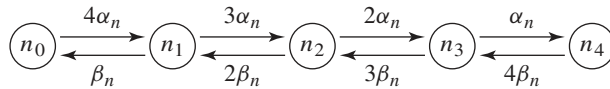### 3.2. Application of the law of large numbers to the Hodgkin–Huxley model

Classically, the Hodgkin–Huxley model is the set of nonlinear differential equations

$$C_m \frac{dV}{dt} = I - g_L(V - V_L) - g_{Na}m^3h(V - V_{Na}) - g_K n^4(V - V_K),$$
$$\frac{dm}{dt} = (1 - m)\alpha_m(V) - m\beta_m(V),$$
$$\frac{dh}{dt} = (1 - h)\alpha_h(V) - h\beta_h(V),$$
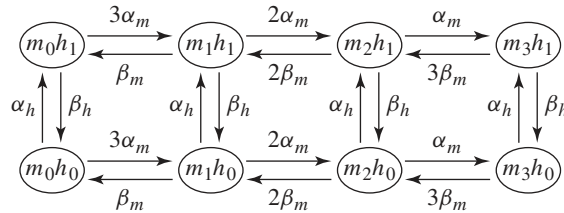$$\frac{dn}{dt} = (1 - n)\alpha_n(V) - n\beta_n(V),$$

where I is the input current, $C_m = 1\,\mu\text{F}/\text{cm}^2$ is the capacitance corresponding to the lipid bilayer of the membrane, $g_L = 0.3\,\text{mS}/\text{cm}^2$, $g_{Na} = 120\,\text{mS}/\text{cm}^2$, and $g_K = 36\,\text{mS}/\text{cm}^2$ are maximum conductances, and $V_L = 10.6\,\text{mV}$, $V_{Na} = 115\,\text{mV}$, and $V_K = -12\,\text{mV}$ are resting potentials, respectively for the leak, sodium, and potassium currents. The model was introduced in [15] to explain the ionic mechanisms behind action potentials in the squid giant axon. The functions α_x and β_x for $x = m, n, h$ are opening and closing rates for the voltage-gated ion channels (see [15] for their expressions). The dynamics of this dynamical system can be very complex, as shown in [13], but for our purpose let us describe schematically only the behaviour of this system as the parameter I is varied (see [29] for more details). First, for all I, there exists a unique equilibrium point. For $0 < I < I_1 \approx 9.8\,\mu\text{A}/\text{cm}^2$, this equilibrium point is stable, and, for $I_0 < I < I_1$, where $I_0 \approx 6.3\,\mu\text{A}/\text{cm}^2$, this equilibrium coexists with a stable limit cycle and possibly many unstable limit cycles. At $I = I_1$ and $I = I_2$ there occur two Hopf bifurcations. For $I_1 < I < I_2 \approx 153\,\mu\text{A}/\text{cm}^2$, the equilibrium point is unstable and coexists with a stable limit cycle. For $I > I_2$, there are no more limit cycles, and the equilibrium point is stable. This bifurcation structure can be roughly interpreted as follows: for small I, the system converges to an equilibrium point, and, for sufficiently large I, the system admits a large amplitude periodic solution, corresponding to an infinite sequence of action potentials or spikes, and the spiking frequency is modulated by the input current I.

There are two stochastic interpretations of the Hodgkin–Huxley model, involving either a multistate Markov model or a two-state gating model. We now present them briefly and we apply our theorems to each of them. A comparison of the deterministic limits obtained for these models is provided in Appendix A and establishes an equivalence between these deterministic versions as soon as their initial conditions satisfy a combinatorial relationship. This question is further studied in [19], where the reduction of the law of jump Markov processes to invariant manifolds is investigated.

3.2.1. *Multistate Markov model.* This model has two types of ion channel: one for sodium and the other for potassium. The kinetic scheme describing the Markov jump process for one single potassium channel is

$$n_0 \underset{\beta_n}{\overset{4\alpha_n}{\rightleftarrows}} n_1 \underset{2\beta_n}{\overset{3\alpha_n}{\rightleftarrows}} n_2 \underset{3\beta_n}{\overset{2\alpha_n}{\rightleftarrows}} n_3 \underset{4\beta_n}{\overset{\alpha_n}{\rightleftarrows}} n_4$$

and for the sodium channel is

$$
\begin{array}{cccc}
m_0h_1 \underset{\beta_m}{\overset{3\alpha_m}{\rightleftarrows}} m_1h_1 & \underset{2\beta_m}{\overset{2\alpha_m}{\rightleftarrows}} m_2h_1 & \underset{3\beta_m}{\overset{\alpha_m}{\rightleftarrows}} m_3h_1 \\
\alpha_h \updownarrow \beta_h \quad \alpha_h \updownarrow \beta_h \quad \alpha_h \updownarrow \beta_h \quad \alpha_h \updownarrow \beta_h \\
m_0h_0 \underset{\beta_m}{\overset{3\alpha_m}{\rightleftarrows}} m_1h_0 & \underset{2\beta_m}{\overset{2\alpha_m}{\rightleftarrows}} m_2h_0 & \underset{3\beta_m}{\overset{\alpha_m}{\rightleftarrows}} m_3h_0
\end{array}
$$

We recall that every single ion channel process is independent within a subpopulation and between subpopulations. All the coefficients in these two schemes are actually functions of the membrane potential, and can be found in [15]. The state spaces are

$$E_1 = \{n_0, n_1, n_2, n_3, n_4\}, \qquad E_2 = \{m_0h_1, m_1h_1, m_2h_1, m_3h_1, m_0h_0, m_1h_0, m_2h_0, m_3h_0\}.$$

The open states are respectively n_4 ($r_1 = 5$) and m_3h_1 ($r_2 = 8$). The proportion of open potassium channels is denoted by $u_N^{(1)} := \{e_N^{(1)}\}_{n_4}$ and the proportion of open sodium channels by $u_N^{(2)} := \{e_N^{(2)}\}_{m_3h_1}$. In this model, the membrane potential dynamic is given by

$$\dot{V}_N(t) = -g_{Na} u_N^{(2)}(t)(V_N(t) - V_{Na}) - g_K u_N^{(1)}(t)(V_N(t) - V_K) - g_L(V(t) - V_L) + I,$$

where $I \in \mathbb{R}$ is a constant applied current. With the notation of Section 2.1, $f(v, u^{(1)}, u^{(2)}) = -g_{Na}u^{(2)}(v - V_{Na}) - g_K u^{(1)}(V_N(t) - V_K) - g_L(v - V_L) + I$ and, for example, $\alpha_{k,l}^{(1)}(v) = 3\alpha_n(v)$ if $k = n_1$ and $l = n_2$, and $\alpha_{k,l}^{(2)}(v) = 2\beta_m(v)$ if $k = m_2h_0$ and $l = m_1h_0$.

Applying Theorem 2.1, we obtain a deterministic limit for the stochastic Hodgkin–Huxley model when $N \to \infty$, provided that we choose the initial conditions appropriately:

$$\dot{v} = -g_{Na}e^{(2)}(t)(v(t) - V_{Na}) - g_K e^{(1)}(t)(v(t) - V_K) - g_L(v(t) - V_L) + I,$$

$$\dot{\boldsymbol{g}}_k^{(j)} = \sum_{1 \le i \le r_j, i \ne k} \alpha_{i,k}^{(j)}(v)\boldsymbol{g}_i^{(j)} - \alpha_{k,i}^{(j)}(v)\boldsymbol{g}_k^{(j)} \quad \text{for all } 1 \le j \le 2 \text{ and } 1 \le k \le r_j,$$

$$V(0) = v_0, \qquad \boldsymbol{g}^{(j)}(0) = \boldsymbol{g}_0^{(j)},$$

with $e^{(j)} = \boldsymbol{g}_{r_j}^{(j)}$, where the rate functions $\alpha_{m,p}^{(j)}$ are given by the above schemes for $j \in \{1, 2\}$.

3.2.2. *Two-state gating model.* Another way of building a stochastic Hodgkin–Huxley model is to consider that the channels can be decomposed into independent subunits, called gates. Each gate can be either open (state 1) or closed (state 0):

$$0 \underset{\beta_z(V_N)}{\overset{\alpha_z(V_N)}{\rightleftarrows}} 1$$
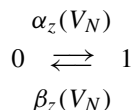
with $z \in \{m, n, h\}$. A channel is open when all gates are open. For the sodium channel, there are

two distinct types of gate, and one for the potassium channel. Thus, in this situation, the number, q, of subpopulations refers to the number of types of gate, so $q = 3$ and $E_1 = E_2 = E_3 = \{0, 1\}$. If we denote by $u_N^{(z)}(t)$ the proportion of open gates z for $z \in \{m, n, h\}$, the membrane potential dynamic is then given by

$$\dot{V}_N(t) = -g_{Na}(u_N^{(m)}(t))^3 u_N^{(h)}(t)(V_N(t) - V_{Na})$$
$$- g_K(u_N^{(n)}(t))^4(V_N(t) - V_{Na}) - g_L(V_N(t) - V_L) + I,$$

which corresponds to $f(v, u^{(m)}, u^{(h)}, u^{(n)}) = -g_{Na}(u^{(m)})^3 u^{(h)}(v - V_{Na}) - g_K(u^{(n)})^4(v - V_{Na}) - g_L(v - V_L) + I$. In Figure 1, we give a sample trajectory of this two-state gating stochastic Hodgkin–Huxley system.
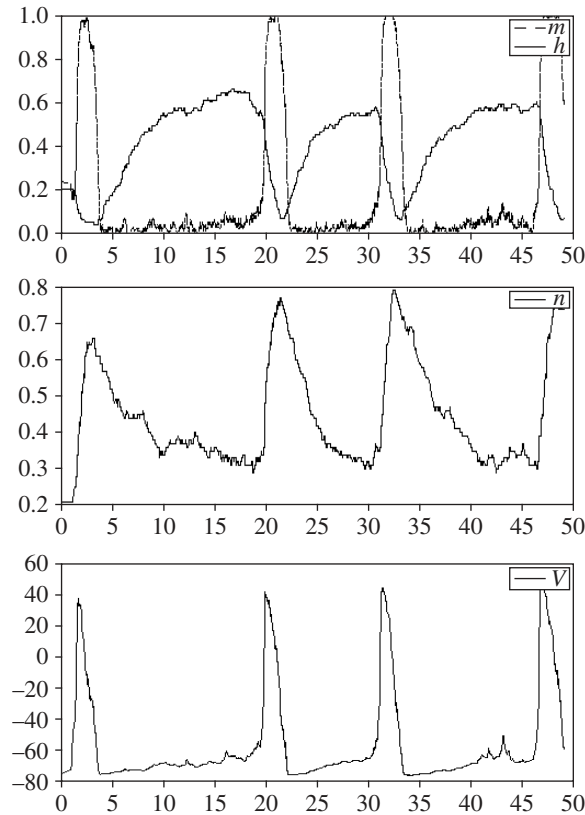

FIGURE 1: Illustrative sample trajectory of a two-state gating stochastic Hodgkin–Huxley model with $N = 20$ (cf. Section 3.2). *Top*: variables m and h for the sodium channel (without unit). *Middle*: variable n for the potassium channel (without unit). *Bottom*: variable V for membrane potential (unit: mV). Abscissa: time (arbitrary units). Since m, n, and h correspond to proportions of open gates, if one of them is equal to 1 then all the corresponding gates are open. An increase in the membrane potential V causes an increase in the proportion of open m (sodium) gates, which in turn implies an increase of V. This positive feedback results in a spike initiation. Meanwhile, a further increase of V is followed by a decrease of the deactivation variable h, which closes the sodium channels. This inhibition effect acts at a slower time scale, enabling a decrease of V. This decrease is strengthened by the dynamic of variable n (proportion of open potassium gates).

Applying Theorem 2.1 gives the classical formulation of the four-dimensional Hodgkin–Huxley model:

$$\dot{v} = -g_{Na}(u^{(m)}(t))^3 u^{(h)}(t)(v(t) - V_{Na}) - g_K(u^{(n)}(t))^4(v(t) - V_{Na}) - g_L(v(t) - V_L) + I,$$

$$\dot{u}^{(z)}(t) = (1 - u^{(z)}(t))\alpha_z(v(t)) - u^{(z)}(t)\beta_z(v(t)), \qquad z \in \{m, n, h\}.$$

3.3. Exponential convergence speed

We illustrate by numerical simulations the upper bound obtained in (2.1) for the stochastic Hodgkin–Huxley model with a two-state gating scheme. The number of sodium channels, N_{Na}, and potassium channels, N_K, are proportional to the area S of the membrane patch. Thus, instead of N, S will denote the size parameter. For the squid giant axon, the estimated densities for the ion channels used in the simulations are $\rho_{Na} = 60 \, \mu\text{m}^{-2}$ and $\rho_{N_K} = 18 \, \mu\text{m}^{-2}$.

We now present the results of numerical simulations of

$$P_S(T, \Delta) = \mathbf{P}\left[\sup_{0 \le t \le T} \|V_S(t) - v(t)\|^2 + \sum_{j=1}^{q} \|e_S^{(j)}(t) - g^{(j)}(t)\|^2 > \Delta\right]$$

using Monte Carlo simulations. We recall from (2.1) that

$$\limsup_{S \to \infty} \frac{1}{S} \log P_S(T, \Delta) \le -\frac{\Delta e^{-BT^2}}{CT} = C(T).$$

In Figure 2, the simulation estimations of $C_S(T) = 1/S \log P_S(T, \Delta)$ are shown for different values of T and S, and can be compared to the theoretical bound $C(T)$. Simulations are made without input current, meaning that the stochastic solution is supposed to fluctuate around the equilibrium point of the deterministic system in a neighbourhood of size proportional to $S^{-1/2}$. When S increases, the simulation curve $C_S(T)$ is expected to pass below the theoretical bound $C(T)$. For higher input currents, still subthreshold, but close to the bifurcation I_c, channel noise will induce spontaneous action potentials. For appropriate Δ, the probability $P_S(T, \Delta)$
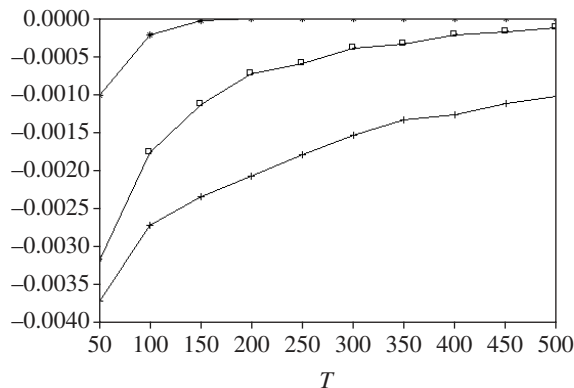

FIGURE 2: Simulation results for the Hodgkin–Huxley model with a 'two-state gating' scheme and input current $I = 0$: this figure shows the quantity $1/S \ln P_S(T, \Delta)$ as a function of T, where S is the area of the patch, and is thus proportional to N. Here $S = 250 \, \mu\text{m}^2$ (*stars*) (corresponding to $N_{Na} = 15\,000$ and $N_K = 5000$), $S = 500 \, \mu\text{m}^2$ (*squares*), and $S = 750 \, \mu\text{m}^2$ (*crosses*). The lines are plotted as a guide for the eye.

can be interpreted as the probability that the first spontaneous action potential (SAP) occurs before time T. Thus, the convergence speed bound gives an upper bound of the cumulative distribution function of this first SAP time.

For higher input currents $I > I_c$, the deterministic solution will be attracted by a stable limit cycle, which corresponds to repetitive action potentials. In this case, channel noise may introduce a jitter in the spiking times. Thus, if one considers the supremum of the errors between the stochastic and the deterministic solutions, this supremum will be quite large (approximately the size of an action potential) as soon as the difference between the stochastic spiking times and the deterministic ones is of order of the time course of an action potential (2 ms). Thus, the supremum of the difference is not appropriate here and we will see in the following section how to quantify the impact of channel noise on the spiking frequency.

3.4. Application of the central limit theorems

In this section we show how to investigate the fluctuations around a stable fixed point (sub-threshold fluctuations) and the fluctuations around a stable limit cycle (firing rate fluctuations) using Theorem 2.2. Let us consider a class of two-dimensional models, corresponding to Example 2.1. This class contains reductions of the previous two-state gating Hodgkin–Huxley model, or other models such as the Morris–Lecar model [25]. Consider the process, using the notation of Example 2.1,

$$\begin{pmatrix} Y_N \\ P_N \end{pmatrix} = \begin{pmatrix} \sqrt{N}(V_N - V) \\ \sqrt{N}(e_N - g) \end{pmatrix}$$

with initial conditions $(P_N(0), Y_N(0)) = (0, 0)$. Then the two-dimensional process $Z_N = (P_N, Y_N)$ converges in law, as $N \to \infty$, towards the process $Z = (P, Y)$, whose characteristic function is given by

$$\mathrm{E}[\exp(\mathrm{i}(\theta_1 P(t) + \theta_2 Y(t)))] = \exp(\theta_1^2 A_t + \theta_2^2 B_t + \theta_1 \theta_2 C_t).$$

This characteristic function is obtained by specializing the differential equation of Theorem 2.2. Thus, defining Σ_s to be the square root matrix of

$$\Gamma_s := \begin{pmatrix} A'_s & C'_s/2 \\ C'_s/2 & B'_s \end{pmatrix} \quad \text{for } 0 \leq s \leq t,$$

Z can be written as a Gaussian diffusion process:

$$Z_t = \int_0^t \Sigma_s \, dW_s,$$

where W is a standard two-dimensional Brownian motion. The condition that the matrix Γ_s admits a real square root matrix can be reduced to $A'_s + B'_s \leq 0$ because we can show that $\det(\Gamma_s) = A'_s B'_s - (C'_s)^2/4 = 0$ for all $s \geq 0$. This condition is thus always satisfied because $A'_0 + B'_0 \leq 0$, A'_s and B'_s have the same sign, and (A'_s, B'_s, C'_s) cannot cross $(0, 0, 0)$ by uniqueness of the solution of $z' = Mz$ (satisfied by y'). The computation of the matrix Σ_s gives

$$\Sigma_s = \frac{\sqrt{-2(A'_s + B'_s)}}{A'_s + B'_s} \Gamma_s.$$

From the equation for the characteristic function obtained in Theorem 2.2, we find that the triple $y = (A, B, C)$ is a solution of the system $\dot{y}_t = M_t y_t + E_t$, defined as

$$\begin{pmatrix} \dot{A}_t \\ \dot{B}_t \\ \dot{C}_t \end{pmatrix} = \begin{pmatrix} 2b'_u & 0 & b'_v \\ 0 & 2f'_v & f'_u \\ 2f'_u & 2b'_v & b'_u + f'_v \end{pmatrix} \begin{pmatrix} A_t \\ B_t \\ C_t \end{pmatrix} + \begin{pmatrix} -\frac{1}{2}\lambda(V, u) \\ 0 \\ 0 \end{pmatrix}$$

with initial conditions $(0, 0, 0)$, and $\lambda(v, u) = \sqrt{(1 - u)\alpha(v) + u\beta(v)}$. The partial derivatives f'_v, f'_u, b'_v, and b'_u and λ are evaluated at the deterministic solution (V_t, g_t) of (D).

We remark that, if J is the Jacobian matrix of the two-dimensional deterministic system (D) at the point (V_t, g_t), and if its spectrum is $\mathrm{sp}(J) = \{\lambda_1, \lambda_2\}$, then the spectrum of M is $\mathrm{sp}(M) = \{2\lambda_1, 2\lambda_2, \lambda_1 + \lambda_2\}$. Two different situations can be considered.

- Starting from a fixed point (V_0, g_0) of the deterministic system, the matrix $M_t = M(V_t, g_t)$ and the vector $E_t = E(V_t, g_t)$ are constant. We can derive an explicit analytical solution diagonalizing the matrix M. The time evolution for the variance and covariance of the difference between the deterministic and stochastic solutions then depends on the stability (λ_1, λ_2) of the considered fixed point. Noise-induced small oscillations may appear if the eigenvalues are complex conjugates.

- Around a stable limit cycle (periodic firing), M_t and E_t are T-periodic functions. Using suitable coordinates and following Floquet's theory (see [3]), stability would be given by the spectrum of the solver $R(T)$: $(A_0, B_0, C_0) \rightarrow (A_T, B_T, C_T)$. As explained in [17], even if the real parts of the eigenvalues of the Jacobian matrix are strictly negative for all times, unstable solutions may exist. In Section 3.5 we numerically investigate the fluctuations around a stable limit cycle for the Morris–Lecar system.

If we consider

$$\begin{pmatrix} \tilde{Y}_N \\ \tilde{P}_N \end{pmatrix} = \begin{pmatrix} \sqrt{N}(\tilde{V}_N - V) \\ \sqrt{N}(\tilde{g}_N - g) \end{pmatrix},$$

where $(\tilde{V}_N, \tilde{g}_N)$ is the Langevin approximation, then the moment equations, written for the linearized version around the deterministic solution, give the same matrix Γ_s at the limit $N \rightarrow \infty$. But, for finite N, the linearized process is not Gaussian (see Appendix B). Thus, our mathematical result can be directly related to the simulation results obtained in [34]: in this paper simulations of two neuron models with a large number of stochastic ion channels are made, and the fluctuations of the membrane potential below threshold exhibit approximately Gaussian distributions, but only for a certain range of resting potentials. For smaller resting potentials, the shape of the distribution remained unclear as it was more difficult to compute. Our approach shows that, at finite N, for any range of the resting potentials, the distribution is non-Gaussian. However, when $N \rightarrow \infty$, the distribution tends to a Gaussian, which corresponds to the approximate Gaussian distribution observed in the simulations of [34].

3.5. Quantifying the effect of channel noise on neural coding

In the next subsections we show how our results can be applied to quantify the effect of ion channel stochasticity on neural coding. The impact of channel noise on frequency coding is investigated in Section 3.5.1 and on spike timing coding in Section 3.5.2. We close this section with some remarks concerning non-Markovian processes arising when considering synaptic transmission in Section 3.5.3.

3.5.1. *Numerical study of the variance of the spiking rate for the Morris–Lecar model.* In this subsection, applying Theorem 2.2 to the Morris–Lecar system, we investigate the impact of channel noise on the variance of the firing frequency. The Morris–Lecar system was introduced in [25] to account for various oscillating states in the barnacle giant muscle fiber. We denote by $X = (V, m, n)$ the solution of

$$C_m \frac{dV}{dt} = I - g_L(V - V_L) - g_{Ca}m(V - V_{Ca}) - g_K n(V - V_K) := F_v(X), \qquad (3.1)$$

$$\frac{dm}{dt} = \lambda_m(V)(M_\infty(V) - m) := F_m(X), \qquad (3.2)$$

$$\frac{dn}{dt} = \lambda_n(V)(N_\infty(V) - n) := F_n(X), \qquad (3.3)$$

where $\lambda_m(V) = \cosh((V - V_1)/2V_2)$, $\lambda_n(V) = \phi_n \cosh((V - V_3)/2V_4)$, $M_\infty(V) = (1 + \tanh[(V - V_1)/V_2])/2$, and $N_\infty(V) = (1 + \tanh[(V - V_3)/V_4])/2$. We introduce, as in the previous sections, a stochastic version X_N of this model with stochastic ion channels, replacing the differential equation for m and n by birth-and-death processes with voltage-dependent opening rates $\alpha_m = \lambda_m M_\infty$ and $\alpha_n = \lambda_n N_\infty$ and closing rates $\beta_n = \lambda_n(1 - N_\infty)$. According to the parameters of the model, the deterministic system (3.1)–(3.3) may have a stable limit cycle X^{LC} for some values of $I \in [I_{\min}, I_{\max}]$ (see [25]). This corresponds to a phenomenon of regular spiking, characterized by its rate. Assuming that the time length of a spike is almost constant, we suggest a proxy for this spiking rate:

$$r(T) := \frac{1}{T} \int_0^T \phi_{th}(X(s)) \, ds,$$

where ϕ_{th} is a sigmoid threshold function. In a similar way, we define the stochastic spiking rate by

$$r_N(T) := \frac{1}{T} \int_0^T \phi_{th}(X_N(s)) \, ds.$$

As a candidate for ϕ_{th}, we choose $\phi_{th}(V) := e^{c(V - V_{th})}/(1 + e^{c(V - V_{th})})$, where c and V_{th} are two parameters.

A consequence of the central limit theorem for X_N is weak convergence,

$$\sqrt{N}[r_N(T) - r(T)] \Rightarrow R(T) = \frac{1}{T} \int_0^T Z(s) \nabla \phi_{th}(X(s)) \, ds,$$

where Z is the weak limit of $\sqrt{N}[X_N - X]$,

$$Z(s) = \int_0^s \Sigma(u) \, dW_u,$$

and $R(T)$ is a Gaussian random variable with zero mean. For simplicity, we consider the case where ϕ_{th} is only a function of the membrane potential V. Then the variance of $R(T)$ is

$$\sigma_R^2(T) = E[R(T)^2] = \frac{2}{T^2} \int_0^T \int_0^s S_v(s') \phi'_{th}(V(s')) \, ds' \phi'_{th}(V(s)) \, ds, \qquad (3.4)$$

where $S_v(s) = \Sigma_{1,1}(s)$ is the variance of $\sqrt{N}(V_N(s) - V(s))$.
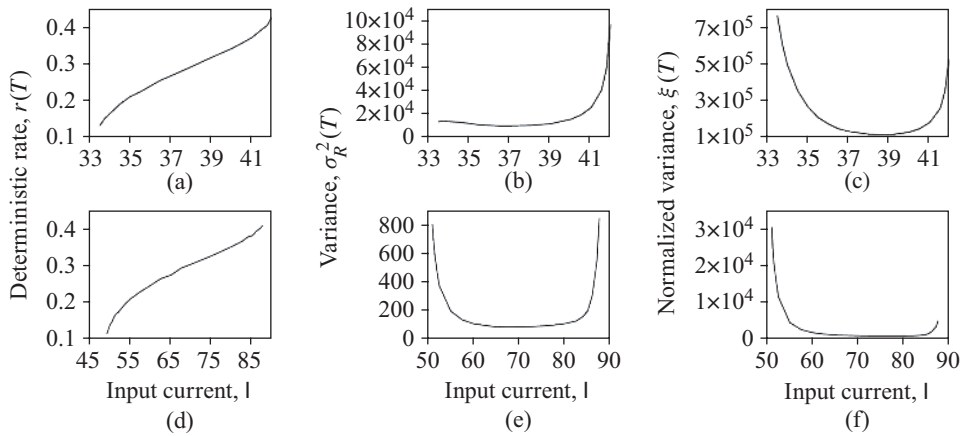
FIGURE 3: Impact of channel noise on the spiking rate. *Top row*: class I regime. *Bottom row*: class II regime. (a) and (d) show the deterministic rate $r(T)$ versus the input current I (μA/cm^2). (b) and (e) show the variance $\sigma_R^2(T)$ versus the input current I (μA/cm^2). (c) and (f) show the normalized variance $\xi(T)$ versus the input current I (μA/cm^2). Parameters: for all figures, $T = 2000$ ms, $c = 10$, $V_{\text{th}} = 0$ mV, $C_m = 20\,\mu$F/cm^2, $V_1 = 0$ mV, $V_2 = 15$ mV, $V_3 = 10$ mV, $g_{Ca} = 4$ mS/cm^2, $g_K = 8$ mS/cm^2, $g_L = 2$ mS/cm^2, $V_K = -70$ mV, $V_L = -50$ mV, $V_{Ca} = 100$ mV, $\phi_n = 0.1$, and $V_4 = 20$ mV for class II and $V_4 = 10$ mV for class I.

To numerically estimate the variance $\sigma_R^2(T)$, the first step is to numerically determine the limit cycle, then solve the moment equations (Appendix C) and immediately deduce $\Sigma(s)$. Thus, the variance σ_R^2 can be computed using (3.4) without any stochastic simulation. In Figure 3 we show our numerical results, where we plot in (c)–(f), as a function of the input current I, the normalized variance $\xi(T)$ defined as $\xi(T) := \sigma_R^2(T)/r(T)^2$.

Remarks. The value of $\xi(T)$ depends on a combination of the linear stability along the cycle and the variance of the noise (which is multiplicative) along the cycle. If we want to have the quantity $\mathrm{E}[(r_N(T) - r(T))^2]/r(T)^2$ of order 1 then the number, N, of channels should be of order $\xi(T)$. Interestingly, this gives much smaller values for the class II regime compared to the class I regime (see [16] for more details about these regimes). In both cases, it corresponds to a reasonably small number of channels when I is not too close to bifurcation points.

3.5.2. *Impact of channel noise on latency coding in the Morris–Lecar model.* Whereas frequency coding requires an integration of the input signal over a relatively long time, individual spike time coding does not require such an integration. The time to the first spike, called latency, depends on the value of the suprathreshold input. Thus, it may have an interpretation in terms of neural coding, and it has been shown in several sensory systems [39] that the first spike latency carries information. For example, a recent study [12] concerning the visual system suggests that it allows the retina to rapidly transfer new spatial information. Impacts of external noise on latency coding have been investigated in numerical studies [27] with stochastic simulations. We apply Theorem 2.3 to the Morris–Lecar model to investigate the impact of internal channel noise on the first spike time. We choose the parameters (see Figure 3) to obtain a class I neuron model

in the excitable regime. In this setting, there exists a unique steady state $X^* = (V^*, m^*, n^*)$. Starting from this equilibrium point, the impact of an input at $t = t_0$ is equivalent to an instantaneous shift of the membrane potential $V^* \to V^* + A$, where $A > 0$ is the amplitude of this shift. Eventually, the system goes back to its steady state, but if A is higher than a threshold A_{th} then a spike is emitted before going back to the steady state, whereas if A is lower than A_{th}, no spike is emitted. For $A > A_{\text{th}}$, we define the latency time $T(A)$ as the elapsed time between t_0 and the spike. More precisely, let $X_A(t) = (V_A(t), m_A(t), n_A(t))$ for $t \geq t_0$ be the solution of the Morris–Lecar equations with initial conditions $X(t_0) = (V^* + A, m^*, n^*)$. We define a spike as a passage of the membrane potential $V_A(t)$ through a threshold V_{th}. Then, with $t_0 = 0$ for simplicity, the latency time $T(A)$ can be written as $T(A) := \inf\{t \geq 0; V_A(t) > V_{\text{th}}\}$. As shown in Figure 4(a), the more $A > A_{\text{th}}$ is close to A_{th}, the longer is the latency time $T(A)$. The same setting can be extended to the stochastic case, defining a random variable $T_N(A)$. Applying Theorem 2.3, with $\phi(V, m, n) = V_{\text{th}} - V$, we express the variance $P(A)$ of the limit
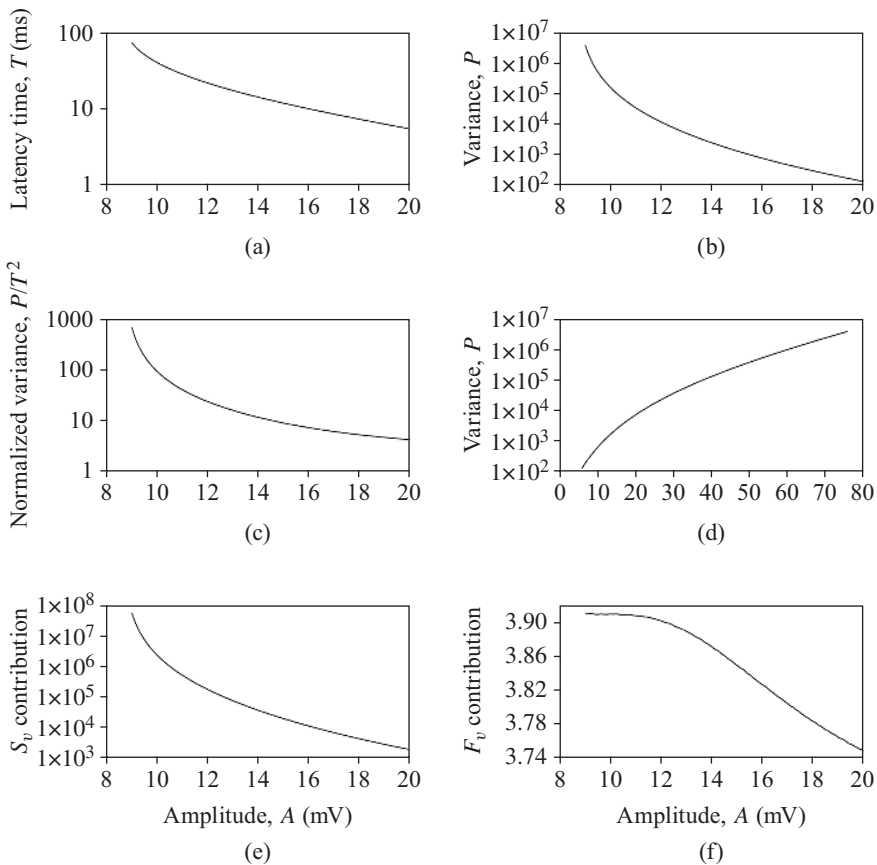

FIGURE 4: Impact of channel noise on latency coding. (a) Latency time $T(A)$ versus amplitude A. (b) Variance $P(A)$ versus amplitude A. (c) Normalized variance $P(A)/T(A)^2$ versus amplitude A. (d) Variance $P(A)$ versus latency time $T(A)$. (e) Variance $S_v(A)$ versus amplitude A. (f) Crossing speed $F_v(X(T(A)))$ versus amplitude A. The parameters used are the same as those used in Figure 3, class I, with input current $I = 32 \, \mu\text{A/cm}^2$.

of $\sqrt{N}(T_N(A) - T(A))$ as $N \to \infty$:

$$P(A) = \frac{S_v(T(A))}{F_v(X(T(A)))^2}. \tag{3.5}$$

In (3.5), $S_v(T(A))$ is the variance of the V-component Y of Z, where we recall that Z is the limit of $\sqrt{N}(X_N - X)$ (see Theorem 2.2). The value of $S_v(T(A))$ is obtained from the numerical integration of the moment equations (Appendix C). The results are displayed in Figure 4, where the variance $P(A)$ and a normalized variance $P(A)/T(A)^2$ are plotted against the amplitude A (Figure 4(b)). In Figure 4(d) the variance $P(A)$ is plotted against the latency time $T(A)$. From (3.5), it appears that $P(A)$ is determined by two distinct contributions: the variance $S_v(T(A))$ (Figure 4(e)) and the crossing speed $F(X(T(A)))$ (Figure 4(f)), which does not actually influence much the variance $P(A)$. One way to interpret the results is the following: if N is large, of order $P(A)$, then $E[(T_N(A) - T(A))^2]$ is of order 1. Thus, as an illustration, in order to keep $E[(T_N(A) - T(A))^2]$ of order 1, the required number of channels would be of order 10^2 for a latency time of 10 ms and of order 10^5 for latency time of 60 ms.

3.5.3. *Synaptic transmission and non-Markovian processes.* In Section 3.5.1, the quantity of interest was the firing frequency. However, the synaptic transmission between neuron 1 and neuron 2 has its own time scale. Therefore, neuron 2's input, called the post-synaptic potential $\Psi^{1 \to 2}$, may be modelled as a functional of neuron 1's membrane potential $\{V_1(t)\}_{t \geq 0}$. Although synaptic transmission is presumably a nonlinear process, we can consider as a first approximation (cf. [21]) that the process of interest is obtained directly by the convolution of the process V_1 with some kernel $K^{1 \to 2}$:

$$\Psi^{1 \to 2}(t) = \int_0^t K^{1 \to 2}(t, s) V_1(s) \, ds.$$

The mathematical analysis of the impact of channel noise on this variable can be done in light of Theorems 2.1 and 2.2. Using the general notation for the stochastic process and its deterministic limit, we define $\Psi_N(t) = \int_0^t K(t, s) X_N(s) \, ds$ and $\Psi(t) = \int_0^t K(t, s) X(s) \, ds$.

Law of large numbers. Define $S_N(T) = \sup_{s \in [0,T]} |\Psi_N(t) - \Psi(t)|^2$. Clearly, using the Cauchy–Schwarz inequality, with the same notation as in Theorem 2.1,

$$P(S_N(T) > \Delta) \leq P_N(T, \eta(T)^{-1}\Delta)$$

with

$$\eta(T) = T \sup_{t \in [0,T]} \int_0^t |K(t, s)|^2 \, ds.$$

The convergence of Ψ_N to Ψ is thus a direct consequence of Theorem 2.1. The convergence speed is exponential, as in Theorem 2.1.

Gaussian fluctuations. We know from Theorem 2.2 that $\sqrt{N}(X_N - x)$ converges weakly to the diffusion $Z(t) = \int_0^t \Sigma(u) \, dW_u$. As a consequence, $\Omega_N = \sqrt{N}(\Psi_N - \Psi)$ also converges weakly to the following process:

$$\Omega(t) = \int_0^t K(t, s) \left(\int_0^s \Sigma(u) \, dW_u \right) ds.$$

Using integration by parts, we can rewrite

$$\Omega(t) = \int_0^t J(t, s) \, dW_s,$$

where

$$J(t, s) = \int_s^t K(t, u) \, du \, \Sigma(s).$$

The process Ω is Gaussian and we can easily compute its variance as $\int_0^t J(t, s)^2 \, ds$. However, it is non-Markovian, and some issues concerning the first hitting times of such processes are addressed in [35].

4. Proof of the law of large numbers

In this section we give the proof for Theorem 2.1. This proof is inspired from [1], except for the exponential martingale bound. In order to simplify the notation and to make the arguments clearer and more intuitive, we write the proof for the case of a single channel type with state space $\{0, 1\}$ and transition rates given by the scheme

$$\begin{array}{c} \alpha(V_N) \\ 0 \; \rightleftarrows \; 1. \\ \beta(V_N) \end{array}$$

In this case, the stochastic model (S_N^0) is

$$\dot{V}_N(t) = f(V_N(t), u_N(t)), \qquad V_N(0) = V_0,$$

$$u_N(t) = \frac{1}{N} \sum_{k=1}^N \delta_1(u_t^{(k)}), \qquad u_N(0) = u_0^{(N)},$$

where $u_t^{(k)} : 0 \to 1$ with rate $\alpha(V_N(t))$ and $1 \to 0$ with rate $\beta(V_N(t))$ for all $1 \leq k \leq N$.

The deterministic solution (v, u) satisfies

$$\dot{v}(t) = f(v(t), u(t)), \qquad \dot{u}(t) = (1 - u(t))\alpha(v(t)) - u(t)\beta(v(t)),$$

$$v(0) = v_0, \qquad u(0) = u_0.$$

In this section and Section 5, we use the variable u instead of g, as we are working with a simplified version of the model. In order to complete the proof, a few slight changes in the notation are necessary:

- in order to work with more general jump Markov processes with finite state space, essentially all the expressions of the form $\delta_0(u)\alpha(v) - \delta_1(u)\beta(v)$ should be replaced by

$$\sum_{i \neq j} \alpha_{i,j}(v, u)\delta_{e_i}(u) - \alpha_{j,i}(v, u)\delta_{e_j}(u);$$

- in order to include q different channel types (different ions), we should just write the same arguments for all the q processes $\{e_N^{(j)}(t)\}$ for $1 \leq j \leq q$ and include all the $\|e_N^{(j)}(t) - e^{(j)}(t)\|$ for $1 \leq j \leq q$ in the function $f(t)$ of Gronwall's lemma in Section 3.4.

4.1. Decomposition into a martingale part and a finite-variation part

Decomposition. We decompose the difference between the stochastic and deterministic processes as a sum of a martingale part, M_N, and a finite-variation part, Q_N:

$$[u_N(t) - u_N(0)] - [u(t) - u(0)] = M_N(t) + \int_0^t Q_N(s) \, ds,$$

where we define

$$Q_N(t) = \frac{1}{N} \sum_{i=1}^{N} [\delta_0(u_t^{(i)})\alpha(V_N(t)) - \delta_1(u_t^{(i)})\beta(V_N(t))] - \dot{u}(t),$$

$$M_N(t) = [u_N(t) - u(t)] - [u_N(0) - u(0)] - \int_0^t Q_N(s)\,\mathrm{d}s.$$

Let $\{F_t\}_{t \geq 0}$ denote the filtration of σ-algebras indexed by time generated by all the stochastic processes under consideration for a particular N.

Lemma 4.1. *As defined above, $(M_N(t))$ is an $\{F_t\}$-martingale.*

Proof. For $h > 0$, define $\Delta M_N(t, h) = (1/h)\,\mathrm{E}[M_N(t + h) - M_N(t) \mid F_t]$. Then

$$\Delta M_N(t, h) = \frac{1}{h}\frac{1}{N} \sum_{i=1}^{N} \mathrm{E}[\delta_1(u_{t+h}^{(i)}) \mid F_t] - \mathrm{E}[\delta_1(u_t^{(i)}) \mid F_t]$$

$$- \frac{1}{h} \mathrm{E}\left[\int_t^{t+h} \left[\frac{1}{N}\sum_{i=1}^{N} \delta_0(u_s^{(i)})\alpha(V_s) - \delta_1(u_s^{(i)})\beta(V_s)\right]\mathrm{d}s \,\bigg|\, F_t\right]$$

$$- \frac{1}{h}[u(t+h) - u(t)] + \frac{1}{h}\int_t^{t+h} \dot{u}(s)\,\mathrm{d}s.$$

The last line clearly converges to 0 as $h \to 0$, and the two first terms compensate as $h \to 0$. So we have

$$\lim_{h \to 0} \frac{1}{h}\mathrm{E}[M_N(t+h) - M_N(t) \mid F_t] = 0.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}[M_N(t+s) \mid F_t]|_{s=0} = 0.$$

By dominated convergence we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{E}[M_N(t+s) \mid F_t]|_{s=s_0} = \mathrm{E}\left[\frac{\mathrm{d}}{\mathrm{d}u}\mathrm{E}[M_{t+s_0+u} \mid F_{t+s_0}]|_{u=0} \,\bigg|\, F_t\right] = 0.$$

Finally,

$$\mathrm{E}[M_N(t+h) \mid F_t] = \text{constant} = M_N(t).$$

4.2. Martingale bound

In this section we want to obtain a bound in probability for the martingale part. We introduce the jump measure and the associated compensator.

We define two random measures on $(0, T] \times \{0, 1\}$:

- the jump measure,

$$\kappa_i = \sum_{t \in (0,T],\, u_t^{(i)} \neq u_{t-}^{(i)}} \delta_{(t, u_t^{(i)})};$$

- the compensator,

$$\nu_i(\mathrm{d}t, \mathrm{d}y) = [\beta(V_N(t))\delta_1(u_{t-}^{(i)})\delta_0(y) + \alpha(V_N(t))\delta_0(u_{t-}^{(i)})\delta_1(y)]\,\mathrm{d}t.$$

We can rewrite $Q_N(s)$ and $M_N(t)$ as

$$\int_0^t Q_N(s)\,\mathrm{d}s = \frac{1}{N}\sum_{i=1}^N \int_{(0,T]\times\{0,1\}} (\delta_1(y) - \delta_1(u_{t-}^{(i)}))\nu_i(\mathrm{d}s,\,\mathrm{d}y) - \int_0^t \dot{u}(s)\,\mathrm{d}s,$$

$$M_N(t) = \frac{1}{N}\sum_{i=1}^N \int_{(0,T]\times\{0,1\}} (\delta_1(y) - \delta_1(u_{t-}^{(i)}))(\kappa_i - \nu_i)(\mathrm{d}s,\,\mathrm{d}y).$$

Then we have the following proposition.

Proposition 4.1. *Let $T > 0$, $\varepsilon > 0$, and $\delta > 0$. Then there exists N_0 such that, for all $N \geq N_0$,*

$$\mathrm{P}\left[\sup_{0\leq t\leq T} M_N(t)^2 \geq \delta\right] \leq \varepsilon.$$

Proof. Let us first recall that from standard results about residual processes (see [18]) we have

$$\mathrm{E}[M_N(t)^2] = \frac{1}{N^2}\sum_{i=1}^N \mathrm{E}\left[\int_{(0,T]\times\{0,1\}} (\delta_1(y) - \delta_1(u_{t-}^{(i)}))^2(\kappa_i - \nu_i)(\mathrm{d}s,\,\mathrm{d}y)\right]$$

$$= \frac{1}{N^2}\sum_{i=1}^N \mathrm{E}\left[\int_{(0,T]} \beta(V_N(s))\delta_1(u_{s-}^{(i)}) + \alpha(V_N(s))\delta_0(u_{s-}^{(i)})\,\mathrm{d}s\right].$$

Therefore, we can obtain a bound for $\mathrm{E}[M_N(t)^2]$:

$$\mathrm{E}[M_N(t)^2] \leq C_1 \frac{t}{N}\max(\|\alpha\|_\infty, \|\beta\|_\infty),$$

where $\|\alpha\|_\infty$ and $\|\beta\|_\infty$ are finite because α and β are continuous and assumption (H1) holds. We then use Chebychev's inequality and Doob's inequality for L^2 martingales:

$$\mathrm{P}\left[\sup_{0\leq t\leq T} M_N(t)^2 \geq \delta\right] \leq \frac{1}{\delta}\mathrm{E}\left[\sup_{0\leq t\leq T} M_N(t)^2\right] \leq \frac{4}{\delta}\mathrm{E}[M_N(t)^2]$$

and $\mathrm{E}[M_N(t)^2] \leq \varepsilon\delta/4$ for all $N \geq N_0$. ∎

In order to obtain a better estimate for the convergence rate, we derive here an exponential bound for the martingale part. Our proof is inspired from techniques developed in [5].

Proposition 4.2. *Let $T > 0$ and $\eta > 0$. There exists a constant C_η such that, for all $\delta \in (0, \eta C_\eta T)$,*

$$\mathrm{P}\left[\sup_{0\leq t\leq T} |M_N(t)| \geq \delta\right] \leq 2\exp\left(-\frac{\delta^2 N}{2C_\eta T}\right).$$

Proof. We define, for $x = (u, v)$ and $\theta \in \mathbb{R}$,

$$m_N(x, \theta) = \int_R e^{\theta y}\lambda_N(x)\mu_N(x,\,\mathrm{d}y) = N\lambda(x)[e^{\theta/N}\mu^+(x) + e^{-\theta/N}\mu^-(x)],$$

$$\phi_N(x, \theta) = \int_R [e^{\theta y} - 1 - \theta y]\lambda_N(x)\mu_N(x,\,\mathrm{d}y)$$

$$= \int_0^1 \frac{\partial^2 m_N}{\partial\theta^2}(x, r\theta)\theta^2(1 - r)\,\mathrm{d}r.$$

The second equality stems from an integration by parts argument. Also, if $|\theta| < N\eta$,

$$\left|\frac{\partial^2 m_N}{\partial \theta^2}(x, r\theta)\right| = \left|N\lambda(x)\frac{1}{N^2}[e^{r\theta/N}\mu^+(x) + e^{-r\theta/N}\mu^-(x)]\right| \leq \frac{C_\eta}{N}.$$

So, $|\phi_N(x, \theta)| \leq C_\eta \theta^2/2N$. Let us define

$$Z_N^\varepsilon(t) = \exp\left(\varepsilon M_N(t) - \int_0^t \phi_N((u_N(s), V_N(s)), \varepsilon)\,\mathrm{d}s\right),$$

$(Z_N^\varepsilon(t))$ is a martingale thanks to Doléans' formula (cf. [28]):

$$Z_N^\varepsilon(t) = 1 + \int_0^t \int_R Z_N^\varepsilon(s^-)[e^{\varepsilon y} - 1](\mu - \nu)(\mathrm{d}s, \mathrm{d}y).$$

Then we note that $\tau = \inf\{t; M_N(t) > \delta\}$. On $\{\tau \leq t\}$, $Z_N^\varepsilon(\tau) \geq \exp(\delta\varepsilon - t\varepsilon^2 C_\eta/2N)$. By the optional stopping theorem,

$$\mathrm{E}[Z_N^\varepsilon(\min(t, \tau))] = \mathrm{E}[Z_N^\varepsilon(0)] \geq \mathrm{E}[Z_N^\varepsilon(\tau)\mathbf{1}_{\{\tau \leq t\}}] \geq \mathrm{P}(\tau \leq t)\exp\left(\delta\varepsilon - \frac{t\varepsilon^2 C_\eta}{2N}\right).$$

So, $\mathrm{P}[\sup_{0 \leq t \leq T} M_N(t) > \delta] = \mathrm{P}[\tau \leq T] \leq \exp(-\delta\varepsilon + T\varepsilon^2 C_\eta/2N)$.

Finally, when $\delta \in (0, \eta C_\eta T)$, with $\varepsilon = \delta N/C_\eta t$, and applying the same argument to $-M_N(t)$, we obtain the result.

4.3. Finite-variation part

In this section we use the Lispchitz property of α and β to provide a bound for the finite-variation part, in order to apply Gronwall's lemma later.

Lemma 4.2. *There exists $C_1 > 0$ independent of N such that*

$$|Q_N(t))| \leq C(|u_N(t) - u(t)| + |V_N(t) - v(t)|).$$

Proof. We have

$$Q_N(t) = \frac{1}{N}\sum_{i=1}^N \delta_0(u_t^{(i)})\alpha(V_N(t)) - (1 - u(t))\alpha(v(t))$$

$$- \frac{1}{N}\sum_{i=1}^N \delta_1(u_t^{(i)})\beta(V_N(t)) - u(t)\beta(v(t)).$$

Let us start with the second term of the difference, called $Q^{1\to 0}$:

$$Q^{1\to 0} = \frac{1}{N}\sum_{i=1}^N \delta_1(u_t^{(i)})\beta(V_N(t)) - u(t)\beta(v_t)$$

$$= \frac{1}{N}\sum_{i=1}^N \delta_1(u_t^{(i)})\beta(V_N(t)) - u(t)\beta(V_N(t)) + u(t)(\beta(V_N(t)) - \beta(v(t)))$$

$$= \beta(V_N(t))(u_N(t) - u(t)) + \underbrace{u(t)}_{\in[0,1]}(\beta(V_N(t)) - \beta(v(t))).$$

Then,

$$|Q^{1 \to 0}| \leq \|\beta\|_\infty |u_N(t) - u(t)| + K_\beta |V - v(t)|,$$

where K_β is the Lipschitz coefficient of β. We do the same for the other term of the difference:

$$|Q^{0 \to 1}| \leq \|\alpha\|_\infty |u_N(t) - u(t)| + K_\alpha |V_N(t) - v(t)|.$$

So the proof is complete, with $C_1 = \max(\|\alpha\|_\infty, \|\beta\|_\infty, K_\alpha, K_\beta)$.

If more general transition rates $\alpha(v, u)$ and $\beta(v, u)$ depend on v and u, we would need to replace $\|\alpha\|_\infty$ and $\|\beta\|_\infty$ respectively by $\|\alpha\|_\infty + K_\alpha^{(u)}$ and $\|\beta\|_\infty + K_\beta^{(u)}$, where $K_\alpha^{(u)}$ and $K_\beta^{(u)}$ are the Lipschitz coefficients associated with the second variable u.

4.4. Proof of Theorem 2.1

Law of large numbers. We want to apply Gronwall's lemma to the function

$$f(t) = |V_N(t) - v(t)|^2 + |u_N(t) - u(t)|^2.$$

From the previous section we have a good control on the martingale term and the following estimate.

Corollary 4.1. *There exists $C_2 > 0$ independent of N such that*

$$|u_N(t) - u(t)|^2 \leq 4 \left[|u_N(0) - u(0)|^2 + C_2 T \int_0^t |u_N(s) - u(s)|^2 \, ds \right.$$
$$\left. + C_2 T \int_0^t |V_N(s) - v(s)|^2 \, ds + M_N(t)^2 \right].$$

Proof. As $u_N(t) - u(t) = u_N(0) - u(0) + M_N(t) + \int_0^t Q_N(s) \, ds$ and $(x + y + z + w)^2 \leq 4(x^2 + y^2 + z^2 + w^2)$, the result is a direct application of the previous lemma and the Cauchy–Schwarz inequality.

We now need to work on $|V_N(t) - v(t)|^2$, using hypothesis (H1), with

$$K_1 = \sup_N \sup_{s \in [0,T]} \left| \frac{\partial f}{\partial v}(V_N(s), u_N(s)) \right| \quad \text{and} \quad K_2 = \sup_N \sup_{s \in [0,T]} \left| \frac{\partial f}{\partial u}(V_N(s), u_N(s)) \right|.$$

Between the jumps, we have

$$\frac{d}{dt}(|V_N(t) - v(t)|^2) = 2(f(V_N(t), u_N(t)) - f(v(t), u(t)))(V_N(t) - v(t)).$$

Thus,

$$|V_N(t) - v(t)|^2 = 2 \int_0^t [f(V_N(s), u_N(s)) - f(v(s), u(s))](V_N(s) - v(s)) \, ds$$
$$+ |V_N(0) - v_0|^2$$
$$\leq |V_N(0) - v_0|^2 + 2K_1 \int_0^t |V_N(s) - v(s)|^2 \, ds$$
$$+ 2K_2 \int_0^t |u_N(s) - u(s)||V_N(s) - v(s)| \, ds$$
$$\leq |V_N(0) - v_0|^2 + 2K_1 \int_0^t |V_N(s) - v(s)|^2 \, ds$$
$$+ K_2 \int_0^t |u_N(s) - u(s)|^2 \, ds + K_2 \int_0^t |V_N(s) - v(s)|^2 \, ds,$$

where we have successively used the Cauchy–Schwarz inequality and the fact that $ab \leq \frac{1}{2}(a^2 + b^2)$. Putting this inequality together with the corollary we obtain

$$f(t) \leq A + B \int_0^t f(s) \, ds,$$

where $B = B(T) = \max(2K_1(T) + K2(T), C_2 T)$ does not depend on N and is linear with respect to T if (H2) holds, and

$$A = |u_N(0) - u_0|^2 + |V_N(0) - v_0|^2 + K_A \sup_{0 \leq s \leq T} M_s^2.$$

If we control the initial conditions then, with the control we have on the martingale part from Proposition 4.1, A can be chosen arbitrarily small (with high probability) and we can conclude with Gronwall's lemma.

Exponential convergence speed. If the initial conditions are the same for the stochastic and deterministic models, we actually have exponentially fast convergence, thanks to the exponential bound for the martingale part given by Proposition 4.2: there exists a constant $C_m > 0$ such that

$$\limsup_{N \to \infty} \frac{1}{N} \log P\left[\sup_{0 \leq t \leq T} |V_N(t) - v(t)|^2 + |u_N(t) - u(t)|^2 > \Delta \right] \leq -\frac{\Delta e^{-B(T)T}}{2K_A C_m T}.$$

Indeed, from Gronwall's lemma we deduce that

$$P\left[\sup_{0 \leq t \leq T} |V_N(t) - v(t)|^2 + |u_N(t) - u(t)|^2 > \Delta \right] \leq P\left[\sup_{0 \leq t \leq T} |M_s|^2 > \frac{\Delta e^{-B(T)T}}{K_A} \right].$$

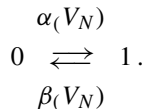Then, from Proposition 4.2 we know that there exists a constant C_m such that

$$\frac{1}{N} \log P\left[\sup_{0 \leq t \leq T} |M_s|^2 > \delta^2 \right] \leq \frac{\log(2)}{N} - \frac{\delta^2}{2C_m T}.$$

We complete the proof with $\delta^2 = \Delta e^{-B(T)T}/K_A$.

5. Proof of the central limit theorems

As before, we write the proofs for the case of a single channel type with state space $\{0, 1\}$ and transition rates given by the scheme

$$0 \underset{\beta(V_N)}{\overset{\alpha(V_N)}{\rightleftarrows}} 1 \, .$$

5.1. Functional central limit theorem

Let $b(u, v) = (1 - u)\alpha(v) - u\beta(v)$, and let (V_N, u_N) be the solution of the simplified stochastic model (S_N) and (V, u) be the solution of the deterministic model (D) introduced in the Example 2.1. Consider the process

$$\begin{pmatrix} P_N \\ Y_N \end{pmatrix} = \begin{pmatrix} \sqrt{N}(u_N - u) \\ \sqrt{N}(V_N - V) \end{pmatrix}.$$

If the initial conditions satisfy $(P_N(0), Y_N(0)) = (0, 0)$, the two-dimensional process (P_N, Y_N) converges in law, as $N \to \infty$, towards the process (P, Y), with characteristic function

$$E[\exp(i(\theta_1 P(t) + \theta_2 Y(t)))] = \exp(\theta_1^2 A(t) + \theta_2^2 B(t) + \theta_1 \theta_2 C(t)).$$

The functions A, B, and C are solutions of the system

$$\begin{pmatrix} A' \\ B' \\ C' \end{pmatrix} = \begin{pmatrix} 2b'_u & 0 & b'_v \\ 0 & 2f'_v & f'_u \\ 2f'_u & 2b'_v & b'_u + f'_v \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} + \begin{pmatrix} -\frac{1}{2}\lambda(V, u) \\ 0 \\ 0 \end{pmatrix} \quad (5.1)$$

with initial conditions $(0, 0, 0)$, and with $\lambda(v, u) = \sqrt{(1 - u)\alpha(v) + u\beta(v)}$.

Proof of Theorem 2.2. We adapt the proof given in [22]: we prove the convergence of characteristic functions plus tightness. Let

$$\phi_N(t, \theta) = E[\exp(i(\theta_1 P_N(t) + \theta_2 Y_N(t)))].$$

Let us also define $Z_N = (u_N - u, V_N - V)$, $X_N = (u_N, V_N)$, $X = (u, V)$, and $h(x, y) = \exp(i\sqrt{N}(\theta_1 x + \theta_2 y))$. Then

$$\begin{aligned} \phi(t, \theta) - 1 &= E[h(Z_N(t)) - h(Z_N(0))] \\ &= \int_0^t E\bigg[N\lambda(X_N(s)) \int_{E_N} \{h(w - u(s), V_N(s) - V(s)) \\ &\qquad\qquad - h(Z_N(s))\}\mu(X_N(s), \, dw) \\ &\qquad - b(X(s))h'_x(Z_N(s)) + (f(X_N(s)) - f(X(s)))h'_y(Z_N(s)) \bigg] \, ds. \end{aligned}$$

So $\phi_N(t, \theta) - 1 = G_N(\theta, t) + H_N(\theta, t)$ with

$$G_N(\theta, t) = \int_0^t E[\Omega_N(s)\{(e^{i\theta_1\sqrt{N}/N} - 1)\mu_+ + (e^{-i\theta_1\sqrt{N}/N} - 1)\mu_-\}] \, ds,$$

$$\Omega_N(s) = N\lambda(X_N(s))h(Z_N(s)), \qquad \mu_{+/-} = \mu_{+/-}(X_N(s)),$$

$$H_N(\theta, t) = \int_0^t i E[-\theta_1\sqrt{N}b(X(s))h(Z_N(s)) + \theta_2\sqrt{N}\{f(X_N(s)) - f(X(s))\}] \, ds.$$

Then in order to use the asymptotic development of e^x when $x \to 0$, we introduce the function $K(u) = e^{iu} - 1 - iu + u^2/2$. Then, knowing that $\mu_+ + \mu_- = 1$,

$$G_N(\theta, t) = \int_0^t E\bigg[\Omega_N(s)\bigg\{i\frac{\theta_1}{\sqrt{N}}(\mu_+ - \mu_-)(X_N(s)) - \frac{\theta_1^2}{2N} + K\bigg(\frac{\theta_1}{\sqrt{N}}\bigg)\bigg\}\bigg] \, ds.$$

Since $b(x) = \lambda(x)(\mu_+(x) - \mu_-(x))$, we have

$$\begin{aligned} G_N(\theta, t) = \int_0^t E[i\theta_1\sqrt{N}b(X_N(s))h(Z_N(s))] \, ds &+ \int_0^t E\bigg[-\frac{\theta_1^2}{2}\lambda(X_N(s))h(Z_N(s))\bigg] \, ds \\ &+ \int_0^t E\bigg[NK\bigg(\frac{\theta_1}{\sqrt{N}}\bigg)h(Z_N(s))\bigg] \, ds. \end{aligned}$$

Therefore,

$$\phi_N(t,\theta) - 1 = \int_0^t \mathrm{E}\left[-\frac{1}{2}\theta_2^2\lambda(X_N(s))h(Z_N(s))\right]\mathrm{d}s$$

$$+ \int_0^t \mathrm{E}[h(Z_N(s))\mathrm{i}\theta_1\sqrt{N}\{b(X_N(s)) - b(X(s))\}]\,\mathrm{d}s \tag{5.2}$$

$$+ \int_0^t \mathrm{E}[h(Z_N(s))\mathrm{i}\theta_2\sqrt{N}\{f(X_N(s)) - f(X(s))\}]\,\mathrm{d}s \tag{5.3}$$

$$+ \int_0^t \mathrm{E}\left[h(Z_N(s))NK\left(\frac{\theta_1}{\sqrt{N}}\right)\lambda(X_N(s))\right]\mathrm{d}s. \tag{5.4}$$

Using the derivatives of b and f, and the convergence of X_N to X, we can make a development of the sum (5.2) + (5.3):

$$(5.2) + (5.3) = \int_0^t \mathrm{E}[h(Z_N)\mathrm{i}\sqrt{N}\{(u_N - u)(\theta_1 b_u' + \theta_2 f_u') + (V_N - V)(\theta_1 b_v' + \theta_2 f_v')\}]\,\mathrm{d}s$$
$$+ \varepsilon_N(t,\theta),$$

where we have dropped the s, and where b_u', b_v', f_u', and f_v' are taken at $X_N(s)$.

Noting that $h(Z_N)\mathrm{i}\sqrt{N}(u_N - u) = h_x'(Z_N)$ and $h(Z_N)\mathrm{i}\sqrt{N}(V_N - V) = h_y'(Z_N)$, we have

$$(5.2) + (5.3) = \int_0^t \mathrm{E}[h_x'(Z_N)(\theta_1 b_u' + \theta_2 f_u') + h_y'(Z_N)(\theta_1 b_v' + \theta_2 f_v')]\,\mathrm{d}s.$$

The term in (5.4) converges to 0 as $N \to \infty$ by dominated convergence since $K(u)/u^2$ is bounded and converges to 0.

As we have the convergence in Theorem 2.1 of X_N to X, we obtain the convergence of $\phi_N(t,\theta)$ to $\Psi(t,\theta)$, satisfying

$$\frac{\partial\Psi}{\partial t}(t,\theta) = -\frac{1}{2}\theta_1^2\lambda(X(t))\Psi(t,\theta) + (\theta_1 b_u'(X(t)) + \theta_2 f_u'(X(t)))\frac{\partial\Psi}{\partial\theta_1}$$

$$+ (\theta_1 b_v'(X(t)) + \theta_2 f_v'(X(t)))\frac{\partial\Psi}{\partial\theta_2}.$$

Tightness stems from the Markov property and the following estimate obtained in the proof of Theorem 2.1:

$$\mathrm{P}\left[\sup_{0\leq t\leq T}\|\sqrt{N}[V_N(t) - v(t)]\|^2 + \sum_{j=1}^q \|\sqrt{N}[e_N^{(j)}(t) - g^{(j)}(t)]\|^2 > \Delta\right]$$

$$\leq \exp\left(-\frac{(\Delta/N)N\mathrm{e}^{-B(T)T}}{CT}\right).$$

The announced convergence in law follows.

To solve the partial differential equation, we set $\Psi(t,\theta) = \exp(\theta_1^2 A(t) + \theta_1\theta_2 C(t) + \theta_2^2 B(t))$. Then, substituting into the initial equation, and identifying the coefficients, we obtain system (5.1).

Proof of Theorem 2.3. (This proof is inspired from [9].) The convergence of X_N to X almost surely (a.s.), uniformly on finite time intervals, obtained in Theorem 2.1, implies that $\tau_N \to \tau$ a.s. In order to apply Theorem 2.2, let us introduce Z_N through the following decomposition:

$$\sqrt{N}(\phi(X(\tau)) - \phi(X(\tau_N))) = \sqrt{N}\left[\phi\left(X(\tau_N) + \frac{1}{\sqrt{N}}Z_N(\tau_N)\right) - \phi(X(\tau_N))\right]$$
$$- \sqrt{N}\phi(X_N(\tau_N)),$$

where we have used the fact that $\phi(X(\tau)) = 0$. As $N \to \infty$, we claim that the right-hand side converges in law to $\nabla\phi(X(\tau))Z(\tau)$ since $\sqrt{N}\phi(X_N(\tau_N))$ converges in law to 0. Indeed, as $\phi(X_N(\tau_N)) \leq 0$ and $\phi(X_N(\tau_N^-)) \geq 0$,

$$|\sqrt{N}\phi(X_N(\tau_N))| \leq |\sqrt{N}(\phi(X_N(\tau_N)) - \phi(X_N(\tau_N^-)))|.$$

There exists θ_N on the line between $X_N(\tau_N)$ and $X_N(\tau_N^-)$ such that

$$|\sqrt{N}(\phi(X_N(\tau_N)) - \phi(X_N(\tau_N^-)))| = |\nabla\phi(\theta_N)(Z_N(\tau_N) - Z_N(\tau_N^-))|,$$

which converges in law to 0 since $Z_N \to Z$ and Z is continuous. The claim follows. By continuity, $\phi(X(\tau)) = 0$, so that $\sqrt{N}(\phi(X(\tau)) - \phi(X(\tau_N)))$ is asymptotic to

$$-\nabla\phi(X(\tau))F(X(\tau))\sqrt{N}(\tau_N - \tau).$$

Thus, $\sqrt{N}(\tau_N - \tau)$ converges in law to $\pi(\tau)$. To complete the proof, we remark that

$$\sqrt{N}(X_N(\tau_N) - X(\tau)) = Z_N(\tau_N) + \sqrt{N}(X(\tau_N) - X(\tau)),$$

which converges in law to $Z(\tau) + \pi(\tau)F(X(\tau))$.

5.2. Langevin approximation

In this case Theorem 2.4 can be written as follows. Let (V_N, u_N) be the solution of the stochastic model (S_N). Then, the process

$$R_N(t) = \sqrt{N}\left(u_N(t) - u_N(0) - \int_0^t b(u_N(s), V_N(s))\,ds\right)$$

converges in law, as $N \to \infty$, towards the process $R(t)$ defined as a stochastic integral:

$$R(t) = \int_0^t \sqrt{(1 - u(s))\alpha(V(s)) + u(s)\beta(V(s))}\,dB_s,$$

where B is a standard Brownian motion and $(u(t), V(t))$ is the unique solution of

$$\dot{V} = f(V, u), \qquad \dot{u} = (1 - u)\alpha(V) - u\beta(V),$$
$$u(0) = u_0 = u_N(0), \qquad V(0) = V_0 = V_N(0), \quad \text{for all } N.$$

This result provides the following degenerate diffusion approximation $(\tilde{V}_N, \tilde{u}_N)$ for sufficiently large N:

$$d\tilde{V}_N(t) = f(\tilde{V}_N(t), \tilde{u}_N(t))\,dt,$$
$$d\tilde{u}_N(t) = [(1 - \tilde{u}_N(t))\alpha(\tilde{V}_N(t)) - \tilde{u}_N(t)\beta(\tilde{V}_N(t))]\,dt + \sigma_N(\tilde{u}_N(t), \tilde{V}_N(t))\,dB_t,$$
$$\sigma_N(u)^2 = \frac{1}{N}[(1 - u)\alpha(v) + u\beta(v)] = \frac{1}{N}\lambda(v, u).$$

Let

$$g(u, v) = \lambda_N(u, v)\left[\frac{1}{N^2}\mu^+(u, v) + \frac{1}{N^2}\mu^-(u, v)\right] = (1 - u)\alpha(v) + u\beta(v).$$

Note that in the multidimensional case, the real-valued function g above becomes a $d \times d$ matrix. Since the different channel types j are supposed to be independent, this matrix would be block diagonal, with blocks of size r_j, thus assuring the independence of the q (r_j-dimensional Brownian motions) $W^{(j)}$. The blocks of size r_j are given by the matrix $G^{(j)}$, and arise from the calculation of the covariances:

$$G_{i,k}^{(j)}(x) = N\lambda_N(x)\int_E z_i z_k \mu_N(x, z)\, dz.$$

Proof of Theorem 2.4. The tightness property follows from the inequality

$$P\left[\sup_{s \leq T} |R_N(s)| > \delta\right] \leq \frac{TN}{\delta^2}\|g\|_\infty.$$

Just as in the proof of Theorem 2.2, let us define $\phi_N(t, \theta) = E[e^{i\theta R_N(t)}]$, the characteristic function of R_N. Let $h(M_N(t)) = e^{i\theta R_N(t)}$, $\sqrt{N}M_N(t) = R_N(t)$, $\psi(u) = (e^{iu} - 1 - iu + u^2/2)/u^2$, and $\xi(u) = e^{iu} - 1 - iu = u^2\psi(u) - u^2/2$. We then have

$$\begin{aligned}
\phi_N(t, \theta) - 1 &= E[h(M_N(t))] - h(0) \\
&= \int_0^t E\left[\lambda_N(s)\int_{E_N} h(w - u_N(s) + M_N(s)) - h(M_N(s))\right. \\
&\qquad\qquad \left. - (w - u_N(s))h'(M_N(s))\mu_N(s, \, dw)\right] ds \\
&= \int_0^t E\left[e^{i\theta R_N(s)}\lambda_N(s)\int_{E_N}\xi(\theta\sqrt{N}(w - u_N(s)))\mu_N(s, \, dw)\right] ds \\
&= -\int_0^t E\left[\frac{1}{2}e^{i\theta R_N(s)}\lambda_N(s)\int_{E_N} N\theta^2(w - u_N(s))^2\mu_N(s, \, dw)\right] ds \\
&\quad + \int_0^t E\left[e^{i\theta R_N(s)}\lambda_N(s)\int_{E_N} N\theta^2(w - u_N(s))^2\right. \\
&\qquad\qquad\qquad \left. \times \psi(\sqrt{N}\theta(w - u_N(s)))\mu_N(s, \, dw)\right] ds,
\end{aligned}$$

where $\lambda_N(s)$ stands for $\lambda_N(u_N(s), V_N(s))$ and $\mu_N(s, \, dw)$ stands for $\mu_N(u_N(s), V_N(s), \, dw)$. The second term in the last equality, call it $K_N(\theta)$, converges to 0 as $N \to \infty$ by dominated convergence, and because $\psi(\sqrt{N}\theta(w - u_N(s))) = \psi(\pm\theta/\sqrt{N}) \to 0$ as $\lim_{u \to 0}\psi(u) = 0$. So we have

$$\begin{aligned}
\phi_N(t, \theta) - 1 &= -\int_0^t E\left[\frac{1}{2}e^{i\theta R_N(s)}\theta^2 g(u_N(s), V_N(s))\right] ds + K_N(\theta) \\
&= -\frac{1}{2}\int_0^t \theta^2 g(u(s), V(s))\phi_N(s, \theta)\, ds \\
&\quad + \frac{1}{2}\int_0^t \theta^2 E[(g(u(s), V(s)) - g(u_N(s), V_N(s)))e^{i\theta R_N(s)}]\, ds + K_N(\theta).
\end{aligned}$$

Again, the second term in the last equality, call it $J_N(\theta)$, converges to 0 as $N \to \infty$, because of the convergence of u_N and V_N to u and V (cf. Theorem 2.1).

By Gronwall's lemma we conclude that $\phi_N(t, \theta) \to \phi(t, \theta)$ with

$$\phi(t, \theta) = \exp\left(-\frac{1}{2}\theta^2 \int_0^t g(u(s), V(s)) \, ds\right).$$

Proof of Theorem 2.5. We want to prove that the process Z has the same law as the limit as $N \to \infty$ of the difference between the Langevin approximation linearized around the deterministic solution and the deterministic solution itself, scaled by $\sqrt{N}$. We write it in the general case, not in dimension two only, as above. First we identify the equations satisfied by the moments of Z starting from the equation satisfied by the characteristic function. We make the ansatz

$$\psi(t, \theta) = e^{-\theta \Gamma(t) \theta^\top /2}.$$

The matrix Γ_t corresponds to the variance/covariance matrix. We substitute this expression into the equation satisfied by ψ as given in Theorem 2.2:

$$\frac{\partial \Psi}{\partial t} = \sum_{j=1}^q \left\{ \sum_{l \in L} \sum_{k=1}^{r_j} \theta_k^{(j)} \frac{\partial b_k^{(j)}}{\partial x_l} \frac{\partial \Psi}{\partial \theta_l} - \frac{1}{2} \sum_{k,l=1}^{r_j} \theta_k^{(j)} \theta_l^{(j)} G_{k,l}^{(j)} \Psi \right\}$$
$$+ \sum_{m=1}^p \sum_{l \in L} \theta_m \frac{\partial f^m}{\partial x_l} \frac{\partial \Psi}{\partial \theta_l}.$$

The ensemble of indices L can be written as $L = L_v \cup L_u$, where $L_v = \{1 \le m \le p\}$ and $L_u = \{(j, k); 1 \le j \le q, 1 \le k \le r_j\}$. To identify the equations satisfied by Γ_{ab}, we distinguish the following cases:

- for $a \in L_v$ and $b \in L_v$,

$$\frac{1}{2}\Gamma'_{ab} = \sum_{l \in L}\left[\frac{\partial f^a}{\partial x_l}\Gamma_{bl} + \frac{\partial f^b}{\partial x_l}\Gamma_{al}\right];$$

- for $a \in L_v$ and $b \in L_u$, $b = (j, k)$,

$$\frac{1}{2}\Gamma'_{ab} = \sum_{l \in L}\left[\frac{\partial b_k^{(j)}}{\partial x_l}\Gamma_{al} + \frac{\partial f^a}{\partial x_l}\Gamma_{bl}\right];$$

- for $a \in L_u$, $a = (j, k)$, and $b \in L_v$,

$$\frac{1}{2}\Gamma'_{ab} = \sum_{l \in L}\left[\frac{\partial b_k^{(j)}}{\partial x_l}\Gamma_{bl} + \frac{\partial f^b}{\partial x_l}\Gamma_{al}\right];$$

- for $a \in L_u$, $a = (j, k)$, and $b \in L_v$, $b = (j', k')$,

$$\frac{1}{2}\Gamma'_{ab} = \sum_{l \in L}\left[\frac{\partial b_k^{(j)}}{\partial x_l}\Gamma_{bl} + \frac{\partial b_{j',k'}}{\partial x_l}\Gamma_{al}\right] + \frac{1}{2}G_{k,k'}^{(j)}\mathbf{1}_{\{j=j'\}}.$$

We then write the equations satisfied by $K^{(N)}(t) = \sqrt{N}(\tilde{X}_N(t) - X(t)) = (Y_N^m, P_N^{j,k})$, where $\tilde{X}_N$ is the Langevin approximation defined in Section 2.4, and where X is the deterministic limit:

$$
\begin{aligned}
\mathrm{d}Y_N^m &= \sqrt{N}(f^m(\tilde{X}_N) - f^m(X))\,\mathrm{d}t, \\
\mathrm{d}P_N^{j,k} &= \sqrt{N}(b_k^{(j)}(\tilde{X}_N) - b_k^{(j)}(X))\,\mathrm{d}t + \sigma^{(j)}(\tilde{X}_N)\,\mathrm{d}W_t^j.
\end{aligned}
$$

When we linearize around the deterministic solution, we obtain the following equations:

$$
\mathrm{d}\tilde{Y}_N^m = \sum_{l \in L} \frac{\partial f^m}{\partial x_l} K_l^{(N)}\,\mathrm{d}t,
$$

$$
\mathrm{d}\tilde{P}_N^{j,k} = \sum_{l \in L} \frac{\partial b_k^{(j)}}{\partial x_l} K_l^{(N)}\,\mathrm{d}t + \left(\sum_{k'=1}^{r_j} \sigma_{k,k'}^{(j)}(X) + \frac{1}{\sqrt{N}} \Omega_{k,k'}^{(j)} \right) \mathrm{d}W_t^{j,k'},
$$

where the terms $\Omega_{k,k'}^{(j)}/\sqrt{N}$ come from the linearization of $\sigma_{k,k'}^{(j)}(\tilde{X}_N)$; we do not need to specify them here because they go to 0 as $N \to \infty$.

It is now clear that the moment equations for this linear diffusion system converge to the system satisfied by Γ_{ab} as $N \to \infty$.

Appendix A. Comparison between two deterministic limits of different stochastic Hodgkin–Huxley models

We want to compare the two following deterministic systems, with f, α, β continuously differentiable functions, α and β nonnegative, and k an integer greater than or equal to 1:

$$
\frac{\mathrm{d}V}{\mathrm{d}t} = f(V, u^k), \qquad \frac{\mathrm{d}u}{\mathrm{d}t} = (1-u)\alpha(V) - u\beta(V), \tag{A.1}
$$

$$
\begin{aligned}
\frac{\mathrm{d}\hat{V}}{\mathrm{d}t} &= f(\hat{V}, x_k), \\
\frac{\mathrm{d}x_j}{\mathrm{d}t} &= (k-j+1)x_{j-1}\alpha(\hat{V}) + (j+1)x_{j+1}\beta(\hat{V}) - x_j(j\beta(\hat{V}) + (k-j)\alpha(\hat{V}))
\end{aligned} \tag{A.2}
$$

for all $0 \le j \le k$.

System (A.1) corresponds to the classical 'Hodgkin–Huxley' model, with only two variables for simplicity, and system (A.2) is a $(k+2)$-dimensional system, where x_j, $0 \le j \le k$, is the proportion of channels in state j, and $j = k$ is the *open* state.

Proposition A.1. *Let $V_0 \in \mathbb{R}$ and $u_0 \in [0,1]$. If the following conditions on the initial values are satisfied:*

$$
V(0) = \hat{V}(0) = V_0
$$

$$
\text{and} \quad C_k^j u(0)^{k-j}(1-u(0))^j = x_{k-j}(0) = C_k^j u_0^{k-j}(1-u_0)^j \quad \text{for all } 0 \le j \le k,
$$

then, for all $t \ge 0$, $V(t) = \hat{V}(t)$ (same potential) and $u(t)^k = x_k(t)$ (the proportion of open channels is $u(t)^k$).

Moreover, for all $1 \le j \le k$ and all $t \ge 0$, $x_{k-j}(t) = C_k^j u(t)^{k-j}(1-u(t))^j$.

Proof. Consider (V, u) to be the unique solution of (A.1) for $V(0) = V_0$ and $u(0) = u_0$. Let $y_j(t) = C_k^j u(t)^{k-j}(1-u(t))^j$, $0 \le j \le k$. Then $(V, y_k, \ldots, y_0)$ is a solution of (A.2)

(we just need to compute y'_j and write it as a function of y_{j-1} and y_{j+1}). As the initial values are equal (by hypothesis), $x_{k-j}(0) = C_k^j u_0^{k-j}(1-u_0)^j = y_{k-j}(0)$ and, by uniqueness, $(V, y_k, \ldots, y_0) = (\hat{V}, x_k, \ldots, x_0)$ for all $t \geq 0$.

Remark. The result is essentially the same for more complicated Markov schemes (e.g. the sodium multistate Markov model).

Appendix B. Moment equations for the linearized Langevin approximation

From Theorem 2.2 we can build a diffusion approximation $(\tilde{V}_N, \tilde{e}_N)$ of the stochastic hybrid process (V_N, e_N) given in the Example 2.1:

$$d\tilde{V}_N(t) = f(\tilde{V}_N(t), \tilde{e}_N(t)) \, dt,$$

$$d\tilde{u}_N(t) = b(\tilde{u}_N(t), \tilde{V}_N(t)) \, dt + \sqrt{\frac{\lambda(\tilde{u}_N(t), \tilde{V}_N(t))}{N}} \, dB_t,$$

$$b(v, u) = [(1-u)\alpha(v) - u\beta(v)],$$

$$\lambda(v, u) = [(1-u)\alpha(v) + u\beta(v)].$$

We want to write the moment equations for the linearized version of

$$\begin{pmatrix} \tilde{P}_N \\ \tilde{Y}_N \end{pmatrix} = \begin{pmatrix} \sqrt{N}(\tilde{e}_N - e) \\ \sqrt{N}(\tilde{V}_N - V) \end{pmatrix}$$

with (V, e) being the deterministic solution. The linearized equations are given by

$$dY_N^L = (f'_V Y_N^L + f'_u P_N^L) \, dt,$$

$$dP_N^L = (b'_V Y_N^L + b'_u P_N^L) \, dt + \left[\sqrt{\lambda_t} + \frac{1}{2\sqrt{N\lambda_t}} (\lambda'_V Y_N^L + \lambda'_u P_N^L) \right] dB_t,$$

with $\lambda_t = \lambda(V(t), u(t))$. We define $m_1^N = E[Y_N^L]$, $m_2^N = E[P_N^L]$, $S_1^N = E[(Y_N^L - m_1)^2]$, $S_2^N = E[(P_N^L - m_2)^2]$, and $C_{12}^N = E[(Y_N^L - m_1)(P_N^L - m_2)]$. Then we have the following system of five equations:

$$\frac{dm_1^N}{dt} = f'_V m_1 + f'_u m_2,$$

$$\frac{dm_2^N}{dt} = b'_V m_1 + b'_u m_2,$$

$$\frac{dS_1^N}{dt} = 2f'_V S_1 + 2f'_u C_{12},$$

$$\frac{dS_2^N}{dt} = 2b'_u S_2 + 2b'_V C_{12} + \left(\sqrt{\lambda_t} + \frac{1}{2\sqrt{N\lambda}} (\lambda'_N m_1 + \lambda'_e m_2) \right)^2$$
$$+ \left(\frac{\lambda'_V}{2\sqrt{N\lambda_t}} \right)^2 S_1 + \left(\frac{\lambda'_u}{2\sqrt{N\lambda_t}} \right)^2 S_2 + 2\frac{\lambda'_V \lambda'_u}{4N\lambda_t} C_{12},$$

$$\frac{dC_{12}^N}{dt} = b'_V S_1 + f'_u S_2 + (f'_V + b'_u) C_{12}.$$

At the limit $N \to \infty$, and with $A = -2S_2$, $B = -2S_1$, and $C = -C_{12}$, this system is the same as the one found in the application of Theorem 2.2 in Section 3.

Appendix C. Moment equations for the Morris–Lecar system

The moment equations used in Sections 3.5.1 and 3.5.2 are the following linear nonhomogeneous system of differential equations:

$$
\begin{bmatrix} S_m \\ S_n \\ S_v \\ C_{mv} \\ C_{nv} \\ C_{mn} \end{bmatrix} = M(t) \begin{bmatrix} S_m \\ S_n \\ S_n \\ C_{mv} \\ C_{nv} \\ C_{mn} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

with

$$
M(t) = \begin{bmatrix}
2\dfrac{\partial F_m}{\partial m} & 0 & 0 & 2\dfrac{\partial F_m}{\partial V} & 0 & 0 \\
0 & 2\dfrac{\partial F_n}{\partial n} & 0 & 0 & 2\dfrac{\partial F_n}{\partial V} & 0 \\
0 & 0 & 2\dfrac{\partial F_v}{\partial V} & 2\dfrac{\partial F_v}{\partial m} & 02\dfrac{\partial F_v}{\partial n} & 0 \\
\dfrac{\partial F_v}{\partial m} & 0 & \dfrac{\partial F_m}{\partial V} & \dfrac{\partial F_v}{\partial V} + \dfrac{\partial F_m}{\partial m} & 0 & \dfrac{\partial F_v}{\partial n} \\
0 & \dfrac{\partial F_v}{\partial n} & \dfrac{\partial F_n}{\partial V} & 0 & \dfrac{\partial F_v}{\partial V} + \dfrac{\partial F_n}{\partial n} & \dfrac{\partial F_v}{\partial m} \\
0 & 0 & 0 & \dfrac{\partial F_n}{\partial V} & \dfrac{\partial F_m}{\partial V} & \dfrac{\partial F_m}{\partial m} + \dfrac{\partial F_n}{\partial n}
\end{bmatrix},
$$

where all the functions are evaluated at $X(t) = (V(t), m(t), n(t))$, the solution of (3.1)–(3.3), and $B_1(t) = (1 - m(t))\alpha_m(V(t)) + m(t)\beta_m(V(t))$ and $B_2(t) = (1 - n(t))\alpha_n(V(t)) + n(t)\beta_n(V(t))$.

References

[1] AUSTIN, T. D. (2008). The emergence of the deterministic Hodgkin–Huxley equations as a limit from the underlying stochastic ion-channel mechanism. *Ann. Appl. Prob.* **18,** 1279–1325.
[2] BLOM, H. A. P. AND LYGEROS, J. (eds) (2006). Stochastic Hybrid Systems (Lecture Notes Control Inf. Sci. **337**). Springer.
[3] CHICONE, C. (1999). *Ordinary Differential Equations with Applications*. Springer, New York.
[4] CRONIN, J. (1987). *Mathematical Aspects of Hodgkin–Huxley Neural Theory*. Cambridge University Press.
[5] DARLING, R. W. R. AND NORRIS, J. R. (2005). Structure of large random hypergraphs. *Ann. Appl. Prob.* **15,** 125–152.
[6] DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Statist. Soc. B* **46,** 353–388.
[7] DEFELICE, L. J. AND ISAAC, A. (1993). Chaotic states in a random world: relationship between the nonlinear differential equations of excitability and the stochastic properties of ion channels. *J. Statist. Phys.* **70,** 339–354.
[8] DESTEXHE, A., MAINEN, Z. F. AND SEJNOWSKI, T. J. (1994). Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *J. Comput. Neuroscience* **1,** 195–230.

[9] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. John Wiley, New York.

[10] FAISAL, A. A., SELEN, L. P. J. AND WOLPERT, D. M. (2008). Noise in the nervous system. *Nature Rev. Neurosci.* **9,** 292–303.

[11] FOX, R. F. AND LU, Y.-N. (1994). Emergent collective behaviour in large numbers of globally coupled independently stochastic ion channels. *Phys. Rev. E* **49,** 3421–3431.

[12] GOLLISCH, T. AND MEISTER, M. (2008). Rapid neural coding in the retina with relative spike latencies. *Science* **319,** 1108–1111.

[13] GUCKENHEIMER, J. AND OLIVA, R. A. (2002). Chaos in the Hodgkin–Huxley Model. *SIAM J. Appl. Dynam. Systems* **1,** 105–114.

[14] HESPANHA, J. P. (2005). A model for stochastic hybrid systems with application to communication networks. *Nonlinear Anal.* **62,** 1353–1383.

[15] HODGKIN, A. L. AND HUXLEY, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiology* **117,** 500–544.

[16] IZHIKEVICH, E. M. (2006). *Dynamical Systems in Neuroscience*. MIT Press, Cambridge, MA.

[17] JOSIĆ, K. AND ROSENBAUM, R. (2008). Unstable solutions of nonautonomous linear differential equations. *SIAM Rev.* **50,** 570–584.

[18] KALLENBERG, O. (1997). *Foundations of Modern Probability*. Springer, New York.

[19] KEENER, J. P. (2009). Invariant manifold reductions for markovian ion channel dynamics. *J. Math. Biol.* **58,** 447–57.

[20] KOURETAS, P., KOUTROUMPAS, K., LYGEROS, J. AND LYGEROU, Z. (2006). Stochastic hybrid modeling of biochemical processes. In *Stochastic Hybrid Systems*, CRC Press, Boca Raton, FL, pp. 221–248.

[21] KRAUSZ, H. I. AND FRIESEN, W. O. (1977). The analysis of nonlinear synaptic transmission. *J. General Physiology* **70,** 243–265.

[22] KURTZ, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl Prob.* **8,** 344–356.

[23] LYGEROS, J. *et al.* (2008). Stochastic hybrid modeling of DNA replication across a complete genome. *Proc. Nat. Acad. Sci. USA* **105,** 12295–12300.

[24] MAINEN, Z. F., JOERGES, J., HUGUENARD, J. R. AND SEJNOWSKI, T. J. (1995). A model of spike initiation in neocortical pyramidal neurons. *Neuron* **15,** 1427–1439.

[25] MORRIS, C. AND LECAR, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **35,** 193–213.

[26] PAKDAMAN, K., THIEULLEN, M. AND WAINRIB, G. (2010). Diffusion approximation of birth–death processes: comparison in terms of large deviations and exit points. *Statist. Prob. Lett.* **80,** 1121–1127.

[27] PANKRATOVA, E. V., POLOVINKIN, A. V. P. AND MOSEKILDE, E. (2005). Resonant activation in a stochastic Hodgkin–Huxley model: interplay between noise and suprathreshold driving effects. *Europ. Phys. J. B* **45,** 391–397.

[28] PROTTER, P. F. (2004). *Stochastic Integration and Differential Equations*, 2nd edn. Springer, Berlin.

[29] RINZEL, J. AND MILLER, R. N. (1980). Numerical calculation of stable and unstable periodic solutions to the Hodgkin–Huxley equations. *Math. Biosci.* **49,** 27–59.

[30] ROWAT, P. (2007). Interspike interval statistics in the stochastic Hodgkin–Huxley model: coexistence of gamma frequency bursts and highly irregular firing. *Neural Computation* **19,** 1215–1250.

[31] SEGUNDO, J. P. *et al.* (1994). Noise and the neurosciences: a long history, a recent revival and some theory. In *Origins: Brain and Self Organization*, Lawrence Erlbaum, Hillsdale, pp. 299–331.

[32] SHUAI, J. W. AND JUNG, P. (2003). Optimal ion channel clustering for intracellular calcium signaling. *Proc. Nat. Acad. Sci. USA* **100,** 506–512.

[33] SKAUGEN, E. AND WALLŒ, L. (1979). Firing behaviour in a stochastic nerve membrane model based upon the Hodgkin–Huxley equations. *Acta Physiology Scand.* **107,** 343–363.

[34] STEINMETZ, P. N. *et al.* (2000). Subthreshold voltage noise due to channel fluctuations in active neuronal membranes. *J. Comput. Neurosci.* **9,** 133–148.

[35] TOUBOUL, J. AND FAUGERAS, O. (2008). A characterization of the first hitting time of double integral processes to curved boundaries. *Adv. Appl. Prob.* **40,** 501–528. (Correction: **40** (2009), 309.)

[36] TUCKWELL, H. C. (1987). Diffusion approximations to channel noise. *J. Theoret. Biol.* **127,** 427–438.

[37] VANDENBERG, C. A. AND BEZANILLA, F. (1991). A sodium channel gating model based on single channel, macroscopic ionic, and gating currents in the squid giant axon. *Biophys. J.* **60,** 1511–1533.

[38] VAN KAMPEN, N. G. (1981). *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam.

[39] VANRULLEN, R., GUYONNEAU, R. AND THORPE, S. J. (2005). Spike times make sense. *Trends Neurosci.* **28,** 1–4.

[40] WHITE, J. A., RUBINSTEIN, J. T. AND KAY, A. R. (2000). Channel noise in neurons. *Trends Neurosci.* **23,** 131–137.