The Journal of Mathematical Neuroscience
a SpringerOpen Journal

RESEARCH   **Open Access**

CrossMark

Finite-Size Effects on Traveling Wave Solutions to Neural Field Equations

Eva Lang[1,2] · **Wilhelm Stannat**[1,2]

Abstract Neural field equations are used to describe the spatio-temporal evolution of the activity in a network of synaptically coupled populations of neurons in the continuum limit. Their heuristic derivation involves two approximation steps. Under the assumption that each population in the network is large, the activity is described in terms of a population average. The discrete network is then approximated by a continuum. In this article we make the two approximation steps explicit. Extending a model by Bressloff and Newby, we describe the evolution of the activity in a discrete network of finite populations by a Markov chain. In order to determine finite-size effects—deviations from the mean-field limit due to the finite size of the populations in the network—we analyze the fluctuations of this Markov chain and set up an approximating system of diffusion processes. We show that a well-posed stochastic neural field equation with a noise term accounting for finite-size effects on traveling wave solutions is obtained as the strong continuum limit.

1 Introduction

The analysis of networks of neurons of growing size quickly becomes involved from a computational as well as from an analytic perspective when one tracks the spiking activity of every neuron in the network. It can therefore be useful to zoom out from the microscopic perspective and identify a population activity as an average over a certain group of neurons. In the heuristic derivation of such population models it is usually assumed that each of the populations in the network is infinite, such that, in

---

✉ W. Stannat
   stannat@math.tu-berlin.de

   E. Lang
   lang@math.tu-berlin.de

1   Institut für Mathematik, Technische Universität Berlin, Berlin, 10623, Germany

2   Bernstein Center for Computational Neuroscience, Berlin, 10115, Germany

the spirit of the law of large numbers, the description of the activity in each population reduces to a description of the mean. By considering a spatially extended network and letting the density of populations go to infinity, neural field equations are obtained as the continuum limit of these models. Here we consider the voltage-based neural field equation, which is a nonlocal evolution equation of the form

$$\frac{\partial}{\partial t}u(x,t) = -u(x,t) + w * F\big(u(\cdot,t)\big)(x), \quad x \in \mathbb{R}, t \geq 0, \tag{1}$$

where $u(x,t)$ describes the average membrane potential in the population at x at time t, $w : \mathbb{R} \to [0,\infty)$ is a kernel describing the strengths of the synaptic connections between the populations, and the gain function $F : \mathbb{R} \to [0,1]$ relates the potential to the activity in the population.

Neural field equations were first introduced by Amari [1] and Wilson and Cowan [2, 3] and have since been used extensively to study the spatio-temporal dynamics of the activity in coupled populations of neurons. While they are of a relatively simple form, they exhibit a variety of interesting spatio-temporal patterns. For an overview see for example [4–8]. In this article we will concentrate on traveling wave solutions, modeling the propagation of activity, which were proven to exist in [9].

The communication of neurons is subject to noise. It is therefore crucial to study stochastic versions of (1). While several sources of noise have been identified on the single neuron level, it is not clear how noise translates to the level of populations. Since neural field equations are derived as mean-field limits, the usual effects of noise should have averaged out on this level. However, the actual finite size of the populations leads to deviations from the mean-field behavior, suggesting finite-size effects as an intrinsic source of noise.

The (heuristic) derivation of neural field equations involves two approximation steps. First, the local dynamics in each population is reduced to a description of the mean activity. Second, the discrete network is approximated by a continuum. In this article we make these two approximation steps explicit. In order to describe deviations from the mean-field behavior for finite population sizes, we set up a Markov chain to describe the evolution of the activity in the finite network, extending a model by Bressloff and Newby [10]. The transition rates are chosen in such a way that we obtain the voltage-based neural network equation in the infinite-population limit. We analyze the fluctuations of the Markov chain in order to determine a stochastic correction term describing finite-size effects. In the case of fluctuations around traveling wave solutions, we set up an approximating system of diffusion processes and prove that a well-posed stochastic neural field equation is obtained in the continuum limit.

In order to derive corrections to the neural field equation accounting for finite-size effects, in [11], Bressloff (following Buice and Cowan [12]) sets up a continuous-time Markov chain describing the evolution of the activity in a finite network of populations of finite size N. The rates are chosen such that in the limit as $N \to \infty$ one obtains the usual activity-based network equation. He then carries out a van Kampen system size expansion of the associated master equation in the small parameter $1/N$ to derive deterministic corrections of the neural field equation in the form of coupled differential equations for the moments. To first order, the finite-size effects can be characterized as Gaussian fluctuations around the mean-field limit.

The model is considered from a mathematically rigorous perspective by Riedler and Buckwar in [13]. They make use of limit theorems for Hilbert-space valued piecewise deterministic Markov processes recently obtained in [14] as an extension of Kurtz's convergence theorems for jump Markov processes to the infinite-dimensional setting. They derive a law of large numbers and a central limit theorem for the Markov chain, realizing the double limit (number of neurons per population to infinity and continuum limit) at the same time. They formally set up a stochastic neural field equation, but the question of well-posedness is left open.

In [10], Bressloff and Newby extend the original approach of [11] by including synaptic dynamics and consider a Markov chain modeling the activity coupled to a piecewise deterministic process describing the synaptic current (see also Sect. 6.4 in [8] for a summary). In two different regimes, the model covers the case of Gaussian-like fluctuations around the mean-field limit as derived in [11], as well as a situation in which the activity has Poisson statistics as considered in [15].

Here we consider the question how finite-size effects can be included in the voltage-based neural field equation. We take up the approach of describing the dynamics of the activity in a finite-size network by a continuous-time Markov chain and motivate a choice of jump rates that will lead to the voltage-based network equation in the infinite-population limit. We derive a law of large numbers and a central limit theorem for the Markov chain. Instead of realizing the double limit as in [13], we split up the limiting procedure, which in particular allows us to insert further approximation steps. We follow the original approach by Kurtz to determine the limit of the fluctuations of the Markov chain. By linearizing the noise term around the traveling wave solution, we obtain an approximating system of diffusion processes. After introducing correlations between populations lying close together (cf. Sect. 5.1) we obtain a well-posed $L^2(\mathbb{R})$-valued stochastic evolution equation, with a noise term approximating finite-size effects on traveling waves, which we prove to be the strong continuum limit of the associated network. Further results concerning (stochastic) stability of travelling waves and a corresponding multiscale analysis for the stochastic neural field equations derived in this paper are contained in the references [16, 17], see also the PhD-thesis [18].

The article is structured as follows. We recall how population models can be derived heuristically in Sect. 2 and summarize the work on the description of finite-size effects that can be found in the literature so far. In Sect. 3 we introduce our Markov chain model for determining finite-size effects in the voltage-based neural field equation and prove a law of large numbers and a central limit theorem for our choice of jump rates. We use it to set up a diffusion approximation with a noise term accounting for finite-size effects on traveling wave solutions in Sect. 4. Finally, in Sect. 5, we prove that a well-posed stochastic neural field equation is obtained in the continuum limit.

1.1 Assumptions on the Parameters

As usual, we take the gain function $F : \mathbb{R} \to [0, 1]$ to be a sigmoid function, for example $F(x) = \frac{1}{1+e^{-\gamma(x-\kappa)}}$ for some $\gamma > 0$, $0 < \kappa < 1$. In particular we assume that

(i) $F \geq 0$, $\lim_{x \downarrow -\infty} F(x) = 0$, $\lim_{x \uparrow \infty} F(x) = 1$,

(ii) $F(x) - x$ has exactly three zeros $0 < a_1 < a < a_2 < 1$,

(iii) $F \in C^3$ and F', F'' and F''' are bounded,

(iv) $F' > 0$, $F'(a_1) < 1$, $F'(a_2) < 1$, $F'(a) > 1$.

Our assumptions on the synaptic kernel w are the following:

(i) $w \in C^1$,

(ii) $w(x, y) = w(|x - y|) \geq 0$ is nonnegative and homogeneous,

(iii) $\int_{-\infty}^{\infty} w(x)\, dx = 1$, $w_x \in L^1$.

Assumption (iv) on F implies that a_1 and a_2 are stable fixed points of (1), while a is an unstable fixed point. It has been shown in [9] that under these assumptions there exists a unique monotone traveling wave solution to (1) connecting the stable fixed points (and, in [19], that traveling wave solutions are necessarily monotone). That is, there exist a unique wave profile $\hat{u} : \mathbb{R} \to [0, 1]$ and a unique wave speed $c \in \mathbb{R}$ such that $u^{\mathrm{TW}}(t, x) = u_t^{\mathrm{TW}}(x) := \hat{u}(x - ct)$ is a solution to (1), i.e.

$$-c\partial_x u_t^{\mathrm{TW}}(x) = \partial_t u_t^{\mathrm{TW}}(x) = -u_t^{\mathrm{TW}}(x) + \int_{-\infty}^{\infty} w(x - y) F\left(u_t^{\mathrm{TW}}(y)\right) dy$$

$$= -u_t^{\mathrm{TW}}(x) + w * F\left(u_t^{\mathrm{TW}}\right)(x),$$

and

$$\lim_{x \to -\infty} \hat{u}(x) = a_1, \qquad \lim_{x \to \infty} \hat{u}(x) = a_2.$$

As also pointed out in [9], we can without loss of generality assume that $c \geq 0$. Note that $\hat{u}_x \in L^2(\mathbb{R})$ since in the case $c > 0$

$$\int \hat{u}_x^2(x)\, dx = \int \hat{u}_x(x) \frac{1}{c} \left(\hat{u}(x) - w * F(\hat{u})(x)\right) dx$$

$$\leq \frac{1}{c} \left(\|\hat{u}\|_\infty + \|F\|_\infty \int w(x)\, dx\right) \int \hat{u}_x(x)\, dx = \frac{1}{c}(a_2 + 1)(a_2 - a_1),$$

and in the case $c = 0$

$$\int \hat{u}_x^2(x)\, dx = \int \hat{u}_x(x) w_x * F(\hat{u})(x)\, dx \leq \|w_x\|_1 (a_2 - a_1).$$

2 Finite-Size Effects in Population Models

2.1 Population Models

In population models, or firing rate models, instead of tracking the spiking activity of every neuron in the network, neurons are grouped together and the activity is identified as a population average. We start by giving a heuristic derivation of population models, distinguishing as usual between an activity-based and a voltage-based regime.

We consider a population of N neurons. We say that a neuron is 'active' if it is in the process of firing an action potential such that its membrane potential is larger than some threshold value κ. If Δ is the width of an action potential, then a neuron is active at time t if it fired a spike in the time interval $(t - \Delta, t]$. We define the population activity at a given time t as the proportion of active neurons,

$$a^N(t) = \frac{\text{\# neurons that are active at time } t}{N} \in \left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\right\}. \qquad (2)$$

We assume that all neurons in the population are identical and receive the same input. If the neurons fire independently from each other, then, for a constant input current I,

$$a^N(t) \xrightarrow{N \to \infty} F(I),$$

where $F(I)$ is the probability that a neuron receiving constant stimulation I is active. In the infinite-population limit, the population activity is thus related to the input current via the function F, called the *gain function*. Sometimes one also defines F as a function of the potential u, assuming that the potential is proportional to the current as in Ohm's law. F is typically a nonlinear function. It is usually modeled as a sigmoid, for example

$$F(x) = \frac{1}{1 + e^{-\gamma(x-\kappa)}}$$

for some $\gamma > 0$ and some threshold $\kappa > 0$, imitating the threshold-like nature of spiking activity.

Sometimes a *firing rate* is considered instead of a probability. We define the population firing rate λ^N as

$$\lambda^{\delta,N}(t) = \frac{\text{\# spikes in the time interval } (t - \delta, t]}{\delta N}.$$

If $\delta = \Delta$, then $\Delta \lambda^{\delta,N}(t) = a^N(t)$. At constant potential u, $\lim_{\delta \to 0} \lim_{N \to \infty} \lambda^{\delta,N}(t) = \lambda(u)$, where $\lambda(u)$ is the single neuron firing rate. Note that $\lambda \leq \frac{1}{\Delta}$. The firing rate is related to the probability $F(u)$ via

$$F(u) \approx \lambda(u)\Delta.$$

If the stimulus varies in time, then the activity may track this stimulus with some delay such that

$$a(t + \tau_a) = F\big(u(t)\big)$$

for some time constant τ_a. Taylor expansion of the left-hand side gives an approximate description of the (infinite-population) activity in terms of the differential equation

$$\tau_a \dot{a}(t) = -a(t) + F\big(u(t)\big), \qquad (3)$$

to which we refer as the *rate equation*.

We now consider a network of P populations, each consisting of N neurons. We assume that each presynaptic spike in population j at time s causes a postsynaptic potential

$$h(t - s) = \frac{1}{N} w_{ij} \frac{1}{\tau_m} e^{-\frac{1}{\tau_m}(t-s)}$$

in population i at time t. Here the (w_{ij}) are weights characterizing the strength of the synaptic connections between populations i and j, and τ_m is the membrane time constant, describing how fast the membrane potential relaxes back to its resting value.

Under the assumption that all inputs add up linearly, the potential in population i at time t is given as

$$u_i^N(t) = \sum_{j=1}^{P} w_{ij} \int_{-\infty}^{t} \frac{1}{\tau_m} e^{-\frac{1}{\tau_m}(t-s)} a_j^N(s) \, ds,$$

or, in differential form,

$$\tau_m \dot{u}_i^N(t) = -u_i^N(t) + \sum_{j=1}^{P} w_{ij} a_j^N(t). \tag{4}$$

In the infinite-population limit we therefore obtain the coupled system

$$\tau_m \dot{u}_i(t) = -u_i(t) + \sum_{j=1}^{P} w_{ij} a_j(t),$$
$$\tau_a \dot{a}_i(t) = -a_i(t) + F\big(u_i(t)\big), \quad i = 1, \dots P. \tag{5}$$

The behavior of (5) depends on the two time constants, τ_m and τ_a. Two different regimes can be identified in which the model can be reduced to just one of the two variables, u or a.

2.1.1 Case 1: $\tau_m \gg \tau_a \to 0$

Setting $\tau_a = 0$ yields $a_i(t) = F(u_i(t))$ and thus (5) reduces to

$$\tau_m \dot{u}_i(t) = -u_i(t) + \sum_{j=1}^{P} w_{ij} F\big(u_j(t)\big), \tag{6}$$

which we will call the *voltage-based neural network equation*.

2.1.2 Case 2: $\tau_a \gg \tau_m \to 0$

Setting $\tau_m = 0$ yields $u_i(t) = \sum_{j=1}^{P} w_{ij} a_j(t)$ and thus (5) reduces in this case to

$$\tau_a \dot{a}_i(t) = -a_i(t) + F\left(\sum_{j=1}^{P} w_{ij} a_j(t)\right), \tag{7}$$

which we will call the *activity-based neural network equation*.

Throughout this paper we will be interested in the former regime of the voltage-based neural network equation, thereby assuming τ_a small. Our aim is to derive stochastic differential equations in a first step for the infinite-population limit $N \to \infty$ and subsequently for the continuum limit $P \to \infty$, describing the asymptotic fluctuations.

2.2 Finite-Size Effects

The rate equation discussed in the previous subsection depends crucially on the way, how activity is measured. Starting from the definition (2) for given $\delta > 0$, the value will depend on the relation between the value of δ and the width Δ of an action potential. In [10], a model for the evolution of the activity in a network of finite populations is set up, based on the definition

$$a_j^{\delta,N}(t) = \frac{\# \text{ spikes in } (t - \delta, t] \text{ in population } j}{\delta N}$$

of the activity in population j. Here, δ is a time window of variable size. If δ is chosen as the width of an action potential Δ, one obtains the original notion of the activity, $\Delta a^{\Delta,N} = a^N$. Note that the number of spikes in the time interval $(t - \delta, t]$ is limited by $n_{\max} := N \vee [\frac{\delta N}{\Delta}]$.

The dynamics of $a^{\delta,N}$ is modeled as a Markov chain with state space $\{0, \frac{1}{\delta N}, \ldots, \frac{n_{\max}}{\delta N}\}^P$ and jump rates

$$\begin{aligned}
q_a^N\left(x, x + \frac{1}{\delta N}e_i\right) &= \frac{1}{\tau_a}\delta N \lambda\big(u_i^N(t)\big) \quad \text{if } x(i) < \frac{n_{\max}}{\delta N}, \\
q_a^N\left(x, x - \frac{1}{\delta N}e_i\right) &= \frac{1}{\tau_a}\delta N x(i).
\end{aligned} \tag{8}$$

Here e_i denotes the ith unit vector, $\lambda(u)$ is the firing rate at potential u, related to the probability $F(u)$ via $\Delta \lambda(u) = F(u)$, and u^N evolves according to (4)

$$\dot{u}_i^N(t) = \frac{1}{\tau_m}\left(-u_i^N(t) + \sum_{j=1}^P w_{ij}a_j^N(t)\right).$$

The idea for these transition rates is that the activation rate should be proportional to $\lambda(u)$, while the inactivation rate should be proportional to the activity itself.

The following two regimes are then considered in [10].

2.2.1 Case 1: $\delta = 1$, $\tau_a \gg \tau_m \to 0$

In the first regime, the size of the time window δ is fixed, say $\delta = 1$. If $\tau_a \gg \tau_m \to 0$, then as in Sect. 2.1, $u_i^N(t) = \sum_{j=1}^P w_{ij}a_j^{\delta,N}(t)$. The description of the Markov chain can thus be closed in the variables $a_i^{\delta,N}$, leading to the model already considered in [11]. In the limit $N \to \infty$ one obtains the activity-based network equation

$$\tau_a \dot{a}_i^N(t) = -a_i^N(t) + \lambda\left(\sum_{j=1}^P w_{ij}a_j^N(t)\right). \tag{9}$$

By formally approximating to order $\frac{1}{N}$ in the associated master equation, they derive a stochastic correction to (9), leading to the diffusion approximation

$$da_i^{\delta,N}(t) \approx \frac{1}{\tau_a}\left(-a_i^{\delta,N}(t) + \lambda\left(w_{ij}\sum_{j=1}^{P} a_j^{\delta,N}(t)\right)\right)dt$$

$$+ \frac{1}{\sqrt{\tau_a N}}\left(a_i^{\delta,N}(t) + \lambda\left(\sum_{j=1}^{P} w_{ij}a_j^{\delta,N}(t)\right)\right)^{\frac{1}{2}}dB_j(t)$$

for independent Brownian motions B_j.

In [13], Riedler and Buckwar rigorously derive a law of large numbers and a central limit theorem for the sequence of Markov chains as N tends to infinity. Note that the nature of the jump rates is such that the process has to be 'forced' to stay in its natural domain $[0, \frac{n_{\max}}{N}]$ by setting the jump rate to 0 at the boundary. As they point out, this discontinuous behavior is difficult to deal with mathematically. They therefore have to slightly modify the model and allow the activity to be larger than $\frac{n_{\max}}{N}$. They embed the Markov chain into $L^2(D)$ for a bounded domain $D \subset \mathbb{R}^d$ and derive the LLN in $L^2(D)$ and the CLT in the Sobolev space $H^{-\alpha}(D)$ for some $\alpha > d$.

2.2.2 Case 2: $\delta = \frac{1}{N}$, $\tau_m \gg \tau_a$

In the second regime, the size of the time window δ goes to 0 as N goes to infinity such that $\delta N = 1$. In this case,

$$a_i^{\delta,N}(t) = \frac{\text{\# spikes in } (t-\delta, t] \text{ in pop. } i}{\delta N} \approx \frac{\lambda(u_i^N(t))\delta N}{\delta N} = \lambda(u_i^N(t)).$$

The corresponding Markov chain has state space $\{0, 1, 2, \ldots, n_{\max}\}^P$. Since $\tau_a \ll \tau_m$, changes in $u_i(t)$ are slow in comparison to changes of the activity, and for fixed voltage u, the Markov chain in fact has a stationary distribution that is approximately Poisson with rate $\lambda(u)$ (see e.g. [10]). This corresponds to the regime considered in [15].

Note that we cannot expect to have a deterministic limit for the activity as $N \to \infty$ in this case, so we do not have a deterministic limit for the voltage. We will therefore consider a third regime with $\delta > 0$ fixed in this paper and study its asymptotic behavior in the large population limit $N \to \infty$.

2.2.3 Case 3: $\delta = \Delta$, $\tau_m \gg \tau_a \to 0$

We can expect in this case to have a law of large numbers, i.e. a deterministic limit for the activity. Since $a_i(t) \approx F(u_i(t))$ we also obtain a law of large numbers for the voltage. In order to derive an appropriate diffusion approximation, we will have to reconsider the transition rates of the Markov chain (8).

Going back to our original definition of the activity let us fix the time window δ to be the length of an action potential Δ. We assume that the potential evolves slowly, $\tau_m \gg 0$. Speeding up time, we define

$$\tilde{u}_i^N(t) = u_i^N(t\tau_m).$$

Then

$$\tilde{u}_i^N(t) = \sum_{j=1}^{P} w_{ij} \int_{-\infty}^{t\tau_m} \frac{1}{\tau_m} e^{-\frac{1}{\tau_m}(t\tau_m - s)} a_j^N(s)\, ds$$

$$= \sum_{j=1}^{P} w_{ij} \int_{-\infty}^{t} e^{-(t-s)} a_j^N(s\tau_m)\, ds.$$

For some large n,

$$\tilde{u}_i^N(t) \approx \sum_{j=1}^{P} w_{ij} \sum_{k=-\infty}^{[tn]-1} e^{-(t-\frac{k}{n})} \int_{\frac{k}{n}}^{\frac{k+1}{n}} a_j^N(s\tau_m)\, ds.$$

The potentials $\tilde{u}_i^N$ therefore only depend on the time-averaged activities given for $\frac{k}{n} \le t < \frac{k+1}{n}$ as

$$\tilde{a}_i^N(t) = n \int_{\frac{k}{n}}^{\frac{k+1}{n}} a_i^N(s\tau_m)\, ds.$$

We have

$$\tilde{u}_i^N(t) \approx \sum_{j=1}^{P} w_{ij} \sum_{k=-\infty}^{[tn]-1} \frac{1}{n} e^{-(t-\frac{k}{n})} \tilde{a}_j^N\left(\frac{k}{n}\right) \approx \sum_{j=1}^{P} w_{ij} \int_{-\infty}^{t} e^{-(t-s)} \tilde{a}_j^N(s)\, ds.$$

Since the relaxation rate τ_a approaches zero, the time-averaged activity $\tilde{a}_i^N(t)$ approximates $F(\tilde{u}_i^N(t))$ and thus

$$\tilde{a}_i^N(t) \approx F\big(\tilde{u}_i^N(t)\big) \approx F\left(\sum_{j=1}^{P} w_{ij} \int_{-\infty}^{t} e^{-(t-s)} \tilde{a}_j^N(s)\, ds\right), \tag{10}$$

with equality in the limit $N \to \infty$. Differentiating (10) w.r.t. t therefore yields

$$\frac{d}{dt}\tilde{a}_i(t) = \frac{d}{dt} F\big(\tilde{u}_i(t)\big)$$

$$= F'\big(\tilde{u}_i(t)\big)\left(-\tilde{u}_i(t) + \sum_{j=1}^{P} w_{ij} F\big(\tilde{u}_j(t)\big)\right)$$

$$= F'\big(F^{-1}\big(\tilde{a}_i(t)\big)\big)\left(-F^{-1}\big(\tilde{a}_i(t)\big) + \sum_{j=1}^{P} w_{ij}\tilde{a}_j(t)\right). \tag{11}$$

If $N < \infty$, then the finite size of the populations causes deviations from (11). In order to determine these finite-size effects, in the next section we will set up a Markov chain $X^{P,N}$ to describe the evolution of the time-averaged activity $\tilde{a}^N$.

3 A Markov Chain Model for the Activity

We will model the time-averaged activity by a Markov chain $X^{P,N}$ on the state space $E^{P,N} = \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}^P$. To this end consider the following jump rates:

$$
\begin{aligned}
q^{P,N}\left(x, x + \frac{1}{N}e_i\right) &= NF'\left(F^{-1}(x_i)\right)\left(-F^{-1}(x_i) + \sum_{j=1}^{P} w_{ij}x_j\right)_+ , \\
q^{P,N}\left(x, x - \frac{1}{N}e_i\right) &= NF'\left(F^{-1}(x_i)\right)\left(-F^{-1}(x_i) + \sum_{j=1}^{P} w_{ij}x_j\right)_- ,
\end{aligned}
\tag{12}
$$

for $x \in E^{P,N}$ with $x_i \in \{\frac{1}{N}, \ldots, 1 - \frac{1}{N}\}$. Here, we used the notation $x_+ := x \vee 0$, $x_- := (-x) \vee 0$, and e_i denotes the ith unit vector.

Note that, for $N \in \mathbb{N}$ large, the interior $(E^{P,N})^\circ := \{\frac{1}{N}, \ldots, 1 - \frac{1}{N}\}^P$ will be invariant, hence the Markov chain, when started in $(E^{P,N})^\circ$ stays away from the boundary. Indeed, for large $N \in \mathbb{N}$ it follows that

$$
q^{P,N}\left(x, x + \frac{1}{N}e_i\right) = 0 \qquad \left(\text{resp. } q^{P,N}\left(x, x - \frac{1}{N}e_i\right) = 0\right)
$$

for $x \in (E^{P,N})^\circ$ with $x_i = 1 - \frac{1}{N}$ (resp. $x_i = \frac{1}{N}$), because $F^{-1}(1 - \frac{1}{N}) > 1 - \frac{1}{N} \geq \sum_{j=1}^{P} w_{ij}x_j$, hence $q(x, x + \frac{1}{N}e_i) = 0$, (resp. $F^{-1}(\frac{1}{N}) < \frac{1}{N} \leq \sum_{j=1}^{P} w_{ij}x_j$, hence $q(x, x - \frac{1}{N}e_i) = 0$) for large N, using the limiting behavior $\lim_{x \uparrow 1} F^{-1}(x) = \infty$ (resp. $\lim_{x \downarrow 0} F^{-1}(x) = -\infty$).

The idea behind the choice of our rates is the following: the time-averaged activity tends to jump up (down) if the potential in the population, which is approximately given by $F^{-1}(x_i)$, is lower (higher) than the input from the other populations, which is given by $\sum_{j=1}^{P} w_{ij}x_j$. The probability that the activity jumps down (up) when the potential is lower (higher) than the input is assumed to be negligible. The jump rates are proportional to the difference between the two quantities, scaled by the factor $F'(F^{-1}(x_i))$. They are therefore higher in the sensitive regime where $F' \gg 1$, that is, where small changes in the potential have large effects on the activity. If $a_i^N = F(u_i^N)$ in all populations i, then the system is in balance.

We will see in Proposition 1 below that the Markov chain converges to the solution of (11) for increasing population size $N \uparrow \infty$.

In [11] a different choice of jump rates was suggested in analogy to (8):

$$
\begin{aligned}
\tilde{q}\left(x, x + \frac{1}{N}e_i\right) &= NF'\left(F^{-1}(x_i)\right) \sum_{j=1}^{P} w_{ij}x_j, \\
\tilde{q}\left(x, x - \frac{1}{N}e_i\right) &= NF'\left(F^{-1}(x_i)\right) F^{-1}(x_i).
\end{aligned}
$$

Also this choice leads to (11) in the limit. With these rates, however, the jump rates are high in regions where the activity is high. Since, as explained above, one should

think of the Markov chain as governing a slowly varying time-averaged activity, (12) seems like a more natural choice.

The generator of $Q^{P,N}$ of $X^{P,N}$ is given for bounded measurable $f : E^{P,N} \to \mathbb{R}$ by

$$
\begin{aligned}
Q^{P,N} f(x) = N \sum_{k=1}^{P} F'\big(F^{-1}(x_k)\big) \Bigg(& \bigg(-F^{-1}(x_k) + \sum_{j=1}^{P} w_{kj} x_j \bigg)_+ \\
& \times \bigg(f\Big(x + \tfrac{1}{N} e_k\Big) - f(x) \bigg) \\
& + \bigg(-F^{-1}(x_k) + \sum_{j=1}^{P} w_{kj} x_j \bigg)_- \bigg(f\Big(x - \tfrac{1}{N} e_k\Big) - f(x) \bigg) \Bigg).
\end{aligned}
$$

Let N_0 be such that for $N \geq N_0$ the interior $(E^{P,N})^\circ$ is an invariant set for $X^{P,N}$.

The jump rates out of the interval $[\frac{1}{N_0}, 1 - \frac{1}{N_0}]$ are 0.

Proposition 1 *Let X^P be the (deterministic) Feller process on $[1/N_0, 1 - 1/N_0]^P$ with generator*

$$
L^P f(x) = \sum_{k=1}^{P} F'\big(F^{-1}(x_k)\big) \bigg(-F^{-1}(x_k) + \sum_{j=1}^{P} w_{kj} x_j \bigg) \partial_k f(x).
$$

If $X^{P,N}(0) \xrightarrow[N\to\infty]{d} X^P(0)$, then $X^{P,N} \xrightarrow[N\to\infty]{d} X^P$ on the space of càdlàg functions $D([0, \infty), [1/N_0, 1 - 1/N_0]^P)$ with the Skorohod topology (where $\xrightarrow{d}$ denotes convergence in distribution).

Proof By a standard theorem on the convergence of Feller processes (cf. [20], Thm. 19.25) it is enough to prove that for $f \in \mathcal{C}^\infty([1/N_0, 1 - 1/N_0]^P)$ there exist bounded measurable f_N such that $\|f_N - f\|_\infty \xrightarrow{N\to\infty} 0$ and $\|Q^{P,N} f_N - L^P f\|_\infty \xrightarrow{N\to\infty} 0$.

Let thus $f \in \mathcal{C}^\infty([1/N_0, 1 - 1/N_0]^P)$ and set $f_N(x) = f(\frac{[x_1 N]}{N}, \dots, \frac{[x_P N]}{N})$. Then it is easy to see that

$$
Q^{P,N} f_N(x) \xrightarrow{N\to\infty} L^P f(x)
$$

uniformly in x. $\square$

4 Diffusion Approximation

We are now going to approximate $X^{P,N}$ by a diffusion process. To this end, we follow the standard approach due to Kurtz and derive a central limit theorem for the fluctuations of $X^{P,N}$. This will give us a candidate for a stochastic correction term to (11).

4.1 A Central Limit Theorem

The general theory of continuous-time Markov chains implies the following semi-martingale decomposition:

$$X_k^{P,N}(t) = X_k^{P,N}(0) + \int_0^t Q^{P,N} \pi_k(X^{P,N}(s)) \, ds + M_k^{P,N}(t),$$

where $\pi_k : (0, 1)^P \to (0, 1)$, $x \mapsto x_k$, is the projection onto the kth coordinate, $Q^{P,N} \pi_k(X^{P,N}(s)) = \sum_{y \in E^{P,N}} q^{P,N}(X^{P,N}(s), y) y_k$ is the (signed) kernel $Q^{P,N}$ applied to the function π_k evaluated at the state $X^{P,N}(s)$, and

$$M_k^{P,N}(t) := X_k^{P,N}(t) - X_k^{P,N}(0) - \int_0^t Q^{P,N} \pi_k(X^{P,N}(s)) \, ds$$

is a zero mean square-integrable martingale (with right-continuous sample paths having left limits in $t > 0$). Moreover,

$$\left(M_k^{P,N}(t)\right)^2 - \int_0^t \sum_{y \in E^{P,N}} q^{P,N}(X_k^{P,N}(s), y)(X_k^{P,N}(s) - y_k)^2 \, ds$$

is a martingale too (see [21]). The general theory of stochastic processes states that for any square-integrable right-continuous martingale M_t, $t \geq 0$, having left limits in $t > 0$, there exists a unique increasing previsible process $\langle M \rangle_t$ starting at 0, such that $M_t^2 - \langle M \rangle_t$, $t \geq 0$, is a martingale too. We can therefore conclude that

$$\langle M^{P,N} \rangle_t = \int_0^t \sum_{y \in E^{P,N}} q^{P,N}(X_k^{P,N}(s), y)(X_k^{P,N}(s) - y_k)^2 \, ds.$$

The bracket process will be crucial to determine the limit of the fluctuations as seen in the next proposition.

Proposition 2

$$\left(\sqrt{N} M_k^{P,N}\right)$$

$$\xrightarrow[N \to \infty]{d} \left(\int_0^t \left(F'(F^{-1}(X_k^P(s))) \left| -F^{-1}(X_k^P(s)) + \sum_{j=1}^P w_{kj} X_j^P(s) \right| \right)^{\frac{1}{2}} dB_k(s) \right)$$

on $\mathcal{D}([0, \infty), \mathbb{R}^P)$, where B is a P-dimensional standard Brownian motion, and X^P is the Feller process from Proposition 1.

Proof Using

$$\langle M_k^{P,N} \rangle_t = \int_0^t \sum_{y \in E^{P,N}} q^{P,N}(X^{P,N}(s), y)(X_k^{P,N}(s) - y_k)^2 \, ds$$

$$= \frac{1}{N} \int_0^t F'(F^{-1}(X_k^{P,N}(s))) \left(\left(-F^{-1}(X_k^{P,N}(s)) + \sum_{j=1}^P w_{kj} X_j^{P,N}(s) \right)_+$$

$$+ \left(-F^{-1}\big(X_k^{P,N}(s)\big) + \sum_{j=1}^{P} w_{kj} X_j^{P,N}(s) \right)_{-} \right) ds$$

$$= \frac{1}{N} \int_0^t F'\big(F^{-1}\big(X_k^{P,N}(s)\big)\big) \left| -F^{-1}\big(X_k^{P,N}(s)\big) + \sum_{j=1}^{P} w_{kj} X_j^{P,N}(s) \right| ds.$$

Thus,

$$\big\langle \sqrt{N} M_k^{P,N} \big\rangle_t \xrightarrow{N \to \infty} \int_0^t F'\big(F^{-1}\big(X_k^P(s)\big)\big) \left| -F^{-1}\big(X_k^P(s)\big) + \sum_{j=1}^{P} w_{kj} X_j^P(s) \right| ds$$

in probability. For $k \neq l$,

$$\big\langle M_k^{P,N}, M_l^{P,N} \big\rangle_t = \int_0^t \sum_y q^{P,N}\big(X^{P,N}(s), y\big)\big(X_k^{P,N}(s) - y_k\big)\big(X_l^{P,N}(s) - y_l\big) ds = 0,$$

since for y with $q^{P,N}(X^{P,N}(s), y) > 0$ at least one of $X_k^{P,N}(s) - y_k$ and $X_l^{P,N}(s) - y_l$ is always 0. Now

$$E\left(\sup_t \sqrt{N} \big\| M^{P,N}(t) - M^{P,N}(t-) \big\|_2 \right) \leq \frac{1}{\sqrt{N}} \xrightarrow{N \to \infty} 0,$$

and the statement follows by the martingale central limit theorem; see for example Theorem 1.4, Chap. 7 in [22]. □

This suggests to approximate $X^{P,N}$ by the system of coupled diffusion processes

$$da_k^{P,N}(t)$$

$$= F'\big(F^{-1}\big(a_k^{P,N}(t)\big)\big) \left(-F^{-1}\big(a_k^{P,N}(t)\big) + \sum_{j=1}^{P} w_{kj} a_j^{P,N}(t) \right) dt$$

$$+ \frac{1}{\sqrt{N}} \left(F'\big(F^{-1}\big(a_k^{P,N}(t)\big)\big) \left| -F^{-1}\big(a_k^{P,N}(t)\big) + \sum_{j=1}^{P} w_{kj} a_j^{P,N}(t) \right| \right)^{\frac{1}{2}} dB_k(t),$$

$1 \leq k \leq P$.

Using Itô's formula, we formally obtain an approximation for $u_k^{P,N} := F^{-1}(a_k^{P,N})$,

$$du_k^{P,N}(t) = \left(-u_k^{P,N}(t) + \sum_{j=1}^{P} w_{kj} F\big(u_j^{P,N}(t)\big) \right.$$

$$\left. - \frac{1}{2N} \frac{F''(u_k^{P,N}(t))}{F'(u_k^{P,N}(t))^2} \left| -u_k^{P,N}(t) + \sum_{j=1}^{P} w_{kj} F\big(u_j^{P,N}(t)\big) \right| \right) dt$$

$$+ \frac{1}{\sqrt{N}} \left(\frac{|-u_k^{P,N}(t) + \sum_{j=1}^{P} w_{kj} F(u_j^{P,N}(t))|}{F'(u_k^{P,N}(t))} \right)^{\frac{1}{2}} dB_k(t). \qquad (13)$$

Since the square root function is not Lipschitz continuous near 0, we cannot apply standard existence theorems to obtain a solution to (13) with the full multiplicative noise term. Instead we will linearize around a deterministic solution to the neural field equation and approximate to a certain order of $\frac{1}{\sqrt{N}}$.

4.2 Fluctuations around the Traveling Wave

Let $\bar{u}$ be a solution to the neural field equation (1). To determine the finite-size effects on $\bar{u}$, we consider a spatially extended network, that is, we look at populations distributed over an interval $[-L, L] \subset \mathbb{R}$ and use the stochastic integral derived in Proposition 2 to describe the local fluctuations on this interval.

Let $m \in \mathbb{N}$ be the density of populations on $[-L, L]$ and consider $P = 2mL$ populations located at $\frac{k}{m}$, $k \in \{-mL, -mL+1, \ldots, mL-1\}$. We choose the weights w_{kl} as a discretization of the integral kernel $w : \mathbb{R} \to [0, \infty)$,

$$w_{kl}^m = \int_{\frac{l}{m}}^{\frac{l+1}{m}} w\left(\frac{k}{m} - y\right) dy, \quad -mL \leq k, l \leq mL - 1. \tag{14}$$

Since we think of the network as describing only a section of the actual domain $\mathbb{R}$, we add to each population an input $F(\bar{u}_t(-L))$ and $F(\bar{u}_t(L))$, respectively, at the boundaries with corresponding weights

$$
\begin{aligned}
w_k^{m,+} &= \int_L^\infty w\left(\frac{k}{m} - y\right) dy, \\
w_k^{m,-} &= \int_{-\infty}^{-L} w\left(\frac{k}{m} - y\right) dy.
\end{aligned}
\tag{15}
$$

Fix a population size $N \in \mathbb{N}$. Set $\bar{u}_k(t) = \bar{u}(\frac{k}{m}, t)$ and for $u \in \mathbb{R}^P$,

$$\hat{b}_k^m(t, u) = -u_k(t) + \sum_{l=-mL}^{mL-1} w_{kl}^m F\big(u_l(t)\big) + w_k^{m,+} F\big(\bar{u}(L, t)\big) + w_k^{m,-} F\big(\bar{u}(-L, t)\big).$$

We write

$$u_k = \bar{u}_k + v_k \tag{16}$$

and assume that v_k is of order $1/\sqrt{N}$. Linearizing (13) around $(\bar{u}_k)$ we obtain the approximation

$$du_k(t) = \hat{b}_k^m(t, u)\, dt + \frac{1}{\sqrt{N F'(\bar{u}_k(t))}} \big|\hat{b}_k^m(t, \bar{u})\big|^{\frac{1}{2}}\, dB_k(t)$$

to order $1/\sqrt{N}$.

Note that $\hat{b}_k^m(t, \bar{u}) \approx \partial_t \bar{u}_k(t) = 0$ for a stationary solution $\bar{u}$, with equality if $\bar{u}$ is constant. The finite-size effects are hence of smaller order. Since the square root function is not differentiable at 0 we cannot expand further.

However, the situation is different if we linearize around a moving pattern. We consider the traveling wave solution $u_t^{\mathrm{TW}}(x) = \hat{u}(x - ct)$ to (1) and we assume without loss of generality that $c > 0$. Then $\hat{b}_k^m(t, \bar{u}) \approx \partial_t \bar{u}_k(t) = -c \partial_x u_t^{\mathrm{TW}} < 0$. This monotonicity property allows us to approximate to order $1/N$ in (13). Indeed, note that since $\hat{u}$ and F are increasing,

$$
\begin{aligned}
-\hat{b}_k^m\big(t, u_t^{\mathrm{TW}}\big) \\
= u_t^{\mathrm{TW}}\left(\frac{k}{m}\right) - \sum_l w_{kl}^m F\left(u_t^{\mathrm{TW}}\left(\frac{l}{m}\right)\right) \\
\quad - w_k^{m,+} F\big(u_t^{\mathrm{TW}}(L)\big) - w_k^{m,-} F\big(u_t^{\mathrm{TW}}(-L)\big) \\
\geq u_t^{\mathrm{TW}}\left(\frac{k}{m}\right) - \int_{-L}^{\infty} w\left(\frac{k}{m} - y\right) F\big(u_t^{\mathrm{TW}}\big)(y)\,dy \\
\quad - \int_{-\infty}^{-L} w\left(\frac{k}{m} - y\right) F\big(u_t^{\mathrm{TW}}(-L)\big)\,dy \\
= c\hat{u}_x\left(\frac{k}{m} - ct\right) - \int_{-\infty}^{-L} w\left(\frac{k}{m} - y\right)\big(F\big(u_t^{\mathrm{TW}}(-L)\big) - F\big(u_t^{\mathrm{TW}}(y)\big)\big)\,dy \\
\geq c\hat{u}_x\left(\frac{k}{m} - ct\right) - \big(F\big(u_t^{\mathrm{TW}}(-L)\big) - F(a_1)\big) \\
\xrightarrow{L \to \infty} c\hat{u}_x\left(\frac{k}{m} - ct\right) > 0.
\end{aligned}
\tag{17}
$$

So for L large enough, $-\hat{b}_k^m(t, u_t^{\mathrm{TW}}) > 0$ and we have, using Taylor's formula and (16),

$$
\begin{aligned}
\left(\frac{|\hat{b}_k^m(t, u)|}{F'(u_k(t))}\right)^{\frac{1}{2}} = \left(\frac{-\hat{b}_k^m(t, u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}}(\frac{k}{m}))}\right)^{\frac{1}{2}} + \frac{1}{2\sqrt{-\hat{b}_k^m(t, u_t^{\mathrm{TW}}) F'(u_t^{\mathrm{TW}}(\frac{k}{m}))}} \\
\times \left(\frac{F''(u_t^{\mathrm{TW}}(\frac{k}{m}))}{F'(u_t^{\mathrm{TW}}(\frac{k}{m}))} \hat{b}_k^m\big(t, u_t^{\mathrm{TW}}\big) v_k(t)\right. \\
\left. + v_k(t) - \sum_l w_{kl} F'\left(u_t^{\mathrm{TW}}\left(\frac{l}{m}\right)\right) v_l(t)\right) + O\left(\frac{1}{N}\right).
\end{aligned}
$$

As a possible diffusion approximation in the case of traveling wave solutions we therefore obtain the system of stochastic differential equations

$$
\begin{aligned}
du_k(t) = \left(\hat{b}_k^m(t, u) + \frac{1}{2N} \frac{F''(u_t^{\mathrm{TW}}(\frac{k}{m}))}{F'(u_t^{\mathrm{TW}}(\frac{k}{m}))^2} \hat{b}_k^m\big(t, u_t^{\mathrm{TW}}\big)\right) dt \\
+ \frac{1}{\sqrt{N}}\left[\left(\frac{-\hat{b}_k^m(t, u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}}(\frac{k}{m}))}\right)^{\frac{1}{2}} + \frac{1}{2\sqrt{-\hat{b}_k^m(t, u_t^{\mathrm{TW}}) F'(u_t^{\mathrm{TW}}(\frac{k}{m}))}}\right.
\end{aligned}
$$

$$\times \left(\frac{F''(u_t^{\mathrm{TW}}(\frac{k}{m}))}{F'(u_t^{\mathrm{TW}}(\frac{k}{m}))} \hat{b}_k^m(t, u_t^{\mathrm{TW}}) v_k(t) \right.$$

$$\left. + v_k(t) - \sum_l w_{kl} F'\left(u_t^{\mathrm{TW}}\left(\frac{l}{m}\right)\right) v_l(t) \right) \Bigg] dB_k(t), \tag{18}$$

for which there exists a unique solution as we will see in the next section.

5 The Continuum Limit

In this section we take the continuum limit of the network of diffusions (18), that is, we let the size of the domain and the density of populations go to infinity in order to obtain a stochastic neural field equation with a noise term describing the fluctuations around the deterministic traveling wave solution due to finite-size effects.

We thus have to deal with functions that 'look almost like the wave' and choose to work in the space $\mathcal{S} := \{u : \mathbb{R} \to \mathbb{R} : u - \hat{u} \in L^2\}$. Note that since for $u_1, u_2 \in \mathcal{S}$, $\|u_1 - u_2\| < \infty$, the L^2-norm induces a topology on $\mathcal{S}$.

5.1 A Word on Correlations

Recall the definition of the Markov chain introduced in Sect. 3. Note that as long as we allow only single jumps in the evolution, meaning that there will not be any jumps in the activity in two populations at the same time, the martingales associated with any two populations will be uncorrelated, yielding independent driving Brownian motions in the diffusion limit (cf. Proposition 2).

This only makes sense for populations that are clearly distinguishable. In order to determine the fluctuations around traveling wave solutions, we consider spatially extended networks of populations. The population located at $x \in \mathbb{R}$ is to be understood as the ensemble of all neurons in the ε-neighborhood $(x - \varepsilon, x + \varepsilon)$ of x for some $\varepsilon > 0$. If we consider two populations located at $x, y \in \mathbb{R}$ with $|x - y| < 2\varepsilon$, then they will overlap. Consequently, simultaneous jumps will occur, leading to correlations between the driving Brownian motions.

Thus the Markov chain model (and the associated diffusion approximation) is only appropriate as long as the distance between the individual populations is large enough. When taking the continuum limit, we therefore adapt the model by introducing correlations between the driving Brownian motions of populations lying close together.

5.2 The Stochastic Neural Field Equation

We start by defining the limiting object. For $u \in \mathcal{S}$ and $t \in [0, T]$ set

$$b(t, u)(x) = -u(x) + \int_{-\infty}^{\infty} w(x - y) F(u(y)) \, dy = -u(x) + w * F(u)(x).$$

Let $\mathcal{W}^Q$ be a (cylindrical) Q-Wiener process on L^2 with covariance operator $\sqrt{Q}$ given as $\sqrt{Q}h(x) = \int_{-\infty}^{\infty} q(x, y)h(y) \, dy$ for some symmetric kernel $q(x, y)$ with

$q(x, \cdot) \in L^2 \cap L^1$ for all $x \in \mathbb{R}$ and $\sup_{x \in \mathbb{R}}(\|q(x, \cdot)\| + \|q(x, \cdot)\|_1) < \infty$. (Details on the theory of Q-Wiener processes can be found in [23, 24].) We assume that the dispersion coefficient is given as the multiplication operator associated with σ : $[0, T] \times \mathcal{S} \to L^2(\mathbb{R})$, which we also denote by σ, where σ is Lipschitz continuous with respect to the second variable uniformly in $t \leq T$, that is, we assume that there exists $L_\sigma > 0$ such that, for all $u_1, u_2 \in \mathcal{S}$ and $t \in [0, T]$,

$$\|\sigma(t, u_1) - \sigma(t, u_2)\| \leq L_\sigma \|u_1 - u_2\|. \tag{19}$$

The correlations are described by the kernel q. For f, g in $L^2(\mathbb{R})$,

$$E\big(\langle f, \mathcal{W}_t^Q \rangle \langle g, \mathcal{W}_t^Q \rangle\big) = \int f(x) Q g(x) \, dx$$
$$= \int f(x) \int q(x, z) \int q(z, y) g(y) \, dy \, dz \, dx,$$

so formally,

$$`E\big(\mathcal{W}_t^Q(x) \mathcal{W}_t^Q(y)\big) = E\big(\langle \delta_x, \mathcal{W}_t^Q \rangle \langle \delta_y, \mathcal{W}_t^Q \rangle\big) = q * q(x, y)',$$

where we denote by $q * q(x, y)$ the integral $\int q(x, z) q(z, y) \, dz$. We could for example take

$$q(x, y) = q(x - y) = \frac{1}{2\varepsilon} \mathbb{1}_{(-\varepsilon, \varepsilon)}(x - y) \tag{20}$$

for some small $\varepsilon > 0$ (cf. Sect. 5.1).

The definition of the stochastic integral $\int_0^t \sigma(s, u_s) \, d\mathcal{W}_s^Q(x)$ w.r.t. the Q-Wiener process $\mathcal{W}^Q$ requires in particular that the operator $\sigma(t, v) \circ Q^{\frac{1}{2}}$ is Hilbert–Schmidt. In the following, let L_2^0 be the space of all linear operators L on L^2 for which $L \circ Q^{\frac{1}{2}}$ is Hilbert–Schmidt and denote with $\|L\|_{L_2^0}$ the Hilbert–Schmidt norm of $L \circ Q^{\frac{1}{2}}$.

With the above assumptions on Q and σ it then follows that $\sigma(t, v) \in L_2^0$ since by Parseval's identity

$$\|\sigma(t, v)\|_{L_2^0}^2 = \sum_k \|\bar\sigma(t, v) Q^{\frac{1}{2}} e_k\|_2^2$$
$$= \int \sigma(t, v)^2(x) \|q(x, \cdot)\|^2 \, dx$$
$$\leq \sup_x \|q(x, \cdot)\|^2 \|\sigma(t, v)\|^2 < \infty.$$

Note that, for uncorrelated noise (i.e. $Q = E$), this is not the case. Therefore, in [13] Riedler and Buckwar derive the central limit theorem in the Sobolev space $H^{-\alpha}$. Splitting up the limiting procedures, $N \to \infty$ and continuum limit, allows us to incorporate correlations and finally to work in the more natural function space L^2.

Proposition 3 *For any initial condition $u^0 \in \mathcal{S}$, the stochastic evolution equation*

$$
\begin{aligned}
du_t(x) &= \left(-u_t + w * F(u_t) + \frac{1}{2N} \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \partial_t u_t^{\mathrm{TW}} \right) dt \\
&\quad + \sigma(t, u_t) \, d\mathcal{W}_t^Q(x),
\end{aligned}
\tag{21}
$$
$$
u_0 = u^0,
$$

has a unique strong $\mathcal{S}$-valued solution. u has a continuous modification. For any $p \geq 2$,

$$
E \left(\sup_{t \in [0,T]} \left\| u_t - u_t^{\mathrm{TW}} \right\|^p \right) < \infty.
$$

For a proof see for example Prop. 6.5.1 in [18].

5.3 Embedding of the Diffusion Processes

As a next step we embed the systems of coupled diffusion processes (18) into $L^2(\mathbb{R})$. Let $m \in \mathbb{N}$ be the population density and $L^m \in \mathbb{N}$ the length of the domain with $L^m \uparrow \infty$ as $m \to \infty$. For $k \in \{-mL^m, -mL^m + 1, \ldots, mL^m - 1\}$ set $I_k^m = [\frac{k}{m}, \frac{k+1}{m})$ and $J_k^m = (\frac{k}{m} - \frac{1}{4m}, \frac{k}{m} + \frac{1}{4m})$, and let

$$
W_k^m(t) = 2m \langle \mathcal{W}_t^Q, \mathbb{1}_{J_k^m} \rangle
$$

be the average of $\mathcal{W}_t^Q$ on the interval J_k^m. Then the W_k^m are one-dimensional Brownian motions with covariances

$$
E\left(W_k^m W_l^m \right) = 4m^2 \langle \sqrt{Q} \mathbb{1}_{J_k^m}, \sqrt{Q} \mathbb{1}_{J_l^m} \rangle = 4m^2 \int_{J_k^m} \int_{J_l^m} q * q(y, z) \, dy \, dz.
$$

Note that the Brownian motions are independent as long as $m < \frac{1}{4\varepsilon}$.

For $m \in \mathbb{N}$ let $\hat{\sigma}^m : [0, T] \times \mathbb{R}^P \to \mathbb{R}^P$ and assume that there exists $L_{\hat{\sigma}^m} > 0$ such that, for any $t \in [0, T]$ and $u_1, u_2 \in \mathbb{R}^P$,

$$
\left\| \hat{\sigma}^m(t, u_1) - \hat{\sigma}^m(t, u_2) \right\|_2 \leq L_{\hat{\sigma}^m} \| u_1 - u_2 \|_2.
$$

Consider the system of coupled stochastic differential equations

$$
\begin{aligned}
du_k^m(t) &= \left[\hat{b}_k^m \left(t, (u^m) \right) + \frac{1}{2N} \frac{F''(u_t^{\mathrm{TW}}(\frac{k}{m}))}{F'(u_t^{\mathrm{TW}}(\frac{k}{m}))^2} \hat{b}_k^m \left(t, \left(u_t^{\mathrm{TW}} \left(\frac{k}{m} \right) \right) \right) \right. \\
&\quad \left. + w_k^{m,+} F\left(u_t^{\mathrm{TW}}(L^m) \right) + w_k^{m,-} F\left(u_t^{\mathrm{TW}}(-L^m) \right) \right] dt \\
&\quad + \hat{\sigma}_k^m \left(t, u^m(t) \right) dW_k^m(t), \quad -mL^m \leq k \leq mL^m - 1,
\end{aligned}
$$

with weights as in (14) and (15).

We identify $u = (u_k)_{-mL^m \le k \le mL^m - 1} \in \mathbb{R}^P$ with its piecewise constant interpolation as an element of L^2 via the embedding

$$\iota^m(u) = \sum_{k=-mL^m}^{mL^m-1} u_k \mathbb{1}_{I_k^m}.$$

For $u \in \mathcal{C}(\mathbb{R})$ set

$$\pi^m(u) = \sum_{k=-mL^m}^{mL^m-1} u\left(\frac{k}{m}\right) \mathbb{1}_{I_k^m}.$$

Then $u_t^m := \iota^m((u_k^m(t))_k)$ satisfies

$$
\begin{aligned}
du_t^m(x) = {} & b^m\big(t, u_t^m\big)(x)\,dt \\
& + \frac{1}{2N} \pi^m\left(\frac{F''(u_t^{\text{TW}})}{F'(u_t^{\text{TW}})^2}\right) b^m\big(t, \pi^m(u_t^{\text{TW}})\big)\,dt \\
& + \sigma^m\big(t, u_t^m\big) \circ \Phi^m\,d\mathcal{W}_t^Q(x),
\end{aligned}
\tag{22}
$$

where $b^m : [0, T] \times L^2(\mathbb{R}) \to L^2(\mathbb{R})$ and $\Phi^m : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ are given as

$$
\begin{aligned}
b^m(t, u) = {} & -u_t + \sum_k \Bigg[\int_{-L^m}^{L^m} w\left(\frac{k}{m} - y\right) F(u_t)(y)\,dy \\
& + w_k^{m,+} F\big(u_t^{\text{TW}}(L^m)\big) + w_k^{m,-} F\big(u_t^{\text{TW}}(-L^m)\big) \Bigg] \mathbb{1}_{I_k^m}, \\
\Phi^m(u) = {} & 2m \sum_{k=-mL^m}^{mL^m-1} \langle u, \mathbb{1}_{J_k^m} \rangle \mathbb{1}_{I_k^m},
\end{aligned}
$$

and where $\sigma^m : [0, T] \times L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is such that, for $u \in \mathbb{R}^P$, $\sigma^m(t, \iota^m(u)) = \hat{\sigma}_k^m(t, u)$ on I_k^m. We assume joint continuity and Lipschitz continuity in the second variable uniformly in m and $t \le T$, that is, there exists $L_\sigma > 0$ such that for $u_1, u_2 \in L^2(\mathbb{R})$ and $t \le T$,

$$\big\| \sigma^m(t, u_1) - \sigma^m(t, u_2) \big\| \le L_\sigma \| u_1 - u_2 \|.$$

Proposition 4 *For any initial condition $u^0 \in L^2(\mathbb{R})$ there exists a unique strong L^2-valued solution u^m to* (22). *u^m admits a continuous modification. For any $p \ge 2$,*

$$E\left(\sup_{t \le T} \| u_t^m \|^p \right) < \infty.$$

Proof Again we check that the drift and diffusion coefficients are Lipschitz continuous. Note that

$$\sum_k \frac{1}{m} w\left(\frac{k}{m}, y\right) = \int_{-L^m}^{L^m} w(x - y)\, dx + \sum_k \int_{\frac{k}{m}}^{\frac{k+1}{m}} w\left(\frac{k}{m} - y\right) - w(x - y)\, dx$$

$$\leq \|w\|_1 + \sum_k \int_{\frac{k}{m}}^{\frac{k+1}{m}} \int_{\frac{k}{m}}^{\frac{k+1}{m}} \left|w_x(z - y)\right| dz\, dx$$

$$= 1 + \sum_k \frac{1}{m} \int_{\frac{k}{m}}^{\frac{k+1}{m}} \left|w_x(z - y)\right| dz$$

$$\leq 1 + \frac{1}{m}\|w_x\|_1. \tag{23}$$

Therefore, for $u_1, u_2 \in L^2(\mathbb{R})$,

$$\left\|b^m(t, u_1) - b^m(t, u_2)\right\|_2^2$$

$$\leq 2\|u_1 - u_2\|^2 + 2\|F'\|_\infty^2 \int_{-L^m}^{L^m} \sum_k \frac{1}{m} w\left(\frac{k}{m} - y\right)\left(u_1(y) - u_2(y)\right)^2 dy$$

$$\leq 2\|u_1 - u_2\|^2 + 2\|F'\|_\infty^2 \left(1 + \frac{1}{m}\|w_x\|_1\right)\|u_1 - u_2\|^2$$

and for an orthonormal basis (e_k) of $L^2(\mathbb{R})$ we obtain, using Parseval's identity,

$$\left\|\left(\sigma^m(t, u_1) - \sigma^m(t, u_2)\right) \circ \Phi^m\right\|_{L_2^0}^2$$

$$= \sum_k \left\|\left(\sigma^m(t, u_1) - \sigma^m(t, u_2)\right) \sum_l 2m\langle e_k, \sqrt{Q}\mathbb{1}_{J_l^m}\rangle \mathbb{1}_{I_l^m}\right\|^2$$

$$= \int \left(\sigma^m(t, u_1) - \sigma^m(t, u_2)\right)^2(x) \sum_l 4m^2 \int \left(\int_{J_l^m} q(z, y)\, dy\right)^2 dz\, \mathbb{1}_{I_l^m}(x)\, dx$$

$$\leq \int \left(\sigma^m(t, u_1) - \sigma^m(t, u_2)\right)^2(x) \sum_l 2m \int \int_{J_l^m} q(z, y)^2\, dy\, dz\, \mathbb{1}_{I_l^m}(x)\, dx$$

$$= \int \left(\sigma^m(t, u_1) - \sigma^m(t, u_2)\right)^2(x) \sum_l 2m \int_{J_l^m} \|q(y, \cdot)\|^2\, dy\, \mathbb{1}_{I_l^m}(x)\, dx$$

$$\leq \sup_x \|q(x, \cdot)\|^2 L_\sigma^2 \|u_1 - u_2\|^2. \qquad \square$$

5.4 Convergence

We are now able to state the main convergence result. We will need the following assumption on the kernel w.

Assumption 5 There exists $C_w > 0$ such that, for $x \geq 0$,

$$\int_x^\infty w(y)\,dy \leq C_w w(x). \tag{24}$$

That assumption is satisfied for classical choices of w such as $w(x) = \frac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}$ or $w(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}$.

Theorem 6 *Fix $T > 0$. Let u and u^m be the solutions to* (21) *and* (22)*, respectively. Assume that*

(i) $\sup_k \sup_{x \in I_k^m} \|2m\sqrt{Q}\mathbb{1}_{J_k^m} - q(x,\cdot)\| \xrightarrow{m \to \infty} 0$,
(ii) *for any $u : [0,T] \to \mathcal{S}$ with $\sup_{t \leq T}\|u_t - \hat{u}\| < \infty$,*

$$\sup_{t \leq T}\big\|\sigma^m(t, u_t \mathbb{1}_{(-L^m, L^m)}) - \sigma(t, u_t)\big\| \xrightarrow{m \to \infty} 0.$$

Then, for any initial conditions $u_0^m \in L^2(\mathbb{R})$, $u_0 \in \mathcal{S}$ such that

$$\big\|u_0^m - u_0\big\|_{L^2((-L^m, L^m))} \xrightarrow{m \to \infty} 0,$$

and for all $p \geq 2$,

$$\mathbb{E}\bigg(\sup_{t \in [0,T]} \big\|u_t^m - u_t\big\|_{L^2((-L^m, L^m))}^p\bigg) \xrightarrow{m \to \infty} 0.$$

We postpone the proof to the Appendix.

Remark 7 Let $\varepsilon > 0$. The kernel $q(x,y) = \frac{1}{2\varepsilon}\mathbb{1}_{(x-\varepsilon, x+\varepsilon)}(y)$ satisfies assumption (i) of the theorem. Indeed, note that, for x, z with $|x - z| \leq \frac{1}{m}$, $|\{y : \mathbb{1}_{(x-\varepsilon, x+\varepsilon)}(y) \neq \mathbb{1}_{(z-\varepsilon, z+\varepsilon)}(y)\}| \leq |z - x| \leq \frac{1}{m}$. Therefore we obtain, for all k and for any $x \in I_k^m$,

$$\big\|2m\sqrt{Q}\mathbb{1}_{J_k^m} - q(x,\cdot)\big\|_2^2 = 4m^2 \int_{-\infty}^\infty \bigg(\int_{J_k^m} q(z,y) - q(x,y)\,dz\bigg)^2 dy$$

$$\leq 2m \int_{-\infty}^\infty \int_{J_k^m} \big(q(z,y) - q(x,y)\big)^2 dz\,dy$$

$$\leq 2\int_{J_k^m} \frac{1}{4\varepsilon^2}\,dz \leq \frac{1}{4\varepsilon^2 m} \xrightarrow{m \to \infty} 0.$$

Theorem 6 can now be applied to the asymptotic expansion of the fluctuations around the traveling wave derived in Sect. 4.2. In order to ensure that the diffusion coefficients are in $L^2(\mathbb{R})$, we cut off the noise outside a compact set $\{\partial_x u_t^{\text{TW}} \geq \delta\}$, $\delta > 0$. Note that the neglected region moves with the wave such that we always retain the fluctuations in the relevant regime away from the fixed points.

Theorem 8 *Assume that the wave speed is strictly positive, $c > 0$. Fix $\delta > 0$. The diffusion coefficients as derived in Sect.* 4.2,

$$\sigma(t, u) = \frac{1}{\sqrt{N}} \big(\alpha(t) + \beta(t) \big(u - u_t^{\mathrm{TW}} \big)$$
$$- \gamma(t) w * \big(F'\big(u_t^{\mathrm{TW}} \big) \big(u - u_t^{\mathrm{TW}} \big) \big) \big) \mathbb{1}_{\{\partial_x u_t^{\mathrm{TW}} \geq \delta\}},$$

$$\sigma^m(t, u) = \frac{1}{\sqrt{N}} \big(\alpha^m(t) + \beta^m(t) \big(u - \pi^m\big(u_t^{\mathrm{TW}} \big) \big)$$
$$- \gamma^m(t) \pi^m \big(w * \big(\pi^m \big(F'\big(u_t^{\mathrm{TW}} \big) \big) \big(u - \pi^m \big(u_t^{\mathrm{TW}} \big) \big) \big) \big) \big) \mathbb{1}_{\{\partial_x u_t^{\mathrm{TW}} \geq \delta\}},$$

where

$$\alpha(t) = \sqrt{\frac{| - u_t^{\mathrm{TW}} + w * F(u_t^{\mathrm{TW}})|}{F'(u_t^{\mathrm{TW}})}} = \sqrt{\frac{c \partial_x u_t^{\mathrm{TW}}}{F'(u_t^{\mathrm{TW}})}},$$

$$\beta(t) = \frac{1}{2\sqrt{c \partial_x u_t^{\mathrm{TW}} F'(u_t^{\mathrm{TW}})}} \left(- \frac{F''(u_t^{\mathrm{TW}}(x))}{F'(u_t^{\mathrm{TW}}(x))} c \partial_x u_t^{\mathrm{TW}} + 1 \right),$$

$$\gamma(t) = \frac{1}{2\sqrt{c \partial_x u_t^{\mathrm{TW}} F'(u_t^{\mathrm{TW}})}},$$

$$\alpha^m(t) = \sqrt{\frac{-b^m(t, \pi^m(u_t^{\mathrm{TW}}))}{\pi^m(F'(u_t^{\mathrm{TW}}))}} \mathbb{1}_{[-L^m, L^m)},$$

$$\beta^m(t) = \frac{1}{2\sqrt{-b^m(t, \pi^m(u_t^{\mathrm{TW}})) \pi^m(F'(u_t^{\mathrm{TW}}))}}$$
$$\times \left(- \frac{\pi^m(F''(u_t^{\mathrm{TW}}))}{\pi^m(F'(u_t^{\mathrm{TW}}))} \big(-b^m\big(t, \pi^m\big(u_t^{\mathrm{TW}} \big) \big) \big) + 1 \right) \mathbb{1}_{[-L^m, L^m)},$$

$$\gamma^m(t) = \frac{1}{2\sqrt{-b^m(t, \pi^m(u_t^{\mathrm{TW}})) \pi^m(F'(u_t^{\mathrm{TW}}))}} \mathbb{1}_{[-L^m, L^m)},$$

are jointly continuous and Lipschitz continuous in the second variable with Lipschitz constant uniform in m and $t \leq T$, and satisfy condition (ii) *of Theorem* 6.

For a proof see Thm. 6.5.5 in [18].

6 Summary and Conclusions

In this paper we have investigated the derivation of neural field equations from detailed neuron models. Following a common approach in the literature, we started from a phenomenological Markov chain model that yields the neural field equation in the

infinite-population limit and allows one to obtain stochastic or deterministic corrections for networks of finite populations. As one novelty to the existing literature we considered a new choice of jump rates, given in the activity-based setting as

$$q^{P,N}\left(x, x \pm \frac{1}{N}e_i\right) = N\left(-x_i + F\left(\sum_{j=1}^{P} w_{ij}x_j\right)\right)_{\pm}, \qquad (25)$$

and showed that it leads to qualitatively different results in dynamical states with high (resp. low) activity in comparison with the rates

$$\begin{aligned} \tilde{q}\left(x, x + \frac{1}{N}e_i\right) &= NF\left(\sum_{j=1}^{P} w_{ij}x_j\right), \\ \tilde{q}\left(x, x - \frac{1}{N}e_i\right) &= Nx_i, \end{aligned} \qquad (26)$$

considered in the literature so far (see for example [11, 13]).

The following two major reasons make us believe that the rates (25) provide a better description of finite-size effects in neural fields.

1. The neural field equations evolve on a different time scale than single neuron activity. In the activity-based setting they describe a coarse-grained time-averaged population activity and in the voltage-based setting they are derived under the assumption that the activity in each population is slowly varying such that one can assume that it is related to the population potential via $a = F(u)$. The rates (25) reflect this property more than the rates (26), in particular in regimes with high (resp. low) activity.
2. Fluctuations due to finite-size effects should be strong where $F' \gg 1$, and therefore small changes in the synaptic input lead to comparably large changes in the population activity. They should become small, however, where $F' \ll 1$, so in particular around the stable fixed points of the system. This is captured in the jump rates (25), whereas in case (26), fluctuations are high where the activity is high.

The corrections to the neural field equation derived in this setting are of a qualitatively different form than the ones previously considered. As a consequence, one should reconsider effects of intrinsic noise due to finite-size effects as studied for example in [25]. In particular the lowest order correction to the moment equations, or the additive part of the stochastic correction, vanishes when linearizing around a stationary solution. The noise is therefore of smaller order than previously assumed.

This is not the case for stable moving patterns like traveling fronts or pulses, suggesting the movement as a main source of noise. We have for the first time rigorously derived a well-posed stochastic continuum neural field equation with an additive noise term that can be used to study the finite-size effects on these kinds of solutions. It is particularly suitable for the analysis of traveling wave solutions, since the monotonicity of the solution allows one to consider a multiplicative noise term as derived in Sect. 4.2.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

EL performed the analysis and wrote a first version of the manuscript. WS contributed mathematical discussions and rewrote versions of the manuscript. All authors read and approved the final manuscript.

Appendix: Proof of Theorem 6

Set $v_t = u_t - u_t^{\mathrm{TW}}$ and $v_t^m = u_t^m - \pi^m(u_t^{\mathrm{TW}})$. Note that

$$
\int_{-L^m}^{L^m} \left(\pi^m\left(u_t^{\mathrm{TW}}\right)(x) - u_t^{\mathrm{TW}}(x)\right)^2 dx
$$

$$
\leq \sum_k \int_{\frac{k}{m}}^{\frac{k+1}{m}} \left(\int_{\frac{k}{m}}^{\frac{k+1}{m}} \partial_x u_t^{\mathrm{TW}}(z)\, dz\right)^2 dx
$$

$$
\leq \frac{1}{m^2} \sum_k \int_{\frac{k}{m}}^{\frac{k+1}{m}} \left(\partial_x u_t^{\mathrm{TW}}(z)\right)^2 dz \leq \frac{1}{m^2} \|\hat{u}_x\|^2. \tag{27}
$$

For the proof of the theorem it therefore suffices to show that

$$
E\left(\sup_{t\leq T} \|v_t - v_t^m\|_2^p\right) \xrightarrow{m\to\infty} 0,
$$

since this will imply that

$$
E\left(\sup_{t\leq T} \|u_t - u_t^m\|_{L^2((-L^m, L^m))}^p\right)
$$

$$
\leq const \times \left[E\left(\sup_{t\leq T} \|u_t^{\mathrm{TW}} - \pi^m(u_t^{\mathrm{TW}})\|_{L^2((-L^m, L^m))}^p\right) + E\left(\sup_{t\leq T} \|v_t - v_t^m\|_2^p\right)\right]
$$

$$
\xrightarrow{m\to\infty} 0.
$$

By Itô's formula,

$$
\frac{1}{2} d \|v_t^m - v_t\|_2^2
$$

$$
= \left\langle b^m\left(t, v_t^m + \pi^m\left(u_t^{\mathrm{TW}}\right)\right) - b\left(t, v_t + u_t^{\mathrm{TW}}\right) - \pi^m\left(\partial_t u_t^{\mathrm{TW}}\right) + \partial_t u_t^{\mathrm{TW}}, v_t^m - v_t\right\rangle dt
$$

$$+ \frac{1}{2N}\left\langle \pi^m\left(\frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2}\right)b^m\left(t, \pi^m(u_t^{\mathrm{TW}})\right) - \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2}\partial_t u_t^{\mathrm{TW}}, v_t^m - v_t\right\rangle dt$$

$$+ \frac{1}{2}\left\|\sigma^m\left(t, v_t^m + \pi^m(u_t^{\mathrm{TW}})\right)\circ\Phi^m - \sigma\left(t, v_t + u_t^{\mathrm{TW}}\right)\right\|_{L_2^0}^2 dt$$

$$+ \left\langle v_t^m - v_t, \left(\sigma^m\left(t, v_t^m + \pi^m(u_t^{\mathrm{TW}})\right)\circ\Phi^m - \sigma\left(t, v_t + u_t^{\mathrm{TW}}\right)\right)d\mathcal{W}_t^Q\right\rangle.$$

In order to finally apply Gronwall's lemma, we estimate the terms one by one.

A.1 The Drift

We start by regrouping the terms in a suitable way. We have

$$\left\|v_t^m - v_t + b^m\left(t, v_t^m + \pi^m(u_t^{\mathrm{TW}})\right) - b\left(t, v_t + u_t^{\mathrm{TW}}\right) - \pi^m\left(\partial_t u_t^{\mathrm{TW}}\right) + \partial_t u_t^{\mathrm{TW}}\right\|^2$$

$$= \int_{-\infty}^{\infty}\left[\sum_k\left(\int_{-L^m}^{L^m} w\left(\frac{k}{m} - y\right)F\left(v_t^m(y) + \pi^m(u_t^{\mathrm{TW}})(y)\right)dy\right.\right.$$

$$+ \int_{L^m}^{\infty} w\left(\frac{k}{m} - y\right)F\left(u_t^{\mathrm{TW}}(L^m)\right)dy$$

$$+ \int_{-\infty}^{-L^m} w\left(\frac{k}{m} - y\right)F\left(u_t^{\mathrm{TW}}(-L^m)\right)dy\right)\mathbb{1}_{I_k^m}(x)$$

$$- \int_{-\infty}^{\infty} w(x - y)F\left(v_t(y) + u_t^{\mathrm{TW}}(y)\right)dy$$

$$\left.- \pi^m\left(w * F(u_t^{\mathrm{TW}})\right) + w * F(u_t^{\mathrm{TW}})(x)\right]^2 dx$$

$$\leq 6\int_{-\infty}^{\infty}\left[\sum_k\int_{-L^m}^{L^m} w\left(\frac{k}{m} - y\right)\left(F\left(v_t^m(y) + \pi^m(u_t^{\mathrm{TW}})(y)\right)\right.\right.$$

$$\left.\left.- F\left(v_t(y) + u_t^{\mathrm{TW}}(y)\right)\right)dy\,\mathbb{1}_{I_k^m}(x)\right]^2$$

$$+ \left[\sum_k\int_{-L^m}^{L^m}\left(w\left(\frac{k}{m} - y\right) - w(x - y)\right)\left(F\left(v_t(y) + u_t^{\mathrm{TW}}(y)\right)\right.\right.$$

$$\left.\left.- F\left(u_t^{\mathrm{TW}}(y)\right)\right)dy\,\mathbb{1}_{I_k^m}(x)\right]^2$$

$$+ \left[\sum_k\int_{L^m}^{\infty} w\left(\frac{k}{m} - y\right)\left(F\left(u_t^{\mathrm{TW}}(L^m)\right) - F\left(u_t^{\mathrm{TW}}(y)\right)\right)dy\,\mathbb{1}_{I_k^m}(x)\right]^2$$

$$+ \left[\sum_k\int_{-\infty}^{-L^m} w\left(\frac{k}{m} - y\right)\left(F\left(u_t^{\mathrm{TW}}(-L^m)\right) - F\left(u_t^{\mathrm{TW}}(y)\right)\right)dy\,\mathbb{1}_{I_k^m}(x)\right]^2$$

$$+ \left[\left(\int_{L^m}^{\infty} + \int_{-\infty}^{-L^m} \right) w(x-y) \left(F\left(v_t(y) + u_t^{\text{TW}}(y) \right) - F\left(u_t^{\text{TW}}(y) \right) \right) dy \right]^2$$

$$+ \left[\int_{-L^m}^{L^m} w(x-y) \left(F\left(v_t(y) + u_t^{\text{TW}}(y) \right) \right. \right.$$

$$\left. \left. - F\left(u_t^{\text{TW}}(y) \right) \right) dy \mathbb{1}_{(-\infty, -L^m) \cup [L^m, \infty)}(x) \right]^2 dx$$

$$=: 6(S_1 + S_2 + S_3 + S_4 + S_5 + S_6).$$

Using the Cauchy–Schwarz inequality we get

$$S_1 \leq \|F'\|_{\infty}^2 \int_{-L^m}^{L^m} \sum_k \frac{1}{m} w\left(\frac{k}{m} - y \right)$$

$$\times \left(v_t^m(y) + \pi^m \left(u_t^{\text{TW}} \right)(y) - v_t(y) - u_t^{\text{TW}}(y) \right)^2 dy$$

$$\overset{(23)}{\leq} 2\left(1 + \frac{1}{m} \|w_x\|_1 \right) \|F'\|_{\infty}^2$$

$$\times \left(\left\| v_t^m - v_t \right\|_2^2 + \int_{-L^m}^{L^m} \left(\pi^m \left(u_t^{\text{TW}} \right)(y) - u_t^{\text{TW}}(y) \right)^2 dy \right).$$

With (27) it follows that

$$S_1 \leq 2\left(1 + \frac{1}{m} \|w_x\|_1 \right) \|F'\|_{\infty}^2 \left(\left\| v_t^m - v_t \right\|_2^2 + \frac{1}{m^2} \|\hat{u}_x\|_2^2 \right).$$

Another application of the Cauchy–Schwarz inequality yields

$$S_2 = \sum_k \int_{\frac{k}{m}}^{\frac{k+1}{m}} \left(\int_{-L^m}^{L^m} \left(w\left(\frac{k}{m} - y \right) - w(x-y) \right) \right.$$

$$\left. \times \left(F\left(v_t(y) + u_t^{\text{TW}}(y) \right) - F\left(u_t^{\text{TW}}(y) \right) \right) dy \right)^2 dx$$

$$\leq \sum_k \frac{1}{m} \left(\int_{-L^m}^{L^m} \int_{\frac{k}{m}}^{\frac{k+1}{m}} |w_x(z-y)| \, dz \left(F\left(v_t(y) + u_t^{\text{TW}}(y) \right) - F\left(u_t^{\text{TW}}(y) \right) \right) dy \right)^2$$

$$\leq \sum_k \frac{1}{m} \int_{-L^m}^{L^m} \int_{\frac{k}{m}}^{\frac{k+1}{m}} |w_x(z-y)| \, dz \, dy$$

$$\times \int_{-L^m}^{L^m} \int_{\frac{k}{m}}^{\frac{k+1}{m}} |w_x(z-y)| \, dz \left(F\left(v_t(y) + u_t^{\text{TW}}(y) \right) - F\left(u_t^{\text{TW}}(y) \right) \right)^2 dy$$

$$\leq \frac{1}{m^2} \|w_x\|_1^2 \|F'\|_{\infty}^2 \|v_t\|_2^2.$$

Using integration by parts, (23), and assumption (24), we obtain

$$
S_3 = \sum_k \frac{1}{m} \left(\underbrace{\left[-\int_y^\infty w\left(z - \frac{k}{m}\right) dz \left(F\left(u_t^{\mathrm{TW}}(L^m)\right) - F\left(u_t^{\mathrm{TW}}(y)\right)\right) \right]_{y=L^m}^\infty}_{=0} \right.
$$

$$
\left. - \int_{L^m}^\infty \int_y^\infty w\left(z - \frac{k}{m}\right) dz\, F'\left(u_t^{\mathrm{TW}}(y)\right) \partial_x u_t^{\mathrm{TW}}(y)\, dy \right)^2
$$

$$
\leq C_w^2 \sum_k \frac{1}{m} \left(\int_{L^m}^\infty w\left(y - \frac{k}{m}\right) F'\left(u_t^{\mathrm{TW}}(y)\right) \partial_x u_t^{\mathrm{TW}}(y)\, dy \right)^2
$$

$$
\leq C_w^2 \left(1 + \frac{1}{m} \|w_x\|_1 \right) \|F'\|_\infty^2 \int_{L^m}^\infty \left(\partial_x u_t^{\mathrm{TW}}(y) \right)^2 dy.
$$

Analogously,

$$
S_4 \leq C_w^2 \left(1 + \frac{1}{m} \|w_x\|_1 \right) \|F'\|_\infty^2 \int_{-\infty}^{-L^m} \left(\partial_x u_t^{\mathrm{TW}}(y) \right)^2 dy.
$$

Finally we observe that

$$
S_5 \leq \|F'\|_\infty^2 \left(\int_{L^m}^\infty + \int_{-\infty}^{-L^m} \right) v_t^2(y)\, dy
$$

and

$$
S_6 \leq \|F'\|_\infty^2 \left(\int_{L^m}^\infty + \int_{-\infty}^{-L^m} \right) \left(\int_{-L^m}^{L^m} w(x-y) |v_t(y)|\, dy \right)^2 dx
$$

$$
\leq \|F'\|_\infty^2 \left(\int_{L^m}^\infty + \int_{-\infty}^{-L^m} \right) \left(w * |v_t|(x) \right)^2 dx.
$$

Finally we consider

$$
S_7 := \left\| \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \partial_t u_t^{\mathrm{TW}} - \pi^m \left(\frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right) b^m \left(t, \pi^m \left(u_t^{\mathrm{TW}} \right) \right) \right\|^2
$$

$$
\leq 4 \left[\left(\int_{-\infty}^{-L^m} + \int_{L^m}^\infty \right) \left(\frac{F''(u_t^{\mathrm{TW}})(y)}{F'(u_t^{\mathrm{TW}})^2(y)} \partial_t u_t^{\mathrm{TW}}(y) \right)^2 dy \right.
$$

$$
+ \left\| \left(\frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} - \pi^m \left(\frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right) \right) \partial_t u_t^{\mathrm{TW}} \mathbb{1}_{(-L^m, L^m)} \right\|^2
$$

$$
+ \left\| \pi^m \left(\frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right) \left(\partial_t u_t^{\mathrm{TW}} - \pi^m \left(\partial_t u_t^{\mathrm{TW}} \right) \right) \mathbb{1}_{(-L^m, L^m)} \right\|^2
$$

$$
+ \left\| \pi^m \left(\frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right) \left(\pi^m \left(w * F\left(u_t^{\mathrm{TW}}\right) \right) - \pi^m \left(w * \pi^m \left(F\left(u_t^{\mathrm{TW}}\right) \right) \right) \right) \right.
$$

$$- \sum_k \left(w_k^{m,+} F\left(u_t^{\mathrm{TW}}(L^m)\right) + w_k^{m,-} F\left(u_t^{\mathrm{TW}}\right)(-L^m) \right) \mathbb{1}_{I_k^m} \right) \Bigg\|^2 \Bigg]$$

$$= 4(S_{7,1} + S_{7,2} + S_{7,3} + S_{7,4}).$$

We have

$$S_{7,2} \le \left\| \frac{F^{(3)}(\hat{u})\hat{u}_x}{(F'(\hat{u}))^2} - \frac{2(F''(\hat{u}))^2 \hat{u}_x}{(F'(\hat{u}))^3} \right\|_\infty^2 \frac{1}{m^2} \left\| \partial_t u_t^{\mathrm{TW}} \right\|^2$$

and, as in (27),

$$S_{7,3} \le \left\| \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right\|_\infty^2 \left\| \left(\partial_t u_t^{\mathrm{TW}} - \pi^m \left(\partial_t u_t^{\mathrm{TW}} \right) \right) \mathbb{1}_{(-L^m, l^m)} \right\|^2$$

$$\le \frac{1}{m^2} c^2 \left\| \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right\|_\infty^2 \| \hat{u}_{xx} \|^2.$$

The last summand satisfies, using (23) and (27),

$$S_{7,4} \le 3 \left\| \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right\|_\infty^2$$

$$\times \left[\sum_k \frac{1}{m} \left(\int_{-L^m}^{L^m} w\left(\frac{k}{m} - y \right) \left(F\left(u_t^{\mathrm{TW}}(y)\right) - \pi^m \left(F\left(u_t^{\mathrm{TW}}\right)\right)(y) \right)^2 dy \right)^2 \right.$$

$$\left. + S_3 + S_4 \right]$$

$$\le 3 \left\| \frac{F''(u_t^{\mathrm{TW}})}{F'(u_t^{\mathrm{TW}})^2} \right\|_\infty^2 \left[\left(1 + \frac{1}{m} \| w_x \|_1 \right) \| F' \|_\infty^2 \frac{1}{m^2} \| \hat{u}_x \|^2 + S_3 + S_4 \right].$$

A.2 The Itô Correction

We have

$$\left\| \sigma^m\left(t, v_t^m + \pi^m\left(u_t^{\mathrm{TW}}\right)\right) \circ \Phi^m - \sigma\left(t, v_t + u_t^{\mathrm{TW}}\right) \right\|_{L_2^0}^2$$

$$\le 3 \Big(\left\| \left(\sigma^m\left(t, v_t^m + \pi^m\left(u_t^{\mathrm{TW}}\right)\right) - \sigma^m\left(t, \left(v_t + u_t^{\mathrm{TW}}\right) \mathbb{1}_{(-L^m, L^m)}\right)\right) \circ \Phi^m \right\|_{L_2^0}^2$$

$$+ \left\| \left(\sigma^m\left(t, \left(v_t + u_t^{\mathrm{TW}}\right) \mathbb{1}_{(-L^m, L^m)}\right) - \sigma\left(t, v_t + u_t^{\mathrm{TW}}\right)\right) \circ \Phi^m \right\|_{L_2^0}^2$$

$$+ \left\| \sigma\left(t, v_t + u_t^{\mathrm{TW}}\right) \circ \Phi^m - \sigma\left(t, v_t + u_t^{\mathrm{TW}}\right) \right\|_{L_2^0}^2 \Big)$$

$$=: 3(S_8 + S_9 + S_{10}).$$

Let (e_k) be an orthonormal basis of $L^2(\mathbb{R})$. Note that by Parseval's identity

$$
\sum_k \left(\Phi^m (\sqrt{Q} e_k) \right)^2
$$

$$
= \sum_k \left(\sum_l 2m \langle \sqrt{Q} e_k, \mathbb{1}_{J_l^m} \rangle \mathbb{1}_{I_l^m} \right)^2
$$

$$
= \sum_l \sum_k 4m^2 \langle \sqrt{Q} \mathbb{1}_{J_l^m}, e_k \rangle^2 \mathbb{1}_{I_l^m} = \sum_l 4m^2 \| \sqrt{Q} \mathbb{1}_{J_l^m} \|^2 \mathbb{1}_{I_l^m}
$$

$$
= \sum_l 4m^2 \int \left(\int_{J_l^m} q(x,y)\,dy \right)^2 dx\, \mathbb{1}_{I_l^m}
$$

$$
\leq \sum_l 2m \int \int_{J_l^m} q^2(x,y)\,dy\,dx\, \mathbb{1}_{I_l^m} \leq \sup_x \| q(x,\cdot) \|^2 \mathbb{1}_{[-L^m, L^m)}.
$$

Thus,

$$
\begin{aligned}
S_8 &= \sum_k \int_{-\infty}^{\infty} \big(\sigma^m \big(t, v_t^m + \pi^m (u_t^{\mathrm{TW}}) \big) \\
&\quad - \sigma^m \big(t, (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)} \big) \big)^2(x) \big(\Phi^m (\sqrt{Q} e_k) \big)^2(x)\,dx \\
&\leq \sup_x \| q(x,\cdot) \|^2 L_\sigma^2 \big\| v_t^m + \pi^m (u_t^{\mathrm{TW}}) - (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)} \big\|^2 \\
&\stackrel{(27)}{\leq} 2 \sup_x \| q(x,\cdot) \|^2 L_\sigma^2 \left(\| v_t^m - v_t \|^2 + \frac{1}{m^2} \| \hat{u}_x \|^2 \right).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
S_8 &= \sum_k \int_{-\infty}^{\infty} \big(\sigma^m \big(t, v_t^m + \pi^m (u_t^{\mathrm{TW}}) \big) \\
&\quad - \sigma^m \big(t, (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)} \big) \big)^2(x) \big(\Phi^m (\sqrt{Q} e_k) \big)^2(x)\,dx \\
&\leq \sup_x \| q(x,\cdot) \|^2 L_\sigma^2 \big\| v_t^m + \pi^m (u_t^{\mathrm{TW}}) - (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)} \big\|^2 \\
&\stackrel{(27)}{\leq} 2 \sup_x \| q(x,\cdot) \|^2 L_\sigma^2 \left(\| v_t^m - v_t \|^2 + \frac{1}{m^2} \| \hat{u}_x \|^2 \right)
\end{aligned}
$$

and

$$
\begin{aligned}
S_9 &= \big\| \big(\sigma^m \big(t, (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)} \big) - \sigma \big(t, v_t + u_t^{\mathrm{TW}} \big) \big) \circ \Phi^m \big\|_{L_2^0}^2 \\
&\leq \sup_x \| q(x,\cdot) \|^2 \big\| \sigma^m \big(t, (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)} \big) - \sigma \big(t, v_t + u_t^{\mathrm{TW}} \big) \big\|^2.
\end{aligned}
$$

Using Parseval's identity again we get

$$
\begin{aligned}
S_{10} &= \int_{-\infty}^{\infty} \sigma\left(t, v_t + u_t^{\text{TW}}\right)^2(x) \\
&\quad \times \sum_k \left(\sum_l 2m \langle \sqrt{Q} e_k, \mathbb{1}_{J_l^m} \rangle \mathbb{1}_{I_l^m}(x) - \sqrt{Q} e_k(x) \right)^2 dx \\
&= \sum_l \int_{\frac{l}{m}}^{\frac{l+1}{m}} \sigma\left(t, v_t + u_t^{\text{TW}}\right)^2(x) \\
&\quad \times \sum_k \left(2m \int_{J_l^m} \int_{-\infty}^{\infty} \left(q(z,y) - q(x,y)\right) e_k(y)\, dy\, dz \right)^2 dx \\
&\quad + \left(\int_{-\infty}^{-L^m} + \int_{L^m}^{\infty} \right) \sigma\left(t, v_t + u_t^{\text{TW}}\right)^2(x) \sum_k \left(\sqrt{Q} e_k(x)\right)^2 dx \\
&\leq \sum_l \int_{\frac{l}{m}}^{\frac{l+1}{m}} \sigma\left(t, v_t + u_t^{\text{TW}}\right)^2(x) \left\| 2m \int_{J_l^m} \left(q(z,\cdot) - q(x,\cdot)\right) dz \right\|^2 dx \\
&\quad + \left(\int_{-\infty}^{-L^m} + \int_{L^m}^{\infty} \right) \sigma\left(t, v_t + u_t^{\text{TW}}\right)^2(x) \sum_k \left(\int_{-\infty}^{\infty} q(x,y) e_k(y)\, dy \right)^2 dx \\
&\leq \left\| \sigma\left(t, v_t + u_t^{\text{TW}}\right) \right\|^2 \sup_l \sup_{x \in I_l^m} \left\| 2m \int_{J_l^m} q(z,\cdot)\, dz - q(x,\cdot) \right\|^2 \\
&\quad + \sup_x \|q(x,\cdot)\|^2 \left(\int_{-\infty}^{-L^m} + \int_{L^m}^{\infty} \right) \sigma\left(t, v_t + u_t^{\text{TW}}\right)^2(x)\, dx.
\end{aligned}
$$

A.3 Application of Gronwall's Lemma

We use K, K_1, K_2, $\tilde{K}$, etc. to denote suitable constants that may differ from step to step. Summarizing the previous steps and using Young's inequality we arrive at

$$
\begin{aligned}
\frac{1}{2} d \|v_t^m - v_t\|^2 \\
\leq \Big[&-\|v_t^m - v_t\|^2 + \frac{1}{2} \|v_t^m - v_t\|^2 \\
&+ \frac{1}{2} \big\| v_t^m - v_t + b^m\left(t, v_t^m + \pi^m\left(u_t^{\text{TW}}\right)\right) - b\left(t, v_t + u_t^{\text{TW}}\right) \\
&- \pi^m\left(\partial_t u_t^{\text{TW}}\right) + \partial_t u_t^{\text{TW}} \big\|_2^2 + \frac{1}{4N} \|v_t^m - v_t\|^2 \\
&+ \frac{1}{4N} \left\| \frac{F''(u_t^{\text{TW}})}{F'(u_t^{\text{TW}})^2} \partial_t u_t^{\text{TW}} - \pi^m \left(\frac{F''(u_t^{\text{TW}})}{F'(u_t^{\text{TW}})^2} \right) b^m\left(t, \pi^m\left(u_t^{\text{TW}}\right)\right) \right\|^2 \Big] dt
\end{aligned}
$$

$$+ \frac{1}{2} \left\| \sigma^m \big(t, v_t^m + \pi^m \big(u_t^{\mathrm{TW}}\big)\big) \circ \Phi^m - \sigma \big(t, v_t + u_t^{\mathrm{TW}}\big) \right\|_{L_2^0}^2 dt$$

$$+ \left\langle v_t^m - v_t, \big(\sigma^m \big(t, v_t^m + \pi^m \big(u_t^{\mathrm{TW}}\big)\big) \circ \Phi^m - \sigma \big(t, v_t + u_t^{\mathrm{TW}}\big)\big) d\mathcal{W}_t^Q \right\rangle$$

$$\leq K_1 \left\| v_t^m - v_t \right\|^2 dt + K_2 R(t, v_t, m) \, dt + dM_t,$$

where

$$R(t, v_t, m) = \frac{1}{m^2} \big(\|v_t\|^2 + \|\hat{u}_x\|^2 + \|\hat{u}_{xx}\|^2 \big)$$

$$+ \left(\int_{L^m}^{\infty} + \int_{-\infty}^{-L^m} \right) \big(v_t^2(x) + (w * |v_t|)^2(x)$$

$$+ \sigma \big(t, v_t + u_t^{\mathrm{TW}}\big)^2(x) + \big(\partial_x u_t^{\mathrm{TW}}(x)\big)^2 \big) \, dx$$

$$+ \left\| \sigma \big(t, v_t + u_t^{\mathrm{TW}}\big) \right\|^2 \sup_k \sup_{x \in I_k^m} \left\| 2m\sqrt{Q} \mathbb{1}_{J_k^m} - q(x, \cdot) \right\|^2$$

$$+ \left\| \sigma^m \big(t, (v_t + u_t^{\mathrm{TW}}) \mathbb{1}_{(-L^m, L^m)}\big) - \sigma \big(t, v_t + u_t^{\mathrm{TW}}\big) \right\|^2$$

and

$$M_t = \int_0^t \left\langle v_s^m - v_s, \big(\sigma^m \big(s, v_s^m + \pi^m \big(u_s^{\mathrm{TW}}\big)\big) \circ \Phi^m - \sigma \big(s, v_s + u_s^{\mathrm{TW}}\big)\big) d\mathcal{W}_s^Q \right\rangle$$

is a martingale with quadratic variation process

$$[M]_t = \int_0^t \sum_k \langle v_s^m - v_s, \big(\sigma^m \big(s, v_s^m + \pi^m \big(u_s^{\mathrm{TW}}\big)\big) \circ \Phi^m$$

$$- \sigma \big(s, v_s + u_s^{\mathrm{TW}}\big)\big) \circ \sqrt{Q} e_k \rangle^2 \, ds$$

$$\leq \int_0^t \left\| v_s^m - v_s \right\|^2 \left\| \big(\sigma^m \big(s, v_s^m + \pi^m \big(u_s^{\mathrm{TW}}\big)\big) \circ \Phi^m \right.$$

$$\left. - \sigma \big(s, v_s + u_s^{\mathrm{TW}}\big)\big) \right\|_{L_2^0}^2 \, ds. \tag{28}$$

Applying Itô's formula to the real-valued stochastic process $\|v_t^m - v_t\|^2$ we obtain for $p \geq 2$

$$d \left\| v_t^m - v_t \right\|^p$$

$$= \frac{p}{2} \left\| v_t^m - v_t \right\|^{p-2} d \left\| v_t^m - v_t \right\|^2$$

$$+ \frac{p(p-2)}{8} \left\| v_t^m - v_t \right\|^{p-4} d \big[\left\| v_t^m - v_t \right\|^2 \big]_t$$

$$\overset{(28)}{\leq} K_1 p \left\| v_t^m - v_t \right\|^p dt + K_2 p \left\| v_t^m - v_t \right\|^{p-2} R(t, v_t, m) \, dt$$

$$+ p \left\| v_t^m - v_t \right\|^{p-2} dM_t + \frac{p(p-2)}{2} \left\| v_t^m - v_t \right\|^{p-2}$$

$$\times \left\| \left(\sigma^m \left(t, v_t^m + \pi^m \left(u_t^{\mathrm{TW}} \right) \right) \right) \circ \Phi^m - \sigma \left(t, v_t + u_t^{\mathrm{TW}} \right) \right\|_{L_2^0}^2 dt.$$

Estimating the last term as above and using Young's inequality we obtain

$$d \left\| v_t^m - v_t \right\|^p$$

$$\leq \tilde{K}_1 \left\| v_t^m - v_t \right\|^p dt + \tilde{K}_2 \left\| v_t^m - v_t \right\|^{p-2} R(t, v_t, m) \, dt$$

$$+ p \left\| v_t^m - v_t \right\|^{p-2} dM_t$$

$$\leq \left(\tilde{K}_1 + \tilde{K}_2 \frac{p-2}{p} \right) \left\| v_t^m - v_t \right\|^p dt + \tilde{K}_2 \frac{2}{p} R(t, v_t, m)^{\frac{p}{2}} dt$$

$$+ p \left\| v_t^m - v_t \right\|^{p-2} dM_t.$$

Integrating, maximizing over $t \leq T$, and taking expectations we get

$$E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^p \right)$$

$$\leq \left\| v_0^m - v_0 \right\|^p + \left(\tilde{K}_1 + \tilde{K}_2 \frac{p-2}{p} \right) E \left(\sup_{t \leq T} \int_0^t \left\| v_s^m - v_s \right\|^p ds \right)$$

$$+ \tilde{K}_2 \frac{2}{p} E \left(\sup_{t \leq T} \int_0^t R(s, v_s, m)^{\frac{p}{2}} ds \right) + p E \left(\sup_{t \leq T} \int_0^t \left\| v_s^m - v_s \right\|^{p-2} dM_s \right)$$

$$\leq \left\| v_0^m - v_0 \right\|^p + \left(\tilde{K}_1 + \tilde{K}_2 \frac{p-2}{p} \right) \int_0^T E \left(\sup_{s \leq t} \left\| v_s^m - v_s \right\|^p \right) dt$$

$$+ \tilde{K}_2 \frac{2}{p} E \left(\int_0^T R(s, v_s, m)^{\frac{p}{2}} ds \right) + p E \left(\sup_{t \leq T} \int_0^t \left\| v_s^m - v_s \right\|^{p-2} dM_s \right). \quad (29)$$

We estimate the last term using the Burkholder–Davis–Gundy inequality, (28), and Young's inequality:

$$p E \left(\sup_{t \leq T} \int_0^t \left\| v_s^m - v_s \right\|^{p-2} dM_s \right)$$

$$\leq K E \left(\int_0^T \left\| v_s^m - v_s \right\|_2^{2p-2} \right.$$

$$\left. \times \left\| \left(\sigma^m \left(s, v_s^m + \pi^m \left(u_s^{\mathrm{TW}} \right) \right) \right) \circ \Phi^m - \sigma \left(s, v_s + u_s^{\mathrm{TW}} \right) \right\|_{L_2^0}^2 ds \right)^{\frac{1}{2}}$$

$$\leq K E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^{p-1} \right.$$

$$\times \left(\int_0^T \left\| \left(\sigma^m \left(s, v_s^m + \pi^m \left(u_s^{\mathrm{TW}} \right) \right) \right) \circ \Phi^m - \sigma \left(s, v_s + u_s^{\mathrm{TW}} \right) \right\|_{L_2^0}^2 ds \right)^{\frac{1}{2}} \right)$$

$$\leq \frac{1}{2} E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^p \right)$$

$$+ \tilde{K}_3 E \left(\int_0^T \left\| \left(\sigma^m \left(s, v_s^m + \pi^m \left(u_s^{\mathrm{TW}} \right) \right) \right) \circ \Phi^m - \sigma \left(s, v_s + u_s^{\mathrm{TW}} \right) \right\|_{L_2^0}^2 ds \right)^{\frac{p}{2}}.$$

Bringing the first summand to the left-hand side of (29) this implies that

$$E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^p \right)$$

$$\leq 2 \left\| v_0^m - v_0 \right\|^p + 2 \left(\tilde{K}_1 + \tilde{K}_2 \frac{p-2}{p} \right) \int_0^T E \left(\sup_{s \leq t} \left\| v_s^m - v_s \right\|^p \right) dt$$

$$+ \tilde{K}_2 \frac{4}{p} E \left(\int_0^T R(s, v_s, m)^{\frac{p}{2}} ds \right)$$

$$+ 2 \tilde{K}_3 E \left(\int_0^T \left\| \left(\sigma^m \left(t, v_t^m + \pi^m \left(u_t^{\mathrm{TW}} \right) \right) \right) \circ \Phi^m - \sigma \left(t, v_t + u_t^{\mathrm{TW}} \right) \right\|_{L_2^0}^2 ds \right)^{\frac{p}{2}}.$$

We estimate the last term as before and obtain

$$E \left(\int_0^T \left\| \left(\sigma^m \left(s, v_s^m + \pi^m \left(u_s^{\mathrm{TW}} \right) \right) \right) \circ \Phi^m - \sigma \left(s, v_s + u_s^{\mathrm{TW}} \right) \right\|_{L_2^0}^2 ds \right)^{\frac{p}{2}}$$

$$\leq K E \left(\int_0^T \left\| v_s^m - v_s \right\|^2 + R(s, v_s, m) \, ds \right)^{\frac{p}{2}}$$

$$\leq \tilde{K} E \left(\int_0^T \sup_{s \leq t} \left\| v_s^m - v_s \right\|^p dt + \int_0^T R(s, v_s, m)^{\frac{p}{2}} ds \right).$$

Altogether we arrive at

$$E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^p \right) \leq 2 \left\| v_0^m - v_0 \right\|_2^p + \hat{K}_1 E \int_0^T R(t, v_t, m)^{\frac{p}{2}} dt$$

$$+ \hat{K}_2 \int_0^T E \left(\sup_{s \leq t} \left\| v_s^m - v_s \right\|^p \right) dt.$$

An application of Gronwall's lemma yields

$$E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^p \right) \leq K \left(\left\| v_0^m - v_0 \right\|^p + E \left(\sup_{t \leq T} R(t, v_t, m)^{\frac{p}{2}} \right) \right).$$

The sequence of continuous functions $f^m : [0, T] \to \mathbb{R}$

$$f^m(t) = \left(\left(\int_{L^m}^{\infty} + \int_{-\infty}^{-L^m} \right) v_t^2(x) + \left(w * |v_t| \right)^2 (x) \right.$$
$$\left. + \sigma \left(t, v_t + u_t^{\text{TW}} \right)^2 (x) + \left(\partial_x u_t^{\text{TW}}(x) \right)^2 dx \right)^{\frac{p}{2}}$$

is decreasing and converges pointwise to 0 since all the integrands are in $L^2(\mathbb{R})$. By Dini's theorem the convergence is uniform. This together with the facts that $\|\sigma(t, v_t)\|_2^2 \leq K(1 + \|v_t\|^2)$ and $E(\sup_{t \leq T} \|v_t\|^2) < \infty$ by Proposition 3, assumptions (i) and (ii), and dominated convergence implies that

$$E \left(\sup_{t \leq T} R(t, v_t, m)^{\frac{p}{2}} \right)$$
$$\leq K \left(\frac{1}{m^p} \left(E \left(\sup_{t \leq T} \|v_t\|^p \right) + \|\hat{u}_x\|^p + \|\hat{u}_{xx}\|^p \right) + E \left(\sup_{t \leq T} f^m(t) \right) \right.$$
$$+ E \left(\sup_{t \leq T} \left\| \sigma^m \left(t, \left(v_t + u_t^{\text{TW}} \right) \mathbb{1}_{(-L^m, L^m)} \right) \circ \Phi^m - \sigma \left(t, v_t + u_t^{\text{TW}} \right) \right\|^p \right)$$
$$+ E \left(\sup_{t \leq T} \left\| \sigma \left(t, v_t + u_t^{\text{TW}} \right) \right\|^p \right) \sup_k \sup_{x \in I_k^m} \left\| 2m\sqrt{Q} \mathbb{1}_{J_k^m} - q(x, \cdot) \right\|^p \right) \xrightarrow{m \to \infty} 0$$

and hence

$$E \left(\sup_{t \leq T} \left\| v_t^m - v_t \right\|^p \right) \xrightarrow{m \to \infty} 0.$$

Publisher's Note

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

References

1. Amari S. Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybern. 1977;27:77–87.
2. Wilson HR, Cowan JD. Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J. 1972;12:1–24.
3. Wilson HR, Cowan JD. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik. 1973;13:55–80.
4. Ermentrout GB, Terman DH. Mathematical foundations of neuroscience. New York: Springer; 2010.
5. Coombes S, beim Graben P, Potthast R, Wright J. Neural fields—theory and applications. Berlin: Springer; 2014.
6. Ermentrout GB. Neural networks as spatio-temporal pattern-forming systems. Rep Prog Phys. 1998;61:353–430.

7. Bressloff PC. Spatiotemporal dynamics of continuum neural fields. J Phys A. 2011;45:033001.
8. Bressloff PC. Waves in neural media. New York: Springer; 2014.
9. Ermentrout GB, McLeod JB. Existence and uniqueness of travelling waves for a neural network. Proc R Soc Edinb A. 1993;123:461–78.
10. Bressloff PC, Newby JM. Metastability in a stochastic neural network modeled as a velocity jump Markov process. SIAM J Appl Dyn Syst. 2013;12(3):1394–435.
11. Bressloff PC. Stochastic neural field theory and the system-size expansion. SIAM J Appl Math. 2009;70(5):1488–521.
12. Buice MA, Cowan JD. Field-theoretic approach to fluctuation effects in neural networks. Phys Rev E. 2007;75:051919.
13. Riedler MG, Buckwar E. Laws of large numbers and Langevin approximations for stochastic neural field equations. J Math Neurosci. 2013;3:1.
14. Riedler MG, Thieullen M, Wainrib G. Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models. Electron J Probab. 2012;17:55.
15. Buice MA, Cowan JD, Chow CC. Systematic fluctuation expansion for neural network activity equations. Neural Comput. 2010;22(2):377–426.
16. Lang E. Multiscale analysis of traveling waves in stochastic neural fields. SIAM J Appl Dyn Syst. 2016;15(3):1581–614.
17. Lang E, Stannat W. L^2-Stability of traveling wave solutions to nonlocal evolution equations. J Differ Equ. 2016;261(8):4275–97.
18. Lang E. Traveling waves in stochastic neural fields [PhD thesis]. Technische Universität Berlin; 2016. doi:10.14279/depositonce-5019.
19. Chen F. Travelling waves for a neural network. Electron J Differ Equ. 2003;2003:13.
20. Kallenberg O. Foundations of modern probability. 2nd ed. New York: Springer; 2002.
21. Stroock DW. An introduction to Markov processes. Berlin: Springer; 2014.
22. Ethier SN, Kurtz TG. Markov processes: characterization and convergence. New York: Wiley; 1986.
23. Prévôt C, Röckner M. A concise course on stochastic partial differential equations. Berlin: Springer; 2007.
24. Da Prato G, Zabczyk J. Stochastic equations in infinite dimensions. 2nd ed. Cambridge: Cambridge University Press; 2014.
25. Bressloff PC. Metastable states and quasicycles in a stochastic Wilson–Cowan model of neuronal population dynamics. Phys Rev E. 2010;82:051903.