Dipartimento di Matematica, Università di Trento
Preprint UTM - 725

Large deviation principle for stochastic FitzHugh-Nagumo equations on networks

Stefano BONACCORSI
Elisa MASTROGIACOMO
Dipartimento di Matematica, Università di Trento,
Via Sommarive 14, 38050 Povo (Trento), Italia
stefano.bonaccorsi@unitn.it

1. Introduction

Electric propagation in neurons has been studied since the 50s, starting with the now classical Hodgkin-Huxley model [13] for the diffusion of the transmembrane electrical potential in a neuronal cell. It is widely accepted in the literature that the internal diffusion of such a potential is described by a semilinear equation $\dot{u} = (cu_x)_x + f(u)$. Until not long ago, such problems were discussed mainly for isolated neurons, which are described by a finite or semi-infinite interval, compare with [5, 6]. Recently, new mathematical tools have been developed in order to analyse the behaviour of models describing a whole network of interacting neurons. Among other papers dealing with this case (usually modeled as a graph with m edges and n nodes), which is intended to be a simplified model for a large region of the brain, let us mention a series of recent papers by Mugnolo et al. [4, 16], where the well-posedness of the isolated system is studied.

In this paper we study a system of nonlinear diffusion equations on a finite network in the presence of a Wiener noise acting on the system. We propose to identify every node with a soma (body of the neuron cell) while edges are equivalent cylinders which models the interactions between different neurons; we allow a rather general nonlinear drift term in the cable equation, which includes, in particular, dissipative functions of the FitzHugh-Nagumo type (i.e. $f(u) = -u(u-1)(u-a)$) proposed in various models of neurophysiology (see e.g. the monograph [14] for more details).

In our neural network, every cell is subject to a chaotic stream of excitatory and inhibitory action potentials coming from the surroundings, and arriving continuously and randomly in time; the membrane's potential in the soma thus follows an Ornstein-Uhlenbeck type process, perturbed by the diffusion of electrical potential due to the presence of other conductors (the different cables which connect it to other neurons).

Several papers deal with the stochastic behaviour of neurons near the threshold level a, both from a theoretical and a numerical point of view. Our interest is mainly set in studying the small noise asymptotic of the system. This is motivated by recent researches in *in-vivo* neuronal activities. Cortical cells detect and employ large signals excursions in the generator potential that exceeds that of the background noise (i.e., they behave like large deviation detectors). In our model, we can estimate the probability that some of the neurons develop an action potential (active impulse propagation along the axon) in presence of a background stochastic noise. We will show that this probability decays exponentially with the intensity of the noise. We refer to [20] for experimental evidences and biological applications.

2. The mathematical setting

Our formulation of the problem reflects the construction of diffusions on graphs as proposed – among others – by von Below and Nicaise [2]; we obtain that the diffusion is governed by an operator matrix A and the problem can be modeled by a stochastic differential equation on an Hilbert space H. Generation properties of A can be obtained (see for instance [16]) by means of variational methods (based on the theory of sesquilinear Dirichlet forms): then we are allowed to use the semigroups methods for SDEs of Da Prato and Zabczyk [8] in order to solve our problem. In the last part of this section, we finally introduce the Large Deviation Principle (LDP for short) and discuss the main result of the paper.

2.1. **Diffusion on network.** The network is identified with the underlying graph G, described by a set of n vertices $\mathsf{v}_1, \dots, \mathsf{v}_n$ and m oriented edges $\mathsf{e}_1, \dots, \mathsf{e}_m$ which we assume to be normalized, i.e., $\mathsf{e}_j = [0,1]$. The graph is described by the *incidence matrix* $\Phi = \Phi^+ - \Phi^-$, where $\Phi^+ = (\phi^+_{ij})_{n\times m}$ and $\Phi^- = (\phi^-_{ij})_{n\times m}$ are given by

$$\phi^-_{ij} = \begin{cases} 1, & \mathsf{v}_i = \mathsf{e}_j(1) \\ 0, & \text{otherwise} \end{cases} \qquad \phi^+_{ij} = \begin{cases} 1, & \mathsf{v}_i = \mathsf{e}_j(0) \\ 0, & \text{otherwise.} \end{cases}$$

The degree of a vertex is the number of edges entering or leaving the node. Let us set

$$\Gamma(\mathsf{v}_i) = \{j \in \{1,\dots,m\} \ : \ \mathsf{e}_j(0) = \mathsf{v}_i \text{ or } \mathsf{e}_j(1) = \mathsf{v}_i\}$$

hence the degree of the vertex v_i is the cardinality $|\Gamma(\mathsf{v}_i)|$ of the set $\Gamma(\mathsf{v}_i)$.

The electrical potential in the network shall be denoted by $\bar u(t,x)$ where $\bar u \in (L^2(0,1))^m$ is the vector $(u_1(t,x),\dots,u_m(t,x))$ and $u_j(t,\cdot)$ is the electrical potential on the edge e_j. We impose a general diffusion equation on every edge

$$\frac{\partial}{\partial t} u_j(t,x) = \frac{\partial}{\partial x}\left(c_j(x)\frac{\partial}{\partial x}u_j(t,x)\right) + f_j(u_j(t,x)), \tag{2.1}$$

for all $(t,x) \in \mathbb{R}_+ \times (0,1)$ and all $j = 1,...,m$, where we are concerned with a family of weights $c_j(x) \in C^1([0,1])$ and nonlinear functions f_j. Precise assumptions on these coefficients will be given later. The generality of the above diffusion is motivated by the discussion in the biological literature, see for example [17], where it is remarked that the basic cable properties are not constant throughout the cell. The above equation shall be endowed with suitable boundary and initial conditions. Initial conditions are given for simplicity at time $t = 0$ of the form

$$u_j(0,x) = u_{j0}(x) \in C([0,1]), \qquad j = 1,...,m. \tag{2.2}$$

Since we are dealing with a diffusion in a network, we require a continuity assumption on every node

$$p_i(t) := u_j(t,\mathsf{v}_i) = u_k(t,\mathsf{v}_i), \qquad t>0,\ j,k \in \Gamma(\mathsf{v}_i),\ i = 1,...,n \tag{2.3}$$

and a stochastic generalized Kirchhoff law in the nodes

$$\frac{\partial}{\partial t} p_i(t) = -d_i p_i(t) + \sum_{j\in\Gamma(\mathsf{v}_i)} \phi_{ij}\mu_j c_j(\mathsf{v}_i)\frac{\partial}{\partial x}u_j(t,\mathsf{v}_i) + b_i \frac{\partial}{\partial t}W(t,\mathsf{v}_i), \tag{2.4}$$

for all $t > 0$ and $i = 1, \dots, n$. Postsynaptic potentials can have graded amplitudes modeled by the constants $\mu_j > 0$ for all $j = 1, ..., m$; finally, $W(t, \mathsf{v}_i)$, $i = 1, ..., n$, represent the stochastic perturbation acting on each node, due to the external surrounding, and $\frac{\partial}{\partial t} W(t, \mathsf{v}_i)$ is the formal time derivative of the process W, which takes a meaning only in integral sense.

Let us state the main assumptions on the data of the problem.

Hypothesis 2.1.
(1) *In (2.1), we assume that $c_j(\cdot)$ belongs to $C^1([0, 1])$, for $j = 1, \dots, m$ and $c_j(x) > 0$ for every $x \in [0, 1]$.*
(2) *There exists constants $\eta \in \mathbb{R}$, $c_0 > 0$ and $s \geq 1$ such that, for $j = 1, \dots, m$, the functions $f_j(u)$ satisfy $f_j(u) + \eta u$ is continuous and decreasing, and $|f_j(u)| \leq c_0(1 + |u|^s)$.*
(3) *In (2.4), we assume that $d_i \geq 0$ for every $i = 1, \dots, n$ and at least one of the coefficients d_i is strictly positive.*
(4) *$\{\mu_j\}_{j=1,\dots,m}$ and $\{b_i\}_{i=1,\dots,n}$ are positive real numbers.*
(5) *Given a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ satisfying the standard assumptions, we let $\{W_t,\ t \geq 0\}$ be a n-dimensional Wiener process defined on it.*

2.2. **The abstract setting.** There is a by-now classical way to treat system (2.1)-(2.2)-(2.3)-(2.4) by formally rewriting the problem as an abstract stochastic differential equation in an infinite dimensional Hilbert space $H = (L^2(0, 1))^m \times \mathbb{R}^n$. We shall denote with σ the diagonal matrix with entries b_i: $b = diag(b_1, \dots, b_n)$. Then, with obvious notation, we denote B the matrix operator on $\mathbb{R}^n$ taking values in $H = (L^2(0, 1))^m \times \mathbb{R}^n$ which acts as $Bp = \binom{0}{bp}$. With this notation we are get to write our system in the form

$$\begin{aligned} \mathrm{d}u(t) &= [Au(t) + F(u(t))]\,\mathrm{d}t + B\,\mathrm{d}W(t), \qquad t \in [0, T] \\ u(0) &= u_0. \end{aligned} \tag{2.5}$$

In the following propositions we read the properties of the coefficents in (2.5) under the assumptions in Hypothesis 2.1. First, following [16], we treat the linear operator A leading the drift part of the equation.

Proposition 2.2. *The linear operator A generates an analytic semigroup of contractions on $H = (L^2(0, 1))^m \times \mathbb{R}^n$. Further, the operator A is invertible and the semigroup $\{e^{tA},\ t \geq 0\}$ generated by A is exponentially bounded, with growth bound given by the strictly negative spectral bound of the operator A.*

Next, let us consider the nonlinearity F. In general, F does not need to be defined on the whole H, as it is, for instance, in the case of FitzHugh-Nagumo nonlinearity. Hence, we introduce a (reflexive) Banach space $X = (C([0, 1]))^m \times \mathbb{R}^n$, continuously embedded in H, such that $X \subset D(F)$ ($D(F)$ being the definition domain of F in H) but large enough to contain the trajectories of the stochastic convolution process.

Proposition 2.3. *The perturbation term F maps X into X, it is uniformly continuous on bounded sets of X and F is m-dissipative on X.*

Further, the part A_X of A in X generates a C_0-analytic semigroup of contractions.

For simplicity, we shall also assume that $u_0 \in X$. Then we introduce the stochastic convolution process $\{W_S(t),\ t \geq 0\}$

$$W_S(t) = \int_0^t e^{(t-s)A} B\,\mathrm{d}W_s$$

and we shall prove the following regularity for the stochastic convolution process.

Theorem 2.4. $W_S \in L^2_{\mathcal{F}}(\Omega; C([0,T];X))$.

Therefore we are in a position to rewrite Equation (2.5), letting $v(t) = u(t) - W_S(t)$, in the following form

$$v(t) = u_0 + \int_0^t Av(s)\,\mathrm{d}s + \int_0^t F(v(s) + z(s))\,\mathrm{d}s \tag{2.6}$$

where $z(t) = W_S(t) \in C([0,T];X)$ is a given trajectory of the stochastic convolution process. Equation (2.6) can be solved by using known results on dissipative systems (see for instance [9]); then one obtains the solution of the abstract version of our problem.

Theorem 2.5. *There exists a unique mild solution $u \in L^2_{\mathcal{F}}(\Omega; C([0,T];X))$ of (2.5) in the sense that the following holds:*

$$\mathbb{P}\left(\int_0^T \|F(u(s))\|_X^2 \,\mathrm{d}s < +\infty\right) = 1 \tag{2.7}$$

and

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))\,\mathrm{d}s + W_S(t) \tag{2.8}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$.

2.3. **The transfer functional and the Large Deviation Principle.** The general idea of LDP is to study the limiting behaviour of a family $\{\nu_\varepsilon,\ \varepsilon > 0\}$ of probability laws, describing the state of a noisy-perturbed system, weakly converging to a point mass δ_p, in terms of a *rate functional.* We aim to characterize how p is "typical" for the behaviour of the system; in other words, given an event Γ for which $p \notin \bar{\Gamma}$, the interest is to study the rate at which $\nu_\varepsilon(\Gamma) \to 0$. Large deviation means that we restrict our interest to events which are "very deviant", in the sense that $\nu_\varepsilon(\Gamma)$ goes to zero exponentially fast.

The standard definition thus reads as follows: $\{\nu_\varepsilon,\ \varepsilon > 0\}$ satisfies a LDP with rate functional J if, for all events Γ,

$$-\inf_{x \in \mathring{\Gamma}} J(x) \le \liminf_{\varepsilon \searrow 0} \varepsilon \log(\nu_\varepsilon(\Gamma)) \le \limsup_{\varepsilon \searrow 0} \varepsilon \log(\nu_\varepsilon(\Gamma)) \le -\inf_{x \in \bar{\Gamma}} J(x).$$

Let us consider the transfer functional $\Phi : C([0,T];X) \to C([0,T];X)$ that associates to every trajectory $z \in C([0,T];X)$ the solution v of the problem (2.6)

$$v(t) = u_0 + \int_0^t Av(s)\,\mathrm{d}s + \int_0^t F(v(s) + z(s))\,\mathrm{d}s.$$

Our aim is to show that this is a continuous mapping; that would play a central rôle in proving a LDP.

Theorem 2.6. *Under our assumptions, the transfer functional Φ is an homeomorphism of $C([0,T];X)$ into itself.*

Corollary 2.7. *Under the same assumptions, let $\Psi(u) = u + \Phi(u)$ for $u \in C([0,T];X)$. Then also Ψ is an homeomorphism of $C([0,T];X)$ into itself.*

$\Psi : C([0,T];X) \to C([0,T];X)$ is, obviously, the transfer functional related to problem (2.5), in the sense that Ψ associates to every trajectory of the stochastic convolution process W_S the corresponding trajectory of the solution u.

To obtain a LDP for (2.5), the main technique is the contraction principle, which requires the continuity of the transfer functional Ψ, compare [10]. Similar results in LDP for semilinear equations in a Banach space X were proved in [21, 12, 8]. We consider problem (2.5) with W replaced by $\sqrt{\varepsilon} W$

$$\begin{aligned} \mathrm{d}u_\varepsilon(t) &= [Au_\varepsilon(t) + F(u_\varepsilon(t))]\,\mathrm{d}t + \sqrt{\varepsilon} B\,\mathrm{d}W(t), \qquad t \in [0,T] \\ u(0) &= u_0. \end{aligned} \tag{2.9}$$

bringing up a family of solutions u_ε; we shall denote by ν_ε the law of u_ε on the space $C([0,T];X)$: then we obtain a LDP for such laws.

Theorem 2.8. *The family of laws ν_ε satisfies the large deviation principle with respect to the following explicit functional $J : C([0,T];X) \to [0,+\infty]$*

$$J(f) = \begin{cases} \dfrac{1}{2} \displaystyle\int_0^T \left| B^{-1}(\dot f(s) - A_X f(s)) \right|_H^2 \mathrm{d}s, & f \in \tilde R, \\ +\infty, & \textit{otherwise} \end{cases}$$

where $\tilde R$ is the subspace of $C([0,T];X)$ defined as

$$\tilde R = \Big\{ f \in C([0,T];X) \mid \exists\, g \in L^2(0,T;H) \ : \ f(t) = e^{tA_X} x + \int_0^t e^{(t-s)A_X} F(f(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A_X} Bg(s)\,\mathrm{d}s \Big\}.$$

The set $\tilde R$ has a natural interpretation in terms of the control problem associated to (2.5):

$$\dot f = (Af + F(f)) + Bg, \qquad f(0) = x.$$

$\tilde R$ contains all and only the trajectories that we can force the system to follow by applying a control g. It is easily seen that for every g there exists a unique $f = f^g$ that solves the control problem. Then, it is possible to define J in terms of g:

$$J(f^g) = \frac{1}{2} \int_0^T |g(t)|^2\,\mathrm{d}t;$$

this formula expresses the minimal energy that the forcing term shall give to the system in order to stay out of the path of the deterministic system. Moreover, it is possible to say that the probability that the system remains in a given subset of trajectories, in the limit for $\varepsilon \to 0$, depends only on the smooth trajectory with minimal L^2-norm.

For every $r_0 > 0$ we let $\tilde R(x, r_0)$ be the subset of $\tilde R$ given by all the functions f that solves the control process with initial condition x and a control g which satisfies $\frac{1}{2}\int_0^T |g(s)|^2\,\mathrm{d}s < r_0^2$. Let further u_0 be such that $\tilde R(u_0, r_0)$ is contained in a bounded subset of $C([0,T];X)$.

Assume now that u_ε is a solution to (2.9); then for every $\delta > 0$ and for every $\gamma > 0$ there exists an ε_0 such that for every $\varepsilon \in (0, \varepsilon_0)$ the following inequality holds

$$\mathbb{P}\left(d_{C([0,T];X)}(u_\varepsilon, \tilde R(u_0, r_0)) > \delta \right) \le \exp\left(-\frac{1}{\varepsilon}\left(r_0^2 - \gamma\right) \right).$$

Remark 2.9. *The quantity on the left is the probability that the distance between u_ε and the set of deterministic trajectories f^g exceeds δ; in biological applications, we are interested in the case when δ represents the* threshold level *for the action potential inside a neuronal network. Consider the potential $p_i(t)$ in the n vertices; then the right hand side of the latter equation estimates the probability that the electrical potential in (some) vertex reaches the treshold level even if this should not occour in the deterministic evolution; in other terms, this quantity estimates the presence of spurious impulses on the network during the time interval* $[0,T]$.

3. The stochastic Cauchy problem

This section contains all the technical tools and proofs necessary to solve the system of stochastic differential equations (2.1)– (2.4).

To start with, let us consider, for a continuous function $\bar u$ defined on the network, the boundary evaluation mapping

$$\Pi : D(\Pi) \subset (L^2(0,1))^m \to \mathbb{R}^n, \qquad \Pi\bar u = (u_{j_1}(\mathsf{v}_1), \dots, u_{j_n}(\mathsf{v}_n)),$$

where $j_k \in \Gamma(\mathsf{v}_k)$ (that is, the vertex v_k belongs to the edge e_{j_k}, compare Section 2.1).

Then we define on the space $H = (L^2(0,1))^m \times \mathbb{R}^n$ the matrix operator $A : D(A) \subset H \to H$ by setting

$$\begin{aligned} &D(A) = \{(\bar u, p) \in D(A_0) \times \mathbb{R}^n \mid \Pi\bar u = p\} \\ &A\ : u = \begin{pmatrix} \bar u \\ p \end{pmatrix} \ \mapsto\ \begin{pmatrix} A_0 & 0 \\ C & D \end{pmatrix} \begin{pmatrix} \bar u \\ p \end{pmatrix} \end{aligned} \tag{3.1}$$

where $A_0 :\ D(A_0) \subset (L^2(0,1))^m \to (L^2(0,1))^m$ is the second order differential operator defined by

$$\begin{aligned} &D(A_0) = \left\{ \bar u \in (H^2(0,1))^m \mid \bar u \text{ satisfy the continuity condition (2.3) } \right\}; \\ &A_0 \bar u = (\partial_x c_j(x) \partial_x u_j)_{j=1,\dots,m} \end{aligned} \tag{3.2}$$

notice that A_0 is not the generator of a semigroup on $(L^2(0,1))^m$, but this holds if we restrict further the domain and consider the operator $A_D :\ D(A_D) \subset (L^2(0,1))^m \to (L^2(0,1))^m$ with Dirichlet boundary conditions

$$\begin{aligned} &D(A_D) = \left\{ \bar u \in (H^2(0,1))^m \mid \Pi\bar u = 0 \right\}; \\ &A_D \bar u = (\partial_x c_j(x) \partial_x u_j)_{j=1,\dots,m}. \end{aligned} \tag{3.3}$$

D is the linear operator on $\mathbb{R}^n$ associated to the matrix

$$D = \operatorname{diag}(-d_1, \dots, -d_n) \tag{3.4}$$

while $C :\ D(C) \subset (L^2(0,1))^m \to (L^2(0,1))^m$ is the feedback operator given by

$$\begin{aligned} &D(C) = (H^1(0,1))^m. \\ &C\bar u = \left(\sum_{j \in \Gamma(\mathsf{v}_i)} \phi_{ij} \mu_j c_j(\mathsf{v}_i) \partial_x \bar u_j(\mathsf{v}_i) \right)_{i=1,\dots,n} \end{aligned} \tag{3.5}$$

The functions $f_j(u)$ which appear in (2.1) are assumed to have a polynomial growth at infinity. We remark that the classical FitzHugh-Nagumo problem requires

$$f_j(u) = u(u-1)(a_j - u) \qquad j = 1, ..., m$$

for some $a_j \in (0,1)$, and satisfies Hypothesis 2.1.2 with

$$\eta \le -\max_j \frac{(a_j^3+1)}{3(a_j+1)}, \qquad s = 3.$$

We set

$$F(u) = \begin{pmatrix} -\left(f_j(u_j)\right)_{j=1,\dots,m} \\ 0 \end{pmatrix} \qquad \text{for } u = \begin{pmatrix} (u_j)_{j=1,\dots,m} \\ p \end{pmatrix} \tag{3.6}$$

and we write our problem in an abstract form

$$\begin{aligned} \mathrm{d}u(t) &= [Au(t) - F(u(t))]\,\mathrm{d}t + B\,\mathrm{d}W(t) \\ u(0) &= u_0 \end{aligned} \tag{3.7}$$

where B is the operator-valued matrix defined by

$$B = \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \operatorname{diag}(b_1,\dots,b_n) \end{pmatrix},$$

and with abuse of notation $W(t)$ is the natural embedding in H of the n-dimensional Wiener process $W(t)$.

3.1. **The operator** A. In this section we study the linear operator A. Our first aim is to prove the following result.

Proposition 3.1. *A generates an analytic semigroup of contractions on* $H = (L^2(0,1))^m \times \mathbb{R}^n$.

Our proof is based on an application of sesquilinear form theory. As a general reference we consider the monograph [18]; for an application to a similar problem, see also [3, 16].

As a consequence of our construction, we obtain that the semigroup $\{e^{tA},\ t \ge 0\}$ generated by A is exponentially bounded, with growth bound given by the strictly negative spectral bound of the operator A. By these results, we conclude the proof of Proposition 2.2.

Sketch of the proof. Let us consider the following form:

$$a(u,v) = \sum_{j=1}^m \int_0^1 (c_j \bar{u}_j' \bar{v}_j')\mu_j \,\mathrm{d}x + \sum_{i=1}^n d_i u(\mathsf{v}_i) v(\mathsf{v}_i) \tag{3.8}$$

on the domain

$$V = \left\{ u = (\bar{u}, \alpha) \in (H^1(0,1))^m \times \mathbb{R}^n \right\} \tag{3.9}$$

and let we introduce the norm

$$\|u\|_a = \sqrt{a(u,u) + \|u\|_V^2}, \quad u \in V.$$

Here $\|\cdot\|_V$ denotes the norm induced by H on V so that V endowed wtih $\|\cdot\|_V$ becomes an Hilbert space. We recall that a form $a : V \times V \to \mathbb{R}$ is continuous if it satisfies

$$\|a(u,v)\| \le M \|u\|_a \|v\|_a$$

for some positive constant M. We can associate to a form a an unbounded operator $\mathbf{A}$, defined on a linear subspace $D(\mathbf{A})$ of H; the domain $D(\mathbf{A})$ is given by the elements $u \in D(a)$ such that there exists $v \in H$ with $a(u,\phi) = \langle v, \phi \rangle$ for every $\phi \in D(a)$; then we set $\mathbf{A}u = v$.

The thesis follows by proving that a is densely defined, closed, positive, symmetric and continuous and that the operator associated to the form a given in (3.8) is $(A, D(A))$, i.e. the operator defined by means of (3.2)–(3.5). According to [18], this implies that the operator A is dissipative, self-adjoint and it generates a contraction semigroup e^{tA} on H.

□

Let us consider the space $X = (C([0,1])^m \times \mathbb{R}^n$ and the part of A in X; our aim is to prove that also this operator A_X behaves well. As opposite to the Hilbert case of the previous proposition, which resembles standard results available in the literature, this result needs some more care in the proof.

Proposition 3.2. *The operator A_X generates a bounded analytic semigroup.*

Proof. The proof is based on the techniques of decoupling the domain of the operator matrix A first introduced by J.K. Engel in [11].

As a necessary tool we introduce the Dirichlet mapping $D^{A_0,\Pi}$ associated to the operator A_0, i.e., $D^{A_0,\Pi}$ maps every vector $p \in \mathbb{R}^n$ into the (unique) solution of the abstract Dirichlet problem

$$\begin{cases} A_0 \bar{u} = 0 \\ \Pi \bar{u} = p. \end{cases}$$

Then we consider the operator matrix

$$N = \begin{pmatrix} I & -D^{A_0,\Pi} \\ 0 & I \end{pmatrix}$$

which is an isomorphism on X. With some computation one obtains that A is similar to the operator matrix

$$\tilde{A} = \begin{pmatrix} A_D - D^{A_0,\Pi} C & -D^{A_0,\Pi}(D + CD^{A_0,\Pi}) \\ C & D + CD^{A_0,\Pi} \end{pmatrix}$$

with diagonal domain $D(\tilde{A}) = D(A_D) \times \mathbb{R}^n$ where A_D is the operator introduced in (3.3). The similarity transformation is expressed by the operator matrix N.

Since $\tilde{A}$ can be seen as the sum of a diagonal operator matrix which generates an analytic semigroup and a relatively bounded perturbation (with relative bound 0)

$$\tilde{A} = \begin{pmatrix} A_D & 0 \\ 0 & D + CD^{A_0,\Pi} \end{pmatrix} + \begin{pmatrix} -D^{A_0,\Pi} C & -D^{A_0,\Pi}(D + CD^{A_0,\Pi}) \\ C & 0 \end{pmatrix},$$

we conclude that $\tilde{A}$ is the generator of an analytic semigroup of angle $\frac{\pi}{2}$ by [1, Theorem 3.7.23].

Now let us denote by $(A_\infty, D(A_\infty))$ the part of A in $(L^\infty(0,1))^m \times \mathbb{R}^n$. Clearly $D(A_X) \subset D(A_\infty)$ and $A_X = A_\infty$ on $D(A_\infty)$ so that $\sigma(A_X) \subset \sigma(A)$. Moreover, it has been proved in [16, Proposition 6.4] that the spectrum of $D(A_\infty)$ coincides with the spectrum of A; thus $\sigma(A_X) \subset \sigma(A)$. It follows that A_X is invertible and the semigroup is bounded. This completes the proof.

□

3.2. **The stochastic convolution process.** This section is devoted to prove a result concerning the spatial regularity of the stochastic convolution process $W_S(t)$, that, in our setting, can be written as

$$W_S(t) = \int_0^t e^{(t-s)A} B \, \mathrm{d}W_s.$$

Let us recall that $X := (C[0,1])^m \times \mathbb{R}^n$; our next result shows that W_S belongs to the space $L^2_{\mathcal{F}}(\Omega; C([0,T]; X))$ of X-valued, adapted continuous processes Y on the time interval $[0,T]$ such that

$$\mathbb{E} \sup_{t\in[0,T]} |Y(t)|_X^2 < \infty.$$

Lemma 3.3. *For all $t \geq 0$, the stochastic convolution $\{W_S(t),\ t \in [0,T]\}$ belongs to the space $L^2_{\mathfrak{F}}(\Omega; C([0,T];X))$.*

Proof. Let us recall that the linear operator A with domain

$$D(A) = \left\{ \begin{pmatrix} \bar{u} \\ p \end{pmatrix} \in (H^2(0,1))^m \times \mathbb{R}^n \ : \ \Pi\bar{u} = p \right\}$$

generates a C_0-analytic semigroup of contractions on H. Then we introduce the interpolation spaces $H_\theta = (H, D(A))_{\theta,2}$ for $\theta \in (0,1)$. By classical interpolation theory (see e.g. [15]) it results that, for $\theta < 1/4$, $H_\theta = (H^{2\theta}(0,1))^m \times \mathbb{R}^n$ while for $\theta > 1/4$ the definition of H_θ involves boundary conditions, that is

$$H_\theta = \left\{ \begin{pmatrix} \bar{u} \\ p \end{pmatrix} \in (H^{2\theta}(0,1))^m \times \mathbb{R}^n \ : \ \Pi\bar{u} = p \right\}.$$

Therefore, one has $(0,x) \in H_\theta$ for any $x \in \mathbb{R}^n$ and any $\theta < 1/4$. Furthermore, for $\theta > 1/2$, one also has $H_\theta \subset (C[0,1])^m \times \mathbb{R}^n$ by Sobolev's embedding theorem. Moreover, for all $u \in H_\theta$ and $\theta + \gamma \in (0,1)$, one has

$$|e^{tA}u|_{\theta+\gamma} \leq t^{-\gamma}|u|_\theta e^{\omega_A t},$$

where ω_A is the spectral bound of the operator A.

Then the result follows from an application of the factorization lemma of Da Prato and Zabczyk, compare for instance [9, Proposition A.1.1], once we notice that

$$\int_0^t s^{-2\gamma}\|e^{sA}B\|_{HS}^2 \,\mathrm{d}s$$

is finite for every $\gamma < 1/2$.

□

We now consider a large deviation principle for the solution of the linear equation with additive noise

$$\begin{aligned} \mathrm{d}z_\varepsilon(t) &= Az_\varepsilon(t)\,\mathrm{d}t + \sqrt{\varepsilon}B\,\mathrm{d}W_t, \qquad t \in [0,T], \\ z_\varepsilon(0) &= 0. \end{aligned} \tag{3.10}$$

It is known that the solution of this problem is given by

$$z_\varepsilon(t) = \sqrt{\varepsilon}W_S(t) \tag{3.11}$$

and we are interested in the properties of the law μ_ε of $z_\varepsilon(\cdot)$ in the space $C([0,T];X)$. The following result is proved in [19].

Proposition 3.4. *$\{\mu_\varepsilon\}_{\varepsilon>0}$ fulfills the LDP with respect to the rate functional I defined by*

$$I(f) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \|B^{-1}(\dot{f}(s) - A_X f(s))\|_H^2 \,\mathrm{d}s, & f \in R, \\ +\infty, & \textit{otherwise} \end{cases}$$

where R is the subspace of $C([0,T];X)$ defined as

$$R = \left\{ f \in C([0,T];X) \mid \exists\, g \in L^2(0,T;H) \ : \ f(t) = \int_0^t e^{(t-s)A_X} Bg(s)\,\mathrm{d}s \right\}.$$

3.3. **The semilinear equation.** We are concerned with the abstract stochastic differential equation (2.5); we first verify that the nonlinear term F satisfies suitable regularity properties.

Proposition 3.5. *$X \subset D(F)$ and the part F_X of F in X:*

$$\begin{aligned} F_X(x) &= F(x), \qquad x \in D(F_X) \\ D(F_X) &= \{x \in D(F) \cap X \ : \ F(x) \in X\} \end{aligned}$$

is quasi dissipative and uniformly continuous on bounded subsets of X.

Proof. Fix $\delta > 0$ and $R > 0$; let $u, v \in X$ with $\|u\|_X \le R$, $\|v\|_X \le R$: we aim to prove that there exists $\epsilon > 0$ such that $\|u - v\|_X \le \epsilon$ implies $\|F(u) - F(v)\|_X \le \delta$. Now, notice that the norm $\|F(u) - F(v)\|_X$ is bounded by the maximum of the norms $\|f_j(u_j(\cdot)) - f_j(v_j(\cdot))\|_\infty$; hence it is sufficient to notice that $f_j : \mathbb{R} \to \mathbb{R}$ is uniformly continuous on bounded subsets and that the bounds on u and v holds also for u_j and v_j: this concludes the first part of the proof.

Further the dissipativity of F is a well known result in the theory of dissipative mappings, compare [7].

□

From the theory of dissipative mappings (see for instance [8, Appendix D]) we recall the definition of the Yosida approximations F_α, $\alpha > 0$, for F:

$$\begin{aligned} F_\alpha(u) &= F(J_\alpha(u)) = \frac{1}{\alpha}(J_\alpha(u) - u), \quad \text{where} \\ J_\alpha(u) &= (I - \alpha F)^{-1}(u), \qquad u \in X. \end{aligned}$$

In this section we fix z in $C([0,T];X)$ and consider the following integral equation:

$$f(t) = e^{tA_X} x + z(t) + \int_0^t e^{(t-s)A_X} F(f(s))\,\mathrm{d}s, \qquad x \in X. \tag{3.12}$$

It is possible to prove that Eq. (3.12) has a unique solution $f \in C([0,T];X)$ for all $z \in C([0,T];X)$, compare [8, Theorem 7.13].

Theorem 3.6. *In our assumptions, let z be a function of $C([0,T];X)$ and x an element of X. Then Eq. (3.12) has a unique solution $f \in C([0,T];X)$.*

For the sake of completeness, we give below a sketch of the proof.

Proof. Setting $v(t) = f(t) - z(t)$ we are concerned with the equation

$$v(t) = e^{tA_X} x + \int_0^t e^{(t-s)A_X} F(v(s) + z(s))\,\mathrm{d}s, \qquad x \in X. \tag{3.13}$$

We introduce the family of approximating equations defined in terms of the Yosida approximations F_α, for any $\alpha > 0$, and we notice that

$$v_\alpha(t) = e^{tA_X} x + \int_0^t e^{(t-s)A_X} F_\alpha(v_\alpha(s) + z(s))\,\mathrm{d}s, \qquad x \in X \tag{3.14}$$

has a unique solution $v_\alpha \in C([0,T];X)$ for every z in the same space. Moreover, since $\|e^{tA_X}\|_{L(X)} \le 1$ and the Yosida approximations F_α are bounded by F, it holds that

$$\|v_\alpha(t)\|_X \le \|x\|_X + \int_0^t \|F(z(s))\|_X\,\mathrm{d}s. \tag{3.15}$$

Finally, when we consider the difference $v_\alpha - v_\beta$ we get the estimate

$$\frac{d}{dt}\|v_\alpha(t) - v_\beta(t)\|_X \le \|F(J_\alpha(v_\alpha(t))) - F(v_\alpha(t))\|_X + \|F(J_\beta(v_\beta(t))) - F(v_\beta(t))\|_X.$$

Since $\|v_\alpha(t)\|_X$ is bounded by (3.15), F is bounded on bounded subsets and $\|J_\alpha(v_\alpha) - v_\alpha\| \le \alpha\|F(v_\alpha)\|$, by letting $\alpha, \beta \to 0$ it is possible to prove the convergence of the sequence $\{v_\alpha\}$ in $C([0,T];X)$ to a function v that solves Eq. (3.13).

□

3.4. **The transfer functional.** In this section we consider the functional $\Phi : C([0,T];X) \to C([0,T];X)$ that associates to every forcing term z the solution of Eq. (3.12). From the existence theorem 3.6 we notice that Φ is a bijection from $C([0,T];X)$ into itself. We can go further and prove the following result.

Theorem 3.7. *Φ is an homeomorphism on $C([0,T];X)$.*

Proof. Our argument is divided in two steps. In the first part we prove the continuity of Φ, then in the second we shall be concerned with Φ^{-1}.

Suppose first that the non linear term F is locally Lipschitz on X. To show continuity of Φ, fix a point z_1 of $C_0([0,T];X)$, and a bounded subset B around z_1. Then, there exists a bounded Borel subset $C \subset X$ such that $z(t) \in C$ for any $z \in B$ and $t \in [0,T]$. Since F is locally Lipschitz, we can suppose, with no loss of generality, that F is Lipschitz on $C \subset X$, with Lipschitz constant equal to Λ.

Let $z_2 \in B$ and denote by v_1 and v_2 the solutions $\Phi(z_1)$ and $\Phi(z_2)$ respectively. Then

$$v_1(t) - v_2(t) = z_1(t) - z_2(t) + \int_0^t e^{(t-s)A_X}[F(v_1(s)) - F(v_2(s))]\,\mathrm{d}s$$

therefore, using the estimate on e^{tA_X} proved in Proposition 3.2 we get

$$\|v_1(t) - v_2(t)\|_X \le \|z_1(t) - z_2(t)\|_X + C_{\Lambda,A_X}\int_0^t \|v_1(s) - v_2(s)\|_X\,\mathrm{d}s;$$

now the thesis follows by applying Gronwall's lemma.

We now proceed with the general case. We shall prove that given $z_1 \in C([0,T];X)$ and $\delta > 0$, then for every $z_2 \in C([0,T];X)$ with $\|z_1 - z_2\|_X \le \delta$ it holds $\|v_1 - v_2\|_X \le \epsilon$ for a constant ϵ which depends only on δ and z_1.

Let v_i, $i = 1,2$ be the solution of (3.13) and $v_{i,\alpha}$ the solution of the approximating problem (3.14) with forcing term z_i; notice that, since z_1 is given, there exists M large enough with $\sup\|z_i(t)\|_X \le M$, for $i = 1,2$. Now, we can write

$$\|v_1(t) - v_2(t)\| \le \|v_1(t) - v_{1,\alpha}(t)\| + \|v_{1,\alpha}(t) - v_{2,\alpha}(t)\| + \|v_2(t) - v_{2,\alpha}(t)\|. \tag{3.16}$$

We shall begin with the first term, and the last one is treated similarly. We can estimate the distance $\|v_1(t) - v_{1,\alpha}(t)\|$ in the following way

$$\frac{d}{dt}\|v_1(t) - v_{1,\alpha}(t)\| \le \omega_A\|v_1(t) - v_{1,\alpha}(t)\| + \|F(J_\alpha(v_1(t))) - F(v_1(t))\|$$

which leads, using the modulus of continuity* $K_F(\alpha)$ of F on $B(0,\alpha M)$, to the estimate

$$\|v_1(t) - v_{1,\alpha}(t)\| \le C_{\omega_A,T} K_F(\alpha).$$

Therefore, for α small enough, this term is bounded by $\epsilon/3$.

It remains to treat the second term in the right-hand side of (3.16), i.e., $\|v_{1,\alpha}(t) - v_{2,\alpha}(t)\|$. However, since F_α is a Lipschitz continuous mapping, it follows from the first part of the proof that it is possible to choose δ small enough such that also this term is bounded by $\epsilon/3$. Collecting the above estimates we obtain the thesis.

We now proceed by proving that Φ^{-1} is continuous. Let f_1, f_2 belong to $C([0,T];X)$ with $\|f_1 - f_2\| \le \delta$ and define $z_i = \Phi^{-1}(f_i)$; then we estimate

$$\begin{aligned}
\|z_1 - z_2\|_X &\le \|f_1 - f_2\|_X + \left\| \int_0^t e^{(t-s)A_X} [F(f_1(s)) - F(f_2(s))] \, \mathrm{d}s \right\|_X \\
&\le \|f_1 - f_2\|_X + M \int_0^t \|F(f_1(s)) - F(f_2(s))\|_X \, \mathrm{d}s \\
&\le \delta + MTK_F(\delta) =: \epsilon
\end{aligned}$$

and the last term of the previous inequality tends to zero as $\delta \to 0$.

□

3.5. **Large deviation principle.** We are ready to prove the LDP for the solutions of the family of semilinear equations

$$\begin{aligned}
\mathrm{d}u_\varepsilon(t) &= [Au_\varepsilon(t) + F(u_\varepsilon(t))] \, \mathrm{d}t + \sqrt{\varepsilon} B \, \mathrm{d}W(t), \qquad t \in [0,T] \\
u(0) &= u_0,
\end{aligned} \tag{3.17}$$

as stated in Theorem 2.8. We shall denote by ν_ε the law of $u_\varepsilon(\cdot)$ on $C([0,T];X)$; then, by the definition of the transfer functional Φ, the law ν_ε is given by $\nu_\varepsilon = \Phi \circ \mu_\varepsilon$, where the measure μ_ε is the law of the Ornstein-Uhlenbeck process z_ε defined in (3.11). Recall that the family $\{\mu_\varepsilon\}$ satisfies a LDP with rate functional $I(x)$ as stated in Proposition 3.4. It is known, compare [8, Proposition 12.3], that for a continuous Φ, the family ν_ε satisfy a LDP with rate function $J(f) = I(\Phi^{-1}(f))$, so the thesis follows from Theorem 3.7.

□

References

1. W. Arendt and C. J. K. Batty and M. Hieber and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser: Basel, 2001.
2. J. von Below and S. Nicaise, *Dynamical interface transition in ramified media with diffusion*, Comm. Partial Differential Equations **21** (1996), no. 1-2, 255–279.
3. S. Bonaccorsi and F. Confortola and E. Mastrogiacomo, *Optimal control of stochastic differential equations with dynamical boundary conditions*, to appear on J. Math. Analysis and Applications, 2008. Preprint available at http://eprints.biblio.unitn.it/archive/00001228/01/UTM709.pdf
4. S. Cardanobile and D. Mugnolo, *Analysis of a FitzHugh-Nagumo-Rall model of a neuronal network*, Math. Methods Appl. Sci. **30** (2007), no. 18, 2281–2308.

*A modulus of continuity of a function F on a space Ξ is a function $K : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{r \searrow 0} K(r) = 0$ and $\|F(x) - F(y)\| \le K(|x-y|)$ for any $x, y \in D(F)$

5. E. A. Coayla-Teran and P. M. Dias de Magalhes, *Weak solutions for stochastic FitzHugh-Nagumo equations*, Stochastic Anal. Appl. **21** (2003), no. 2, 443–463.
6. ———, *Stochastic FitzHugh-Nagumo equations in a time dependent domain*, Random Oper. Stoch. Equ. **15** (2007), no. 1, 49–64.
7. G. Da Prato, *Applications croissantes et équations d'évolution dans les espaces de Banach*, Academic Press, London-New York, 1976.
8. G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge UP, 1992.
9. ———, *Ergodicity for infinite-dimensional systems*, Cambridge UP, 1996.
10. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer, New York, 1998.
11. K.-J. Engel, *Spectral theory and generator property for one-sided coupled operator matrices*, Semigroup Forum **58** (1999), no. 2, 267–295.
12. M. Fantozzi, *Large deviations for semilinear differential stochastic equations with dissipative nonlinearities*, Stochastic Anal. Appl. **21** (2003), no. 1, 127–139.
13. A.L. Hodgkin and A.F. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, J. Physiol. **117** (1952), no. 2, 500–544.
14. J. Keener and J. Sneyd, *Mathematical physiology*, Springer, New York, 1998.
15. A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Progress in Nonlinear Differential Equations and their Applications, 16, Birkhäuser Verlag, Basel, 1995.
16. D. Mugnolo and S. Romanelli, *Dynamic and generalized Wentzell node conditions for network equations*, Math. Methods Appl. Sci. **30** (2007), no. 6, 681–706.
17. J. D. Murray, *Mathematical Biology: I. An Introduction*, Springer-Verlag, 3rd ed., 2002.
18. E. M. Ouhabaz, *Analysis of heat equations on domains*, Princeton University Press, Princeton, NJ, 2005.
19. S. Peszat, *Large deviation principle for stochastic evolution equations*, Probab. Theory Related Fields **98** (1994), no. 1, 113–136.
20. D. L. Ringach and B. J. Malone, *The Operating Point of the Cortex: Neurons as Large Deviation Detectors*, J. Neurosci. **27** (2007), no. 29, 7673–7683.
21. W. Smoleński and R. Sztencel and J. Zabczyk, *Large deviation estimates for semilinear stochastic equations*, in *Proceedings IFIP Confereence on Stochastic Differential Systems*, no. 98 in LNiCIS, pages 218–231. Eisenach, 1986.