



Optimal Control of the FitzHugh–Nagumo Stochastic Model with Nonlinear Diffusion

Francesco Cordonì¹ · Luca Di Persio¹

Accepted: 28 November 2020 / Published online: 19 January 2021
© The Author(s) 2021

Abstract

We consider the existence and first order conditions of optimality for a stochastic optimal control problem inspired by the celebrated FitzHugh–Nagumo model, with nonlinear diffusion term, perturbed by a linear multiplicative Brownian-type noise. The main novelty of the present paper relies on the application of the *rescaling method* which allows us to reduce the original problem to a random optimal one.

Keywords FitzHugh–Nagumo model · Stochastic process · Optimal control · m -accretive operator · Cauchy problem

Mathematics Subject Classification 47H06 · 60H15 · 91G80 · 93E20

1 Introduction

Consider the following problem

$$\begin{cases} dv(t, \xi) = (\Delta\gamma(v(t, \xi)) - I^{ion}(v(t, \xi)) - f(\xi)v(t, \xi) + F(t, \xi)) dt + v(t, \xi)dW(t), \xi \in \mathcal{O} \\ v(0, \xi) = v_0(\xi), \\ \gamma(v(t, \xi)) = 0, \text{ on } (0, T) \times \partial\mathcal{O}. \end{cases} \quad (1)$$

$\gamma : \mathbb{R} \rightarrow \mathbb{R}$ being a monotone, increasing continuous function, $v = v(t, \xi)$ represents the *transmembrane electrical potential*, $\mathcal{O} \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded and open set with smooth boundary $\partial\mathcal{O}$. We indicate with Δ_ξ the Laplacian operator with respect to the spatial variable ξ , while ε and δ are positive constants representing phenomenological coefficients, $f(\xi)$ is a given external forcing term, while

✉ Francesco Cordonì
francescogiuseppe.cordonì@univr.it

Luca Di Persio
luca.dipersio@univr.it

¹ Department of Computer Science, University of Verona, Strada le Grazie, 15, Verona 37134, Italy

I^{ion} is the *Ionic current* and, according with the FitzHugh–Nagumo model, it equals $I^{ion}(v) = v(v - a)(v - 1)$, $v_0, w_0 \in L^2(\mathcal{O})$, namely it represents a cubic non-linearity. Also F is a bounded term needed to treat the general controlled equation in next section.

Equation (1) with linear diffusion, i.e. $\gamma(x) = x$, is the well-known FitzHugh–Nagumo (FHN) equation. FHN equation is a reaction–diffusion equation, first introduced by Hodgkin and Huxley in [32] and then simplified in [31,35]. The model has been proposed to provide a rigorous, yet simplified, analysis of electrical impulses dynamics along a nerve axon, see, e.g., [38], where the propagation of the transmembrane potential on the nerve axon is represented by a cubic nonlinear reaction term, possibly perturbed by a noisy one, see, e.g., [12,22,38,41].

The random perturbation represents the effect of noisy input currents within neurons, their source being the random opening/closing actions of ion channels, see, e.g., [41]. Moreover, in two-dimensional and three-dimensional settings, Eq. (1) plays also a relevant role in statistical mechanics, under the name of *Ginzburg–Landau* equation, as well as concerning phase transition models of *Ginzburg–Landau* type, see, e.g., [27].

The general case where γ is a monotone function corresponds to an anomalous–diffusive FitzHugh–Nagumo (FHN) equation, see [33], also describing phase transitions in porous media, see, e.g., [36,40].

Remark 1.1 In what follows we shall focus on the mathematical setting behind the Stochastic FitzHugh–Nagumo (FHN) model, without entering into details about the neuro-biological justification of parameters characterizing it. Appropriate details, as well as in depth analysis of the existing literature on the subject, will be provided later.

We assume that W is a $H^{-1} := H^{-1}(\mathcal{O})$ -cylindrical Wiener processes, such that

$$W(t, \xi) = \sum_{n=1}^{\infty} \mu_n e_n \beta_n(t), \quad t \geq 0, \xi \in \mathcal{O},$$

where $\{\beta_n\}_{n \geq 1}$ is a sequence of mutually independent standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, while $\{e_n\}_{n \geq 1}$ is an orthonormal basis in H^{-1} and $\mu_n \in \mathbb{R}$.

Since the Laplacian operator Δ_ξ is a linear operator in $L^2(\mathcal{O})$, and $-\Delta_\xi$ is self-adjoint, then there exists a complete orthonormal system $\{\bar{e}_k\}_{k \geq 1}$ in $L^2(\mathcal{O})$ of eigenfunctions of $-\Delta_\xi$, and we shall indicate the corresponding sequence of eigenvalues denoted by $\{\bar{\lambda}_k\}_{k \geq 1}$. Therefore, we have

$$\Delta_\xi \bar{e}_k = -\bar{\lambda}_k \bar{e}_k, \quad k \in \mathbb{N}.$$

Also, we set

$$Gv = I^{ion}(v) = v(v - a)(v - 1), \tag{2}$$

and note that G is monotonically nondecreasing.

The present paper addresses the problem of existence and uniqueness of a strong solution, in a sense to be better specified in a while, to Eq. (1). We stress that this is

not a trivial problem as the nonlinear operator $\Delta\gamma$ is naturally defined on the space H^{-1} whereas the nonlinear polynomial perturbation I^{ion} is not m -accretive on the same space. In order to solve above problem we will transform the original equation, via a *rescaling transformation*, to a random PDE. It turns out that the existence and uniqueness of transformed random PDE can be treated by the theory of nonlinear semigroup in L^1 .

We will further consider the problem of existence of an optimal control for the nonlinear FHN equation. Again, in order to solve the problem we will apply a *rescaling transformation* to obtain a corresponding random PDE. As already emerged in [12,22], the nonlinear polynomial term implies that standard minimization argument does not apply. Therefore, existence of an optimal control is achieved using *Ekeland’s variational principle*. First order conditions of optimality are given in terms of dual stochastic backward equation, see, e.g. [12,17], whereas, due to the applied *rescaling transformation* are expressed in terms of a random backward dual equation which allows to simplify the setting also giving more insights on the derived optimal controller.

It is worth stressing that the present work continues the investigation of optimal control problem for a stochastic FHN system, generalizing further the result presented in [12,22]. It must be stressed that the techniques used in the present work, although presenting some similarity with [12,22], as for instance the usage of the Ekeland principle to treat the cubic nonlinearity typical of the FHN equation, are in general different. In fact, the nonlinear term γ poses several difficulties since its natural state space is H^{-1} . As already mentioned, the cubic nonlinear term arising in the FHN equation is not sufficiently regular in such a space. Therefore, differently to the previous works [12,22], existence and uniqueness of a solution for the main equation is non trivial. The nonlinear diffusion γ also affects the main technique used in proving the existence of an optimal control since a suitable transformation to reduce the problem to a random equation must be applied.

The present paper is structured as follows: Sect. 1.1 introduces main notation used thorough the paper. Section 2 addresses the problem of proving existence and uniqueness for the state equation whereas in Sect. 3 the problem of the existence for an optimal control is considered.

1.1 Main Notations

In what follows we will denote by $|\cdot|$, resp. $\langle \cdot, \cdot \rangle$, the norm, resp. scalar product, on \mathbb{R}^d . Also, $L^p(\mathcal{O}) =: L^p$, for $1 \leq p \leq \infty$, is the standard space of p -Lebesgue measurable function over the domain $\mathcal{O} \subset \mathbb{R}^d$, with corresponding norm defined as $|\cdot|_p$. For the case $p = 2$, we will further denote by $\langle \cdot, \cdot \rangle_2$ the scalar product in L^2 . The space $H^1(\mathcal{O}) =: H^1$ is the Sobolev space $\{u \in L^2 : \partial_\xi u(\xi) \in L^2\}$, endowed with the standard norm $\|u\|_{H^1}^2 := \int_{\mathcal{O}} (|u|^2 + |\nabla u|^2) d\xi$. The dual of the space H^1 will be denoted as H^{-1} equipped with corresponding norm $|\cdot|_{-1}$.

Similarly, we will denote by $W^{n,p}(\mathcal{O}) =: W^{n,p}$, $n \in \mathbb{N}$, $1 \leq p \leq \infty$, the standard Sobolev space of p -integrable functions with p -integrable n -order derivatives. Coherently, $W^{1,p}([0, T]; H^{-1})$ will be the space of absolutely continuous function

$u : [0, T] \rightarrow H^{-1}$ such that both u and $\frac{d}{dt}u \in L^p([0, T]; H^{-1})$. Further, given a Banach space X , $L^p([0, T]; X)$ is the space of X -valued Bochner p -integrable functions on the interval $[0, T]$. Also, $C([0, T]; X)$, resp. $C^1([0, T]; X)$, denotes the space of continuous, resp. continuously differentiable, functions $u : [0, T] \rightarrow X$.

We shall also introduce $C_W([0, T]; H)$ the space of all H^{-1} -valued (\mathcal{F}_t) -adapted processes such that $X \in C([0, T]; L^2(\Omega; H^{-1}))$, that is X satisfies

$$\sup_{t \in [0, T]} \mathbb{E}|X(t)|_{-1}^2 < \infty.$$

In an analogous manner $L^2_W([0, T]; H^{-1})$ is the space of all H^{-1} -valued (\mathcal{F}_t) -adapted processes such that $X \in L^2([0, T]; L^2(\Omega; H^{-1}))$, that is X satisfies

$$\int_0^T \mathbb{E}|X(t)|_{-1}^2 dt < \infty.$$

At last $L^2_W(\Omega; C([0, T]; H^{-1}))$ denotes the space of all H^{-1} -valued (\mathcal{F}_t) -adapted and continuous processes such that

$$\mathbb{E} \sup_{t \in [0, T]} |X(t)|_{-1}^2 < \infty.$$

Above definition are still in place if instead of H^{-1} we consider a general Hilbert space H . It is also known that there is a natural embedding of $L^2_W(\Omega; C([0, T]; H^{-1}))$ into the space $C_W([0, T]; H^{-1})$, see, e.g. [13, Chapter 1].

We therefore can rewrite Eq. (1) as

$$\begin{cases} dX(t) - [\Delta(\gamma(X(t))) - G(X(t)) - fX(t) + F]dt = XdW(t), \\ X(0) = x_0 \in H^{-1}, \quad t \in [0, T], \end{cases} \quad (3)$$

We will assume the following to hold.

Hypothesis 1 (i) $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ with $\gamma(0) = 0$ is a continuous and differentiable function such that

$$0 < C_1 \leq \gamma'(x) \leq C_2 < \infty, \quad \forall x \in \mathbb{R};$$

- (ii) $G : \mathbb{R} \rightarrow \mathbb{R}$ is continuous monotonically non-decreasing and $G(0) = 0$; moreover G is locally Lipschitzian;
- (iii) $F \in L^\infty((0, T) \times \mathcal{O})$, \mathbb{P} -a.s. and it is progressively measurable w.r.t. $(0, T) \times \Omega \times \mathcal{B}(\mathcal{O})$; $f \in L^\infty(\mathcal{O})$, and $f \geq 0$ a.e. in \mathcal{O} ;
- (iv) W is a $H^{-1} := H^{-1}(\mathcal{O})$ -cylindrical Wiener processes, that is,

$$W = \sum_{j=1}^{\infty} \mu_j e_j \beta_j,$$

with

$$\sum_{j=1}^{\infty} \mu_j^2 |e_j|_{L^\infty(\mathcal{O})}^2 < \infty,$$

see, [13, p. 22].

Then, we can state the notion of solution to Eq. (3) that we will consider in subsequent analysis, see [13, p. 50].

Definition 1.1 Let $x \in H^{-1}$, we say that the process

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

is a solution to (3) if $X(t) : [0, T] \rightarrow H^{-1}$ is \mathbb{P} -a.s. continuous and $\forall t \in [0, T]$

$$X(t) = x + \int_0^t (\Delta(\gamma(X(s))) - G(X(s)) + fX(s) + F(s)) ds + \int_0^t X(s) dW(s).$$

2 Existence for the State Equation

The main problem in proving existence and uniqueness for a solution to Eq. (3) is that the operator G is not m -accretive on the space H^{-1} and so basic existence results in [11,13] are not applicable in the present case. It turns out that the proper space one has to consider to successfully treat Eq. (3) is the space L^1 , which, in turn, is not the proper one if one has to deal with SPDEs such as (3).

To overcome such a stalemate, we follow [9,10]. In particular, we apply the transformation $X = e^W y$, which allows to reduce the stochastic equation (3) to a random PDE that can be treated with analytical techniques. In fact, the random equation can be successfully solved by exploiting the theory of nonlinear semigroup in L^1 . As noted in [10], we have still to face the problem that, because of the non regularity of the term W , the general theory cannot be applied straightforward to the resulting random PDE. Therefore, for $\epsilon > 0$, we shall consider a suitable sequence of regular approximations W_ϵ of W , to first establish a priori estimates for solutions y_ϵ of the associated W_ϵ -approximating problem, and then to show that, in the limit $\epsilon \rightarrow 0$, we obtain both existence and uniqueness of the solution for the original equation.

The following theorem constitutes the main result existence of this section.

Theorem 2.1 *Let $x \in H^{-1} \cap L^1$ with $\gamma(x) \in H^1_0$, then there is a unique strong solution to Eq. (3) $X = e^W y$ which satisfies*

$$Xe^{-W} \in W^{1,2}([0, T]; H^{-1}) \cap L^\infty([0, T] \times \mathcal{O}), \quad \mathbb{P} - a.s.$$

In order to prove Theorem 2.1 we need some auxiliary lemmas. In particular, let us then introduce the transformation

$$X(t) = e^{W(t)} y(t), \quad t \geq 0, \tag{4}$$

so that by an application of the Itô formula we obtain the random equation

$$\begin{aligned} \frac{\partial}{\partial t} y + e^{-W} G(e^W y) - e^{-W} \Delta \gamma(e^W y) + fy + \mu y &= e^{-W} F \\ y(0, \xi) &= x(\xi), \quad \xi \in \mathcal{O}, \\ y(t) &\in H_0^1(\mathcal{O}), \quad t \in (0, T) \end{aligned} \tag{5}$$

with

$$\mu = \frac{1}{2} \sum_{n=1}^{\infty} \mu_n^2 e_n^2,$$

see, e.g. [7,9,10].

Following [10], we prove the existence of a unique strong solution to Eq. (5) by first considering an approximating problem. In particular, let us denote by $\beta^\epsilon(t) := (\beta * \rho_\epsilon)(t)$, where $\rho_\epsilon(t) = \frac{1}{\epsilon} \rho\left(\frac{t}{\epsilon}\right)$ is a standard mollifier and $\rho \in C_0^\infty$, then we have that $\beta^\epsilon \in C^1([0, T]; \mathbb{R})$. Setting

$$W_\epsilon(t, \xi) = \sum_{n=1}^{\infty} \mu_n e_n \beta^\epsilon(t), \quad t \geq 0, \xi \in \mathcal{O}.$$

we thus have that $W_\epsilon \in C^1([0, T] \times \mathcal{O})$. Moreover

$$W_\epsilon(t, \xi) \rightarrow W(t, \xi) \quad \text{uniformly in } (t, \xi) \in [0, T] \times \xi,$$

as $\epsilon \rightarrow 0$.

For each $\epsilon > 0$, let us thus consider the approximating equation associated to (5)

$$\begin{aligned} \frac{\partial}{\partial t} y_\epsilon + e^{-W_\epsilon} G_\epsilon(e^{W_\epsilon} y_\epsilon) - e^{-W_\epsilon} \Delta \left(\gamma(e^{W_\epsilon} y_\epsilon + \epsilon e^{W_\epsilon} y_\epsilon) \right) + fy + \mu y_\epsilon &= e^{-W_\epsilon} F \\ y_\epsilon(0, \xi) &= x(\xi), \quad \xi \in \mathcal{O}, \\ y_\epsilon(t) &\in H_0^1(\mathcal{O}), \quad t \in (0, T) \end{aligned} \tag{6}$$

where G_ϵ is the Yosida approximation of G , that is

$$G_\epsilon := \frac{1}{\epsilon} \left(I - (I + \epsilon G)^{-1} \right), \quad \epsilon > 0. \tag{7}$$

Note that, G_ϵ is monotonically non-decreasing, Lipschitzian and

$$\lim_{\epsilon \rightarrow 0} G_\epsilon(z) = G(z), \quad \forall z \in \mathbb{R}$$

uniformly on compacts.

Defining $z_\epsilon := e^{W_\epsilon} y_\epsilon$, Eq. (6) becomes

$$\begin{aligned} \frac{\partial}{\partial t} z_\epsilon + G_\epsilon(z_\epsilon) - \Delta(\gamma(z_\epsilon) + \epsilon z_\epsilon) + fz + \left(\mu - \frac{\partial}{\partial t} W_\epsilon\right) z_\epsilon &= F_\epsilon, \quad \text{in } (0, T) \times \mathcal{O}, \\ z_\epsilon(0, \xi) &= x(\xi), \quad \xi \in \mathcal{O}, \\ \gamma(z_\epsilon(t)) + \epsilon z_\epsilon(t) &\in H_0^1(\mathcal{O}), \quad t \in (0, T) \end{aligned} \tag{8}$$

where $F_\epsilon := e^{-W_\epsilon} F$.

Lemma 2.2 *Let $x \in H^{-1} \cap L^1$ with $\gamma(x) \in H_0^1$, then for each $\epsilon > 0$ Eq. (6) has a unique solution such that*

$$y_\epsilon \in W^{1,\infty}([0, T]; H^{-1}) \cap L^\infty(0, T; H_0^1).$$

Proof Let us first prove existence and uniqueness of a solution to Eq. (8) in the space H^{-1} . For a fixed $\epsilon > 0$, let us define the operator $A : D(A) \subset H^{-1} \rightarrow H^{-1}$ as

$$\begin{aligned} Az &= -\Delta(\gamma(z) + \epsilon z) + fz + G_\epsilon(z) + \mu z, \\ D(A) &= \left\{ z \in L^2 : \gamma(z) \in H_0^1 \right\}, \end{aligned} \tag{9}$$

We equip the space H^{-1} with the scalar product

$$\langle y, z \rangle_{-1} := {}_{H^1} \langle (-\Delta)^{-1} y, z \rangle_{H^{-1}},$$

where $(-\Delta)^{-1} y = x$ indicates the solution to the Dirichlet problem $-\Delta x = y$ in \mathcal{O} , $x \in H_0^1$.

Taking into account that G_ϵ is Lipschitz continuous in L^2 and since

$$z \mapsto -\Delta(\gamma(z) + \epsilon z),$$

is m -accretive in the space H^{-1} , see, e.g., [4, p. 68], we have that, for a suitable $\alpha = \alpha_\epsilon$, it holds

$$\langle (A + \alpha I)z - (A + \alpha I)\bar{z}, z - \bar{z} \rangle_{-1} \geq 0$$

which implies $(A + \alpha I)$ to be accretive in H^{-1} .

Moreover, for $\lambda > 0$ sufficiently large, we also have $\mathcal{R}((\lambda + \alpha)I + A) = H^{-1}$, so that A is quasi- m -accretive. In other words, for $f \in H^{-1}$ the equation

$$(\lambda + \alpha)(-\Delta)^{-1} z + \gamma(z) + \epsilon z + (-\Delta)^{-1}(G_\epsilon(z) + fz - \mu z) = (-\Delta)^{-1} \tilde{f}, \tag{10}$$

has a unique solution in $z \in L^2$. Indeed, introducing the operators

$$B : L^2 \rightarrow L^2, \quad Bz := \gamma(z),$$

and

$$\begin{aligned} \Gamma &: L^2 \rightarrow L^2, \\ \Gamma z &= (\lambda + \alpha) (-\Delta)^{-1} z + (-\Delta)^{-1} (G_\epsilon(z) + fz - \mu z), \end{aligned}$$

we see that Eq. (10) can be rewritten as

$$\epsilon z + Bz + \Gamma z = (-\Delta)^{-1} \tilde{f}. \tag{11}$$

Since B is m -accretive and Γ is m -accretive and continuous in L^2 , it follows, see, e.g., [4, p. 104], that $\mathcal{R}(\epsilon I + B + \Gamma) = L^2$, so that Eq. (11) admits a unique solution z in L^2 . Moreover, since $\gamma(z) + \epsilon z \in H^1$ and the inverse map of $z \mapsto \gamma(z) + \epsilon z$ is Lipschitz, then $z \in D(A)$. It follows that, applying [10, Lemma A.1, Corollary A.2], see also [4, Sect. 4], z_ϵ is a strong solution to Eq. (8) in $W^{1;\infty}([0, T]; H^{-1})$. In addition, by [10, Corollary A.2], we also have

$$\gamma(z_\epsilon) + \epsilon z_\epsilon - (-\Delta)^{-1} G_\epsilon(z) \in L^\infty(0, T; H_0^1),$$

and

$$\left| (-\Delta)^{-1} G_\epsilon(z) \right|_2 \leq C_\epsilon |z|_{-1},$$

so that, since $z_\epsilon \in W^{1;\infty}([0, T]; H^{-1})$, we obtain

$$\gamma(z_\epsilon) + \epsilon z_\epsilon \in L^\infty(0, T; L^2),$$

and consequently $z_\epsilon \in L^\infty(0, T; L^2)$. Moreover we have that

$$(-\Delta)^{-1} G_\epsilon(z_\epsilon) \in L^\infty(0, T; H_0^1),$$

which implies that $\gamma(z_\epsilon) + \epsilon z_\epsilon \in L^\infty(0, T; H_0^1)$ and consequently $z_\epsilon \in L^\infty(0, T; H_0^1)$. □

Lemma 2.3 *Taking $x \in D(A)$, then $y_\epsilon \in L^\infty((0, T) \times \mathcal{O})$, and it holds*

$$\sup_\epsilon \{ \|y_\epsilon\|_{L^\infty((0,T) \times \mathcal{O})} \} \leq C(1 + \|x\|_\infty). \tag{12}$$

Proof Let $\alpha \in C^1([0, T])$, such that $\alpha(0) = 0$ and $\alpha' \geq 0$. Then, defining $M := (1 + \|x\|_\infty)$, we have

$$\begin{aligned} \frac{\partial}{\partial t} & (y_\epsilon - M - \alpha(t)) + e^{-W_\epsilon} \left(G_\epsilon \left(e^{W_\epsilon} (y_\epsilon) \right) - G_\epsilon \left(e^{W_\epsilon} (M + \alpha(t)) \right) \right) \\ & - f(y_\epsilon - M - \alpha(t)) - e^{-W_\epsilon} \Delta \left(\gamma(e^{W_\epsilon} y_\epsilon) + \epsilon e^{W_\epsilon} y_\epsilon \right) \\ & + e^{-W_\epsilon} \Delta \left(\gamma \left(e^{W_\epsilon} (M + \alpha(t)) \right) + \epsilon e^{-W_\epsilon} (M + \alpha(t)) \right) \\ & + \mu(y_\epsilon - M - \alpha(t)) = \tilde{F}_\epsilon - \alpha', \end{aligned}$$

with

$$\begin{aligned} \tilde{F}_\epsilon := & -e^{-W_\epsilon} G_\epsilon(e^{W_\epsilon}(M + \alpha(t))) - \mu(M + \alpha(t)) - f(M + \alpha(t)) \\ & + e^{-W_\epsilon} \Delta \gamma \left(e^{W_\epsilon}(M + \alpha(t)) \right) + \epsilon(M + \alpha(t))e^{-W_\epsilon} \Delta(e^{W_\epsilon}) + e^{-W_\epsilon} F, \end{aligned}$$

with α such that $F_\epsilon - \alpha' \leq 0$.

Following [10, Lemma 3.3], we first assume that

$$\frac{\partial}{\partial t} y_\epsilon, \quad \Delta \left(\gamma(e^{W_\epsilon} y_\epsilon) + \epsilon e^{W_\epsilon} y_\epsilon \right) \in L^1((0, T) \times \mathcal{O}), \tag{13}$$

then, denoting by

$$\begin{aligned} J(t) := & - \int_{\mathcal{O}} e^{-W_\epsilon} \left[\left(\Delta(\gamma(e^{W_\epsilon} y_\epsilon) + \epsilon e^{W_\epsilon} y_\epsilon) \right) \right. \\ & \left. - \Delta \left(\gamma \left(e^{W_\epsilon}(M + \alpha(t)) \right) + \epsilon e^{W_\epsilon}(M + \alpha(t)) \right) \right] \text{sign} (y_\epsilon - M - \alpha(t))^+ d\xi, \end{aligned}$$

we have

$$\int_0^t J(s) ds \geq -(\gamma'(e^{|W|_\infty} M) + 1)e^{|W|_\infty} (|\Delta W|_\infty + |\nabla W|_\infty^2) \int_0^t |(y_\epsilon - (M + \alpha(s)))^+|_1 ds.$$

Moreover, by Hypothesis 1, it follows that G_ϵ is monotone, so that

$$\begin{aligned} & \int_0^t \int_{\mathcal{O}} e^{-W_\epsilon} \left(G_\epsilon(e^{W_\epsilon}(y_\epsilon)) - G_\epsilon(e^{W_\epsilon}(M + \alpha(t))) \right) \text{sign}(e^{-W_\epsilon} (y_\epsilon - M - \alpha(s)))^+ ds d\xi \\ & = \int_0^t \int_{\mathcal{O}} e^{-W_\epsilon} \left(G_\epsilon \left(e^{W_\epsilon}(y_\epsilon) \right) - G_\epsilon \left(e^{W_\epsilon}(M + \alpha(t)) \right) \right) \text{sign} \left(G_\epsilon(e^{-W_\epsilon} y_\epsilon) \right. \\ & \quad \left. - G_\epsilon \left(e^{-W_\epsilon}(M + \alpha(s)) \right) \right)^+ ds d\xi \geq 0, \end{aligned}$$

and since

$$\begin{aligned} & \int_{\mathcal{O}} \frac{\partial}{\partial t} (y_\epsilon - M - \alpha(s)) \text{sign} (y_\epsilon - M - \alpha(s))^+ d\xi \\ & = \frac{d}{dt} |(y_\epsilon(t) - M - \alpha(t))^+|_1, \text{ a.e. } t \in (0, T), \end{aligned}$$

by [10, Lemma 3.3], we conclude that

$$|(y_\epsilon(t) - M - \alpha(t))^+|_1 = 0,$$

if $\tilde{F}_\epsilon \leq \alpha'$ a.e. in $(0, T) \times \mathcal{O}$. Moreover, for a suitable α , it also holds

$$y_\epsilon \leq M + \alpha(t), \quad \text{a.e. in } (0, T) \times \mathcal{O},$$

and

$$y_\epsilon \geq -M - \alpha(t), \quad \text{a.e. in } (0, T) \times \mathcal{O},$$

and inequalities (12) follows.

Using the approximating scheme described in [10, Lemma 3.3], we have (12) without requiring the condition (13), and the claim follows. \square

Lemma 2.4 *Let $x \in D(A)$, then there exists an increasing function $C : [0, \infty) \rightarrow (0, \infty)$ such that*

$$\sup_{t \in [0, T]} |y_\epsilon(t)|_2^2 + \int_0^T \int_{\mathcal{O}} |\nabla \gamma(e^{W_\epsilon} y_\epsilon)|^2 d\xi ds \leq C(1 + |x|_\infty), \quad \forall \epsilon \in (0, 1].$$

Proof In what follows we will use the following

$$\begin{aligned} \int_{\mathcal{O}} j(y_\epsilon(t)) d\xi &= \int_0^t \left\langle H^{-1} \left\langle \frac{dy_\epsilon}{ds}(s), \gamma(y_\epsilon(s)) \right\rangle_{H_0^1}, ds \right\rangle + \int_{\mathcal{O}} j(x) d\xi, \\ H^{-1} \left\langle \Delta \gamma(e^{W_\epsilon} y_\epsilon), e^{-W_\epsilon} \gamma(y_\epsilon) \right\rangle &= - \int_{\mathcal{O}} \nabla \gamma(e^{W_\epsilon} y_\epsilon) \cdot \nabla (e^{-W_\epsilon} \gamma(y_\epsilon)) d\xi, \end{aligned}$$

with $j(r) = \int_0^r \gamma(s) ds, r \in \mathbb{R}^+$.

Thus, multiplying Eq. (6) by $\gamma(y_\epsilon)$ and integrating over $(0, t) \times \mathcal{O}$ we obtain

$$\begin{aligned} \int_{\mathcal{O}} j(y_\epsilon(t)) d\xi + \int_0^t \int_{\mathcal{O}} \left[\nabla \gamma(e^{W_\epsilon} y_\epsilon) + \epsilon \nabla (e^{W_\epsilon} y_\epsilon) \right] \cdot \nabla (\gamma(y_\epsilon) e^{-W_\epsilon}) d\xi ds \\ \leq \int_{\mathcal{O}} j(x) d\xi - \int_0^t \int_{\mathcal{O}} \left(e^{-W_\epsilon} G_\epsilon(e^{W_\epsilon} y_\epsilon) + f y_\epsilon - e^{-W_\epsilon} F \right) \gamma(y_\epsilon) d\xi ds. \end{aligned} \tag{14}$$

Concerning the last integral in the right hand side of Eq. (14), using Assumption 1 (i) on γ we obtain

$$\begin{aligned} \int_0^t \int_{\mathcal{O}} \left(e^{-W_\epsilon} G_\epsilon(e^{W_\epsilon} y_\epsilon) + f y_\epsilon \right) \gamma(y_\epsilon) d\xi ds \\ \leq C \int_0^t \int_{\mathcal{O}} \left(e^{-W_\epsilon} G_\epsilon(e^{W_\epsilon} y_\epsilon) + f y_\epsilon \right) y_\epsilon d\xi ds. \end{aligned} \tag{15}$$

Using estimate (15) it follows, recalling that G_ϵ is the Yosida approximant of G and using the monotonicity of γ and G_ϵ that

$$\begin{aligned}
 & \int_0^t \int_{\mathcal{O}} \left(e^{-W_\epsilon} G_\epsilon \left(e^{W_\epsilon} y_\epsilon \right) + f y_\epsilon \right) \gamma \left(y_\epsilon \right) d\xi ds \\
 & \geq C \int_0^t \int_{\mathcal{O}} e^{-W_\epsilon} G_\epsilon \left(e^{W_\epsilon} y_\epsilon \right) y_\epsilon d\xi ds \\
 & \geq -C \int_0^t \int_{\mathcal{O}} |f| |y_\epsilon|^2 d\xi ds .
 \end{aligned} \tag{16}$$

From the boundedness on F and assumption on γ in hypothesis 1 (i) – (ii), we obtain

$$\int_0^t \int_{\mathcal{O}} F \gamma \left(y_\epsilon \right) d\xi ds \leq C(1 + |x|_\infty^2) , \tag{17}$$

for a positive constant C independent of ϵ .

The other terms in Eq. (14) can be studied as done in [10, Lemma 3.3], so that the claim follows by Lemma 2.3. □

Lemma 2.5 *There is a unique solution to Eq. (5) with*

$$\begin{aligned}
 y & \in W^{1,2} \left([0, T]; H^{-1} \right) \cap L^\infty \left((0, T) \times \mathcal{O} \right) , \\
 \gamma \left(e^W y \right) & \in L^2 \left(0, T; H_0^1 \right) .
 \end{aligned} \tag{18}$$

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

Proof Let us first prove uniqueness. Let y_1 and y_2 be two solutions to Eq. (5), and let $\bar{y} := y_1 - y_2$. Then it holds

$$\begin{aligned}
 & \frac{\partial}{\partial t} \bar{y} + e^{-W} \left(G \left(e^W y_1 \right) - G \left(e^W y_2 \right) \right) + f \bar{y} \\
 & \quad - e^{-W} \Delta \left(\gamma \left(e^W y_1 \right) - \gamma \left(e^W y_2 \right) \right) + \mu \bar{y} = 0 \quad \text{in } (0, T) \times \mathcal{O} \\
 & \bar{y}(0, \xi) = 0, \quad \xi \in \mathcal{O} .
 \end{aligned} \tag{19}$$

We can rewrite Eq. (19) as

$$\begin{aligned}
 & \frac{\partial}{\partial t} \bar{y} + (-\Delta)(\bar{y}\eta) = -e^{-W} \left(G \left(e^W y_1 \right) - G \left(e^W y_2 \right) \right) - e^W \Delta \left(e^{-W} \right) \bar{y}\eta \\
 & \quad - f \bar{y} - 2\nabla \left(e^{-W} \right) \cdot \nabla \left(e^W \bar{y}\eta \right) - \mu \bar{y} = 0 ,
 \end{aligned} \tag{20}$$

where we have denoted for short

$$\eta := \begin{cases} \frac{\left(\gamma \left(e^W y_1 \right) - \gamma \left(e^W y_2 \right) \right)}{e^W \bar{y}} \{ (t, \xi) : \bar{y}(t, \xi) \neq 0 \} , \\ 0 \{ (t, \xi) : \bar{y}(t, \xi) = 0 \} . \end{cases}$$

Multiplying Eq. (20) by $(-\Delta)^{-1}\bar{y}$, we obtain

$$\begin{aligned} & \frac{1}{2}|\bar{y}|_{-1}^2 + \int_0^t \int_{\mathcal{O}} \eta \bar{y}^2 ds d\xi \\ &= \int_0^t \int_{\mathcal{O}} e^{-W} \left(G(e^W y_1) - G(e^W y_2) \right) (-\Delta)^{-1} \bar{y} ds d\xi \\ & \quad - \int_0^t \int_{\mathcal{O}} e^W \Delta(e^{-W}) \bar{y} \eta (-\Delta)^{-1} \bar{y} ds d\xi \\ & \quad - 2 \int_0^t \int_{\mathcal{O}} \nabla(e^{-W}) \cdot \nabla(e^W \bar{y} \eta) (-\Delta)^{-1} \bar{y} ds d\xi \\ & \quad - \int_0^t \int_{\mathcal{O}} f \bar{y} (-\Delta)^{-1} \bar{y} ds d\xi - \int_0^t \int_{\mathcal{O}} \mu \bar{y} (-\Delta)^{-1} \bar{y} ds d\xi. \end{aligned} \tag{21}$$

Concerning the first integral in the right hand side of Eq. (21), notice that, for $\alpha \in [0, 1]$ it holds

$$G(e^W y_1) - G(e^W y_2) = G'(\alpha e^W y_1 + (1 - \alpha)e^W y_2) e^W \bar{y}.$$

Moreover, since G is locally Lipschitz, we have

$$\left| \int_{\mathcal{O}} e^{-W} \left(G(e^W y_1) - G(e^W y_2) \right) (-\Delta)^{-1} \bar{y} d\xi \right| \leq C |\bar{y}|_2 |\bar{y}|_{-1},$$

whereas other terms can be treated as in [10, Theorem 2.2]. So that, we have

$$\frac{d}{dt} |\bar{y}|_{-1}^2 \leq C |\bar{y}|_{-1}^2, \text{ a.e. } t > 0,$$

from which it follows that $\bar{y} = 0$, and, by Lemma 2.4, it holds

$$|y(t)|_{\infty} + \int_0^t \int_{\mathcal{O}} |\nabla \gamma(y(s))|^2 d\xi ds \leq C(1 + |x|_{\infty}),$$

so that, see [10, Theorem 2.2], we further have

$$y \in W^{1,2}([0, T]; H^{-1}) \cap L^{\infty}((0, T) \times \mathcal{O}).$$

As regard existence, by Lemmas 2.3 and 2.4, we have that $(\gamma(e^{W_{\epsilon}} y_{\epsilon}))$ is bounded in $L^2(0, T; H_0^1)$, (y_{ϵ}) is bounded in $L^{\infty}(0, T; L^2) \cap L^{\infty}((0, T) \times \mathcal{O}) \cap L^2(0, T; H_0^1)$, and $\left(\frac{dy_{\epsilon}}{dt}\right)$ is bounded in $L^2(0, T; H^{-1})$. Thus, by Aubin compactness theorem, (y_{ϵ}) is compact in each $L^2(0, T; L^2(\mathcal{O}))$. It follows that, for fixed $\omega \in \Omega$, along a sub-

quence, which we still denote by $\{\epsilon\} \rightarrow 0$ for the sake of clarity, we have

$$\begin{aligned}
 y_\epsilon &\rightarrow y \text{ strongly in } L^2((0, T); L^2), \\
 &\text{weak-star in } L^2((0, T); L^2), \\
 &\text{strongly in } L^\infty((0, T) \times \mathcal{O}), \\
 &\text{weakly in } L^2((0, T); H_0^1), \\
 \gamma(e^{W_\epsilon} y_\epsilon) &\rightarrow \eta \text{ weakly in } L^2((0, T); H_0^1), \\
 \frac{dy_\epsilon}{dt} &\rightarrow \frac{dy}{dt} \text{ weakly in } L^2((0, T); H^{-1}), \\
 W_\epsilon &\rightarrow W \text{ in } C((0, T) \times \mathcal{O}).
 \end{aligned}
 \tag{22}$$

Since the map $z \mapsto \gamma(z)$ is maximal monotone, by (22) we have that $\eta = \gamma(e^W y)$. Then, since it holds

$$|(1 + \epsilon G)^{-1}(e^{W_\epsilon} y_\epsilon) - e^{W_\epsilon} y_\epsilon| \leq \epsilon |G_\epsilon(e^{W_\epsilon} y_\epsilon)| \leq C\epsilon \text{ a.e. in } (0, T) \times \mathcal{O},$$

and

$$(1 + \epsilon G)^{-1}(e^{W_\epsilon} y_\epsilon) \rightarrow y \text{ strongly in } L^2((0, T); L^2) \text{ and a.e. in } (0, T) \times \mathcal{O},$$

then, for $\epsilon \rightarrow 0$, we get

$$G_\epsilon(e^{W_\epsilon} y_\epsilon) \rightarrow G(y) \text{ weakly in } L^2((0, T); H_0^1) \text{ and a.e. in } (0, T) \times \mathcal{O}. \tag{23}$$

Thus, again from the fact that $G : \mathbb{R} \rightarrow \mathbb{R}$ is maximal monotone it follows that it is also closed and therefore we have that $\zeta = G(e^W y)$.

Therefore, by letting $\epsilon \rightarrow 0$, from Eq. (6) we obtain

$$\begin{aligned}
 \frac{dy}{dt} + e^{-W} G(e^W y) + fy - e^{-W} \Delta \gamma(e^W y) + \mu y &= e^{-W} F, \text{ in } (0, T) \times \mathcal{O}, \\
 y(0) &= x.
 \end{aligned}$$

Then, by the uniqueness result already proved, we also have that the sequence (y_ϵ) is independent of $\omega \in \Omega$, implying that y is (\mathcal{F}_t) -adapted, ending the proof. \square

We can finally prove that it exists a unique strong solution X to Eq. (3) which satisfies

$$Xe^{-W} \in W^{1,2}([0, T]; H^{-1}) \text{ } \mathbb{P} - a.s.$$

Proof of Theorem 2.1 Using [9, Lemma 8.1] we have the equivalence between the stochastic PDE 3 and the random PDE 5 via the rescaling transformation 4, so that existence and uniqueness of a solution X in the sense of Definition 1.1 follows by Lemma 2.5. \square

3 The Optimal Control Problem

In this section we will focus the attention to a controlled version of Eq. (1). We denote by $\mathcal{X} = L^2_{ad}((0, T) \times \mathcal{O})$ the space of all \mathcal{F}_t -adapted processes $u : [0, T] \rightarrow \mathbb{R}^d$, and we consider the following optimal control problem

$$\text{Minimize } \mathbb{E} \left[\int_0^T \int_{\mathcal{O}} |v(t, \xi) - v_1(\xi)|^2 + \frac{\alpha}{2} |u(t, \xi)|^2 d\xi dt + \int_{\mathcal{O}} |v(T, \xi) - v_2(\xi)|^2 d\xi \right], \tag{P}$$

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

$$\begin{cases} \partial_t v(t, \xi) - \Delta \gamma(v(t, \xi)) + I^{ion}(v(t, \xi)) + f(\xi)v(t, \xi) = u(t, \xi) + v(t, \xi)\partial_t W(t), & \text{in } (0, T) \times \mathcal{O} \\ v(0, \xi) = v_0(\xi), \quad \xi \in \mathcal{O}, \\ v(t, \xi) = 0, & \text{on } (0, T) \times \partial\mathcal{O}. \end{cases} \tag{24}$$

Here

$$\begin{aligned} \mathcal{U} := & \left\{ u \in L^2_{ad}((0, T) \times \mathcal{O} \times \Omega) : \right. \\ & \left. |u(t, \xi, \omega)| \leq M \quad \text{a.e. } (t, \xi, \omega) \in (0, T) \times \mathcal{O} \times \Omega \right\}, \end{aligned}$$

$M > 0$ being a suitable constant, while $v_1, v_2 \in L^2(\Omega)$ \mathcal{F}_0 -adapted and $\alpha > 0$ are given.

In what follows we are going to treat the problem (P) by a *rescaling procedure* which allows us to reduce it to a random optimal control problem. In the current section the following Hypothesis will be assumed to hold.

Hypothesis 2 (i) $G \in C^1(\mathbb{R})$, G' is locally Lipschitz.

Theorem 3.1 *Let Hypotheses 1 and 2 hold, then, for T sufficiently small, there exists at least one optimal pair (u^*, v^*) solution to problem (P).*

Proof As in Sect. 2, we will apply the rescaling transformation $y := e^{-W}v$ so that the optimal control problem (P) reads,

$$\begin{aligned} \text{Minimize } \mathbb{E} \left[\int_0^T \int_{\mathcal{O}} \left| e^W y(t, \xi) - v_1(\xi) \right|^2 + \frac{\alpha}{2} |u(t, \xi)|^2 d\xi dt \right] \\ + \mathbb{E} \left[\int_{\mathcal{O}} \left| e^W y(T, \xi) - v_2(\xi) \right|^2 d\xi \right], \end{aligned} \tag{P2}$$

subject to

$$\begin{aligned} \partial_t y - e^{-W} \Delta \gamma(e^W y) + e^{-W} G(e^W y) + fy + \mu y = e^{-W} u, & \text{ in } (0, T) \times \mathcal{O}, \\ y = 0, & \text{ in } (0, T) \times \partial\mathcal{O}. \end{aligned} \tag{25}$$

Due to the cubic nonlinear term G , existence and uniqueness of an optimal control cannot be established by standard minimization arguments. In order to overcome this problem Ekeland’s variational principle can be exploited in order to obtain a nearly optimal solution to the above control problem.

Applying Ekeland’s variational principle, see, e.g. [28] or also [12,22], there exists a sequence $\{u_\epsilon\} \subset \mathcal{U}$ such that, defining for short,

$$\Psi(u) := \mathbb{E} \left[\int_0^T \int_{\mathcal{O}} |e^W y(t, \xi) - v_1(\xi)|^2 + \frac{\alpha}{2} |u(t, \xi)|^2 d\xi dt \right] + \mathbb{E} \left[\int_{\mathcal{O}} |e^W y((T, \xi) - v_2(\xi)|^2 d\xi \right],$$

it follows

$$\begin{aligned} \Psi(u_\epsilon) &\leq \inf \{ \Psi(u) ; u \in \mathcal{U} \} + \epsilon, \\ \Psi(u_\epsilon) &\leq \Psi(u) + \sqrt{\epsilon} |u_\epsilon - u|, \quad \forall u \in \mathcal{U}, \end{aligned} \tag{26}$$

or equivalently it holds

$$u_\epsilon = \arg \min_{u \in \mathcal{U}} \{ \Psi(u) + \sqrt{\epsilon} |u_\epsilon - u| \}. \tag{27}$$

In particular, the optimal pair (y_ϵ, u_ϵ) solves

$$\begin{aligned} \text{Minimize } &\mathbb{E} \left[\int_0^T \int_{\mathcal{O}} |e^W y(t, \xi) - v_1(\xi)|^2 + \frac{\alpha}{2} |u(t, \xi)|^2 d\xi dt \right] \\ &+ \mathbb{E} \left[\int_{\mathcal{O}} |e^W y((T, \xi) - v_2(\xi)|^2 d\xi \right] + \sqrt{\epsilon} |u_\epsilon - u|_{\mathcal{U}}. \end{aligned} \tag{28}$$

A variational argument alike the one carried out in [12,22], allow us to associate the the optimal control problem (26) the dual backward problem

$$\begin{cases} \partial_t y_\epsilon - e^{-W} \Delta \gamma(e^W y_\epsilon) + e^{-W} G(e^W y_\epsilon) + f y_\epsilon + \mu y_\epsilon = \frac{1}{\alpha} e^{-W} (p_\epsilon + \theta_\epsilon) \\ \partial_t p_\epsilon + e^W \gamma'(e^W y_\epsilon) \Delta (e^{-W} p_\epsilon) + e^W G'(e^W y_\epsilon) p_\epsilon - f p_\epsilon - \mu p_\epsilon = 2(e^W y_\epsilon - v_1) \\ y_\epsilon(0) = y_0, \quad p_\epsilon(T) = 2(e^W y_\epsilon(T) - v_2), \end{cases} \tag{29}$$

where $|\theta_\epsilon|_{L^2(\Omega \times \mathcal{O} \times (0, T))} \leq \sqrt{\epsilon}$. Indeed, by (26), it follows $\Psi'(u_\epsilon) = \theta_\epsilon$, yielding (29). Existence and uniqueness for a solution to Eq. (25) follows from Lemma 2.5 whereas standard arguments yield, using assumption 2, that it exists a unique solution $p_\epsilon \in L^\infty((0, T) \times \mathcal{O} \times \Omega)$, see [5].

By Eq. (29) we have $\mathbb{P} - a.s.$,

$$\begin{aligned} \partial_t (y_\epsilon - y_\lambda) - e^{-W} \Delta \left(\gamma(e^W y_\epsilon) - \gamma(e^W y_\lambda) \right) \\ - e^{-W} \left(G(e^W y_\epsilon) - G(e^W y_\lambda) \right) + f(y_\epsilon - y_\lambda) + \mu(y_\epsilon - y_\lambda) = \quad (30) \\ = \frac{1}{\alpha} e^{-W} ((p_\epsilon - p_\lambda) + (\theta_\epsilon - \theta_\lambda)) . \end{aligned}$$

Therefore multiplying Eq. (30) by $(y_\epsilon - y_\lambda)$ and integrating over \mathcal{O} , we therefore obtain in virtue of Lemma 2.4,

$$\begin{aligned} \frac{1}{2} |y_\epsilon - y_\lambda|_2^2 + \int_0^t \int_{\mathcal{O}} |\nabla (\gamma(y_\epsilon) - \gamma(y_\lambda))|^2 d\xi ds \\ = - \int_0^t \int_{\mathcal{O}} \left(G(e^W y_\epsilon) - G(e^W y_\lambda) \right) (y_\epsilon - y_\lambda) d\xi ds \\ - \int_0^t \int_{\mathcal{O}} f(y_\epsilon - y_\lambda) (y_\epsilon - y_\lambda) d\xi ds \\ - \int_0^t \int_{\mathcal{O}} \mu(y_\epsilon - y_\lambda) (y_\epsilon - y_\lambda) d\xi ds \\ + \int_0^t \int_{\mathcal{O}} \frac{e^{-W}}{\alpha} (p_\epsilon - p_\lambda) (y_\epsilon - y_\lambda) d\xi ds \\ + \int_0^t \int_{\mathcal{O}} \frac{e^{-W}}{\alpha} (\theta_\epsilon - \theta_\lambda) (y_\epsilon - y_\lambda) d\xi ds . \quad (31) \end{aligned}$$

The last two terms in Eq. (31) can be treated exploiting the Young inequality, yielding

$$\begin{aligned} \left| \int_{\mathcal{O}} \frac{e^{-W}}{\alpha} (p_\epsilon - p_\lambda) (y_\epsilon - y_\lambda) d\xi \right| &\leq C |p_\epsilon - p_\lambda|_2^2 |y_\epsilon - y_\lambda|_2^2 \\ &\leq C_1 |p_\epsilon - p_\lambda|_2^2 + C_2 |y_\epsilon - y_\lambda|_2^2, \\ \left| \int_{\mathcal{O}} \frac{e^{-W}}{\alpha} (\theta_\epsilon - \theta_\lambda) (y_\epsilon - y_\lambda) d\xi \right| &\leq C |\theta_\epsilon - \theta_\lambda|_2^2 |y_\epsilon - y_\lambda|_2^2 \\ &\leq C |y_\epsilon - y_\lambda|_2^2 + \epsilon + \lambda, \quad (32) \end{aligned}$$

as to obtain

$$\begin{aligned} \frac{1}{2} |y_\epsilon(t) - y_\lambda(t)|_2^2 + \int_0^t |\nabla (\gamma(y_\epsilon) - \gamma(y_\lambda))|_2^2 ds \\ \leq C_1 |y_\epsilon(t) - y_\lambda(t)|_2^2 + C_2 |p_\epsilon(t) - p_\lambda(t)|_2^2 + \epsilon + \lambda . \quad (33) \end{aligned}$$

Applying the Gronwall lemma and taking the mean value, we obtain using assumption 1 on γ

$$\mathbb{E}|y_\epsilon(t) - y_\lambda(t)|_2^2 + \mathbb{E} \int_0^t |\nabla(y_\epsilon - y_\lambda)|_2^2 ds \leq C \mathbb{E} \int_0^t |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda. \tag{34}$$

Regarding the second equation in (29), we obtain

$$\begin{aligned} & \partial_t (p_\epsilon - p_\lambda) + e^W \left(\gamma' \left(e^W y_\epsilon \right) \Delta \left(e^{-W} p_\epsilon \right) - \gamma' \left(e^W y_\lambda \right) \Delta \left(e^{-W} p_\lambda \right) \right) \\ & \quad + e^W \left(G' \left(e^W y_\epsilon \right) p_\epsilon - G' \left(e^W y_\lambda \right) p_\lambda \right) - f(p_\epsilon - p_\lambda) - \mu(p_\epsilon - p_\lambda) \\ & = 2 \left(e^W (y_\epsilon - y_\lambda) \right). \end{aligned} \tag{35}$$

Then, multiplying Eq. (35) by $(p_\epsilon - p_\lambda)$ and integrating over \mathcal{O} , we obtain

$$\begin{aligned} \frac{1}{2} |p_\epsilon(t) - p_\lambda(t)|_2^2 &= |e^W (y_\epsilon(T) - y_\lambda(T))|_2^2 \\ & \int_t^T \int_{\mathcal{O}} e^W \left(\gamma' \left(e^W y_\epsilon \right) \Delta \left(e^{-W} p_\epsilon \right) - \gamma' \left(e^W y_\lambda \right) \Delta \left(e^{-W} p_\lambda \right) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds \\ & \quad + \int_t^T \int_{\mathcal{O}} e^W \left(G' \left(e^W y_\epsilon \right) p_\epsilon(s) - G' \left(e^W y_\lambda \right) p_\lambda(s) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds \\ & \quad - \int_t^T \int_{\mathcal{O}} (f + \mu)(p_\epsilon(s) - p_\lambda(s))^2 d\xi ds \\ & \quad - \int_t^T \int_{\mathcal{O}} 2 \left(e^W (y_\epsilon - y_\lambda) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds. \end{aligned}$$

Rearranging terms above, we further have

$$\begin{aligned} \frac{1}{2} |p_\epsilon(t) - p_\lambda(t)|_2^2 &= |e^W (y_\epsilon(T) - y_\lambda(T))|_2^2 \\ & \int_t^T \int_{\mathcal{O}} e^W \gamma' \left(e^W y_\epsilon \right) \left(\Delta \left(e^{-W} p_\epsilon \right) - \Delta \left(e^{-W} p_\lambda \right) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds \\ & \quad + \int_t^T \int_{\mathcal{O}} e^W \Delta \left(e^{-W} p_\lambda \right) \left(\gamma' \left(e^W y_\epsilon \right) - \gamma' \left(e^W y_\lambda \right) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds \\ & \quad + \int_t^T \int_{\mathcal{O}} e^W G' \left(e^W y_\epsilon \right) (p_\epsilon(s) - p_\lambda(s))^2 d\xi ds \\ & \quad + \int_t^T \int_{\mathcal{O}} e^W p_\lambda(s) \left(G' \left(e^W y_\epsilon \right) - G' \left(e^W y_\lambda \right) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds \\ & \quad - \int_t^T \int_{\mathcal{O}} (f + \mu) (p_\epsilon(s) - p_\lambda(s))^2 d\xi ds \\ & \quad - \int_t^T \int_{\mathcal{O}} 2 \left(e^W (y_\epsilon - y_\lambda) \right) (p_\epsilon(s) - p_\lambda(s)) d\xi ds. \end{aligned} \tag{36}$$

Taking into account that $u \in \mathcal{U}$, it follows exploiting Lemma 2.3 and using the Young inequality,

$$\begin{aligned}
 |p_\epsilon(t) - p_\lambda(t)|_2^2 &\leq |y_\epsilon(T) - y_\lambda(T)|_2^2 \\
 &C \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + C \int_t^T |p_\lambda(s)|_2 |p_\epsilon(s) - p_\lambda(s)|_2 ds \\
 &+ C \int_t^T |y_\epsilon(s) - y_\lambda(s)|_2 |p_\lambda(s)|_2 |p_\epsilon(s) - p_\lambda(s)|_2 ds \\
 &\leq C \left(\int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \int_t^T |y_\epsilon(s) - y_\lambda(s)|_2^2 ds \right),
 \end{aligned}
 \tag{37}$$

where we have used C to denote possibly different constants as to simplify notation.

Taking the expectation in Eq. (37) and combining it with Eq. (34), we thus have

$$\begin{aligned}
 &\mathbb{E}|y_\epsilon(t) - y_\lambda(t)|_2^2 + \mathbb{E}|p_\epsilon(t) - p_\lambda(t)|_2^2 \\
 &\leq C \left(\mathbb{E} \int_0^t |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda \right) \\
 &+ C \left(\mathbb{E} \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \mathbb{E} \int_t^T |y_\epsilon(s) - y_\lambda(s)|_2^2 ds \right),
 \end{aligned}
 \tag{38}$$

so that, if T is small enough, we can infer that

$$\mathbb{E}|y_\epsilon(t) - y_\lambda(t)|_2^2 + \mathbb{E}|p_\epsilon(t) - p_\lambda(t)|_2^2 \leq C (\epsilon + \lambda), \tag{39}$$

implying that (y_ϵ, p_ϵ) is a Cauchy sequence, therefore ,along a subsequence still denoted by $\{\epsilon\} \rightarrow 0$ for the sake of clarity, we have

$$\begin{aligned}
 y_\epsilon &\rightarrow y \quad \text{weakly in } L^2((0, T); H_0^1), \\
 p_\epsilon &\rightarrow p \quad \text{weakly in } L^2((0, T) \times \mathcal{O}), \\
 u_\epsilon &:= \frac{1}{\alpha} e^{-W} (p_\epsilon + \theta_\epsilon) \rightarrow u^* \quad \text{weakly in } L^2((0, T) \times \mathcal{O} \times \Omega).
 \end{aligned}
 \tag{40}$$

Further, arguments analogous to the ones used in the proof of Lemma 2.5, it follows from the fact that G is maximal monotone, that for $\epsilon \rightarrow 0$, we get

$$G \left(e^{W_\epsilon} y_\epsilon \right) \rightarrow G(e^W y) \quad \text{weakly in } L^2 \left((0, T); H_0^1 \right). \tag{41}$$

Letting then $\epsilon \rightarrow 0$ in the first equation in (29), we have, denoting by y^* the solution corresponding to the optimal control u^* ,

$$\partial_t y^* - e^{-W} \Delta \gamma \left(e^W y^* \right) - e^{-W} G \left(e^W y^* \right) + \mu y^* = e^{-W} u^*,$$

hence, since Ψ is lower-semicontinuous, previous computations give us:

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

and the claimed existence result follows. □

Theorem 3.2 (Necessary condition of optimality) *Let be (v^*, u^*) an optimal pair for problem (P), then if $\alpha > 0$ it holds*

$$u^*(t, \xi) = \frac{1}{\alpha} P_U \left(-e^{-W} p(t, \xi) \right), \quad \text{a.e. on } (0, T) \times \mathcal{O} \times \Omega,$$

where p is the solution to the dual backward Eq. (45) and

$$P_U(v) = \begin{cases} M & v \geq M, \\ v & |v| \leq M \\ -M & v \leq -M. \end{cases} \tag{42}$$

Remark 3.1 We would like to underline that in literature about stochastic control problem, the first order conditions of optimality (the *Pontryagin maximum principle*) are expressed in terms of dual stochastic backward equation, see, e.g., [12,17]. Here, instead, optimality conditions are given in terms of a random backward dual equation which allows to simplify the setting also giving more insights on the derived optimal controller.

Proof We provide the result exploiting the rescaling transformation $y := e^{-W} X$, hence proving necessary condition for the problem (P2).

Let (y^*, u^*) be an optimal pair for problem (P2), therefore we have that for any $u \in \mathcal{U}$, defining $u^\lambda := u^* + \lambda \bar{u} = u^* + \lambda(u - u^*)$, $\lambda \geq 0$, by the optimality of u^* it must hold,

$$\frac{1}{\lambda} (\Psi(u^\lambda) - \Psi(u^*)) \geq 0.$$

By the Gâteaux differentiability of Ψ it follows, taking the limit as $\lambda \rightarrow 0$,

$$\begin{aligned} & \mathbb{E} \left[\int_0^T \int_{\mathcal{O}} \left(e^W y^*(t, \xi) - v_1(\xi) \right) z(t) + \alpha u^* \bar{u} d\xi dt \right] \\ & + \mathbb{E} \left[\int_{\mathcal{O}} \left(e^W y^*(T, \xi) - v_2(\xi) \right) z(T) d\xi \right] \geq 0, \end{aligned} \tag{43}$$

being z the solution to the system in variation defined as

$$\begin{cases} \dot{z}(t) - \Delta (\gamma'(e^W y^*)z(t)) + G'(e^W y^*)z(t) + \mu z(t) = e^{-W} u^*, \\ \gamma'(e^W y^*)z(t) \in H_0^1(\mathcal{O}), \quad t \in (0, T) \\ z(0) = 0. \end{cases} \tag{44}$$

Therefore, introducing the backward dual system

$$\begin{aligned} \dot{p}(t) &= -\Delta\gamma' \left(e^W y^* \right) p - G' \left(e^W y^* \right) p_\lambda(t) + \mu p(t) + 2 \left(e^W y^* - v_1 \right), \\ p(T) &= 2 \left(y^*(T) - v_2 \right), \end{aligned} \tag{45}$$

and exploiting Eqs. (44) and (45) together with Eq. (43), we have

$$\mathbb{E} \left[\int_0^T \int_{\mathcal{O}} \left\langle e^{-W} p(t) + \alpha u^*, \bar{u} \right\rangle dt \right] \geq 0, \tag{46}$$

which gives

$$u^*(t, \xi) = \frac{1}{\alpha} P_U \left(-e^{-W} p(t, \xi) \right), \quad \text{a.e. on } (0, T) \times \mathcal{O} \times \Omega,$$

where P_U is the projection operator defined in (42). □

Theorem 3.3 (The bang–bang principle) *Let be (v^*, u^*) an optimal pair for problem (P) and let $\alpha = 0$, then it holds*

$$u^* = \begin{cases} -M & \text{if } p > 0 \\ \in [-m, M] & \text{if } p = 0 \\ M & \text{if } p < 0. \end{cases} \tag{47}$$

where p is the solution to the dual backward Eq. (45).

Proof Proceeding as in Theorem 3.2 with obtain the equivalent of Eq. (46) to be

$$\mathbb{E} \left[\int_0^T \int_{\mathcal{O}} \left\langle e^{-W} p(t), \bar{u} \right\rangle dt \right] \geq 0, \tag{48}$$

which yields Eq. (47), and the claim follows. □

Remark 3.2 By (47) it follows that if $|v^* - v_1| > 0$ a.e. on $(0, T) \times \mathcal{O} \times \Omega$ then the optimal controller u^* is a *bang-bang* controller, namely $|u^*| = M$ a.e. on $(0, T) \times \mathcal{O} \times \Omega$.

Funding Open Access funding provided by Università degli Studi di Verona

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alberverio, S., Di Persio, L.: Some stochastic dynamical models in neurobiology: recent developments. *Eur. Commun. Math. Theor. Biol.* **14**, 44–53 (2011)
2. Alberverio, S., Di Persio, L., Mastrogiacomo, E.: Small noise asymptotic expansions for stochastic PDE's. I. The case of a dissipative polynomially bounded non linearity. *Tohoku Math. J.* **63**(4), 877–898 (2011)
3. Barbu, V.: *Analysis and Control of Nonlinear Infinite-Dimensional Systems*. Mathematics in Science and Engineering, 190. Academic Press Inc., Boston (1993)
4. Barbu, V.: *Nonlinear Differential Equations of Monotone Types in Banach Spaces*. Springer Monographs in Mathematics. Springer, New York (2010)
5. Barbu, V.: *Mathematical Methods in Optimization of Differential Systems*, vol. 310. Springer Science & Business Media, New York (2012)
6. Barbu, V., Iannelli, M.: Optimal control of population dynamics. *J. Optim. Theory Appl.* **102**(1), 1–14 (1999)
7. Barbu, V., Röckner, M.: On a random scaled porous media equation. *J. Differ. Equ.* **251**(9), 2494–2514 (2011)
8. Barbu, V., Precupanu, T.: *Convexity and Optimization in Banach Spaces*. Springer Science & Business Media, New York (2012)
9. Barbu, V., Röckner, M.: An operatorial approach to stochastic partial differential equations driven by linear multiplicative noise. *J. Eur. Math. Soc.* **17**–7, 1789–1815 (2015)
10. Barbu, V., Röckner, M.: Nonlinear Fokker–Planck equations driven by Gaussian linear multiplicative noise. *J. Differ. Equ.* **265**(10), 4993–5030 (2018)
11. Barbu, V., Da Prato, G., Röckner, M.: Existence of strong solutions for stochastic porous media equation under general monotonicity conditions. *Ann. Probab.* **37**(2), 428–452 (2009)
12. Barbu, V., Cordoni, F., Di Persio, L.: Optimal control of stochastic FitzHugh–Nagumo equation. *Int. J. Control* **89**, 1–25 (2015)
13. Barbu, V., Da Prato, G., Röckner, M.: *Stochastic Porous Media Equations*, vol. 2163. Springer, New York (2016)
14. Benilan, P., Brezis, H., Crandall, M.G.: A semilinear equation in $L^1(\mathbb{R}^N)$. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* **2**(4), 523–555 (1975)
15. Bonaccorsi, S., Marinelli, C., Ziglio, G.: Stochastic FitzHugh–Nagumo equations on networks with impulsive noise. *Electron. J. Probab.* **13**, 1362–1379 (2008)
16. Bonaccorsi, S., Mastrogiacomo, E.: Analysis of the Stochastic FitzHugh–Nagumo system. *Inf. Dim. Anal. Quantum Probab. Relat. Top.* **11**(3), 427–446 (2008)
17. Breiten, T., Kunisch, K.: Riccati-based feedback control of the monodomain equations with the Fitzhugh–Nagumo model. *SIAM J. Control Optim.* **52**(6), 4057–4081 (2014)
18. Casas, E., Ryll, C., Tröltzsch, F.: Sparse optimal control of the Schlögl and FitzHugh–Nagumo systems. *Comput. Methods Appl. Math.* **13**(4), 415–442 (2013)
19. Cordoni, F., Di Persio, L.: Small noise asymptotic expansion for the infinite dimensional Van der Pol oscillator. *Int. J. Math. Models Method Appl. Sci.* **9**, 202–210 (2015)
20. Cordoni, F., Di Persio, L.: Stochastic reaction–diffusion equations on networks with dynamic time-delayed boundary conditions. *J. Math. Anal. Appl.* **451**(1), 583–603 (2017)
21. Cordoni, F., Di Persio, L.: Gaussian estimates on networks with dynamic stochastic boundary conditions. *Infinite Dimens. Anal. Quant. Probab. Relat. Top.* **20**(01), 1750001 (2017)
22. Cordoni, F., Di Persio, L.: Optimal control for the stochastic FitzHugh–Nagumo model with recovery variable. *Evol. Equ. Control Theory* **7**(4), 571–585 (2018)
23. Da Prato, G.: *Kolmogorov Equations for Stochastic PDEs*. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser Verlag, Basel (2004)
24. Da Prato, G., Zabczyk, J.: *Ergodicity for Infinite-Dimensional Systems*. London Mathematical Society Lecture Note Series, 229. Cambridge University Press, Cambridge (1996)
25. Da Prato, G., Zabczyk, J.: *Second Order Partial Differential Equations in Hilbert Spaces*. London Mathematical Society Lecture Note Series, 293. Cambridge University Press, Cambridge (2012)
26. Da Prato, G., Zabczyk, J.: *Stochastic Equations in Infinite Dimensions*, vol. 152. Cambridge University Press, Cambridge (2014)
27. Diehl, H.W.: The theory of boundary critical phenomena. *Int. J. Mod. Phys. B* **11**(30), 3503–3523 (1997)

28. Ekeland, I.: On the variational principle. *J. Math. Anal. Appl.* **47**(2), 324–353 (1974)
29. Fuhrman, M., Orrieri, C.: Stochastic maximum principle for optimal control of a class of nonlinear SPDEs with dissipative drift. *SIAM J. Control Optim.* **54**(1), 341–371 (2016)
30. Fuhrman, M., Tessitore, G.: Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. *Ann. Probab.* **30**(3), 1397–1465 (2002)
31. FitzHugh, R.: Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* **1**, 445–466 (1961)
32. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**(4), 500–544 (1952)
33. Iqbal, N., Ranchao, W., Liu, B.: Pattern formation by super-diffusion in FitzHugh–Nagumo model. *Appl. Math. Comput.* **313**, 245–258 (2017)
34. Marinelli, C., Di Persio, L., Ziglio, G.: Approximation and convergence of solutions to semilinear stochastic evolution equations with jumps. *J. Funct. Anal.* **264**(12), 2784–2816 (2013)
35. Nagumo, J., Arimoto, S., Yoshizawa, S.: An active pulse transmission line simulating nerve axon. *Proc. Inst. Radio Eng.* **50**(10), 2061–2070 (1962)
36. Pankratov, L.: Homogenization of the ginzburg-landau heat flow equation in a porous medium. *Appl. Anal.* **69**(1–2), 31–45 (1998)
37. Pitchaiah, S., Armaou, A.: Output feedback control of the FitzHugh–Nagumo equation using adaptive model reduction. In: *Proceedings of the 49th IEEE Conference on Decision and Control*. Atlanta, GA (2010)
38. Ringkvist, M., Zhou, Y.: On the dynamical behaviour of FitzHugh–Nagumo systems: revisited. *Nonlinear Anal. Theory Methods Appl.* **71**(7–8), 2667–2687 (2009)
39. Scarpa, L.: Optimal distributed control of a stochastic Cahn–Hilliard Equation. *SIAM J. Control Optim.* **57**(5), 3571–3602 (2019)
40. Siddheshwar, P.G., Vanishree, R.K.: Lorenz and Ginzburg–Landau equations for thermal convection in a high-porosity medium with heat source. *Ain Shams Eng. J.* (2016). <https://doi.org/10.1016/j.asej.2016.11.007>
41. Tuckwell, H.C.: Random perturbations of the reduced Fitzhugh–Nagumo equation. *Phys. Scr.* **46**(6), 481 (1992)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.