# Optimal control of stochastic FitzHugh-Nagumo equation

Viorel Barbu[1] Francesco Cordoni[2] and Luca Di Persio[3]
[1] A.I. Cuza University, Iasi
Carol I Blvd, Nr.11, 700506 Iasi, Romania
e-mail: vb41@uaic.ro
[2] University of Trento - Department of Mathematics
Via Sommarive, 14 Trento, 38123-ITALY
e-mail: francesco.cordoni@unitn.it
[3] University of Verona - Department of Computer Science
Strada le Grazie, 15 Verona, 37134-ITALY
e-mail: luca.dipersio@univr.it , corresponding author

**Abstract**

This paper is concerned with existence and uniqueness of solution for the the optimal control problem governed by the stochastic FitzHugh-Nagumo equation driven by a Gaussian noise. First order conditions of optimality are also obtained.

## 1 Introduction

Consider here the reaction-diffusion equation

$$\begin{cases} dX(t,\xi) - \Delta_\xi X(t,\xi)dt + f(X(t,\xi))dt = \sqrt{Q}dW(t) + F(t,\xi)\,dt\,(t,\xi) \in [0,T]\times\mathcal{O}\,, \\ X(t,\xi)|_{\partial\mathcal{O}} = 0, \quad t\in[0,T]\,, \\ X(0,\xi) = x(\xi), \quad \xi\in\mathcal{O},\ \ x\in L^2(\mathcal{O}) \end{cases}, \tag{1.1}$$

in a probability space $(\Omega,\mathcal{F},\mathbb{P})$, where $f(u) = u(u-a)(u-b)$, $\forall\, u \in \mathbb{R}$, $\mathcal{O}\subset\mathbb{R}^d$, $d=1,2,3$ is a bounded and open set with smooth boundary $\partial\mathcal{O}$, $W(t)$ is a cylindrical Wiener process and $Q\in\mathcal{L}(L^2(\mathcal{O}),L^2(\mathcal{O}))$ (the space of linear and continuous operator from $L^2(\mathcal{O})$ into itself equipped with the operator norm) is a self-adjoint positive operator with $TrQ < \infty$. Here $a$, $b \in L^\infty([0,T] \times \mathcal{O})$ and $x \in L^2(\mathcal{O})$ are given. Also $F \in L^2([0,T] \times \mathcal{O})$. We shall denote by $(\mathcal{F}_t)_{t\geq 0}$ the natural filtration induced by $W(t)$. Equation (1.1) can be rewritten as follows

$$\begin{cases} dX(t) + AX(t)dt + f(X(t))dt = \sqrt{Q}dW(t) + F(t)\,dt, \quad t \in [0,T]\,, \\ X(0) = x, \quad x \in L^2(\mathcal{O}) \end{cases}, \tag{1.2}$$

$A$ being the Laplace operator $-\Delta_\xi$ with domain $D(A) := H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$. In the special case $a, b \in \mathbb{R}$, (1.1) is the dimensionless form of the celebrated *FitzHugh-Nagumo* equation, see, e.g., [1] and reference therein, perturbed by a coloured Gaussian noise $\sqrt{Q}\dot{W}$. Its deterministic counterpart has been introduced by FitzHugh (1922–2007) and Nagumo, see [18, 20] in order to model the conduction of electrical impulses in a nerve axon. In particular $X$ is the nerve membrane potential and $F := -V + I$ where $V$ is the ion concentration and $I$ is the applied current. The Gaussian perturbation is the effect of random input currents in neurons and their source is the random opening or closing of ion channels, see, e.g. [23]. In 2-D and 3-D equation (1.1) is relevant in statistical mechanics where it is called *Ginzburg-Landau* equation and also in phase transition models of *Ginzburg-Landau* type, see, e.g. [15]. We would like to underline that nonlinear potential of the form $f(u) = u(u-a)(u-b)$ arising here are specific for diffusion processes in excitable media or for phase transition.

In what follows we will study the optimal control problem for (1.1) providing an existence and uniqueness result as well as the first order necessary conditions for optimality, namely the maximum principle. In Sec. 2 we shall prove the well-posedness of problem (1.1), see [9] for other results of this type.

The existence of a solution to optimal control problem (P) will be proved under suitable conditions on time interval $[0,T]$ and the cost functional in Sec. 2. It should be mentioned that there exists a large literature concerning the optimal control problems governed by the deterministic FitzHugh-Nagumo equation, see, e.g., [10, 19], while to the best of our knowledge, the stochastic case that we are interested in, lacks of such results. The motivation is that existence of an optimal control for the stochastic problem we consider here is quite a delicate problem which cannot be solved with standard optimization arguments which require the weak lower semicontinuity of cost functional in the control basic space and a more subtle argument based on

Eckelands's variational principle was used. The existence result we obtain here is the main novelty of this work. To prove the existence of an optimal control an essential property of nonlinear function $f$ is that it is ultimately monotonically increasing, that is outside a bounded interval.

We shall use the basic notions and standard notation $L^p(\mathcal{O})$, $1 \le p \le \infty$ and $H^k(\mathcal{O})$, $k = 1, 2$, $H^1_0(\mathcal{O})$ for spaces of Lebesgue $p-$integrable functions on $\mathcal{O}$ and respectively, Sobolev spaces on $\mathcal{O}$. The norm in $L^p(\mathcal{O})$ will be denoted by $|\cdot|_p = \|\cdot\|_{L^p(\mathcal{O})}$ and the scalar product in $L^2(\mathcal{O})$ by $\langle \cdot, \cdot \rangle_2$. Given a Banach space $Y$ we shall denote by $|\cdot|_Y$ its norm. By $C([0,T];Y)$ we denote the space of $Y-$valued continuous functions on $[0,T]$ and by $L^p([0,T];Y)$ the space of $p-$integrable $Y-$valued functions on $[0,T]$. By $W^{1,p}([0,T];Y)$, $1 \le p \le \infty$ we shall denote the space of absolutely continuous functions $u : [0,T] \to Y$ such that $\frac{du}{dt} \in L^p([0,T];Y)$.

We shall use the standard notations, see, e.g. [11], for spaces of processes defined in probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\ge 0}, W\right)$. $C_W([0,T];L^2(\mathcal{O}))$ is the space of all $L^2(\mathcal{O})-$valued $(\mathcal{F}_t)_{t\ge 0}-$adapted process such that $u \in C([0,T];L^2(\Omega, L^2(\mathcal{O})))$. Similarly, $L^2_W([0,T];H^1_0(\mathcal{O}))$ is the space of all $(\mathcal{F}_t)_{t\ge 0}-$adapted processes $u \in L^2([0,T];L^2(\Omega, H^1_0(\mathcal{O})))$.

We denote by $W_A$ the stochastic convolution defined by

$$W_A(t) := \int_0^t e^{-(t-s)A}\sqrt{Q}dW(s), \quad \forall\, t \ge 0\,.$$

In the following we shall assume that

$$\mathbb{E}\left[\sup_{(t,\xi)\in[0,T]\times\mathcal{O}} |W_A(t,\xi)|^{2m}\right] < \infty\,, m \in [1,2]\,. \tag{1.3}$$

Sufficient conditions for (1.3) to hold are given in [11, Th.2.13]. We refer to [7] for standard results on convex analysis which will be used in the following.

## 2 Existence for equation (1.1)

**Definition 2.0.1.** We say that the function $X \in C_W([0,T];L^2(\mathcal{O}))$ is called a *mild solution* to (1.1) if $X(t) : [0,T] \to L^2(\mathcal{O})$ is continuous $\mathbb{P}-$a.s., $\forall\, t \in [0,T]$ and it satisfies the stochastic integral equation

$$X(t) = e^{-At}x - \int_0^t e^{-(t-s)A}\left(f(X(s)) - F(s)\right)ds + W_A(t), \quad \forall\, t \in [0,T]\,.$$

**Theorem 2.1.** *Assume that assumption* (1.3) *holds and that* $x \in H_0^1(\mathcal{O})$. *Then there exists a unique solution* $X$ *to* (1.1) *which satisfies*

$$X \in C_W\left([0,T]; H_0^1(\mathcal{O})\right) \cap L^2_W\left([0,T]; H^2(\mathcal{O})\right) \cap L^2\left(\Omega; C\left([0.T]; H_0^1(\mathcal{O})\right)\right) .$$

We note in particular that assumption (1.3) holds if $Q = A^{-\frac{\gamma}{2}}$, with $\gamma > \frac{d}{2} - 1$, see, e.g., [11, Prop.4.3].

If we define the stochastic process $y := X - W_A$, then equation (1.1) reduces to the random parabolic equation

$$\begin{cases} y_t(t,\xi) - \Delta_\xi y(t,\xi) + y^3(t,\xi) + f_1(t,\xi)y^2(t,\xi) + f_2(t,\xi)y(t,\xi) = f_3(t,\xi) \text{ in } [0,T]\times\mathcal{O}\,, \\ y(t,\xi) = 0 \text{ in } [0,T]\times\partial\mathcal{O}\,, \\ y(0,\xi) = y_0(\xi)\,, \xi\in\mathcal{O} \end{cases} ,\tag{2.1}$$

where $f_1,\ f_2 \in L^\infty([0,T]\times\mathcal{O})$, $f_3 \in L^2([0,T]\times\mathcal{O})$ are $(\mathcal{F}_t)_{t\geq 0}-$adapted $L^2(\mathcal{O})-$valued processes on $[0,T]$. More precisely $f_1 = a + 3W_A$, $f_2 = b + 3W_A^2 + 2W_A$, $f_3 = -W_A^3 - 9W_A^2 - bW_A + F$.

The following proposition states an existence and uniqueness result for equation (2.1)

**Proposition 2.2.** *Assume* $x \in H_0^1(\mathcal{O})$. *Then there is a unique solutions to equation* (2.1) *satisfying* $\mathbb{P}-a.s.$

$$y \in C\left([0,T]; H_0^1(\mathcal{O})\right) \cap L^2\left([0,T]; H^2(\mathcal{O})\right) . \tag{2.2}$$

*Moreover the process* $t \mapsto y(t)$ *is* $(\mathcal{F}_t)_{t\geq 0}-$*adapted.*

*Proof.* Let us consider, for fixed $\omega\in\Omega$, the set

$$K = \left\{ y \in C\left([0,T^*]; L^2(\mathcal{O})\right) \ : \ \|y\|_{L^\infty\left([0,T^*];H_0^1(\mathcal{O})\right)} \leq R, \quad 0 \leq T^* \leq T \right\} ,\tag{2.3}$$

where $R$ is a positive real constant and $T^*$ has to be chosen later on.

The set $K$ is closed in $C\left([0,T^*]; L^2(\mathcal{O})\right)$ and therefore it is a complete metric space when equipped with the metric

$$\rho(y,v) = \sup_{t\in[0,T^*]} |y(t) - v(t)|_2. \tag{2.4}$$

Let $z \in K$ and let us consider the operation $F : K \to K$ defined by $Fz = y$, where $y$ is solution to

$$\begin{cases} y_t(t,\xi) - \Delta_\xi y(t,\xi) + y^3(t,\xi) = -f_1(t,\xi)z^2(t,\xi) - f_2(t,\xi)z(t,\xi) + f_3(t,\xi) \text{ in } [0,T]\times\mathcal{O}\,, \\ y(0,\xi) = x(\xi) \text{ in } \mathcal{O}\,, \\ y(t,\xi) = 0\,, \quad (t,\xi)\in[0,T]\times\partial\mathcal{O} \end{cases} .\tag{2.5}$$

By standard existence and uniqueness results, see, e.g., [4], problem (2.5) has a unique solution

$$y \in C\left([0,T^*]; H_0^1(\mathcal{O})\right) \cap L^2\left([0,T^*]; H^2(\mathcal{O})\right) , \quad \mathbb{P}-a.s. ,$$
$$y_t \in L^2\left([0,T]; L^2(\mathcal{O})\right) , \quad \mathbb{P}-a.s. ,$$

and by the Sobolev embedding theorem the following estimate holds

$$\begin{aligned} &\|y(t)\|^2_{H_0^1(\mathcal{O})} + \int_0^t |\Delta_\xi y(s)|_2^2 ds + \int_0^t |y(s)|_6^6 ds \le \\ &\le C_1 \left( \int_0^t \int_{\mathcal{O}} \left( f_1^2 z^4 + f_2^2 z^2 + f_3^2 \right) d\xi\, ds + \|x\|^2_{H_0^1(\mathcal{O})} \right) . \end{aligned} \tag{2.6}$$

By multiplying (2.5) by $y$, respectively $\Delta y$, and integrating on $(0,t)\times\mathcal{O}$ it also follows that

$$\begin{aligned} &\|y\|^2_{C\left([0,T^*];H_0^1(\mathcal{O})\right)} + \int_0^t \left( |y(s)|_6^6 + |y(s)|^2_{H^2(\mathcal{O})} \right) ds \le \\ &\le C_2 \int_0^t \int_{\mathcal{O}} \left( |z|^4 + |z|^2 + 1 \right) d\xi\, ds \le C_3 T^*(R^4+R^2+1) , \end{aligned} \tag{2.7}$$

because $\|y\|_{H^2(\mathcal{O})} \le C\|\Delta y\|_2$ and by the Sobolev embedding theorems $H_0^1(\mathcal{O}) \subset L^6(\mathcal{O})$. This yields

$$\|y\|_{C\left([0,T^*],H_0^1(\mathcal{O})\right)} \le C_3 \sqrt{T^*(R^4+R^2+1)} ,$$

and so for $T^*$ small enough we have that $y = Fz \in K$. Hence $F$ maps $K$ into itself. Moreover $F$ is a contraction on $K$ under the metric (2.4). Indeed we have by (2.5)

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}|y(t)-\bar{y}(t)|_2^2 + \|y(t)-\bar{y}(t)\|^2_{H_0^1(\mathcal{O})} \le C \int_{\mathcal{O}} \left( |z-\bar{z}||y-\bar{y}|(|z|+|\bar{z}|+1) \right) d\xi \le \\ &\le C\left( |z(t)-\bar{z}(t)|_2 |y(t)-\bar{y}(t)|_3 (|z(t)|_6 + |\bar{z}(t)|_6 + 1) \right) , \quad a.e. \ \ t\in[0,T] , \end{aligned}$$

the last being implied by the Hölder inequality, namely, $\left|\int_{\mathcal{O}} uvz\right| \le |u|_2|v|_3|z|_6$ , therefore, we have

$$\begin{aligned} &|y(t)-\bar{y}(t)|_2^2 + \int_0^t \|y(s)-\bar{y}(s)\|^2_{H_0^1(\mathcal{O})} ds \\ &\le C(R+1)\left( \int_0^t |z(s)-\bar{z}(s)|_2^2 ds \right)^{\frac{1}{2}} \left( \int_0^t |y(s)-\bar{y}(s)|_3^2 ds \right)^{\frac{1}{2}} \le \\ &\le \frac{C^2}{4}(R+1)^2 \int_0^t |z(s)-\bar{z}(s)|_2^2 ds + \int_0^t |y(s)-\bar{y}(s)|^2_{H_0^1(\mathcal{O})} ds , \end{aligned} \tag{2.8}$$

so that

$$\rho(y, \bar{y}) \leq \frac{C}{2}(R+1)T^{*}\rho(z, \bar{z})\,,$$

and taking $T^{*} < \frac{2}{C(R+1)}$, we have that $F$ is a contraction on $K$. Then by the Banach fixed point theorem on $[0, T^{*}]$, there is exists a unique solution $y$ to (2.5) providing that $T^{*} \in [0, T]$ is sufficiently small.

Let us now show by contradiction, that such a solution exists on a fixed interval $[0, T]$. Indeed if $[0, T^{*}]$ is the maximal interval on which $y$ exists, by (2.5) we have , as mentioned above, that the following estimate holds

$$\begin{aligned} &\|y(t)\|^{2}_{H^{1}_{0}(\mathcal{O})} + \int_{0}^{t} \left(\|y(s)\|^{2}_{H^{2}(\mathcal{O})} + |y(s)|^{6}_{6}\right) ds \leq \\ &\leq C \int_{0}^{t} \left(|y(s)|^{4}_{4} + |y(s)|^{2}_{2} + \|f_{3}\|^{2}_{2}\right) ds, \quad \forall\, t \in [0, T^{*}]\,. \end{aligned} \tag{2.9}$$

Taking into account that

$$|y|^{2}_{2} + |y|^{4}_{4} \leq \epsilon |u|^{6}_{6} + C_{\epsilon}, \quad \forall\, \epsilon > 0\,,$$

we get by (2.9) that

$$\|u(t)\|_{H^{1}_{0}(\mathcal{O})} + \int_{0}^{t} \|u(s)\|^{2}_{H^{2}(\mathcal{O})} ds + \int_{0}^{t} |u(s)|^{6}_{6} ds \leq C, \quad \forall\, t \in [0, T^{*}]\,,$$

where $C$ is independent of $T^{*}$. Therefore we also have that

$$|\frac{d}{dt} y(t)|_{2} \leq C_{1}, \quad \forall\, t \in [0, T^{*}]\,,$$

and the limit $\lim_{t \to T^{*}} y(t) = y(T^{*})$ exists with $u(T^{*}) \in H^{1}_{0}(\mathcal{O})$. Then we can apply the above local existence result, extending $y$ as a solution to (2.1) on $[T^{*}, T^{*} + \delta]$, which contradicts the assumption that $[0, T^{*}]$ is the maximal interval of existence, hence $T^{*} = T$.

Since the right hand side of (3.3) where $z = y$ is in $L^{2}(0, T; L^{2}(\mathcal{O}))$, we infer that

$$y \in C\left([0, T]; H^{1}_{0}(\mathcal{O})\right) \cap L^{2}\left([0, T]; H^{2}(\mathcal{O})\right)\,, \quad \mathbb{P} - a.s.\,, \tag{2.10}$$

moreover, since the contraction principle implies that the limit $y = \lim_{n \to \infty} y_{n}$ belongs to $C\left([0, T]; L^{2}(\mathcal{O})\right)$, where $y_{n} = F(y_{n-1})$ are $\left(\mathcal{F}_{t}\right)_{t \geq 0}$ −adapted, we can conclude that $y$ is in fact an $\left(\mathcal{F}_{t}\right)_{t \geq 0}$ −adapted process and so $y$ satisfies (2.2), as claimed. □

*Proof of Theorem 2.1(continued).* We set $X := y + W_A$, where $y$ is the solution to (2.1) given by Proposition 2.2, that is

$$\begin{cases} y_t - \Delta y + y^3 + ay^2 + by + 3W_A y^2 + 3W_A y^2 + 2W_A y = F - W_A^3 - aW_A^2 - bW_A \text{ in } [0,T] \times \mathcal{O}\,, \\ y = 0 \text{ on } [0,T] \times \partial\mathcal{O}\,, \\ y(0,\xi) = x(\xi)\,,\, \xi \in \mathcal{O} \end{cases}, \tag{2.11}$$

By assumption (1.3) on $W_A$ we see that

$$\begin{aligned} &\mathbb{E}\left[\sup_{(t,\xi)\in[0,T]\times\mathcal{O}} \left(|f_1|^{2m} + |f_2|^{2m}\right)\right] < \infty, \quad m = 1, 2\,, \\ &\mathbb{E}\left[\|f_3\|_{L^2([0,T]\times\mathcal{O})}\right] < \infty\,. \end{aligned} \tag{2.12}$$

Taking into account (2.11) and (2.1) we get that

$$y \in L^2\left(\Omega; C\left([0,T]; L^2(\mathcal{O})\right) \cap L^2_W\left([0,T]; H^1_0(\mathcal{O}) \cap H^2(\mathcal{O})\right)\right)\,, \tag{2.13}$$

which implies (2.2) as claimed. □

# 3 The optimal control of stochastic FitzHugh-Nagumo equation

Let $U$ be a real Hilbert space with the norm $|\cdot|_U$ and $B \in L\left(U; L^2(\mathcal{O})\right)$. We shall denote by $\mathcal{U}$ the space of all $\left(\mathcal{F}_t\right)_{t\geq 0}$ −adapted processes $u : [0,T] \to U$ s.t. $\mathbb{E}\left[\int_0^T |u(t)|_U^2 dt\right] < \infty$. The space $\mathcal{U}$ is a Hilbert space with the norm $|u|_{\mathcal{U}} = \left(\mathbb{E}\left[\int_0^T |u(t)|_U^2 dt\right]\right)^{\frac{1}{2}}$ and scalar product

$$\langle u, v\rangle_{\mathcal{U}} = \left(\mathbb{E}\left[\int_0^T \langle u(t), v(t)\rangle_U dt\right]\right)^{\frac{1}{2}}, \quad \forall\, u,\, v \in \mathcal{U}\,,$$

where $\langle\cdot,\cdot\rangle_U$ is the scalar product of $U$.

Consider the functions $g,\, g_0 : \mathbb{R} \to \mathbb{R}$ and $h : U \to \bar{\mathbb{R}} := ]-\infty, \infty]$, which satisfy the following conditions

**(i)** $g,\, g_0 \in C^1\left(L^2(\mathcal{O})\right)$ and $Dg,\, Dg_0 \in Lip\left(L^2(\mathcal{O}); L^2(\mathcal{O})\right)$ (where $D$ stands for the Fréchet differential) and $Lip\left(L^2(\mathcal{O}); L^2(\mathcal{O})\right)$ is the the space of Lipschitz continuous function from $L^2(\mathcal{O})$ to $L^2(\mathcal{O})$ with the norm defined denoted $\|\cdot\|_{Lip(L^2(\mathcal{O}))}$.

**(ii)** $h$ is convex, lower-semicontinuous and $(\partial h)^{-1} \in Lip(U)$ where $\partial h : U \to U$ is the subdifferential of $h$ (see, e.g. [7, p. 82]). Moreover assume that $\exists\ \alpha_1 > 0$ and $\alpha_2 \in \mathbb{R}$ s.t. $h(u) \geq \alpha_1 |u|_U^2 + \alpha_2$, $\forall\ u \in U$. We set $L = \|(\partial h)^{-1}\|_{Lip(U)}$ (Here $Lip(U)$ is the space of Lipschitz operators on $U$.

We consider the following optimal control problem

$$\text{Minimize}\,\mathbb{E}\left[\int_0^T (g(X(t)) + h(u(t)))\,dt\right] + \mathbb{E}\left[g_0(X(T))\right]\,, \tag{P}$$

subject to $u \in \mathcal{U}$ and

$$\begin{cases} dX(t) - \Delta_\xi X(t)dt + f(X(t))dt = \sqrt{Q}dW(t) + Bu(t)dt + f_0 dt\,, \text{ in } [0,T]\times\mathcal{O}\,, \\ X = 0 \text{ on } [0,T]\times\partial\mathcal{O}\,, \\ X(0) = x \text{ in } \mathcal{O}\,, \end{cases}, \tag{3.1}$$

where $f_0 \in L^\infty([0,T]\times\mathcal{O})$.

In the following we shall assume both (2.12) and $Tr[QA] < \infty$, where $A$ is as above the Laplace operator with domain $H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$.

**Theorem 3.1.** *Let $x \in H_0^1(\mathcal{O})$. Then there exists $C^* > 0$ independent of $x$ such that for $LT + \|Dg_0\|_{Lip(L^2(\mathcal{O}))} < C^*$ there is a unique solution $(u^*, X^*)$ to problem* (P)*.*

*Proof.* The proof is based on Ekeland's variational principle already used in a similar deterministic context (See, e.g. [6]). Namely, we consider the function $\Psi : \mathcal{U} \to \bar{\mathbb{R}}$ defined by

$$\Psi(u) = \mathbb{E}\left[\int_0^T (g(X^u(t)) + h(u(t)))\,dt\right] + \mathbb{E}\left[g_0(X^u(T))\right]\,,$$

where $X^u$ is the solution to (3.1). It is easily seen by (2.1) that $\Psi$ is lower-semicontinuous and $\Psi(u) \to +\infty$ as $|u|_{\mathcal{U}} \to +\infty$.

If $\Psi$ is weakly lower continuous on $\mathcal{U}$ this is sufficient for the existence of a minimum of $\Psi$ on $\mathcal{U}$. In the deterministic case, that is, if $Q = 0$ the weak lower continuity of $\Psi$ is a direct consequence of compactness of the map $u \mapsto X^u$ from $\mathcal{U}$ to $C([0,T]; L^2(\mathcal{O}))$ which is not the case here, that is, this map is not compact from $\mathcal{U}$ to $L^2(\Omega; C([0,T]; L^2(\mathcal{O})))$. So the existence in problem (P) does not follows by standard minimization techniques. However, by the

Ekeland variational principle, see, e.g., [16], there is a sequence $\{u_\epsilon\} \subset \mathcal{U}$ such that

$$\begin{aligned} &\Psi(u_\epsilon) \le \inf\{\Psi(u)\,; u \in \mathcal{U}\} + \epsilon\,, \\ &\Psi(u_\epsilon) \le \Psi(u) + \sqrt{\epsilon}\,|u_\epsilon - u|_{\mathcal{U}}\,, \quad \forall\, u \in \mathcal{U}\,. \end{aligned} \tag{3.2}$$

In other words,

$$u_\epsilon = \arg\min_{u\in\mathcal{U}}\{\Psi(u) + \sqrt{\epsilon}\,|u_\epsilon - u|_{\mathcal{U}}\}\,.$$

Hence $(X^{u_\epsilon}, u_\epsilon)$ is a solution to the optimal control problem

$$\begin{aligned} \min\Bigg\{ &\mathbb{E}\left[\int_0^T (g(X^u(t) + h(u(t)))\,dt\right] + \mathbb{E}\left[g_0\left(X^u(T)\right)\right] + \\ &+ \sqrt{\epsilon}\left(\mathbb{E}\left[\int_0^T |u(t) - u_\epsilon(t)|_U^2\,dt\right]\right)^{\frac{1}{2}} ; u \in \mathcal{U}\Bigg\}\,. \end{aligned} \tag{3.3}$$

The latter means that for all $v \in \mathcal{U}$ and $\lambda > 0$

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \left(g(X^{u_\epsilon+\lambda v}(t) + h((u_\epsilon + \lambda v)(t))\right)dt\right] + \mathbb{E}\left[g_0(X^{u_\epsilon+\lambda v}(T))\right] + \\ &+ \lambda\sqrt{\epsilon}\left(\mathbb{E}\left[\int_0^T |v(t)|_U^2\,dt\right]\right)^{\frac{1}{2}} \le \\ &\le \mathbb{E}\left[\int_0^T (g(X_\epsilon(t)) + h(u_\epsilon(t)))\,dt\right] + \mathbb{E}\left[g_0(X_\epsilon(T))\right]\,. \end{aligned}$$

This yields

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \langle Dg(X_\epsilon(t)), Z^v(t)\rangle_2\,dt\right] + \mathbb{E}\left[\int_0^T h'(u_\epsilon(t), v(t))dt\right] + \\ &+ \mathbb{E}\left[\langle Dg_0(X_\epsilon(T)), Z^v(T)\rangle_2\right] + \sqrt{\epsilon}\left(\mathbb{E}\left[\int_0^T |v(t)|_U^2 dt\right]\right)^{\frac{1}{2}} \le 0\,, \quad \forall\, v \in \mathcal{U}\,, \end{aligned} \tag{3.4}$$

where $Z^v$ solves the system in variations associated with (3.1), that is

$$\begin{cases} \frac{\partial}{\partial t} Z^v - \Delta Z^v + f'(X_\epsilon)Z^v = Bv \ \text{ in } \ [0,T] \times \mathcal{O}\,, \\ Z^v(0) = 0 \ \text{ in } \ \mathcal{O}\,, \\ Z^v = 0 \ \text{ on } \ [0,T] \times \partial\mathcal{O}\,, \end{cases} \tag{3.5}$$

and $h' : U \times U \to \mathbb{R}$ is the directional derivatives of $h$, see, e.g., [7, p.81], namely

$$h'(u_\epsilon, v) = \lim_{\lambda \downarrow 0} \frac{h(u_\epsilon + \lambda v) - h(u_\epsilon)}{\lambda}, \quad \forall\, v \in U\,.$$

We associate with (3.1) the dual stochastic backward equation

$$\begin{cases} dp_\epsilon + \Delta p_\epsilon dt - f'(X_\epsilon) p_\epsilon dt = \kappa_\epsilon \sqrt{Q} dW(t) + Dg(X_\epsilon) dt \text{ in } [0,T] \times \mathcal{O}\,, \\ p_\epsilon(T) = -Dg_0(X_\epsilon(T)) \text{ in } \mathcal{O}\,, \\ p_\epsilon = 0 \text{ on } [0,T] \times \partial\mathcal{O}\,, \end{cases} \qquad . \tag{3.6}$$

It is well-known that equation (3.6) has a unique solution $(p_\epsilon, \kappa_\epsilon)$ satisfying

$$\begin{aligned} & p_\epsilon \in L^\infty_W\left([0,T]; L^2(\mathcal{O})\right) \cap L^2_W\left([0,T]; H^1_0(\mathcal{O})\right)\,, \\ & k_\epsilon \in L^2_W\left([0,T]; L^2(\mathcal{O})\right)\,, \end{aligned}$$

(See, e.g., [17, Prop. 4.3] or [22]). By Itô's formula we have

$$d\left\langle p_\epsilon, Z^v \right\rangle_2 = \left\langle dp_\epsilon, Z^v \right\rangle_2 + \left\langle p_\epsilon, dZ^v \right\rangle_2\,,$$

and this yields

$$\mathbb{E}\left[\int_0^T \left\langle Dg(X_\epsilon(t)), Z^v(t)\right\rangle_2 dt\right] + \mathbb{E}\left[\left\langle Dg_0(X_\epsilon(T)), Z^v(T)\right\rangle_2\right] = 0\,,$$

and substituiting in (3.4), we obtain, $\forall\, v \in \mathcal{U}$, the following inequality

$$\begin{aligned} & \mathbb{E}\left[\int_0^T h'(u_\epsilon(t), v(t)) dt\right] + \sqrt{\epsilon}\left(\mathbb{E}\left[\int_0^T |v(t)|_U^2 dt\right]\right)^{\frac{1}{2}} \leq \\ & \leq \mathbb{E}\left[\int_0^T \left\langle B^* p_\epsilon(t), v(t)\right\rangle_U dt\right]\,. \end{aligned}$$

Let $G(u) := \mathbb{E}\left[\int_0^T h(u(t)) dt\right]$, then its subdifferential $\partial G : \mathcal{U} \to \mathcal{U}$, evaluated in $u_\epsilon$ is given by

$$\partial G(u_\epsilon) = \left\{ v^* \in \mathcal{U}\,:\, \langle v, v^* \rangle_{\mathcal{U}} \leq \mathbb{E}\left[\int_0^T h'(u_\epsilon(t), v(t)) dt\right], \forall\, v \in \mathcal{U} \right\}.$$

(See, e.g., [7, p.81]). Then we infer that

$$u_\epsilon(t) = (\partial h)^{-1}\left(B^* p_\epsilon(t) + \sqrt{\epsilon}\tilde{\theta}_\epsilon\right),\ t \in [0,T]\,, \quad \mathbb{P} - a.s.\,,$$

where $\tilde{\theta}_\epsilon \in \mathcal{U}$ and $|\tilde{\theta}_\epsilon|_{\mathcal{U}} \le 1$, $\forall\, \epsilon > 0$.

Therefore, we have shown that

$$\begin{array}{l} u_\epsilon = (\partial h)^{-1}\left(B^* p_\epsilon + \theta_\epsilon\right)\,, \|\theta_\epsilon\|_{L^2([0,T]\times\Omega;U)} \le \sqrt{\epsilon}\,, \\ dp_\epsilon + \Delta p_\epsilon dt - f'(X_\epsilon) p_\epsilon dt = Dg_\epsilon(X_\epsilon)dt + \kappa_\epsilon \sqrt{Q} dW(t) \ \text{ in } [0,T]\times\mathcal{O}\,, \\ p_\epsilon(T) = -Dg_0(X_\epsilon(T)) \ \text{ in } \ \mathcal{O}\,, \\ p_\epsilon = 0 \ \text{ in } \ [0,T]\times\partial\mathcal{O}\,, \end{array} \qquad (3.7)$$

By (3.2) and by assumptions (ii) it follows also that $\{u_\epsilon\}_{\epsilon>0}$ is bounded in $\mathcal{U}$. Moreover, by (3.1) we have that

$$\begin{cases} dX_\epsilon(t) - \Delta X_\epsilon(t)dt + f(X_\epsilon(t))dt = \sqrt{Q}dW(t) + f_0 dt + Bu_\epsilon(t)dt\,, \ \text{in } \ [0,T]\times\mathcal{O}\,, \\ X_\epsilon = 0 \ \text{ on } [0,T]\times\partial\mathcal{O}\,, \\ X_\epsilon(0) = x \ \text{ in } \mathcal{O} \end{cases}, \qquad (3.8)$$

which by (3.3), assumption (ii) and exploiting the Itô formula, implies that

$$\mathbb{E}\left[\int_0^T \left(|X_\epsilon(t)|_2^2 + |\nabla X_\epsilon(t)|_2^2 + L^{-1}|u_\epsilon(t)|_U^2\right) dt\right] \le C\,, \quad \forall\, \epsilon > 0\,. \qquad (3.9)$$

Moreover by (3.8) and (3.2), again using the Itô formula applied to $|X|_2^2$, we have that $\forall\ \epsilon > 0$

$$\begin{array}{l} \mathbb{E}\left[\sup_{t\in[0,T]} |X_\epsilon(t)|_2^2\right] + \mathbb{E}\left[\int_0^T |X_\epsilon(t)|^2_{H_0^1(\mathcal{O})}\, dt\right] + \mathbb{E}\left[\int_0^T |X_\epsilon(t)|_2^4\, dt\right] \le \\ \le C(1+|x|_2^2) \end{array}. \qquad (3.10)$$

If we now apply the Itô formula in (3.8) to the function $X \to \frac{1}{2}|X|^2_{H_0^1(\mathcal{O})}$, taking into account that $Tr[QA] < \infty$ and that

$$-\int_{\mathcal{O}} f(X_\epsilon)\Delta X_\epsilon d\xi \ge ab\int_{\mathcal{O}} |\nabla X_\epsilon|^2 d\xi - \int_{\mathcal{O}} |\Delta X_\epsilon||X_\epsilon|^2 d\xi\,,$$

we obtain by (3.10) that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_\epsilon(t)|^2_{H_0^1(\mathcal{O})}\right] + \mathbb{E}\left[\int_0^T |\Delta X_\epsilon(t)|_2^2 dt\right] \le C(1+|x|^2_{H_0^1(\mathcal{O})})\,. \qquad (3.11)$$

Similarly, by (3.7) we obtain that

$$\frac{1}{2}d|p_\epsilon(t)|_2^2 - \int_{\mathcal{O}} |\nabla p_\epsilon(t)|^2 d\xi - \int_{\mathcal{O}} f'(X_\epsilon)p_\epsilon^2(t)d\xi = \\ = \int_{\mathcal{O}} Dg(X_\epsilon(t))p_\epsilon(t)d\xi + \frac{1}{2}\int_{\mathcal{O}} |\kappa_\epsilon|^2 d\xi + \int_{\mathcal{O}} p_\epsilon \kappa_\epsilon \sqrt{Q} dW(t)\,.$$

which yields

$$\mathbb{E}\left[\sup_{t\in[0,T]} |p_\epsilon(t)|_2^2\right] + \mathbb{E}\left[\int_0^T |p_\epsilon(t)|^2_{H_0^1(\mathcal{O})}\,dt\right] + \mathbb{E}\left[\int_0^T\int_{\mathcal{O}} |X_\epsilon|^2|p_\epsilon|^2 d\xi\,dt\right] \\ + \mathbb{E}\left[\int_0^T |\kappa_\epsilon(t)|_2^2 dt\right] \leq C + \mathbb{E}\left[|X_\epsilon(T)|_2^2\right] \leq C\,, \quad \forall\, \epsilon > 0\,. \tag{3.12}$$

(Here and everywhere in the following we shall denote by $C$ several positive constants independent of $\epsilon$). In particular, it follows by (3.7) and (3.12) that $\{u_\epsilon\}_{\epsilon>0}$ is bounded in $L^2(\Omega; L^\infty([0,T];U))$.

Equation (3.8) implies that

$$\frac{\partial}{\partial t}(X_\epsilon(t) - X_\lambda(t)) - \Delta(X_\epsilon(t) - X_\lambda(t)) + (f(X_\epsilon(t)) - f(X_\lambda(t))) = \\ = BB^*(p_\epsilon(t) - p_\lambda(t)) + B(\theta_\epsilon(t) - \theta_\lambda(t))\,. \tag{3.13}$$

In virtue of (3.12) this yields

$$\frac{1}{2}|X_\epsilon(t) - X_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|^2_{H_0^1(\mathcal{O})}\,ds \leq \\ \leq -\int_0^t\int_{\mathcal{O}} (f(X_\epsilon(s)) - f(X_\lambda(s)))(X_\epsilon(s) - X_\lambda(s))\,d\xi\,ds \\ + L\int_0^t |p_\epsilon(s) - p_\lambda(s)|_2 |X_\epsilon(s) - X_\lambda(s)|_2 ds \\ + C\int_0^t |\theta_\epsilon(s) - \theta_\lambda(s)|_U |X_\epsilon(s) - X_\lambda(s)|_2 ds\,, \quad \forall\, t \in [0,T]\,,$$

where $L = \|(\partial h)^{-1}\|_{Lip}$.

We further have

$$(f(X_\epsilon) - f(X_\lambda))(X_\epsilon - X_\lambda) = f'(\alpha X_\epsilon + (1-\alpha)X_\lambda)(X_\epsilon - X_\lambda)^2\,,$$

where $\alpha \in [0,1]$ and assuming that $0 < a < b$,

$$f'(u) \geq 0 \ \text{ for } \ u \notin [0,b]\,,$$
$$|f'(u)| \leq C \ \text{ for } \ u \in [0,b]\,,$$

then

$$-\int_0^t \int_{\mathcal{O}} (f(X_\epsilon) - f(X_\lambda))\,(X_\epsilon - X_\lambda)\,d\xi\,ds \leq C\int_0^t |X_\epsilon(s) - X_\lambda(s)|_2^2 ds\,, \quad \forall\,\epsilon,\,\lambda > 0\,,$$

which yields, for $t \in [0,T]$

$$\begin{aligned} &|X_\epsilon(t) - X_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|^2_{H_0^1(\mathcal{O})}\,ds \leq \\ &\leq C\left(L\int_0^t |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \int_0^t\int_{\mathcal{O}} (X_\epsilon(s) - X_\lambda(s))^2 d\xi ds + \epsilon + \lambda\right)\,. \end{aligned} \tag{3.14}$$

Applying Gronwall's lemma in (3.14), we have

$$\begin{aligned} &|X_\epsilon(t) - X_\lambda(t)|_2^2 + \frac{1}{2}\int_0^t |X_\epsilon(s) - X_\lambda(s)|^2_{H_0^1(\mathcal{O})}\,ds \leq \\ &\leq C\left(L\int_0^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda\right)\,, \quad \forall\,\epsilon\,,\,\lambda > 0\,,\,t \in [0,T]\,. \end{aligned} \tag{3.15}$$

Similarly we get by (2.9) and the Itô formula

$$
\begin{aligned}
&|p_\epsilon(t)-p_\lambda(t)|_2^2+\int_t^T |p_\epsilon(s)-p_\lambda(s)|^2_{H^1_0(\mathcal{O})}\,ds+\frac{1}{2}\int_t^T|\kappa_\epsilon(s)-\kappa_\lambda(s)|_2^2ds=\\
&=|Dg_0(X_\epsilon(T))-Dg_0(X_\lambda(T))|_2^2+\\
&\quad+\int_t^T\int_{\mathcal{O}}\left(f'(X_\epsilon(s))p_\epsilon(s)-f'(X_\lambda(s))p_\lambda(s)\right)(p_\epsilon(s)-p_\lambda(s))d\xi ds+\\
&\quad-\int_t^T\left\langle \kappa_\epsilon(s)-\kappa_\lambda(s))\sqrt{Q}dW(s),X_\epsilon(s)-X_\lambda(s)\right\rangle_2=\\
&=\int_t^T\int_{\mathcal{O}}f'(X_\epsilon(s))(p_\epsilon(s)-p_\lambda(s))^2d\xi\,ds+\\
&\quad-\int_t^T\int_{\mathcal{O}}\left(f'(X_\epsilon(s)-f'(X_\lambda(s))\right)(p_\epsilon(s)-p_\lambda(s))d\xi\,ds+\\
&\quad-\int_t^T\left\langle \kappa_\epsilon(s)-\kappa_\lambda(s))\sqrt{Q}dW(s),X_\epsilon(s)-X_\lambda(s)\right\rangle_2+\\
&\quad+|Dg_0\left(X_\epsilon(T)\right)-Dg_0\left(X_\lambda(T)\right)|_2^2\le\\
&\le C\left(\int_t^T\int_{\mathcal{O}}(|X_\epsilon|+1)(p_\epsilon(s)-p_\lambda(s))^2d\xi\,ds\right)+\\
&\quad+\left(\int_t^T\int_{\mathcal{O}}\left(X_\epsilon(s)-X_\lambda(s)\right)(p_\epsilon(s)-p_\lambda(s))(1+|X_\epsilon|+|X_\lambda|)|p_\epsilon|d\xi\,ds\right)+\\
&\quad-\int_t^T\left\langle \kappa_\epsilon(s)-\kappa_\lambda(s))\sqrt{Q}dW(s),X_\epsilon(s)-X_\lambda(s)\right\rangle_2+\\
&\quad+\|Dg_0\|_{Lip(L^2(\mathcal{O}))}|X_\epsilon(T)-X_\lambda(T)|_2^2,\quad t\in[0,T]\,,\mathbb{P}-a.s..
\end{aligned}
\tag{3.16}
$$

Proceeding as above, we also have

$$
\begin{aligned}
&\int_{\mathcal{O}}|X_\epsilon(s)||p_\epsilon(s)-p_\lambda(s)|^2d\xi\le|p_\epsilon(s)-p_\lambda(s)|_4|p_\epsilon(s)-p_\lambda(s)|_2|X_\epsilon(s)|_4\le\\
&\le\frac{1}{2}|p_\epsilon(s)-p_\lambda(s)|^2_{H^1_0(\mathcal{O})}+\frac{1}{2}|p_\epsilon(s)-p_\lambda(s)|_2^2|X_\epsilon(s)|_4^2\,.
\end{aligned}
\tag{3.17}
$$

Moreover, exploiting both the Hölder and the interpolation inequality, we obtain

$$
\begin{aligned}
&\int_{\mathcal{O}} |X_\epsilon - X_\lambda|\,|p_\epsilon - p_\lambda|\,(1+|X_\epsilon|+|X_\lambda|)|p_\epsilon| d\xi \leq \\
&\leq |X_\epsilon - X_\lambda|_4\,|p_\epsilon - p_\lambda|_4 \left(\int_{\mathcal{O}} (1+|X_\epsilon|+|X_\lambda|)^2 |p_\epsilon|^2 d\xi\right)^{\frac{1}{2}} \leq \\
&\leq |X_\epsilon - X_\lambda|_2^{\frac{1}{2}}\,|X_\epsilon - X_\lambda|_6^{\frac{1}{2}}\,|p_\epsilon - p_\lambda|_2^{\frac{1}{2}}\,|p_\epsilon - p_\lambda|_6^{\frac{1}{2}} \times \\
&\quad \times \left(\int_{\mathcal{O}} (1+|X_\epsilon|+|X_\lambda|)^2 |p_\epsilon|^2 d\xi\right)^{\frac{1}{2}} \leq \\
&\leq |X_\epsilon - X_\lambda|_2^{\frac{1}{2}}\,|X_\epsilon - X_\lambda|_{H_0^1(\mathcal{O})}^{\frac{1}{2}}\,|p_\epsilon - p_\lambda|_2^{\frac{1}{2}}\,|p_\epsilon - p_\lambda|_{H_0^1(\mathcal{O})}^{\frac{1}{2}} \times \\
&\quad \times \left(\int_{\mathcal{O}} (1+|X_\epsilon|+|X_\lambda|)^2 |p_\epsilon|^2 d\xi\right)^{\frac{1}{2}} \leq \alpha\left(|X_\epsilon - X_\lambda|_{H_0^1(\mathcal{O})}^2 + |p_\epsilon - p_\lambda|_{H_0^1(\mathcal{O})}^2\right) + \\
&\quad + \frac{C}{\alpha}\left(|X_\epsilon - X_\lambda|_2^2 + |p_\epsilon - p_\lambda|_2^2\right)\left(\int_{\mathcal{O}} (1+|X_\epsilon|+|X_\lambda|)^2 |p_\epsilon|^2 d\xi\right),
\end{aligned}
\tag{3.18}
$$

where $\alpha$ is arbitrary small. Substituting now (3.17), (3.18) into (3.14), (3.16), we obtain $\mathbb{P}-$a.s.

$$
\begin{aligned}
&|X_\epsilon(t) - X_\lambda(t)|_2^2 + |p_\epsilon(t) - p_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \\
&\quad + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_{H_0^1(\mathcal{O})}^2\, ds + \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds \leq \\
&\leq C\left(L\int_0^t |p_\epsilon(s) - p_\lambda(s)|_2^2\, ds + \epsilon + \lambda\right) + C\int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2\, |X_\epsilon(s)|_4^2 ds + \\
&\quad + \|Dg_0\|_{Lip} |X_\epsilon(T) - X_\lambda(T)|_2^2 + \\
&\quad + \frac{C}{\alpha}\left(\int_t^T |X_\epsilon(s) - X_\lambda(s)|_2^2 + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 T_{\epsilon,\lambda}(s) ds\right) + \\
&\quad - \int_t^T \left\langle \kappa_\epsilon(s) - \kappa_\lambda(s))\sqrt{Q}dW(s), X_\epsilon(s) - X_\lambda(s)\right\rangle_2, \quad \forall\, t\in[0,T],
\end{aligned}
\tag{3.19}
$$

where

$$T_{\epsilon,\lambda} := \int_{\mathcal{O}} (1+|X_\epsilon|+|X_\lambda|)^2 |p_\epsilon|^2 d\xi\,.$$

We note that the process $r \mapsto \int_t^r \left\langle (\kappa_\epsilon - \kappa_\lambda)\sqrt{Q}dW(s), X_\epsilon(s) - X_\lambda(s)\right\rangle_2$ is a local martingale on $[t,T]$, hence by the Burkholder-Davis-Gundy inequality, see, e.g., [13, p.58], we have for all $r \in [t,T]$

$$\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[t,T]} \left|\int_t^r \left\langle (\kappa_\epsilon(s)-\kappa_\lambda(s))\sqrt{Q}dW(s), X_\epsilon(s)-X_\lambda(s)\right\rangle_2\right|\right] \leq \\
&\leq C\left(\mathbb{E}\left[\int_0^r |\kappa_\epsilon(s)-\kappa_\lambda(s)|_2^2 |X_\epsilon(s)-X_\lambda(s)|_2^2 ds\right]\right)^{\frac{1}{2}} \leq \\
&\leq C\mathbb{E}\left[\sup_{s\in[t,r]} |X_\epsilon(s)-X_\lambda(s)|_2^2\right] + \frac{1}{2}\mathbb{E}\left[\int_t^r |\kappa_\epsilon(s)-\kappa_\lambda(s)|_2^2 ds\right],
\end{aligned}$$

and by (3.18) we get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[t,T]} \left(|X_\epsilon(s)-X_\lambda(s)|_2^2 + |p_\epsilon(s)-p_\lambda(s)|_2^2\right)\right] \\
&\quad + \mathbb{E}\left[\int_0^T |X_\epsilon(s)-X_\lambda(s)|^2_{H_0^1(\mathcal{O})} ds + \int_t^T |p_\epsilon(s)-p_\lambda(s)|_2^2 ds\right] \\
&\quad + \mathbb{E}\left[\int_t^T |\kappa_\epsilon(s)-\kappa_\lambda(s)|_2^2 ds\right] \leq \\
&\leq \|Dg_0\| \mathbb{E}\left[|X_\epsilon(T)-X_\lambda(T)|_2^2\right] + C\left(L\mathbb{E}\left[\int_0^T |p_\epsilon(s)-p_\lambda(s)|_2^2\, ds\right] + \epsilon + \lambda\right) \\
&\quad + C\mathbb{E}\left[\sup_{s\in[t,T]} |X_\epsilon(s)-X_\lambda(s)|_2^2\right] \\
&\quad + C\mathbb{E}\left[\int_t^T \left(|p_\epsilon(s)-p_\lambda(s)|_2^2 + |X_\epsilon(s)-X_\lambda(s)|_2^2\right)\left(|X_\epsilon(s)|_4^2 + T_{\epsilon,\lambda}(s)\right) ds\right].
\end{aligned} \tag{3.20}$$

Taking into account estimates (3.10), (3.11) and (3.15), from (3.20) we have

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[t,T]}\left(|X_\epsilon(s)-X_\lambda(s)|_2^2+|p_\epsilon(s)-p_\lambda(s)|_2^2\right)\right]\\
&\quad+\mathbb{E}\left[\int_0^T|X_\epsilon(s)-X_\lambda(s)|^2_{H_0^1(\mathcal{O})}ds+\int_t^T|p_\epsilon(s)-p_\lambda(s)|_2^2ds\right]\\
&\quad+\mathbb{E}\left[\int_t^T|\kappa_\epsilon(s)-\kappa_\lambda(s)|_2^2ds\right]\leq\\
&\leq\tilde{C}\left(L\mathbb{E}\left[\int_0^T|p_\epsilon(s)-p_\lambda(s)|_2^2\,ds\right]\right)\\
&\quad+\tilde{C}\left(\mathbb{E}\left[\int_t^T|p_\epsilon(s)-p_\lambda(s)|_2^2\left(|X_\epsilon(s)|_4^2+T_{\epsilon,\lambda}(s)\right)ds\right]\right)\\
&\quad+\tilde{C}\|Dg_0\|_{Lip}\mathbb{E}\left[|X_\epsilon(T)-X_\lambda(T)|_2^2\right]+\tilde{C}(\epsilon+\lambda)\,.
\end{aligned}
\tag{3.21}
$$

where $\tilde{C}$ is a positive constant independent of $\epsilon$ and $\lambda$. It follows that if $\tilde{C}(LT+\|Dg_0\|_{Lip})<1$, then, for any $t\in[0,T]$,

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[t,T]}\left(|X_\epsilon(s)-X_\lambda(s)|_2^2+|p_\epsilon(s)-p_\lambda(s)|_2^2\right)\right]\\
&\quad+\mathbb{E}\left[\int_0^T|X_\epsilon(s)-X_\lambda(s)|^2_{H_0^1(\mathcal{O})}ds+\int_t^T|p_\epsilon(s)-p_\lambda(s)|_2^2ds\right]\\
&\quad+\mathbb{E}\left[\int_t^T|\kappa_\epsilon(s)-\kappa_\lambda(s)|_2^2ds\right]\leq\\
&\leq C\mathbb{E}\left[\int_t^T|p_\epsilon(s)-p_\lambda(s)|_2^2\left(|X_\epsilon(s)|_4^2+T_{\epsilon,\lambda}(s)\right)ds\right]+C(\epsilon+\lambda)\,.
\end{aligned}
\tag{3.22}
$$

Let us define for $j\in\mathbb{N}$

$$
\Omega_j:=\left\{\omega\in\Omega\,:\sup_{\epsilon}\sup_{t\in[0,T]}\left(|X_\epsilon(t)|_2^2+|X_\epsilon(t)|^2_{H_0^1(\mathcal{O})}+|X_\epsilon(t)|_4^2+|p_\epsilon(t)|_2^2\right)dt\leq j\right\}\,,
$$

then estimates (3.10) and (3.11) implies that

$$
\mathbb{P}\left(\Omega_j\right)\geq 1-\frac{C}{j}\,,\quad\forall\,j\in\mathbb{N}\,,
$$

for some constant $C$ independent of $\epsilon$.

If we set $X_\epsilon^j := \mathbb{1}_{\Omega_j} X_\epsilon$, $p_\epsilon^j := \mathbb{1}_{\Omega_j} p_\epsilon$ and $\kappa_\epsilon^j := \mathbb{1}_{\Omega_j}\kappa_\epsilon$, then such quantities satisfy the system (3.7)-(3.8), with $\mathbb{1}_{\Omega_j}\sqrt{Q}dW$. The latter means that estimate (3.22) still holds in this context, so that we have

$$\begin{aligned} &\mathbb{E}\left[\sup_{s\in[t,T]}|X_\epsilon^j(s)-X_\lambda^j(s)|_2^2+\sup_{s\in[t,T]}|p_\epsilon^j(t)-p_\lambda^j(t)|_2^2\right] \\ &\quad+\mathbb{E}\left[\int_t^T|p_\epsilon^j(s)-p_\lambda^j(s)|^2_{H_0^1(\mathcal{O})}ds\right]+\mathbb{E}\left[\int_t^T|(\kappa_\epsilon(s)-\kappa_\lambda(s))\chi_j|_2^2 ds\right]\leq \\ &\leq C_j\int_t^T\mathbb{E}\left[|p_\epsilon^j(s)-p_\lambda^j(s)|_2^2\right]ds+C\left(\epsilon+\lambda\right)\,,\quad j\in\mathbb{N}\,. \end{aligned} \tag{3.23}$$

By Gronwall's lemma we get, for any $t\in[0,T]$

$$\mathbb{E}\left[\sup_{s\in[t,T]}|X_\epsilon^j(s)-X_\lambda^j(s)|_2^2+\sup_{s\in[t,T]}|p_\epsilon^j(s)-p_\lambda^j(s)|_2^2\right]\leq C(\epsilon+\lambda)e^{C_jT}\,, \tag{3.24}$$

where $C_j=C(j^3+1)$, hence, for $\epsilon\to 0$ and all $j\in\mathbb{N}$, we obtain

$$\begin{aligned} &X_\epsilon^j\to X^j \quad \text{in} \quad L^2\left(\Omega_j;L^2\left([0,T]\times\mathcal{O}\right)\right)\,,\\ &p_\epsilon^j\to p^j \quad \text{in} \quad L^2\left(\Omega_j;L^2\left([0,T]\times\mathcal{O}\right)\right)\,, \end{aligned} \tag{3.25}$$

where $\to$ means strong convergence. By estimates (3.10) and (3.12) it follows that taking related subsequences, still denoted by $\epsilon$, we have

$$\begin{aligned} &X_\epsilon\rightharpoonup X^* &&\text{in } L^2\left([0,T]\times\Omega;H_0^1(\mathcal{O})\right)\,,\\ &p_\epsilon\rightharpoonup p^* &&\text{in } L^\infty\left([0,T];L^2\left(\Omega\times\mathcal{O}\right)\right)\,,\\ &p_\epsilon\rightharpoonup p^* &&\text{in } L^2\left([0,T]\times\Omega\times\mathcal{O}\right)\,,\\ &p_\epsilon\rightharpoonup p^* &&\text{in } L^2\left([0,T]\times\Omega;H_0^1(\mathcal{O})\right)\,,\\ &u_\epsilon\rightharpoonup u^* &&\text{in } L^\infty\left([0,T];L^2\left(\Omega;U\right)\right)\,, \end{aligned} \tag{3.26}$$

where $\rightharpoonup$ means weak (respectively, weak-star) convergence, so we have for $\epsilon\to 0$

$$X_\epsilon\to X^*\,,\quad p_\epsilon\to p^*\,, a.e. \text{ in } [0,T]\times\Omega_j\times\mathcal{O}\,. \tag{3.27}$$

By (3.10) we see that

$$\mathbb{E}\int_0^T\int_{\mathcal{O}}|f(X_\epsilon(s,\xi))|^{\frac{4}{3}}d\xi ds\leq C\,,\quad \forall\epsilon>0\,.$$

Since $\{f(X_\epsilon)\}$ is bounded in $L^{\frac{4}{3}}([0,T]\times\Omega\times\mathcal{O})$, then it is weakly compact in $L^1([0,T]\times\Omega\times\mathcal{O})$ and by (3.27) we have that for a subsequence $\{\epsilon\}\to 0$,

$$f(X_\epsilon)\to f(X^*)\,,\quad a.e.\ \text{in}\ [0,T]\times\Omega\times\mathcal{O}\,,$$

which, in virtue of (3.27) and since

$$\mathbb{P}(\Omega_j)\geq 1-\frac{C}{j}\,,\forall\, j\in\mathbb{N}_0\,,$$

we have

$$f(X_\epsilon)\to f(X^*)\qquad \text{in}\ L^1([0,T]\times\Omega_j\times\mathcal{O})\,. \tag{3.28}$$

Then, letting $\epsilon\to 0$ in (3.8), we obtain

$$\begin{cases} dX^*(t)-\Delta X^*(t)dt+f(X^*(t))dt=\sqrt{Q}dW(t)+Bu^*(t)dt \ \text{in}\ [0,T]\times\mathcal{O}\,,\\ X^*=0\ \text{on}\ [0,T]\times\partial\mathcal{O}\,,\\ X^*(0)=x\ \text{in}\ \mathcal{O}\end{cases}\,.$$

Taking into account that $\Psi$ is weakly lower semicontinuous in $\mathcal{U}$ we infer by (3.2) that

$$\Psi(u^*)=\inf\{\Psi(u);u\in\mathcal{U}\}\,,$$

therefore $(X^*,u^*)$ is optimal for the problem (P) and the proof of existence is therefore complete. □

Concerning the uniqueness for the optimal pair $(X^*,u^*)$ given by Th. 3.1, we have that it follows by the same argument via the maximum principle result for problem (P), namely one has the following result.

**Theorem 3.2.** *Let $(X^*,u^*)$ be optimal in problem (P), then*

$$u^*=(\partial h)^{-1}(B^*p)\,,a.e.\,t\in[0,T]\,, \tag{3.29}$$

*where $p$ is the solution to the backward stochastic equation*

$$\begin{cases} dp+\Delta pdt+f'(X^*)pdt=g'(X)dt+\kappa\sqrt{Q}dW(t)\qquad in\ [0,T]\times\mathcal{O}\,,\\ p(T)=-Dg_0(X^*(T))\ \ in\ \mathcal{O}\,,\\ p_\epsilon=0\ \ on\ [0,T]\times\partial\mathcal{O}\end{cases}\,, \tag{3.30}$$

*Proof.* If $(X^*, u^*)$ is optimal for the problem (P), then by the same argument used to prove Th. 3.1, see (3.4), we have

$$\begin{aligned}&\mathbb{E}\left[\int_0^T \langle Dg(X^*(t)), Z^v(t)\rangle_2\, dt\right] + \mathbb{E}\left[\int_0^T h'(u^*(t), v(t))dt\right] \\ &+ \mathbb{E}\left[\langle Dg_0(X^*(T)), Z^v(T)\rangle_2\right] \le 0\,, \quad \forall\, v \in \mathcal{U}\,,\end{aligned} \tag{3.31}$$

where $Z^v$ is solution to equation (3.5) with $X_\epsilon$ replaced by $X^*$. This implies as above that (3.29) holds. □

*The uniqueness in* (P). If $(X^*, u^*)$ is optimal in (P) then it satisfies systems (1.1), (3.29) and (3.31), so that arguing as in the proof of Th. 3.1, the same set of estimates implies that the previous system has at most one solution if $LT + \|Dg_0\|_{Lip} < C^*$, where $C^*$ is sufficiently small. □

*Remark* 3.3. Theorems 3.1 and 3.2 remain true if assumption (i) is relaxed to

**(i)'** $Dg_0 \in Lip(L^2(\mathcal{O}))$, $g = g(t, y) : [0, T] \times L^2(\mathcal{O}) \to \mathbb{R}$ is of class $C^1$ in $y$, $D_y g \in C([0, T] \times L^2(\mathcal{O}))$, and $\sup_{t\in[0,T]} \|D_y\, g(t, y)\|_{Lip(L^2(\mathcal{O}))} < \infty$.

*Remark* 3.4. As clear from the previous proof the constant $C^*$ arising in conditions of Theorem 3.1 depends of $f$ and $g$ only and, as mentioned earlier, it is independent of initial data $x$.

# 4 An example

Roughly speaking the control objective in system (1.1) is to drive the potential $X$ to track a given trajectory $X^1$ and an end potential $X^0$. This can be reformulated as the optimal control problem

$$\text{Minimize} \quad \mathbb{E}\left[\int_0^T \alpha|u(t)|_2^2 + |X(t) - X^1(t)|_2^2 dt\right] + \lambda\mathbb{E}\left[|X(T) - X^0|_2^2\right]\,, \tag{4.1}$$

subject to

$$u \in L^2_W([0, T]; L^2(\mathcal{O}))\,, \quad m \le u \le M\, a.e. \text{ on } [0, T] \times \mathcal{O}\,. \tag{4.2}$$

$$\begin{cases} dX(t) - \Delta X(t)dt + f(X(t)) = \sqrt{Q}dW(t) + u(t)dt + f_0 dt \text{ in } [0, T] \times \mathcal{O}\,, \\ X(0) = x \text{ in } \mathcal{O}\,, \\ X = 0 \text{ on } [0, T] \times \partial\mathcal{O} \end{cases}\,, \tag{4.3}$$

where $f(u) = u(u-a)(u-b)$, $\alpha$, $\lambda > 0$, $0 < m < M < \infty$ and the functions $X^1 \in L^2([0,T]; L^2(\mathcal{O}))$, $X^0 \in L^2(\mathcal{O})$ are given.

As mentioned above, the physical significance of the problem is the following: find an optimal current $u$ applied to a nerve axon in such a way that the resulting potential $X$ flows closely to a specified regime $X^0 = X^0(t,\xi)$ during the time interval [0,T], and such that it is near to a given potential $X_0$ at the final time $T$ (on these lines see also [21]).

Problem (4.1)-(4.3) is of the form (P) where

$$g(t,X) = |X - X^1(t)|_2^2 \quad , \quad g_0(X) = \lambda |X - X^0|_2^2 \,, \tag{4.4}$$

and $h : L^2(\mathcal{O}) \to ]-\infty, +\infty]$ is defined by

$$h(u) = \begin{cases} \alpha |u|_2^2 & \text{if } u \in U_0 \,, \\ \infty & \text{if } u \notin U_0 \end{cases} \,, \tag{4.5}$$

$$U_0 := \left\{ u \in L^2(\mathcal{O}) : \quad m \le u \le M \quad a.e. \text{ in } \mathcal{O} \right\} .$$

We have $\partial h(u) = 2\alpha u + N_{U_0}(u)$, where $N_{U_0}$ (the normal cone to $U_0$) is given by

$$N_{U_0}(u) = \left\{ v \in L^2(\mathcal{O}) : \quad \begin{array}{l} v(\xi) = 0 \text{ if } u(\xi) \in (m,M)\,; \\ v(\xi) \ge 0 \text{ if } u(\xi) = M\,; \\ v(\xi) \le 0 \text{ if } u(\xi) = m \end{array} \right\} .$$

Then $(\partial h)^{-1}(v) = 2\alpha P_{U_0}(v)$ where $P_{U_0} : L^2(\mathcal{O}) \to U_0$ is the projection operator

$$P_{U_0}(v)(\xi) := \begin{cases} M & \text{if } v(\xi) \ge M \,, \\ m & \text{if } v(\xi) \le m \,, \\ v(\xi) & \text{if } m < v(\xi) < M \,, \end{cases} \qquad \xi \in \mathcal{O} .$$

By by Theorem 3.1, there exists a constant $C^* > 0$ such that if $\alpha T + \lambda < C^*$ the problem (4.1)-(4.3) has a unique solution $(X^*, u^*)$ given by

$$\begin{cases} u^* = 2\alpha P_{U_0}(p^*) \quad \text{in } [0,T] \times \mathcal{O}, \quad \mathbb{P} - a.s. \\ dp^*(t) + \Delta p^*(t) dt - f'(X^*(t)) p^*(t) dt = 2(X^*(t) - X^1) dt + \kappa \sqrt{Q} dW(t) \text{ in } [0,T] \times \mathcal{O} \,, \\ p^*(T) = -2\lambda (X^*(T) - X^0) \quad \text{in } \mathcal{O} \,, \\ p^* = 0 \quad \text{on } [0,T] \times \partial \mathcal{O} \end{cases} . \tag{4.6}$$

# 5 The linear multiplicative noise perturbation

We briefly discuss here the case where the Gaussian perturbation is proportional with the nerve membrane potential. The neuron impulse dynamic is better described by the equation

$$\begin{cases} dX(t) - \Delta X(t)dt + f(X(t))dt = X(t)dW(t)\,dt + F(t)dt \quad \text{in} \quad [0,T]\times\mathcal{O}\,, \\ X(0) = x \quad \text{in} \quad \mathcal{O}\,, \\ X = 0 \quad \text{on} \quad [0,T]\times\partial\mathcal{O}\,, \end{cases} , \qquad (5.1)$$

where

$$W(t,\xi) = \sum_{j=1}^{\infty} \mu_j e_i(\xi)\beta_j(t)\,, \quad t \geq 0\,, \;\; \xi \in \mathcal{O}\,,$$

$\mu_j \in \mathbb{R}$ and $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis in $L^2(\mathcal{O})$ of eigenfunctions for $A$ corresponding to eigenvalues $\lambda_j$.

By the scaling transformation $X = e^W y$ equation (5.1) reduces to the random differential equation (see, e.g. [8])

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y + (\mu - \Delta W)y - 2\nabla W \cdot \nabla y = e^{-W}F \quad \text{in} \quad [0,T]\times\mathcal{O}\,, \\ y(0,\xi) = x(\xi) \;\; \xi \in \mathcal{O}\,, \\ y = 0 \quad \text{on} \quad [0,T]\times\partial\mathcal{O}\,, \end{cases} , \qquad (5.2)$$

where $\mu = \frac{1}{2}\sum_{j=1}^{\infty} \mu_j^2 e_j$.

We shall assume that

$$\sum_{j=1}^{\infty} \mu_j^2 |e_j|_{\infty}^2 < \infty\,. \qquad (5.3)$$

Arguing as in Proposition 2.2 it follows by (5.3) that (5.2) has a unique solution $y$ satisfying (2.13) and this implies that $X = e^W y$ is a strong solution to (5.1) which satisfies condition of Theorem 2.1. We omit the details.

As regards the corresponding optimal control problem P governed by the equation

$$\begin{cases} dX(t) - \Delta_\xi X(t)dt + f(X(t))dt = X(t)dW(t) + Bu(t)dt + f_0 dt\,, \; \text{in} \;\; [0,T]\times\mathcal{O}\,, \\ X(0) = x \; \text{in} \;\; \mathcal{O}\,, X = 0 \; \text{on} \;\; [0,T]\times\partial\mathcal{O}\,, \end{cases} , \qquad (5.4)$$

the existence of an optimal control pair $(X^*, u^*)$ follows as Theorem 3.1 by Eckeland variational principle (3.2) taking however in account that the corresponding dual backward equation (3.6) is in this case

$$\begin{cases} dp_\epsilon + \Delta p_\epsilon dt - f'(X_\epsilon)p_\epsilon dt + \kappa_\epsilon dt = \kappa_\epsilon dW(t) + Dg(X_\epsilon)dt \ \text{ in } [0,T]\times\mathcal{O}\,, \\ p_\epsilon(T) = -Dg_0(X_\epsilon(T)) \ \text{ in } \ \mathcal{O}\,, \\ p_\epsilon = 0 \ \text{ on } \ [0,T]\times\partial\mathcal{O}\,, \end{cases} \tag{5.5}$$

The details are left to the reader.

# References


[1] S. Albeverio and L. Di Persio, Some stochastic dynamical models in neurobiology: recent developments Europena communications in mathematical and theoretical biology, vol. 14 , 2011 , pp. 44-53

[2] S. Albeverio, L. Di Persio and E. Mastrogiacomo. Small noise asymptotic expansions for stochastic PDE's, I. The case of a dissipative polynomially bounded non linearity. Tohoku Mathematical Journal Vol.63, No.4 (2011): 877-898.

[3] S. Albeverio, L. Di Persio and E. Mastrogiacomo. Invariant measures for stochastic differential equations on networks. Proc. Sympos. Pure Math. Vol. 87, 2013.

[4] V. Barbu. Nonlinear differential equations of monotone types in Banach spaces. Springer Monographs in Mathematics. Springer, New York, 2010.

[5] V. Barbu. Analysis and control of nonlinear infinite-dimensional systems. Mathematics in Science and Engineering, 190. Academic Press, Inc., Boston, MA, 1993.

[6] V. Barbu and M. Iannelli. Optimal control of population dynamics. Journal of optimization theory and applications, Vol.102, No.1,(1999) 1-14.

[7] V. Barbu and T. Precupanu. Convexity and optimization in Banach spaces. Springer Science & Business Media, 2012.

[8] V. Barbu and M. Röckner. On a random scaled porous media equation. Journal of Differential Equations 251.9 (2011): 2494-2514.

[9] S. Bonaccorsi, C. Marinelli and G. Ziglio. Stochastic FitzHugh-Nagumo equations on networks with impulsive noise. Electron. J. Probab, Vol.13 (2008): 1362-1379.

[10] E. Casas, C. Ryll and F. Tröltzsch. Sparse optimal control of the Schlögl and FitzHugh–Nagumo systems. Computational Methods in Applied Mathematics Vol.13, No.4 (2013): 415-442.

[11] G. Da Prato. Kolmogorov equations for stochastic PDEs. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser Verlag, Basel, 2004.

[12] G. Da Prato and J. Zabczyk. Second order partial differential equations in Hilbert spaces. London Mathematical Society Lecture Note Series, 293. Cambridge University Press, Cambridge, 2012

[13] G. Da Prato and J. Zabczyk. Ergodicity for infinite-dimensional systems. London Mathematical Society Lecture Note Series, 229. Cambridge University Press, Cambridge, 1996.

[14] L. Di Persio and G. Ziglio. Gaussian estimates on networks with applications to optimal control, (2011).

[15] H.W. Diehl. The theory of boundary critical phenomena. International Journal of Modern Physics B, Vol.11, No.30 (1997): 3503-3523.

[16] I. Ekeland. On the variational principle. Journal of Mathematical Analysis and Applications, Vol.47, No.2 (1974): 324-353.

[17] M. Fuhrman and G. Tessitore. Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. Annals of probability (2002): 1397-1465.

[18] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. Biophysical Journal, Vol.1, (1961): 445–466

[19] K. Kunisch and M. Wagner. Optimal control of the bidomain system (III): Existence of minimizers and first-order optimality conditions. ESAIM: Mathematical Modelling and Numerical Analysis, Vol.47, No.04 (2013): 1077-1106.

[20] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. Proceedings of the Institute of Radio Engineers, Vol.50, No.10 (1962): 2061–2070.

[21] S. Pitchaiah, A. Armaou. Output feedback control of the FitzHugh-Nagumo equation using adaptive model reduction, Proceedings of the 49th IEEE Conference on Decision and Control, Atlanta, GA, 2010.

[22] G. Tessitore. Existence, uniqueness and space regularity of the adapted solutions of a backward SPDE. Stochastic Analysis and Applications, Vol.14, No.4 (1996): 461-486.

[23] H.C. Tuckwell. Random perturbations of the reduced Fitzhugh-Nagumo equation. Physica Scripta, Vol.46, No.6 (1992): 481.