

# AVERAGING FOR A FULLY COUPLED PIECEWISE-DETERMINISTIC MARKOV PROCESS IN INFINITE DIMENSIONS

ALEXANDRE GENADOT\* \*\* AND

MICHÈLE THIEULLEN,\* \*\*\* *Université Pierre et Marie Curie*

## Abstract

In this paper we consider the generalized Hodgkin–Huxley model introduced in Austin (2008). This model describes the propagation of an action potential along the axon of a neuron at the scale of ion channels. Mathematically, this model is a fully coupled piecewise-deterministic Markov process (PDMP) in infinite dimensions. We introduce two time scales in this model in considering that some ion channels open and close at faster jump rates than others. We perform a slow–fast analysis of this model and prove that, asymptotically, this ‘two-time-scale’ model reduces to the so-called averaged model, which is still a PDMP in infinite dimensions, for which we provide effective evolution equations and jump rates.

*Keywords:* Piecewise-deterministic Markov process; fully coupled system; averaging principle; reaction diffusion equation; slow–fast system; Markov chain; neuron model; Hodgkin–Huxley model

2010 Mathematics Subject Classification: Primary 60B12; 60J75; 35K57  
Secondary 92C20; 92C45

## 1. Introduction

The Hodgkin–Huxley model has been one of the most studied models in neuroscience since its creation in the early 1950s by Hodgkin and Huxley [20]. This deterministic model was first created in order to describe the propagation of an action potential or nerve impulse along the axon of a neuron of the giant squid. In this paper we will consider a mathematical generalization of the Hodgkin–Huxley model investigated numerically in [12] and analytically in [1]. This model, called the spatial stochastic Hodgkin–Huxley model in the sequel, comprises a partial differential equation (PDE) describing the evolution of the variation of the potential across the membrane, coupled with a continuous-time Markov chain which describes the dynamics of ion channels that are present all along the axon. It is a piecewise-deterministic Markov process (PDMP) in infinite dimensions (see [2] and [23] for a PDMP in infinite dimensions, and also [4]–[6], [15], and the references therein for a PDMP in finite dimensions). Moreover, it is fully coupled in the sense that the evolution of the potential depends on the kinetics of ion channels and vice versa. The role of ion channels is fundamental because they allow and amplify the propagation of an action potential. We introduce different time scales in this model

---

Received 14 June 2011; revision received 5 March 2012.

\* Postal address: Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, Paris 6, UMR 7599, Case courrier 188, 4 Place Jussieu, 75252 Paris Cedex 05, France.

This work was supported by the Agence Nationale de la Recherche through the project MANDy, Mathematical Analysis of Neuronal Dynamics, ANR-09-BLAN-0008-01.

\*\* Email address: [algenadot@gmail.com](mailto:algenadot@gmail.com)

\*\*\* Email address: [michele.thieullen@upmc.fr](mailto:michele.thieullen@upmc.fr)

by considering that some ion channels open and close at faster rates than others. We perform a slow–fast analysis of this model and prove that, asymptotically, it reduces to the so-called averaged model, which is still a PDMP in infinite dimensions, for which we provide effective evolution equations and jump rates. We thus reduce the complexity of the original model by simplifying the kinetics of ion channels. To the best of the authors' knowledge, no averaging results are available for a PDMP in infinite dimensions.

Even though the Hodgkin–Huxley model was first introduced in neurophysiology, its paradigm is commonly used to describe the behavior of cardiac cells; see, for instance, [14] and [17]. It was at first deterministic. In order to include more variability, and to better describe the axon itself and a certain class of phenomena, stochastic versions of this model have been proposed since the 1990s (see, for instance, [3], [9], [12], and [16]), introducing in the model two kinds of randomness. One is to include in the Hodgkin–Huxley equations an external noise so as to model the multitude of impulses received by one single neuron through its connections to the extraordinarily complex system which is the brain. The other is to introduce in the Hodgkin–Huxley equations the intrinsic noise of the neuron due to the presence of ion channels which are intrinsically stochastic entities. The spatial stochastic Hodgkin–Huxley model which deals with the intrinsic noise of neurons pertains to the latter class.

The generalized Hodgkin–Huxley equations of [1] also belong to a larger class of stochastic systems describing the evolution of a macroscopic variable, here the membrane potential, coupled with a Markovian dynamic modeling the kinetics of smaller entities, here ion channels. Between two successive jumps of the Markovian kinetics, the macroscopic variable follows a deterministic evolution. These models are known mathematically as PDMPs and are often called stochastic hybrid systems in the applications. The theory of PDMPs in finite dimensions (dealing with ordinary differential equations coupled with Markovian kinetics) has been developed in [7] and [8]. The extension to infinite dimensions (dealing with PDEs coupled with Markovian kinetics) has been performed in [2]. The spatial stochastic Hodgkin–Huxley model belongs to the class of PDMPs in infinite dimensions since the equation describing the evolution of the membrane potential is a PDE.

Reducing the complexity of a model is of first importance since it allows us to perform more tractable mathematics and numerical simulations, and, in this way, to expand our understanding of the model. The reduction can be achieved by two different means. The first is to consider that the number or the population of ion channels is large enough to apply a law-of-large-numbers argument so that we can replace the empirical proportion of ion channels in a given state by the averaged theoretical proportion of ion channels in this state. This is achieved by letting the number of ion channels go to  $\infty$ ; the corresponding law of large numbers was proved in [1] and in the context of PDMPs in [26]. The second way to reduce the complexity of the model, the one we will utilize in this paper, is to consider that the model has two intrinsic time scales. It is a quite natural approach in this context; see, for instance, [13] and [19, p. 48].

In the present paper the number of ion channels is fixed and we consider that some ion channels have a faster dynamic than others. We then perform a slow–fast mathematical analysis of this 'two-time-scale' system. We obtain a reduced system in which the intrinsic variability of ion channels is still taken into account, but with a simplified kinetic. Unlike the main result of [1], after reduction, we remain here at a stochastic level. The methods used in this paper are known mathematically under the name of stochastic averaging. As far as the authors are aware, these methods have never been applied to a PDMP in infinite dimensions, which we deal with in the present paper. For PDMPs in finite dimensions, a theory of averaging was developed in [11] and a central limit theorem was proved in [23].

In the infinite-dimensional framework which we consider in this paper we have been able to reduce the complexity of a fully coupled PDMP in infinite dimensions by using averaging methods. The natural step after this averaging result is to prove the central limit theorem associated with it. We aim to investigate this in a future work.

We conclude this introduction with the plan of the paper. In Section 2 we introduce the spatial stochastic Hodgkin–Huxley model and the formalism of a PDMP in infinite dimensions. Then we introduce the two time scales in the model and prove a crucial result for the sequel. We complete the section by presenting the main assumptions of the model and our main result. Section 3 is devoted to the proof of our main result. In Section 4 we apply our result to a Hodgkin–Huxley-type model as an example. In Appendices A–C we respectively provide an important lemma proved in [1], two results about tightness, and the parameter values used in the simulation presented in Section 4.

**1.1. Basic notation**

We set  $I = [0, 1]$ . Define

$$\|f\|_H = \sqrt{\int_I (f(x))^2 + (f'(x))^2 dx} \quad \text{and} \quad \|f\|_{L^2(I)} = \sqrt{\int_I (f(x))^2 dx}.$$

In this paper we will work with the following triplet of Banach spaces:  $H_0^1(I) \subset L^2(I) \subset H^{-1}(I)$ . The space  $H_0^1(I)$  denotes the completion of the set of  $C^\infty$  functions with compact support on  $I$  with respect to the norm  $\|\cdot\|_H$ , and will be abbreviated to  $H$ . The space  $L^2(I)$  endowed with the norm  $\|\cdot\|_{L^2(I)}$  is the usual space of measurable and square-integrable functions with respect to the Lebesgue measure on  $I$ . The space  $H^{-1}(I)$  is the dual space of  $H$  and will be abbreviated to  $H^*$ .

We recall here some basic results about these three spaces. The spaces  $L^2(I)$  and  $H$  are two separable Hilbert spaces endowed with their usual scalar products, respectively denoted by  $(\cdot, \cdot)_{L^2(I)}$  and  $(\cdot, \cdot)$ . We denote by  $(\cdot, \cdot)$  the duality pairing between  $H$  and  $H^*$ .

For an integer  $k \geq 1$ , we define the following functions on  $I$ :

$$e_k(\cdot) = \frac{\sqrt{2}}{\sqrt{1 + (k\pi)^2}} \sin(k\pi\cdot), \quad f_k(\cdot) = \sqrt{2} \sin(k\pi\cdot).$$

The families  $\{f_k, k \geq 1\}$  and  $\{e_k, k \geq 1\}$  are Hilbert bases of  $L^2(I)$  and  $H$ , respectively. The Laplacian with zero Dirichlet boundary conditions has spectral decomposition

$$\Delta u = - \sum_{k \geq 1} (k\pi)^2 (u, f_k)_{L^2(I)} f_k$$

in  $L^2(I)$  for  $u$  in the domain  $\mathcal{D}_{L^2(I)}(\Delta) = \{u \in L^2(I); \sum_{k \geq 1} k^4 (u, f_k)_{L^2(I)}^2 < \infty\}$  and

$$\Delta u = - \sum_{k \geq 1} (k\pi)^2 (u, e_k) e_k$$

in  $H$  for  $u$  in the domain  $\mathcal{D}(\Delta) = \{u \in H; \sum_{k \geq 1} k^4 (u, e_k)^2 < \infty\}$ . We refer the reader to [18, Chapter 1, Section 1.3] for more details.

We denote by  $C_P$  the (Poincaré) constant such that, for all  $u \in H$ , we have

$$\sup_I |u| \leq C_P \|u\|_H.$$

The embeddings  $H \subset L^2(I) \subset H^*$  are continuous and dense. Moreover, for any  $h \in L^2(I)$  and any  $u \in H$ ,  $\langle h, u \rangle := (h, u)_{L^2(I)}$ .

We say that a function  $f : H \mapsto \mathbb{R}$  has a Fréchet derivative in  $u \in H$  if there exists a bounded linear operator  $T_u : H \mapsto \mathbb{R}$  such that

$$\lim_{h \rightarrow 0} \frac{f(u + h) - f(u) - T_u(h)}{\|h\|_H} = 0.$$

We then write  $df[u]/du$  for the operator  $T_u$ . For example, the square of the  $\|\cdot\|_H$  norm and the Dirac distribution in  $x \in I$  are Fréchet differentiable on  $H$ . For all  $u \in H$ ,

$$\frac{d\|\cdot\|_H^2}{du}[u](h) = 2(u, h), \quad \frac{d\delta_x}{du}[u](h) = h(x),$$

for all  $h \in H$ . Fréchet differentiation is stable under summation and multiplication.

## 2. Statement of the model and results

### 2.1. The spatial stochastic Hodgkin–Huxley model

We introduce here the stochastic Hodgkin–Huxley equations considered in [1]. Basically, this spatial stochastic Hodgkin–Huxley model describes the propagation of an action potential along an axon. The axon is the part of a neuron which transmits the information received from the soma to another neuron over long distances: the length of the axon is large relative to the size of the soma. Along the axon are the ion channels which amplify and allow the propagation of the received impulse. For mathematical convenience, we assume that the axon is a segment of length 1 and we denote it by  $I = [0, 1]$ . We assume that there are  $N \geq 1$  ion channels and that they are regularly placed along the axon at location  $i/N$  for  $i \in \mathbb{N} \cap NI$ , where  $I$  denotes the interior of  $I$ . This distribution of the channels is certainly unrealistic, but we proceed with the assumption in order to fix the ideas and be in accordance with [1]. However, the mathematics is the same if we consider any finite subset of  $I$  instead of  $(1/N)(\mathbb{N} \cap NI)$ . Each ion channel can be in a state  $\xi \in E$ , where  $E$  is a finite state space. For example, for the Hodgkin–Huxley model, a state can be ‘receptive to sodium ions and open’ (see [19, p. 487]).

The ion channels switch between states according to a continuous-time Markov chain whose jump intensities depend on the local potential of the membrane—this is why the model is said to be fully coupled. For any states  $\xi, \zeta \in E$ , we define by  $\alpha_{\xi, \zeta}$  the jump intensity (or rate function) from the state  $\xi$  to the state  $\zeta$ . It is a real-valued function of a real variable and assumed to be, as is its derivative, Lipschitz continuous. We assume moreover that  $0 \leq \alpha_{\xi, \zeta} \leq \alpha^+$  for any  $\xi, \zeta \in E$  and that either  $\alpha_{\xi, \zeta}$  is constant, equal to 0, or is strictly positive and bounded below by a strictly positive constant  $\alpha_-$ . That is, the nonzero rate functions are bounded below and above by strictly positive constants. For a given channel, the rate function describes the rate at which it switches from one state to another.

A possible configuration of all the  $N$  ion channels is denoted by an element  $r = (r(i), i \in \mathbb{N} \cap NI)$  of  $\mathcal{R} = E^{\mathbb{N} \cap NI}$ :  $r(i)$  is the state of the channel at location  $i/N$  for  $i \in \mathbb{N} \cap NI$ . The channels, or stochastic processes  $r(i)$ , are assumed to evolve independently over infinitesimal time scales. Denoting by  $u_t(i/N)$  the local potential at location  $i/N$  at time  $t$ , we have

$$\Pr(r_{t+h}(i) = \zeta \mid r_t(i) = \xi) = \alpha_{\xi, \zeta} \left( u_t \left( \frac{i}{N} \right) \right) h + o(h). \tag{1}$$

For any  $\xi \in E$ , we also define the maximal conductance  $c_\xi$  and the steady state potentials  $v_\xi$  of a channel in state  $\xi$ , both of which are constants, with  $C_\xi$  positive.

The membrane potential  $u_t(x)$  along the axon evolves according to the following PDE:

$$\partial_t u_t = \Delta u_t + \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\dot{i}} c_\xi 1_\xi(r_t(i)) \left( v_\xi - u_t \left( \frac{i}{N} \right) \right) \delta_{i/N}. \tag{2}$$

We will assume the zero Dirichlet boundary conditions for this PDE (clamped axon). We are interested in the process  $(u_t, r_t)_{t \in [0, T]}$ .

We recall here the result of [1] which states that there exists a stochastic process satisfying (1) and (2). Let  $u_0$  be in  $\mathcal{D}(\Delta)$  such that  $\min_{\xi \in E} v_\xi \leq u_0 \leq \max_{\xi \in E} v_\xi$ , the initial potential of the axon. Let  $q_0 \in \mathcal{R}$  be the initial configuration of the ion channels.

**Proposition 1.** ([1].) *Fix  $N \geq 1$ , and let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \Pr)$  be a filtered probability space satisfying the usual conditions. There exists a pair  $(u_t, r_t)_{0 \leq t \leq T}$  of càdlàg adapted stochastic processes such that each sample path of  $u$  is a continuous map from  $[0, T]$  to  $H$  and  $r_t$  is in  $\mathcal{R}$  for all  $t \in [0, T]$ , which satisfies the following properties.*

- (Regularity) *The map  $t \mapsto \partial_t u_t$  lies in  $L^2([0, T], H^*)$  almost surely.*
- (Dynamic: PDE)

$$\partial_t u_t = \Delta u_t + \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\dot{i}} c_\xi 1_\xi(r_t(i)) \left( v_\xi - u_t \left( \frac{i}{N} \right) \right) \delta_{i/N}$$

*for all  $t \in [0, T]$ , Pr-almost surely.*

- (Dynamic: jump)

$$\Pr(r_{t+h}(i) = \zeta \mid r_t(i) = \xi) = \alpha_{\xi, \zeta} \left( u_t \left( \frac{i}{N} \right) \right) h + o(h).$$

- (Initial condition: PDE)  $u_0$  given.
- (Initial condition: jump)  $q_0$  given.
- (Boundary conditions: PDE only)  $u_t(0) = u_t(1) = 0$  for all  $t \in [0, T]$ .

Moreover, there exists a constant  $C$  such that  $\|u_t\|_\infty \leq C$  for all  $t \in [0, T]$  and  $\omega$ .

In [1] the author proved that, when the number of ion channels increases to  $\infty$ , the above model converges, in a sense, toward a deterministic model. In our paper, unlike in [1], we work with a fixed number of ion channels but introduce two time scales into the model in order to perform a slow–fast analysis.

### 2.2. The spatial stochastic Hodgkin–Huxley model as a PDMP in infinite dimensions

Buckwar and Riedler [2] extended the theory of finite-dimensional PDMPs introduced in [7] and [8] to PDMPs in infinite dimensions. The results stated here come from Sections 2.3 and 3.1 of [2], adapted to our particular notation.

For all  $r \in \mathcal{R}$ , we define on  $H$  the function

$$G_r(u) = \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\dot{i}} c_\xi 1_\xi(r(i)) \left( v_\xi - u \left( \frac{i}{N} \right) \right) \delta_{i/N},$$

which is  $H^*$ -valued. We do not include the dependence on  $N$  in the notation of  $G_r(u)$  since, unlike in [1],  $N$  is a fixed parameter here.

The stochastic process  $(u_t, r_t)_{t \in [0, T]}$  can be described in the following way. We start with a given initial potential  $v_0$  and ion channels configuration  $q_0$ . Then the  $u$  component evolves according to the evolution equation

$$\partial_t u_t = \Delta u_t + G_{q_0}(u_t)$$

until the first jump of the  $r$  component at time  $\tau_1$  which occurs according to the transition rate function  $\Lambda : H \times \mathcal{R} \rightarrow \mathbb{R}_+$  given for  $(u, r) \in H \times \mathcal{R}$  by

$$\Lambda(u, r) = \sum_{i \in \mathbb{N} \cap N} \sum_{\xi \neq r(i)} \alpha_{r(i), \xi} \left( u \left( \frac{i}{N} \right) \right)$$

according to (1). A new value  $q_1$  for  $r$  is then chosen according to a transition measure  $Q$  from  $H \times \mathcal{R}$  to  $\mathcal{P}(\mathcal{R})$  (the set of all probability measures on  $\mathcal{R}$ ). This transition measure is also given by the jump distribution of the ion channels. For  $(u, r) \in H \times \mathcal{R}$  and  $r' \in \mathcal{R}$ , which differs from  $r$  only by the component  $r(i_0)$ ,

$$Q(u, r)(\{r'\}) = \frac{\alpha_{r(i_0), r'(i_0)}(u(i_0/N))}{\Lambda(u, r)}.$$

If  $r'$  differs from  $r$  by two or more components then  $Q(u, r)(\{r'\}) = 0$ .

Then  $u$  evolves according to the ‘updated’ evolution equation

$$\partial_t u_t = \Delta u_t + G_{q_1}(u_t)$$

with initial condition  $u_{\tau_1}$  and Dirichlet boundary conditions, until the second jump of the  $r$  component. And so on. This description justifies the term ‘piecewise-deterministic process’: the dynamic of the process is indeed purely deterministic between the jumps.

As shown in Theorem 6 of [2], the equations of Proposition 1 define a standard piecewise-deterministic process (PDP) in the sense of Definition 3 of [2]. We recall briefly here what this means technically.

- For a given ion channel configuration  $r \in \mathcal{R}$ , the PDE

$$\partial_t u_t = \Delta u_t + G_r(u_t)$$

with zero Dirichlet boundary conditions admits a unique global (weak) solution continuous on  $H$  for every initial value  $u_0 \in H$ , denoted by  $\psi_r(t, x)$ .

- The number of state switches in the  $(r_t)_{t \in [0, T]}$  component during any finite time interval is finite almost surely.
- The transition rate  $\Lambda$  from one state of  $\mathcal{R}$  to another is a measurable function from  $H \times \mathcal{R}$  to  $\mathbb{R}_+$  and, for all  $(u, r) \in H \times \mathcal{R}$ , the transition rate as a function of time is integrable over every finite time interval but diverges to  $\infty$  when it is integrated over  $\mathbb{R}_+$  (the expected number of jumps tends to  $\infty$  when the time horizon increases).
- The transition measure  $Q$  from one state of  $\mathcal{R}$  to another is such that the mapping  $(u, r) \rightarrow Q(u, r)(R)$  is measurable for every  $R \subset \mathcal{R}$  and  $Q(u, r)(\{r\}) = 0$  for all  $(u, r) \in H \times \mathcal{R}$ .

In fact, the stochastic process  $(u_t, r_t)_{t \in [0, T]}$  is more than a PDP, it is a piecewise-deterministic Markov process.

**Theorem 1.** 1. *There exists a filtered probability space satisfying the usual conditions such that our infinite-dimensional standard PDP is a homogeneous Markov process on  $H \times \mathcal{R}$ .*

2. *Locally bounded measurable functions on  $H \times \mathcal{R}$  which are absolutely continuous as maps  $t \mapsto f(\psi_r(t, x), r)$  for all  $(x, r)$  are in the domain  $\mathcal{D}(\mathcal{A})$  of the extended generator. The extended generator is given for almost all  $t$  by*

$$\mathcal{A}f(u_t, r_t) = \frac{d}{dt} f(u_t, r_t) + \mathcal{B}f(u_t, r_t),$$

where

$$\mathcal{B}f(u_t, r_t) = \sum_{i \in \mathbb{N} \cap N} \sum_{\zeta \in E} [f(u_t, r_t(r_t(i) \rightarrow \zeta)) - f(u_t, r_t)] \alpha_{r_t(i), \zeta} \left( u_t \left( \frac{i}{N} \right) \right)$$

and  $r_t(r_t(i) \rightarrow \zeta)$  is the component of  $\mathcal{R}$  with  $r_t(r_t(i) \rightarrow \zeta)(j)$  equal to  $r_t(j)$  if  $j \neq i$  and equal to  $\zeta$  if  $j = i$ . The notation  $\frac{d}{dt} f(u_t, r_t)$  means that the function  $s \mapsto f(u'_s, r)$  is differentiated at  $s = t$ , where  $u'$  is the solution of the PDE of Proposition 1 such that  $u'_t = u_t$ , with the channel state  $r$  held fixed at  $r_t$ .

**Remark 1.** It is not usual to define a generator along only the paths  $(u_t, r_t)_{t \geq 0}$  as we do. For a generator, we would expect an analytical expression of the form  $\mathcal{A}f(u, r)$  for any  $(u, r) \in H \times \mathcal{R}$ . It is in fact possible to obtain such an expression for a more regular function  $f$ . The part of the generator related to the continuous Markov chain  $r$  takes the form  $\mathcal{B}f(u, r)$  with the only change that the derivative  $\frac{d}{dt} f(u, r)$  can then be expressed thanks to the Fréchet derivative of  $f$ ; see [2] for more details. We will not use this refinement in the sequel.

Knowing that a stochastic process is Markovian is crucial for its study since a lot of mathematical tools have been developed for Markovian processes; see, for example, [10].

### 2.3. Singularly perturbed model and main results

2.3.1. *Introduction of two time scales into the model.* We now introduce two time scales into the model. Indeed, we consider that the Hodgkin–Huxley model is a slow–fast system: some states of the ion channels communicate faster than others; see, for example, [19, p. 48]. Mathematically, this leads to the introduction of a small parameter  $\varepsilon > 0$  in our equation. For the states which communicate at a faster rate, we say that they communicate at the usual rate divided by  $\varepsilon$ .

We can then consider different classes of states or partition the state space  $E$  into states which communicate at a high rate. This kind of description is very classical; see, for example, [11]. We regroup our states into classes which form a partition of the state space  $E$ , i.e.

$$E = E_1 \sqcup \dots \sqcup E_l,$$

where  $l \in \{1, 2, \dots\}$  is the number of classes. Inside a class  $E_j$ , the states communicate faster at jump rates of order  $1/\varepsilon$ . States of different classes communicate at the usual rate of order 1. For fixed  $\varepsilon > 0$ , we denote by  $(u^\varepsilon, r^\varepsilon)$  the two-time-scale modification of the PDMP introduced in the previous section. Its generator is given by, for  $f \in \mathcal{D}(\mathcal{A})$ ,

$$\mathcal{A}f(u_t^\varepsilon, r_t^\varepsilon) = \frac{d}{dt} f(u_t^\varepsilon, r_t^\varepsilon) + \mathcal{B}^\varepsilon f(u_t^\varepsilon, r_t^\varepsilon),$$

where  $\mathcal{B}^\varepsilon$  is the component of the generator related to the continuous-time Markov chain  $r^\varepsilon$ .

According to our slow–fast description we have a two-time-scale decomposition of this generator, i.e.

$$\mathcal{B}^\varepsilon = \frac{1}{\varepsilon} \mathcal{B} + \hat{\mathcal{B}},$$

where the ‘fast’ generator  $\mathcal{B}$  is given by

$$\begin{aligned} &\mathcal{B} f(u_t^\varepsilon, r_t^\varepsilon) \\ &= \sum_{i \in \mathbb{N} \cap N\varepsilon} \sum_{j=1}^l 1_{E_j}(r_t^\varepsilon(i)) \sum_{\zeta \in E_j} [f(u_t^\varepsilon, r_t^\varepsilon(r_t^\varepsilon(i) \rightarrow \zeta)) - f(u_t^\varepsilon, r_t^\varepsilon)] \alpha_{r_t^\varepsilon(i), \zeta} \left( u_t^\varepsilon \left( \frac{i}{N} \right) \right) \end{aligned}$$

and the ‘slow’ generator  $\mathcal{B}$  is given by

$$\begin{aligned} &\hat{\mathcal{B}} f(u_t^\varepsilon, r_t^\varepsilon) \\ &= \sum_{i \in \mathbb{N} \cap N\varepsilon} \sum_{j=1}^l 1_{E_j}(r_t^\varepsilon(i)) \sum_{\zeta \notin E_j} [f(u_t^\varepsilon, r_t^\varepsilon(r_t^\varepsilon(i) \rightarrow \zeta)) - f(u_t^\varepsilon, r_t^\varepsilon)] \alpha_{r_t^\varepsilon(i), \zeta} \left( u_t^\varepsilon \left( \frac{i}{N} \right) \right). \end{aligned}$$

For fixed  $y \in \mathbb{R}$  and  $g: \mathbb{R} \times E \rightarrow \mathbb{R}$ , we denote by  $\mathcal{B}_j(y)$ ,  $j \in \{1, \dots, l\}$ , the generator

$$\mathcal{B}_j(y)g(\xi) = 1_{E_j}(\xi) \sum_{\zeta \in E_j} [g(y, \zeta) - g(y, \xi)] \alpha_{\xi, \zeta}(y),$$

which will be of interest in the sequel.

2.3.2. *Uniform boundedness.* Proposition 2 below is a crucial result for the proof of our main result. Its proof follows the argument developed in [1], but, for the sake of completeness and given the intensive use of this proposition in the sequel, we provide a short proof.

**Proposition 2.** *For any  $T > 0$ , there exists a deterministic constant  $C > 0$  independent of  $\varepsilon \in (0, 1]$  such that*

$$\sup_{t \in [0, T]} \|u_t^\varepsilon\|_H \leq C \quad \text{almost surely.}$$

In the proof of this proposition we will use the following representation of a solution of the PDE part of our PDMP. We say that  $u^\varepsilon$  is a mild solution to (2) if

$$u_t^\varepsilon = e^{\Delta t} u_0 + \int_0^t e^{\Delta(t-s)} G_{r_s^\varepsilon}(u_s^\varepsilon) ds, \quad t \geq 0, \tag{3}$$

almost surely, where  $(e^{\Delta t}, t \geq 0)$  is the semigroup associated to  $\Delta$  in  $H$ . For any  $u \in H$ ,

$$e^{\Delta t} u = \sum_{k \geq 1} e^{-(k\pi)^2 t} (u, e_k) e_k, \quad t > 0.$$

Let us mention that the mild formulation (3) also holds for a bounded stopping time  $\tau$ ; this will be useful later in the paper.

*Proof of Proposition 2.* We work  $\omega$  by  $\omega$ . Using the mild formulation (3), we see that

$$\begin{aligned} \|u_t^\varepsilon\|_H &\leq \|e^{\Delta t} u_0\|_H + \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\varepsilon} \left\| \int_0^t c_\xi 1_\xi(r_s^\varepsilon(i)) v_\xi e^{\Delta(t-s)} \delta_{i/N} ds \right\|_H \\ &\quad + \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\varepsilon} \left\| \int_0^t c_\xi u_s^\varepsilon \left( \frac{i}{N} \right) e^{\Delta(t-s)} \delta_{i/N} ds \right\|_H. \end{aligned}$$



By Lemma 2 in Appendix A, the first of the above sums is bounded by a fixed constant which can be chosen independently of  $t \in [0, T]$ . For the second sum, we break the integral into two. For all  $\gamma > 0$ , by Lemma 2 we can choose  $\eta > 0$  so small that

$$\left\| \int_{t-\eta}^t c_\xi u_s^\varepsilon \left( \frac{i}{N} \right) e^{\Delta(t-s)} \delta_{i/N} ds \right\|_H \leq \frac{1}{2} \gamma \max_{s \in [0, t]} \left| c_\xi u_s^\varepsilon \left( \frac{i}{N} \right) \right|.$$

Since  $|c_\xi u_s^\varepsilon(i/N)| \leq \max c_\xi \|u_s^\varepsilon\|_\infty \leq \max c_\xi C_P \|u_s^\varepsilon\|_H$ , choosing  $\gamma$  such that  $\gamma \max c_\xi C_P < 1$  yields

$$\left\| \int_{t-\eta}^t c_\xi u_s^\varepsilon \left( \frac{i}{N} \right) e^{\Delta(t-s)} \delta_{i/N} ds \right\|_H \leq \frac{1}{2} \max_{s \in [0, t]} \|u_s^\varepsilon\|_H.$$

Now the second estimates of Lemma 2 give

$$\begin{aligned} \left\| \int_0^{t-\eta} c_\xi u_s^\varepsilon \left( \frac{i}{N} \right) e^{\Delta(t-s)} \delta_{i/N} ds \right\|_H &\leq C_2(\eta) \max_{\xi \in E} |c_\xi| \int_0^{t-\eta} \|u_s^\varepsilon\|_\infty ds \\ &\leq C_P C_2(\eta) \max_{\xi \in E} |c_\xi| \int_0^{t-\eta} \max_{\sigma \in [0, s]} \|u_\sigma^\varepsilon\|_H ds. \end{aligned}$$

Reassembling the above inequalities, the proof is completed by a classical application of Gronwall’s lemma. We finally find that  $\max_{s \in [0, t]} \|u_t^\varepsilon\|_H$  is bounded by a fixed constant independent of  $t \in [0, T]$ ,  $\omega$ , and  $\varepsilon \in (0, 1]$ .

**2.3.3. Main assumptions and averaged equation.** We make the following assumptions. For any fixed  $y \in \mathbb{R}$  and any  $j \in \{1, \dots, l\}$ , the fast generator  $\mathcal{B}_j(y)$  is weakly irreducible on  $E_j$ , i.e. it has a unique quasistationary distribution which we denote by  $\mu_j(y)$ . This quasistationary distribution is assumed to be, as is its derivative, Lipschitz continuous in its argument  $y$  (see Section 2.3.4 for more details).

Following [27, Chapter 7, p. 167], the states in  $E_j$  can be considered equivalent. For any  $i \in \mathbb{N} \cap N\hat{I}$ , we define the new stochastic process  $(\bar{r}_t^\varepsilon)_{t \geq 0}$  by  $\bar{r}_t^\varepsilon(i) = j$  when  $r_t^\varepsilon(i) \in E_j$  and abbreviate  $E_j$  to  $j$ . We then have an aggregate process  $\bar{r}^\varepsilon(i)$  with values in  $\{1, \dots, l\}$ . It is not a Markov process, but we have the following proposition.

**Proposition 3.** ([27].) *For any  $y \in \mathbb{R}$ ,  $i \in \mathbb{N} \cap N\hat{I}$ , and  $g : \{1, \dots, l\} \rightarrow \mathbb{R}$ ,  $\bar{r}^\varepsilon(i)$  converges weakly when  $\varepsilon$  goes to 0 to a Markov process  $\bar{r}(i)$  generated by*

$$\bar{\mathcal{B}}(y)g(\bar{r}(i)) = \sum_{j=1}^l 1_j(\bar{r}(i)) \sum_{k=1, k \neq j}^l (g(k) - g(j)) \sum_{\xi \in E_k} \sum_{\zeta \in E_j} \alpha_{\zeta, \xi}(y) \mu_j(\zeta).$$

We then average the function  $G_r(u)$  against the quasi-invariant distributions. That is, we consider that each pack of states  $E_j$  has reached its stationary behavior. For any  $\bar{r} \in \bar{\mathcal{R}} = \{1, \dots, l\}^{\mathbb{N} \cap N\hat{I}}$ , we define the averaged function by

$$F_{\bar{r}}(u) = \frac{1}{N} \sum_{i \in \mathbb{N} \cap N\hat{I}} \sum_{j=1}^l 1_j(\bar{r}(i)) \sum_{\zeta \in E_j} c_\zeta \mu_j \left( u \left( \frac{i}{N} \right) \right) (\zeta) \left( v_\zeta - u \left( \frac{i}{N} \right) \right) \delta_{i/N}.$$

We therefore call the PDE

$$\partial_t u_t = \Delta u_t + F_{\bar{r}_t}(u_t) \tag{4}$$

with zero Dirichlet boundary condition and initial conditions  $u_0$  and  $\bar{q}_0$  (where  $\bar{q}_0$  is the aggregation of the initial channel configuration  $q_0$ ) the averaged equation of (2). In (4), the process  $(\bar{r}_t)_{t \in [0, T]}$  evolves, each coordinate independently over infinitesimal time scales, according to the averaged jump rates between the subsets  $E_j$  of  $E$ .

**Proposition 4.** *For any  $T > 0$ , there exists a probability space satisfying the usual conditions such that (4) defines a PDMP  $(u_t, \bar{r}_t)_{t \in [0, T]}$  in infinite dimensions in the sense of [2]. Moreover, there exists a constant  $C$  such that*

$$\sup_{t \in [0, T]} \|u_t\|_H \leq C$$

and  $u \in \mathcal{C}([0, T], H)$  almost surely.

*Proof.* Equation (4) is of the form (2) except that the indicator function is replaced by the probabilities  $\mu_j$ . This does not change the mathematics, and the arguments developed in [1] or [2] still apply. We refer the reader to [1] and [2] for more details.

We now give our main result, which states that approximation (4) is valid.

**Theorem 2.** *The stochastic process  $u^\varepsilon$  satisfying (2) converges in law to a solution of (4) when  $\varepsilon$  goes to 0.*

More precisely, we will prove the following proposition.

**Proposition 5.** *The family  $(u^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $\mathcal{C}([0, T], H)$  and any accumulation point  $u$  verifies*

$$(u_t, \phi)_{L^2(I)} = (u_0, \phi)_{L^2(I)} + \int_0^t (u_s, \phi'')_{L^2(I)} ds + \int_0^t \langle F_{\bar{r}_s}(u_s), \phi \rangle ds \tag{5}$$

for all  $\phi$  in  $\mathcal{C}_0^2(I)$ . Moreover, the accumulation point is unique up to indistinguishability.

2.3.4. *Plan.* To make the proof of Proposition 5 easier to read, we will consider in a first step that all the states communicate at a fast rate. This is called the all-fast case. We will prove Proposition 5 in detail in this case and then give the key points for the validity of the proof in the general case. In the all-fast case there is a unique class  $E$ , and the generator has the simplest form

$$\mathcal{B}f(u, r) = \sum_{i \in \mathbb{N} \cap N\dot{I}} \sum_{\zeta \in E} [f(u, r(r(i) \rightarrow \zeta)) - f(u, r)] \alpha_{r(i), \zeta} \left( u \left( \frac{i}{N} \right) \right). \tag{6}$$

We make the following assumptions: for any  $i \in \mathbb{N} \cap N\dot{I}$  with fixed  $u \in H$ , the Markov process  $r(i)$  has a unique stationary distribution  $\mu(u(i/N))$ . Then the process  $(r(i), i \in \mathbb{N} \cap N\dot{I})$  has the following stationary distribution:

$$\mu(u) = \bigotimes_{i \in \mathbb{N} \cap N\dot{I}} \mu \left( u \left( \frac{i}{N} \right) \right).$$

We assume that the rate functions are Lipschitz continuous. The average (generalized) function is then

$$F(u) = \int_{\mathcal{R}} G_r(u) \mu(u)(dr) = \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\dot{I}} c_\xi \mu \left( u \left( \frac{i}{N} \right) \right) (\xi) \left( v_\xi - u \left( \frac{i}{N} \right) \right) \delta_{i/N} \tag{7}$$

if  $u$  is held fixed.

Is this hypothesis reasonable? For the most classical model in neuroscience, the rate functions are indeed, as well as their derivatives, Lipschitz continuous. See, for example, the rate functions in classical Hodgkin–Huxley model [20] (see also Appendix C).

Let us briefly show that in this case the ergodic measure  $\mu$  is Lipschitz continuous in its argument. Let  $i$  be in  $\mathbb{N} \cap NI$ , and let the membrane potential at  $i/N$  be fixed and equal to  $x \in \mathbb{R}$ ; the state space of the continuous-time Markov chain  $r(i)$  is  $E$  which is a finite set. We set  $|E| = m$ . We can then enumerate the elements of  $E$ :  $E = \{\xi_1, \dots, \xi_m\}$ . The generator of  $r(i)$  is then the matrix

$$J(x) = (\alpha_{\xi_i, \xi_j}(x))_{1 \leq i, j \leq m}$$

with  $\alpha_{\xi_i, \xi_i}(x) = -\sum_{j \neq i} \alpha_{\xi_i, \xi_j}(x)$ . Therefore, if we assumed that our continuous-time Markov chain is ergodic, it has an invariant measure of the form

$$\mu(x)(\xi) = \frac{\sum \prod \alpha_{\xi, \xi'}(x)}{\sum \prod \alpha_{\zeta, \zeta'}(x)}.$$

The sums and products involved here are calculated over subsets of  $E$  which are given by resolving the equation  $\mu(x)^\top J = 0$ . Then  $\mu$  is Lipschitz continuous in  $x$  since each  $\alpha_{\xi, \zeta}$  is Lipschitz continuous in  $x$  and bounded.

The assumption of independence of each coordinate of the process  $r$  over infinitesimal time scales is important because it leads to a simple expression for the invariant measure of the process  $r$  and simplifies the mathematics. It is possibly unrealistic from the biological point of view but is often assumed in mathematical neuroscience.

Our plan to prove the main result (2) in the all-fast case is the following. We want to prove the convergence in law of a family of càdlàg stochastic processes with values in a Hilbert space to another, here deterministic, process. To this end, we use a well-known strategy:

- prove the tightness of the family,
- identify the limit.

In our case, tightness (see Section 3.1) will follow roughly from the uniform boundedness of the family  $(u^\varepsilon, \varepsilon \in (0, 1])$  obtained in Proposition 2. The identification of the limit will follow on one hand from tightness and on the other hand from some classical argument in the theory of averaging for stochastic processes; see, for example, [24]. We will introduce a Poisson equation which will enable us to control some problematic terms (see Proposition 6) and identify the limit (Section 3.1). We will then examine the general case with multiple classes (Section 3.2), focusing on resolving the points in the analysis that can lead to difficulties.

### 3. Proof of the main result

#### 3.1. The all-fast case

We want to prove Proposition 5. Let us recall that in the all-fast case  $F_{\bar{r}}$  in (5) reduces to  $F$  in (7).

Let us outline the strategy of the proof, which follows the plan given in Section 2.3.4.

*Step 1: tightness.*

- (i) Thanks to Proposition 2, use the Markov inequality to prove that Aldous’s condition and property (14) given in Theorems 3 and 4 (in Appendix B) hold.

- (ii) For  $t \in [0, T]$  fixed, find the finite-dimensional space  $L_{\delta, \eta}$  of (15) (see Theorem 4 in Appendix B) using a truncation of the solution  $u_t^\varepsilon$  in the mild formulation. Bound  $E(d(u_t^\varepsilon, L_{\delta, \eta}) > \eta)$  and use the Markov inequality.
- (iii) Obtain tightness in  $\mathbb{D}([0, T], H)$  by Theorems 3 and 4 (in Appendix B) and conclude that tightness in  $\mathcal{C}([0, T], H)$  holds as well thanks to the regularity of the solution  $u^\varepsilon$  for any  $\varepsilon \in (0, 1]$ .

*Step 2: identification of the limit.*

- (i) Begin by showing that the following property is sufficient:

$$\lim_{\varepsilon \rightarrow 0} E \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds \right| = 0 \quad \text{for all } \phi \in \mathcal{C}_0^2(I).$$

To prove this property, use the tightness of the family  $(u^\varepsilon, \varepsilon \in (0, 1])$ .

- (ii) Replace the term  $\langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle$  by  $\mathcal{B}f(u_s^\varepsilon, r_s^\varepsilon)$ , where  $f$  is the solution of a Poisson equation. Show that this solution exists and has some nice properties.
- (iii) Write the semimartingale expansion of  $\int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds$ . Bound its martingale part using the bound for the bracket of the martingale.
- (iv) Bound its finite-variation part using the nice properties of  $f$  and the uniform bound from Proposition 2.
- (v) Gather these bounds together to show that  $E \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds \right|^2$  is bounded by a constant times  $\varepsilon$ .
- (vi) Conclude that (5) holds and implies uniqueness up to indistinguishability.

*Step 1(i) (property (14) and Aldous’s condition from Appendix B).* We begin with the Aldous condition. Let  $\tau$  be a stopping time, and let  $\theta$  be such that  $\tau + \theta < T$ . We need to control the size of  $\|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H$ . At first we work deterministically. We break the difference  $\|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H^2$  into three pieces using the mild formulation of  $u^\varepsilon$  in (3):

$$\begin{aligned} \|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H^2 &\leq 4\|e^{\Delta(\tau+\theta)}u_0 - e^{\Delta\tau}u_0\|_H^2 + 4\left\| \int_\tau^{\tau+\theta} e^{\Delta(\tau+\theta-s)}G_{r_s^\varepsilon}(u_s^\varepsilon) ds \right\|_H^2 \\ &\quad + 4\left\| \int_0^\tau (e^{\Delta(\tau+\theta-s)} - e^{\Delta(\tau-s)})G_{r_s^\varepsilon}(u_s^\varepsilon) ds \right\|_H^2. \end{aligned}$$

Let us show now that the supremum over  $\theta \in (0, \eta)$  of each term on the right-hand side of the above inequality goes to 0 with  $\eta$  (we take  $\eta > 0$ ). We start with the third term:

$$\left\| \int_0^\tau (e^{\Delta(\tau+\theta-s)} - e^{\Delta(\tau-s)})G_{r_s^\varepsilon}(u_s^\varepsilon) ds \right\|_H^2.$$

Note that

$$\begin{aligned} &(e^{\Delta(\tau+\theta-s)} - e^{\Delta(\tau-s)})G_{r_s^\varepsilon}(u_s^\varepsilon) \\ &= \sum_{k \geq 1} e^{-(k\pi)^2(\tau-s)}(e^{-(k\pi)^2\theta} - 1)\frac{1}{N} \\ &\quad \times \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap \dot{I}} c_\xi 1_\xi(r_s^\varepsilon(i)) \left( v_\xi - u_s^\varepsilon \left( \frac{i}{N} \right) \right) (1 + (k\pi)^2) e_k \left( \frac{i}{N} \right) e_k, \end{aligned}$$

where we have used the fact that  $\langle e_k, \delta_{i/N} \rangle = (1 + (k\pi)^2)e_k(i/N)$  (recall that  $e_k = \sqrt{2} \sin(k\pi \cdot) / \sqrt{1 + (k\pi)^2}$ ). Thus, we obtain

$$\begin{aligned} & \left\| \int_0^\tau (e^{\Delta(\tau+\theta-s)} - e^{\Delta(\tau-s)}) G_{r_s^\varepsilon}(u_s^\varepsilon) ds \right\|_H^2 \\ & \leq 4|E|^2 \max_{\xi \in E} c_\xi^2 \left( \max_{\xi \in E} v_\xi^2 + C^2 \right) \sum_{k \geq 1} \left( \int_0^\tau e^{-(k\pi)^2(\tau-s)} (e^{-(k\pi)^2\theta} - 1) \sqrt{1 + (k\pi)^2} ds \right)^2 \\ & = 4|E|^2 \max_{\xi \in E} c_\xi^2 \left( \max_{\xi \in E} v_\xi^2 + C^2 \right) \sum_{k \geq 1} (e^{-(k\pi)^2\theta} - 1)^2 (1 + (k\pi)^2) \frac{(1 - e^{-(k\pi)^2\tau})^2}{(k\pi)^4} \\ & \leq 4|E|^2 \max_{\xi \in E} c_\xi^2 \left( \max_{\xi \in E} v_\xi^2 + C^2 \right) \sum_{k \geq 1} (e^{-(k\pi)^2\theta} - 1)^2 \frac{1 + (k\pi)^2}{(k\pi)^4}, \end{aligned}$$

where  $C$  is the constant of Proposition 2. By dominated convergence, we see that

$$\lim_{\eta \rightarrow 0} \sup_{\theta \in (0, \eta)} \sum_{k \geq 1} (e^{-(k\pi)^2\theta} - 1)^2 \frac{1 + (k\pi)^2}{(k\pi)^4} \leq \lim_{\eta \rightarrow 0} \sum_{k \geq 1} (e^{-(k\pi)^2\eta} - 1)^2 \frac{1 + (k\pi)^2}{(k\pi)^4} = 0.$$

For the second term, as before, we can show that

$$\begin{aligned} & \left\| \int_\tau^{\tau+\theta} e^{\Delta(\tau+\theta-s)} G_{r_s^\varepsilon}(u_s^\varepsilon) ds \right\|_H^2 \\ & \leq 4|E|^2 \max_{\xi \in E} c_\xi^2 \left( \max_{\xi \in E} v_\xi^2 + C^2 \right) \sum_{k \geq 1} \left( \int_\tau^{\tau+\theta} e^{-(k\pi)^2(\tau+\theta-s)} \sqrt{1 + (k\pi)^2} ds \right)^2 \\ & \leq 4|E|^2 \max_{\xi \in E} c_\xi^2 \left( \max_{\xi \in E} v_\xi^2 + C^2 \right) \sum_{k \geq 1} (1 + (k\pi)^2) \frac{(1 - e^{-(k\pi)^2\theta})^2}{(k\pi)^4}. \end{aligned}$$

By dominated convergence, we see that

$$\lim_{\eta \rightarrow 0} \sup_{\theta \in (0, \eta)} \sum_{k \geq 1} (1 + (k\pi)^2) \frac{(1 - e^{-(k\pi)^2\theta})^2}{(k\pi)^4} \leq \lim_{\eta \rightarrow 0} \sum_{k \geq 1} (1 + (k\pi)^2) \frac{(1 - e^{-(k\pi)^2\eta})^2}{(k\pi)^4} = 0.$$

For the first term, we have, by the Bessel–Parseval equality,

$$\begin{aligned} \|e^{\Delta(\tau+\theta)} u_0 - e^{\Delta\tau} u_0\|_H^2 &= \left\| \sum_{k \geq 1} e^{-(k\pi)^2\tau} (e^{-(k\pi)^2\theta} - 1) (u_0, e_k) e_k \right\|_H^2 \\ &= \sum_{k \geq 1} e^{-2(k\pi)^2\tau} (e^{-(k\pi)^2\theta} - 1)^2 (u_0, e_k)^2 \\ &\leq \sum_{k \geq 1} (e^{-(k\pi)^2\theta} - 1)^2 (u_0, e_k)^2. \end{aligned}$$

By dominated convergence, we see that

$$\lim_{\eta \rightarrow 0} \sup_{\theta \in (0, \eta)} \sum_{k \geq 1} (e^{-(k\pi)^2\theta} - 1)^2 (u_0, e_k)^2 \leq \lim_{\eta \rightarrow 0} \sum_{k \geq 1} (e^{-(k\pi)^2\eta} - 1)^2 (u_0, e_k)^2 = 0.$$

Combining the results obtained for the three terms we have

$$\lim_{\eta \rightarrow 0} \sup_{\theta \in (0, \eta)} \|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H^2 = 0$$

uniformly in  $\varepsilon \in (0, 1]$ . Therefore, for all  $M, \delta > 0$ , using the Chebyshev inequality, we can choose  $\eta$  so small that

$$\sup_{\varepsilon \in (0, 1]} \sup_{\theta \in (0, \eta)} \Pr(\|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H \geq M) \leq \sup_{\varepsilon \in (0, 1]} \sup_{\theta \in (0, \eta)} \frac{E(\|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H^2)}{M^2} < \delta.$$

The first condition, property (14) of Theorem 4 given in Appendix B, is verified by the application of the Markov inequality. Indeed, for any  $t \in [0, T]$  and  $\delta > 0$ , by Proposition 2, there exists a constant  $C > 0$  independent of  $\varepsilon \in (0, 1]$  and  $t \in [0, T]$  such that

$$\sup_{\varepsilon \in (0, 1]} \Pr(\|u_t^\varepsilon\|_H > \rho) \leq \frac{1}{\rho} \sup_{\varepsilon \in (0, 1]} E \|u_t^\varepsilon\|_H \leq \frac{C}{\rho} < \delta$$

for any  $\rho > 0$  large enough.

*Step 1(ii) (truncation).* We have to show that, for any  $\delta, \eta > 0$ , we can find a  $\varepsilon_0 > 0$  and a space  $L_{\delta, \eta}$  such that

$$\sup_{\varepsilon \in (0, \varepsilon_0]} \Pr(d(u_t^\varepsilon, L_{\delta, \eta}) > \eta) \leq \delta,$$

where  $d(u_t^\varepsilon, L_{\delta, \eta}) = \inf_{v \in L_{\delta, \eta}} \|u_t^\varepsilon - v\|_H$ . We fix  $t \in [0, T]$ . Recalling the mild representation (3), we have

$$u_t^\varepsilon = e^{\Delta t} u_0 + \int_0^t \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N i} c_\xi 1_\xi(r_s^\varepsilon(i)) \left( v_\xi - u_s^\varepsilon \left( \frac{i}{N} \right) \right) e^{\Delta(t-s)} \delta_{i/N} ds,$$

where, using the explicit expression of the semigroup  $e^{\Delta t}$ ,  $e^{\Delta(t-s)} \delta_{i/N}$  is equal to

$$\sum_{k \geq 1} e^{-(k\pi)^2(t-s)} (1 + (k\pi)^2) e_k \left( \frac{i}{N} \right) e_k.$$

We define, for  $f \in H$  and  $x \in I$ ,

$$e_p^{\Delta t} f := \sum_{k=1}^p e^{-(k\pi)^2 t} (f, e_k) e_k,$$

$$e_p^{\Delta t} \delta_x := \sum_{k=1}^p e^{-(k\pi)^2 t} (1 + (k\pi)^2) e_k(x) e_k,$$

and

$$u_t^{\varepsilon, p} := e_p^{\Delta t} u_0 + \int_0^t \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N i} c_\xi 1_\xi(r_s^\varepsilon(i)) \left( v_\xi - u_s^\varepsilon \left( \frac{i}{N} \right) \right) e_p^{\Delta(t-s)} \delta_{i/N} ds.$$

For any  $\delta, \eta > 0$ , we can find  $p$  independent of  $\varepsilon$  such that  $\|u_t^\varepsilon - u_t^{\varepsilon, p}\|_H \leq C' \eta \delta$  with the constant  $C'$  deterministic and independent of  $\delta, \eta$ , and  $\varepsilon \in (0, 1]$ . Indeed, we can easily show, as in step 1(i), that

$$\|u_t^\varepsilon - u_t^{\varepsilon, p}\|_H^2 \leq 2 \left\| \sum_{k=p+1}^\infty e^{-(k\pi)^2 t} (u_0, e_k) e_k \right\|_H^2 + 2C' \sum_{k=p+1}^\infty (1 + (k\pi)^2) \frac{(1 - e^{-(k\pi)^2 t})^2}{(k\pi)^4},$$

which is independent of  $\varepsilon \in (0, 1]$ . The convergence of each series, uniformly in  $\varepsilon$ , enables us to choose a suitable  $p$  independent of  $\varepsilon$ . Let us define  $L_p = \text{span}\{e_i, 1 \leq i \leq p\}$ . We choose  $L_{\delta,\eta} = L_p$ . Since  $u_t^{\varepsilon,p} \in L_{\delta,\eta} = L_p$ ,

$$E(d(u_t^\varepsilon, L_{\delta,\eta})) \leq E(\|u_t^\varepsilon - u_t^{\varepsilon,p}\|_H) \leq C'\eta\delta.$$

Markov's inequality gives

$$\Pr(d(u_t^\varepsilon, L_{\delta,\eta}) > \eta) \leq C'\delta$$

with  $\delta$  independent of  $\varepsilon \in (0, 1]$ .

*Step 1(iii) (tightness).* Using Theorems 3 and 4 in Appendix B,  $(u^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $\mathbb{D}([0, T], H)$ . Since we know that, for each  $\varepsilon \in (0, 1]$ ,  $u^\varepsilon$  is in  $\mathcal{C}([0, T], H)$ , we know in fact that  $(u^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $\mathcal{C}([0, T], H)$ .

*Step 2(i) (reduction of the problem).* We know that the family  $(u^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $\mathcal{C}([0, T], H)$ . We still denote by  $u^\varepsilon$  a converging subsequence of  $u^\varepsilon$  and we denote by  $u$  the corresponding accumulation point. We want to show that, for all  $\phi$  in  $\mathcal{C}_0^2(I)$ ,

$$(u_t, \phi)_{L^2(I)} = (u_0, \phi)_{L^2(I)} + \int_0^t (u_s, \phi'')_{L^2(I)} ds + \int_0^t \langle F(u_s), \phi \rangle ds$$

almost surely. We will show that

$$E \left| (u_t, \phi)_{L^2(I)} - (u_0, \phi)_{L^2(I)} - \int_0^t (u_s, \phi'')_{L^2(I)} ds - \int_0^t \langle F(u_s), \phi \rangle ds \right| = 0. \tag{8}$$

Note that

$$\begin{aligned} E \left| (u_t^\varepsilon, \phi)_{L^2(I)} - (u_0, \phi)_{L^2(I)} - \int_0^t (u_s^\varepsilon, \phi'')_{L^2(I)} ds - \int_0^t \langle F(u_s^\varepsilon), \phi \rangle ds \right| \\ = E \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds \right|. \end{aligned}$$

Since the function

$$h(v) = \left| (v_t, \phi)_{L^2(I)} - (v_0, \phi)_{L^2(I)} - \int_0^t (v_s, \phi'')_{L^2(I)} ds - \int_0^t \langle F(v_s), \phi \rangle ds \right|$$

is continuous on  $\mathcal{C}([0, T], H)$ , and the arguments  $(u^\varepsilon, \varepsilon \in (0, 1])$  and  $u$  are uniformly bounded in  $\varepsilon$  in  $\mathcal{C}([0, T], H)$ , by convergence in law we have

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} E \left| (u_t^\varepsilon, \phi)_{L^2(I)} - (u_0, \phi)_{L^2(I)} - \int_0^t (u_s^\varepsilon, \phi'')_{L^2(I)} ds - \int_0^t \langle F(u_s^\varepsilon), \phi \rangle ds \right| \\ = E \left| (u_t, \phi)_{L^2(I)} - (u_0, \phi)_{L^2(I)} - \int_0^t (u_s, \phi'')_{L^2(I)} ds - \int_0^t \langle F(u_s), \phi \rangle ds \right|. \end{aligned}$$

Therefore, (8) will follow if  $\lim_{\varepsilon \rightarrow 0} E \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds \right| = 0$ . We will show more precisely that

$$\lim_{\varepsilon \rightarrow 0} E \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds \right|^2 = 0.$$

Step 2(ii) (Poisson equation). We introduce the following Poisson equation on  $f : H \times \mathcal{R} \rightarrow \mathbb{R}$ :

$$\mathcal{B}f(u, r) = \langle G_r(u) - F(u), \phi \rangle, \tag{9}$$

$$\int_{\mathcal{R}} f(u, r) \mu(u)(dr) = 0.$$

Here  $\mathcal{B}$  is the generator given in (6):

$$\mathcal{B}f(u, r) = \sum_{i \in \mathbb{N} \cap N\hat{i}} \sum_{\zeta \in E} [f(u, r(r(i) \rightarrow \zeta)) - f(u, r)] \alpha_{r(i), \zeta} \left( u \left( \frac{i}{N} \right) \right).$$

**Proposition 6.** *The Poisson equation (9) has a unique solution  $f$  which is measurable and locally Lipschitz continuous in its first variable with respect to the  $\|\cdot\|_H$  norm. For all fixed  $r \in \mathcal{R}$ , the map  $u \in H \mapsto f(u, r)$  has a Fréchet derivative denoted by  $df[u, r]/du$ . Moreover,  $\sup_{s \in [0, T]} |f(u_s^\varepsilon, r_s^\varepsilon)|$  and  $\sup_{s \in [0, T]} \|df[u_s^\varepsilon, r_s^\varepsilon]/du\|_{H^*}$  are both bounded almost surely by constants independent of  $\varepsilon$ .*

In order to prove the above proposition, we need the following lemma.

**Lemma 1.** *Let  $B \subset H$  be a bounded domain, and let  $r$  be in  $\mathcal{R}$ . The map  $u \in H \mapsto \langle G_r(u) - F(u), \phi \rangle$  is*

1. *bounded on  $B$  by a constant independent of  $r$ ,*
2. *Lipschitz continuous on  $B$  with Lipschitz constant independent of  $r$ ,*
3. *Fréchet differentiable on  $B$  with Fréchet derivative in  $u \in H$  denoted by  $df[u, r]/du$ . Moreover, the Fréchet derivative is bounded in the  $\|\cdot\|_{H^*}$  norm uniformly in  $u \in B$  and  $r \in \mathcal{R}$ .*

*Proof.* Recall that

$$\langle G_r(u) - F(u), \phi \rangle = \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\hat{i}} c_\xi \left( 1_\xi(r(i)) - \mu \left( u \left( \frac{i}{N} \right) \right) (\xi) \right) \left( v_\xi - u \left( \frac{i}{N} \right) \right) \phi \left( \frac{i}{N} \right).$$

Since, for  $\xi \in E$ ,  $\mu(\cdot)(\xi)$  is bounded and Lipschitz continuous on  $\mathbb{R}$ , points 1 and 2 follow. For point 3, we note that the map  $u \in H \mapsto \langle G_r(u) - F(u), \phi \rangle$  is a linear combination of the Fréchet differentiable functions  $u \in H \mapsto \mu(u(i/N))(\xi)$  and  $u \mapsto u(i/N)$ , and of the product of these two functions. The Fréchet derivative in  $u \in H$  against  $h \in H$  is given by

$$-\frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N\hat{i}} c_\xi \left[ \mu' \left( u \left( \frac{i}{N} \right) \right) (\xi) \left( v_\xi - u \left( \frac{i}{N} \right) \right) + \left( 1_\xi(r(i)) - \mu \left( u \left( \frac{i}{N} \right) \right) (\xi) \right) \right] \times \phi \left( \frac{i}{N} \right) h \left( \frac{i}{N} \right),$$

which gives us the boundedness property uniformly in  $r \in \mathcal{R}$  and  $u \in B$ .

*Proof of Proposition 6.* For  $u \in H$  fixed,  $\mathcal{B}(u, \cdot)$  is an operator on  $\mathbb{R}^{\mathcal{R}}$  which is a space of finite dimension. The Fredholm alternative in such a space is  $\text{Im}(\mathcal{B}) = (\ker(\mathcal{B}^*))^\perp$ . Moreover, by the ergodicity assumption, for any fixed  $u \in H$ ,  $\ker(\mathcal{B}^*(u, \cdot)) = \text{span}(\mu(u))$ .



Therefore, (9) with  $u \in H$  fixed has a solution if and only if

$$\int_{\mathcal{R}} \mu(u)(dr) \langle G_r(u) - F(u), \phi \rangle = 0.$$

This latter equality holds by the definition of  $F$  and the bilinearity of the duality product  $\langle \cdot, \cdot \rangle$ . Moreover, we can then always choose  $f$  satisfying  $\int_{\mathcal{R}} f(u, r) \mu(u)(dr) = 0$  by taking its projection on  $(\ker(\mathcal{B}^*))^\perp$ . Thus, we have a solution  $f(u, \cdot)$  for any  $u \in H$  fixed. Uniqueness of  $f$  follows easily from the condition  $\int_{\mathcal{R}} f(u, r) \mu(u)(dr) = 0$ .

Recall that, for all  $\varepsilon \in (0, 1]$ ,  $u^\varepsilon \in B$ , where  $B = \{u \in H; \|u\| \leq C\}$  and  $C$  is the deterministic constant given by Proposition 2. Then the desired properties of  $f$  follow from Lemma 1 and the fact that, for a bounded domain  $B \in H$ , each function  $u \in B \mapsto \alpha_{r(i), \zeta}(u(i/N))$  is bounded below and above by strictly positive constants, Lipschitz continuous, and Fréchet differentiable with Fréchet derivative uniformly bounded in  $u, \zeta$ , and  $r$ .

*Step 2(iii) (bound the martingale part).* Our previous result implies that  $f \in \mathcal{D}(\mathcal{A})$  and

$$M_t^\varepsilon = f(u_t^\varepsilon, r_t^\varepsilon) - f(u_0, r_0) - \int_0^t \mathcal{A}f(u_s^\varepsilon, r_s^\varepsilon) ds$$

defines a square-integrable martingale; see, for example, [10, Chapter 4, Proposition 1.7].

**Proposition 7.** *We have*

$$\langle M^\varepsilon \rangle_t = \frac{1}{\varepsilon} \int_0^t \sum_{i \in \mathbb{N} \cap N\hat{I}} \sum_{\zeta \neq r_s(i)} [f(u_s^\varepsilon, r_s(r_s(i) \rightarrow \zeta)) - f(u_s^\varepsilon, r_s^\varepsilon)]^2 \alpha_{r_s(i), \zeta} \left( u_s^\varepsilon \left( \frac{i}{N} \right) \right) ds.$$

*Proof.* The proof is classical in stochastic calculus (see, for example, [10, Chapter 1, Problem 29]). The Itô and the Dynkin formulae give two distinct decompositions of the squared process  $f^2(u_t^\varepsilon, r_t^\varepsilon)_{t \geq 0}$  in a semimartingale. The uniqueness of the Doob–Meyer decomposition of a semimartingale enables us to identify the bracket of our martingale.

We have the following semimartingale decomposition:

$$\int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds = \varepsilon f(u_t^\varepsilon, r_t^\varepsilon) - \varepsilon f(u_0, r_0) - \varepsilon \int_0^t \frac{d}{ds} f(u_s^\varepsilon, r_s^\varepsilon) ds - \varepsilon M_t^\varepsilon. \tag{10}$$

Recall that we are interested in bounding  $E|\int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds|^2$ . As  $\sup_{s \in [0, T]} |f(u_s^\varepsilon, r_s^\varepsilon)|$  is bounded independently of  $\varepsilon \in (0, 1]$ , denoting by  $\varepsilon g$  one of the two first terms of (10) we have  $E|\varepsilon g|^2 \leq C'\varepsilon^2$ , where  $C'$  is a constant independent of  $\varepsilon \in (0, 1]$ . For the martingale term, by the Itô isometry,

$$E|M_t^\varepsilon|^2 = E\langle M^\varepsilon \rangle_t \leq \frac{1}{\varepsilon} C'$$

since  $\varepsilon \langle M^\varepsilon \rangle_t$  is bounded uniformly in  $t \in [0, T]$  and  $\varepsilon \in (0, 1]$  thanks to the bounds on  $f$  and each  $\alpha_{\xi, \zeta}$ . More precisely, we have

$$\begin{aligned} \langle M^\varepsilon \rangle_t &= \frac{1}{\varepsilon} \int_0^t \sum_{i \in \mathbb{N} \cap N\hat{I}} \sum_{\zeta \neq r_s(i)} [f(u_s^\varepsilon, r_s(r_s(i) \rightarrow \zeta)) - f(u_s^\varepsilon, r_s^\varepsilon)]^2 \alpha_{r_s(i), \zeta} \left( u_s^\varepsilon \left( \frac{i}{N} \right) \right) ds \\ &\leq \frac{4T}{\varepsilon} \alpha^+ N |E| \sup_{s \in [0, T]} |f(u_s^\varepsilon, r_s^\varepsilon)|^2. \end{aligned}$$

*Step 2(iv) (bound the finite-variation part).* It remains to bound the third term of our semimartingale decomposition.

**Proposition 8.** *There exists a constant  $C$  independent of  $\varepsilon \in (0, 1]$  such that*

$$\int_0^T \left| \frac{d}{dt} f(u_t^\varepsilon, r_t^\varepsilon) \right| dt \leq C$$

almost surely.

*Proof.* First recall the meaning of the notation  $df(u_t^\varepsilon, r_t^\varepsilon)/dt$  given in Theorem 1. Recall that the map  $u \mapsto f(u, r)$  for  $r \in \mathcal{R}$  fixed is Fréchet differentiable with Fréchet derivative in  $u$  denoted by  $df[u, r]/du$ , which is a bounded linear form on  $H$ . By the Riesz representation theorem, there exists  $f_u(u, r) \in H$  such that

$$\frac{df}{du}[u, r](h) = (f_u(u, r), h)_H \quad \text{for all } h \in H.$$

Moreover, the correspondence is isometric:  $\|df[u, r]/du\|_{H^*} = \|f_u(u, r)\|_H$ . We then have (see Theorem 4 of [2])

$$\frac{d}{dt} f(u_t^\varepsilon, r_t^\varepsilon) = \langle \Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon), f_u(u_t^\varepsilon, r_t^\varepsilon) \rangle.$$

By Proposition 6, there exists a constant  $C_1$  independent of  $\varepsilon \in (0, 1]$  and  $t \in [0, T]$  such that

$$\|f_u(u_t^\varepsilon, r_t^\varepsilon)\|_H = \left\| \frac{df}{du}[u_t^\varepsilon, r_t^\varepsilon] \right\|_{H^*} \leq C_1.$$

Therefore,

$$\left| \frac{d}{dt} f(u_t^\varepsilon, r_t^\varepsilon) \right| \leq C_1 \|\Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon)\|_{H^*}.$$

It remains to show that  $\|\Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon)\|_{H^*}$  is bounded uniformly in  $\varepsilon$  and  $t \in [0, T]$ . For  $\phi \in H$ , we have, denoting by  $D$  the derivative with respect to  $x$ ,

$$\begin{aligned} & | \langle \Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon), \phi \rangle | \\ &= \left| -\langle Du_t^\varepsilon, D\phi \rangle + \frac{1}{N} \sum_{\xi \in E} \sum_{i \in \mathbb{N} \cap N_i} c_\xi 1_\xi(r(i)) \left( v_\xi - u_t^\varepsilon \left( \frac{i}{N} \right) \right) \phi \left( \frac{i}{N} \right) \right| \\ &\leq |(Du_t^\varepsilon, D\phi)_{L^2(I)}| + |E| \max_{\xi \in E} c_\xi \left( \max_{\xi \in E} |v_\xi| + C \right) C_P \|\phi\|_H \\ &\leq \|Du_t^\varepsilon\|_{L^2(I)} \|D\phi\|_{L^2(I)} + |E| \max_{\xi \in E} c_\xi \left( \max_{\xi \in E} |v_\xi| + C \right) C_P \|\phi\|_H \\ &\leq \|u_t^\varepsilon\|_H \|\phi\|_H + |E| \max_{\xi \in E} c_\xi \left( \max_{\xi \in E} |v_\xi| + C \right) C_P \|\phi\|_H \\ &\leq \left( C + |E| \max_{\xi \in E} c_\xi \left( \max_{\xi \in E} |v_\xi| + C \right) C_P \right) \|\phi\|_H, \end{aligned}$$

where  $C$  is the deterministic constant given by Proposition 2. Thus,

$$\|\Delta u_t^\varepsilon + G_{r_t^\varepsilon}(u_t^\varepsilon)\|_{H^*} \leq C + |E| \max_{\xi \in E} c_\xi \left( \max_{\xi \in E} |v_\xi| + C \right) C_P.$$

This completes the proof.

Step 2(v) (aggregate bounds). Assembling all the bounds of the different terms we see that

$$E \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F(u_s^\varepsilon), \phi \rangle ds \right|^2 \leq C\varepsilon,$$

with the constant  $C$  independent of  $\varepsilon \in (0, 1]$ . It remains to let  $\varepsilon$  go to 0 to conclude the proof of the identification of the limit.

Step 2(vi) (uniqueness). We consider  $u^1$  and  $u^2$  to be two possible accumulation points verifying, for  $i = 1, 2$ ,

$$u_t^i = e^{\Delta t} u_0 + \int_0^t e^{\Delta(t-s)} F(u_s^i) ds$$

for all  $t \in [0, T]$  almost surely. We can easily show that, for any  $t \in [0, T]$ ,

$$E \left( \sup_{s \in [0, t]} \|u_s^1 - u_s^2\|_H \right) \leq C \int_0^t E \left( \sup_{l \in [0, s]} \|u_l^1 - u_l^2\|_H \right) dl,$$

with the constant  $C$  independent of  $t \in [0, T]$ . Therefore, by application of Gronwall’s lemma we obtain  $E(\sup_{t \in [0, T]} \|u_t^1 - u_t^2\|_H) = 0$ .

### 3.2. The general case

We now consider the original case with different state classes  $E_1, \dots, E_l$ . We first prove the tightness of the family  $((u^\varepsilon, \bar{r}^\varepsilon), \varepsilon \in (0, 1])$  in  $\mathbb{D}([0, T], H \times \{1, \dots, l\}^{|\mathbb{N} \cap N \setminus I|})$  (see Section 2.3.3 for the definition of  $\bar{r}^\varepsilon$ ). We apply a comparison argument. We note that

$$\begin{aligned} \Pr(r_{t+h}^\varepsilon(i) = \zeta \mid r_t^\varepsilon(i) = \xi) &= \begin{cases} \frac{1}{\varepsilon} \alpha_{\xi, \zeta} \left( u_t \left( \frac{i}{N} \right) \right) h + o(h) & \text{if } \xi \text{ and } \zeta \text{ are in the same class } E_k, \\ \alpha_{\xi, \zeta} \left( u_t \left( \frac{i}{N} \right) \right) h + o(h) & \text{otherwise.} \end{cases} \end{aligned}$$

For  $\xi, \zeta \in E$ , we set

$$\lambda_{\xi, \zeta} = \begin{cases} \frac{1}{\varepsilon} \alpha^+ & \text{if } \xi \text{ and } \zeta \text{ are in the same class } E_k, \\ \alpha^+ & \text{otherwise.} \end{cases}$$

We denote by  $r^{\varepsilon, \text{Max}}$  the associated jump process with constant rates  $\lambda_{\xi, \zeta}$ . We have the following stochastic domination:

$$\Pr(r_{t+h}^\varepsilon(i) = \zeta \mid r_t^\varepsilon(i) = \xi) \leq \Pr(r_{t+h}^{\varepsilon, \text{Max}}(i) = \zeta \mid r_t^{\varepsilon, \text{Max}}(i) = \xi).$$

We construct the process  $\bar{r}^{\varepsilon, \text{Max}}$  as  $\bar{r}^\varepsilon$  by aggregation. The sequence  $(\bar{r}^{\varepsilon, \text{Max}}, \varepsilon \in (0, 1])$  is tight in  $\mathbb{D}([0, T], \{1, \dots, l\}^{|\mathbb{N} \cap N \setminus I|})$ ; see, for instance, [27, Chapter 7, Theorem 7.4]. Therefore, by comparison, the sequence  $(\bar{r}^\varepsilon, \varepsilon \in (0, 1])$  is also tight in  $\mathbb{D}([0, T], \{1, \dots, l\}^{|\mathbb{N} \cap N \setminus I|})$ ; see, for instance, [21, Chapter 7, Theorem 7.4]. Moreover, the sequence  $(u^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $\mathbb{D}([0, T], H)$  by applying the arguments developed for the ‘all-fast’ case. Endowing  $H \times \{1, \dots, l\}^{|\mathbb{N} \cap N \setminus I|}$  with the product topology, we see that the sequence  $((u^\varepsilon, \bar{r}^\varepsilon), \varepsilon \in (0, 1])$  is tight in  $\mathbb{D}([0, T], H \times \{1, \dots, l\}^{|\mathbb{N} \cap N \setminus I|})$ .

We must now deal with the identification of the limit. There are six steps in the proof of the identification of the limit, we have to check that these six steps generalize to the general case (jumping slow–fast case).

In step 2(i), by the same arguments, we find that it is sufficient to show that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left| \int_0^t \langle G_{r_s^\varepsilon}(u_s^\varepsilon) - F_{\bar{r}_s^\varepsilon}(u_s^\varepsilon), \phi \rangle ds \right|^2 = 0.$$

In step 2(ii), the Poisson equation becomes

$$\mathcal{B}f(u, r) = \langle G_r(u) - F_{\bar{r}}(u), \phi \rangle, \tag{11}$$

where  $\mathcal{B}$  is now the ‘fast’ part of the generator ( $f$  does not depend on  $\bar{r}$  explicitly because  $\bar{r}$  is constructed from  $r$ ). A configuration  $r$  for the ion channels is now

$$r \in E_{j_1} \times \cdots \times E_{j_{N-1}},$$

where  $(j_1, \dots, j_{N-1}) \in \{1, \dots, l\}^{N-1}$  (noting that  $|\mathbb{N} \cap N\dot{I}| = N - 1$ ). That is, each channel is in one of the  $E_j$  classes for  $j \in \{1, \dots, l\}$ . Then, since, for fixed  $u$ , the quasistationary measure associated to the class  $E_{j_k}$  for  $k \in \{1, \dots, N - 1\}$  is  $\mu_{j_k}(u)$ , we find that the kernel of  $\mathcal{B}^*$  is spanned by

$$\{\mu_{j_1}(u) \times \cdots \times \mu_{j_{N-1}}(u), (j_1, \dots, j_{N-1}) \in \{1, \dots, l\}^{N-1}\}$$

when  $u$  is held fixed. Then, by the Fredholm alternative and the definition of the averaged function  $F_{\bar{r}}(u)$ , the Poisson equation (11) has a solution. Uniqueness follows by the projection condition:

$$\int_{\mathcal{R}} f(u, r) \mu_{j_1}(u)(dr) \times \cdots \times \mu_{j_{N-1}}(u)(dr) = 0 \quad \text{for all } (j_1, \dots, j_{N-1}) \in \{1, \dots, l\}^{N-1}.$$

The proofs of the last four steps are then exactly the same as in the ‘all-fast’ case.

### 4. Example

In this section we give a concrete example where our result allows us to reduce the complexity of a neuronal model staying nevertheless at a stochastic level.

We consider a usual Hodgkin–Huxley model but, to be very straightforward in the application of our result, we consider only the sodium current, i.e. all the ion channels are sodium channels. It is still a case of interest since sodium channels are involved in the increasing phase of an action potential. At a fixed potential, the kinetic of a sodium channel is described by Figure 1, which represents the states and the jump rates of a continuous-time Markov chain denoted by  $r^\varepsilon(i)$  for the channel at location  $i/N$  for  $i \in \mathbb{N} \cap N\dot{I}$ . A sodium channel can be in eight different states, denoted by  $m_i h_j$  for  $i \in \{0, 1, 2, 3\}$  and  $j \in \{0, 1\}$ ; the state  $m_3 h_1$  is the only ‘open’ state for the sodium channel (see [19, p. 487] for more details). We simply write  $a_m, b_m$  and  $a_h, b_h$  for  $a_m(u), b_m(u)$  and  $a_h(u), b_h(u)$  (see Appendix C for more details on the rate functions). Here  $m$  is the fast variable and  $h$  is the slow variable for the kinetic of a sodium channel.

Each channel can therefore be in one of these eight states divided into two classes:  $E = E_0 \sqcup E_1$ , where  $E_0 = \{m_0 h_0, m_1 h_0, m_2 h_0, m_3 h_0\}$  and  $E_1 = \{m_0 h_1, m_1 h_1, m_2 h_1, m_3 h_1\}$ . Inside these two classes the states communicate at fast rates and transition between these two

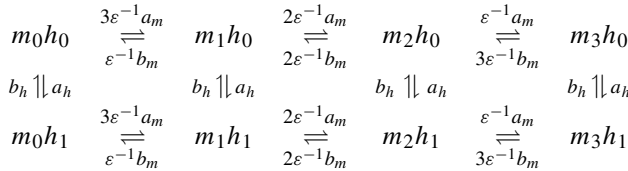


FIGURE 1: Kinetic of sodium channels.

classes occurs at slow rates. A sodium channel is open if and only if it is in the state  $m_3h_1$ . We then define  $c_{m_3h_1} = c_{Na}$ ,  $v_{m_3h_1} = v_{Na}$  and  $c_\xi = 0$ ,  $v_\xi = 0$  when  $\xi \in E \setminus \{m_3h_1\}$ .

Equation (2), which describes the evolution of the potential for a number  $N$  of ion channels, is given here by

$$\partial_t u^\varepsilon = K \Delta u^\varepsilon + \frac{1}{N} \sum_{i \in \mathbb{N} \cap N\mathbb{I}} c_{Na} 1_{m_3h_1}(r^\varepsilon(i)) \left( v_{Na} - u \left( \frac{i}{N} \right) \right) \delta_{i/N}. \tag{12}$$

The global variable  $u$  represents the difference in potential between the inside and the outside of the axon membrane. We see that (12) is indeed of the form that we have studied in this paper. The constant  $K$  is related to the radius and the internal conductivity of the axon (see Appendix C).

Let us compute the different generators and quasi-invariant measures associated to this model according to our description. The two ‘fast’ generators are the same and are given for  $u(i/N)$  fixed (again, we do not make the dependence explicit) by

$$\mathcal{B}_j = \begin{pmatrix} -3a_m & 3a_m & 0 & 0 \\ b_m & -b_m - 2a_m & 2a_m & 0 \\ 0 & 2b_m & -2b_m - a_m & a_m \\ 0 & 0 & 3b_m & -3b_m \end{pmatrix}$$

with  $j = 0$  or  $1$ . We can compute the associated quasi-invariant measure  $\mu_j(u(i/N))$  ( $j = 0$  or  $1$ ); the only term of interest to us is

$$\mu_1 \left( u \left( \frac{i}{N} \right) \right) (m_3h_1) = \left( 1 + \frac{b_m(u(i/N))}{a_m(u(i/N))} \right)^{-3}.$$

The reader familiar with the classical Hodgkin–Huxley model will note that  $\mu_1(u(i/N))(m_3h_1)$  is in fact equal to the steady-state function associated to the ordinary differential equation describing the motion of the  $m$ -gates in the classical deterministic Hodgkin–Huxley description; see, for example, [20].

The asymptotic aggregated Markov chain  $\bar{r}$  is valued in the two-state space  $\{0, 1\}$ . According to Proposition 3, its generator is

$$\begin{pmatrix} -a_h(u(i/N)) & a_h(u(i/N)) \\ b_h(u(i/N)) & -b_h(u(i/N)) \end{pmatrix}.$$

Therefore, according to Theorem 2, the reduced model is described by the following PDE coupled with the continuous-time Markov chain  $\bar{r}$ :

$$\partial_t u = K \Delta u + \frac{1}{N} \sum_{i \in \mathbb{N} \cap N\mathbb{I}} 1_{\bar{r}(i)} c_{Na} \mu_1 \left( u \left( \frac{i}{N} \right) \right) (m_3h_1) \left( v_{Na} - u \left( \frac{i}{N} \right) \right) \delta_{i/N}. \tag{13}$$

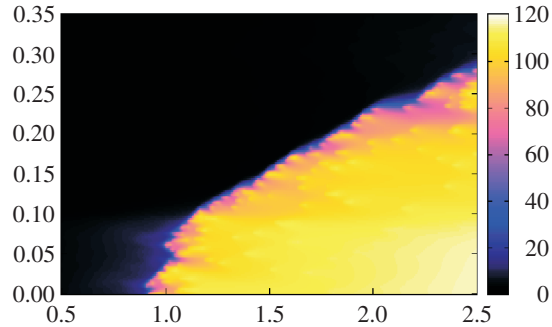


FIGURE 2: Simulation of the action potential against space (*vertical axis*) and time (*horizontal axis*). The averaged model (13) is displayed for a number of ion channels fixed to 250. We see what we expect from the PDE: a traveling wave connecting the initial state 0 to the stable state  $v_{Na}$ .

We display a realization of the averaged PDMP in Figure 2. To perform the numerical simulation (in C), we extended Riedler's work (cf. [25]) to our particular framework. In [25], an algorithm (Algorithm A2) was proposed to simulate the trajectory of a PDMP in finite dimensions. The idea is to simulate the jumping part of the PDMP and to use an accurate method to simulate the ordinary differential equation constituting the deterministic part of the PDMP. The kinetics of the jumping component for PDMPs in finite and infinite dimensions have the same form. Here we simulate the jumping part of our infinite-dimensional PDMP following [25] and simulate the PDE between successive jumps by a deterministic scheme. For the PDE, we used an explicit finite difference Euler scheme in space and time. Simulating the jumps of an inhomogeneous time Markov chain is classical and there exist a lot of efficient numerical schemes for PDEs. Therefore, the described method is natural. This argument is of course not a proof but gives an heuristic interpretation of our approach.

## 5. Concluding remarks

In this paper we applied the methods of stochastic averaging to a fully coupled piecewise-deterministic Markov process (PDMP) in infinite dimensions. To the best of our knowledge, it is the first time that such a problem has been considered. The two main difficulties are due to the infinite-dimensional setting and the fact that the system is fully coupled. Regarding the fully coupled property, our task was made easier by the uniform bound on the macroscopic component of the PDMP and on the jump rates of the continuous-time Markov chain. As an application, we showed how to reduce the complexity of a generalized Hodgkin–Huxley model which can be used as a paradigm in the modeling of the propagation of an action potential for neurons but also for cardiac cells.

Let us make some comments about a possible extension of our work to other forms of generalized Hodgkin–Huxley models. Firstly, to choose that the ion channels are uniformly distributed along the axon ( $i/N$  for  $i \in \mathbb{N} \cap N\dot{I}$ ) is just a simple way to fix ideas. We could also choose any distribution in a finite subset  $\mathcal{N}$  of  $\dot{I}$  without changing anything to our work except the notation.

Secondly, in [2], the authors noted that it can be of interest to replace the Dirac point masses  $\delta_{i/N}$  of the spatial stochastic Hodgkin–Huxley model by functions of the form  $|U(i/N)|^{-1} \times 1_{U(i/N)}(\cdot)$ , where  $U$  is a neighborhood of the point  $i/N$ . Indeed, the concentration of ions is essentially homogeneous in a spatially extended domain around an open channel. Along the

same line, we could replace the Dirac point masses  $\delta_{i/N}$  by any smoother approximation  $\phi_{i/N}$ . The mathematics would become a little easier. The methods presented in this paper apply to both cases.

The next step after the present averaging result is to prove the associated central limit theorem. That is, to give more precision on how the original model converges in law to the averaged model.

Another line of work would be to consider that not only some ion channels but also the membrane potential have a faster kinetic. It seems that this situation is closer to the real behavior of the nerve impulse [19], and is related to the study of invariant measures for PDMPs and the necessity to find explicit forms or bounds for these invariant measures. These will be the object of a future work.

### Appendix A. Austin’s lemma about the semigroup $e^{\Delta t}$

We recall here a lemma in [1]. Note that the last part of statement 1 of the following lemma is contained in the proof of the corresponding lemma in [1].

**Lemma 2.** *Let  $y$  be in the interior of  $I$ . Then  $e^{\Delta t} \delta_y$  is a smooth function on  $I$  vanishing at the endpoints for any  $t > 0$ . Furthermore, the following statements hold.*

1. *There exists some constant  $C_1 > 0$ , depending on  $T$  but otherwise not on  $t \in [0, T]$ , such that, for any continuous function  $u : [0, t] \rightarrow \mathbb{R}$ , the function*

$$I \rightarrow \mathbb{R} : x \mapsto \int_0^t u(s) e^{\Delta(t-s)} \delta_y(x) \, ds$$

*is in  $H$  and satisfies the estimate*

$$\left\| \int_0^t u(s) e^{\Delta(t-s)} \delta_y(\cdot) \, ds \right\|_H \leq C_1 \|u\|_\infty.$$

*Moreover, for any  $\eta > 0$ , we can choose  $\varepsilon > 0$  so small that*

$$\left\| \int_{t-\varepsilon}^t u(s) e^{\Delta(t-s)} \delta_y(\cdot) \, ds \right\|_H \leq \frac{1}{2} \eta \|u\|_\infty.$$

2. *For any fixed  $\varepsilon > 0$ , there exists some constant  $C_2(\varepsilon)$ , depending on  $\varepsilon$  and  $T$  but otherwise not on  $t \in [0, T]$ , such that, for any continuous function  $u : [0, t] \rightarrow \mathbb{R}$ , we have*

$$\left\| \int_0^{t-\varepsilon} u(s) e^{\Delta(t-s)} \delta_y(\cdot) \, ds \right\|_H \leq C_2(\varepsilon) \int_0^{t-\varepsilon} |u(s)| \, ds$$

*for any  $t \in [0, T]$ .*

### Appendix B. Two theorems about tightness in Hilbert spaces

We recall here two important results from [22] about tightness in Hilbert spaces with the notation and choice of spaces adapted to our setting.

**Theorem 3.** (General criterion for tightness.) *Let us assume that  $(u^\varepsilon, \varepsilon \in (0, 1])$  satisfies Aldous’s condition, which means that, for any  $\delta, M > 0$ , there exist  $\eta, \varepsilon_0 > 0$  such that, for all stopping times  $\tau$  such that  $\tau + \eta < T$ ,*

$$\sup_{\varepsilon \in (0, \varepsilon_0]} \sup_{\theta \in (0, \eta)} \Pr(\|u_{\tau+\theta}^\varepsilon - u_\tau^\varepsilon\|_H \geq M) \leq \delta.$$

If, moreover, for each  $t \in [0, T]$  fixed, the family  $(u_t^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $H$  then  $(u^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $\mathbb{D}([0, T], H)$ .

**Theorem 4.** (Tightness in a Hilbert space.)  $H$  is a separable Hilbert space endowed with a basis  $\{e_k, k \geq 1\}$ . We define, for  $k \geq 1$ ,

$$L_k = \text{span}\{e_i, 1 \leq i \leq k\}.$$

Then  $(u_t^\varepsilon, \varepsilon \in (0, 1])$  is tight in  $H$  if and only if, for any  $\delta, \eta > 0$ , there exist  $\rho, \varepsilon_0 > 0$  and  $L_{\delta,\eta} \subset \{L_k, k \geq 1\}$  such that

$$\sup_{\varepsilon \in (0, \varepsilon_0]} \Pr(\|u_t^\varepsilon\|_H > \rho) \leq \delta, \tag{14}$$

$$\sup_{\varepsilon \in (0, \varepsilon_0]} \Pr(d(u_t^\varepsilon, L_{\delta,\eta}) > \eta) \leq \delta, \tag{15}$$

where  $d(u_t^\varepsilon, L_{\delta,\eta}) = \inf_{v \in L_{\delta,\eta}} \|u_t^\varepsilon - v\|_H$ .

### Appendix C. Functions and data for the simulation

For the simulation, the jump rate functions are given for real  $u$  by

$$\begin{aligned} a_m(u) &= \frac{0.1(25 - u)}{e^{2.5-0.1u} - 1}, \\ b_m(u) &= 4e^{-u/18}, \\ a_h(u) &= 0.07e^{-u/20}, \\ b_h(u) &= \frac{1}{e^{3-0.1u} + 1}. \end{aligned}$$

The maximal conductance associated to sodium ions is  $c_{Na} = 120 \text{ mS/cm}^2$  and the potential at rest is  $v_{Na} = 115 \text{ mV}$ . The constant  $K$  is given by  $K = a/2R$ , where  $a$  is the radius of the axon,  $a = 0.0238 \text{ cm}$ , and  $R$  is the internal resistance of the axon,  $R = 34.5 \Omega \text{ cm}$ —these data are classical; see, for example, [20]. We have used an input on the potential equal to  $\mu = 6.7$  when in the segment  $[0, 0.1]$  of the axon.

### Acknowledgement

The authors warmly thank the anonymous referee for his/her helpful comments, suggestions, and corrections that led to a significant improvement of the paper.

### 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] BUCKWAR, R. AND RIEDLER, M. G. (2011). An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *J. Math. Biol.* **63**, 1051–1093.
- [3] CHOW, C. C. AND WHITE, J. A. (1996). Spontaneous action potentials due to channel fluctuations. *Biophys. J.* **71**, 3013–3021.
- [4] COSTA, O. L. V. AND DUFOUR, F. (2008). Stability and ergodicity of piecewise deterministic Markov processes. *SIAM J. Control Optimization* **47**, 1053–1077.
- [5] COSTA, O. L. V. AND DUFOUR, F. (2011). Singular perturbation for the discounted continuous control of piecewise deterministic Markov processes. *Appl. Math. Optimization* **63**, 357–384.
- [6] CRUDU, A., DEBUSSCHE, A., MULLER, A. AND RADULESCU, O. (2012). Convergence of stochastic gene networks to hybrid piecewise deterministic processes. To appear in *Ann. Appl. Prob.*



- [7] DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Statist. Soc. B* **46**, 353–388.
- [8] DAVIS, M. H. A. (1993). *Markov Models and Optimization* (Monogr. Statist. Appl. Prob. **49**). Chapman and Hall, London.
- [9] 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.
- [10] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.
- [11] FAGGIONATO, A., GABRIELLI, D. AND CRIVELLARI, M. R. (2010). Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors. *Markov Process. Relat. Fields* **16**, 497–548.
- [12] FAISAL, A. A., WHITE, J. A. AND LAUGHLIN, S. B. (2005). Ion-channel noise places limits on the miniaturization of the brain's wiring. *Current Biol.* **15**, 1143–1149.
- [13] FITZHUGH, R. (1960). Thresholds and plateaus in the Hodgkin-Huxley nerve equations. *J. Gen. Physiol.* **43**, 867–896.
- [14] FLAIM, S. N., GILLES, W. R. AND McCULLOCH, A. D. (2006). Contributions of sustained  $I_{Na}$  and  $I_{Kv43}$  to transmural heterogeneity of early repolarization and arrhythmogenesis in canine left ventricular myocytes. *Am. J. Physiol. Heart Circ. Physiol.* **291**, 2617–2629.
- [15] FONBONA, J., GUÉRIN, H. AND MALRIEU, F. (2012). Quantitative estimates for the long time behavior of a PDMP describing the movement of bacteria. Submitted.
- [16] FOX, R. F. (1997). Stochastic versions of the Hodgkin-Huxley equations. *Biophys. J.* **72**, 2068–2074.
- [17] GREENSTEIN, J. L., HINCH, R. AND WINSLOW, R. L. (2006). Mechanisms of excitation-contraction coupling in an integrative model of the cardiac ventricular myocyte. *Biophys. J.* **90**, 77–91.
- [18] HENRY, D. (1981). *Geometric Theory of Semilinear Parabolic Equations* (Lecture Notes Math. **840**). Springer, Berlin.
- [19] HILLE, B. (1992). *Ionic Channels of Excitable Membranes*, 2nd edn. Sinauer, Sunderland, MA.
- [20] HODGKIN, A. L. AND HUXLEY, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544.
- [21] JACOD, J. AND SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin.
- [22] MÉTIVIER, M. (1984). Convergence faible et principe d'invariance pour des martingales à valeurs dans des espaces de Sobolev. *Ann. Inst. H. Poincaré Prob. Statist.* **20**, 329–348.
- [23] PAKDAMAN, K., THIEULLEN, M. AND WAINRIB, G. (2012). Asymptotic expansion and central limit theorem for multiscale piecewise deterministic Markov processes. To appear in *Stoch. Process Appl.*
- [24] PAVLIOTIS, G. A. AND STUART, A. M. (2008). *Multiscale Methods* (Texts Appl. Math. **53**). Springer, New York.
- [25] RIEDLER, M. G. (2011). Almost sure convergence of numerical approximations for piecewise deterministic Markov processes. Preprint. Available at <http://arxiv.org/abs/1112.1190v1>.
- [26] RIEDLER, M. G., THIEULLEN, M. AND WAINRIB, G. (2012). Limit theorems for infinite-dimensional piecewise deterministic processes and applications to stochastic neuron models. Submitted.
- [27] YIN, G. G. AND ZHANG, Q. (1997). *Continuous-Time Markov Chains and Applications*. Springer, Berlin.