



# A System of Interacting Neurons with Short Term Synaptic Facilitation

A. Galves<sup>1</sup> · E. Löcherbach<sup>2</sup> · C. Pouzat<sup>3</sup> · E. Presutti<sup>4</sup>

Received: 4 March 2019 / Accepted: 10 December 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

In this paper we present a simple microscopic stochastic model describing short term plasticity within a large homogeneous network of interacting neurons. Each neuron is represented by its membrane potential and by the residual calcium concentration within the cell at a given time. Neurons spike at a rate depending on their membrane potential. When spiking, the residual calcium concentration of the spiking neuron increases by one unit. Moreover, an additional amount of potential is given to all other neurons in the system. This amount depends linearly on the current residual calcium concentration within the cell of the spiking neuron. In between successive spikes, the potentials and the residual calcium concentrations of each neuron decrease at a constant rate. We show that in this framework, short time memory can be described as the tendency of the system to keep track of an initial stimulus by staying within a certain region of the space of configurations during a short but macroscopic amount of time before finally being kicked out of this region and relaxing to equilibrium. The main technical tool is a rigorous justification of the passage to a large population limit system and a thorough study of the limit equation.

**Keywords** Systems of spiking neurons · Short term plasticity · Piecewise deterministic Markov processes · Mean-field interaction · Biological neural nets · Interacting particle systems · Hawkes processes

---

Communicated by Eric A. Carlen.

---

✉ A. Galves  
galves@usp.br

E. Löcherbach  
eva.locherbach@univ-paris1.fr

C. Pouzat  
christophe.pouzat@parisdescartes.fr

E. Presutti  
errico.presutti@gmail.com

<sup>1</sup> Universidade de São Paulo, São Paulo, Brazil

<sup>2</sup> Université de Paris 1 Panthéon Sorbonne, Paris, France

<sup>3</sup> Université Paris Descartes, Paris, France

<sup>4</sup> Gran Sasso Science Institute, L'Aquila, Italy

## 1 Introduction

In this paper we present a simple microscopic stochastic model describing short term plasticity within a large network of interacting neurons. In this framework it is possible to describe short time memory of the system in a precise mathematical way. Namely, short time memory can be seen as the tendency of the system to keep track of an initial stimulus by staying within a certain region of the space of configurations during a short but macroscopic amount of time before finally being kicked out of this region and relaxing to equilibrium.

In our model, the successive times at which the neuron emits an action potential are described by a point process. The stochastic spiking intensity of the neuron, i.e., the infinitesimal probability of emitting an action potential during the next time unit, conditionally on the past, depends on the past history of the neuron and it is affected by the activity of other neurons in the network, either in an excitatory or an inhibitory way.

Short term synaptic plasticity (STP) refers to a change in the synaptic efficacy on timescales which are of the order of milliseconds, that is, comparable to the timescale of the spiking activity of the network. We express this through the fact that the stochastic spiking intensity also depends on the synaptic efficacy of the neuron at that time. This synaptic efficacy changes over time as a function of the residual calcium concentration within the cell. In our model the residual calcium concentration increases by one unit any time the neuron spikes and decreases at a constant rate in between successive spikes.

Since at least the last 2 decades, many papers have been devoted to STP. Probably starting with Markram and Tsodyks [13] and Tsodyks et al. [21], a lot of these papers propose relatively simple phenomenological models and study, mostly numerically, their properties. Kistler and van Hemmen [12] consider a deterministic model which is an adaptation of the model of Tsodyks and Markram [20] to the spike response model. They work within a homogenous strongly connected network and study the impact of STP on the stability of limit cycles. Our model, though stochastic, is close to this. Several recent papers are devoted to the study of the effect of STP on working memory, see Barak and Tsodyks [1], Mongillo et al. [14] and the recent article by Seeholzer et al. [18]. Finally, for a recent survey on STP, we refer the interested reader to the Scholarpedia article [22], and for a rather complete survey on the biological aspects, to Zucker and Regehr [23].

Our model can be seen as a huge system of interacting pairs of coupled Hawkes processes. Hawkes processes provide good models for systems of spiking neurons by the structure of their intensity processes and have been widely studied, see for instance Chevallier et al. [4], Chornoboy et al. [5], Hansen et al. [9], Reynaud-Bouret et al. [16] and Ditlevsen and Löcherbach [6].

In our model we make the following basic mathematical assumptions. First of all, we work within a mean-field system in which each neuron interacts with all other neurons in a homogenous way. Second, we only consider excitatory synapses. Finally, our model of STP only describes facilitation, not depression. In this framework we study in a rigorous way the intermediate time behavior of the process. This is the content of our main result, Theorem 2.5. The main step of the proof of this theorem is a rigorous justification of the passage to a large population limit model.

Describing short term memory as the tendency of the system to stay within a certain region of the state space representing some initial stimulus is not a new idea, neither is the idea that

a proof of this behavior should be done through the analysis of the limit system. Similar ideas appear already in Barak and Tsodyks [1], Mongillo et al. [14] and Seeholzer et al. [18], see also the recent paper by Schmutz et al. [17]. Nevertheless to the best of our knowledge our paper is the first in which these results are rigorously mathematically proved.

*Organisation of the paper* This paper is organised as follows. In Sect. 2, we introduce our model and state the main results of the paper. This section is complimented by a simulation study. The proofs are given in Sects. 3–7.

## 2 Overview of the Paper

### 2.1 Notation

The following notation is used throughout the paper.

- If  $M$  is a counting measure on  $\mathbb{R}_+ \times \mathbb{R}_+$ , we shall use the following notation for integration over semi-closed boxes.  $\int_0^t \int_0^\infty f(s, z) M(ds, dz) = \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} 1_{[0, t]}(s) f(s, z) M(ds, dz)$ , for any positive measurable function  $f$ .
- If  $g$  is a probability measure on  $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$  and  $\int_{\mathbb{R}_+} |f| dg < \infty$ , then we write  $\int_0^\infty f(r) g(dr)$  for the integral  $\int_{\mathbb{R}_+} f dg$ .
- For any bounded function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  we write  $\|f\|_\infty = \sup_{x \in \mathbb{R}_+} |f(x)|$ .

### 2.2 Description of the Model

We consider a system of  $N$  interacting neurons with membrane potentials  $(U_t^N(1), \dots, U_t^N(N))$  together with the corresponding residual calcium concentration within each neuron  $(R_t^N(1), \dots, R_t^N(N))$ . Each neuron  $i$ , independently of the others, spikes at rate  $\varphi(U_{t-}^N(i))$ . When spiking, it gives an additional amount  $\alpha R_{t-}^N(i)/N$  of potential to all neurons. ( $R$  is dimensionless and  $\alpha$  is a potential).  $\alpha > 0$  is a measure of the interaction strength, and the interaction is modulated by the current value of the calcium concentration of the spiking neuron. At the same time, the residual calcium concentration of the spiking neuron is increased by 1. This models the short term plasticity. In between successive spikes, the potential of each neuron decreases at rate  $\beta > 0$ , and their residual calcium concentrations decrease at constant rate  $\lambda > 0$ .  $\beta$  and  $\lambda$  are homogeneous to the inverse of a time.

To define the process, consider a family of i.i.d. Poisson random measures  $(\mathbf{M}^i(ds, dz))_{i \geq 1}$  on  $\mathbb{R}_+ \times \mathbb{R}_+$  having intensity measure  $ds dz$  each. Here,  $ds$  is homogeneous to a time and  $dz$  to the inverse of a time. Finally we consider an i.i.d. family  $(U_0^N(i), R_0^N(i))_{i=1, \dots, N}$  of  $\mathbb{R}_+^2$ -valued random variables, independent of the Poisson measures and distributed according to some probability measure  $\eta_0(du, dr)$  on  $\mathbb{R}_+^2$ . Then the system of interaction neurons is represented by the Markov process  $(U_t^N, R_t^N) = (U_t^N(1), \dots, U_t^N(N), R_t^N(1), \dots, R_t^N(N))$  taking values in  $\mathbb{R}_+^{2N}$  and solving, for  $i = 1, \dots, N$ , for  $t \geq 0$ ,

$$\begin{aligned} U_t^N(i) &= U_0^N(i) - \beta \int_0^t U_s^N(i) ds + \frac{\alpha}{N} \sum_{j=1}^N \int_0^t \int_0^\infty R_{s-}^N(j) 1_{\{z \leq \varphi(U_{s-}^N(j))\}} \mathbf{M}^j(ds, dz), \\ R_t^N(i) &= R_0^N(i) - \lambda \int_0^t R_s^N(i) ds + \int_0^t \int_0^\infty 1_{\{z \leq \varphi(U_{s-}^N(i))\}} \mathbf{M}^i(ds, dz). \end{aligned} \quad (2.1)$$

The coefficients of this system are the positive constants  $\alpha, \beta, \lambda > 0$  together with the spiking rate function  $\varphi$ . The generator of the process  $(U_t^N, R_t^N)$  is given for any smooth test function  $f : \mathbb{R}_+^{2N} \rightarrow \mathbb{R}$  by

$$\begin{aligned} A^N f(u, r) = & - \sum_{i=1}^N \left( \beta u^i \partial_{u^i} f(u, r) + \lambda r^i \partial_{r^i} f(u, r) \right) \\ & + \sum_{i=1}^N \varphi(u^i) \left[ f(\pi^i(u, r)) - f(u, r) \right], \end{aligned}$$

where

$$\pi^i(u, r)(j) := (u^j + \alpha r^i / N, r^j) \text{ if } j \neq i, \pi^i(u, r)(i) := (u^i + \alpha r^i / N, r^i + 1).$$

Notice that (2.1) is close to the system studied in [12], when taking the Heaviside function as a spiking rate (that is, spiking does only occur when reaching a fixed deterministic threshold, but when hitting this threshold, it occurs with certainty, that is, at rate  $= \infty$ ). In the present paper, following the classical reference Brillinger and Segundo [3], we will however suppose most of the time that

**Assumption 2.1**  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is bounded and Lipschitz continuous with Lipschitz constant  $L_\varphi$ . Moreover we have  $\varphi(0) = 0$ ,  $\varphi(x) > 0$  for all  $x > 0$ .

Under minimal regularity assumptions on the spiking rate, if we work at a fixed system size  $N$ , this process will die out in the long run as shows the following

**Theorem 2.2** *Grant Assumption 2.1. If  $\varphi$  is differentiable in 0, then the system stops spiking almost surely. As a consequence, the unique invariant measure of the process  $(U_t^N, R_t^N)$  is given by the Dirac measure  $\delta_{(\mathbf{0}, \mathbf{0})}$ , where  $\mathbf{0} \in \mathbb{R}^N$  denotes the all-zero vector in  $\mathbb{R}^N$ .*

This situation might however change if we consider large-population limits of the system.

## 2.3 Large Population Limits

In Sect. 5 we show that the solution  $(U_t^N, R_t^N)_{t \geq 0}$  behaves, for  $N$  large, as  $N$  independent copies of the solution  $(U_t, R_t)_{t \geq 0}$  of the following *nonlinear*, in the sense of McKean, SDE

$$\begin{aligned} U_t &= U_0 - \beta \int_0^t U_s ds + \alpha \int_0^t \mathbb{E}[\varphi(U_s) R_s] ds, \\ R_t &= R_0 - \lambda \int_0^t R_s ds + \int_0^t \int_0^\infty 1_{\{z \leq \varphi(U_{s-})\}} \mathbf{M}(ds, dz). \end{aligned} \quad (2.2)$$

In the above formula,  $(U_0, R_0)$  is an  $\eta_0$ -distributed random variable, independent of a Poisson measure  $\mathbf{M}(ds, dz)$  on  $\mathbb{R}_+ \times \mathbb{R}_+$  having intensity measure  $ds dz$ . Concerning the law  $\eta_0$  of the initial condition, in the sequel we impose

**Assumption 2.3**  $\eta_0(du, dr) = \delta_{u_0}(du) g_0(dr)$ , for some fixed  $u_0 > 0$ . Here,  $g_0$  is a probability measure on  $\mathbb{R}_+$  such that  $\int_0^\infty r^2 g_0(dr) < \infty$ .

In what follows, under Assumption 2.3, we shall write  $r_0 := \int_0^\infty r g_0(dr)$ . Under this condition, our Theorems 5.2 and 5.3 in Sect. 5 below provide an explicit coupling showing that the finite system is close to the limit system.

## 2.4 Modeling Short Term Memory

For smooth spiking rate functions  $\varphi$ , by Proposition 2.2, the finite size system has only one invariant state corresponding to extinction of the system. In the large-population limit however, this situation changes for suitable choices of the form of the spiking rate function. More precisely, we suppose that

**Assumption 2.4**  $\varphi$  is non-decreasing and differentiable with  $\varphi'(0) > 0$ . Moreover, there exists a constant  $D > 0$  such that the equation  $D\varphi^2(x) = x$  possesses exactly three solutions  $x_0 = 0 < x_1 < x_2$  in  $[0, \infty[$ . ( $D$  is homogeneous to a potential times a time squared).

Let us write  $u^{max} = x_2, r^{max} = \frac{1}{\lambda}\varphi(u^{max})$ . We show in Proposition 6.2 below that  $(u^{max}, r^{max})$  is an attracting equilibrium of the limit system (2.2), for suitable choices of  $\alpha, \beta$  and  $\lambda$ .

Suppose now we observe a huge system of interacting neurons which is undergoing synaptic plasticity modulated by the residual calcium concentrations within each neuron. Hence,  $N$  is big without being infinite. We expose the system to some initial stimulus pushing it into the vicinity of the attracting non-trivial equilibrium point  $(u^{max}, r^{max})$  of the limit system. At time 0, this stimulus is switched off, and we start observing the system, evolving according to (2.1). Since this point is attracting and  $N$  large, the system is attracted to a small neighbourhood of  $(u^{max}, r^{max})$  and stays in this neighbourhood for a while. We interpret this transient behavior as an expression of *short term memory*. Of course, in the long run, the system will finally get kicked out of this neighbourhood and start rapidly decaying towards the all-zero state. These ideas are formalised in the following theorem.

**Theorem 2.5** Grant Assumptions 2.1 and 2.3 and suppose moreover that  $\int_0^\infty e^{ur} g_0(dr) < \infty$  for some  $u > 0$ . Fix some  $T > 1$ .

1. There exist positive constants  $C_T$  and  $c$  where  $C_T$  depends only on the parameters of the model and on  $T$  and  $c$  only on the parameters of the model with the following properties. For any  $\varepsilon > 0$ , there exists  $N_0$  such that for all  $N \geq N_0$ , for every  $1 \leq i \leq N$ ,

$$IP \left( \sup_{s \leq T} \left( |U_s^N(i) - U_s| + \left| \frac{1}{N} \sum_{j=1}^N R_s^N(j) - \mathbb{E} R_s \right| \right) \geq N^{-1/5} \varepsilon \right) \leq C_T e^{-c\varepsilon^2 \sqrt{N}}. \quad (2.3)$$

2. Grant moreover Assumption 2.4 and suppose that

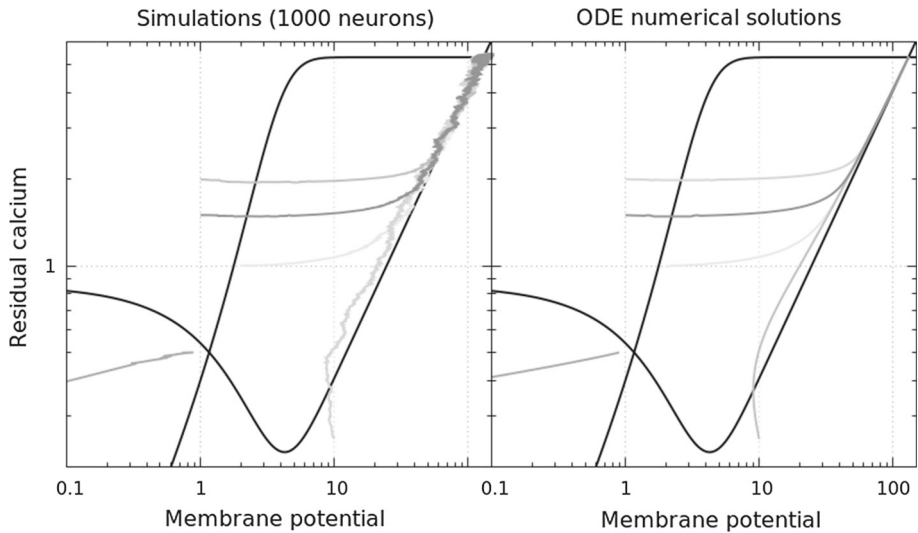
$$\alpha \geq D\beta\lambda. \quad (2.4)$$

Then the equation  $x = \frac{\alpha}{\beta\lambda}\varphi^2(x)$  possesses three solutions  $x_0 = 0 < x_1 < x_2$ . We put  $u^{max} = x_2, r^{max} = \frac{1}{\lambda}\varphi(u^{max})$ . The points  $(0, 0)$  and  $(u^{max}, r^{max})$  are locally attracting equilibria of the dynamical system

$$u_t = u_0 - \beta \int_0^t u_s ds + \alpha \int_0^t \varphi(u_s) r_s ds, \quad r_t = r_0 - \lambda \int_0^t r_s ds + \int_0^t \varphi(u_s) ds. \quad (2.5)$$

3. Suppose  $(u_0, r_0)$  belongs to the domain of attraction of  $(u^{max}, r^{max})$  and  $U_0^N(i) = u_0, R_0^N(i) = r_0$  for all  $i, N$ . Finally let, for  $\varepsilon > 0$ ,

$$t_1 = t_1(\varepsilon) = \inf\{t : |u_t - u^{max}| + |r_t - r^{max}| < \varepsilon\}.$$



**Fig. 1** Phase plots on a log–log scale. Left, 5 trajectories (gray lines) of the mean residual calcium versus the mean membrane potential obtained by simulating a network of 1000 neurons from 5 different initial states (see the main text for simulation details). Right, 5 trajectories solutions of the limit ODE system (2.5) with the same parameters and initial states (gray lines). On both plots the black curves show the null-cline of the mean membrane potential (V shaped) and of the mean residual calcium (inverted L shape)

Then for all  $N \geq N_0$ , for all  $1 \leq i \leq N$ ,

$$IP \left( |U_{t_1}^N(i) - u^{max}| \geq 2\varepsilon \text{ or } \left| \frac{1}{N} \sum_{j=1}^N R_{t_1}^N(j) - r^{max} \right| \geq 2\varepsilon \right) \leq C_{t_1} e^{-c\varepsilon^2 \sqrt{N}}.$$

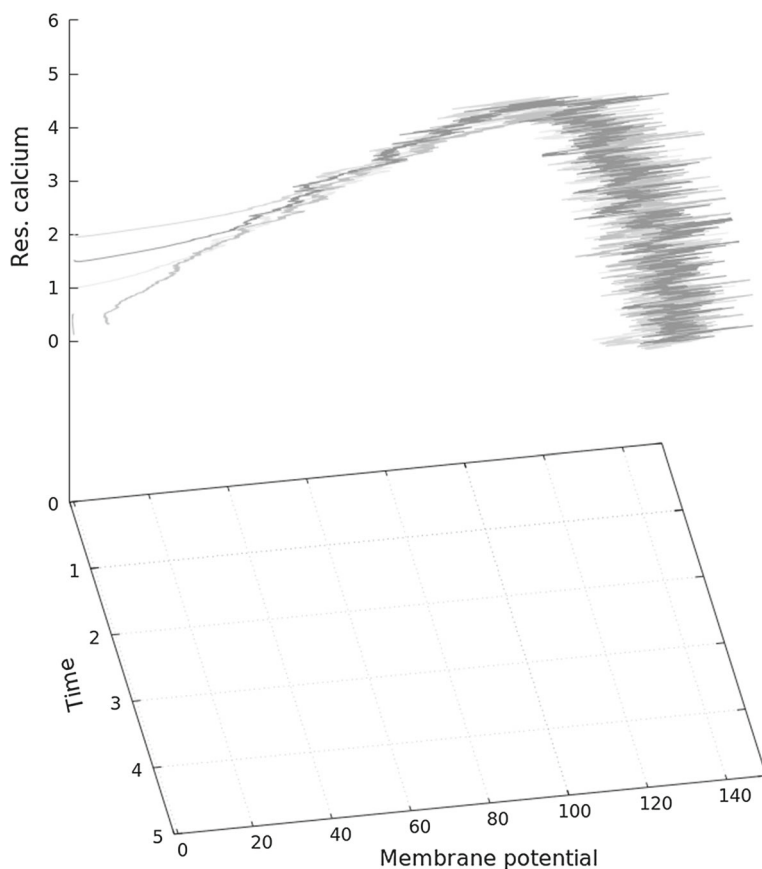
## 2.5 An Example with a Simulation Study

We consider spiking rate functions of sigmoid type which are defined in terms of a parameter  $a > 1$  satisfying  $4a < 1 + e^a$  by

$$\varphi(x) = \frac{4a}{1 + e^{-(x-a)}} - \frac{4a}{1 + e^a}, x \geq 0.$$

The point  $a$  is the inflexion point of  $\varphi$ , and it is easy to see that there exist  $x_1 \in ]0, a[$  and  $x_2 \in ]a, \infty[$  with  $\varphi^2(x_i) = x_i, i = 1, 2$ . Thus, Assumption 2.4 is satisfied with  $D = 1$ .

Figure 1 illustrates 5 trajectories of the mean residual calcium versus the mean membrane potential obtained by simulating a network of 1000 neurons from 5 different initial states on the left side. The null-clines corresponding to a null membrane potential derivative (V shaped) and a null residual calcium derivative (inverted L shaped) are also drawn. On the right side the numerical solution of the corresponding ODE system (2.5) with the same initial conditions as the right side are shown. A custom developed C code implementing Ogata's thinning method (see [15]) was used for the simulations with the Xoroshiro128+ pseudo-random number generator of Blackman and Vigna, see [2]. The ODE system (2.5) was numerically solved using the ode program of the open source GNU plotutils package (<https://www.gnu.org/software/plotutils/>). The default method—Runge–Kutta–Fehlberg



**Fig. 2** 3D plot of 5 trajectories of the mean residual calcium and the mean membrane potential obtained by simulating a network of 1000 neurons from 5 different initial states

with adaptive time steps—was used. The network parameters were:  $\alpha = 107.78$ ;  $\beta = 50$ ;  $\lambda = 2.16$  (rounded to the second decimal);  $a = 3$ . The parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  are chosen such that  $\alpha = D\beta\lambda$ . The specific choice of  $\alpha$ ,  $\beta$ ,  $\lambda$  was arbitrary and guided by aesthetic reasons. For the network simulations the initial states were obtained by drawing the membrane potential and the residual calcium of each neuron from a uniform distribution centred on a user set mean value with a range set to 10% of the mean. The (membrane potential, residual calcium) pairs were: (2, 1); (1, 2); (10, 0.25); (0.75, 0.5); (1, 1.5). The initial values of the ODE numerical solution were set to these mean values. The phase plots shown on Fig. 1 use a log–log scale. The trajectory starting from (0.75, 0.5) moves towards the origin: the network activity dies quickly in that case. All the other trajectories reach quickly (in less than 5 time units) the fixed point corresponding to the upper-right intersection of the 2 null-clines.

Figure 2 shows the simulated trajectories of same network in 3D using linear scales. All the codes and instructions required to reproduce these simulations and figures can be found at the following address: [https://plmlab.math.cnrs.fr/xtof/interacting\\_neurons\\_with\\_stp](https://plmlab.math.cnrs.fr/xtof/interacting_neurons_with_stp).

## 2.6 Constants

In the whole paper,  $C$  stands for a (large) finite constant and  $c$  stands for a (small) positive constant. Their values may change from line to line. They are allowed to depend only on  $\varphi, \alpha, \beta, \lambda$  and  $\eta_0$ , the law of the initial condition. Any other dependence will be indicated in subscript. For example,  $C_T$  is a finite constant depending only on  $\varphi, \alpha, \beta, \lambda, \eta_0$  and  $T$ . The letter  $K$  will be reserved to denote a bound on  $\|\varphi\|_\infty$ .

## 3 Well-Posedness of the Particle System and Proof of Theorem 2.2

**Proposition 3.1** *Under Assumption 2.1 a path-wise unique strong Markov process  $(U_t^N, R_t^N)$  exists which is solution of (2.1) for all  $t \geq 0$ .*

**Proof** This follows from Theorem 9.1 in Chapter IV of Ikeda and Watanabe (1989) [10].  $\square$

We now proceed with the

**Proof of Theorem 2.2** The proof works along the lines of the proof of Theorem 2.3 of Duarte and Ost [7].

Firstly, it is immediate to show that almost surely the process comes back to a compact set infinitely often. This follows from the boundedness of  $\varphi$  and the back driving force induced by the two drift terms  $-\beta U_t^N(i)dt$  and  $-\lambda R_t^N(i)dt$ .

Let then  $T_1$  be the first jump time of the system. Then, as in Proposition 3.1 of [7], for any  $c > 0$ ,

$$\inf_{u_0 \in [0, c]^N} \mathbb{P}(T_1 = \infty | U_0^N = u_0) \geq e^{-\frac{N}{\beta} \int_0^c \frac{\varphi(x)}{x} dx} > 0.$$

Since the process  $U_t^N$  almost surely comes back to a suitable compact set  $[0, c]^N$ , the assertion then follows by a Borel-Cantelli argument. The details are in [7].  $\square$

## 4 Well-Posedness of the Limit Equation

This section is devoted to the study of existence and uniqueness of the limit Eq. (2.2). A proof of the well-posedness of the limit system (2.2)

$$\begin{aligned} U_t &= U_0 - \beta \int_0^t U_s ds + \alpha \int_0^t \mathbb{E}[\varphi(U_s) R_s] ds, \\ R_t &= R_0 - \lambda \int_0^t R_s ds + \int_0^t \int_0^\infty 1_{\{z \leq \varphi(U_{s-})\}} \mathbf{M}(ds, dz), \end{aligned}$$

is not immediate due to the presence of the product term  $\mathbb{E}[\varphi(U_s) R_s] ds$  in the first line of the above system leading to non-Lipschitz terms. Our Assumption 2.3 has been introduced to cope with this problem. Indeed, observe that under Assumption 2.3, the limit process  $U_t$  is a deterministic process, and we shall write  $U_t = u_t$  to highlight this fact. Notice that in this case, the spike counting process of a typical neuron in the limit population

$$t \rightarrow \int_0^t \int_0^\infty 1_{\{z \leq \varphi(u_s)\}} \mathbf{M}(dz, ds)$$

is an inhomogeneous Poisson process of rate  $\varphi(u_t)$  at time  $t$ .



**Proposition 4.1** *Grant Assumption 2.3 and suppose only that  $\varphi$  is Lipschitz, satisfying  $\varphi(x) \leq C\sqrt{x}$  for all  $x \geq x_0$ , for some fixed  $x_0 > 0$ . Then a path-wise unique process  $(u_t, R_t)$  exists which is solution of (2.2).*

**Proof** Writing  $r_t = \mathbb{E} R_t$ , the system  $(u_t, r_t)$  solves, since  $U_0 = u_0$  is deterministic,

$$du_t = -\beta u_t dt + \alpha \varphi(u_t) r_t dt, dr_t = -\lambda r_t + \varphi(u_t) dt.$$

We first show that any solution is non-exploding. For that sake, suppose w.l.o.g. that  $u_0 > x_0$ . Since  $\varphi(x) \leq C\sqrt{x}$  for all  $x \geq x_0$ , we use the change of variables  $g_t := \sqrt{u_t}$  for all  $t \leq \tau := \inf\{s : u_s \leq x_0\}$ . Then, for all  $t \leq \tau$ ,

$$dg_t \leq -\frac{\beta}{2} g_t dt + C \frac{\alpha}{2} r_t dt, dr_t \leq -\lambda r_t + C g_t dt.$$

The rhs of the above inequality defines a linear equation which can be solved explicitly implying the existence of a non-exploding solution.

To prove uniqueness of the solution, suppose that  $(u'_t, R'_t)$  is another solution, starting from the same initial conditions. Due to the first part of the proof, we know that for any  $T > 0$ , the functions  $r_t = \mathbb{E} R_t$ ,  $r'_t = \mathbb{E} R'_t$ ,  $\varphi(u_t)$  and  $\varphi(u'_t)$  are bounded on  $[0, T]$ , say by a constant  $K_T$ . Then, by the Lipschitz continuity of  $\varphi$  with Lipschitz constant  $L_\varphi$ ,

$$|u_t - u'_t| \leq \beta \int_0^t |u_s - u'_s| ds + \alpha L_\varphi K_T \int_0^t |u_s - u'_s| ds + \alpha K_T \int_0^t |\mathbb{E} R_s - \mathbb{E} R'_s| ds$$

and

$$|\mathbb{E} R_t - \mathbb{E} R'_t| \leq \lambda \int_0^t |\mathbb{E} R_s - \mathbb{E} R'_s| ds + L_\varphi \int_0^t |u_s - u'_s| ds,$$

implying that

$$|\mathbb{E} R_t - \mathbb{E} R'_t| + |u_t - u'_t| \leq C_T \int_0^t [|\mathbb{E} R_s - \mathbb{E} R'_s| + |u_s - u'_s|] ds,$$

and thus  $R_t = R'_t$  almost surely and  $u_t = u'_t$  for all  $t \leq T$ , whence the uniqueness of the solution.  $\square$

## 5 Convergence of the Particle System to the Limit Equation

We now show that under Assumptions 2.1 and 2.3 the finite system (2.1) converges to the limit Eq. (2.2) in a certain sense. Recall that we suppose that  $(U_0^N(i), R_0^N(i))_{1 \leq i \leq N}$  are i.i.d.  $\eta_0$ -distributed random variables.

### 5.1 A Priori Bounds

In the sequel we shall use a priori upper bounds on the processes of residual calcium concentrations. Recall that in this part of the paper we work under the assumption that  $\varphi$  is bounded and that  $K = \|\varphi\|_\infty$ . We introduce

$$J_t(i) := \int_0^t \int_0^\infty 1_{\{z \leq K\}} \mathbf{M}^i(ds, dz), 1 \leq i \leq N.$$

Then  $J_t(i)$ ,  $i \geq 1$ , are i.i.d. standard Poisson processes of rate  $K$  each, and each  $R_t^N(i)$  is stochastically dominated by

$$R_t^N(i) \leq \check{R}_t^N(i) := e^{-\lambda t} R_0^N(i) + e^{-\lambda t} \int_0^t e^{\lambda s} dJ_s(i). \quad (5.1)$$

For each  $1 \leq i \leq N$ , the process  $\check{R}^N(i)$  is a Markov process with generator

$$\check{A}g(r) = -\lambda r g'(r) + K(g(r+1) - g(r)),$$

for sufficiently smooth test functions  $g$ . Moreover, the processes  $\check{R}^N(i)$ ,  $1 \leq i \leq N$ , are i.i.d. Taking a Lyapunov function  $V(x) = |x|$ ,  $V(x) = x^2$  or  $V(x) = e^{ux}$ , respectively, it is easy to see that for any such choice of  $V$  there exist suitable positive constants  $c, d, r_0$ , such that

$$\check{A}V(r) \leq c1_{\{r \leq r_0\}} - dV(r).$$

A standard Lyapunov argument then implies that for all  $t \geq 0$ ,

$$\mathbb{E}_r V(\check{R}_t^N(i)) \leq e^{-dt} r + \frac{c}{d},$$

and hence

$$\sup_{t \geq 0} \mathbb{E} V(R_t^N(i)) \leq \sup_{t \geq 0} \mathbb{E} V(\check{R}_t^N(i)) < \infty, \quad (5.2)$$

provided  $\mathbb{E}(V(R_0^N(i))) < \infty$ .

In particular, under our assumptions,

$$\sup_{t \geq 0} \mathbb{E}(R_t^N(i)) \leq \kappa_R := \mathbb{E}(R_0^N(1)) + \frac{K}{\lambda}, \quad (5.3)$$

implying that

$$\sup_{t \geq 0} \mathbb{E}(U_t^N(i)) \leq \kappa_U := \mathbb{E}(U_0^N(1)) + \frac{\alpha K}{\beta} \kappa_R. \quad (5.4)$$

## 5.2 Tightness in Skorokhod Space

We consider a probability distribution  $\eta_0$  on  $\mathbb{R}_+^2$  such that  $\int_{\mathbb{R}_+^2} (u+r)\eta_0(du, dr) < \infty$ , and for each  $N \geq 1$ , the unique solution  $(U_t^N, R_t^N)_{t \geq 0}$  to (2.1) starting from some i.i.d.  $\eta_0$ -distributed initial conditions  $(U_0^N(i), R_0^N(i))$ . We want to show that the sequence of processes  $(U_t^N(i), R_t^N(i))_{t \geq 0}$  is tight in  $D(\mathbb{R}_+, \mathbb{R}_+^2)$ , for any  $i \geq 1$ . Here, the set  $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+^2)$  of càdlàg functions on  $\mathbb{R}_+$  taking values in  $\mathbb{R}_+^2$  is endowed with the topology of the Skorokhod convergence on compact time intervals, see [11].

**Proposition 5.1** *Grant Assumption 2.1 and let  $(U_0^N(i), R_0^N(i))_{1 \leq i \leq N}$  be i.i.d. satisfying that  $\mathbb{E}(U_0^N(i)) < \infty$  and  $\mathbb{E}(R_0^N(i)) < \infty$ .*

- (i) *The sequence of processes  $(U_t^N(1), R_t^N(1))_{t \geq 0}$  is tight in  $D(\mathbb{R}_+, \mathbb{R}_+^2)$ .*
- (ii) *The sequence of empirical measures  $\hat{\mu}_N = N^{-1} \sum_{i=1}^N \delta_{(U_t^N(i), R_t^N(i))_{t \geq 0}}$  is tight in  $\mathcal{P}(\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+^2))$ .*

**Proof** Point (ii) follows from point (i) and the exchangeability of the system, see [19, Proposition 2.2-(ii)]. We thus only prove (i). To show that the family  $((U_t^N(1), R_t^N(1)))_{t \geq 0, N \geq 1}$  is tight in  $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+^2)$ , we use the criterion of Aldous, see [11, Theorem 4.5 page 356]. It is sufficient to prove that

(a) for all  $T > 0$ , all  $\varepsilon > 0$ ,  $\lim_{\delta \downarrow 0} \limsup_{N \rightarrow \infty} \sup_{(S, S') \in A_{\delta, T}} \mathbb{P}(|U_{S'}^N(1) - U_S^N(1)| + |R_{S'}^N(1) - R_S^N(1)| > \varepsilon) = 0$ , where  $A_{\delta, T}$  is the set of all pairs of stopping times  $(S, S')$  such that  $0 \leq S \leq S' \leq S + \delta \leq T$  a.s.,

(b) for all  $T > 0$ ,  $\lim_{L \uparrow \infty} \sup_N \mathbb{P}(\sup_{t \in [0, T]} (|U_t^N(1)| + |R_t^N(1)|) \geq L) = 0$ .

To check (a), consider  $(S, S') \in A_{\delta, T}$  and write

$$\begin{aligned} \mathbb{E}|U_{S'}^N(1) - U_S^N(1)| &\leq \beta \int_S^{S'} \mathbb{E}|U_s^N(1)| ds + \frac{\alpha}{N} \sum_{j=1}^N \int_S^{S'} \mathbb{E}[\varphi(U_s^N(j)) R_s^N(j)] ds \\ &\leq \beta \int_S^{S'} \mathbb{E}|U_s^N(1)| ds + \alpha K \kappa_R \delta \leq \beta \kappa_U \delta + \alpha K \kappa_R \delta \rightarrow 0 \end{aligned}$$

as  $\delta \rightarrow 0$ . The expression  $\mathbb{E}|R_{S'}^N(1) - R_S^N(1)|$  is treated analogously. Moreover, (b) immediately follows from the a priori bounds (5.3) and (5.4), since  $\kappa_R$  and  $\kappa_U$  do not depend on  $N$ .  $\square$

At this point, one usually concludes the proof that the sequence of processes  $(U_t^N(i), R_t^N(i))_{i \geq 0}$  converges weakly to  $(U_t, R_t)_{t \geq 0}$  in  $D(\mathbb{R}_+, \mathbb{R}_+^2)$  by showing that any possible limit point of the sequence is necessarily solution of the limit equation. Classically, this is shown by proving that any limit law must be solution of the associated martingale problem. Uniqueness of the solution of the martingale problem then implies the desired convergence.

In our specific situation however, we are able to identify any possible limit thanks to a coupling argument that we shall present in the next subsection. This coupling argument has another advantage. It enables us to give a precise rate of convergence.

### 5.3 A Coupling Approach

We propose a coupling approach, which is inspired by the ideas presented in Sznitman [19]. The non-Lipschitz term  $\mathbb{E}[\varphi(U_s) R_s] ds$  appearing in the limit system (2.2) demands however to adapt this approach to the present situation. Throughout this section we work under Assumption 2.3 implying that  $U_0^N(i) = u_0$  for all  $N$  and  $1 \leq i \leq N$ . As a consequence, coming back to (2.1), we see that

$$U_t^N(i) = U_t^N(1) =: U_t^N$$

for all  $1 \leq i \leq N$ , that is, the membrane potential processes of all neurons within the system are all equal, and only the values of the calcium concentrations differ. We can therefore rephrase (2.1) as

$$\begin{aligned} U_t^N &= u_0 - \beta \int_0^t U_s^N ds + \frac{\alpha}{N} \sum_{j=1}^N \int_0^t \int_0^\infty R_{s-}^N(j) 1_{\{z \leq \varphi(U_{s-}^N)\}} \mathbf{M}^j(ds, dz), \\ R_t^N(i) &= R_0^N(i) - \lambda \int_0^t R_s^N(i) ds + \int_0^t \int_0^\infty 1_{\{z \leq \varphi(U_{s-}^N)\}} \mathbf{M}^i(ds, dz), 1 \leq i \leq N. \end{aligned} \quad (5.5)$$

In order to control the speed of convergence to the limit system, we now first replace (5.5) by an approximating system which is given as follows.

$$\begin{aligned}\tilde{U}_t^N &= u_0 - \beta \int_0^t \tilde{U}_s^N ds + \alpha \int_0^t \varphi(\tilde{U}_s^N) \left( \frac{1}{N} \sum_{j=1}^N \tilde{R}_s^N(j) \right) ds, \\ \tilde{R}_t^N(i) &= R_0^N(i) - \lambda \int_0^t \tilde{R}_s^N(i) ds + \int_0^t \int_0^\infty 1_{\{z \leq \varphi(\tilde{U}_{s-}^N)\}} \mathbf{M}^i(ds, dz),\end{aligned}\quad (5.6)$$

where  $\mathbf{M}^i(ds, dz)$  is the Poisson random measure driving the dynamics of  $R^N(i)$ .

Our aim is to show that (5.6) is close to the original system (5.5). To do so, we introduce the distance

$$\Delta_t^N = \mathbb{E} \left( \left[ |\tilde{U}_t^N - U_t^N| + |\tilde{R}_t^N(1) - R_t^N(1)| \right] \right). \quad (5.7)$$

**Theorem 5.2** *Grant Assumptions 2.1 and 2.3. Fix  $T > 1$ . Then for all  $t \leq T$ ,*

$$\Delta_t^N \leq C_T N^{-1/2}. \quad (5.8)$$

**Proof** We work on the fixed time interval  $[0, T]$  and we suppose w.l.o.g. that  $KT \geq 1$ . Rewrite first the equation of  $U_t^N$  in the following way. For  $\tilde{\mathbf{M}}^j(ds, dz) := \mathbf{M}(ds, dz) - ds dz$ , we have

$$\begin{aligned}U_t^N &= u_0 - \beta \int_0^t U_s^N ds + \frac{\alpha}{N} \sum_{j=1}^N \int_0^t \int_0^\infty R_{s-}^N(j) 1_{\{z \leq \varphi(U_{s-}^N)\}} \tilde{\mathbf{M}}^j(ds, dz) \\ &\quad + \alpha \int_0^t \varphi(U_s^N) \left( \frac{1}{N} \sum_{j=1}^N R_s^N(j) \right) ds \\ &=: u_0 - \beta \int_0^t U_s^N ds + \alpha \int_0^t \varphi(U_s^N) \left( \frac{1}{N} \sum_{j=1}^N R_s^N(j) \right) ds + M_t^{N,1},\end{aligned}$$

where

$$M_t^{N,1} := \frac{\alpha}{N} \sum_{j=1}^N \int_0^t \int_0^\infty R_{s-}^N(j) 1_{\{z \leq \varphi(U_{s-}^N)\}} \tilde{\mathbf{M}}^j(ds, dz). \quad (5.9)$$

Therefore, writing

$$\bar{R}_t^N := \frac{1}{N} \sum_{j=1}^N \tilde{R}_t^N(j), \quad (5.10)$$

we have for a constant  $C$  that might change from one occurrence to another

$$\begin{aligned}|\tilde{U}_t^N - U_t^N| &\leq \beta \int_0^t |\tilde{U}_s^N - U_s^N| ds + \alpha \int_0^t \frac{1}{N} \sum_j |\varphi(U_s^N) R_s^N(j) \\ &\quad - \varphi(\tilde{U}_s^N) \tilde{R}_s^N(j)| ds + |M_t^{N,1}|\end{aligned}$$

$$\begin{aligned} &\leq \beta \int_0^t |\tilde{U}_s^N - U_s^N| ds + \alpha \int_0^t |\varphi(U_s^N) - \varphi(\tilde{U}_s^N)| \bar{R}_s^N ds \\ &\quad + C \int_0^t \frac{1}{N} \sum_j |R_s^N(j) - \tilde{R}_s^N(j)| ds + |M_t^{N,1}|, \end{aligned} \quad (5.11)$$

where we have used the boundedness of  $\varphi$ .

In the above expression, we have to control the term  $\alpha \int_0^t |\varphi(U_s^N) - \varphi(\tilde{U}_s^N)| \bar{R}_s^N ds$ . For that sake, fix a constant  $a \geq \sup_{s \geq 0} \mathbb{E} \check{R}_s^N(1) + 1$ . Notice that by (5.2), applied with  $V(x) = x^2$ , we can choose  $a$  such that it does not depend on  $T$ . Then, by the Lipschitz continuity of  $\varphi$ ,

$$\begin{aligned} &\alpha \int_0^t |\varphi(U_s^N) - \varphi(\tilde{U}_s^N)| \bar{R}_s^N ds \\ &\leq a\alpha L_\varphi \int_0^t |U_s^N - \tilde{U}_s^N| ds + \alpha \int_0^t |\varphi(U_s^N) - \varphi(\tilde{U}_s^N)| \bar{R}_s^N 1_{\{\bar{R}_s^N \geq a\}} ds. \end{aligned}$$

Since  $\varphi$  is bounded, the last expression in the above term is bounded by

$$\alpha \|\varphi\|_\infty \int_0^t \bar{R}_s^N 1_{\{\bar{R}_s^N \geq a\}} ds.$$

We use Hölder's inequality to obtain

$$\mathbb{E}(\bar{R}_s^N 1_{\{\bar{R}_s^N \geq a\}}) \leq \left( \mathbb{E}[(\bar{R}_s^N)^2] \right)^{1/2} \left( \mathbb{P}(\bar{R}_s^N \geq a) \right)^{1/2}.$$

(5.2) applied with  $V(x) = x^2$  and Chebyshev's inequality yield, by independence of the processes  $\check{R}^N(i)$ ,  $1 \leq i \leq N$ ,

$$\begin{aligned} \mathbb{P}(\bar{R}_s^N > a) &\leq \mathbb{P}\left(\sum_{i=1}^N \check{R}_s^N(i) > Na\right) \leq \mathbb{P}\left(\sum_{i=1}^N \check{R}_s^N(i) - \mathbb{E} \check{R}_s^N(i) > N\right) \\ &\leq \frac{1}{N} \text{Var}(\check{R}_s^N(1)) \leq \frac{C}{N}, \end{aligned}$$

since  $a \geq \mathbb{E} \check{R}_s^N(i) + 1$ . Therefore

$$\alpha \|\varphi\|_\infty \int_0^t \mathbb{E}[\bar{R}_s^N 1_{\{\bar{R}_s^N \geq a\}}] ds \leq CtN^{-1/2}.$$

We conclude that for all  $t \leq T$ ,

$$\begin{aligned} \mathbb{E}(|\tilde{U}_t^N - U_t^N|) &\leq \beta \int_0^t \mathbb{E}|\tilde{U}_s^N - U_s^N| ds + a\alpha L_\varphi \int_0^t \mathbb{E}|U_s^N - \tilde{U}_s^N| ds \\ &\quad + CtN^{-1/2} + C \int_0^t \mathbb{E}|R_s^N(1) - \tilde{R}_s^N(1)| ds + \mathbb{E}|M_t^{N,1}| \\ &\leq C \int_0^t \Delta_s^N ds + \frac{CT}{\sqrt{N}} + \mathbb{E}|M_t^{N,1}|, \end{aligned} \quad (5.12)$$

where we have used the exchangeability of the system.

To deal with the martingale part, write

$$M_t^{N,1} = \frac{\alpha}{N} \sum_{j=1}^N M_t^{N,1,j}, \quad \text{for } M_t^{N,1,j} = \int_0^t \int_0^\infty R_{s-}^N(j) 1_{\{z \leq \varphi(U_{s-}^N)\}} \tilde{\mathbf{M}}^j(ds, dz).$$

Notice that  $M_t^{N,1,j}$  and  $M_t^{N,1,k}$  almost surely never jump together for  $j \neq k$ , hence, using once more (5.2) with  $V(x) = x^2$ ,

$$\mathbb{E}(M_t^{N,1})^2 = \frac{\alpha^2}{N^2} \sum_{j=1}^N \mathbb{E}(M_t^{N,1,j})^2 \leq CN^{-1} \int_0^t \mathbb{E} \left( (R_s^N(1))^2 \varphi(U_s^N) \right) ds \leq C \frac{t}{N}.$$

As a consequence, for all  $t \leq T$ ,

$$\mathbb{E}|M_t^{N,1}| \leq CT^{1/2}N^{-1/2}.$$

Finally, using once again the Lipschitz continuity of  $\varphi$ ,

$$\mathbb{E} \left( |\tilde{R}_t^N(1) - R_t^N(1)| \right) \leq \lambda \int_0^t \mathbb{E} |\tilde{R}_s^N(1) - R_s^N(1)| ds + C \int_0^t \mathbb{E} |\tilde{U}_s^N - U_s^N| ds,$$

whence

$$\Delta_t^N \leq C(\alpha, K, g_0)TN^{-1/2} + C \int_0^t \Delta_s^{N,a} ds,$$

implying that

$$\Delta_t^N \leq CTN^{-1/2}e^{Ct}$$

for all  $t \leq T$ , which concludes the proof.  $\square$

We are now going to control the distance between the approximating system (5.6) and the limit system (2.2). Recall that

$$\bar{R}_t^N := \frac{1}{N} \sum_{i=1}^N \tilde{R}_t^N(i).$$

Then

$$d\bar{R}_t^N = -\beta \bar{R}_t^N dt + \int_0^\infty 1_{\{z \leq \varphi(\tilde{U}_{t-}^N)\}} \bar{\mathbf{M}}^N(dt, dz), \quad \text{where } \bar{\mathbf{M}}^N = \frac{1}{N} \sum_{i=1}^N \mathbf{M}^i.$$

Compensating each Poisson random measure, this yields

$$d\bar{R}_t^N = -\beta \bar{R}_t^N dt + \varphi(\tilde{U}_t^N)dt + dM_t^{N,2},$$

where

$$M_t^{N,2} = \frac{1}{N} \sum_{i=1}^N \int_0^t \int_0^\infty 1_{\{z \leq \varphi(\tilde{U}_{s-}^N)\}} \tilde{\mathbf{M}}^i(ds, dz) \quad (5.13)$$

is a square integrable martingale. We can thus rewrite (5.6) as

$$\tilde{U}_t^N = u_0 - \beta \int_0^t \tilde{U}_s^N ds + \alpha \int_0^t \varphi(\tilde{U}_s^N) \bar{R}_s^N ds, \quad (5.14)$$

$$\bar{R}_t^N = \bar{R}_0^N - \lambda \int_0^t \bar{R}_s^N ds + \int_0^t \varphi(\tilde{U}_s^N) ds + M_t^{N,2}. \quad (5.15)$$

Moreover, since  $U_0 = u_0$  is deterministic, writing  $r_t := \mathbb{E} R_t$  and recalling that  $u_t$  is deterministic, the dynamics of the limit equation boils down to

$$\begin{cases} du_t = -\beta u_t dt + \alpha \varphi(u_t) r_t dt \\ dr_t = -\lambda r_t dt + \varphi(u_t) dt \end{cases}. \quad (5.16)$$

We obtain

**Theorem 5.3** *Grant the assumptions of Theorem 5.2. Fix  $T > 1$ . Then for all  $t \leq T$*

$$\mathbb{E} \left[ |u_t - \tilde{U}_t^N| + |r_t - \bar{R}_t^N| \right] \leq C_T N^{-1/2}$$

and consequently for the original particle system,

$$\mathbb{E}[|u_t - U_t^N|] + \mathbb{E} \left[ \left| \frac{1}{N} \sum_{j=1}^N R_t^N(j) - r_t \right| \right] \leq C_T N^{-1/2}.$$

**Proof** (5.16) together with (5.14) implies

$$\begin{aligned} |u_t - \tilde{U}_t^N| &\leq \beta \int_0^t |u_s - \tilde{U}_s^N| ds + \alpha \int_0^t |\varphi(\tilde{U}_s^N) \bar{R}_s^N - \varphi(u_s) r_s| ds \\ &\leq \beta \int_0^t |u_s - \tilde{U}_s^N| ds + C \int_0^t |\tilde{U}_s^N - u_s| r_s ds + \alpha K \int_0^t |r_s - \bar{R}_s^N| ds. \end{aligned} \quad (5.17)$$

We pass to expectation and use the fact that  $r_s$  is deterministic and bounded by  $r_s \leq r_0 + \frac{K}{\lambda}$ , to deduce from this that

$$\mathbb{E}|u_t - \tilde{U}_t^N| \leq C \int_0^t \left( \mathbb{E}|u_s - \tilde{U}_s^N| + \mathbb{E}|r_s - \bar{R}_s^N| \right) ds.$$

Moreover,

$$|r_t - \bar{R}_t^N| \leq |r_0 - \bar{R}_0^N| + \lambda \int_0^t |r_s - \bar{R}_s^N| ds + L_\varphi \int_0^t |\tilde{U}_s^N - u_s| ds + |M_t^{N,2}|. \quad (5.18)$$

We use that  $M_t^{N,2}$  has quadratic variation

$$\langle M^{N,2} \rangle_t = \frac{1}{N} \int_0^t \varphi(\tilde{U}_s^N) ds \leq \frac{Kt}{N}$$

to obtain

$$\mathbb{E}|r_t - \bar{R}_t^N| \leq \mathbb{E}|r_0 - \bar{R}_0^N| + \lambda \int_0^t \mathbb{E}|r_s - \bar{R}_s^N| ds + C \int_0^t \mathbb{E}|\tilde{U}_s^N - u_s| ds + C_T \frac{1}{\sqrt{N}}, \quad (5.19)$$

implying the assertion, since  $\mathbb{E}|r_0 - \bar{R}_0^N| \leq C N^{-1/2}$ .  $\square$

Theorems 5.2 and 5.3 enable us finally to conclude that

**Theorem 5.4** *Grant the assumptions of Theorem 5.2. Then we have that for all  $i \geq 1$ , the sequence of processes  $(U_t^N(i), R_t^N(i))_{t \geq 0}$  converges weakly to  $(u_t, R_t)_{t \geq 0}$  in  $D(\mathbb{R}_+, \mathbb{R}_+^2)$ .*

**Proof** Take any weakly convergent subsequence of  $(U_t^N(i), R_t^N(i))$  and call  $Z = (V, S)$  its weak limit. We suppose that  $(V, S)$  is defined on a filtered probability space  $(\Omega', \mathcal{A}', (\mathcal{F}_t')_{t \geq 0}, P')$ , where

$$\mathcal{F}_t' = \bigcap_{s > t} \mathcal{F}_s^0, \mathcal{F}_s^0 = \sigma(V_t, S_t, t \leq s).$$

Theorems 5.2 and 5.3 imply that  $V = u$  almost surely (since  $u$  is deterministic). Moreover it is straightforward to show that the limit  $(u, S)$  must be solution of the following martingale problem. For any smooth and bounded test function  $\psi$ , any  $s_1 < s_2 < \dots < s_k \leq s < t$ , for continuous and bounded test functions  $\psi_i$ , we have

$$\mathbb{E} \left[ \left( \psi(S_t) - \psi(S_s) - \lambda \int_s^t \psi'(S_v) S_v dv - \int_s^t \varphi(u_v) [\psi(S_v + 1) - \psi(S_v)] dv \right) \prod_{i=1}^k \psi_i(Z_{s_i}) \right] = 0. \quad (5.20)$$

By [11, Theorem II.2.42, page 86], and using the right-continuity of  $S$ , this implies that  $S$  is a  $(P', (\mathcal{F}_t')_{t \geq 0})$ -semi-martingale with characteristics  $B = -\lambda \int_0^\cdot S_s ds$ ,  $\nu(ds, dx) = \varphi(u_s) ds \delta_1(dx)$ ,  $C_t = 0$ . Moreover, [11, Theorem III.2.26, page 157] implies that there exists a Poisson random measure  $\pi$  defined on  $(\Omega', \mathcal{A}', (\mathcal{F}_t')_{t \geq 0}, P')$ , such that  $S$  is solution of

$$S_t = S_0 - \lambda \int_0^t S_v dv + \int_0^t \int_0^\infty 1_{\{z \leq \varphi(u_s)\}} \pi(ds, dz),$$

where  $S_0$  is  $g_0$ -distributed. In other words,  $S \stackrel{\mathcal{L}}{=} R$ . Hence any weak limit has the same law, implying the weak convergence of  $(U_t^N(i), R_t^N(i))$ .  $\square$

## 6 Stationary Solutions of the Limit Equation

For smooth spiking rate functions  $\varphi$ , by Theorem 2.2, the finite size system has only one invariant state corresponding to extinction of the system. The limit system can however display several invariant states as we show now, including persistent behavior where the spiking activity of the system survives forever.

Recall that passing to expectation and writing  $u_t = \mathbb{E}U_t$ ,  $r_t = \mathbb{E}R_t$ , we have reduced the limit system to

$$\begin{cases} du_t = -\beta u_t dt + \alpha \varphi(u_t) r_t dt \\ dr_t = -\lambda r_t dt + \varphi(u_t) dt \end{cases}. \quad (6.1)$$

Any stationary solution  $(u^*, r^*)$  of (6.1) must satisfy

$$\lambda r^* = \varphi(u^*) \text{ and } \beta u^* = \alpha \varphi(u^*) r^* = \frac{\alpha}{\lambda} \varphi^2(u^*), \quad (6.2)$$

implying that

$$u^* = \frac{\alpha}{\beta \lambda} \varphi^2(u^*). \quad (6.3)$$

Of course,  $(0, 0)$  is always a stationary solution since  $\varphi(0) = 0$ . However, for suitable choices of

$$\kappa := \frac{\alpha}{\beta \lambda}$$

( $\kappa$  is homogeneous to a potential times a time squared) and of the form of  $\varphi$ , also other stationary solutions appear, some of them being attracting. Let us come back to the example already presented in Sect. 2.



**Example 6.1** We consider spiking rate functions of sigmoid type. They are defined in terms of a parameter  $a > 1$  satisfying that

$$\frac{4a}{1+e^a} < 1.$$

Let

$$\varphi(x) = \frac{4a}{1+e^{-(x-a)}} - \frac{4a}{1+e^a}, x \geq 0.$$

Then  $\varphi(0) = 0$ ,  $0 < \varphi'(0) < 1$ ,  $\varphi'' > 0$  on  $[0, a[$  and  $\varphi'' < 0$  on  $]a, \infty[$ , thus  $a$  is the inflexion point of  $\varphi$ . We have that

$$\varphi(a) = 4a \left( \frac{1}{2} - \frac{1}{1+e^a} \right) \geq 4a \left( \frac{1}{2} - \frac{1}{4} \right) = a,$$

since  $\frac{4a}{1+e^a} < 1$  and  $a > 1$  imply that  $\frac{1}{1+e^a} < \frac{1}{4}$ . Notice that  $\varphi^2(0) = 0$ ,  $(\varphi^2)'(0) = 0$  and  $\varphi^2(a) = a^2 > a$  (recall that  $a > 1$ ). This implies that for some  $x_1 \in ]0, a[$  we have  $\varphi^2(x_1) = x_1$ . Finally, boundedness of  $\varphi^2$  implies the existence of a second point  $x_2 > a$  with  $\varphi^2(x_2) = x_2$ .

Our Assumption 2.4 implies the existence of at least two non trivial solutions of  $\kappa\varphi^2(x) = x$  in  $]0, \infty[$ , for all  $\kappa \geq D$ . This means that two non-trivial stationary solutions for (6.1) exist. Let  $u^{max} > 0$  be the maximal solution of  $\kappa\varphi^2(x) = x$  in  $]0, \infty[$  and write  $r^{max} := \frac{1}{\lambda}\varphi(u^{max})$ . The point  $(u^{max}, r^{max})$  is locally attracting as shows the following proposition.

**Proposition 6.2** Grant Assumption 2.4 and suppose that  $\kappa \geq D$ . Then  $(0, 0)$  and  $(u^{max}, r^{max})$  are locally attracting equilibria of (6.1).

**Proof** Write  $\Phi(u) = u/\varphi(u)$ . Then  $\{r = \frac{1}{\lambda}\varphi(u)\}$  is the null-cline of  $r_t$  and  $\{r = \frac{\beta}{\alpha}\Phi(u), u > 0\} \cup \{r = u = 0\}$  the null-cline of  $u_t$ . We suppose w.l.o.g. that  $\kappa\varphi^2(x) = x$  possesses exactly two non trivial solutions  $0 < x_1 < x_2 = u^{max}$  in  $]0, \infty[$ . Then we have five regions

$$R_1 = \left\{ (u, r) : u \leq x_1, \frac{1}{\lambda}\varphi(u) \leq r \leq \frac{\beta}{\alpha}\Phi(u) \right\}, R_2 = \left\{ (u, r) : r \leq \frac{1}{\lambda}\varphi(u) \wedge \frac{\beta}{\alpha}\Phi(u) \right\},$$

$$R_3 = \left\{ (u, r) : u \geq x_2, \frac{1}{\lambda}\varphi(u) \leq r \leq \frac{\beta}{\alpha}\Phi(u) \right\}, R_4 = \left\{ (u, r) : r \geq \frac{1}{\lambda}\varphi(u) \vee \frac{\beta}{\alpha}\Phi(u) \right\}$$

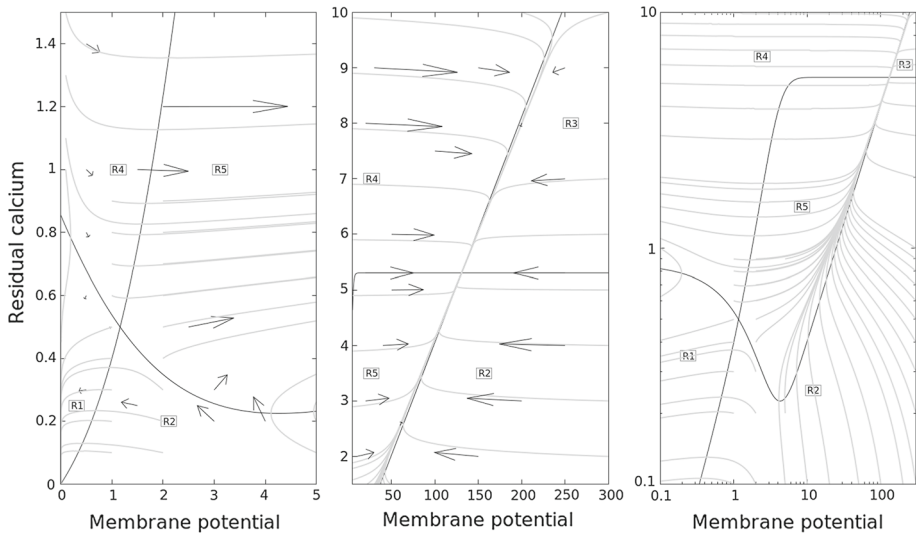
and finally

$$R_5 = \left\{ (u, r) : x_1 \leq u \leq x_2, \frac{\beta}{\alpha}\Phi(u) \leq r \leq \frac{1}{\lambda}\varphi(u) \right\},$$

see Fig. 3. On  $R_1 \cup R_3$ , both  $u_t$  and  $r_t$  decrease. On  $R_2$ , only  $r_t$  increases while  $u_t$  decreases. Thus, for sufficiently small  $u_0$ , trajectories starting within  $R_2$  will enter  $R_1$  after some time and then be attracted to  $(0, 0)$ . The other trajectories starting within  $R_2$  will either enter  $R_5$  or  $R_3$ . In both cases, they finish being attracted to  $(u^{max}, r^{max})$ . The same argument applies to trajectories starting in  $R_4$ , where  $u_t$  increases, while  $r_t$  decreases.  $\square$

Let us finally shortly discuss the role of  $\kappa$  on the number of stationary limit states.

**Proposition 6.3** Grant Assumption 2.4. There exists  $\kappa_c \leq D$  such that for all  $\kappa < \kappa_c$ , no non-trivial solution to (6.3) exist, and for all  $\kappa > \kappa_c$ , at least two non-trivial solutions exist.



**Fig. 3** Phase portrait of the dynamical system given in (6.1) with  $\varphi$  as in Example 6.1 and  $\alpha = 107.78$ ;  $\beta = 50$ ;  $\lambda = 2.16$ ;  $a = 3$ . Left: Plot showing the unstable equilibrium  $(x_1, r_1)$ ,  $r_1 = \frac{1}{\lambda}\varphi(x_1)$ . The black curve starting from  $(0, 0)$  is the null-cline of  $r$ , the other black curve the null-cline of  $u$ . Trajectories starting in  $R_1$  are attracted to  $(0, 0)$ . Middle: Plot showing the stable equilibrium  $(x_2, r_2)$ , the upper black curve is the null-cline of  $r$ . The arrows represent the tangent vectors of the flow, their length correspond to  $1/100$  of their true length. Right: Plot on a log–log scale showing both equilibria

## 7 Deviation Inequalities and Proof of Theorem 2.5

We are now able to conclude the paper with the proof of Theorem 2.5. In what follows we shall use the martingales  $M^{N,1}$  and  $M^{N,2}$  which have been defined in (5.9) and (5.13). We also introduce

$$M_t^{N,3} = \frac{1}{N} \sum_{i=1}^N \int_0^t \int_{\mathbb{R}_+} |1_{\{z \leq \varphi(U_{s-}^N)\}} - 1_{\{z \leq \varphi(\tilde{U}_{s-}^N)\}}| \tilde{\mathbf{M}}^i(ds, dz).$$

Finally we use the definition of  $\tilde{R}_t^N$ , given in (5.10) above.

The first item of Theorem 2.5 will then be a consequence of

**Theorem 7.1** Fix  $T > 1$  and grant the assumptions of Theorem 5.2. Suppose moreover that  $\int_0^\infty e^{ur} g_0(dr) < \infty$  for some  $u > 0$ . Then for any  $R > 0$ , there exists a constant  $C_R$  depending only on the parameters of the model and on  $R$ , such that

$$\begin{aligned} & \sup_{s \leq T} \left( |U_s^N - u_s| + \left| \frac{1}{N} \sum_{i=1}^N R_s^N(i) - r_s \right| \right) \\ & \leq (C_R)^{T+1} \left( \int_0^T 1_{\{\tilde{R}_s^N > R\}} \tilde{R}_s^N ds + |\tilde{R}_0^N - r_0| + \sup_{s \leq T} |M_s^{N,1}| + \sup_{s \leq T} |M_s^{N,2}| + \sup_{s \leq T} |M_s^{N,3}| \right). \end{aligned} \quad (7.1)$$

Consequently for all  $\varepsilon > 0$  there exists  $N_0 = N_0(T, \varepsilon)$ , such that for all  $N \geq N_0$ , we obtain the following deviation bound

$$IP \left( \sup_{s \leq T} \left( |U_s^N - u_s| + \left| \frac{1}{N} \sum_{i=1}^N R_s^N(i) - r_s \right| \right) \geq N^{-1/5} \varepsilon \right) \leq C_T e^{-c\varepsilon^2 \sqrt{N}}. \quad (7.2)$$

**Proof** In the sequel we shall suppose w.l.o.g. that  $KT \geq 1$ .

*Step 1* Fix some  $0 < t_0 < T$  and let  $t > 0$  such that  $t_0 + t \leq T$ . Then Eq. (5.11) implies, for any  $R > 0$ ,

$$\begin{aligned} \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N| &\leq |U_{t_0}^N - \tilde{U}_{t_0}^N| + t \cdot (\beta + \alpha L_\varphi R) \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N| \\ &+ \alpha K \int_0^T 1_{\{\tilde{R}_s^N > R\}} \tilde{R}_s^N ds + Ct \frac{1}{N} \sum_{i=1}^N \sup_{t_0 \leq s \leq t_0+t} |R_s^N(i) - \tilde{R}_s^N(i)| + \sup_{t_0 \leq s \leq t_0+t} |M_s^{N,1}|. \end{aligned} \quad (7.3)$$

Applying the same argument to  $R^N$  and  $\tilde{R}^N$ , we also have that

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sup_{t_0 \leq s \leq t_0+t} |R_s^N(i) - \tilde{R}_s^N(i)| &\leq \frac{1}{N} \sum_{i=1}^N |R_{t_0}^N(i) - \tilde{R}_{t_0}^N(i)| \\ &+ \lambda t \left( \frac{1}{N} \sum_{i=1}^N \sup_{t_0 \leq s \leq t_0+t} |R_s^N(i) - \tilde{R}_s^N(i)| \right) \\ &+ \sup_{t_0 \leq s \leq t_0+t} |M_s^{N,3}| + L_\varphi t \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N|. \end{aligned}$$

Choosing  $s_1$  sufficiently small such that  $\lambda s_1 \leq \frac{1}{2}$  and also  $2L_\varphi s_1 \leq 1$ , we deduce from this that for all  $t \leq s_1$ ,

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sup_{t_0 \leq s \leq t_0+t} |R_s^N(i) - \tilde{R}_s^N(i)| &\leq \frac{2}{N} \sum_{i=1}^N |R_{t_0}^N(i) - \tilde{R}_{t_0}^N(i)| + 2 \sup_{t_0 \leq s \leq t_0+t} |M_s^{N,3}| \\ &+ \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N|, \end{aligned} \quad (7.4)$$

and replacing in (7.3), we obtain

$$\begin{aligned} \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N| &\leq |U_{t_0}^N - \tilde{U}_{t_0}^N| + \frac{2Ct}{N} \sum_{i=1}^N |R_{t_0}^N(i) - \tilde{R}_{t_0}^N(i)| \\ &+ t \cdot (\beta + \alpha L_\varphi R + C) \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N| \\ &+ \alpha K \int_0^T 1_{\{\tilde{R}_s^N > R\}} \tilde{R}_s^N ds + 2Ct \sup_{s \leq t} |M_s^{N,3}| + \sup_{s \leq t} |M_s^{N,1}|. \end{aligned}$$

Consequently we may choose  $s_2 \leq s_1$  such that

$$s_2 \cdot (\beta + \alpha L_\varphi R + C) \leq \frac{1}{2} \text{ and also } 2Cs_2 \leq 1.$$

Therefore, for all  $t \leq s_2$ , for a suitable constant  $C$ ,

$$\begin{aligned} \sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N| &\leq 2|U_{t_0}^N - \tilde{U}_{t_0}^N| + \frac{2}{N} \sum_{i=1}^N |R_{t_0}^N(i) - \tilde{R}_{t_0}^N(i)| \\ &\quad + C \left( \int_0^T 1_{\{\bar{R}_s^N > R\}} \bar{R}_s^N ds + \sup_{t_0 \leq s \leq t_0+t} |M_s^{N,3}| + \sup_{t_0 \leq s \leq t_0+t} |M_s^{N,1}| \right), \end{aligned}$$

and replacing this once more in (7.4), we finally obtain for all  $t_0, t$  such that  $t_0 + t \leq T$ ,

$$\begin{aligned} &\sup_{t_0 \leq s \leq t_0+t} |U_s^N - \tilde{U}_s^N| + \frac{1}{N} \sum_{i=1}^N \sup_{t_0 \leq s \leq t_0+t} |R_s^N(i) - \tilde{R}_s^N(i)| \\ &\leq C|U_{t_0}^N - \tilde{U}_{t_0}^N| + \frac{C}{N} \sum_{i=1}^N |R_{t_0}^N(i) - \tilde{R}_{t_0}^N(i)| \\ &\quad + C \left( \int_0^T 1_{\{\bar{R}_s^N > R\}} \bar{R}_s^N ds + \sup_{s \leq T} |M_s^{N,3}| + \sup_{s \leq T} |M_s^{N,1}| \right) \end{aligned} \quad (7.5)$$

for all  $t \leq s_2$ .

*Step 2* We apply the above inequality (7.5) with  $t_0 = 0, t_0 = s_2, \dots, t_0 = \ell s_2$ , for  $\ell = \lceil T/s_2 \rceil$ , where  $\lfloor x \rfloor$  denotes the integer part of  $x > 0$ , and with  $t = s_2$  (except in the last case  $t_0 = \ell s_2$  where we take  $t = T - \ell s_2$ ). Write for short

$$Y_T := C \left( \int_0^T 1_{\{\bar{R}_s^N > R\}} \bar{R}_s^N ds + \sup_{s \leq T} |M_s^{N,3}| + \sup_{s \leq T} |M_s^{N,1}| \right)$$

and

$$\Delta(k) := \sup_{ks_2 \leq s \leq (k+1)s_2 \wedge T} |U_s^N - \tilde{U}_s^N| + \frac{1}{N} \sum_{i=1}^N \sup_{ks_2 \leq s \leq (k+1)s_2 \wedge T} |R_s^N(i) - \tilde{R}_s^N(i)|, 0 \leq k \leq \ell.$$

(7.5) implies that

$$\Delta(k+1) \leq C\Delta(k) + Y_T, \text{ for all } k < \ell$$

(choose  $t_0 = (k+1)s_2$  in (7.5)). Since  $U_0^N = \tilde{U}_0^N = u_0$  and  $R_0^N(i) = \tilde{R}_0^N(i)$ , for all  $1 \leq i \leq N$ , we have moreover that  $\Delta(0) \leq Y_T$ . Therefore,

$$\Delta(\ell) \leq (1 + C + \dots C^\ell) Y_T \leq C^{\ell+1} Y_T$$

(we suppose w.l.o.g. that  $C \geq 1$ ). Finally,  $\ell \leq T/s_2$  and  $s_2 = s_2(R)$  not depending on  $T$  imply that, for  $C_R = C^{\frac{1}{s_2(R)}}$  depending only on the parameters of the model and on  $R$ , but not on  $T$ ,

$$\begin{aligned} &\sup_{s \leq T} |U_s^N - \tilde{U}_s^N| + \frac{1}{N} \sum_{i=1}^N \sup_{s \leq T} |R_s^N(i) - \tilde{R}_s^N(i)| \leq \\ &\leq (C_R)^{T+1} \left( \int_0^T 1_{\{\bar{R}_s^N > R\}} \bar{R}_s^N ds + \sup_{s \leq T} |M_s^{N,3}| + \sup_{s \leq T} |M_s^{N,1}| \right). \end{aligned} \quad (7.6)$$

This is the first part of (7.1).

*Step 3* We now turn to the second step in our approximation procedure. It is treated analogously to the two previous steps. We have by (5.17), for all  $t_0, t$  such that  $t_0 + t \leq T$ ,

$$\begin{aligned} \sup_{t_0 \leq s \leq t_0+t} |\tilde{U}_s^N - u_s| &\leq |\tilde{U}_{t_0}^N - u_{t_0}| + t \left( \beta + C \left( r_0 + \frac{K}{\lambda} \right) \right) \sup_{t_0 \leq s \leq t_0+t} |\tilde{U}_s^N - u_s| \\ &\quad + \alpha K t \sup_{t_0 \leq s \leq t_0+t} |r_s - \bar{R}_s^N|, \end{aligned}$$

where we have used that for all  $s \leq T$ ,  $r_s \leq r_0 + \frac{K}{\lambda}$ .

Analogously, (5.18) implies that

$$\begin{aligned} \sup_{t_0 \leq s \leq t_0+t} |\bar{R}_s^N - r_s| &\leq |\bar{R}_{t_0}^N - r_{t_0}| + \lambda t \sup_{t_0 \leq s \leq t_0+t} |r_s - \bar{R}_s^N| \\ &\quad + L_\varphi t \sup_{t_0 \leq s \leq t_0+t} |\tilde{U}_s^N - u_s| + \sup_{s \leq T} |M_s^{N,2}|. \end{aligned}$$

Putting the two together and using similar arguments as in the first step, for a choice of  $t_1 \leq s_2$  not depending on  $T$  and sufficiently small,

$$\sup_{t_0 \leq s \leq t_0+t} |\tilde{U}_s^N - u_s| + \sup_{t_0 \leq s \leq t_0+t} |\bar{R}_s^N - r_s| \leq C |\tilde{U}_{t_0}^N - u_{t_0}| + C |\bar{R}_{t_0}^N - r_{t_0}| + C \sup_{s \leq T} |M_s^{N,2}|. \quad (7.7)$$

Iterating the above inequality for  $t_0 = 0, \dots, t_0 = kt_1$ , with  $k = [T/t_1]$  as in Step 2. above implies then

$$\sup_{s \leq T} |\tilde{U}_s^N - u_s| + \sup_{s \leq T} |\bar{R}_s^N - r_s| \leq C^{T+1} |\bar{R}_0^N - r_0| + C^{T+1} \sup_{s \leq T} |M_s^{N,2}|. \quad (7.8)$$

Finally, (7.8) together with (7.6) imply (7.1).

*Step 4* We use a large deviations upper bound to obtain

$$\mathbb{P}(|\bar{R}_0^N - r_0| \geq \varepsilon N^{-1/4}) \leq e^{-c\sqrt{N}},$$

for  $\varepsilon > 0$ , due to our assumptions on the law  $g_0$  of  $R_0^N(1)$ .

Moreover, we use similar arguments as those in the proof of Theorem 5.2 to deduce that for any  $\eta > 0$ ,

$$\mathbb{P} \left( \int_0^T 1_{\{\bar{R}_s^N > R\}} \bar{R}_s^N ds \geq \eta \right) \leq \frac{C}{\eta} \int_0^T \sqrt{\mathbb{P}(\bar{R}_s^N > R)} ds.$$

We wish to use again a large deviations upper bound to control  $\mathbb{P}(\bar{R}_s^N > R)$ . Since  $\int_0^\infty e^{ur} g_0(dr) < \infty$ , applying (5.2) with  $V(x) = e^{ux}$  gives

$$\sup_{t \geq 0} \mathbb{E} e^{u \check{R}_t^N(i)} < \infty. \quad (7.9)$$

Therefore we may choose  $R$  such that  $\log \left( \sup_{t \geq 0} \mathbb{E} \left( e^{u \check{R}_t^N(i)} \right) \right) - uR := -\chi < 0$ . Then

$$\mathbb{P}(\bar{R}_s^N > R) \leq \mathbb{P} \left( \sum_{i=1}^N \check{R}_s^N(i) > NR \right) \leq e^{-NuR} \left( \mathbb{E} \left( e^{u \check{R}_s^N(i)} \right) \right)^N \leq e^{-\chi N},$$

implying that

$$\mathbb{P} \left( \int_0^T 1_{\{\bar{R}_s^N > R\}} \bar{R}_s^N ds \geq \eta \right) \leq \frac{C}{\eta} T e^{-\frac{1}{2} \chi N}.$$

To deal with the square integrable martingale terms  $M^{N,i}$ ,  $1 \leq i \leq 3$ , we rely on the Bernstein inequality of [8, Theorem 3.3]). For any square integrable purely discontinuous martingale  $M$ , writing

$$[M]_t = \sum_{s \leq t} (\Delta M_s)^2, \quad \text{and} \quad \langle M \rangle_t \quad \text{for its predictable compensator,}$$

and putting, for a fixed  $a$ ,

$$H_t^a := \sum_{s \leq t} (\Delta M_s)^2 1_{\{|\Delta M_s| > a\}} + \langle M \rangle_t,$$

we have that

$$IP(M_t^* \geq z, H_t^a \leq L) \leq 2 \exp \left( -\frac{1}{2} \frac{z^2}{L} \psi \left( \frac{az}{L} \right) \right), \quad (7.10)$$

where  $\psi(x) = (1 + x/3)^{-1}$  and  $M_t^* = \sup_{s \leq t} |M_s|$ .

Observing that  $M^{N,2}$  and  $M^{N,3}$  have jumps bounded by  $\frac{1}{N}$  and quadratic variation bounded by

$$\langle M^{N,i} \rangle_T \leq \frac{KT}{N}, \quad i = 2, 3,$$

we apply (7.10) with  $a = \frac{1}{N}$  and  $L = \frac{KT}{N}$  to obtain, for  $i = 2, 3$ ,

$$IP \left( \sup_{s \leq T} |M_s^{N,i}| \geq z \right) \leq 2 \exp \left( -\frac{1}{2} \frac{z^2}{KT} \psi \left( \frac{z}{KT} \right) N \right).$$

Choosing  $z = \frac{\varepsilon KT}{N^{1/4}}$ , for some  $0 < \varepsilon < 1$ , this yields, for  $i = 2, 3$ ,

$$IP \left( \sup_{s \leq T} |M_s^{N,i}| \geq \frac{\varepsilon KT}{N^{1/4}} \right) \leq 2 \exp \left( -\frac{1}{2} \varepsilon^2 \psi(1) \sqrt{N} KT \right) \leq 2 \exp \left( -\frac{1}{2} \varepsilon^2 \psi(1) \sqrt{N} \right), \quad (7.11)$$

where we have used that  $KT \geq 1$ , that  $\psi(\frac{z}{KT}) \geq \psi(1)$ , since  $\psi$  is decreasing, and  $\frac{z}{KT} \leq 1$ .

We now treat  $M_t^{N,1}$ . In the following we shall apply (7.10) with  $a = 0$ . Then

$$H_T^0 \leq \frac{\alpha^2}{N^2} \sum_{j=1}^N \int_0^T \int_{\mathbb{R}_+} (R_{s-}^N(j))^2 1_{\{z \leq \varphi(U_{s-}^N)\}} \mathbf{M}^j(ds, dz) + \frac{\alpha^2 K}{N^2} \sum_{j=1}^N \int_0^T (R_s^N(j))^2 ds.$$

Using that  $R_s^N(j) \leq R_0^N(j) + J_T(j)$ , we get

$$\begin{aligned} & \frac{\alpha^2}{N^2} \sum_{j=1}^N \int_0^T \int_{\mathbb{R}_+} (R_{s-}^N(j))^2 1_{\{z \leq \varphi(U_{s-}^N)\}} \mathbf{M}^j(ds, dz) \\ & \leq \frac{2\alpha^2}{N} \left( \frac{1}{N} \sum_{j=1}^N \left[ (R_0^N(j))^2 + (J_T(j))^2 \right] J_T(j) \right). \end{aligned}$$

This implies that  $H_T^0$  possesses exponential moments; that is, there exists some  $\nu > 0$  such that  $\mathbb{E} e^{\nu H_T^0} < \infty$ . Moreover, since  $\sup_s \mathbb{E} (R_s^N(j))^2 < \infty$ ,  $\mathbb{E} H_T^0 \leq 2 \frac{\alpha^2 K}{N} T \sup_t \mathbb{E} (R_s^N(j))^2 =$

$CT/N$ . Therefore, using a large deviations upper bound, choosing  $c_T = CT + 1$ , we deduce that for a suitable constant  $C$ ,

$$\mathbb{P}(H_T^0 \geq c_T/N) \leq Ce^{-cN}.$$

Applying (7.10) with  $L = c_T/N$ ,  $z^2 = \varepsilon^2 c_T N^{-1/2}$  and  $a = 0$ , we deduce from this that

$$\mathbb{P}(\sup_{s \leq T} |M_s^{N,1}| \geq \varepsilon \sqrt{c_T} N^{-1/4}) \leq Ce^{-\frac{1}{2}\varepsilon^2 \sqrt{N}} + Ce^{-cN} \leq Ce^{-c\varepsilon^2 \sqrt{N}}.$$

Consequently, choosing e.g.  $\eta = \varepsilon N^{-1/4}$ ,

$$\begin{aligned} \mathbb{P}\left(\sup_{s \leq T} \left(|U_s^N - u_s| + \left|\frac{1}{N} \sum_{i=1}^N R_s^N(i) - r_s\right|\right) \geq N^{-1/4}(C_R)^{T+1}[2\varepsilon + \varepsilon T^{1/2} + 2KT\varepsilon]\right) \\ \leq CT e^{-c\varepsilon^2 \sqrt{N}}. \end{aligned}$$

Taking  $N_0$  such that  $N_0^{-1/4}(C_R)^{T+1}[2 + T^{1/2} + 2KT] \leq N_0^{-1/5}$ , we obtain the assertion.  $\square$

We close this section with the proof of the second item of Theorem 2.5.

**Proof of the third item of Theorem 2.5** Since  $(u^{max}, r^{max})$  is locally attracting,  $t_1$  is finite (and deterministic). We apply the first item of Theorem 2.5 with  $T = t_1$ . It implies that for all  $N \geq N_0$ ,

$$\begin{aligned} \mathbb{P}\left(|U_{t_1}^N - u_{t_1}| + \left|\frac{1}{N} \sum_{i=1}^N R_{t_1}^N(i) - r_{t_1}\right| \geq \varepsilon\right) \\ \leq C_{t_1} e^{-c_{t_1} \varepsilon^2 \sqrt{N}}, \end{aligned}$$

which implies the assertion.  $\square$

**Acknowledgements** The authors thank two anonymous referees for useful remarks and careful reading. AG and EL thank the Gran Sasso Science Institute (GSSI) for hospitality and support. This research is part of USP project *Mathematics, computation, language and the brain* and of FAPESP project *Research, Innovation and Dissemination Center for Neuromathematics* (Grant 2013/07699-0). AG is partially supported by CNPq fellowship (Grant 311 719/2016-3).

## References

- Barak, O., Tsodyks, M.: Persistent activity in neural networks with dynamic synapses. *PLoS Comput. Biol.* **3**(2) (2007). <https://doi.org/10.1371/journal.pcbi.0030035>
- Blackman, D., Vigna, S.: Scrambled Linear Pseudorandom Number Generators. *CoRR*, [arXiv:1805.01407](https://arxiv.org/abs/1805.01407) (2018)
- Brillinger, D., Segundo, J.P.: Empirical examination of the threshold model of neuron firing. *Biol. Cybern.* **35**, 213–220 (1979)
- Chevallier, J., Caceres, M.J., Doumic, M., Reynaud-Bouret, P.: Microscopic approach of a time elapsed neural model. *Math. Mod. Methods Appl. Sci.* **25**(14), 2669–2719 (2015)
- Chornoboy, E., Schramm, L., Karr, A.: Maximum likelihood identification of neural point process systems. *Biol. Cybern.* **59**, 265–275 (1988)
- Ditlevsen, S., Löcherbach, E.: Multi-class oscillating systems of interacting neurons. *Stoch. Proc. Appl.* **127**, 1840–1869 (2017)
- Duarte, A., Ost, G.: A model for neuronal activity in the absence of external stimuli. *Markov Process. Relat. Fields* **22**, 37–52 (2016)
- Dzhaparidze, K., van Zanten, J.H.: On Bernstein-type inequalities for martingales. *Stoch. Process. Appl.* **93**(1), 109–117 (2001)

9. Hansen, N., Reynaud-Bouret, P., Rivoirard, V.: Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21**(1), 83–143 (2015)
10. Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*. North Holland, Amsterdam (1989)
11. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin (2003)
12. Kistler, W.M., van Hemmen, L.: Short-term synaptic plasticity and network behavior. *Neural Comput.* **11**, 1579–1594 (1999)
13. Markram, H., Tsodyks, M.: Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* **382**(6594), 807–810 (1996)
14. Mongillo, G., Barak, O., Tsodyks, M.: Synaptic theory of working memory. *Science* **319**, 1543 (2008)
15. Ogata, Y.: On Lewis' simulation method for point processes. *IEEE Trans. Inf. Theory* **IT-27**, 23–31 (1981)
16. Reynaud-Bouret, P., Rivoirard, V., Grammont, F., Tuleau-Malot, C.: Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *J. Math. Neurosci. (JMN)* **4**, 1–41 (2014)
17. Schmutz, V., Gerstner, W., Schwaiger, T.: Mesoscopic population equations for spiking neural networks with synaptic short-term plasticity. [arXiv:1812.09414](https://arxiv.org/abs/1812.09414) (2018)
18. Seeholzer, A., Deger, M., Gerstner, W.: Stability of working memory in continuous attractor networks under the control of short-term plasticity. *bioRxiv* 424515 (2018). <https://doi.org/10.1101/424515>
19. Sznitman, A.-S.: Topics in propagation of chaos. In: *École d'Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math.* Springer, Berlin, pp. 165–251 (1991)
20. Tsodyks, M., Markram, H.: The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci.* **94**(2), 719–723 (1997)
21. Tsodyks, M., Pawelzik, K., Markram, H.: Neural networks with dynamic synapses. *Neural Comput.* **10**(4), 821–835 (1998)
22. Tsodyks, M., Wu, S.: Short-term synaptic plasticity. *Scholarpedia* **8**(10), 3153 (2013)
23. Zucker, R.S., Regehr, W.G.: Short-term synaptic plasticity. *Annu. Rev. Physiol.* **64**, 355–405 (2002)

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