



# Optimal Control of Mean Field Equations with Monotone Coefficients and Applications in Neuroscience

Antoine Hocquet<sup>1</sup> · Alexander Vogler<sup>1</sup>

Accepted: 12 August 2021 / Published online: 9 September 2021  
© The Author(s) 2021

## Abstract

We are interested in the optimal control problem associated with certain quadratic cost functionals depending on the solution  $X = X^\alpha$  of the stochastic mean-field type evolution equation in  $\mathbb{R}^d$

$$dX_t = b(t, X_t, \mathcal{L}(X_t), \alpha_t)dt + \sigma(t, X_t, \mathcal{L}(X_t), \alpha_t)dW_t, \quad X_0 \sim \mu \quad (\mu \text{ given}), \quad (1)$$

under assumptions that enclose a system of FitzHugh–Nagumo neuron networks, and where for practical purposes the control  $\alpha_t$  is deterministic. To do so, we assume that we are given a drift coefficient that satisfies a one-sided Lipschitz condition, and that the dynamics (2) satisfies an almost sure boundedness property of the form  $\pi(X_t) \leq 0$ . The mathematical treatment we propose follows the lines of the recent monograph of Carmona and Delarue for similar control problems with Lipschitz coefficients. After addressing the existence of minimizers via a martingale approach, we show a maximum principle for (2), and numerically investigate a gradient algorithm for the approximation of the optimal control.

**Keywords** Optimal control · McKean–Vlasov equations · FitzHugh–Nagumo neurons · Stochastic differential equations · Gradient descent

**Mathematics Subject Classification** 93E20 · 92B20 · 65K10

---

✉ Antoine Hocquet  
antoine.hocquet@tu-berlin.de  
Alexander Vogler  
vogler@math.tu-berlin.de

<sup>1</sup> Technische Universität Berlin Institut für Mathematik, Straße des 17. Juni 136, 10623 Berlin, Germany

# 1 Introduction

## Motivations

Based on a modification of a model by van der Pol, FitzHugh [17] proposed in 1961 the following system of equations in order to describe the dynamics of a single neuron subject to an external current  $I$ :

$$\begin{aligned} \dot{v} &= v - \frac{1}{3}v^3 - w + I \\ \dot{w} &= c(v + a - bw) \end{aligned} \tag{2}$$

for some constants  $a, b, c > 0$ , where the unknowns  $v, w$  correspond respectively to the so-called voltage and recovery variables (see also Nagumo [19]). In presence of interactions, one has to enlarge the previous pair by an additional unknown  $y$  that counts a fraction of open channels (synaptic channels), and which is sometimes referred to as gating variable.

When it comes to an interacting network of neurons, it is customary to assume that the corresponding graph is fully connected, which is arguably a good approximation at small scales [24]. This implies that all the neurons in the given network add a contribution to the interaction terms in the equation. Precisely, for a population of size  $N \in \mathbb{N}$ , the state at time  $t$  of the  $i$ -th neuron is described by the three-dimensional vector

$$X_t^i = (v_t^i, w_t^i, y_t^i), \quad i = 1, \dots, N,$$

and one is led to study the system of  $3N$  stochastic differential equations:

$$\left\{ \begin{aligned} dv_t^i &= \left( v_t^i - \frac{(v_t^i)^3}{3} - w_t^i + I_t \right) dt + \sigma_{ext} dW_t^i \\ &\quad - \frac{1}{N} \sum_{j=1}^N J(v_t^i - V_{rev}) y_t^j dt - \frac{1}{N} \sum_{j=1}^N \sigma^J (v_t^i - V_{rev}) y_t^j dB_t^i \\ dw_t^i &= c(v_t^i + a - bw_t^i) dt, \\ dy_t^i &= (\bar{a}S(v_t^i)(1 - y_t^i) - \bar{b}y_t^i) dt + \sigma^{y^i}(v^i) d\tilde{B}_t^i. \end{aligned} \right. \tag{3}$$

In the above,  $W^i, B^i, \tilde{B}^i$  are i.i.d. Brownian motions modelling independent sources of noise with respective intensities  $\sigma^J, \sigma_{ext}, \sigma^{y^i}(v^i) > 0$ . The last of these intensities depends on the solution, through the formula

$$\sigma^{y^i}(v) = \chi(y) \sqrt{\bar{a}S(v)(1 - y) + \bar{b}y} \tag{4}$$

with given constants  $\bar{a}, \bar{b} > 0$  and some smooth cut-off function  $\chi : \mathbb{R} \rightarrow \mathbb{R}$  supported in  $(0, 1)$ .

In this model, the voltage variable  $v^i$  describes the membrane potential of the  $i$ -th neuron in the network, while the recovery variable  $w^i$  models the dynamics of corresponding ion channels, which can influence the membrane potential of neuron  $i$  by opening and closing, depending on  $v^i$ .

Without interaction term (i.e. when  $J = \sigma^J = 0$ ), equation (3) reduces to a network of  $N$  independent neurons, where the dynamics of each neuron is modelled by the FitzHugh–Nagumo equation (2).

When interaction is present ( $J \neq 0$ ), the model describes the situation where each neuron of the network affects its adjacent neurons by releasing chemical transmitters, causing particular ion channels of adjacent neurons to open. This induces a current to the adjacent neuron, affecting its membrane potential. In this extended model, the gating variable  $y^i$  models a fraction of open ion channels in the adjacent neurons of neuron  $i$ , and thus ought to be a number between 0 and 1 (hence the cut-off  $\chi(y^i)$  in (4)). Loosely speaking,  $y^i$  should be thought as the output contribution of the neuron  $i$ .

Depending on the fraction  $y^i$  of open channels, the induced current for an adjacent neuron with membrane potential  $v$  is given by  $-J(v - V_{rev})y^i$ , where the constant  $V_{rev}$  describes the membrane potential, at which there is no net current flow. The coupling strength  $J$  originally refers to the (mean of) the maximum conductance, which is typically affected by noise coming from the environment. This explains the diffusion part of the interaction term in Eq. (3) for  $\sigma^J \neq 0$ . For further details we refer to [2].

The dynamics of the gating variable  $y^i$  in Eq. (3) depends on some physical constants, which we will now briefly introduce:

- $S(v^i)$  refers to the concentration of chemical transmitters, released by neuron  $i$ ; explicitly for  $v \in \mathbb{R}$

$$S(v) = \frac{T_{max}}{1 + e^{-\lambda(v - V_T)}} \tag{5}$$

where  $T_{max}$  is a given maximal concentration and  $\lambda^{-1} > 0$ ,  $V_T > 0$ , are constants setting the steepness, resp. the value, at which  $S(v)$  is half-activated (for typical values, see for instance [13]);

- $\bar{a}, \bar{b} > 0$  correspond to some rise and decay rates, respectively.

For a better understanding of the interaction, we included a small illustration in Fig. 1.

In this representation, a rapid increase of the membrane potential of the neuron  $i$  will cause it to release chemical messengers into the (synaptic) cleft between the neuron  $i$  and the adjacent neuron  $j$ , which in turn will bind to receptors of the neuron  $j$ . The receptors will cause ion channels of neuron  $j$  to open, thus neuron  $i$  induces the opening of a fraction  $y^i$  of ion channels at the dendrites of the neuron  $j$ . As already mentioned, the resulting current from  $i$  to  $j$  affecting the neuron  $j$  is then given by  $-J(v^j - V_{rev})y^i$ . For a thorough presentation of (3) and its applications in the field of neurosciences, we refer for instance to the monograph of Ermentrout and Terman [16].

When it comes to monitoring neural activity in the brain, one does typically not measure single neuron activity, but consider more macroscopic measurements like, e.g., local

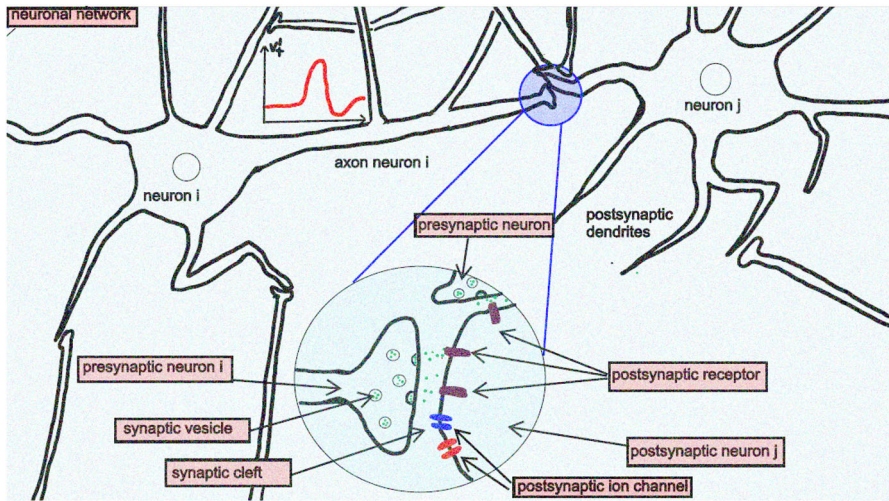


Fig. 1 Synaptic dynamics

field potentials (LFPs). The LFP refers to the electric potential in the extracellular space around neurons and it is influenced by all ionic processes around the electrode. It is assumed that action potentials have a limited impact to the LFP and that it is strongly influenced by synaptic currents. Although we are aware of the fact that the correct interpretation of measurements like the intracerebral local field potential is highly non trivial, for our type of application it is reasonable to assume that measurements only depend on the distribution of our network, rather than on activities of single neurons from the network. Furthermore, external stimulation acts homogeneously on every neuron of the network, that is every neuron of the network receives the same external input  $I$ . We therefore model the dynamics of a typical neuron as a controlled mean-field type equation and admissible controls need to be independent of the state of the individual neuron, hence we will consider deterministic controls. In the following we will formulate our mean-field model and the set of admissible controls will be defined in Sect. 2.2.

**Propagation of Chaos**

The system (3) has the generic form

$$\begin{cases} dX_t^{N,i} = b(t, X_t^{N,i}, \bar{\mu}_{X_t^N}, \alpha_t)dt + \sigma(t, X_t^{N,i}, \bar{\mu}_{X_t^N}, \alpha_t)dW_t^i, & t \in [0, T], \\ X_0^{N,i} \sim u_0, \end{cases} \quad (6)$$

for  $i = 1, \dots, N$ , where  $u_0$  is a probability measure on  $\mathbb{R}^d$ ,  $(\alpha_t)$  is a control and  $\bar{\mu}_{X_t^N}$  denotes the empirical measure

$$\bar{\mu}_{X_t^N} := \frac{1}{N} \sum_{k=1}^N \delta_{X_t^{N,k}}.$$

For  $N \rightarrow \infty$ , one is naturally pushed to investigate the convergence in law of the solutions of (6) towards the probability measure  $\mu = \mathcal{L}(X|\mathbb{P})$ , where  $X$  solves

$$\begin{cases} dX_t = b(t, X_t, \mathcal{L}(X_t), \alpha_t)dt + \sigma(t, X_t, \mathcal{L}(X_t), \alpha_t)dW_t, & t \in [0, T], \\ X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d), \end{cases} \tag{7}$$

and where  $b, \sigma$  are the coefficients obtained by substituting expectations in (6) in place of empirical means. In the context of (3), a first mathematical investigation of such convergence is due to Baladron, Fasoli, Faugeras and Touboul [2] (see also the clarification notes [6]). In this direction, the authors show that the sequence of symmetric probability measures

$$\mu_N := \mathcal{L}((X^{N,1}, \dots, X^{N,N})|\mathbb{P})$$

is  $\mu$ -chaotic. Namely, for each integer  $k \leq N$  and every finite sequence of bounded and continuous functions  $\phi_i : C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ , for  $i = 1, \dots, k$ , it holds

$$\lim_{N \rightarrow \infty} \langle \mu_N, \phi_1 \otimes \dots \otimes \phi_k \otimes 1 \otimes \dots \otimes 1 \rangle = \prod_{i=1}^k \langle \mu, \phi_i \rangle,$$

where  $\langle \mu, \phi \rangle := \int \phi d\mu$  and the symbol “ $\otimes$ ” denotes the usual tensor product, i.e.  $\phi_i \otimes \phi_j(f, g) := \phi_i(f)\phi_j(g)$ , for any  $f, g \in C([0, T]; \mathbb{R}^d)$ . This situation is usually referred to as “propagation of chaos”. Although we could not find any literature concerning propagation of chaos results in the general non Lipschitz case, propagation of chaos for the system (3) was investigated in [6].

**Mean-Field Limit and Control**

In this regard, taking  $N \gg 1$  guarantees that a “good enough” approximation of (3) is given by the mean-field limit (7), where the corresponding coefficients  $(b, \sigma) : [0, T] \times \mathbb{R}^3 \times \mathcal{P}(\mathbb{R}^3) \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}^{3 \times 3}$ , are given by

$$b(t, x, \mu, \alpha) = \begin{pmatrix} v - \frac{v^3}{3} - w + \alpha \\ c(v + a - bw) \\ \bar{a}S(v)(1 - y) - \bar{b}y \end{pmatrix} + \begin{pmatrix} -J(v - V_{rev}) \int_{\mathbb{R}^3} z_3 \mu(dz) \\ 0 \\ 0 \end{pmatrix}, \tag{8}$$

for  $x = (v, w, y)$ , and

$$\sigma(t, x, \mu, \alpha) = \begin{pmatrix} \sigma_{ext} - \sigma^J(v - V_{rev}) \int_{\mathbb{R}^3} z_3 \mu(dz) & 0 \\ 0 & 0 \\ 0 & \chi(y) \sqrt{\bar{a}S(v)(1 - y) + \bar{b}y} \end{pmatrix}. \tag{9}$$

In this paper, we concentrate our attention on the optimal control problem associated with a cost functional of the form

$$J : \mathbb{A} \rightarrow \mathbb{R}, \quad \alpha \mapsto \mathbb{E} \left( \int_0^T f(t, X_t^\alpha, \mathcal{L}(X_t^\alpha), \alpha_t) dt + g(X_T^\alpha, \mathcal{L}(X_T^\alpha)) \right), \tag{10}$$

for suitable functions  $f$  and  $g$ , and where  $X^\alpha$  is subject to the dynamical constraint (7). The functional cost ought to be minimized over some convex, admissible set of controls  $\mathbb{A}$ .

Because of potential applications in neuroscience, the control of the stochastic FHN model has gained a lot of attention during the last years (see, e.g., [3,11]). The need to introduce random perturbations in the original model is widely justified from a physics perspective (see for instance [12] and the references therein). In [11] the authors investigate a FitzHugh–Nagumo SPDE which results from the continuum limit of a network of coupled FitzHugh–Nagumo equations. We have a similar structure in mind regarding the dependence of the coefficients on the control (namely, the dynamics of the membrane potential depends linearly on the control). Our approach here is however completely different, in that we hinge on the McKean-Vlasov type SDE (7) that originates from the propagation of chaos.

McKean-Vlasov control problems of this type were investigated in the past decade by Bensoussan, Frehse and Yam [4], but also by Carmona and co-authors (see for instance [9]). These developments culminated with the monograph of Carmona and Delarue [8], where a systematic treatment is made (under reasonable assumptions). Other related works include [1,5,7,14]. These results fail however to encompass (7)–(9), due for instance to the lack of Lipschitz property for the drift coefficient.

From the analytic point of view, the FitzHugh–Nagumo model also suffers the fact that the diffusion matrix is degenerate, making difficult to obtain energy estimates for the Kolmogorov equation (see Remark 3.2).

Our objective in this work is twofold. At first, our purpose is to extract some of the qualitative features of FitzHugh–Nagumo system and its mean field limit, in a broader treatment that encloses (3) and (7)–(9). In this sense, our intention is not to deal with the previous models “as such” but instead, we aim to take a step further by dealing with a certain class of equations that has the following peculiarities

- (Monotonicity) – though the drift coefficient in (7) displays a cubic non-linearity, it satisfies the monotonicity condition  $\langle x - x', b(t, x, \mu, \alpha) - b(t, x', \mu, \alpha) \rangle \lesssim |x - x'|^2$ .
- (Structural assumption on dynamics and level set boundedness) – the dynamics of the coupling variable ensures that the boundedness property  $y_t \in [0, 1]$  holds for all times.

- (Interaction with linear growth w.r.t. the unknown) – the drift nonlinearity displays the behaviour  $|b(t, x, \mu, \alpha) - b(t, x, \nu, \alpha)| \lesssim (1 + |x|) W_2(\mu, \nu)$ , where  $W_2$  denotes the usual 2-Wasserstein distance defined in Sect. 2.1.

Under the above setting, we aim to develop and implement direct variational methods, in the spirit of the stochastic approach of Yong and Zhou [25] for classical control problems (note that some work in this direction has been already done by Pfeiffer [21, 22], in a slightly different setting). Second, we aim to derive a Pontryagin maximum principle for mean-field type control problems of the previous form, with a view towards efficient numerical approximations of optimal controls (e.g. gradient descent).

### Organization of the Paper

In Sect. 2 we introduce our assumptions on the coefficients and give the main results. Section 3 is devoted to the well-posedness of the main optimal control problem (Theorem 2.1). In Sect. 4, we show the corresponding maximum principle (Theorem 2.2). Finally, Sect. 5 will be devoted to numerical examples.

## 2 Preliminaries

### 2.1 Notation and Settings

In the whole manuscript, we consider an arbitrary but finite time horizon  $T > 0$ . We fix a dimension  $d \geq 1$ , and denote the scalar product in  $\mathbb{R}^d$  by  $\langle \cdot, \cdot \rangle$ . If  $A, B$  are matrices of the same size, we shall also write  $\langle A, B \rangle$  for their scalar product, namely

$$\langle A, B \rangle := \text{tr}(A^\dagger B)$$

where  $A^\dagger$  is the transposed matrix, and  $\text{tr}$  the trace operator. For a continuously differentiable function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we adopt the suggestive notation  $f_x$  to denote its Jacobian (seen for each  $x \in \mathbb{R}^d$  as an element of the dual of  $\mathbb{R}^d$ ). Given  $h \in \mathbb{R}^d$ , we let

$$f_x(x) \cdot h \tag{11}$$

be the evaluation of  $f_x(x)$  at  $h$ . A similar convention will be used for vector-valued functions.

Throughout the paper, we fix a complete filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  carrying an  $m$ -dimensional Wiener process  $(W_t)_{t \in [0, T]}$ . Given  $p \in [1, \infty)$  and a  $p$ -integrable random variable  $X$ , we denote its usual  $L^p$ -norm by  $\|X\|_p := \mathbb{E}(|X|^p)^{1/p}$ . We further introduce the spaces

$$\begin{aligned} \mathcal{H}^{2,d} &:= \left\{ Z : \Omega \times [0, T] \rightarrow \mathbb{R}^d \mid Z \text{ prog. measurable and } \int_0^T \|Z_t\|_2^2 dt < \infty \right\} \\ \mathcal{S}^{2,d} &:= \left\{ Z : \Omega \times [0, T] \rightarrow \mathbb{R}^d \mid Z \text{ prog. measurable, continuous and } \left\| \sup_{t \in [0, T]} |Z_t| \right\|_2^2 < \infty \right\}. \end{aligned}$$

For  $m \in \mathbb{N}$ , the notations  $\mathcal{S}^{2,d \times m}, \mathcal{H}^{2,d \times m}$  will also be used to denote the corresponding sets of  $d \times m$  matrix-valued processes. Whenever clear from the context, we will omit to indicate dimensions and write  $\mathcal{S}^2$  or  $\mathcal{H}^2$  instead.

We will denote by  $\mathcal{P}(\mathbb{R}^d)$  the set of all probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . For  $p \in [1, \infty)$ ,  $\mu \in \mathcal{P}(\mathbb{R}^d)$  we define the moment of order  $p$ :

$$\mathcal{M}_p(\mu)^p := \int_{\mathbb{R}^d} |x|^p \mu(dx) \in [0, \infty],$$

and we let  $\mathcal{P}_p(\mathbb{R}^d) := \{\mu \in \mathcal{P}(\mathbb{R}^d) \mid \mathcal{M}_p(\mu) < \infty\}$ . By  $W_p$ ,  $p \in [1, \infty)$ , we denote the usual  $p$ -Wasserstein distance on  $\mathcal{P}_p$ , that is for  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$

$$W_p(\mu, \nu)^p := \inf_{\pi \in \Pi(\mu, \nu)} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx \times dy), \tag{12}$$

where  $\Pi(\mu, \nu)$  denotes the set of probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with  $\mu$  and  $\nu$  as respective first and second marginals (we refer to [8, Chap. 5] for a thorough introduction to the subject). Moreover, we recall the following elementary but useful consequence of the previous definition. Let  $\mu, \nu$  be in  $\mathcal{P}_p$ , and assume that there are random variables  $X, Y$  on  $(\Omega, \mathcal{F}, \mathbb{P})$  such that  $X \sim \mu$  and  $Y \sim \nu$ . Then, it holds

$$W_p(\mu, \nu) \leq \mathbb{E}(|X - Y|^p)^{\frac{1}{p}}. \tag{13}$$

Finally, we will shortly recall the notion of L-differentiability. A function  $f : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$  is called L-differentiable at  $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$  if there exists a random variable  $X_0$  with law  $\mu_0$ , such that the lifted function

$$\tilde{f} : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \rightarrow \mathbb{R}, \quad X \mapsto f(\mathcal{L}(X))$$

is Fréchet differentiable at  $X_0$ . The following property for a L-differentiable function  $f$  is well-known (see [8, Chap. 5]): for any  $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ , there exists a  $\mu_0$ -almost everywhere uniquely defined measurable function  $\xi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , such that for all  $X_0 \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$  with  $\mathcal{L}(X_0) = \mu_0$ , it holds  $D\tilde{f}(X_0) = \xi(X_0)$ . We write  $f_\mu(\mu_0)$  to denote the equivalence class of  $\xi$  in  $L^2(\mathbb{R}^d, \mu_0; \mathbb{R}^d)$ . In keeping with the notation (11) on differentials, we will let  $f_\mu(\nu)(x) \cdot h$  be its evaluation (as an element of the dual of  $\mathbb{R}^d$ ) at  $h \in \mathbb{R}^d$ . If  $f_\mu$  is continuous, we call  $f$  continuously L-differentiable.

### 2.2 Controls and Cost Functional

Our controlled dynamics will be given by a McKean-Vlasov type SDE (state equation) of the form (7), where  $X_0 \in L^r(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$  for some fixed  $r \geq 6$  and  $\alpha$  is an admissible control, i.e. for some bounded convex set  $A \subset \mathbb{R}^k$  and throughout the paper,

$$\alpha \in \mathbb{A} := \{\alpha : [0, T] \rightarrow A\}. \tag{14}$$



In the whole manuscript, we assume that we are given continuous running and terminal cost functions

$$\begin{aligned}
 f &: [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R} \\
 g &: \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}
 \end{aligned}$$

which have quadratic growth in the following sense: there exists  $C > 0$  such that for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$ ,  $\alpha \in A$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$

$$\begin{aligned}
 |f(t, x, \mu, \alpha)| &\leq C(1 + |x| + \mathcal{M}_2(\mu) + |\alpha|)^2 \\
 |g(x, \mu)| &\leq C(1 + |x| + \mathcal{M}_2(\mu))^2.
 \end{aligned}$$

We will then consider the cost functional

$$J: \mathbb{A} \rightarrow \mathbb{R}, \quad \alpha \mapsto \mathbb{E} \left( \int_0^T f(t, X_t^\alpha, \mathcal{L}(X_t^\alpha), \alpha_t) dt + g(X_T^\alpha, \mathcal{L}(X_T^\alpha)) \right), \quad (15)$$

where  $X^\alpha$  is subject to the dynamical constraint (7).

### 2.3 Level Set Boundedness

A formal application of Itô Formula reveals that the solutions of the state equation associated with a network of FitzHugh–Nagumo neurons take values in the set

$$\mathcal{C} := \{x = (v, w, y) : 0 \leq y \leq 1\}.$$

This is of course coherent with the intuition that  $y$  is a fraction of open channels. In other words, we have  $\pi(X) \leq 0$  where  $\pi: \mathbb{R}^3 \rightarrow \mathbb{R}$ , is the map  $x \mapsto y(y - 1)$ . Motivated by this example, we will assume in the sequel that we are given a convex function  $\pi \in C^2(\mathbb{R}^d, \mathbb{R})$  such that any solution  $X$  is supported in  $\mathcal{C} \subset \mathbb{R}^d$  for all times, where  $\mathcal{C}$  is the set

$$\mathcal{C} := \{x \in \mathbb{R}^d : \pi(x) \leq 0\}. \quad (16)$$

We suppose moreover that  $\mathcal{C}$  contains at least one element, which for convenience is assumed to be 0. To ensure that the solutions are indeed  $\mathcal{C}$ -valued, we need to assume that  $\pi(X_0) \leq 0$ ,  $\mathbb{P}$ -almost surely. Furthermore we need to make the following compatibility assumptions on  $\pi: \mathbb{R}^d \rightarrow \mathbb{R}$ .

**Assumption 2.1** (structural assumption on the dynamics) For all  $\mu \in \mathcal{P}(\mathbb{R}^d)$ ,  $\alpha \in A$ ,  $t \in [0, T]$  and  $x \in \mathbb{R}^d \setminus \mathcal{C}$ , we have

$$\pi_x(x) \cdot b(t, x, \mu, \alpha) \leq 0, \quad (17)$$

while

$$\text{Im}(\sigma(t, x, \mu, \alpha)) \subset \pi_x(x)^\perp \quad \text{and} \quad \pi_{xx}(x) \cdot (\sigma\sigma^\dagger(t, x, \mu, \alpha)) = 0. \quad (18)$$

**Example 2.1** (*Gating variable boundedness for FitzHugh–Nagumo*) Assumption 2.1 is fulfilled for (7)–(9) and with  $\pi(v, w, y) = y(y - 1)$ , as can be seen as follows. We have the identities (recall the notation (4))

$$\pi_x(x) = (0 \ 0 \ 2y - 1), \quad \pi_{xx}(x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

$$\sigma\sigma^\dagger(t, x, \mu, \alpha) = \begin{pmatrix} \sigma_{ext}^2 + (\sigma^J)^2(v - V_{rev})^2 \left(\int_{\mathbb{R}^3} z_3 \mu(dz)\right)^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma^y(v)^2 \end{pmatrix},$$

Clearly,

$$\pi_x(x)\sigma(t, x, \mu, \alpha) \in \text{Linspan} \left\{ (0 \ 0 \ (2y - 1)\sigma^y(v)) \right\}.$$

But using  $\text{Supp}\chi \subset (0, 1)$ , we find indeed that  $(2y - 1)\sigma^y(v) = 0$  outside  $\mathcal{C}$ . The same argument implies

$$\pi_{xx}(x) \cdot (\sigma\sigma^\dagger(t, x, \mu, \alpha)) = 2\sigma^y(v)^2$$

and the latter vanishes if  $x \notin \mathcal{C}$ , hence (18).

Towards (17), one observes letting  $q = \bar{a}S(v)$  that

$$\pi_x(x) \cdot b(t, x, \mu, \alpha) = -q + (3q + \bar{b})y - 2(q + \bar{b})y^2 = P(y).$$

The polynomial  $P(y)$  has discriminant  $(q - \bar{b})^2$ , hence the roots

$$r_- = \frac{q}{q + \bar{b}}, \quad r_+ = \frac{1}{2},$$

which both lie in the interval  $(0, 1)$ . It follows that  $P(y)$  is negative outside  $\mathcal{C}$ , implying (17).

### 2.4 Regularity Assumptions and Main Results

Besides Assumption 2.1, one needs to make suitable hypotheses on the regularity of the drift and diffusion coefficients. In the sequel, we denote by  $\mathcal{P}_2^{\mathcal{C}}(\mathbb{R}^d)$  the subset of all probability measures in  $\mathcal{P}_2(\mathbb{R}^d)$  which are supported in  $\mathcal{C} = \pi^{-1}((-\infty, 0])$ .

**Assumption 2.2** (*MKV Regularity*) We assume that the coefficients

$$(b, \sigma) : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^d \times \mathbb{R}^{d \times m}$$

are locally Lipschitz. Moreover, there are constants  $L_1, L_2, L_3 > 0$  such that the following properties hold.

(L1) – (regularity of the diffusion coefficient) – The diffusion coefficient  $\sigma$  satisfies the property  $\sup_{0 \leq t \leq T} |\sigma(t, 0, \delta_0, 0)|^2 < \infty$ . Moreover, for all  $t \in [0, T], x \in \mathcal{C}, \alpha \in A$  and  $\mu \in \mathcal{P}_2^{\mathcal{C}}(\mathbb{R}^d)$  we have

$$|\sigma(t, x, \mu, \alpha)|^2 \leq L_1(1 + |\alpha|^2 + |x|^2). \tag{19}$$

For all  $x, x' \in \mathcal{C}, \alpha' \in A$ , it holds

$$|\sigma(t, x, \mu, \alpha) - \sigma(t, x', \mu, \alpha')|^2 \leq L_1(|x - x'|^2 + |\alpha - \alpha'|^2). \tag{20}$$

Finally, if  $x \in \mathcal{C}$  and  $\mu' \in \mathcal{P}_2(\mathbb{R}^d)$ , then

$$|\sigma(t, x, \mu, \alpha) - \sigma(t, x, \mu', \alpha)|^2 \leq L_1(1 + |x|^2)W_2(\mu, \mu')^2. \tag{21}$$

(L2) – (regularity of the drift coefficient) – There exists  $q \in \mathbb{N}$  with  $4q \leq r$ , such that for all  $t \in [0, T], x, x' \in \mathcal{C}, \alpha, \alpha' \in A$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$

$$\begin{aligned} &|b(t, x, \mu, \alpha) - b(t, x', \mu, \alpha')| \\ &\leq \sqrt{L_2}(1 + |x|^{q-1} + |x'|^{q-1} + |\alpha|^{q-1} + |\alpha'|^{q-1} + \mathcal{M}_2(\mu)^2)(|x - x'| + |\alpha - \alpha'|). \end{aligned} \tag{22}$$

In addition,  $b$  satisfies the following Lipschitz property with respect to the Wasserstein distance: for all  $t \in [0, T], x \in \mathcal{C}, \alpha \in A$  and  $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$

$$|b(t, x, \mu, \alpha) - b(t, x, \mu', \alpha)|^2 \leq L_2(1 + |x|^2)W_2(\mu, \mu')^2. \tag{23}$$

(L3) – (monotonicity of the drift) – The drift coefficient  $b$  is such that  $\sup_{0 \leq t \leq T} |b(t, 0, \delta_0, 0)| < \infty$ . Moreover, for all  $t \in [0, T], x \in \mathcal{C}, \alpha \in A$  and  $\mu \in \mathcal{P}_2^{\mathcal{C}}(\mathbb{R}^d)$  it holds

$$\langle x, b(t, x, \mu, \alpha) \rangle \leq L_3(1 + |\alpha|^2 + |x|^2) \tag{24}$$

and if  $x' \in \mathcal{C}, \alpha' \in A$ , then

$$\langle x - x', b(t, x, \mu, \alpha) - b(t, x', \mu, \alpha') \rangle \leq L_3(|x - x'|^2 + |\alpha - \alpha'|^2). \tag{25}$$

**Example 2.2** (Analysis of the FitzHugh–Nagumo model) Let us go back to the settings of (7)–(9) for a coupled system of FitzHugh–Nagumo neurons. Trivially, one has  $\sup_{0 \leq t \leq T} |\sigma(t, 0, \delta_0, 0)| = |\sigma_{ext}| < \infty$ . The map  $v \mapsto S(v)$  being positive and bounded, we further see that the (3, 3)-th entry of  $\sigma$  is Lipschitz, as deduced immediately from the fact that  $\chi$  is supported in  $(0, 1)$ . For the remaining non-trivial component, we have

$$\sigma^{1,2}(x, \mu, \alpha)^2 \leq J(V_{rev} + |v|^2)|\beta(\mu)|$$

where to ease notation we introduce the barycenter  $\beta(\mu)$ , defined as the quantity

$$\beta(\mu) := \int_{\mathbb{R}^3} z_3 \mu(dz_1 \times dz_2 \times dz_3). \tag{26}$$

The condition  $\text{Supp}\mu \subset \mathcal{C}$ , implies trivially that  $|\beta(\mu)| \leq 1$  and thus we obtain (19) for  $L_1 = (V_{rev}J) \vee 1$ . The Lipschitz-type property (20) is shown in a similar fashion.

The Wasserstein-type regularity (21) is hardly more problematic: using the Kantorovitch duality Theorem [8, Prop. 5.3 & Cor. 5.4] and the fact that the projector  $z = (z_1, z_2, z_3) \mapsto z_3$  is Lipschitz, one finds that

$$|\beta(\mu - \mu')| = \left| \int_{\mathbb{R}^3} z_3(\mu - \mu')(dz) \right| \leq W_1(\mu, \mu'). \tag{27}$$

hence

$$|\sigma(x, \mu) - \sigma(x, \mu')| \leq J|v - V_{rev}|W_1(\mu, \mu').$$

As is classical, the 1-Wasserstein distance  $W_1(\mu, \mu')$  can be estimated by  $W_2(\mu, \mu')$ , which in turn implies (21), and thus 2.2.

As for the drift coefficient, since  $b(t, 0, \delta_0, 0)$  is also independent of  $t$ , the supremum condition in 2.2 is clear. Moreover, it has polynomial dependency on the variables  $v, w, y$ , which implies the local Lipschitz property (22) with  $q = 3$ . We also have

$$|b(t, x, \mu, \alpha) - b(t, x, \mu', \alpha)| \leq J|v - V_{rev}||\beta(\mu - \mu')|$$

and we conclude by (27) that 2.2 holds.

To show (24) and (25), it is enough to prove the corresponding bounds when  $c = 0 = \bar{b}$ , since the related contributions are affine linear in the variables. Similarly, by linearity we can let  $w = \alpha = 0$ . But in that case, it holds

$$\langle x, b(t, x, \mu, 0) \rangle \leq v^2 - \frac{v^4}{3} + \bar{a}S(v)(1 - y)y - Jv^2\beta(\mu) + J V_{rev}v\beta(\mu).$$

Observe that, since  $\mu$  is supported inside  $\mathcal{C}$ , one has in particular  $\beta(\mu) \geq 0$ . Consequently, the fourth term in the right hand side can be ignored, showing (24) with  $L_3 = L_3(\bar{a}, |S|_\infty, J, V_{rev}) > 0$ .

Similarly, if  $x' = (v', 0, y') \in \mathbb{R}^3$

$$\begin{aligned} & \langle x - x', b(t, x, \mu, 0) - b(t, x', \mu, 0) \rangle \\ &= (v - v')^2 - \frac{1}{3}(v^3 - v'^3)(v - v') - J(v - v')^2\beta(\mu) \\ & \quad + \bar{a}(1 - y)(y - y')(S(v) - S(v')) - \bar{a}S(v')(y - y')^2 \\ & \leq |S'|_\infty(1 \vee \bar{a})(1 + y^2)(|y - y'|^2 + |v - v'|^2). \end{aligned}$$

It follows that (24) holds with  $L_3 = L_3(\bar{a}, \bar{b}, c, |S|_{\mathcal{C}^1}) > 0$ .

**Assumption 2.3** (*Weak continuity*) For any  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , the function  $A \rightarrow \mathbb{R}$ ,  $\alpha \mapsto f(t, x, \mu, \alpha)$  is convex and the functions  $A \rightarrow \mathbb{R}^d \times \mathbb{R}^{d \times m}$ ,  $\alpha \mapsto (b, \sigma)(t, x, \mu, \alpha)$  are affine. Furthermore, for all  $x \in C([0, T]; \mathbb{R}^d)$  and  $\mu \in C([0, T]; \mathcal{P}_2^C(\mathbb{R}^d))$  the functions

$$\begin{aligned} \mathbb{A} &\rightarrow L^2([0, T]; \mathbb{R}^d), & \alpha &\mapsto b(\cdot, x, \mu, \alpha), \\ \mathbb{A} &\rightarrow L^2([0, T]; \mathbb{R}^{d \times m}), & \alpha &\mapsto \sigma(\cdot, x, \mu, \alpha), \end{aligned}$$

are weakly sequentially continuous, that is for all  $(\alpha^n)_{n \in \mathbb{N}} \subset \mathbb{A}$  with  $\alpha^n \rightharpoonup \alpha$ , for some  $\alpha \in \mathbb{A}$ , it holds  $(b, \sigma)(\cdot, x, \mu, \alpha^n) \rightharpoonup (b, \sigma)(\cdot, x, \mu, \alpha)$ .

**Remark 2.1** The continuity and convexity of  $f(t, x, \mu, \cdot)$  leads to weak lower semi-continuity of the map

$$\mathbb{A} \rightarrow \mathbb{R}, \quad \alpha \mapsto \int_0^T f(t, x_t, \mu_t, \alpha_t) dt,$$

for all  $x \in C([0, T]; \mathbb{R}^d)$  and  $\mu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d))$ .

We can now present our main results. At first, we investigate the existence of an optimal control for the following problem

$$\min_{\alpha \in \mathbb{A}} J(\alpha), \tag{SM}$$

subject to

$$\begin{cases} dX_t = b(t, X_t, \mathcal{L}(X_t), \alpha_t) dt + \sigma(t, X_t, \mathcal{L}(X_t), \alpha_t) dW_t, & t \in [0, T], \\ X_0 \in L^r(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d). \end{cases} \tag{28}$$

**Theorem 2.1** *Under assumptions 2.1–2.3, the problem (SM) is finite and has an optimal control. Namely,  $\inf_{\alpha \in \mathbb{A}} J(\alpha) < \infty$  and there is  $\bar{\alpha} \in \mathbb{A}$ , such that*

$$J(\bar{\alpha}) \leq J(\alpha),$$

for all  $\alpha \in \mathbb{A}$ .

In order to address the corresponding maximum principle, we now introduce further assumptions on our coefficients.

**Assumption 2.4** (*Pontryagin Principle*) The coefficients  $b, \sigma, f$  and  $g$  are continuously differentiable with respect to  $(x, \alpha)$  and continuously L-differentiable with respect to  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ . Furthermore there exist  $A_1, A_2, A_3 > 0$  such that:

(A1) For every  $(s, x, \mu, \alpha) \in [0, T] \times \mathcal{C} \times \mathcal{P}_2^C(\mathbb{R}^d) \times A$ , and each  $y, z \in \mathbb{R}^d$ :

$$\langle b_x(t, x, \mu, \alpha) \cdot z, z \rangle \leq A_1 |z|^2,$$

$$\begin{aligned}
 |b_x(t, x, \mu, \alpha)| &\leq A_1(1 + |x|^{q-1}), \\
 |b_\alpha(t, x, \mu, \alpha)| &\leq A_1, \\
 |b_\mu(t, x, \mu, \alpha)(y)| &\leq A_1(1 + |x|),
 \end{aligned}$$

where  $q$  is the same constant as in 2.2.

(A2) For every  $(s, x, \mu, \alpha) \in [0, T] \times \mathcal{C} \times \mathcal{P}_2^{\mathcal{C}}(\mathbb{R}^d) \times A$ , and  $y \in \mathbb{R}^d$ :

$$\begin{aligned}
 |\sigma_x(t, x, \mu, \alpha)| &\leq A_2, \\
 |\sigma_\alpha(t, x, \mu, \alpha)| &\leq A_2, \\
 |\sigma_\mu(t, x, \mu, \alpha)(y)| &\leq A_2(1 + |x|).
 \end{aligned}$$

(A3) For all  $R > 0$  and every  $(s, x, \mu, \alpha) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$ , such that  $|x| \vee \mathcal{M}_2(\mu) \vee |\alpha| \leq R$  the quantities

$$\begin{aligned}
 &f_x(t, x, \mu, \alpha), f_\alpha(t, x, \mu, \alpha), g_x(x, \mu), \int_{\mathbb{R}^d} |f_\mu(t, x, \mu, \alpha)(y)|^2 \mu(dy), \\
 &\int_{\mathbb{R}^d} |g_\mu(x, \mu)(y)|^2 \mu(dy),
 \end{aligned}$$

are all bounded in norm by  $A_3(1 + R)$ .

**Example 2.3** Again, we investigate the above properties for the setting of a FitzHugh–Nagumo neural network. The property 2.4 depends on the choice of  $f$  and  $g$ , hence we do not discuss it here (it is however clear for the ansatz (47) below). Concerning assumption 2.4 and 2.4 we have

$$b_x(t, x, \mu, \alpha) = \begin{pmatrix} 1 - v^2 - J\beta(\mu) & 1 & 0 \\ c & -cb & 0 \\ \bar{a}S'(v)(1 - y) & 0 & -\bar{a}S(v) - \bar{b} \end{pmatrix},$$

where we recall the notation (26). Using that  $\text{Supp}(\mu) \subset \mathcal{C}$ , together with the boundedness of  $S'(v)$ , this leads to

$$\langle b_x(t, x, \mu, \alpha) \cdot z, z \rangle \leq A_1(b, c, \bar{a}, \bar{b}, |S|_\infty, |S'|_\infty) |z|^2,$$

hence the first estimate. Letting as before  $\beta(\mu) := \int_{\mathbb{R}^3} z_3 \mu(dz)$ , it is easily seen by definition of the L-derivative that

$$\beta_\mu(\mu)(\tilde{x}) \cdot h = h_3 \quad \text{for all } \tilde{x} \text{ and } h \equiv (h_1, h_2, h_3) \in \mathbb{R}^3.$$

In a matrix representation, this gives the following constant value for the L-derivative of the drift coefficient at a given point  $x \equiv (v, w, y) \in \mathbb{R}^3$

$$b_\mu(t, x, \mu, \alpha)(\tilde{x}) = \begin{pmatrix} 0 & 0 & -J(v - V_{rev}) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{for all } \tilde{x} \in \mathbb{R}^3.$$

Thus we have  $|b_\mu(t, x, \mu, \alpha)(\tilde{x})| \leq J \vee (J V_{rev})(1 + |x|)$ , showing the desired property.

Next, we introduce the corresponding adjoint equation, which will be essential for the maximum principle. For a solution  $X \in \mathcal{S}^{2,d}$  of (28) consider the following backward SDE

$$\begin{cases} dP_t = - \left\{ \langle b_x(t, X_t, \mathcal{L}(X_t), \alpha_t), P_t \rangle + \langle \sigma_x(t, X_t, \mathcal{L}(X_t), \alpha_t), Q_t \rangle + f_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \right. \\ \quad \left. - \tilde{\mathbb{E}} \left( b_\mu(t, \tilde{X}_t, \mathcal{L}(X_t), \alpha_t)(X_t), \tilde{P}_t \right) + f_\mu(t, \tilde{X}_t, \mathcal{L}(X_t), \alpha_t)(X_t) \right\} dt + Q_t dW_t \\ P_T = g_x(X_T, \mathcal{L}(X_T)) + \tilde{\mathbb{E}} \left( g_\mu(\tilde{X}_T, \mathcal{L}(X_T))(X_T) \right), \end{cases} \tag{29}$$

where the tilde variables  $\tilde{X}, \tilde{P}$  are independent copies of the corresponding random variables (carried on some arbitrary probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ ), and  $\tilde{\mathbb{E}}$  denotes integration in  $\tilde{\Omega}$  (this convention will be adopted throughout the paper). Herein, we recall that  $\langle \sigma(t, x, \mu, \alpha), q \rangle$  is a synonym for  $\text{tr}(\sigma(t, x, \mu, \alpha)^\dagger q)$ .

A pair of processes  $(P, Q) \in \mathcal{H}^{2,d} \times \mathcal{H}^{2,d \times m}$  will be called a solution to the adjoint equation corresponding to  $X$  if it satisfies (29) for all  $t \in [0, T]$ ,  $\mathbb{P}$ -almost surely.

We are now in position to formulate the maximum principle. For that purpose, we introduce the Hamiltonian, which for each  $x, p \in \mathbb{R}^d, q \in \mathbb{R}^{d \times m}, \mu \in \mathcal{P}_2$  and  $\alpha \in A$ , is the quantity

$$H(t, x, \mu, p, q, \alpha) := \langle b(t, x, \mu, \alpha), p \rangle + \langle \sigma(t, x, \mu, \alpha), q \rangle + f(t, x, \mu, \alpha).$$

**Theorem 2.2** *Let assumptions 2.1–2.4 hold. Let  $\bar{\alpha} \in \mathbb{A}$  be an optimal control for the problem (SM). If  $(P, Q) \in \mathcal{H}^{2,d} \times \mathcal{H}^{2,d \times m}$  is the solution to the corresponding adjoint equation, then we have for Lebesgue-almost every  $t \in [0, T]$*

$$\mathbb{E} (H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \bar{\alpha}_t)) \leq \mathbb{E} (H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha)),$$

for all  $\alpha \in A$ .

It should be noticed that in contrast to the maximum principle stated in [8, Thm. 6.14 p. 548], the maximum principle here is formulated in terms of the expectation for almost every  $t \in [0, T]$  instead of  $dt \otimes \mathbb{P}$ -almost everywhere, since we only consider deterministic controls and thus we only alter the control in deterministic directions.

### 3 Well-Posedness of the Optimal Control Problem

The main purpose of this section is to prove the existence of an optimal control for the stated control problem. For that purpose, we will need to show (among other results) that the state equation (7) is well-posed, and that the solution satisfies uniform moment bounds up to a certain level. Hereafter, we suppose that assumptions 2.1, 2.2 and 2.3 are fulfilled.

### 3.1 Well-Posedness of the State Equation

Our first task is to show that the level-set constraint which was alluded to in Sect. 2.3 is preserved along the flow of solutions. This statement is contained in the next result. The proof is partially adapted from that of [6, Prop. 3.3].

**Lemma 3.1** *For every  $\alpha \in \mathbb{A}$  and  $\mu \in C([0, T]; \mathcal{P}_2^C(\mathbb{R}^d))$  we have that*

$$\mathbb{P}(\pi(X_t^{\alpha, \mu}) \leq 0, \forall t \in [0, T]) = 1 \tag{30}$$

where  $X_t^{\alpha, \mu}$  is the unique solution to

$$\begin{cases} dX_t = b(t, X_t, \mu_t, \alpha_t)dt + \sigma(t, X_t, \mu_t, \alpha_t)dW_t, & t \in [0, T] \\ X_0 \in L^r(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d). \end{cases} \tag{31}$$

**Proof** First, observe that given  $\mu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d))$ , equation (31) has a unique strong solution  $X^\mu$  in  $\mathcal{S}^2$ . Indeed, if we let

$$b^\mu(t, x, \alpha) := b(t, x, \mu, \alpha), \quad \sigma^\mu(t, x, \alpha) := \sigma(t, x, \mu, \alpha),$$

then from Assumption 2.2 we see that  $\sigma^\mu$  is Lipschitz, while 2.2 and 2.2 imply the local Lipschitz continuity and the monotonicity of the drift coefficient  $b^\mu$ . Hence, by standard results on monotone SDEs (see for instance [20, Thm. 3.26 p. 178]) (31) has a unique strong solution, this solution being progressively measurable and square integrable. This proves our assertion.

In order to show (30), consider a family  $(\Psi_\epsilon)_{\epsilon > 0}$  of non-negative and non-decreasing functions in  $C^2(\mathbb{R})$  which for all  $\epsilon > 0$  satisfy:

$$\Psi_\epsilon(x) = 0 \text{ on } (-\infty, 0], \quad \Psi_\epsilon(x) = 1 \text{ on } [\epsilon, \infty), \quad \sup_\epsilon |\Psi_\epsilon|_\infty \leq 1,$$

and such that  $\Psi_\epsilon$  converges pointwise to  $\mathbf{1}_{(0, \infty)}$  as  $\epsilon \rightarrow 0$ . Let  $\tau_n := \inf\{t \geq 0 \text{ s.t. } |X_t| \geq n\}$ . By Itô Formula, we have for each  $n \geq 0$  and  $\epsilon > 0$

$$\begin{aligned} \Psi_\epsilon(\pi(X_{t \wedge \tau_n})) - M_t^\epsilon &= \int_0^{\tau_n \wedge t} (\pi_x(X_s) \cdot b(s, X_s, \mu_s, \alpha_s)) \Psi'_\epsilon(\pi(X_t)) ds \\ &\quad + \frac{1}{2} \int_0^{\tau_n \wedge t} \Psi''_\epsilon(\pi(X_s)) |\pi_x(X_s)^\dagger \sigma(s, X_s, \mu_s, \alpha_s)|^2 ds \\ &\quad + \frac{1}{2} \int_0^{\tau_n \wedge t} \Psi'_\epsilon(\pi(X_s)) \pi_{xx}(X_s) \cdot (\sigma \sigma^\dagger(s, X_s, \mu_s, \alpha_s)) ds, \end{aligned}$$

where we let  $M_t^\epsilon := \sum_{k=1}^m \int_0^{\tau_n \wedge t} \pi_x(X_s) \cdot \sigma^{\cdot, k}(s, X_s, \mu_s, \alpha_s) \Psi'_\epsilon(\pi(X_s)) dW_s^k$ . Since  $\Psi_\epsilon$  is supported on the real positive axis, only the values of  $X$  which satisfy  $\pi(X) > 0$  contribute to the above expression. Hence, making use of Assumption 2.1, we see that



the first term in the previous right hand side is bounded above by 0, while the two last terms simply vanish. We arrive at the relation

$$\mathbb{E} \left( \sup_{t \in [0, T]} \Psi_\epsilon (\pi(X_{t \wedge \tau_n})) \right) \leq 0.$$

Letting first  $n \rightarrow \infty$ , and then  $\epsilon \rightarrow 0$ , we observe by Fatou Lemma that

$$\mathbb{E} \left( \sup_{t \in [0, T]} \mathbf{1}_{(0, \infty)} (\pi(X_t)) \right) = 0,$$

and our claim follows. □

We are now able to prove the existence of a unique solution to equation (7).

**Theorem 3.1** *There exists a unique strong solution to equation (7) in  $\mathcal{S}^2$ , which is supported in  $\mathcal{C}$  for all times. Furthermore, for each  $p \in [2, r]$  and every  $\alpha \in \mathbb{A}$ , the solution satisfies the moment estimate*

$$\mathbb{E} \left( \sup_{t \in [0, T]} |X_t|^p \right) \leq C (\|X_0\|_p, L_1, L_3, p) \left( 1 + \int_0^T |\alpha_t|^p dt \right). \tag{32}$$

where the constant  $C$  depends only upon the indicated quantities.

**Proof** Recall that  $\mathcal{P}_2^{\mathcal{C}}$  denotes the set of probability measures in  $\mathcal{P}_2(\mathbb{R}^d)$  which are supported in  $\mathcal{C} := \pi^{-1}((-\infty, 0])$ . Equipped with the standard Wasserstein distance, it is a closed subset of  $\mathcal{P}_2(\mathbb{R}^d)$ . Indeed, it is standard (see for instance [15]) that given probability measures  $\{\mu_n, n \in \mathbb{N}\}$  and  $\mu$  such that  $\mu_n \Rightarrow \mu$ , then

$$\text{supp} \mu \subset \liminf_{n \rightarrow \infty} (\text{supp} \mu_n) := \left\{ x \in \mathbb{R}^d \mid \limsup_{n \rightarrow \infty} \inf_{y \in \text{supp} \mu_n} |x - y| = 0 \right\},$$

so that our claim follows. Thus, for fixed  $\alpha \in \mathbb{A}$ , we can rightfully consider the operator

$$\Theta : C([0, T]; \mathcal{P}_2^{\mathcal{C}}) \rightarrow C([0, T]; \mathcal{P}_2^{\mathcal{C}}), \quad \mu \mapsto (\mathcal{L}(X_t^{\alpha, \mu}))_{t \in [0, T]},$$

where  $X^\mu = X^{\alpha, \mu}$  is the unique solution to equation (31). Using similar arguments as in [8], the existence of a unique solution to (28) follows if one can show that  $\Theta$  has a unique fixed point. In fact, we are going to show that it is a contraction (for a well-chosen metric). The moment estimate (32) will follow from the fixed point argument, provided one can show that

$$\mathbb{E} \left( \sup_{t \in [0, T]} |X_t^\mu|^p \right) \leq C (\|X_0\|_p, L_1, L_3, p) \left( 1 + \int_0^T |\alpha_t|^p dt \right) \tag{33}$$

where the displayed constant depends on the indicated quantities but not on the particular element  $\mu$  in  $C([0, T]; \mathcal{P}_2^C)$ . We now divide the proof into two steps.

*Step 1: moment bounds.* We need a localization argument. For each  $n > 0$ , introduce  $\tau_n := \inf\{t \in (0, T], |X_t^\mu| > n\}$ , and denote by  $X_t^{\mu,n} := X_{t \wedge \tau_n}^\mu$ . Itô Formula gives

$$\begin{aligned} \frac{1}{p} |X_t^{\mu,n}|^p - N_t^{\mu,n} &= \frac{1}{p} |X_0^{\mu,n}|^p + \int_0^{t \wedge \tau_n} \left\{ \langle X_s^\mu, b(s, X_s^\mu, \mu_s, \alpha_s) \rangle |X_s^\mu|^{p-2} \right. \\ &\quad \left. + \frac{1}{2} |\sigma(s, X_s^\mu, \mu_s, \alpha_s)|^2 |X_s^\mu|^{p-2} + \frac{p-2}{2} |\sigma^\dagger X_s^\mu|^2 |X_s^\mu|^{p-4} \right\} ds \end{aligned} \tag{34}$$

where  $N_t^{\mu,n} := \int_0^{t \wedge \tau_n} |X_s^\mu|^{p-2} \langle X_s^\mu, \sigma(s, X_s^\mu, \mu_s, \alpha_s) dW_s \rangle$  is the corresponding martingale term. Denoting by  $\kappa > 0$  the constant in the Burkholder–Davis–Gundy Inequality, the latter is estimated for each  $t \in [0, T]$  thanks to (19) and Cauchy–Schwarz Inequality as

$$\begin{aligned} \mathbb{E} \left( \sup_{s \in [0,t]} N_s^{\mu,n} \right) &\leq \kappa \mathbb{E} \left( \left( \int_0^{t \wedge \tau_n} |X_s^\mu|^{2p-4} |\sigma(s, X_s^\mu, \mu_s, \alpha_s)^\dagger X_s^\mu|^2 ds \right)^{\frac{1}{2}} \right) \\ &\leq \kappa \sqrt{L_1} \mathbb{E} \left( \sup_{0 \leq s \leq t} |X_s^{\mu,n}|^{\frac{p}{2}} \left( \int_0^t |X_s^{\mu,n}|^{p-2} (1 + |X_s^{\mu,n}|^2 + |\alpha_s|^2 \mathbf{1}_{[0, \tau_n]}(s)) ds \right)^{\frac{1}{2}} \right). \end{aligned}$$

But from Young’s inequality, the previous right hand side is also bounded by

$$\begin{aligned} \frac{1}{2p} \mathbb{E} \left( \sup_{0 \leq s \leq t} |X_s^{\mu,n}|^p \right) \\ + \frac{p\kappa^2 L_1}{2} \mathbb{E} \left( \int_0^t (|X_s^{\mu,n}|^{p-2} + |X_s^{\mu,n}|^p + |X_s^{\mu,n}|^{p-2} |\alpha_s|^2 \mathbf{1}_{[0, \tau_n]}(s)) ds \right). \end{aligned}$$

Define  $\Psi_t^n := \mathbb{E} (\sup_{0 \leq s \leq t} |X_s^{\mu,n}|^p)$ . Taking the expectation in (34), we infer from (24), (19), Young’s inequality  $ab \leq \frac{2}{p} a^{\frac{p}{2}} + \frac{p-2}{p} b^{\frac{p}{p-2}}$  and the previous discussion that

$$\frac{1}{2p} \Psi_t^n \leq \frac{1}{p} \mathbb{E} (|X_0|^p) + C_p (L_1 + L_3) \int_0^t (1 + \Psi_s^n + |\alpha_s|^p) ds$$

for some universal constant  $C_p > 0$ . Applying Gronwall Inequality, we obtain the corresponding moment estimate for  $X^{\mu,n}$ , for each  $n > 0$ . Since none of the constants used in the above computations depend on  $n$ , the localization is removed by letting  $n \nearrow \infty$  and using the monotone convergence theorem.

*Step 2: the fixed point argument.* From Lemma 3.1, it is clear that for all  $t \in [0, T]$ , the probability measure  $\mathbb{P} \circ (X_t^\mu)^{-1}$  is supported in  $\mathcal{C}$ . For simplicity, let  $L := L_1 \vee L_2 \vee L_3$  and introduce the weight

$$\phi_t := \exp(-2Lt), \quad t \in [0, T].$$

Then, Itô Formula gives

$$\begin{aligned}
 d\left(\frac{1}{2}|X_t^\mu - X_t^\nu|^2\phi_t\right) + 2L|X_t^\mu - X_t^\nu|^2\phi_t dt &= \phi_t \langle X_t^\mu - X_t^\nu, b(t, X_t^\mu, \mu_t, \alpha_t) - b(t, X_t^\nu, \mu_t, \alpha_t) \rangle dt \\
 &+ \phi_t \langle X_t^\mu - X_t^\nu, b(t, X_t^\nu, \mu_t, \alpha_t) - b(t, X_t^\nu, \nu_t, \alpha_t) \rangle dt \\
 &+ \phi_t \langle X_t^\mu - X_t^\nu, \sigma(t, X_t^\mu, \mu_t, \alpha_t) - \sigma(t, X_t^\nu, \nu_t, \alpha_t) \rangle dW_t \\
 &+ \frac{1}{2}\phi_t |\sigma(t, X_t^\mu, \mu_t, \alpha_t) - \sigma(t, X_t^\nu, \nu_t, \alpha_t)|^2 dt.
 \end{aligned}
 \tag{35}$$

The first term in the right hand side of (35) is evaluated thanks to (25). For the second term, we use the quadratic growth assumption (23). As for the Itô correction, we can estimate it similarly, using this time Assumption 2.2. With  $M_t := \int_0^t \phi_s \langle X_s^\mu - X_s^\nu, \sigma(s, X_s^\mu, \mu_s, \alpha_s) - \sigma(s, X_s^\nu, \nu_s, \alpha_s) \rangle dW_s$  we get

$$\begin{aligned}
 \frac{1}{2}|X_t^\mu - X_t^\nu|^2\phi_t + 2L \int_0^t |X_s^\mu - X_s^\nu|^2\phi_s ds - M_t &\leq \int_0^t \left\{ (L_1 + L_3)|X_s^\mu - X_s^\nu|^2 + (L_1 + L_2)(1 + |X_s^\nu|^2)W_2(\mu_s, \nu_s)^2 \right\} \phi_s ds \\
 &\leq 2L \left( \int_0^t |X_s^\mu - X_s^\nu|^2\phi_s ds + \left(1 + \sup_{s \in [0,t]} |X_s^\nu|^2\right) \int_0^t W_2(\mu_s, \nu_s)^2\phi_s ds \right).
 \end{aligned}$$

Taking expectations, supremum in  $t$ , then absorbing to the left yields

$$\sup_{0 \leq s \leq t} \mathbb{E} \left( |X_s^\mu - X_s^\nu|^2 \right) \phi_s \leq 4L \left( 1 + \mathbb{E} \left( \sup_{s \in [0,t]} |X_s^\nu|^2 \right) \right) \int_0^t W_2(\mu_s, \nu_s)^2 \phi_s ds$$

Using the estimate (32) with  $p = 2$ , the fact that  $\exp(-2TL) \leq \phi \leq 1$ , inequality (13) and the boundedness of the control state space  $A \subset \mathbb{R}^k$ , we arrive at

$$\sup_{0 \leq s \leq t} W_2(\Theta(\mu)_s, \Theta(\nu)_s)^2 \leq C(\|X_0\|_p, T, L) \int_0^t W_2(\mu_s, \nu_s)^2 ds.$$

Considering the  $k$ -th composition of the map  $\Theta$ , we get

$$\sup_{0 \leq s \leq T} W_2(\Theta^k(\mu)_s, \Theta^k(\nu)_s)^2 \leq \frac{C(\|X_0\|_p, T, L)^k T^k}{k!} \sup_{0 \leq s \leq T} W_2(\mu_s, \nu_s)^2,$$

hence contractivity of  $\Theta^k$  follows for  $k > 0$  large enough and the result then follows from Banach-fixed point theorem. □

We now investigate some regularity of the control-to-state operator, which will be needed in the proof of the optimality principle.

**Lemma 3.2** For  $p \in [2, r]$ , the solution map

$$G: \mathbb{A} \rightarrow \mathcal{S}^p \cap \mathcal{S}^2, \quad \alpha \mapsto X^\alpha$$

is well-defined and Lipschitz continuous. More precisely, there exists a constant  $C(L_1, L_2, L_3, T) > 0$ , such that for all  $\alpha, \beta \in \mathbb{A}$

$$\mathbb{E} \left( \sup_{t \in [0, T]} |X_t^\alpha - X_t^\beta|^2 \right) \leq C(L_1, L_2, L_3, T) \int_0^T |\alpha_t - \beta_t|^2 dt.$$

**Proof** That  $G$  is well-defined follows immediately from Theorem 3.1. Towards Lipschitz-continuity, the property is shown by similar considerations as in the proof of Theorem 3.1. Indeed, fixing  $\alpha, \beta \in \mathbb{A}$  and letting  $M$  be the martingale  $M_t := \int_0^t \langle \sigma(t, X^\alpha, \mathcal{L}(X^\alpha), \alpha) - \sigma(t, X^\beta, \mathcal{L}(X^\beta), \beta), (X^\alpha - X^\beta)dW \rangle$ , then using Itô Formula with assumptions 2.2, 2.2 and 2.2, we arrive at

$$\begin{aligned} & \frac{1}{2} |X_t^\alpha - X_t^\beta|^2 - M_t \\ &= \int_0^t \left\{ \langle X_s^\alpha - X_s^\beta, b(s, X_s^\alpha, \mathcal{L}(X_s^\alpha), \alpha_s) - b(s, X_s^\beta, \mathcal{L}(X_s^\beta), \beta_s) \rangle \right. \\ & \quad + \langle X_s^\alpha - X_s^\beta, b(s, X_s^\beta, \mathcal{L}(X_s^\alpha), \beta_s) - b(s, X_s^\beta, \mathcal{L}(X_s^\beta), \beta_s) \rangle \\ & \quad \left. + \frac{1}{2} |\sigma(t, X_s^\alpha, \mathcal{L}(X_s^\alpha), \alpha_s) - \sigma(t, X_s^\beta, \mathcal{L}(X_s^\beta), \beta_s)|^2 \right\} ds \\ &\leq \int_0^t \left\{ (L_3 + \frac{1}{2} + L_1)(|X^\alpha - X^\beta|^2 + |\alpha - \beta|^2) \right. \\ & \quad \left. + (\frac{L_2}{2} + L_1)(1 + |X^\alpha|^2 + |X^\beta|^2) W_2(\mathcal{L}(X^\alpha), \mathcal{L}(X^\beta))^2 \right\} ds. \end{aligned}$$

Letting  $\kappa > 0$  be the constant in the BDG inequality, the estimate (13) and  $ab \leq \frac{a^2}{4} + b^2$  yield

$$\begin{aligned} & \frac{1}{4} \mathbb{E} \left( \sup_{s \in [0, t]} |X_s^\alpha - X_s^\beta|^2 \right) \\ & \leq C_L (3 + \kappa^2) \left( 2 + \mathbb{E} \left( \sup_{s \in [0, T]} |X_s^\alpha|^2 + |X_s^\beta|^2 \right) \right) \\ & \quad \int_0^t \left\{ \mathbb{E} \left( \sup_{r \in [0, s]} |X_r^\alpha - X_r^\beta|^2 \right) + |\alpha_s - \beta_s|^2 \right\} ds \end{aligned}$$

where  $C_L := \frac{1}{2} \vee L_1 \vee \frac{L_2}{2} \vee L_3$ . The result now follows from the uniform bound (32), together with Gronwall Lemma. □

**Remark 3.1** Since we have  $W_2(\mathcal{L}(X_s^\alpha), \mathcal{L}(X_s^\beta)) \leq \mathbb{E} \left( \sup_{t \in [0, T]} |X_s^\alpha - X_s^\beta|^2 \right)^{\frac{1}{2}}$  we also get the Lipschitz continuity of the map

$$\mathbb{A} \rightarrow \mathcal{P}_2(\mathcal{S}^2), \quad \alpha \mapsto \mathcal{L}(G(\alpha)).$$

**Remark 3.2 (Fokker-Planck equation)** Given the settings of Example 2.2, we define

$$b_0(t, x, \alpha) := \begin{pmatrix} v - \frac{v^3}{3} - w + \alpha \\ c(v + a - bw) \\ \bar{a}S(v)(1 - y) - \bar{b}y \end{pmatrix}, \quad b_1(x, z) := \begin{pmatrix} -J(v - V_{rev})z_3 \\ 0 \\ 0 \end{pmatrix},$$

$$\tilde{\sigma}(x, z) := \begin{pmatrix} \sigma_{ext} - \sigma^J(v - V_{rev})z_3 & 0 \\ 0 & 0 \\ 0 & \chi(y)\sqrt{\bar{a}S(v)(1 - y) + \bar{b}y} \end{pmatrix}.$$

If we assume that the solution to the corresponding mean-field equation has a density  $p(t, x)$  with respect to the 3-dimensional lebesgue measure, then the McKean-Vlasov-Fokker-Planck equation is given by the nonlinear PDE:

$$\partial_t p(t, x) = -\operatorname{div} \left( \left( b_0(t, x, \alpha) + \int_{\mathbb{R}^3} b_1(x, z) p(t, z) dz \right) p(t, x) \right) + \frac{1}{2} \nabla^2 \cdot \left( \left( \int_{\mathbb{R}^3 \times \mathbb{R}^3} \tilde{\sigma}(x, z) \tilde{\sigma}(x, \bar{z})^\dagger p(t, z) p(t, \bar{z}) dz d\bar{z} \right) p(t, x) \right)$$

(see [2]). It is degenerate parabolic because the matrix  $\sigma \tilde{\sigma}^\dagger$  is not strictly positive.

### 3.2 Proof of Theorem 2.1

We now prove the existence of an optimal control for (28). The strategy we use strings along the commonly named “direct method” in the calculus of variations. As a trivial consequence of the assumptions made in Sect. 2.2 and the uniform estimate (32), note at first that our control problem is indeed finite. Next, consider a sequence  $(\alpha^n)_{n \in \mathbb{N}} \subset \mathbb{A}$  realizing the infimum of  $J$  asymptotically, i.e.

$$\lim_{n \rightarrow \infty} J(\alpha^n) = \inf_{\alpha \in \mathbb{A}} J(\alpha).$$

Since  $\mathbb{A} \subset L^2([0, T]; \mathbb{R}^k)$  is bounded and closed, by Banach Alaoglu Theorem there exists an  $\alpha \in L^2([0, T]; \mathbb{R}^k)$  and a subsequence also denoted by  $(\alpha^n)_{n \in \mathbb{N}}$ , such that

$$\alpha^n \rightharpoonup \alpha, \quad \text{weakly in } L^2(0, T; \mathbb{R}^k).$$

Since  $\mathbb{A}$  is also convex, we get  $\alpha \in \mathbb{A}$ , so  $\alpha$  is indeed an admissible control. We now divide the proof into four steps.

*Step 1: tightness.* In the sequel, we denote by  $X^n$  the solution of the state equation (7) with respect to the control  $\alpha^n$ ,  $n \in \mathbb{N}$ . Adding and subtracting in (7), we have

$$\|X_t^n - X_s^n\|_4^4 \leq 4^3 \left\{ \left\| \int_s^t b(r, 0, \delta_0, 0) dr \right\|_4^4 + \left\| \int_s^t b(r, 0, \mathcal{L}(X_r^n), 0) - b(r, 0, \delta_0, 0) dr \right\|_4^4 + \left\| \int_s^t b(r, X_r^n, \mathcal{L}(X_r^n), \alpha_r^n) - b(r, 0, \mathcal{L}(X_r^n), 0) dr \right\|_4^4 + \kappa \left\| \int_s^t |\sigma(r, X_r^n, \mathcal{L}(X_r^n), \alpha_r^n)|^2 dr \right\|_2^2 \right\},$$

where  $\kappa > 0$  is the constant in the BDG inequality. Using the assumptions 2.2, 2.2, 2.2, the fact that  $0 \in \mathcal{C}$  and the basic inequality (13), we obtain that

$$\begin{aligned} \|X_t^n - X_s^n\|_4^4 &\leq 4^3 \left\{ (t-s)^4 \sup_{r \in [0, T]} |b(r, 0, \delta_0, 0)|^4 + (t-s)^4 L_2^2 \sup_{r \in [0, T]} |X_r^n|^4 \right. \\ &\quad \left. + L_2^2 \left\| \int_s^t (1 + |X_r^n|^{q-1} + |\alpha_r^n|^{q-1} + \mathcal{M}_2(\mathcal{L}(X_r^n))^2 (|X_r^n| + |\alpha_r^n|) dr \right\|_4^4 \right. \\ &\quad \left. + \kappa L_1^2 \left\| \int_s^t (1 + |X_r^n|^2 + |\alpha_r^n|^2) dr \right\|_2^2 \right\}. \end{aligned}$$

Using Hölder Inequality, Young Inequality  $ab \leq \frac{q-1}{q} a^{\frac{q}{q-1}} + \frac{1}{q} b^q$ , we arrive at the following estimate, for all  $n \in \mathbb{N}$  and  $0 \leq s \leq t \leq T$

$$\begin{aligned} \mathbb{E} \left( |X_t^n - X_s^n|^4 \right) &\leq C(L, T) \left\{ (t-s)^4 \left[ \sup_{r \in [0, T]} |b(r, 0, \delta_0, 0)|^4 + \mathbb{E} \left( 1 + \sup_{r \in [0, T]} |X_r^n|^{4q} \right) \right] \right. \\ &\quad \left. + C(t-s)^{4/3} (1 + (t-s)) \right\}, \end{aligned}$$

where we used the fact that  $A$  is bounded. Note that the above constants depends upon the indicated quantities, but not on  $n \in \mathbb{N}$ .

Making use of the uniform estimate (32), the Kolmogorov continuity criterion then asserts that the sequence of probability measures  $(\mathbb{P} \circ (X^n)^{-1})_{n \in \mathbb{N}}$ , defined on the space

$$E := \left( C([0, T]; \mathbb{R}^d), \mathcal{B}(C([0, T]; \mathbb{R}^d)) \right)$$

is tight. In the same way, we can prove that the sequence on probability measures  $(\mathbb{P}_n)_{n \in \mathbb{N}} := (\mathbb{P} \circ (X^n, B^n)^{-1})_{n \in \mathbb{N}}$ , with

$$B^n(t) := \int_0^t b(s, X_s^n, \mathcal{L}(X_s^n), \alpha_s^n) ds,$$

is tight on the product space  $E \times E$ , with respect to the product topology, where for two  $E$ -valued random variables  $Z_1, Z_2$  defined on  $(\Omega, \mathcal{A}, \mathbb{P})$ ,  $\mathbb{P} \circ (Z_1, Z_2)^{-1}$  denotes the joint law of  $Z_1$  and  $Z_2$ . Thus by Prokhorov’s theorem there exists a subsequence of  $(\mathbb{P}_n)_{n \in \mathbb{N}}$ , which converges weakly to some probability measure  $\mathbb{P}^*$  on  $E \times E$ .

*Step 2: passage to the limit in the drift.* By Skorokhod’s representation theorem we can then find random variables  $\bar{X}, \bar{B}, (\bar{X}^n)_{n \in \mathbb{N}}, (\bar{B}^n)_{n \in \mathbb{N}}$  defined on some probability space  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$  and with values in  $E \times E$  such that

- $\bar{\mathbb{P}} \circ (\bar{X}^n, \bar{B}^n)^{-1} = \mathbb{P}_n$  for all  $n \in \mathbb{N}$  and  $\bar{\mathbb{P}} \circ (\bar{X}, \bar{B})^{-1} = \mathbb{P}^*$  and
- $\lim_{n \rightarrow \infty} (\bar{X}^n, \bar{B}^n) = (\bar{X}, \bar{B}), \bar{\mathbb{P}}$ -almost surely with respect to the uniform topology.

From (33) and by the definition of  $\mathbb{A}$  we get for any  $p \leq r$

$$\bar{\mathbb{E}} \left( \sup_{0 \leq t \leq T} |\bar{X}_t^n|^p \right) \leq C(p, \|X_0\|_p, L_1, L_2, L_3),$$

for some constant independent of  $n$ . Thus we can conclude by the dominated convergence theorem that

$$W_2(\mathcal{L}(\bar{X}_t^n), \mathcal{L}(\bar{X}_t))^2 \leq \mathbb{E} \left( \sup_{0 \leq t \leq T} |\bar{X}_t^n - \bar{X}_t|^2 \right) \rightarrow 0,$$

as  $n \rightarrow \infty$ . This also implies  $(\mathcal{L}(\bar{X}_t))_{t \in [0, T]} \subset \mathcal{P}_2^C$ , since  $\mathcal{P}_2^C$  is closed.

To identify the almost sure limit  $\bar{B}$ , we first claim that for each  $t \in [0, T]$

$$\bar{B}^n(t) \rightharpoonup \int_0^t b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \alpha_s) ds, \tag{36}$$

weakly in  $L^2(\bar{\Omega}; \mathbb{R}^d)$ . Indeed, by (22) and the dominated convergence theorem we have

$$\bar{\mathbb{E}} \left( \int_0^t |b(s, \bar{X}_s^n, \mathcal{L}(\bar{X}_s^n), \alpha_s^n) - b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \alpha_s^n)|^2 ds \right) \rightarrow 0.$$

Likewise, for  $h \in L^2(\bar{\Omega}; \mathbb{R}^d)$  we have by Assumption 2.3 and dominated convergence

$$\bar{\mathbb{E}} \left( \int_0^t \langle (b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \alpha_s^n) - b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \alpha_s)), h \rangle ds \right) \rightarrow 0,$$

as  $n \rightarrow \infty$ , thus proving our claim.

The desired identification then follows from (36), the Banach-Saks theorem and the uniqueness of the almost sure limit. The processes  $\bar{B}$  and  $\int_0^\cdot b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \alpha_s) ds$  being both continuous pathwise, they are indistinguishable, hence the identity

$$\bar{B}(t) = \int_0^t b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \alpha_s) ds, \tag{37}$$

for all  $t \in [0, T], \bar{\mathbb{P}}$ -almost surely.

*Step 3: identification of the martingale.* Letting  $\sigma\sigma^\dagger(t, x, \mu, \alpha) := \sigma(t, x, \mu, \alpha)\sigma(t, x, \mu, \alpha)^\dagger$  for short, similar arguments as above show that

$$\sigma\sigma^\dagger(t, \bar{X}_t^n, \mathcal{L}(\bar{X}_t^n), \alpha_t^n) \rightarrow \sigma\sigma^\dagger(t, \bar{X}_t, \mathcal{L}(\bar{X}_t), \alpha_t)$$

weakly in  $L^2([0, T] \times \bar{\Omega}; \mathbb{R}^d)$ . Since the process

$$M_t^n := X_t^n - X_0 - B^n(t) = \int_0^t \sigma(s, X_s^n, \mathcal{L}(X_s^n), \alpha_s^n) dW_s$$

is, for each  $n$ , a  $\mathcal{G}_t^n := \sigma(X_s^n | s \leq t)$  martingale under  $\mathbb{P}$ , we can conclude that

$$\bar{M}_t^n := \bar{X}_t^n - X_0 - \bar{B}^n(t)$$

is a  $\bar{\mathcal{G}}_t^n := \sigma(\bar{X}_s^n | s \leq t)$  martingale under  $\bar{\mathbb{P}}$  with quadratic variation

$$\langle \bar{M}^n \rangle_t = \int_0^t \sigma\sigma^\dagger(s, \bar{X}_s^n, \mathcal{L}(\bar{X}_s^n), \alpha_s^n) ds.$$

From the previous considerations, we can conclude that

$$\bar{M}_t^n \rightarrow \bar{X}_t - X_0 - \int_0^t b(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{\alpha}_s) ds =: \bar{M}_t,$$

$\bar{\mathbb{P}}$ -almost surely for all  $t \in [0, T]$ . Thus by the dominated convergence theorem the process  $(\bar{M}_t)_{t \in [0, T]}$  is a  $\bar{\mathcal{G}}_t := \sigma(\bar{X}_s | s \leq t)$  martingale under  $\bar{\mathbb{P}}$  and with standard arguments we also obtain, that  $(\bar{M}_t)_{t \in [0, T]}$  has quadratic variation

$$\langle \bar{M} \rangle_t = \int_0^t \sigma\sigma^\dagger(s, \bar{X}_s, \mathcal{L}(\bar{X}_s), \bar{\alpha}_s) ds.$$

By the martingale representation theorem we can find an extended probability space  $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t \in [0, T]}, \hat{\mathbb{P}})$  with an  $m$ -dimensional brownian motion  $\hat{W}$ , such that the natural extension  $\hat{X}$  of  $\bar{X}$  satisfies  $\hat{\mathbb{P}} \circ (\hat{X}^{-1}) = \bar{\mathbb{P}} \circ (\bar{X}^{-1})$  and

$$\hat{X}_t = X_0 + \int_0^t b(s, \hat{X}_s, \mathcal{L}(\hat{X}_s), \bar{\alpha}_s) ds + \int_0^t \sigma(s, \hat{X}_s, \mathcal{L}(\hat{X}_s), \bar{\alpha}_s) d\hat{W}_s,$$

$\hat{\mathbb{P}}$ -almost surely for all  $t \in [0, T]$ .

*Step 4: end of the proof* It remains to show that the infimum is attained for  $\alpha$ . Due to the uniqueness of equation (7), we have  $\mathbb{P} \circ (X^\alpha)^{-1} = \hat{\mathbb{P}} \circ (\hat{X}^{-1})$ . Using Fatou’s lemma, continuity of  $f, g$ , Assumption 2.3 and Remark 2.1, we obtain

$$\inf_{\tilde{\alpha} \in \mathbb{A}} J(\tilde{\alpha}) = \lim_{n \rightarrow \infty} J(\alpha^n)$$



$$\begin{aligned}
 &\geq \liminf_{n \rightarrow \infty} \mathbb{E} \left( \int_0^T f(t, X_t^n, \mathcal{L}(X_t^n), \alpha_t^n) dt + g(X_T^n, \mathcal{L}(X_T^n)) \right) \\
 &= \liminf_{n \rightarrow \infty} \bar{\mathbb{E}} \left( \int_0^T f(t, \bar{X}_t^n, \mathcal{L}(\bar{X}_t^n), \alpha_t^n) dt + g(\bar{X}_T^n, \mathcal{L}(\bar{X}_T^n)) \right) \\
 &\geq \bar{\mathbb{E}} \left( \int_0^T f(t, \bar{X}_t, \mathcal{L}(\bar{X}_t), \alpha_t) dt + g(\bar{X}_T, \mathcal{L}(\bar{X}_T)) \right) \\
 &= \hat{\mathbb{E}} \left( \int_0^T f(t, \hat{X}_t, \mathcal{L}(\hat{X}_t), \alpha_t) dt + g(\hat{X}_T, \mathcal{L}(\hat{X}_T)) \right) \\
 &= \mathbb{E} \left( \int_0^T f(t, X_t^\alpha, \mathcal{L}(X_t^\alpha), \alpha_t) dt + g(X_T^\alpha, \mathcal{L}(X_T^\alpha)) \right) \\
 &= J(\alpha).
 \end{aligned}$$

This shows that  $\alpha$  has the desired properties, and hence the proof is finished. □

### 4 The Maximum Principle: Proof of Theorem 2.2

In this section, it will be assumed implicitly that assumptions 2.1, 2.2, 2.3 and 2.4 hold. Hereafter, we let  $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}})$  be a copy of the probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . The corresponding expectation map will be denoted by  $\tilde{\mathbb{E}}$ .

#### 4.1 Gâteaux Differentiability

In this subsection we aim to complete Lemma 3.2 by showing the Gâteaux-differentiability of the control-to-state operator

$$G : \mathbb{A} \subset L^p([0, T]; \mathbb{R}^k) \rightarrow \mathcal{S}^2, \quad \alpha \mapsto X^\alpha.$$

The Gâteaux derivative of the solution map will be given by the solution of a mean-field equation with random coefficients. We will deal with this problem in the similar fashion as it is done in [8, Thm. 6.10 p. 544].

**Lemma 4.1** *The solution map  $G$  is Gâteaux-differentiable. Moreover, for each  $\alpha \in \mathbb{A}$ , its derivative in the direction  $\beta \in \mathbb{A}$  is given by*

$$dG(\alpha) \cdot \beta = Z^{\alpha, \beta},$$

where, introducing

$$\begin{aligned}
 B_\mu(t, x, \mu) &:= \iint_{\mathbb{R}^d \times \mathbb{R}^d} b_\mu(t, x, \mathcal{L}(X_t), \alpha_t)(\tilde{x}) \cdot \tilde{y} \mu(d\tilde{x} \times d\tilde{y}) \\
 \Sigma_\mu(t, x, \mu) &:= \iint_{\mathbb{R}^d \times \mathbb{R}^d} \sigma_\mu(t, x, \mathcal{L}(X_t), \alpha_t)(\tilde{x}) \cdot \tilde{y} \mu(d\tilde{x} \times d\tilde{y}),
 \end{aligned}$$

the process  $Z = Z^{\alpha,\beta}$  is characterized as the unique solution to

$$\begin{cases} dZ_t = \{b_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot Z_t + b_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t + B_\mu(t, X_t, \mathcal{L}(X_t, Z_t))\}dt, \\ \quad + \{\sigma_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot Z_t + \sigma_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t + \Sigma_\mu(t, X_t, \mathcal{L}(X_t, Z_t))\}dW_t \\ Z_0 = 0. \end{cases} \tag{38}$$

**Proof** We will start by showing that (38) has a unique solution. For that purpose, we define

$$\mathcal{R} := \left\{ \mu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)), \text{ such that } \mu_t \circ p_1^{-1} = \mathcal{L}(X_t), \forall t \right\},$$

where  $p_1$  denotes the projector onto the first  $d$ -coordinates, namely

$$p_1 : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad (x, y) \mapsto x.$$

Clearly, if  $\mu_t^n$  is a sequence converging weakly to  $\mu_t$  for every  $t \in [0, T]$ , the constraint  $\mu_t^n \circ p_1^{-1} = \mathcal{L}(X_t), \forall t$  remains true for  $\mu$  itself. Since the Wasserstein distance metrizes the weak topology, we see that  $\mathcal{R}$  is closed in  $C([0, T]; \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d))$ . Next, define

$$\Psi : \mathcal{R} \rightarrow \mathcal{R},$$

which maps  $\mu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d))$  to  $(\mathcal{L}(X_t, V_t))_{t \in [0, T]}$ , where  $(V_t)_{t \in [0, T]}$  is the unique solution to

$$\begin{cases} dV_t = b_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot V_t + b_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t + B_\mu(t, X_t, \mu_t)dt \\ \quad + \sigma_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot V_t + \sigma_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t + \Sigma_\mu(t, X_t, \mu_t)dW_t \\ Z_0 = 0. \end{cases} \tag{39}$$

For fixed  $\mu \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d))$  we first need to check the existence of a unique solution  $V$ . But letting

$$\begin{aligned} B(t, \omega, v, \mu, \alpha) &:= b_x(t, X_t(\omega), \mathcal{L}(X_t), \alpha_t) \cdot v + b_\alpha(t, X_t(\omega), \mathcal{L}(X_t), \alpha_t) \\ &\quad \cdot \beta_t + B_\mu(t, X_t(\omega), \mu_t), \\ \Sigma(t, \omega, v, \mu, \alpha) &:= \sigma_x(t, X_t(\omega), \mathcal{L}(X_t), \alpha_t) \cdot v + \sigma_\alpha(t, X_t(\omega), \mathcal{L}(X_t), \alpha_t) \\ &\quad \cdot \beta_t + \Sigma_\mu(t, X_t(\omega), \mu_t), \end{aligned}$$

we have the following properties:

$$\begin{aligned} \langle B(t, \omega, v, \mu, \alpha) - B(t, \omega, v', \mu, \alpha), v - v' \rangle &\leq A_1 |v - v'|^2 \\ \int_0^T \sup_{|v| \leq c} |B(t, \omega, v, \mu, \alpha)| dt &< \infty, \quad \forall c \geq 0, \end{aligned}$$

for all  $t \in [0, T]$  and  $\mathbb{P}$ -almost every  $\omega$ . The first estimate is a result of Assumption 2.4 and the fact that  $\mathbb{P}(X_t \in C, \forall t) = 1$ . The second estimate follows from

$$|B(t, \omega, v, \mu, \alpha)| \leq C \left\{ (1 + |X_t(\omega)|^{q-1})|v| + |\beta_t| + (1 + |X_t(\omega)|) \iint_{\mathbb{R}^d \times \mathbb{R}^d} |y| \mu_t(dx \times dy) \right\},$$

together with the continuity of  $t \mapsto \iint_{\mathbb{R}^d \times \mathbb{R}^d} |y| \mu_t(dx \times dy)$ , and the uniform estimate (32). Using (30) we get with similar arguments

$$|\Sigma(t, \omega, v, \mu, \alpha) - \Sigma(t, \omega, v', \mu, \alpha)| \leq A_2|v - v'|, \\ \int_0^T \sup_{|v| \leq c} |\Sigma(t, \omega, v, \mu, \alpha)|^2 dt < \infty, \quad \forall c \geq 0,$$

for all  $t \in [0, T]$ ,  $\mathbb{P}$ -almost every  $\omega$ . It follows then by classical SDE results that (39) is well-posed. Moreover, adapting the arguments yielding the moment estimates of Theorem 3.1, it is shown *mutatis mutandis* that for  $2 \leq p \leq r$

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} |V_t|^p \right) < \infty.$$

Therefore  $(V_t)$  (and hence  $\Psi(\mu) \equiv \mathcal{L}(X, V)$ ) is uniquely determined by the probability measure  $\mu$ .

We now aim to prove that  $\Psi$  is a contraction, but for that purpose it is convenient to introduce another (stronger) metric. For any  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$  with  $\mu \circ p_1^{-1} = \nu \circ p_1^{-1}$ , we let

$$d(\mu, \nu)^2 := \inf_{m \in \Lambda(\mu, \nu)} \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} |v - w|^2 m(dx \times dv \times dw),$$

where  $\Lambda(\mu, \nu)$  is the set of all probability measures  $m$  on  $(\mathbb{R}^d)^3$  such that for any  $A, B \in \mathcal{B}(\mathbb{R}^d)$

$$m(A \times B \times \mathbb{R}^d) = \mu(A \times B) \quad \text{and} \quad m(A \times \mathbb{R}^d \times B) = \nu(A \times B).$$

That  $d$  is stronger than  $W_2$  can be seen as follows. If  $m$  is any element in  $\Lambda(\mu, \nu)$ , one can define

$$\rho(dx \times dv \times dy \times dw) := m(dx \times dv \times dw) \delta_x(dy)$$

where  $\delta_x$  is the Dirac mass centered at  $x$ . Clearly,  $\rho$  belongs to the set of transport plans  $\Pi(\mu, \nu)$  between  $\mu$  and  $\nu$ , so that in particular

$$\begin{aligned} W_2(\mu, \nu) &= \inf_{\rho \in \Pi(\mu, \nu)} \iiint_{(\mathbb{R}^d)^4} (|x - y|^2 + |v - w|^2) \pi(dx \times dv \times dy \times dw) \\ &\leq \iiint_{(\mathbb{R}^d)^3} |v - w| m(dx \times dv \times dw). \end{aligned}$$

Then, taking the infimum over all such  $m$  yields our conclusion.

Next, let  $m \in \Lambda(\mu, \nu)$ . Using the marginal condition on  $m$ , we have

$$\begin{aligned} &|B_\mu(t, X_t, \mu_t) - B_\mu(t, X_t, \nu_t)| \\ &= \left| \iint_{\mathbb{R}^d \times \mathbb{R}^d} b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(x) \cdot v \mu(dx \times dv) \right. \\ &\quad \left. - \iint_{\mathbb{R}^d \times \mathbb{R}^d} b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot w \nu(dx \times dw) \right| \\ &= \left| \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(x) \cdot v m(dx \times dv \times dw) \right. \\ &\quad \left. - \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot w m(dx \times dv \times dw) \right|. \end{aligned}$$

Thus,

$$\begin{aligned} |B_\mu(t, X_t, \mu_t) - B_\mu(t, X_t, \nu_t)| &= \left| \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(x) \cdot (v - w) \right. \\ &\quad \left. m(dx \times dv \times dw) \right|. \end{aligned}$$

Since  $m$  is arbitrary, we obtain

$$|B_\mu(t, X_t, \mu_t) - B_\mu(t, X_t, \nu_t)| \leq A_1(1 + |X_t|)d(\mu_t, \nu_t),$$

and a similar result can be shown for  $\Sigma_\mu$ . Now, if we equip  $\mathcal{R}$  with a metric  $\delta$  inherited from  $d$ , for instance  $\delta(\mu, \nu) := \sup_{t \in [0, T]} e^{-\gamma t} d(\mu_t, \nu_t)$  for  $\gamma > 0$  large enough, the proof that  $\Psi$  is a contraction follows with simple arguments. Since it is similar to the proof of Theorem 3.1, we omit the details.

Let now  $\alpha, \beta \in \mathbb{A}$  and  $\epsilon > 0$  small enough, such that  $\alpha + \epsilon\beta \in \mathbb{A}$ . By  $X$  we denote the solution of (7) with respect to  $\alpha$  and by  $X^\epsilon$  we denote the solution to (7) with respect to  $\alpha + \epsilon\beta$ . Furthermore for  $\lambda \in [0, 1]$  we introduce  $X^{\lambda, \epsilon} := X + \lambda(X^\epsilon - X)$  and  $\alpha^{\lambda, \epsilon} := \alpha + \lambda\epsilon\beta$ . Note that, since  $\pi$  is convex, we have

$$\pi(X_t + \lambda(X_t^\epsilon - X_t)) = \pi((1 - \lambda)X_t + \lambda X_t^\epsilon) \leq (1 - \lambda)\pi(X_t) + \lambda\pi(X_t^\epsilon) \leq 0, \tag{40}$$

hence  $X_t^{\lambda, \epsilon}$  is supported in  $\mathcal{C}$ .

Next, by Lemma 3.2 we get

$$\mathbb{E} \left( \sup_{\lambda \in [0, 1]} \sup_{t \in [0, T]} |X_t^{\lambda, \epsilon} - X_t|^2 \right) \leq \hat{C}_{L, T} \epsilon^2 \int_0^T |\beta_t|^2 dt.$$

Thus, we can conclude that  $X^{\lambda, \epsilon} \xrightarrow{\epsilon \rightarrow 0} X$  in  $L^2(\Omega, C([0, T]; \mathbb{R}^d))$ , uniformly in  $\lambda$ . By a simple Taylor expansion we get

$$\begin{aligned} & b(t, X_t^\epsilon, \mathcal{L}(X_t^\epsilon), \alpha_t + \epsilon \beta_t) \\ &= b(t, X_t, \mathcal{L}(X_t), \alpha_t) + [b_x]_t^\epsilon \cdot (X_t^\epsilon - X_t) + \epsilon [b_\alpha]_t^\epsilon \cdot \beta_t + \tilde{\mathbb{E}} \left( [b_\mu]^\epsilon \cdot \widetilde{(X_t^\epsilon - X_t)} \right) \end{aligned}$$

where, given  $\varphi = \varphi(t, x, \mu, \alpha)(\tilde{x})$  we use the shorthand notation

$$[\varphi]_t^\epsilon := \int_0^1 \varphi \left( t, X_t^{\lambda, \epsilon}, \mathcal{L}(X_t^{\lambda, \epsilon}), \alpha_t^{\lambda, \epsilon} \right) \left( \tilde{X}_t^{\lambda, \epsilon} \right) d\lambda,$$

with the convention that the last input is ignored whenever  $\varphi$  does not depend on the tilde variable. Similarly, we have

$$\begin{aligned} & \sigma(t, X_t^\epsilon, \mathcal{L}(X_t^\epsilon), \alpha_t + \epsilon \beta_t) \\ &= \sigma(t, X_t, \mathcal{L}(X_t), \alpha_t) + [\sigma_x]_t^\epsilon \cdot (X_t^\epsilon - X_t) + \epsilon [\sigma_\alpha]_t^\epsilon \cdot \beta_t + \tilde{\mathbb{E}} \left( [\sigma_\mu]^\epsilon \cdot \widetilde{(X_t^\epsilon - X_t)} \right). \end{aligned}$$

Thus, for  $\Delta_t^\epsilon := \frac{X_t^\epsilon - X_t}{\epsilon} - Z_t^{\alpha, \beta}$  we have

$$\begin{aligned} d\Delta_t^\epsilon &= \left\{ [b_x]_t^\epsilon \cdot \Delta_t^\epsilon + \tilde{\mathbb{E}} \left( [b_\mu]_t^\epsilon \cdot \tilde{\Delta}_t^\epsilon \right) + ([b_x]_t^\epsilon - b_x(t, X_t, \mathcal{L}(X_t), \alpha_t)) \cdot Z_t^{\alpha, \beta} \right. \\ &\quad + \epsilon ([b_\alpha]_t^\epsilon - b_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t)) \cdot \beta_t \\ &\quad \left. + \tilde{\mathbb{E}} \left( ([b_\mu]_t^\epsilon - b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t)) \cdot \tilde{Z}_t^{\alpha, \beta} \right) \right\} dt \\ &+ \left\{ [\sigma_x]_t^\epsilon \cdot \Delta_t^\epsilon + \tilde{\mathbb{E}} \left( [\sigma_\mu]_t^\epsilon \cdot \tilde{\Delta}_t^\epsilon \right) + ([\sigma_x]_t^\epsilon - \sigma_x(t, X_t, \mathcal{L}(X_t), \alpha_t)) \cdot Z_t^{\alpha, \beta} \right. \\ &\quad + \epsilon ([\sigma_\alpha]_t^\epsilon - \sigma_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t)) \cdot \beta_t \\ &\quad \left. + \tilde{\mathbb{E}} \left( ([\sigma_\mu]_t^\epsilon - \sigma_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t)) \cdot \tilde{Z}_t^{\alpha, \beta} \right) \right\} dW_t. \end{aligned}$$

By Itô formula, (40) and Assumption 2.4, we get

$$d \left( \frac{|\Delta_t^\epsilon|^2}{2} \right) \leq \left\{ A_1 |\Delta_t^\epsilon|^2 + \tilde{\mathbb{E}} \left( |[b_\mu]_t^\epsilon| |\tilde{\Delta}_t^\epsilon| \right) \right\} |\Delta_t^\epsilon|$$

$$\begin{aligned}
 & + \left\{ \mathbb{E} \left( \left| [b_x]_t^\epsilon - b_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \right| |Z_t^{\alpha, \beta}| \right) \right. \\
 & + \mathbb{E} \left( \left| [b_\alpha]_t^\epsilon - b_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \right| |\beta_t| \right) \\
 & + \mathbb{E} \left( \left| [b_\mu]_t^\epsilon - b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t) \right| |\tilde{Z}_t^{\alpha, \beta}| \right) \left. \right\} |\Delta_t^\epsilon| \Big\} dt \\
 & + \left\{ \Delta_t^\epsilon \cdot \left( [\sigma_x]_t^\epsilon \cdot \Delta_t^\epsilon + \mathbb{E} \left( \sigma_\mu^\epsilon(\Delta) \cdot \tilde{\Delta}_t^\epsilon \right) + ([\sigma_x]_t^\epsilon - \sigma_x(t, X_t, \mathcal{L}(X_t), \alpha_t)) \cdot Z_t^{\alpha, \beta} \right. \right. \\
 & + \epsilon([\sigma_\alpha]_t^\epsilon - \sigma_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t)) \cdot \beta_t \\
 & + \mathbb{E} \left( ([\sigma_\mu]_t^\epsilon - \sigma_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t)) \cdot \tilde{Z}_t^{\alpha, \beta} \right) \left. \right) dW_t \Big\} \\
 & + \frac{5}{2} \left\{ A_2 |\Delta_t^\epsilon|^2 + \left( \int_0^1 A_2 (1 + |X_t^{\lambda, \epsilon}|^2) d\lambda \right) \mathbb{E} \left( |\tilde{\Delta}_t^\epsilon|^2 \right) \right. \\
 & + |[\sigma_x]_t^\epsilon - \sigma_x(t, X_t, \mathcal{L}(X_t), \alpha_t)|^2 |Z_t^{\alpha, \beta}|^2 \\
 & + \epsilon^2 |[\sigma_\alpha]_t^\epsilon - \sigma_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t)|^2 |\beta_t|^2 \\
 & + \mathbb{E} \left( |[\sigma_\mu]_t^\epsilon - \sigma_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t)| |\tilde{Z}_t^{\alpha, \beta}| \right)^2 \left. \right\} dt.
 \end{aligned}$$

By Young Inequality, Jensen Inequality and assumption 2.4 we have

$$\begin{aligned}
 \mathbb{E} \left( \left| [b_\mu]_t^\epsilon \right| |\tilde{\Delta}_t^\epsilon| \right) |\Delta_t^\epsilon| & \leq \frac{1}{2} \left( \mathbb{E} \left( | [b_\mu]_t^\epsilon |^2 |\tilde{\Delta}_t^\epsilon|^2 \right) + |\Delta_t^\epsilon|^2 \right) \\
 & \leq \frac{A_1}{2} \left( \int_0^1 (1 + |X_t^{\lambda, \epsilon}|^2) d\lambda \right) \mathbb{E} \left( |\tilde{\Delta}_t^\epsilon|^2 \right) + \frac{1}{2} |\Delta_t^\epsilon|^2.
 \end{aligned}$$

Since  $\epsilon > 0$  is chosen in a way that  $\alpha + \epsilon\beta \in \mathbb{A}$ , we can conclude by the a priori bound (32) and the definition of  $\mathbb{A}$ , that

$$\mathbb{E} \left( \sup_{s \in [0, t]} \mathbb{E} \left( \left| [b_\mu]_s^\epsilon \right| |\tilde{\Delta}_s^\epsilon| \right) \left| \Delta_s^\epsilon \right. \right) \leq C(T, \|X_0\|_p) \mathbb{E} \left( \sup_{s \in [0, t]} |\tilde{\Delta}_s^\epsilon|^2 \right) + \mathbb{E} \left( \sup_{s \in [0, t]} \frac{|\Delta_s^\epsilon|^2}{2} \right),$$

for some constant  $C(T, \|X_0\|_p) > 0$  which does not depend on  $\epsilon$ . By the Burkholder-Davis-Gundy inequality, Young and Jensen inequalities we arrive at

$$\mathbb{E} \left( \sup_{t \in [0, T]} |\Delta_t^\epsilon|^2 \right) \leq I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + C \int_0^T \mathbb{E} \left( \sup_{s \in [0, t]} |\Delta_s^\epsilon|^2 \right) ds,$$

for a constant  $C > 0$  which does not depend on  $\epsilon$  and

$$\begin{aligned}
 I_1 & = \mathbb{E} \left( \int_0^T | [b_x]_t^\epsilon - b_x(t, X_t, \mathcal{L}(X_t), \alpha_t) |^2 |Z_t^{\alpha, \beta}|^2 dt \right) \\
 I_2 & = \epsilon^2 \mathbb{E} \left( \int_0^T | [b_\alpha]_t^\epsilon - b_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) |^2 |\beta_t|^2 dt \right) \\
 I_3 & = \mathbb{E} \left( \int_0^T \mathbb{E} \left( | [b_\mu]_t^\epsilon - b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t) |^2 |\tilde{Z}_t^{\alpha, \beta}|^2 \right) dt \right)
 \end{aligned}$$

and  $I_4, I_5, I_6$  are analogues for  $\sigma$ . We will only show  $I_1 \rightarrow 0$  as  $\epsilon \rightarrow 0$ , the other terms being handled by similar arguments. By assumption 2.4 we have

$$|[b_x]_t^\epsilon - b_x(t, X_t, \mathcal{L}(X_t), \alpha_t)|^4 \leq C(1 + |X_t|^{4q-4} + |X_t^{\lambda, \epsilon}|^{4q-4}).$$

Furthermore we have for any  $p \leq r$  that

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t^{\lambda, \epsilon}|^p \right) \leq C_p \left\{ (1 + \lambda^p) \mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t^\epsilon|^p \right) + \lambda^p \mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t|^p \right) \right\},$$

is bounded from above by some constant that does not depend on  $\epsilon$  for  $\epsilon > 0$  small enough. Since  $X^{\lambda, \epsilon} \rightarrow X$  in  $L^2(\Omega; C([0, T]; \mathbb{R}^d))$ , by the a-priori bound (32), the estimate  $\mathbb{E}(\sup_{t \in [0, T]} |Z_t|^4) < \infty$ , the continuity of  $b_x$  and the dominated convergence theorem, one concludes that  $I_1 \rightarrow 0$  as  $\epsilon \rightarrow 0$ . Similar arguments combined with Gronwall’s lemma finish the proof.  $\square$

As an important consequence, we obtain the following formula for the Gâteaux derivative of the cost functional. Given Lemma 4.1, the next result is proven in the same way as it is done in [8] and thus omitted.

**Corollary 4.1** *The cost functional*

$$J : \mathbb{A} \rightarrow \mathbb{R}$$

is Gâteaux differentiable and its Gâteaux derivative at  $\alpha \in \mathbb{A}$  in direction  $\beta \in \mathbb{A}$  is given by

$$\begin{aligned} dJ(\alpha) \cdot \beta &= \mathbb{E} \left( f_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot Z_t^{\alpha, \beta} + f_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t \right) \\ &+ \mathbb{E} \left( \tilde{\mathbb{E}} \left( f_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t) \cdot \tilde{Z}_t^{\alpha, \beta} \right) \right) \\ &+ \mathbb{E} \left( g_x(X_T, \mathcal{L}(X_T)) \cdot Z_T^{\alpha, \beta} + \tilde{\mathbb{E}} \left( g_\mu(X_T, \mathcal{L}(X_T))(\tilde{X}_T) \cdot \tilde{Z}_T^{\alpha, \beta} \right) \right). \end{aligned}$$

**4.2 Maximum Principle**

For the reader’s convenience, we now rewrite the adjoint equation (29) using Hamiltonian formalism. Recall that for  $x, y, p \in \mathbb{R}^d, q \in \mathbb{R}^{d \times m} \mu \in \mathcal{P}_2$  and  $\alpha \in A$ , we introduced the quantity

$$H(t, x, \mu, p, q, \alpha) := \langle b(t, x, \mu, \alpha), p \rangle + \langle \sigma(t, x, \mu, \alpha), q \rangle + f(t, x, \mu, \alpha).$$

Thus, given a control  $\alpha \in \mathbb{A}$ , one sees that the pair  $(P, Q) \in \mathcal{S}^{2,d} \times \mathcal{H}^{2,d \times m}$  solves the adjoint equation if and only if for all  $t \in [0, T]$ ,  $\mathbb{P}$ -almost surely

$$\begin{cases} dP_t = - \left[ H_x(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha_t) + \tilde{\mathbb{E}} \left( H_\mu(t, \tilde{X}_t, \mathcal{L}(X_t), \tilde{P}_t, \tilde{Q}_t, \alpha_t)(X_t) \right) \right] dt + Q_t dW_t \\ P_T = g_x(X_T, \mathcal{L}(X_T)) + \tilde{\mathbb{E}} \left( g_\mu(\tilde{X}_T, \mathcal{L}(X_T))(X_T) \right). \end{cases} \tag{41}$$

where  $(\tilde{X}, \tilde{P}, \tilde{Q}, \tilde{\alpha})$  is an independent copy of  $(X, P, Q, \alpha)$  on the space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ .

Let us point out that the above coefficients fail to satisfy [8, Assumption MKV BSDE, Chap. 4]. Hence, we first need to address the solvability of the BSDE (41) under the assumptions of Theorem 2.2.

**Lemma 4.2** *Under the assumptions of Theorem 2.2, there exists a unique solution  $(P, Q) \in \mathcal{S}^2 \times \mathcal{H}^{2,d \times m}$  of (41).*

**Proof** Fix  $\alpha \in \mathbb{A}$  and for simplicity, denote by  $H_x(t, \omega, p, q) := H_x(t, X_t(\omega), \mathcal{L}(X_t), p, q, \alpha_t)$  and by  $H_\mu(t, \omega, x, p, q) := H_\mu(t, x, \mathcal{L}(X_t), p, q, \alpha_t)(X_t(\omega))$ . Consider the map  $\Gamma : \mathcal{H}^{2,d} \times \mathcal{H}^{2,d \times m} \rightarrow \mathcal{H}^{2,d} \times \mathcal{H}^{2,d \times m}$  which maps a given pair

$$(Y, Z) \in \mathcal{H}^{2,d} \times \mathcal{H}^{2,d \times m}$$

to the solution  $(P, Q)$  of

$$\begin{cases} dP_t = - \left[ H_x(t, \omega, P_t, Q_t) + \mathbb{E} \left( H_\mu(t, \omega, X_t, Y_t, Z_t) \right) \right] dt + Q_t dW_t \\ P_T = g_x(X_T(\omega), \mathcal{L}(X_T)) + \tilde{\mathbb{E}} \left( g_\mu(\tilde{X}_T, \mathcal{L}(X_T))(X_T) \right), \end{cases} \tag{42}$$

where the expectation is to be understood in the following way:

$$\mathbb{E} \left( H_\mu(t, \omega, X_t, Y_t, Z_t) \right) = \int_{\Omega} H_\mu(t, \omega, X_t(\omega'), Y_t(\omega'), Z_t(\omega')) \mathbb{P}(d\omega').$$

In the following we drop the dependence on  $\omega$  for  $H_\mu$ .

Since the above equation is a standard backward SDE with monotone coefficients, the existence of a solution is well-known by standard results. We will now show that the map  $\Gamma$  is a contraction, when the space  $\mathcal{H}^{2,d} \times \mathcal{H}^{2,d \times m}$  is equipped with the norm

$$\| (P, Q) \|_\gamma := \left( \int_0^T e^{\gamma t} (\|P_t\|_2^2 + \|Q_t\|_2^2) dt \right)^{1/2},$$

for a sufficiently large parameter  $\gamma > 0$ . If we denote by  $(P^1, Q^1), (P^2, Q^2)$  two solutions of (42) for  $(Y^1, Z^1)$  and  $(Y^2, Z^2)$  respectively, then by the backward Itô Formula [20, p. 356] applied to  $e^{\gamma t} |P_t^1 - P_t^2|^2$  we get

$$|P_t^1 - P_t^2|^2 + \mathbb{E} \left( \int_t^T \gamma e^{\gamma(r-t)} |P_r^1 - P_r^2|^2 dr \mid \mathcal{F}_t \right)$$



$$\begin{aligned}
 & +\mathbb{E}\left(\int_t^T e^{\gamma(r-t)}|Q_r^1 - Q_r^2|^2 dr \mid \mathcal{F}_t\right) \\
 & \leq 2\mathbb{E}\left(\int_t^T e^{\gamma(r-t)}\left\{(H_x(t, P_r^1, Q_r^1) - H_x(r, P_r^2, Q_r^2)) \cdot (P_r^1 - P_r^2) + |P_r^1 - P_r^2| \times \right. \right. \\
 & \left. \left. \int_{\Omega} |H_{\mu}(r, X_r(\omega'), Y_r^1(\omega'), Z_r^1(\omega')) - H_{\mu}(r, X_r(\omega'), Y_r^2(\omega'), Z_r^2(\omega'))| \mathbb{P}(d\omega')\right\} dr \mid \mathcal{F}_t\right).
 \end{aligned}
 \tag{43}$$

From assumptions 2.4,2.4, Young’s inequality and Lemma 3.1, we infer that

$$\begin{aligned}
 & \|(H_x(t, \omega, P^1, Q^1) - H_x(t, \omega, P^2, Q^2)) \cdot (P^1 - P^2)\|_1 \\
 & \leq (A_1 + A_2^2)\|P^1 - P^2\|_2^2 + \frac{1}{4}\|Q^1 - Q^2\|_2^2
 \end{aligned}$$

and

$$\begin{aligned}
 & \int_{\Omega} |H_{\mu}(r, X_r(\omega'), Y_r^1(\omega'), Z_r^1(\omega')) - H_{\mu}(r, X_r(\omega'), Y_r^2(\omega'), Z_r^2(\omega'))| \mathbb{P}(d\omega') \\
 & \leq (A_1 \vee A_2)(1 + \|X_r\|_2^2)^{\frac{1}{2}}(\|Y_r^1 - Y_r^2\|_2 + \|Z_r^1 - Z_r^2\|_2).
 \end{aligned}$$

Invoking (32), Cauchy-Schwarz and Young Inequalities, we can conclude that

$$\begin{aligned}
 & \int_0^T \gamma e^{\gamma r} \|P_r^1 - P_r^2\|_2^2 dr + \int_0^T e^{\gamma r} \|Q_r^1 - Q_r^2\|_2^2 dr \\
 & \leq 2(A_1 + A_2^2) \int_0^T e^{\gamma r} \|P_r^1 - P_r^2\|_2^2 dr + \frac{1}{2} \int_0^T e^{\gamma r} \|Q_r^1 - Q_r^2\|_2^2 dr \\
 & + \frac{1}{2} \int_0^T e^{\gamma r} \|P_r^1 - P_r^2\|_2^2 dr + C(A, \|X_0\|_2) \int_0^T e^{\gamma r} (\|Y_r^1 - Y_r^2\|_2^2 + \|Z_r^1 - Z_r^2\|_2^2) dr.
 \end{aligned}$$

For  $\gamma$  large enough this leads to

$$\|\|(P^1 - P^2, Q^1 - Q^2)\|\|_{\gamma}^2 \leq \frac{1}{2}\|\|(Y_r^1 - Y_r^2, Z_r^1 - Z_r^2)\|\|_{\gamma}^2,$$

showing that  $\Gamma$  is a contraction. The conclusion follows. □

The following corollary follows immediately by integration by parts and an application of Fubini Theorem. We therefore omit the proof and refer to [8, Lemma 6.12 p. 547].

**Corollary 4.2** *Let  $(P, Q)$  be a solution to (41), then it holds*

$$\begin{aligned}
 & \mathbb{E}\left(\langle P_T, Z_T^{\alpha,\beta} \rangle\right) = \mathbb{E}\left(\int_0^T \langle P_t, b_{\alpha}(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta \rangle + \langle Q_t, \sigma_{\alpha}(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta \rangle dt\right) \\
 & - \mathbb{E}\left(\int_0^T f_x(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot Z_t^{\alpha,\beta} + \tilde{\mathbb{E}}\left(f_{\mu}(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t) \cdot \tilde{Z}_t^{\alpha,\beta}\right)\right).
 \end{aligned}
 \tag{44}$$

**Remark 4.1** An immediate consequence of (44) is the following formula for the Gâteaux derivative of the cost functional

$$\begin{aligned} dJ(\alpha) \cdot \beta &= \mathbb{E} \left( \int_0^T \left\{ \langle b_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t, P_t \rangle + \langle \sigma_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t, Q_t \rangle \right. \right. \\ &\quad \left. \left. + f_\alpha(t, X_t, \mathcal{L}(X_t), \alpha_t) \cdot \beta_t \right\} dt \right) \\ &= \mathbb{E} \left( \int_0^T H_\alpha(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha_t) \cdot \beta_t dt \right). \end{aligned}$$

An application of Fubini Theorem then leads to the following representation for the gradient of  $J$ :

$$\nabla J(\alpha)_t = \mathbb{E} (H_\alpha(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha_t)) , \quad t \in [0, T]. \tag{45}$$

It is hardly necessary to mention that the formula (45) is of fundamental importance for numerical purposes, see Sect. 5 below.

We are now in position to prove the maximum principle.

**Proof of Theorem 2.2** Let  $\alpha \in \mathbb{A}$  be an optimal control for (SM),  $X$  the corresponding solution to (7) and  $(P, Q)$  the associated solution to (41). For  $\beta \in \mathbb{A}$  we have by the optimality of  $\alpha$

$$dJ(\alpha) \cdot (\beta - \alpha) = \langle \nabla J(\alpha), \beta - \alpha \rangle_{L^2([0, T]; \mathbb{R}^k)} \geq 0.$$

Invoking the convexity of the Hamiltonian (see Assumption 2.3), we get

$$\int_0^T \mathbb{E} \left( H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha_t) - H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \beta_t) \right) dt \geq 0.$$

For any arbitrary measurable set  $C \subset [0, T]$  and  $\tilde{\alpha} \in A$  we can define the admissible control

$$\beta_t = \begin{cases} \tilde{\alpha} & \text{for } t \in C, \\ \alpha_t & \text{otherwise,} \end{cases}$$

hence

$$\int_0^T \mathbf{1}_C(t) \mathbb{E} (H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha_t) - H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \tilde{\alpha})) dt \geq 0.$$

Therefore we get

$$\mathbb{E} (H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \alpha_t) - H(t, X_t, \mathcal{L}(X_t), P_t, Q_t, \tilde{\alpha})) \geq 0,$$

$dt$ -almost everywhere. This proves the theorem. □

### 5 Numerical Examples

In this section we focus on the FitzHugh–Nagumo model with external noise only, i.e. the system of  $3N$  stochastic differential equations:

$$\begin{cases} dv_t^i = \left( v_t^i - \frac{(v_t^i)^3}{3} - w_t^i + \alpha_t - \frac{1}{N} \sum_{j=1}^N \bar{J}(v_t^i - V_{rev})y_t^j \right) dt + \sigma_{ext} dW_t^i \\ dw_t^i = c(v_t^i + a - bw_t^i) dt, \\ dy_t^i = (a_r S(v_t^i)(1 - y_t^i) - a_d y_t^i) dt, \end{cases} \tag{46}$$

where we recall that  $S(v) := T_{max}/[1 + \exp \lambda(v - V_T)]$ .

We are interested in controlling the average membrane potential (called in the following “local field potential”) of a network of FitzHugh–Nagumo neurons into a desired state. Our cost functional is given by

$$\begin{aligned} f(t, x, \mu, \alpha) &:= \left| \int_{\mathbb{R}^3} v \mu(dv \times dw \times dy) - \bar{v}_t \right|^2 \\ g(t, x) &:= 0, \end{aligned} \tag{47}$$

where  $(\bar{v}_t)_t$  is a certain reference profile. We should mention that the average membrane potential will only give an idea about the average activity of the network at each time. For example a high average membrane potential is an indication that a high number of neurons are in the regenerative or active phase, while a low average membrane potential means that a high number of neurons are in the absolute refractory or silent phase.

In the described case the adjoint equation is reduced to

$$\begin{cases} dP_t = - \left\{ \langle b_x(t, X_t, \mathcal{L}(X_t), \alpha_t), P_t \rangle + \tilde{\mathbb{E}} \left( \langle b_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t), \tilde{P}_t \rangle \right) \right. \\ \quad \left. + \tilde{\mathbb{E}} \left( f_\mu(t, X_t, \mathcal{L}(X_t), \alpha_t)(\tilde{X}_t) \right) \right\} dt + Q_t dW_t \\ P_T = 0. \end{cases} \tag{48}$$

In the following section we will give a short introduction on how to solve (48) numerically.

#### 5.1 Numerical Approximation of the Adjoint Equation

In general we consider the following non fully coupled MFFBSDE

$$\begin{cases} dX_t = b(t, X_t, \mathcal{L}(X_t))dt + \sigma(t, X_t, \mathcal{L}(X_t))dW_t \\ dY_t = [f(t, X_t, Y_t) + h(t, X_t, \mathcal{L}(X_t, Y_t))]dt - Z_t dW_t \\ X_0 = \xi \\ Y_T = g(X_T). \end{cases} \tag{49}$$

For the approximation of the forward component we consider an implicit Euler scheme for McKean-Vlasov equations. Since this is standard, we will not go into further details. Concerning the backward component, we consider a scheme similar to the one presented in [10]. We should mention that since we are not dealing with a fully coupled MFFBSDE, our situation is a lot easier to handle than the one treated in [10]. For a given discrete time grid  $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ , we consider the following numerical scheme:

$$\begin{aligned} Y_{t_k}^\pi &= \mathbb{E} \left( Y_{t_{k+1}}^\pi | \mathcal{F}_{t_k} \right) - (t_{k+1} - t_k) \left\{ f(t_k, X_{t_k}^\pi, Y_{t_k}^\pi) + h(t_{k+1}, X_{t_k}^\pi, \mathcal{L}(X_{t_{k+1}}^\pi, Y_{t_{k+1}}^\pi)) \right\} \\ Z_{t_k}^\pi &:= (t_{k+1} - t_k)^{-1} \mathbb{E} \left( Y_{t_{k+1}}^\pi (W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k} \right), \\ Y_{t_n}^\pi &= g(X_{t_n}^\pi), \quad Z_{t_n}^\pi = 0. \end{aligned}$$

For the approximation of the conditional expectation, we make use of the decoupling field mentioned in [8], to write

$$Y_{t_{k+1}}^\pi = u(t_{k+1}, X_{t_{k+1}}^\pi, \mathcal{L}(X_{t_{k+1}}^\pi)) =: \hat{u}(t_{k+1}, X_{t_{k+1}}^\pi).$$

Thus we can represent the conditional expectation in terms of a function  $\tilde{u}$  by

$$\mathbb{E} \left( Y_{t_{k+1}}^\pi | \mathcal{F}_{t_k} \right) = \tilde{u}(t_{k+1}, X_{t_k}^\pi).$$

We approximate  $\tilde{u}(t_{k+1}, \cdot)$  with gaussian radial basis functions, by solving the following minimization problem for fixed nodes  $x_1, \dots, x_L$ :

$$\min_{\alpha} \mathbb{E} \left( \left| Y_{t_{k+1}}^\pi - \sum_{i=1}^L \alpha_i(t_{k+1}) e^{\frac{1}{2\delta} \|X_{t_k}^\pi - x_i\|^2} \right|^2 \right),$$

for  $\alpha = (\alpha_1(t_{k+1}), \dots, \alpha_L(t_{k+1}))^\dagger$ , where  $\delta > 0$  and  $L \in \mathbb{N}$  are fixed. Therefore we initialize our reference points  $x_1, \dots, x_L$  by  $L$  independent realizations of  $X_{t_k}^\pi$ . For  $m$  realizations of  $Y_{t_{k+1}}^\pi$  and  $X_{t_k}^\pi$ , denoted by  $y_{t_{k+1}}^1, \dots, y_{t_{k+1}}^m$  and  $x_{t_{k+1}}^1, \dots, x_{t_{k+1}}^m$  respectively, we then write

$$\begin{aligned} y_{t_{k+1}} &= (y_{t_{k+1}}^1, \dots, y_{t_{k+1}}^m)^\dagger \\ A(t_k) &= (e^{\frac{1}{2\delta} \|x_{t_k}^i - x_j\|^2})_{i=1, \dots, m, j=1, \dots, L}. \end{aligned}$$

Thus we need to minimize

$$\|y_{t_{k+1}} - A(t_k)\alpha(t_{k+1})\|^2.$$

A similar approach for BSDEs can be found in [18]. There is no convergence analysis of this scheme for our assumptions on the coefficients, this should only give an idea how to solve the adjoint equation in practice. Furthermore we should mention, that in the case where only external noise is present, the duality (44) and the resulting gradient representation still holds true for any non adapted solution of (41), i.e. any solution to the random backward ODE arising from equation (41) for  $Q \equiv 0$ . Thus one can also implement a numerical scheme for the adjoint equation, without any conditional expectations involved. For more general diffusion coefficient however this is not true, since the proof of the duality is based on integration by parts and the stochastic integral which appears is not defined if the integrand is not adapted.

### 5.2 Gradient Descent Algorithm

We will now briefly sketch our gradient descent algorithm.

**Algorithm 5.1** *Take an initial control  $\alpha_0 \in \mathbb{A}$ ,  $s_0 > 0$ , and recursively for  $n = 0, 1, \dots$ :*

- (1) *determine  $X^{\alpha_n}$  by solving the state equation with an implicit particle scheme to avoid particle corruption;*
- (2) *solve the adjoint equation for given  $X^{\alpha_n}$  in order to approximate  $(P^{\alpha_n}, Q^{\alpha_n})$ ;*
- (3) *approximate the gradient*

$$\nabla J(\alpha_n)_s = \mathbb{E}\left[\langle b_\alpha(s, X_s^{\alpha_n}, \mathcal{L}(X_s^{\alpha_n}), \alpha_s^n), P_s^{\alpha_n} \rangle + f_\alpha(s, X_s^{\alpha_n}, \mathcal{L}(X_s^{\alpha_n}), \alpha_s^n) \right]$$

*via Monte-Carlo method, where  $(P^{\alpha_n}, Q^{\alpha_n})$  solves the adjoint equation;*

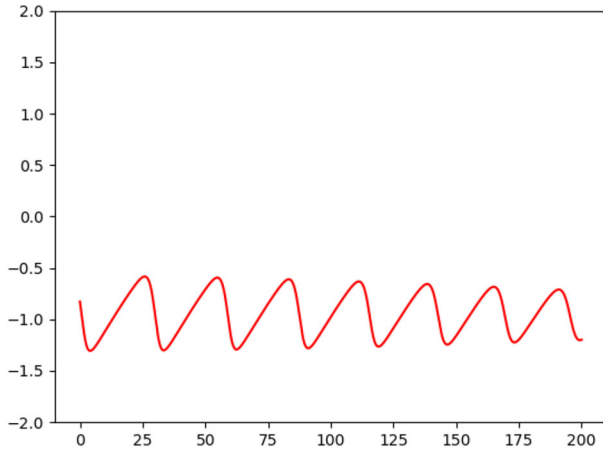
- (4) *update the step size  $s_n$  according to a suitable step size rule (e.g. Armijo-rule), repeating step 5.1–5.1 if necessary*
- (5) *update the control in direction of the steepest descent:  $\alpha_{n+1} := \alpha_n - s_n \nabla J(\alpha_n)$ ;*
- (6) *the algorithm stops if  $\|\nabla J(\alpha_n)\| < \epsilon$*

To compute the expectation term, one is in fact reduced to simulate the solution of the network equation itself and use the particles as samples for the Monte-Carlo simulation.

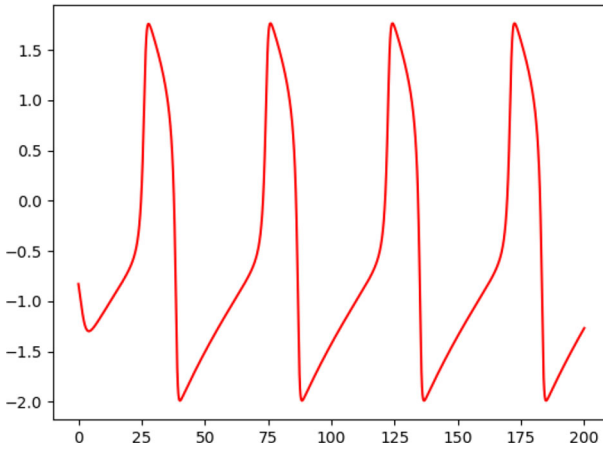
### 5.3 Numerical Examples for Systems of FitzHugh–Nagumo Neurons

Although the solution to the adjoint equation is a 3-dimensional process, in the following we will only plot its first variable, since the other variables are irrelevant for the gradient in our situation.

To illustrate some problems we had with the simulations, we consider the example of the deterministic uncoupled case of equation (46), where  $\bar{J} = 0$  and  $\sigma_{ext} = 0$ .



**Fig. 2** Membrane potential of the solution to (46) for  $\alpha \equiv 0.3251$



**Fig. 3** Reference profile generated by solving (46) for  $\alpha \equiv 0.332$

In the given situation the membrane potential  $v$  becomes highly sensitive to small perturbations of the control at specific times, when we chose the control  $\alpha_t \equiv \alpha$  close to the bifurcation value  $\alpha_c \approx 0.33$  for the supercritical Hopf bifurcation point of the equation, where the fixed point is stable, but there is no periodic orbit, in particular we chose  $\alpha = 0.3251 < \alpha_c \approx 0.33$ . The bifurcation point of the deterministic system can be determined in a similar way as it is done in [23]. This sensitivity can lead to high valued solutions of the corresponding adjoint equation for specific reference profiles. One example is to choose the reference profile as the  $v$ -trajectory of a solution to (46), for a control parameter  $\alpha$  in the limit cycle regime. This situation is illustrated by the Figs. 2, 3 and 4.

The same type of phenomena also occurs in the case of the coupled system of stochastic FitzHugh–Nagumo neurons. Here it can lead to high fluctuations of the

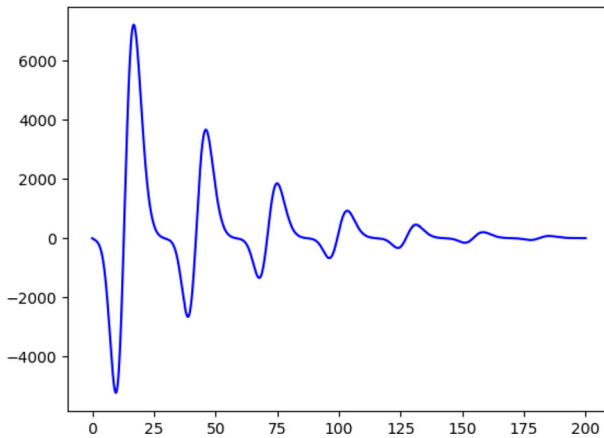


Fig. 4 Solution to the corresponding adjoint equation

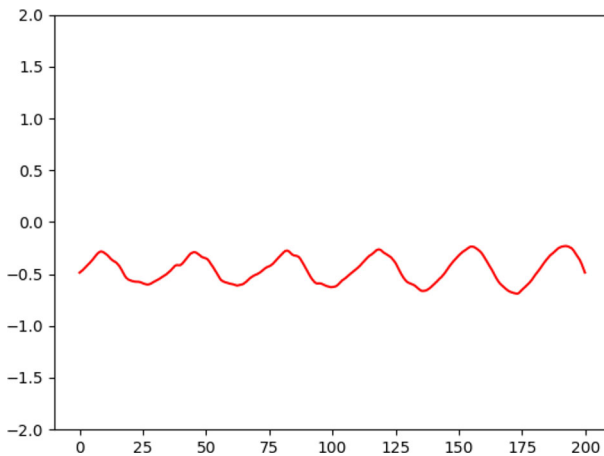
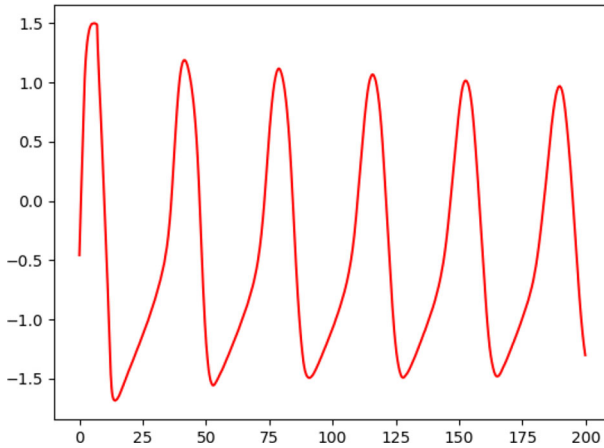


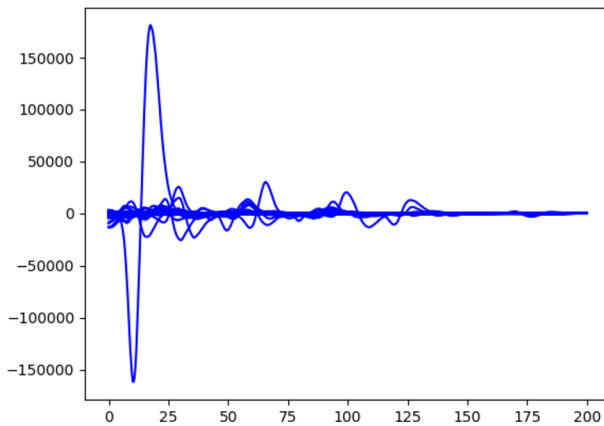
Fig. 5 Local field potential of the solution to (46) for  $\alpha \equiv 0$

sample mean for the adjoint equation, thus a high number of particles is required to compute the expectation of the solution to the adjoint equation. A small illustration is given by the Figs. 5, 6 and 7.

In this example and in the following, the initial states are uniformly distributed on the orbit of a solution to (46) with  $\alpha \equiv 0$ ,  $\sigma_{ext} = 0$  and initial conditions  $V_0 = -0.828$ ,  $w_0 = -0.139$ ,  $y_0 = 0.589$ . The other parameters are given below in Table 1. In our example only external noise is present, so we can approximate the non adapted solution to the BSDE using the simple Euler-scheme with no conditional expectations involved, replacing  $\mathbb{E} \left( Y_{t_{k+1}}^\pi | \mathcal{F}_{t_k} \right)$  by  $Y_{t_{k+1}}^\pi$  in our numerical scheme, to get a good approximation of the gradient. For the sake of completeness we mention that for the approximation of the adapted solution we would use the parameter  $L = 200$  and  $\delta = 2$ .



**Fig. 6** Reference profile chosen as the local field potential of (46) for  $\alpha(t) = 0.8$  if  $t \leq 7$



**Fig. 7** Samples of the solution to the adjoint equation

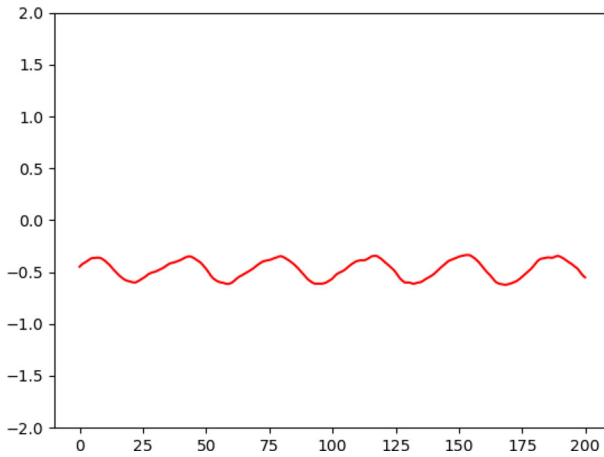
Furthermore we are always using  $N = 1000$  particles for the particle approximation of (46).

### 5.3.1 Control of a Coupled System of FitzHugh–Nagumo Neurons

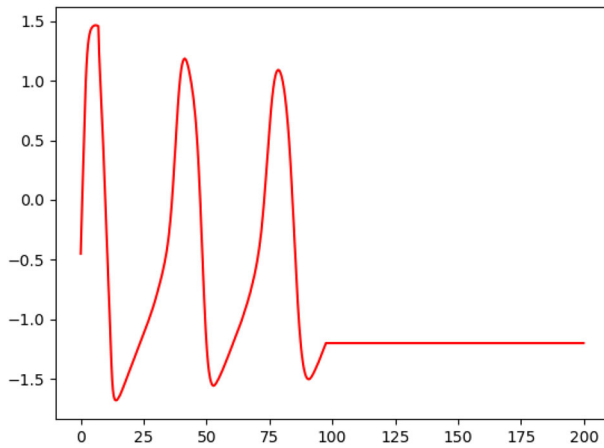
For our first example, we consider a parameter regime where the activity of a large number of neurons of the network at some time  $t$  leads to further activity at a later time, without any external current applied to the system. Therefore we slow down the gating variable, by decreasing the closing rate of the synaptic gates. This way its impact on the network is still high enough, when a large part of the network is excitable again. Figure 8 shows the uncontrolled local field potential in this case (i.e. when  $\alpha \equiv 0$ ).

Our goal is now to increase the activity of the network up to time  $t = 100$  and then control the network back into its resting potential. Up to time  $t = 100$ , the reference profile showed in Fig. 9 shows the local field potential of a network of





**Fig. 8** Uncontrolled local field potential



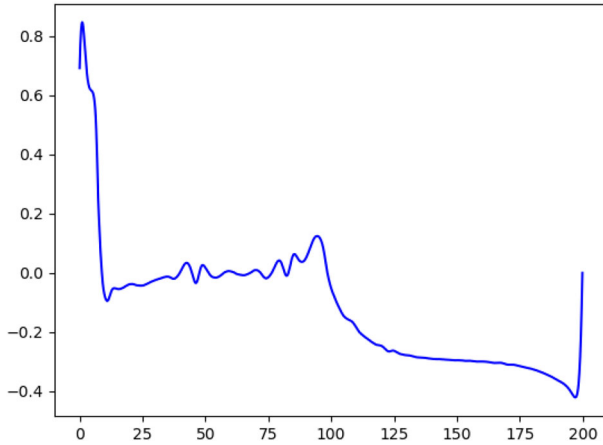
**Fig. 9** Reference profile

coupled FitzHugh–Nagumo neurons, when a constant input current of magnitude 0.8 is applied for a time period of  $\Delta t = 7$  at  $t = 0$ . For times  $t > 100$  it shows the resting potential of a single FitzHugh–Nagumo neuron.

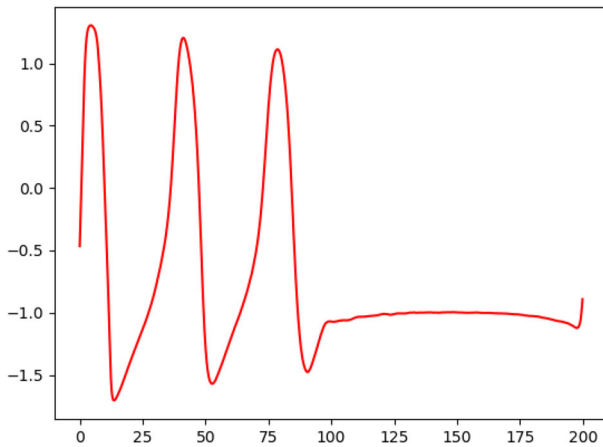
We expect the optimal control to raise the membrane potential for a small time period at  $t = 0$  and then counteract the stimulating effect of the coupling around  $t = 100$ . However this effects should not occur in the uncoupled setting, which we will consider afterwards.

Figures 10 and 11 show the optimal control and the corresponding optimal local field potential. We remind that this might only be locally optimal, since we cannot expect to find a globally optimal control with our gradient descent algorithm.

Since our terminal cost is always zero, i.e.  $g \equiv 0$ , the solution to the adjoint equation will always be zero at terminal time  $T$ . Consequently our gradient will always be zero at time  $T$  and our gradient descent algorithm does not change the control at time  $T$ .



**Fig. 10** Optimal control



**Fig. 11** Local field potential with optimal control

That is why the value of our approximated optimal control will stay at the value of our initial control  $\alpha_0$  at the terminal time. We started with initial control  $\alpha_0 \equiv 0$ , which explains the small peak in Figs. 10 and 11.

### 5.3.2 Control of an Uncoupled System of FitzHugh–Nagumo Neurons

Now we investigate the control problem for the uncoupled equation (46), where  $J = 0$ . Since the reference profile it still the same as in example 5.3.1, we will only present the corresponding optimal control (Fig. 12).

As expected, the control does not need to counteract any stimulating effects for times  $t > 100$ . Furthermore it is not sufficient in the uncoupled case to apply an input current for a small time period at  $t = 0$ , to reach the desired local field potential up to time  $t = 100$  (Table 1).

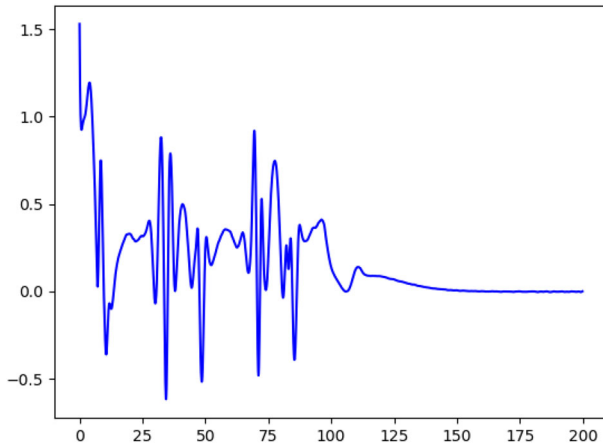


Fig. 12 Optimal control

Table 1 Parameters used for the examples

Time parameters	FitzHugh–Nagumo parameters	Synapse
$t_{end} = 200$	$a = 0.7$	$V_{rev} = 1$
$\Delta t = 0.1$	$b = 0.8$	$\bar{a} = 1$
	$c = 0.08$	$\bar{b} = 0.3$
	$\sigma_{ext} = 0.04$	$T_{max} = 1$
		$\lambda = 0.1$
		$V_{rev} = 1.2$
		$V_T = 2$
		$J = 0.46$
		$\sigma_J = 0$

**Acknowledgements** This work has been funded by Deutsche Forschungsgemeinschaft (DFG) through Grant CRC 910 “Control of self-organizing nonlinear systems: Theoretical methods and concepts of application”, Project (A10) “Control of stochastic mean-field equations with applications to brain networks.” Both authors are thankful to Wilhelm Stannat, Tilo Schwalger and François Delarue for helpful discussions. The anonymous referees are warmly thanked for having helped to improve this manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**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. Albi, G., Choi, Y.-P., Fornasier, M., Kalise, D.: Mean field control hierarchy. *Appl. Math. Optim.* **76**(1), 93–135 (2017)
2. Baladron, J., Fasoli, D., Faugeras, O., Touboul, J.: Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and Fitzhugh-Nagumo neurons. *J. Math. Neurosci.* **2**(1), 1–50 (2012)
3. Barbu, V., Cordoni, F., Di Persio, L.: Optimal control of stochastic Fitzhugh-Nagumo equation. *Int. J. Control* **89**(4), 746–756 (2016)
4. Bensoussan, A., Frehse, J., Yam, P.: *Mean Field Games and Mean Field Type Control Theory*, vol. 101. Springer, Berlin (2013)
5. Frédéric Bonnans, J., Hadikhanloo, S., Pfeiffer, L.: Schauder estimates for a class of potential mean field games of controls. *Appl. Math. Optim.* **83**, 1431–1464 (2019)
6. Bossy, M., Faugeras, O., Talay, D.: Clarification and complement to mean-field description and propagation of chaos in networks of Hodgkin-Huxley and Fitzhugh-Nagumo neurons. *J. Math. Neurosci.* **5**(1), 1–23 (2015)
7. Buckdahn, R., Li, J., Ma, J.: A stochastic maximum principle for general mean-field systems. *Appl. Math. Optim.* **74**(3), 507–534 (2016)
8. Carmona, R., Delarue, F.: *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, Berlin (2018)
9. Carmona, R., Delarue, F., Lachapelle, A.: Control of McKean-Vlasov dynamics versus mean field games. *Math. Financ. Econ.* **7**(2), 131–166 (2013)
10. Chassagneux, J., Crisan, D., Delarue, F.: Numerical method for FBSDE of McKean-Vlasov type. *Ann. Appl. Probab.* **29**(3), 1640–1684 (2019)
11. Cordoni, F., Di Persio, L.: Optimal control for the stochastic Fitzhugh-Nagumo model with recovery variable. *Evol. Equ. Control Theory* **7**(4), 571 (2018)
12. Deco, G., Rolls, E.T., Romo, R.: Stochastic dynamics as a principle of brain function. *Prog. Neurobiol.* **88**(1), 1–16 (2009)
13. Destexhe, A., Mainen, Z.F., Sejnowski, T.J.: Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *J. Comput. Neurosci.* **1**(3), 195–230 (1994)
14. Dos Reis, G., Salkeld, W., Tugaut, J.: Freidlin-Wentzell ldp in path space for McKean-Vlasov equations and the functional iterated logarithm law. *Ann. Appl. Probab.* **29**(3), 1487–1540 (2019)
15. Dudley, R.M.: *Real Analysis and Probability*. Chapman and Hall/CRC, London (2018)
16. Ermentrout, B., Terman, D.: *Mathematical Foundations of Neurosciences*, vol. 35. Springer, Berlin (2010)
17. FitzHugh, R.: Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* **1**(6), 445–466 (1961)
18. Kebiri, O., Neureither, L., Hartmann, C.: Adaptive importance sampling with forward-backward stochastic differential equations. In: *International Workshop on Stochastic Dynamics Out of Equilibrium*, pp. 265–281. Springer, Berlin (2017)
19. Nagumo, J., Arimoto, S., Yoshizawa, S.: An active pulse transmission line simulating nerve axon. *Proc. IRE* **50**(10), 2061–2070 (1962)
20. Pardoux, É., Răşcanu, A.: *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*. Springer, Berlin (2014)
21. Pfeiffer, L.: Optimality conditions for mean-field type optimal control problems. *SFB Rep.* **15**, 2015–015 (2015)
22. Pfeiffer, L.: Numerical methods for mean-field-type optimal control problems. *arXiv preprint arXiv:1703.10001* (2017)
23. Phan, V.L.E.: A study of fixed points and Hopf bifurcation of Fitzhugh-Nagumo model. *Can Tho Univ. J. Sci.* **54**(2), 112–121 (2018)
24. Schwalger, T., Deger, M., Gerstner, W.: Towards a theory of cortical columns: from spiking neurons to interacting neural populations of finite size. *PLoS Comput. Biol.* **13**(4), e1005507 (2017)
25. Yong, J., Zhou, X.Y.: *Stochastic Controls: Hamiltonian Systems and HJB Equations*, vol. 43. Springer, Berlin (1999)