Persistence in a large network of locally interacting neurons

Maximiliano Altamirano, Roberto Cortez, Matthieu Jonckheere, Lasse Leskelä

Abstract
This article presents a biological neural network model driven by inhomogeneous Poisson processes accounting for the intrinsic randomness of synapses. The main novelty is the introduction of local interactions: each firing neuron triggers an instantaneous increase in electric potential to a fixed number of randomly chosen neurons. We prove that, as the number of neurons approaches infinity, the finite network converges to a nonlinear mean-field process characterised by a jump-type stochastic differential equation. We show that this process displays a phase transition: the activity of a typical neuron in the infinite network either rapidly dies out, or persists forever, depending on the global parameters describing the intensity of interconnection. This provides a way to understand the emergence of persistent activity triggered by weak input signals in large neural networks.

Keywords: biological neural network, metastability, mean-field limit, jump process, interacting particle system, phase transition, propagation of chaos, nonlinear Markov process

Contents

1 Introduction
1.1 Biological neural network models .................................................. 2
1.2 Original aspect of our model and main contribution .......................... 3

2 Model description
2.1 Definition as a piecewise deterministic Markov process ...................... 5
2.2 Description as a solution of a stochastic differential equation ................. 5
2.3 Construction using a discrete-time embedded Markov chain ................. 6
2.4 Mean-field process ................................................................. 7

3 Main results
3.1 Long-term behaviour of the finite-neuron system ............................ 7
3.2 Phase transition in the mean field .............................................. 8
3.3 Approximation using the mean field .......................................... 9
3.4 Persistence .......................................................................... 10

*Instituto de Cálculo, Universidad de Buenos Aires/CONICET, Argentina, E-mail: maximiliano.altamirano@ic.fcen.uba.ar.
†Universidad Andrés Bello, Departamento de Matemáticas. E-mail: roberto.cortez.m@unab.cl. Supported by Iniciación Fondecyt Grant 11181082 and by Programa Iniciativa Científica Milenio through Nucleus Millenium Stochastic Models of Complex and Disordered Systems.
‡Instituto de Cálculo, Universidad de Buenos Aires/CONICET, Argentina, E-mail: mjonckhe@dm.uba.ar.
§Aalto University, School of Science, Department of Mathematics and Systems Analysis. Otakaari 1, 02015 Espoo, Finland. E-mail: lasse.leskela@aalto.fi.
1 Introduction

1.1 Biological neural network models

Since the seminal work by Lapicque [16], the body of mathematical literature on biological neural networks has become extensive. Nevertheless, several fundamental questions still remain open due to the extreme complexity inherent to this biological system. Among these, an important one concerns the understanding of the phenomenon of persistence: given a mild input, the network activity is sustained for surprisingly large times. It has been observed in a diverse set of brain regions and organisms, and it is considered as a mechanism for short-term storage [13, 17, 29]. This is one of our main interests here.

Before precisng our contributions, we briefly recall some of the main different efforts of research concerning biological neural networks. On one side of the spectrum, conductance-based models deal in detail with the chemical interchanges between neurons when synapses occur. The Hodgkin-Huxley model is classical among those [14], and there has been a number of recent works within this framework, see [1, 27, 26]. In general, these models lead to complex equations, which can be simplified in some specific cases [25]. On the other side of the spectrum, there are the integrate-and-fire models, Lapicque’s model belonging to this large group, see Burkitt for reviews [4, 5]. Considering the fast, stereotyped nature of neuron synapses, these models focus on the important problem of time distribution of spikes, which are seen as instantaneous events. In simple words, the membrane potential is the magnitude taken into account for each neuron, and it is considered to integrate the potential received by other neurons, until it reaches some threshold—usually fixed—which causes the neuron to fire spikes, received by other neurons in the network. In addition, leaky integrate-and-fire (LIF) models take into account the decay of the membrane potential of neurons towards its resting state, due to the leak of potential across the cellular membrane when not interacting.

Most of the aforementioned works model random effects through an external Gaussian process representing the aggregate effect of other neurons affecting the network. This allows Itô calculus to be applied when making computations and deriving results. There are, however, some several limitations regarding the classical integrate-and-fire models. For instance, these models do not take into account, in general, the intrinsic randomness of neurons, which is an essential feature that cannot be ignored, see for example [23]. Some recent models do include some form of intrinsic randomness by introducing white noise associated to each neuron [11], which provides a partial solution regarding this topic. Another issue lies in the so called avalanche effect caused by the presence of a fixed threshold above which the neurons deterministically fire. More specifically, the neuron potential blows-up in finite time in the mean-field limit (see for instance [6]), which does not properly represent the behaviour of a large population of neurons.
**Linear-nonlinear-Poisson models** (LNP) take a step forward from LIF models, addressing the aforementioned issues in an effective way. The first LNP model was introduced by Chichilnisky [8]. These models represent the behaviour of spikes using inhomogeneous Poisson processes depending on the membrane potential. The model presented by Robert and Touboul [22] belongs to this family and is the main inspiration of our framework. Our model is indeed a modification of the former and belongs to this family as well. LNP models seem to represent neuronal firing more accurately, display a good fit with experimental data [20, 21], and avoid the problem of blow-up. In this context, a well defined limiting behaviour is reached. As underlined in [22], the associated mean-field limit follows a McKean–Vlasov equation where the dynamics of the process depends on mean quantities of the same process.

For diffusive underlying motions, McKean–Vlasov equations associated to potentials have been extensively studied in the last decades. See for instance the seminal paper [2] and many ulterior developments (see [18, 28] and references therein). In most works stemming from statistical physics perspectives, convergence (as \( N \to \infty \)) of a particle system towards its associated mean-field limit is obtained for fixed times. A convergence result of this kind is typically intrinsically linked to the asymptotic independence between particles, and is referred to as propagation of chaos. See also [3] for the use of the Wasserstein distance to prove propagation of chaos in that context. Let us note however that in these references, strong results on the speed of convergence towards hydrodynamic limits are usually proven in the case of a unique stationary measure and/or building on strong regularity property of the limiting PDE. Similarly, the literature of particle systems undergoing energy-preserving binary local (jump-type) interactions and their respective limiting equation is quite extensive, see for instance the seminal work [15], and more recently [19], and the references therein. In this setting, many results regarding convergence to equilibrium and propagation of chaos have been obtained.

In contrast, the local dynamics of the model of the present article involves simultaneous jumps between 3 or more neurons (particles), with no preserved quantities (even on average). Thus, the limiting nonlinear equation that we obtain is of a different nature. For instance, as we shall see, the limit equation can have two stationary distributions, depending on the parameters of the model. This has strong consequences for propagation of chaos and speed of convergence results, making them harder to prove and heavily dependent on model parameters. We remark that multiplicity of stationary distributions is also present in [22], but without any rate of convergence under the stationary regime nor selection principle for the limiting stationary distribution. Some results concerning metastability properties of McKean–Vlasov equations having several stationary distributions are considered in [7].

Consequently, many of the classical techniques used in more physical models cannot be applied directly in the present context. Nevertheless, our propagation of chaos study relies on a recent coupling technique for binary interactions developed in [9], which we adapt and extend to our setting.

### 1.2 Original aspect of our model and main contribution

Following Robert and Touboul [22], we consider both a particle system describing potentials in a finite excitatory network of neurons, and a nonlinear stochastic process approximating the behaviour of a typical neuron in a large network. Our model describes the action potential of a network of \( N \) neurons over time, denoted \( X_t = (X_{1t}, \ldots, X_{Nt}) \). Each neuron \( X_{it} \in \mathbb{R}_+ \) is affected by an exponential decay towards its resting potential, set as 0. Also, some random interactions take place within the network. Those occur when a neuron produces a *spike*—also known as
firing—: it is instantaneously reset to its resting potential and, at the same time, it randomly selects a fixed number \( \kappa \in \mathbb{N} \) of other neurons and gives them an excitatory impulse. Those events take place at a state-dependent rate. In contrast with [22], the dynamics we consider here is local, in the sense that, in each firing event, the firing neuron communicates with a fixed number of other neurons (\( \kappa \)) and not all of them at once, and the magnitude of the spikes is of order 1 and not \( 1/N \). As a consequence, in the limit equation, the firing mechanism gives rise to a nonlinear jump term, and not a drift term as in [22]. We chose this modelling assumption firstly because it allows to model arguably more interesting dynamics from a biological point of view (a natural extension for future work is to consider an underlying graph structure between communicating neurons). Secondly, it exhibits a different (and possibly more meaningful) phase transition concerning the long-time behaviour of the process, which has a natural physical interpretation (see Section 3.2). Last but not least, the mathematical techniques needed to quantify neurons decorrelations are different from the ones in [22] and are possibly generalizable to a larger class of similar models.

The main structure of our discussion is the following. First, we study the large-time behaviour of the finite network (Theorem 3.1). We then characterize the mean-field limit process—which we call nonlinear process—and show its existence and uniqueness (Theorem 3.2). We show that there is a simple phase transition in terms of the parameters of the model: if the transmission of potential is strong enough (in terms of strength and frequency), then the limiting process is active forever; otherwise, it decays to 0 (Theorem 3.3).

We then study propagation of chaos: we prove convergence, for fixed time, of the empirical measure of the finite network towards its mean-field limit, when the number of neurons grows to infinity (Theorem 3.7). To achieve this, a coupling argument is developed, in the same spirit as in [9]. We start by defining a coupling \( Z_t = (Z^1_t, \ldots, Z^N_t) \) such that each \( (Z^i_t)_t \) is a nonlinear process. In a second step, we define a coupling “close” to the first one, but with independent coordinates. Then, some computations with both couplings allow us to prove convergence in terms of the Wasserstein distance and to obtain interesting rates of convergence.

Finally, we show the following persistence phenomenon (Theorem 3.9):

- if the transmission of potential is strong enough, then, even though the finite network activity dies in finite time for any \( N \), it remains active at least for a time of order \( \log(N) \).
- otherwise, the networks activity dies out at a speed independent of \( N \).

The rest of the paper is organised as follows. Section 2 describes the interacting neuron model and its hydrodynamic limit characterised by a McKean–Vlasov equation. Section 3 presents the main results: the behaviour of the network for finite \( N \), the well-posedness of the limiting stochastic differential equation, the phase transition exhibited by the limiting process in terms of global parameters representing the intensity of connection, the convergence as \( N \to \infty \) of the empirical measure of the finite network towards the mean-field, and the persistence result. Section 4 gives detailed proofs. Finally, in Section 5 we mention possible further lines of research that might stem from this work.

## 2 Model description

A system of \( N \) neurons is modelled as an interacting particle system where the potential of each neuron is represented as a stochastic process with values in \( \mathbb{R}_+ = [0, \infty) \). We present three viewpoints to this particle system: piecewise deterministic Markov process (Section 2.1), a
solution of a stochastic differential equation driven by Poisson noise (Section 2.2), and a discrete-time Markov chain obtained by sampling the system at firing instants (Section 2.3). Thereafter, Section 2.4 presents a simplified McKean–Vlasov stochastic differential equation which is used to model systems with a large number of neurons. The model is characterised by four parameters:

- decay rate $\mu > 0$,
- firing rate proportionality constant $\gamma > 0$,
- impulse range $\kappa \in \mathbb{N}$ (number of neurons excited at a firing event),
- impulse magnitude $\rho > 0$.

In what follows, we always assume that the initial condition $X_0 = (X_0^1, \ldots, X_0^N)$ is a collection of i.i.d. copies of a given random variable $Z_0 \in \mathbb{R}_+$. Consequently, $X_t = (X_t^1, \ldots, X_t^N)$ is exchangeable for all $t \geq 0$, in the sense that its distribution is invariant with respect to permuting the neuron indices.

### 2.1 Definition as a piecewise deterministic Markov process

Consider a system of $N$ neurons having potential levels $X_t^1, \ldots, X_t^N \geq 0$ at time $t$. Level zero indicates a resting potential. Each neuron has a decay rate $\mu > 0$ and a firing rate $\gamma > 0$. In the absence of firings, the level of each neuron decays according to $\frac{d}{dt}X_t^i = -\mu X_t^i$. Firings of neuron $i$ are triggered at the events of an inhomogeneous Poisson process with state-dependent intensity $\gamma X_t^i$, independently of other neurons. When a neuron fires, it jumps to zero, and it selects $\kappa \geq 1$ other neurons uniformly at random, which instantly increase their levels by $\rho > 0$ units. We note that, unlike in [22], there is no scaling (in $N$) of any of the model parameters $\kappa, \rho, \mu, \gamma$.

The trajectory of $X_t = (X_t^1, \ldots, X_t^N)$ is a piecewise deterministic continuous-time Markov process [10] on state space $[0, \infty)^N$ with flow $(x, t) \mapsto xe^{-\mu t}$, jump rate $x \mapsto \gamma \|x\|$ where $\|x\| = \sum_i |x_i|$, and jump kernel

$$(x, A) \mapsto \sum_i \sum_{K: K \not\ni i} \frac{x^i}{\|x\|} \left( \frac{N - 1}{\kappa} \right)^{-1} \delta_{x-x^i e_i + \rho e_K}(A),$$

where the second sum on the right is over subsets $K \subset \{1, \ldots, N\}$ of size $\kappa$, $e_i$ denotes the $i$-th unit vector and $e_K = \sum_{j \in K} e_j$.

### 2.2 Description as a solution of a stochastic differential equation

Alternatively, the system trajectory relative to a given initial state can be represented as the a.s. unique strong solution of a system of stochastic differential equations. Hence, for any $i = 1, \ldots, N$, the behaviour of $X_t^i$ can be represented by

$$dX_t^i = -\mu X_t^i dt - X_t^i \sum_{K: K \not\ni i} \int_0^\infty 1(u \leq \gamma X_t^i) \mathcal{N}_K(dt, du, (i-1, i))$$

$$+ \rho \sum_{K: K \not\ni i} \int_0^\infty \int_0^N 1(u \leq \gamma X_t^{[\xi]^i}) \mathcal{N}_K(dt, du, d\xi),$$

(2.1)
in which all sums involving $K$ are taken with respect to subsets of $\{1, \ldots, N\}$ of size $\kappa$, and $N_K(dt, du, d\xi)$, $K \subset \{1, \ldots, N\}$, are mutually independent Poisson random measures on $\mathbb{R}_+ \times \mathbb{R}_+ \times (0, N]$ with common intensity measure

$$
\left(\frac{N - 1}{\kappa}\right)^{-1} dt \, du \, \mathbf{1}(\lceil \xi \rceil \not\in K) d\xi.
$$

Any atom $(T, U, \Xi)$ of the Poisson random measure $N_K$ corresponds to a possible firing where the neurons in $K$ are excited: $T \geq 0$ is the time of the jump, the mark $U \geq 0$ is used to model inhomogeneities, and $\Xi \in (0, N]$ is used to select uniformly the neuron $i = \lceil \Xi \rceil \not\in K$ that fires.

### 2.3 Construction using a discrete-time embedded Markov chain

We now give a third construction for our model using an embedded Markov chain, which is especially well suited for simulating the network.

A system started at state $x = (x_1, \ldots, x_N)$ evolves as follows. Observe first that the total firing rate in state $X_t$ equals $\gamma \|X_t\|$, and as long as firings do not occur, the sum of potentials decays according to $\|X_t\| = \|x\| e^{-\mu t}$. Therefore, the probability that there are no firings during $[0, t]$ is the same as the probability that an inhomogeneous Poisson point process with intensity $\lambda_t = \gamma \|x\| e^{-\mu t}$ has no jumps during $[0, t]$. Because the number of such jumps during $[0, t]$ is Poisson distributed with mean $\int_0^t \lambda_s ds = (\gamma/\mu) \|x\| (1 - e^{-\mu t})$, we see that the distribution of the first firing time $\tau$ is characterised by

$$
\mathbb{P}(\tau > t) = e^{-(\gamma/\mu) \|x\| (1 - e^{-\mu t})}, \quad 0 \leq t < \infty.
$$

In particular, the system never fires with probability $\mathbb{P}(\tau = \infty) = e^{-(\gamma/\mu) \|x\|}$. Another key observation is that the proportions of neuron levels remain constant in the absence of firings. Therefore, given $\tau < \infty$, the label of the first firing neuron $I$ is distributed according to

$$
\mathbb{P}(I = i) = \frac{x_i}{\|x\|}, \quad i = 1, \ldots, N.
$$

Using (2.2)–(2.3), we can formally define the trajectory of a system starting at $X_0 \in \mathbb{R}_+^N$ as follows. We denote by $\hat{X}_k$ the state of the system immediately after the $k$-th firing, and by $\hat{T}_k \in [0, \infty]$ the time of the $k$-th firing, with $\hat{T}_k = \infty$ indicating that there are fewer than $k$ firings in total. The random variables $(\hat{X}_k, \hat{T}_k)$, $k = 0, 1, \ldots$, are recursively defined as follows, starting with $(\hat{X}_0, \hat{T}_0) = (X_0, 0)$. For any $k \geq 1$, set $x = \hat{X}_{k-1}$ and sample the $k$-th waiting time $\tau_k$ from distribution (2.2). If $\tau_k = \infty$, set $(\hat{X}_k, \hat{T}_k) = (0, \infty)$. Otherwise sample the label of the $k$-th firing neuron $I_k$ from distribution (2.3), and then sample a set of excited neurons $K_k$ uniformly at random from the subsets of $\{1, \ldots, N\} \setminus \{I_k\}$ of size $\kappa$. Then set

$$
(\hat{X}_k, \hat{T}_k) = \left( (x - x^I e_{I_k}) e^{-\mu \tau_k} + \rho e_{K_k}, \hat{T}_{k-1} + \tau_k \right),
$$

The full continuous-time trajectory of the system is then defined by

$$
X_t = \hat{X}_k e^{-\mu (t - \hat{T}_k)} \quad \text{for} \quad \hat{T}_k \leq t < \hat{T}_{k+1}.
$$

Equations (2.4)–(2.5) provide a simple way to simulate system trajectories. Sampling from the distribution (2.2) can be done by $\tau = -\frac{1}{\mu} \log((1 - \frac{\mu}{\gamma \|x\|})_+)$ where $\xi$ is exponentially distributed with mean one and $\log 0 = -\infty$, so that $\tau = \infty$ when $\xi \geq \frac{\gamma \|x\|}{\mu}$. We also find that $\lim_{t \to \infty} X_t = 0$ if and only if the total number of firings is finite.
2.4 Mean-field process

When the number of neurons $N$ is large, we may heuristically derive a mean-field approximation of the system as follows. Fix a reference neuron $i$. In the absence of firings, the neuron level decays exponentially according to $\frac{d}{dt}X^i_t = -\mu X^i_t$, and the neuron fires at rate $\gamma X^i_t$. When any other neuron fires, neuron $i$ is excited with probability $\frac{1}{N} \sum_{j < i} X^j_t = \frac{\kappa N - 1}{N - 1}$. Assuming that the empirical sample average $\frac{1}{N} \sum_j X^j_t$ of the neuron levels is close to the mathematical expectation $E[X^i_t]$, the average rate at which neuron $i$ is excited is approximately

$$\frac{\kappa}{N - 1} \sum_{j \neq i} X^j_t \approx \gamma \kappa E[X^i_t].$$

This suggests that the evolution of a reference neuron in a large system can be approximated by the solution of the McKean-Vlasov equation

$$dZ_t = -\mu Z_t dt - Z_t \int_0^\infty 1(u \leq \gamma Z_t)N(dt, du)$$

$$+ \rho \int_0^\infty 1(u \leq \gamma E[Z_t])M(dt, du),$$

in which $N(dt, du)$ and $M(dt, du)$ denote independent Poisson random measures on $\mathbb{R}_+^2$ with intensities $dt du$ and $\kappa dt du$, respectively.

Note that the third term in (2.6), which results from the local firing mechanism, is of jump-type. As mentioned before, this is a key difference from the model in [22], in which the firing mechanism is global and gives rise to a drift term. Notice also that the rate at which $Z_t$ jumps upwards depends on its own expected value; it is in this sense that the process is nonlinear.

3 Main results

3.1 Long-term behaviour of the finite-neuron system

Theorem 3.1 below tells us that in a finite system of neurons, all neuron potentials converge to zero, regardless of initial state and model parameters. This result also implies that the unique invariant probability distribution of the $\mathbb{R}_+^N$-valued Markov process is the Dirac measure at $0 \in \mathbb{R}_+^N$. However, the nature of this convergence will vary drastically, depending on the parameter values. For future reference, we define a reproduction number

$$\theta = \kappa (1 - e^{-\rho \gamma/\mu}),$$

which will play a crucial role in our study. See Section 3.2 for a discussion on the interpretation of this quantity.

Theorem 3.1 (Long-term behaviour of the finite network). For any initial state, the process $X^i_t$ converges according to $\lim_{t \to \infty} X^i_t = 0$ almost surely. Moreover, when $\theta < 1$, the convergence of any particular neuron $i$ occurs exponentially fast (independent of $N$) according to

$$E[\omega(X^i_t, 0)] \leq E[\omega(X^i_0, 0)] e^{-(1-\theta)\mu t}$$

for the metric

$$\omega(x, y) = 1 - e^{-(\gamma/\mu)|x-y|}.$$
The global decay property described in Theorem 3.1 is remarkable because it is valid for arbitrarily large values of the firing rate $\gamma$. To appreciate this feature, we note that in a slightly modified system in which the potentials of firing neurons remain constant instead of being reset to zero, the associated Markov process $(Y_t)_{t \geq 0}$ satisfies $\frac{d}{dt}E[\|Y_t\|] = (\rho\kappa\gamma - \mu)E[\|Y_t\|]$, so that the modified system remains bounded in mean only for $\gamma < \frac{\mu}{\rho\kappa}$.

### 3.2 Phase transition in the mean field

As the number of neurons $N \to \infty$, one expects that any single neuron in the network converges to a nontrivial limit process driven by (2.6). The well-posedness of the SDE (2.6) representing the mean-field approximation is confirmed by the following result.

**Theorem 3.2** (Well-posedness of the mean-field SDE). For any initial state $Z_0 \in \mathbb{R}_+$, there exists a unique strong solution $(Z_t)_{t \geq 0}$ to (2.6).

The main result of this work is the following description of a phase transition of the mean-field limiting process, characterised by the reproduction number $\theta$ defined by (3.1).

**Theorem 3.3** (Phase transition in the mean-field limit). Assume that $E[Z_0] > 0$ and $E[Z_0^2] < \infty$. Then $(Z_t)_{t \geq 0}$ is uniformly integrable, and the following phase transition holds:

- If $\theta < 1$, then $E[Z_t] \to 0$ and $\int_0^\infty E[Z_t] dt < \infty$.
- If $\theta = 1$, then $E[Z_t] \to 0$ and $\int_0^\infty E[Z_t] dt = \infty$.
- If $\theta > 1$, then $\inf_{t \geq 0} E[Z_t] > 0$.

Moreover, for $\theta < 1$, we have $Z_t \to 0$ almost surely, and this convergence is exponentially fast in the metric $\omega$ defined by (3.3), according to

$$E[\omega(Z_t, 0)] \leq E[\omega(Z_0, 0)] e^{-(1-\theta)\mu t}. \quad (3.4)$$

Thus, the mean field exhibits nontrivial long-run behaviour if and only if $\theta > 1$.

A heuristic explanation for this is the following. In analogy with (2.2), we observe that $p = 1 - e^{-\rho\gamma/\mu}$ equals the probability that a neuron which gets excited from the resting potential will eventually fire. In a large system where initially an overwhelming majority of neurons are resting, the number of eventually firing neurons excited by a firing neuron is approximately binomially distributed with success probability $p$ and mean $\theta = \kappa p$. An associated Galton–Watson branching process may survive forever if and only if $\theta > 1$. Therefore:

- When $\theta > 1$, the resets and firings have enough frequency and power to sustain themselves forever.
- If $\theta < 1$, then fewer and fewer firings occur as time grows, and eventually the process dies out due to the exponential decay or a final reset.
- Interestingly enough, in the critical case $\theta = 1$, even though $E[Z_t] \to 0$, the process $Z_t$ never really dies: firings do become more infrequent with time, but, since $\int_0^\infty E[Z_t] dt = \infty$, they never actually stop.
Remark 3.4. The model of Robert and Touboul also exhibits a similar phase transition, but the condition at which it takes place is different from ours. Indeed, in their model, for the case of a linear firing-rate function, the phase transition occurs when the quantity $\theta_c = \kappa \rho \gamma / \mu$ exceeds 1, see [22, Section 7.2.1] (we used the notation of the present paper; note that, because the parameter $\kappa$ has no equivalent in their setting, we have interpreted $\kappa \rho$ as the magnitude of the excitatory impulse in their model, before rescaling by $1/N$). Because $\theta \leq \theta_c$, we see that persistence in our model requires a higher value of $\rho \gamma / \mu$, i.e., the magnitude of the excitatory impulse times the ratio of firing rate to decay rate. This difference is a consequence of the local nature of the firing mechanism we consider here, and it is the most distinctive and interesting feature of our model. However, our upper bounds for the finite network (Theorems 3.7 and 3.9) indicate that $\theta_c$ might have relevance also for the present model.

Remark 3.5. Theorem 3.3 implies that if $\mathbb{E}[Z_0] > 0$ and $\theta > 1$, then any weak limit of $Z_t$ satisfies $\mathbb{E}[Z_\infty] > 0$. We conjecture that for each $\theta > 1$, the process $Z_t$ admits exactly two stationary distributions: the trivial one (Dirac mass at 0), which is reached only when $Z_0 = 0$ almost surely, and a second, nontrivial distribution, towards which $Z_t$ approaches whenever $\mathbb{E}[Z_0] > 0$. This nontrivial stationary distribution could be computed through a fixed point procedure. The construction in Section 2.3 gives a simple way to approximately simulate this distribution.

As a consequence of Theorem 3.3, we obtain the following result regarding the limiting proportion of neurons at the resting potential in the infinite network.

Corollary 3.6. The process $Z_t$ satisfies for all $t \geq 0$

$$
\mathbb{P}(Z_t = 0) = \frac{1}{\kappa} + \left( \mathbb{P}(Z_0 = 0) - \frac{1}{\kappa} \right) e^{-\gamma \kappa \int_0^t \mathbb{E}[Z_s] ds}. \tag{3.5}
$$

Consequently, $\lim_{t \to \infty} \mathbb{P}(Z_t = 0) = \frac{1}{\kappa}$ when $\mathbb{E}[Z_0] > 0$ and $\theta \geq 1$.

3.3 Approximation using the mean field

The following result confirms that the behaviour of a typical neuron is approximated by the nonlinear mean-field SDE (2.6), when the number of neurons $N \to \infty$. The upper bounds on convergence rate are expressed in terms of another fundamental constant,

$$
\theta_c = \kappa \rho \gamma / \mu, \tag{3.6}
$$

which dominates the reproduction number defined by (3.1) according to $\theta \leq \theta_c$. See also Remark 3.4. The subscript $c$ in $\theta_c$ stands for “chaos”.

We recall that, given probability measures $\nu, \tilde{\nu}$ on $\mathbb{R}$, their 1-Wasserstein distance is defined as

$$
W_1(\nu, \tilde{\nu}) = \inf \mathbb{E}[|Y - \tilde{Y}|],
$$

where the infimum is taken over all couplings of $\nu$ and $\tilde{\nu}$, that is, over all possible random pairs $(Y, \tilde{Y})$ where $Y \sim \nu$ and $\tilde{Y} \sim \tilde{\nu}$. $W_1$ is a complete metric on the space of probability measures with finite first moment. Notice that any coupling immediately provides an upper estimate for $W_1$; we will use this fact often.
Theorem 3.7 (Propagation of chaos). Assume that $\kappa = 2$ and $E[(X_0^1)^2] < \infty$. Then there exist a constant $C > 0$ such that for all $N$ and all $t$, the expected Wasserstein-1 distance between $X_t = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i^t}$ and $f_t = \text{Law}(Z_t)$ satisfies

$$E[W_1(f_t, X_t)] \leq \frac{C}{N^{1/3}} \times \begin{cases} 1, & \theta_c < 1, \\ 1 + t, & \theta_c = 1, \\ e^{(\theta_c - 1)\mu t}, & \theta_c > 1. \end{cases}$$

Theorem 3.7 proves propagation of chaos in terms of the expected $W_1$-distance between the empirical measure of the finite network and its mean-field limit. We mention that, besides convergence as $N \to \infty$, the term “propagation of chaos” typically also refers to the asymptotic independence between particles in the system, but this is equivalent to the convergence of its empirical measure, see for instance the classical notes of Sznitman [24]. Rate $N^{-1/3}$ in the upper bound appears reasonable, considering that the convergence of the empirical measure of a sequence of i.i.d. variables towards their common law occurs at rate $N^{-1/2}$ when measured in expected $W_1$-distance, see [12] for details. The rate is uniform in time in the case $\theta_c < 1$.

Theorem 3.7 is proved in Section 4.4 using a coupling argument introduced in [9]. To this end, we will write the SDE (2.1) of a neuron in the finite network, and the mean-field SDE (2.6), in a slightly different way, which is easier to handle and displays similarities between the two processes in a natural fashion.

Remark 3.8. For an arbitrary $\kappa \in \mathbb{N}$ fixed, or even a random $\kappa \in \{1, \ldots, N - 1\}$ with fixed mean and bounded variance, the main idea of the proof still works. However, the full proof requires significant additional notation and a more careful treatment, rendering the argument unnecessarily technical and difficult to follow, while adding little to no new insights. Thus, we decided to treat only the case $\kappa = 2$.

### 3.4 Persistence

We can now combine the previous results and state the following theorem that sheds light on the phenomenon of persistence, in terms of the reproduction number $\theta$ defined in (3.1) and the threshold parameter $\theta_c$ in (3.6).

Theorem 3.9 (Persistence). Assume that $\kappa = 2$ and $E[(X_0^1)^2] < \infty$. If $E[X_0^1] > 0$, then the finite network satisfies the following dichotomy:

- If $\theta > 1$, then there exist constants $c > 0$, $\bar{c} \in \mathbb{R}$, and $\epsilon > 0$ (not depending on $N$) such that for all $N \geq 1$,
  $$E[X_t^1] \geq \epsilon \quad \text{for all } t \leq \bar{c} + c \log N.$$

- If $\theta < 1$, then for the metric $\omega$ defined by (3.3), we have
  $$E[\omega(X_t^1, 0)] \leq E[\omega(X_0^1, 0)] e^{-(1-\theta)\mu t}.$$

  If we additionally assume that $\theta_c < 1$, then
  $$E[X_t^1] \leq E[X_0^1] e^{-(1-\theta_c)\mu t}.$$

Theorem 3.9 states that in the case $\theta > 1$, the activity of the finite network persists for a time of order (at least) $\log N$; but if $\theta < 1$, then the activity dies out at a rate independent of $N$. 

Remark 3.10. We believe that log $N$ is a crude lower estimate on the time it takes for the activity to decay. Numerical simulations suggest that when $\theta > 1$, the activity of the network persists much longer. Moreover, for $N$ mildly large and $\theta > 1$ not too close to 1, the empirical measure of the network appears to stabilize around some nontrivial distribution; this suggests that the decay time grows very fast with $N$, possibly exponentially. This behaviour can also be observed in [22, Section 7.2.1], although the condition at which the phase transition occurs is different, see Remark 3.4.

4 Proofs

In this section we provide detailed proofs for all of our results. By measuring neuron potentials in units of firing amplitude, we may set $\rho = 1$, and by measuring time in units of a neuron half life divided by log(2), we may set $\mu = 1$. To simplify notation, we will hence often set $\rho = \mu = 1$ without loss of generality.

We first obtain a simple formula for the expectation of a function of the process $(X_t)_{t \geq 0}$, which will be used several times. In general fix $\phi : \mathbb{R}^N \to \mathbb{R}$, and let $F_t = \phi(X_t)$. In the absence of firings, $F_t$ evolves according to $\frac{d}{dt} F_t = \sum_i \partial_t \phi(X_t) \frac{d}{dt} X_t = -\mu \sum_i X_i \partial_t \phi(X_t)$. If neuron $i$ fires and excites a set neurons $K$ at time instant $t$, then $F_t$ changes according to $F_t - F_t = \phi(X_t + \rho \sum_{j \in K} e_j - X_t e_i) - \phi(X_t)$. It follows that

$$dF_t = -\mu \sum_i X_i \partial_t \phi(X_t)dt$$

$$+ \sum_i \sum_{K \neq i} \int_{\mathbb{R}^+} 1(u \leq \gamma X_t) \left( \phi(X_t + \rho e_K - X_t e_i) - \phi(X_t) \right) N_{i,K}(dt, du).$$

Taking expectations, we find that

$$\frac{d}{dt} \mathbb{E}[F_t] = -\mu \sum_i \mathbb{E} \left[ X_i \partial_t \phi(X_t) \right]$$

$$+ \left( N - 1 \right)^{-1} \sum_{i} \sum_{K \neq i} \gamma \mathbb{E} \left[ X_t \left( \phi(X_t + \rho e_K - X_t e_i) - \phi(X_t) \right) \right]$$

(4.1)

4.1 Long-run behaviour of finite-neuron system (Theorem 3.1)

Proof of Theorem 3.1. Strong existence of solutions to (2.1) is obtained by truncating the firing rate and letting the truncation parameter approach $\infty$, as in the proof of Lemma 4.1. Uniqueness is standard.

We will next verify that $\|X_t\| = \sum_{i=1}^N X_t^i \to 0$ almost surely. We claim that $H_t = e^{-(\gamma/\mu)\|X_t\|}$ is a submartingale with respect to the filtration $(F_t)_{t \geq 0}$ generated by the marked point processes $(N_i)_{i=1}^N$, where $N_i := \sum_{K \neq i} N_{i,K}(dt, du, (i - 1, i))$. In the absence of firings, each neuron evolves according to $\frac{d}{dt} X_t^i = -\mu X_t^i$, so that $\frac{d}{dt} \|X_t\| = -\mu \|X_t\|$ and $\frac{d}{dt} H_t = -\gamma \mu H_t \frac{d}{dt} \|X_t\| = \gamma \|X_t\| H_t$. If neuron $i$ fires at time instant $t$, we see that $\|X_t\| = \|X_t\| + \rho \kappa - X_t^i$, and it follows that

$$H_t - H_t = \left( e^{-(\gamma/\mu)(\rho \kappa - X_t^i)} - 1 \right) H_t.$$
It follows that for any $0 \leq s \leq t$,

$$H_t - H_s = \int_s^t \gamma \| X_r \| H_r \, dr + \sum_i \int_s^t \int_{\mathbb{R}_+} 1(u \leq \gamma X_r^i) \left( e^{-((\gamma/\mu)(\rho \kappa - X_r^i) - 1)} H_r N_i(dr, du) \right),$$

from which we conclude that

$$H_t - H_s \geq \int_s^t \gamma \| X_r \| H_r \, dr - \sum_i \int_s^t \int_{\mathbb{R}_+} 1(u \leq \gamma X_r^i) H_r N_i(dr, du) \quad (4.2)$$

almost surely. By noting that $\sum_i \int_{\mathbb{R}_+} 1(u \leq \gamma X_r^i) du = \gamma \| X_r \|$, and recalling that the intensity measure of $N_i(dr, du)$ equals $dr, du$, we find that

$$\mathbb{E} \left[ \sum_i \int_s^t \int_{\mathbb{R}_+} 1(u \leq \gamma X_r^i) H_r N_i(dr, du) \mid F_s \right] = \mathbb{E} \left[ \int_s^t \gamma \| X_r \| H_r \, dr \mid F_s \right].$$

The conditional expectation of the right side of (4.2) given $F_s$ hence equals zero, and it follows that $\mathbb{E}[H_t - H_s \mid F_s] \geq 0$ almost surely. Hence $H_t$ is a bounded submartingale, and by Doob’s martingale convergence theorem, $H_t \rightarrow H_\infty$ almost surely, where $H_\infty \in [0, 1]$ is a random variable. Hence $\| X_t \| \rightarrow -\frac{1}{\gamma} \log H_\infty \in [0, \infty]$ almost surely.

Similarly, we find that

$$\| X_t \| - \| X_s \| = -\int_s^t \mu \| X_r \| \, dr + \sum_i \int_s^t \int_{\mathbb{R}_+} 1(u \leq \gamma X_r^i) (\rho \kappa - X_r^i) N_i(dr, du).$$

By taking expectations, we see that

$$\mathbb{E}[\| X_t \|] - \mathbb{E}[\| X_s \|] = -\int_s^t \mu \mathbb{E}[\| X_r \|] \, dr + \mathbb{E} \left[ \sum_i \int_s^t \gamma X_r^i (\rho \kappa - X_r^i) dr \right],$$

so that

$$\frac{d}{dt} \mathbb{E}[\| X_t \|] = (\gamma \rho \kappa - \mu) \mathbb{E}[\| X_t \|] - \gamma \mathbb{E} \left[ \sum_i (X_t^i)^2 \right].$$

By Jensen’s inequality, $\mathbb{E} \left[ \sum_i (X_t^i)^2 \right] \geq N^{-1} \mathbb{E} \left[ \| X_t \|^2 \right] \geq N^{-1} (\mathbb{E}[\| X_t \|])^2$, and it follows that $y_t = \mathbb{E}[\| X_t \|]$ satisfies

$$\frac{d}{dt} y_t \leq (\gamma \rho \kappa - \mu) y_t - \gamma N^{-1} y_t^2.$$
Assuming $\theta < 1$, we will now prove (3.2). By applying (4.1) to $\phi(x) = e^{-(\gamma/\mu)x}$, noting that $\left(\binom{N-1}{\kappa}\right)^{-1}\left(\binom{N-2}{\kappa-1}\right) = \frac{\kappa}{N-1}$, we find that
\[
\frac{d}{dt} \mathbb{E}\left[e^{-(\gamma/\mu)X_t^i}\right] = \gamma \mathbb{E}\left[X_t^i e^{-(\gamma/\mu)X_t^i}\right] + \gamma \mathbb{E}\left[X_t^i \left(1 - e^{-(\gamma/\mu)X_t^i}\right)\right] + \left(\binom{N-1}{\kappa}\right)^{-1} \sum\limits_{j \neq i} \sum\limits_{K \ni j} \gamma \mathbb{E}\left[X_t^j e^{-(\gamma/\mu)(X_t^i + \rho)} - e^{-(\gamma/\mu)X_t^i}\right] + \gamma \mathbb{E}\left[X_t^i\right] - \gamma \left(1 - e^{-\rho \gamma/\mu}\right) \frac{\kappa}{N-1} \sum\limits_{j \neq i} \mathbb{E}\left[X_t^j e^{-(\gamma/\mu)X_t^i}\right] 
\geq \gamma \mathbb{E}\left[X_t^i\right] - \gamma \left(1 - e^{-\rho \gamma/\mu}\right) \frac{\kappa}{N-1} \sum\limits_{j \neq i} \mathbb{E}\left[X_t^j\right] = \gamma (1 - \theta) \mathbb{E}\left[X_t^i\right],
\]
thanks to exchangeability. In light of $(\gamma/\mu)X_t^i \geq 1 - e^{-(\gamma/\mu)X_t^i}$, it follows that
\[
\frac{d}{dt} \mathbb{E}\left[e^{-(\gamma/\mu)X_t^i}\right] \geq \mu(1 - \theta) \left(1 - \mathbb{E}\left[e^{-(\gamma/\mu)X_t^i}\right]\right).
\]
Hence, $g_t = 1 - \mathbb{E}[e^{-(\gamma/\mu)X_t^i}]$ satisfies $\frac{d}{dt} g_t \leq -\mu(1 - \theta) g_t$. The desired inequality (3.2) now follows from Grönwall’s Lemma. \hfill \Box

4.2 Well-posedness of the mean field (Theorem 3.2)

Theorem 3.2 is proved by following the strategy of [22, Theorem 2]. For simplicity, in this section we assume $\rho = \mu = 1$. We first linearise the SDE by fixing a rate function for the firings: given a continuous function $r : \mathbb{R}_{+} \to \mathbb{R}_{+}$, consider the following SDE, with initial condition $Y_0^r = Z_0$:
\[
dY_t^r = -Y_t^r dt - Y_t^r \int_0^\infty 1(u \leq \gamma Y_t^r) \mathcal{N}(dt, du) + \int_0^\infty 1(u \leq \gamma r_t) \mathcal{M}(dt, du),
\]
where $\mathcal{N}$ and $\mathcal{M}$ are as in (2.6). Denote $\|r\|_T := \sup_{t \in [0,T]} |r_t|$.

**Lemma 4.1.** There exists a unique strong solution of (4.3).

**Proof of Lemma 4.1.** Uniqueness is classical. To prove existence, fix a time horizon $T > 0$, a continuous function $r : [0, T] \to \mathbb{R}_{+}$, a threshold $M > 0$, and consider the truncated SDE
\[
dY_t^{r,M} = -Y_t^{r,M} dt - Y_t^{r,M} \int_0^\infty 1\left(u \leq \min(\gamma Y_t^{r,M}, M)\right) \mathcal{N}(dt, du) + \int_0^\infty 1\left(u \leq \gamma r_t\right) \mathcal{M}(dt, du),
\]
with $Y_0^{r,M} = Z_0$. Since the rates are bounded, there is strong existence and uniqueness for (4.4). Consider the event $B_M = \{\gamma Y_t^{r,M} \leq M : \forall t \in [0, T]\}$, which increases with $M$, and notice that for $\tilde{M} \geq M$ we have $Y_t^{r,M} \equiv Y_t^{r,M}$ on $B_M$. Thus, we can define $Y_t^r = \lim_{M \to \infty} Y_t^{r,M}$ on the event $B_\infty = \bigcup_{M > 0} B_M$. It is easy to see that $(Y_t^r)$ solves (4.3) on $B_\infty$. It remains to show that
\( \mathbb{P}(B_\infty) = 1 \), for which it is enough to prove that \( \lim_{M \to \infty} \mathbb{P}(B_M^r) = 0 \). Indeed, since \( \|r\|_T < \infty \), it is not hard to obtain a bound like \( \mathbb{P}(B_M^r) \leq C_T/M \) for some constant \( C_T > 0 \) (for instance, study the solution of (4.3) without resets and exponential decay and compare with \( Y^r_M \)). This proves existence and uniqueness for (4.3) on \([0, T]\) for all \( T > 0 \), which can be easily extended to \([0, \infty)\).

Now consider a functional \( \mathcal{A} : C(\mathbb{R}_+, \mathbb{R}_+) \to C(\mathbb{R}_+, \mathbb{R}_+) \) defined by

\[
(Ar)_t = \mathbb{E}[Y^r_t] \quad \text{where (4.3) solves (4.3).} \tag{4.5}
\]

**Lemma 4.2.** The functional \( \mathcal{A} \) has a unique fixed point.

**Proof of Lemma 4.2.** Fix a time horizon \( T > 0 \), and let \( r, q \in C([0, T], \mathbb{R}_+) \). Let \( (Y^r_t), (Y^q_t) \) be the strong solutions to (4.3) with rate functions \( r \) and \( q \), respectively, with \( Y^r_0 = Y^q_0 = Z_0 \). Call \( h_t = \mathbb{E}[|Y^r_t - Y^q_t|] \), then:

\[
\partial_t h_t = -h_t - \gamma \mathbb{E}[|Y^r_t - Y^q_t|^2] + \gamma \kappa |r_t - q_t| \leq -h_t + \gamma \kappa |r_t - q_t|.
\]

Consequently,

\[
|(Ar)_t - (Aq)_t| \leq h_t \leq \gamma \kappa \int_0^t |r_s - q_s|e^{-(t-s)} ds \leq \gamma \kappa \int_0^t |r_s - q_s| ds.
\]

Iterating this inequality gives

\[
\|A^nr - A^nq\|_T \leq \frac{(\gamma \kappa T)^n}{n!}\|r - q\|_T
\]

for all \( n \in \mathbb{N} \). Thus, for \( n \) large enough, \( A^n \) is a contraction on \( C([0, T], \mathbb{R}_+) \), \( \| \cdot \|_T \), then \( \mathcal{A} \) has a unique fixed point in \( C([0, T], \mathbb{R}_+) \). The extension to \( C(\mathbb{R}_+, \mathbb{R}_+) \) is straightforward.

**Proof of Theorem 3.2.** Define \( Z_t = Y^r_t \), where \( (Y^r_t) \) is the unique strong solution to (4.3) with rate function \( r \in C(\mathbb{R}_+, \mathbb{R}_+) \) chosen as the the unique fixed point of the functional \( \mathcal{A} \) given by (4.5). By definition of \( \mathcal{A} \), we have \( \mathbb{E}[Z_t] = r_t \), thus \( (Z_t) \) solves (2.6).

Uniqueness is straightforward: let \( \bar{Z} \) and \( \check{Z} \) be strong solutions to (2.6) with \( Z_0 = \bar{Z}_0 \). If we define \( r_t := \mathbb{E}[Z_t] \) and \( \tilde{r}_t := \mathbb{E}[\bar{Z}_t] \), then \( \bar{Z} = Y^r \) and \( \check{Z} = Y^{\tilde{r}} \) are strong solutions to (4.3) and is immediate to prove that both \( r \) and \( \tilde{r} \) are fixed points for \( \mathcal{A} \), so \( r = \tilde{r} \) by Lemma 4.2, and hence \( \bar{Z} = \check{Z} \).

### 4.3 Phase transition in the mean field (Theorem 3.3)

**Proof of Theorem 3.3.** Lemma 4.6 below implies that \( \sup_t \mathbb{E}[Z^2_t] < \infty \) when we assume that \( \mathbb{E}[Z^2_0] < \infty \), thus \( (Z_t)_{t \geq 0} \) is uniformly integrable.

We now study phase transition. Let us analyse the evolution of \( h_t = \mathbb{E}[e^{-(\gamma/\mu)Z_t}] \). Observe that \( h_t \in [0, 1] \). In absence of firings, \( e^{-(\gamma/\mu)Z_t} \) evolves according to \( \frac{d}{dt} e^{-(\gamma/\mu)Z_t} = e^{-(\gamma/\mu)\frac{d}{dt} Z_t} = \gamma Z_t e^{-(\gamma/\mu)Z_t} \). In a reset, \( e^{-(\gamma/\mu)Z_t} \) increases by \( 1 - e^{-(\gamma/\mu)Z_t} \), and in an event of excitation, it increases by \( (e^{-\rho/\mu} - 1)e^{-(\gamma/\mu)Z_t} \). Therefore,

\[
d(e^{-(\gamma/\mu)Z_t}) = \gamma Z_t e^{-(\gamma/\mu)Z_t} dt + (1 - e^{-(\gamma/\mu)Z_t}) \int_{\mathbb{R}_+} 1(u \leq \gamma Z_t) \mathcal{N}(dt, du) + (e^{-\rho/\mu} - 1)e^{-(\gamma/\mu)Z_t} \int_{\mathbb{R}_+} 1(u \leq \gamma \mathbb{E}[Z_t]) \mathcal{M}(dt, du),
\]

14
where $\mathcal{N}(dt, du)$ and $\mathcal{M}(dt, du)$ are as in (2.6). By taking expectations, we find that the terms $E[\gamma Z_t e^{-(\gamma/\mu)Z_t}]$ cancel, thus obtaining

$$\frac{d}{dt} h_t = \gamma r_t - \gamma \theta r_t h_t,$$

(4.6)

where $r_t = E[Z_t]$, and we recall that $\theta = \kappa(1-e^{\rho\gamma/\mu})$. This differential equation is solved by

$$h_t = e^{-\gamma \theta R_t} h_0 + \left(1 - e^{-\gamma \theta R_t}\right) \frac{1}{\theta},$$

where $R_t = \int_0^t r_s ds$. Because $R_t$ is nondecreasing, it follows that $h_t$ converges to

$$h_\infty = e^{-\gamma \theta R_\infty} h_0 + \left(1 - e^{-\gamma \theta R_\infty}\right) \frac{1}{\theta},$$

(4.7)

where $R_\infty = \int_0^\infty r_s ds \in [0, \infty]$. Recall that $h_t \in [0, 1]$ for all $t \geq 0$, so $h_\infty \in [0, 1]$ as well. We now prove that $h_\infty = \min(1, \frac{1}{\theta})$. Let us first state some remarks:

1. $h_t$ and $h_\infty$ are convex combinations of $h_0 < 1$ (since $E[Z_0] > 0$) and $\frac{1}{\theta}$.

2. If $R_\infty = \infty$, then $h_\infty = \frac{1}{\theta}$. The converse is also true when $h_0 \neq \frac{1}{\theta}$ (which is the case if $\theta \leq 1$, because $h_0 < 1 \leq \frac{1}{\theta}$).

3. If $R_\infty < \infty$, then $h_\infty = 1$. This is because $R_\infty < \infty$ implies $\lim_{n \to \infty} E[Z_{t_n}] = 0$ for some sequence $t_n \to \infty$. Hence the inequality $e^{-x} \geq 1 - x$ gives $h_\infty = \lim_{n \to \infty} E[e^{-(\gamma/\mu)Z_{t_n}}] \geq \lim_{n \to \infty} E[1 - (\gamma/\mu)Z_{t_n}] = 1$, so $h_\infty = 1$.

Consequently,

- When $\theta < 1$, we must have that $R_\infty < \infty$, because otherwise Remark 2 would imply that $h_\infty = \frac{1}{\theta} > 1$. Remark 3 gives $h_\infty = 1$.

- When $\theta = 1$, then $h_\infty = 1$ regardless of the value of $R_\infty$, thanks to Remarks 2 and 3. But using the converse in Remark 2, we deduce that $R_\infty = \infty$.

- When $\theta > 1$, then $\frac{1}{\theta} < 1$ and Remark 1 implies that $h_\infty < 1$. So, by Remark 3, $R_\infty = \infty$, which in turn implies $h_\infty = \frac{1}{\theta}$ by Remark 2.

This proves that $h_\infty = \min(1, \frac{1}{\theta})$, along with $\int_0^\infty E[Z_t] dt < \infty$ when $\theta < 1$ and $\int_0^\infty E[Z_t] dt = \infty$ when $\theta = 1$. Thus, $\lim_{t \to \infty} E\left[e^{-(\gamma/\mu)Z_t}\right] = 1$ for $\theta \leq 1$, which means that $E\left[\omega(Z_t, 0)\right] \to 0$ for the metric $\omega(x, y) = 1 - e^{-(\gamma/\mu)|x-y|}$. This implies that $1 - e^{-(\gamma/\mu)Z_t} \to 0$ in probability, and hence also $Z_t \to 0$ in probability. Because $(Z_t)_{t \geq 0}$ is uniformly integrable, it follows that $E[Z_t] \to 0$.

For the case $\theta < 1$, we now prove that $\lim_{t \to \infty} Z_t = 0$ almost surely: because $\int_0^\infty r_t dt < \infty$, we deduce that only finitely many atoms $(t, u)$ of $\mathcal{M}$ satisfy $u \leq \gamma r_t$, which implies that the time of the last firing of $Z_t$ is finite almost surely; after that, $Z_t$ will converge to 0 due to the exponential decay or a final reset.

Still assuming that $\theta < 1$, we now prove (3.4): noting that $h_t \leq 1$ and $(\gamma/\mu)Z_t \geq 1 - e^{-(\gamma/\mu)Z_t}$ from (4.6) we obtain

$$\frac{d}{dt} h_t \geq (1 - \theta) \gamma r_t \geq (1 - \theta) \mu(1 - h_t).$$
Therefore, \( g_t = 1 - h_t \) satisfies
\[
\frac{d}{dt}g_t \leq -(1 - \theta)\mu g_t,
\]
from which it follows that \( g_t \leq g_0 e^{-(1-\theta)\mu t} \).
This confirms (3.4) and concludes the proof of the claims concerning \( \theta < 1 \) and \( \theta = 1 \).

Finally, we treat the case \( \theta > 1 \): using Remark 1, we see that
\[
\sup_{t \geq 0} h_t \leq \max \left( h_0, \frac{1}{\theta} \right) < 1.
\]
Since \((\gamma/\mu)E[Z_t] \geq 1 - h_t\), we deduce that \( \inf_{t \geq 0} E[Z_t] > 0 \). The proof is now complete. \( \square \)

**Proof of Corollary 3.6.** Define \( p_t = P(Z_t = 0) = E[1(Z_t = 0)] \). Clearly, in a reset event the indicator \( 1(Z_t = 0) \) increases by \( 1 - 1(Z_t = 0) \), in a firing it increases by \(-1(Z_t = 0)\), while the drift does not affect it. Consequently,
\[
\frac{d}{dt}p_t = E[(1 - 1(Z_t = 0))\gamma Z_t] - E[1(Z_t = 0)\gamma \kappa r_t] = \gamma r_t - \gamma \kappa r_t p_t.
\]
where \( r_t = E[Z_t] \). This differential equation is solved by
\[
p_t = \frac{1}{\kappa} + \left( p_0 - \frac{1}{\kappa} \right) e^{-\gamma \kappa \int_0^t r_s ds},
\]
which is exactly (3.5). Moreover, thanks to Theorem 3.3, we know that \( \int_0^\infty r_s ds = \infty \) when \( E[Z_0] > 0 \) and \( \theta \geq 1 \). Taking limits, we obtain \( \lim_{t \to \infty} p_t = \frac{1}{\kappa} \) in this case. \( \square \)

### 4.4 Convergence to the mean field (Theorem 3.7)

Recall that \( \kappa = 2 \), and for simplicity, in this section we assume that \( \mu = 1 \) and \( \rho = 1 \). Denote \( f_t = \text{Law}(Z_t) \), where \( Z_t \) is the unique solution of (2.6). Let \( N(dt, du) \) and \( M(dt, du, dz) \) be independent Poisson random measures on \( \mathbb{R}^2_+ \) and \( \mathbb{R}^3_+ \) with respective intensities \( dt \, du \) and \( 2dt \, du \, f_t(dz) \), both independent of \( Z_0 \). Consider the SDE
\[
dZ_t = -Z_t \, dt - Z_t \, \int_0^\infty 1(u \leq \gamma Z_t) \, N(dt, du) + \int_0^\infty \int_0^\infty 1(u \leq \gamma z) \, N(dt, du, dz).
\]

**Lemma 4.3.** Equation (4.8) admits a unique strong solution which has the same law as the solution of (2.6).

**Proof.** (4.8) can be reduced to (2.6) in a strong sense: define \( M \) as the point measure on \( \mathbb{R}^2_+ \) with atoms \( (t, ur_t/z) \) for every atom \((t, u, z)\) of \( \hat{M} \) with \( z > 0 \), where \( r_t = \int_0^\infty zf_t(dz) \). It is straightforward to check that \( M(dt, du) \) is a Poisson random measure with intensity measure \( 2dt \, du \) and that (4.8) is exactly (2.6). The claim now follows from this reduction and Theorem 3.2. \( \square \)

Let us now rewrite the dynamics (2.1) as follows: Since the size of \( K \) is \( \kappa = 2 \), we note the poisson processes \( \langle N_K \rangle \) as \( \langle N_{ij} \rangle_{1 \leq i < j \leq N} \), where each \( N_{ij} \) has intensity measure:
\[
\frac{2}{(N-1)(N-2)} \, dt \, du \, 1([\xi] \notin \{i,j\}) \, d\xi
\]
and note \( N_{ij} = N_{ji} \) whenever \( i > j \). As stated before, observe that \( N_{ij} \) provides the randomness for all possible jumps where neurons \( i \) and \( j \) are excited simultaneously, and the particle being
reset corresponds to \([\xi] \in \{1, \ldots, N\}\setminus\{i, j\}\). With these building blocks, for \(i \in \{1, \ldots, N\}\) we define the Poisson random measures

\[
\mathcal{N}^i(dt, du) = \sum_{k,j \neq i} \mathcal{N}^{ij}(dt, du, (i-1, i])
\]

\[
\mathcal{M}^i(dt, du, d\xi) = \sum_{j \neq i} \mathcal{N}^{ij}(dt, du, d\xi).
\]

We note that \(\mathcal{N}^i\) encodes firing instants of neuron \(i\), and \(\mathcal{M}^i\) encodes instants in which neuron \(i\) is excited. Clearly, their respective intensities are

\[
dt du \quad \text{and} \quad 2 \frac{1([\xi] \neq i)}{N-1} dt du d\xi
\]

Thus, (2.1) can be expressed as

\[
dX^i_t = -X^i_t dt - \int_0^\infty 1(u \leq \gamma X^i_t) \mathcal{N}^i dt, du
\]

\[
+ \int_0^\infty \int_0^N 1(u \leq \gamma X^i_t) \mathcal{M}^i dt, du, d\xi,
\]

which closely resembles (4.8). The only difference is that in the third term of (4.8) the variable \(z\) is a sample from \(f_t = \text{Law}(Z_t)\), while in (4.9) this is replaced by \(X^{[\xi]}_t\), which selects randomly a neuron \(j \neq i\); that is, \(X^{[\xi]}_t\) is a sample from the (random) empirical measure \(\bar{X}^i_t\). Here, we employ notations

\[
\bar{x} = \frac{1}{N} \sum_{i=1}^N \delta_{x^i} \quad \text{and} \quad \bar{x}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}.
\]

Motivated by this, for each \(i\) we introduce a measurable mapping \((t, z, u) \mapsto F^i_t(z, u)\) from \(\mathbb{R}_+ \times \mathbb{R}_+^N \times (0, N]\) into \(\mathbb{R}_+\) such that

\[
\text{Law}(F^i_t(z, U)) = f_t
\]

and

\[
\mathbb{E} \left[ F^i_t(z, U) - z^{[U]} \right] = W_1(f_t, \bar{z}^i)
\]

for all \(t \in \mathbb{R}_+, z \in \mathbb{R}_+^N\), and for any random variable \(U\) uniformly distributed on \((0, N]\)\setminus\((i-1, i]\). By recalling that \(\text{Law}(z^{[U]}) = \bar{z}^i\), we see that the pair \((F^i_t(z, U), z^{[U]}\)) constitutes an optimal coupling of \(f_t = \text{Law}(Z_t)\) and the empirical distribution \(\bar{z}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{z^j}\). See [9, Lemma 3] for a proof of existence of such a mapping.

Now, we specify our coupling by defining \(Z_t = (Z^1_t, \ldots, Z^N_t)\) as a solution of

\[
dZ^i_t = -Z^i_t dt - Z^i_t \int_0^\infty 1(u \leq \gamma Z^i_t) \mathcal{N}^i dt, du
\]

\[
+ \int_0^\infty \int_{(0, N]} 1(u \leq \gamma F^i_t(Z_{r}, \xi)) \mathcal{M}^i dt, du, d\xi,
\]

started at \(Z_0^i = X^i_0\) (recall that \(X^1_0, \ldots, X^N_0\) are i.i.d. copies of \(Z_0\)). Here \(\mathcal{N}^i\) and \(\mathcal{M}^i\) are the same Poisson random measures appearing in (4.9). Thanks to Lemma 4.3, (4.11) admits a unique strong solution, and it is a nonlinear process. Notice however that these processes are not independent because they have simultaneous jumps. However, their dependence vanishes as \(N\) grows, according to the following result.
Lemma 4.4. There exists a constant $C > 0$ such that

$$
\mathbb{E} \left[ W_1(f_t, Z_t) \right] \leq \frac{C}{N^{1/3}}.
$$

The same is true with $Z_i$ in place of $Z_t$, for all $i \in \{1, \ldots, N\}$.

Proof of Lemma 4.4. By a direct application of [9, Lemma 7], the following holds for every $k \leq N$ and $t \geq 0$:

$$
\mathbb{E}[W_1(Z_t, f_t)] \leq \frac{kq^k}{N} \left( W_1(f_t^{\otimes k}, \mathcal{L}^k(Z_t)) + \varepsilon_k(f_t) \right) + \frac{l}{N} \left( W_1(f_t^{\otimes l}, \mathcal{L}^l(Z_t)) + \varepsilon_l(f_t) \right) \tag{4.12}
$$

with $N = kq + l$ and $\varepsilon_k(f_t) = \mathbb{E} \left[ W_1(\tilde{Y}_t, f_t) \right]$ where $\tilde{Y}_t = (Y_t^1, \ldots, Y_t^k) \sim f_t^{\otimes k}$ are mutually independent and $f_t$-distributed coordinates, and $\mathcal{L}^k$ denotes the law of the first $k$ coordinates of an exchangeable random vector. Thus, it suffices to find suitable bounds for the quantities $W_1(f_t^{\otimes k}, \mathcal{L}^k(Z_t))$ and $\varepsilon_k(f_t)$. We obtain such a bound for the first quantity with the aid of Lemma 4.5. This lemma states that there is a constant $C > 0$ such that

$$
W_1(f_t^{\otimes k}, \mathcal{L}^k(Z_t)) \leq \frac{Ck}{N}
$$

for any $k \in \{2, \ldots, N\}$.

Let us now bound the other term. Making use of [12, Thm. 1] for $p = 1, r \in (2, +\infty)$, we have:

$$
\varepsilon_k(f_t) = C(r)m_1^{1/r}(f_t) \left( k^{-1/2} + k^{r-1} \right) \leq C'(r)m_1^{1/r}(f_t)k^{-1/2}
$$

Here $m_r(f_t) := \int z^r f_t(dz)$ and in the last inequality we used that $r > 2$. Assume in (4.12) that $k \sim N^{2/3}$. It is not difficult to verify that $l/N = (1 + o(1))N^{-1/3}$. Hence, if these conditions hold, it is sufficient to verify that

$$
W_1(f_t^{\otimes k}, \mathcal{L}^k(Z_t)) + \varepsilon_k(f_t) \leq \frac{C}{N^{1/3}}
$$

If we assume that the $r$-th moment of $f_t$ is uniformly bounded, then

$$
W_1(f_t^{\otimes k}, \mathcal{L}^k(Z_t)) + \varepsilon_k(f_t) \leq \frac{Ck}{N} + C''(r)k^{-1/2}. \tag{4.13}
$$

Recall that $k = (1 + o(1))N^{2/3}$. Then the right hand side of (4.13) is essentially $CN^{-1/3} + C''(r)N^{-1/3}$, which is of order $N^{-1/3}$, so the result holds.

Finally, we note that $m_1(f_t), m_2(f_t), m_3(f_t)$ are uniformly bounded by Lemma 4.6. □

Lemma 4.5. There exists a constant $C > 0$ such that for every $k \in \{2, \ldots, N\}$, and $t \geq 0$,

$$
W_1(f_t^{\otimes k}, \mathcal{L}^k(Z_t)) \leq \frac{Ck}{N}.
$$

Proof of Lemma 4.5. Fix an integer $2 \leq k \leq N$. The idea is to define a second coupling $V_t = (V_t^1, \ldots, V_t^k)$ such that $\mathcal{L}(V_t) = f_t^{\otimes k}$, and then prove that this new coupling is close in Wasserstein distance to $\mathcal{L}^k(Z_t)$.
Recall that each neuron $i$ is associated to two random measures $\mathcal{N}^i$ and $\mathcal{M}^i$ on $\mathbb{R}^2_+ \times (0, N]$ respectively, where

$$\mathcal{N}^i(dt, du) = \sum_{k, j \neq i, j < k} \mathcal{N}_{jk}(dt, du, (i-1, i]),$$

$$\mathcal{M}^i(dt, du, d\xi) = \sum_{j \neq i} \mathcal{N}_{ij}(dt, du, d\xi),$$

with $(\mathcal{N}_{ij})_{i<j}$ being independent processes in $\mathbb{R}^2_+ \times (0, N]$ and $\mathcal{N}_{ij} = \mathcal{N}_{ji}$ if $j < i$.

Furthermore, observe that in every interaction event, there are exactly 3 neurons involved (say $i < j < k$), and the corresponding processes for each potential event involving these neurons are $\mathcal{N}_{ij}(dt, du, (k-1, k]), \mathcal{N}_{jk}(dt, du, (i-1, i]),$ and $\mathcal{N}_{ik}(dt, du, (j-1, j])$. Hence, let $(\tilde{\mathcal{N}}_{ij})_{i,j}$, $(\tilde{\tilde{\mathcal{N}}}_{ij})_{i,j}$, and $(\tilde{\tilde{\tilde{\mathcal{N}}}}_{ij})_{i,j}$ be independent copies of $(\mathcal{N}_{ij})_{i,j}$. Define

$$R^i(dt, du, d\xi) := \sum_{j > k} \left[ \mathbf{1}(\xi \in (k, N]) \mathcal{N}_{ij}(dt, du, d\xi) + \mathbf{1}(\xi \in (0, k]) \tilde{\mathcal{N}}_{ij}(dt, du, d\xi) \right]$$

$$+ \sum_{i < j \leq k} \left[ \mathbf{1}(\xi \in (k, N]) \tilde{\tilde{\mathcal{N!}}}_{ij}(dt, du, d\xi) + \mathbf{1}(\xi \in (0, k]) \tilde{\tilde{\tilde{\mathcal{N}}}}_{ij}(dt, du, d\xi) \right]$$

$$+ \sum_{j < i} \left[ \mathbf{1}(\xi \in (k, N]) \tilde{\tilde{\tilde{\mathcal{N}}}}_{ij}(dt, du, d\xi) + \mathbf{1}(\xi \in (0, k]) \tilde{\tilde{\tilde{\mathcal{N}}}}_{ij}(dt, du, d\xi) \right],$$

and let

$$S^i(dt, du) := \mathcal{N}^i(dt, du).$$

Consequently, defining

$$dV^i_t = -V^i_t dt - V^i_t \int_{\mathbb{R}^+} \mathbf{1}(u \leq \gamma V^i_t) S^i(dt, du)$$

$$+ \int_{\mathbb{R}^+} \int_{(0, N]} \mathbf{1}(u \leq \gamma F^i_t(Z_r, \xi)) R^i(dt, du, d\xi)$$

with $V^i_0 = Z^i_0$, we get the desired coupling. Let us now verify that it remains close to $\mathcal{L}^k(Z_t)$. Observe that

$$W_1(f^{\otimes k}, \mathcal{L}^k(Z_t)) \leq \mathbb{E} \left[ \frac{1}{k} \sum_{i=1}^k |Z^1_t - V^1_t| \right] = \mathbb{E}[||Z^1_t - V^1_t||] := h_t,$$

where we used interchangeability in the last equality. Then, it suffices to bound $h_t$. Let us denote $\mathcal{N}_{ij} := \mathcal{N}_{ij}(dr, du, d\xi)$. For $0 \leq s \leq t$:
\[
\begin{align*}
    h_t &= h_s - \int_s^t h_r \, dr + \mathbb{E} \int_s^t \int_0^{\infty} \Delta(Z^1_r, V^1_r) \, \mathcal{N}^1(\,dr, du) \\
    &\quad + \mathbb{E} \int_s^t \int_0^k \int_0^{\infty} \mathbb{1}(u \leq \gamma F^1_{l_r}(Z_{r-}, \xi)) \left( |Z^1_r + 1 - V^1_r| - |Z^1_r - V^1_r| \right) \sum_{j > k} \mathcal{N}_{1j} \\
    &\quad + \mathbb{E} \int_s^t \int_0^k \int_0^{N} \mathbb{1}(u \leq \gamma F^1_{l_r}(Z_{r-}, \xi)) \left( |Z^1_r + 1 - V^1_r| - |Z^1_r - V^1_r| \right) \sum_{1 < j \leq k} \mathcal{N}_{1j} \\
    &\quad + \mathbb{E} \int_s^t \int_0^k \int_0^{N} \int_0^{k} \mathbb{1}(u \leq \gamma F^1_{l_r}(Z_{r-}, \xi)) \left( |Z^1_r - V^1_r| - |Z^1_r - V^1_r| \right) \\
    &\quad \times \left[ \sum_{j > k} \mathbb{1}(\xi \in (0, k]) \tilde{N}_{1j} + \sum_{1 < j \leq k} \left( \mathbb{1}(\xi \in (k, N]) \tilde{N}_{1j} + \mathbb{1}(\xi \in (0, k]) \tilde{N}_{1j} \right) \right] \\
    &\leq h_s - \int_s^t h_r \, dr + \mathbb{E} \int_s^t \int_0^{\infty} \mathbb{1}(u \leq \gamma F^1_{l_r}(Z_{r-}, \xi)) \\
    &\quad \times \left[ \sum_{j > k} \mathbb{1}(\xi \in (0, k]) \left( \mathcal{N}_{1j} + \tilde{N}_{1j} \right) + \sum_{1 < j \leq k} \left( \mathcal{N}_{1j} + \mathbb{1}(\xi \in (0, k]) \tilde{N}_{1j} \right) \right].
\end{align*}
\]

where \( \mathbb{E} \int_s^t \int_0^{\infty} \Delta(Z^1_r, V^1_r) \, \mathcal{N}^1(\,dr, du) = -\mathbb{E} \int_s^t |Z^1_r - V^1_r|^2 \, dr \) which is negative.

Now, observe that \( \mathbb{E} \int_s^{t+1} |F^1_{l_r}(Z_{r-}, \xi)| \, d\xi = m_1(f_t) \), and denoting \( I(N) \) the intensity of a Poisson process \( \mathcal{N} \), observe that, for each Poisson process in (4.14):

- \( I \left( \mathbb{1}(\xi \in (0, k]) \sum_{j > k} \left( \mathcal{N}_{1j} + \tilde{N}_{1j} \right) \right) = 2 \frac{2(N-k)}{(N-1)(N-2)} \mathbb{1}(\xi \in (1, k]) \, dr \, du \, d\xi \)
- \( I \left( \sum_{1 < j \leq k} \mathcal{N}_{1j} \right) = \frac{2}{(N-1)(N-2)} \left( (k-2) \mathbb{1}(\xi \in (1, k]) + (k-1) \mathbb{1}(\xi \in (k, N]) \right) \, dr \, du \, d\xi \)
- \( I \left( \sum_{1 < j \leq k} \mathbb{1}(\xi \in (k, N]) \tilde{N}_{1j} \right) = \frac{2(k-1)}{(N-1)(N-2)} \mathbb{1}(\xi \in (1, k]) \, dr \, du \, d\xi \)
- \( I \left( \sum_{1 < j \leq k} \mathbb{1}(\xi \in (0, k]) \tilde{N}_{1j} \right) = \frac{2(k-2)}{(N-1)(N-2)} \mathbb{1}(\xi \in (1, k]) \, dr \, du \, d\xi \)

Then, the following inequality holds:

\[
\begin{align*}
    h_t - h_s \leq &\ - \int_s^t h_r \, dr + \frac{4(N-k)(k-1)\gamma}{(N-1)(N-2)} \int_s^t m_1(f_r) \, dr \\
    &\quad + \frac{2(k-2)(k-1)\gamma}{(N-1)(N-2)} \int_s^t m_1(f_r) \, dr \\
    &\quad + \frac{2(k-2)(N-k)\gamma}{(N-1)(N-2)} \int_s^t m_1(f_r) \, dr \\
    &\quad + \frac{2(k-2)(k-1)\gamma}{(N-1)(N-2)} \int_s^t m_1(f_r) \, dr \\
    = &\ - \int_s^t h_r \, dr + \gamma \frac{4[(N-k)(2k-3) + (k-2)(k-1)]}{(N-1)(N-2)} \int_s^t m_1(f_r) \, dr.
\end{align*}
\]
Similarly, observe that
\[ m_t \leq - \int_s^t h_r dr + 8 \gamma \frac{(k-1)}{(N-1)} \int_s^t m_1(f_r) dr. \]
Hence,
\[ \partial_t h_t \leq -h_t + 8 \gamma \frac{(k-1)}{(N-1)} m_1(f_t) \]
Because \( m_1(f_t) \leq M \) is uniformly bounded, Gronwall’s inequality now implies that
\[ h_t \leq 8M \gamma \frac{(k-1)}{(N-1)} (1 - e^{-t^4}), \]
and this gives us the desired result. \( \square \)

**Lemma 4.6.** Denote \( m_r(f_t) = \int z^r f_t(dz) = \mathbb{E}Z_t^r \) for \( f_t = \text{Law}(Z_t) \). Fix an integer \( 1 \leq r \leq 3 \). If \( m_r(f_0) < \infty \), then \( \sup_{t \geq 0} m_r(f_t) < \infty \).

**Proof.** Observe that \( \mathbb{E}\left( (Z_t + \rho)^r - Z_t^r \right) = \sum_{q=0}^{r-1} \binom{r}{q} \rho^{r-q} \mathbb{E}Z_t^q \). Dynkin’s formula applied to (2.6) implies that
\[
\frac{d}{dt} m_r(f_t) = -\mu r m_r(f_t) - \gamma m_{r+1}(f_t) + \kappa \gamma m_1(f_t) \sum_{q=0}^{r-1} \binom{r}{q} \rho^{r-q} m_q(f_t). \tag{4.15}
\]
For \( r = 1 \), (4.15) becomes \( \frac{d}{dt} m_1(f_t) = (\kappa \rho \gamma - \mu) m_1(f_t) - \gamma m_2(f_t) \). Therefore, \( m_2(f_t) \leq m_1(f_t)^2 \)
implies that
\[
\frac{d}{dt} m_1(f_t) \leq (\kappa \rho \gamma - \mu) m_1(f_t) - \gamma m_1(f_t)^2.
\]
This inequality shows that \( \frac{d}{dt} m_1(f_t) \leq 0 \) whenever \( m_1(f_t) \geq \kappa \rho - \mu / \gamma \). Therefore,
\[
c_1 = \sup_{t \geq 0} m_1(f_t) \leq \max\{m_1(f_0), \kappa \rho - \mu / \gamma\}.
\]
Next, (4.15) for \( r = 2 \) implies that
\[
\frac{d}{dt} m_2(f_t) = -2\mu m_2(f_t) - \gamma m_3(f_t) + \kappa \gamma m_1(f_t) \left( \rho^2 + 2 \rho m_1(f_t) \right)
\leq -2\mu m_2(f_t) + \kappa \gamma (c_1 + \rho)^3,
\]
which shows that \( \frac{d}{dt} m_2(f_t) \leq 0 \) whenever \( m_2(t) \geq (2\mu)^{-1} \kappa \gamma (c_1 + \rho)^3 \). Therefore,
\[
c_2 = \sup_{t \geq 0} m_2(f_t) \leq \max\{m_2(f_0), (2\mu)^{-1} \kappa \gamma (c_1 + \rho)^3\}.
\]
Similarly, (4.15) for \( r = 3 \) implies that
\[
\frac{d}{dt} m_3(f_t) = -3\mu m_3(f_t) - \gamma m_4(f_t) + \kappa \gamma m_1(f_t) \left( \rho^3 + 3 \rho^2 m_1(f_t) + 3 \rho m_2(f_t) \right)
\leq -3\mu m_3(f_t) + \kappa \gamma (c_1 \lor c_2 + \rho)^4,
\]
The same argument as for \( r = 2 \) now shows that
\[
c_3 = \sup_{t \geq 0} m_3(f_t) \leq \max\{m_3(f_0), (3\mu)^{-1} \kappa \gamma (c_1 \lor c_2 + \rho)^4\}. \square
\]
Proof of Theorem 3.7. Because \( \frac{1}{N} \sum_{i=1}^{N} \delta(z_i, x_i) \) constitutes a coupling of \( \bar{Z}_t \) and \( \bar{X}_t \), we see that

\[
W_1(f_t, \bar{X}_t) \leq W_1(f_t, \bar{Z}_t) + W_1(\bar{Z}_t, \bar{X}_t) \leq W_1(f_t, \bar{Z}_t) + \frac{1}{N} \sum_{i=1}^{N} |Z_i - X_i^i|,
\]

so due to exchangeability, we find that

\[
\mathbb{E} \left[ W_1(f_t, \bar{X}_t) \right] \leq \mathbb{E} \left[ W_1(f_t, \bar{Z}_t) \right] + h_t,
\]

where \( h_t := \mathbb{E} \left[ |Z_1 - X_1^1| \right] \). By Lemma 4.4, it suffices to estimate \( h_t \).

From (4.9) and (4.11) we see that

\[
\partial_t h_t \leq -h_t - \gamma \mathbb{E} \left[ (Z_t^1 - X_t^1)^2 \right] + 2\gamma \mathbb{E} \int_1^{\theta} \left| F_t^1(Z_t, \xi) - X_t^{[\xi]} \right| \frac{d\xi}{N - 1}.
\]

We just discard the second term, while in the third term we add and subtract \( Z_t^{[\xi]} \). Using (4.10) and exchangeability, we thus obtain

\[
\partial_t h_t \leq -h_t + 2\gamma \mathbb{E} \left[ W_1(f_t, Z_t^1) \right] + 2\gamma \mathbb{E} \left[ \frac{1}{N - 1} \sum_{i=2}^{N} |Z_i^i - X_i^i| \right]
\]

\[
\leq (\theta_c - 1) h_t + \frac{C}{N^{1/3}},
\]

thanks to Lemma 4.4. The claim follows using Gronwall’s lemma. \( \square \)

4.5 Persistence (Theorem 3.9)

Recall that \( \kappa = 2 \). We also assume that \( \rho = \mu = 1 \).

Proof of Theorem 3.9. We first treat the case \( \theta > 1 \), which implies that \( \theta_c > 1 \). Call \( \epsilon = \frac{1}{2} \inf_{\xi \geq 0} \mathbb{E}[Z_t] \); thanks to Theorem 3.3, we know that \( \epsilon > 0 \). Then, for any \( t \geq 0 \), we have

\[ 2\epsilon \leq \mathbb{E}[Z_t] = W_1(f_t, \delta_0) \leq \mathbb{E}[W_1(f_t, \bar{X}_t)] + \mathbb{E}[W_1(\bar{X}_t, \delta_0)] \leq \frac{C e^{(\theta_c-1)t}}{N^{1/3}} + \mathbb{E}[X_t], \]

where in the last step we used Theorem 3.7 and exchangeability. Clearly, \( C e^{(\theta_c-1)t}N^{-1/3} \leq \epsilon \) if and only if \( t \leq \tilde{c} + c \log N \) for \( c = 1/[3(\theta_c - 1)] > 0 \) and \( \tilde{c} = [1/(\theta_c - 1)] \log(\epsilon/C) \in \mathbb{R} \), from which the result follows.

We now treat the case \( \theta < 1 \). The convergence in the metric \( \omega(\cdot, \cdot) \) is exactly (3.2), so there’s nothing to prove. In the case \( 2\gamma = \theta_c < 1 \), from (4.9) it is easy to see that

\[
\frac{d}{dt} \mathbb{E}[X_t^i] = -\mathbb{E}[X_t^i] - \gamma \mathbb{E}[(X_t^i)^2] + \frac{2\gamma}{N - 1} \sum_{i=2}^{N} \mathbb{E}[X_i^i] \leq -(1 - \theta_c) \mathbb{E}[X_t^i],
\]

where in the last step we simply discarded the term \(-\gamma \mathbb{E}[(X_t^1)^2] \leq 0 \), and we used exchangeability. The result now follows from Gronwall’s lemma. \( \square \)
5 Possible further research

Several lines of work can be derived from these results. For instance, we are confident that the process $Z_t$ has a unique nontrivial limit $Z_{\infty}$ in the case $\theta > 1$, see Remark. This is supported by numerical simulations of the finite network for relatively large $N$: the empirical distribution of the process seems to stabilize when one lets the simulation run for a long time (even considering the fact that Theorem 3.1 implies that $X_t$ must eventually decay, see Remark 3.10). It would be very desirable to have a rigorous proof of this phenomenon.

An interesting extension would be to consider the firing range $\kappa$ and/or the magnitude of the excitatory impulse $\rho$ to be random. We believe that such a model should lead essentially to the same results (replacing $\kappa$ and $\rho$ with $E[\kappa]$ and $E[\rho]$, respectively) for sufficiently well behaved $\kappa$ and $\rho$; see Remark 3.8.

Finally, a natural generalization would be the introduction of a spatial structure on the model, for instance by considering a random graph representing the interconnections between neurons. This can possibly lead to interesting results, depending on the underlying structure of the graph.

References

[1] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *Journal of Mathematical Neuroscience*, 2(1):10, 2012.

[2] S. Benachour, B. Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes – i existence, invariant probability, propagation of chaos. *Stochastic Processes and their Applications*, 75(2):173 – 201, 1998.

[3] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probability Theory and Related Fields*, 137(3):541–593, 2007.

[4] A. N. Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological Cybernetics*, 95(1):1–19, 2006.

[5] A. N. Burkitt. A review of the integrate-and-fire neuron model: Ii. inhomogeneous synaptic input and network properties. *Biological Cybernetics*, 95(2):97–112, 2006.

[6] M. J. Cáceres, J. A. Carrillo, and B. Perthame. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *Journal of Mathematical Neuroscience*, 1(1):7, 2011.

[7] J. Carrillo, R. Gvalani, G. Pavliotis, and A. Schlichting. Long-time behaviour and phase transitions for the McKean–Vlasov equation on the torus. *Archive for Rational Mechanics and Analysis*, 235(1):635–690, 2020.

[8] E. Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, 12(2):199–213, 2001.

[9] R. Cortez and J. Fontbona. Quantitative propagation of chaos for generalized Kac particle systems. *Annals of Applied Probability*, 26(2):892–916, 2016.

23
[10] M. H. A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society B, 46*(3):353–388, 1984.

[11] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of McKean–Vlasov type. *Annals of Applied Probability, 25*(4):2096–2133, 2015.

[12] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields, 162*(3-4):707–738, 2015.

[13] E. M. Galloway, N. H. Woo, and B. Lu. Persistent neural activity in the prefrontal cortex: A mechanism by which BDNF regulates working memory? In W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, and S. Belleville, editors, *Essence of Memory*, volume 169 of *Progress in Brain Research*, pages 251–266. Elsevier, 2008.

[14] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology, 117*(4):500, 1952.

[15] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197, Berkeley and Los Angeles, 1956. University of California Press.

[16] L. Lapicque. Recherches quantitatives sur l’excitation electrique des nerfs traitée comme une polarization. *Journal de Physiologie et de Pathologie Generale*, 9:620–635, 1907.

[17] G. Major and D. Tank. Persistent neural activity: Prevalence and mechanisms. *Current Opinion in Neurobiology, 14*(6):675–684, 2004.

[18] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE’s. *Stochastic Processes and their Applications, 95*(1):109 – 132, 2001.

[19] S. Mischler and C. Mouhot. Kac’s program in kinetic theory. *Invent. Math.*, 193(1):1–147, 2013.

[20] J. W. Pillow, L. Paninski, V. J. Uzzell, E. P. Simoncelli, and E. Chichilnisky. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience, 25*(47):11003–11013, 2005.

[21] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. Chichilnisky, and E. P. Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature, 454*(7207):995–999, 2008.

[22] P. Robert and J. Touboul. On the dynamics of random neuronal networks. *Journal of Statistical Physics, 165*(3):545–584, 2016.

[23] E. T. Rolls and G. Deco. *The noisy brain: Stochastic dynamics as a principle of brain function*. Oxford University Press, 2010.

[24] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
[25] J. Touboul. Mean-field equations for stochastic firing-rate neural fields with delays: Derivation and noise-induced transitions. *Physica D: Nonlinear Phenomena*, 241(15):1223–1244, 2012.

[26] J. Touboul. Propagation of chaos in neural fields. *Annals of Applied Probability*, 24(3):1298–1328, 2014.

[27] J. Touboul. Spatially extended networks with singular multi-scale connectivity patterns. *Journal of Statistical Physics*, 156(3):546–573, 2014.

[28] J. Tugaut. Phase transitions of McKean–Vlasov processes in double-wells landscape. *Stochastics*, 86(2):257–284, 2014.

[29] J. Zylberberg and B. W. Strowbridge. Mechanisms of persistent activity in cortical circuits: Possible neural substrates for working memory. *Annual Review of Neuroscience*, 40(1):603–627, 2017.