Particle systems with a singular mean-field self-excitation. Application to neuronal networks.

François Delarue, James Inglis, Sylvain Rubenthaler, Etienne Tanré

PARTICLE SYSTEMS WITH A SINGULAR MEAN-FIELD SELF-EXCITATION. APPLICATION TO NEURONAL NETWORKS.

F. DELARUE, J. INGLIS, S. RUBENTHALER, E. TANRÉ

Abstract. We discuss the construction and approximation of solutions to a nonlinear McKean-Vlasov equation driven by a singular self-excitatory interaction of the mean-field type. Such an equation is intended to describe an infinite population of neurons which interact with one another. Each time a proportion of neurons 'spike', the whole network instantaneously receives an excitatory kick. The instantaneous nature of the excitation makes the system singular and prevents the application of standard results from the literature. Making use of the Skorohod M1 topology, we prove that, for the right notion of a 'physical' solution, the nonlinear equation can be approximated either by a finite particle system or by a delayed equation. As a by-product, we obtain the existence of 'synchronized' solutions, for which a macroscopic proportion of neurons may spike at the same time.

Keywords: McKean nonlinear diffusion process; counting process; propagation of chaos; integrate-and-fire network; Skorohod M1 topology; neuroscience.

1. Introduction

Recently several rigorous studies ([3, 4, 5, 8]) have been concerned with a mean-field equation modeling the behavior of a very large (infinite) network of interacting spiking neurons proposed in [14] (see also [1, 7, 10, 12] and references therein for other types of mean-field models motivated by neuroscience). As a nonlinear SDE in one-dimension the equation for the electrical potential X_t across any typical neuron in the network at time t takes the form

$$X_t = X_0 + \int_0^t b(X_s)ds + \alpha\mathbb{E}(M_t) + W_t - M_t, \qquad t \geqslant 0, \tag{1.1}$$

where $X_0 < 1$ almost surely, $(W_t)_{t\,\geqslant\,0}$ is a standard Brownian motion and b is a Lipschitz function of linear growth. Here α is a parameter in $(0, 1)$ and the process $M = (M_t)_{t\,\geqslant\,0}$ counts the number of times that $X = (X_t)_{t\,\geqslant\,0}$ reaches 1 before time t, so that it is integer-valued (see Section 2 for a precise description). The idea is that when X reaches the threshold 1, M instantly increases by 1 so that X is reset to a value below the threshold, and we say that the neuron has *spiked*. Throughout the article we will write $e(t) := \mathbb{E}(M_t)$.

Equation (1.1) is in fact nontrivial, since the form of the nonlinearity is not regular enough for the application of the standard McKean-Vlasov theory ([13, 17]). Indeed, the problem is that, on the infinitesimal level, the mean-field term in (1.1) reads as $e'(t) = [d/dt]\mathbb{E}(M_t)$, which is by no means regular with respect to the law of X_t. In [8], it is proven that $e'(t) = -(1/2)\partial_y p(t, 1)$, where $p(t, y)dy = \mathbb{P}(X_t \in dy)$ is the marginal density of X_t, which shows how singular the dependence of $e'(t)$ upon the law of X_t is. As such, most of the previous work studying this equation has been focused on the existence of a solution and its properties, bringing to light some non-trivial mechanisms.

The main point is that, for some choices of parameters (α too big for fixed X_0 concentrated close to the boundary), any solution to (1.1) must exhibit what has been described as a 'blow-up' in finite time. More precisely this means that $e'(t)$ (which is the mean-firing rate of the network at time t) must become infinite for some finite t. This was done in [3] by means of a PDE method. Interpreting (1.1) as a description of an infinite network of neurons, a blow-up is thus a time at which a proportion of all the neurons in the network spike at exactly the same time, which we refer to as a synchronization. Despite the interest in this phenomena, up until now it has been unclear how to continue a solution after a blow-up. On the other hand, in [8] it was shown by probabilistic arguments that for other choices of parameters (α small enough for fixed $X_0 = x_0$), (1.1) has a unique solution for all time which does not exhibit the blow-up phenomenon. These two complementary results are made precise in Theorems 2.3 and 2.4 below.

The aim of the present work is to provide further insight into this nonlinear equation by providing two ways of approximating (and moreover constructing) a solution. The first is via the natural particle system associated to (1.1), which describes the behavior of the finite network of neurons. In fact, the introduction of (1.1) in [14] is inspired from this finite dimensional system: it is there asserted that, when the size of the network becomes infinite, neurons become independent and evolve according to (1.1). However, the proof of this fact (which is a propagation of chaos result) is not given. The first of our main objectives is to fill this gap and to rigorously show that any weak limit of the particle system must be a solution to (1.1) (see Theorem 4.4). In particular, we show that the particle system converges to the solution of (1.1) whenever uniqueness holds, in which case propagation of chaos holds as well. Again, due to the irregularity and nature of the particle system, this result is in fact more difficult than it might appear. The second objective is to recover a similar result when approximating the self-interaction in (1.1) by delayed self-interactions (see Theorem 4.6). The motivation for considering the delayed equation (which is still nonlinear) is that it never exhibits a blow-up phenomenon, even with α close to 1, making it easier to handle (see Proposition 3.5).

In both cases, the strategy relies on two ingredients. First, we show that there exists a notion of 'physical' solutions to equation (1.1) for which spikes occur physically, in a 'sequential' way. The interesting feature of 'physical' solutions is that we allow the function $(e(t))_{t \geqslant 0}$ to be discontinuous, but characterize the size of any jumps in a precise way. Second, we show that there is a particularly suitable topology on the space of 'continus à droite avec limites à gauche' paths (càdlàg paths in acronym) for handling both approximations. The point is indeed to prove that the approximating families are tight for the so-called M1 Skorohod topology on the space of càdlàg paths, which is much less popular than the J1 topology, but which turns out to be very convenient for handling non-decreasing càdlàg processes such as the counting process $(M_t)_{t \geqslant 0}$.

As a significant by-product, the paper shows the existence of 'physical' solutions to (1.1) for which the function $(e(t))_{t \geqslant 0}$ may be discontinuous, but where we explicitly specify the size of any jump. This is a completely new fact in the literature, and is of real importance in neuroscience, as the size of the discontinuity of the function e indicates the proportion of neurons that synchronize at any time. The notion of 'physical' solutions together with the existence result thus permits the continuation of the solution after the synchronization and, therefore, allows the circumvention of the blow-up phenomenon experienced in [3] and [8]. In particular, this gives a rigorous framework for investigating the long time behavior of synchronization events, which is a fundamental question in neuroscience. This also raises the question of uniqueness of 'physical' solutions that experience a synchronization. We feel that it must be true, but the question is left open. We refrain from addressing this problem in the paper as it would require additional materials, including a careful discussion about the shape of the solution after some synchronization has occurred. We plan to go back to this question in a future work.

We would finally like to remark that variations of the model we present here could well be of interest in other contexts. In particular in a financial setting, a similar system has indeed been used to model the default rate of a large portfolio ([11, 16]) where a default occurs when a particle reaches a threshold. Our model is however more delicate than the one considered there, since the interactions we consider are more singular and produce the blow-up phenomenon that is not present in their setting.

The organization of the paper is as follows. In Section 2, we discuss the notion of 'physical' solutions to (1.1). The approximating systems are introduced in Section 3, in which we prove that both the associated particle system and the delayed equation are solvable. The main results are exposed in Section 4, where we also give a rough presentation of the M1 topology. Proofs are given in Section 5.

2. The nonlinear equation: background

The central nonlinear equation considered in this article is equation (1.1) where $X_0 < 1$ almost surely, $\alpha \in (0,1)$ and $(W_t)_{t \geqslant 0}$ is a standard Brownian motion, defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with respect to a filtration $(\mathcal{F}_t)_{t \geqslant 0}$ satisfying the usual conditions. The number of spikes of the equation until time t (inclusive) is given by

$$M_t := \sum_{k \geqslant 1} \mathbf{1}_{[0,t]}(\tau_k), \tag{2.1}$$

where the sequence of stopping times $(\tau_k)_{k \geqslant 0}$ is defined by $\tau_0 = 0$ and

$$\tau_k = \inf\big\{t > \tau_{k-1} : X_{t-} + \alpha \Delta e(t) \geqslant 1\big\}, \quad k \geqslant 1. \tag{2.2}$$

We have here used the notation $e(t) := \mathbb{E}(M_t)$, $t \geqslant 0$, and, for a given càdlàg function $f : [0,\infty) \to \mathbb{R}$, $\Delta f(t) := f(t) - f(t-)$, $t \geqslant 0$, which will be fixed throughout the article. The pair (2.1)–(2.2) is highly coupled as the definition of $(M_t)_{t \geqslant 0}$ relies on its own expectation. This asks for a careful description of the notion of a solution.

2.1. The right notion of a solution.

As noted above, the process $M = (M_t)_{t \geqslant 0}$ is intended to count the number of times $X = (X_t)_{t \geqslant 0}$ spikes before time t. At any time t such that $X_{t-} \geqslant 1 - \alpha\Delta e(t)$, the process M registers a new spike (pay attention that the presence of the '$-$' in the condition $X_{t-} \geqslant 1 - \alpha \Delta e(t)$ is crucial for ensuring the càdlàg property of the process). In the case when the mapping e is continuous at point t, the particle spikes if and only if $X_{t-} = 1$. It is then reset to 0 exactly after the spike, that is $X_t = 0$. Whenever e jumps at time t, the jump $\Delta e(t)$ must be of positive size so that, because of the self-interaction, X may spike even if $X_{t-} < 1$. Immediately after a spike occurs, i.e. when $X_{t-} \geqslant 1 - \alpha\Delta e(t)$, X_t is equal to $X_{t-} - 1 + \alpha\Delta e(t)$ and may be strictly positive: it is as if, at time t, the particle is first reset to 0 and then given a kick of magnitude $\alpha\Delta e(t) - (1 - X_{t-})$. Actually, such a description requires some precaution as the kick could force the particle to cross the barrier again at the same time t. This might happen if the kick $\alpha\Delta e(t)$ is greater than or equal to 1. Anyhow, such a phenomenon is expected to be 'non-physical': under the condition $\alpha < 1$, it does not make any sense to allow the system to spike twice (or more) at the same time. The argument for this is discussed at length below when making the connection with the finite particle system. In short, it says that physical spikes occur sequentially.

The fact that the jumps of the process $(M_t)_{t \geqslant 0}$ cannot exceed 1 provides some insight into the sequence of spiking times $(\tau_k)_{k \geqslant 1}$. First, given a solution satisfying $\mathbb{P}(\Delta M_t \leqslant 1) =$

1 for all $t \geqslant 0$, the sequence $(\tau_k)_{k \geqslant 1}$ must be (strictly) increasing: there is no way for two spiking times to coincide if labelled by different indices. Moreover, the sequence $(\tau_k)_{k \geqslant 1}$ cannot accumulate in finite time, as otherwise it would contradict the càdlàg nature of $(X_t)_{t \geqslant 0}$. Indeed, if $\tau_\infty := \lim_{k\to+\infty} \tau_k < +\infty$, then $X_{\tau_\infty -}$ is equal to both $\lim_{k\to+\infty} X_{\tau_k -}$ and $\lim_{k\to+\infty} X_{\tau_k}$, which gives a contradiction since $X_{\tau_k} = X_{\tau_k -} - 1 + \alpha\Delta e(\tau_k) < X_{\tau_k -} - 1 + \alpha$.

It also gives some insight into the jumps of the function e, summarized in the following proposition.

Proposition 2.1. *Assume that the pair* $(X_t, M_t)_{t \geqslant 0}$ *of càdlàg processes is such that*

(1) $(M_t)_{t \geqslant 0}$ *has integrable marginal distributions;*
(2) *for all* $t \geqslant 0$, $\mathbb{P}(\Delta M_t \leqslant 1) = 1$*;*
(3) $\mathbb{P}$*-almost surely,* (1.1), (2.1) *and* (2.2) *hold true.*

Then, for any time $t \geqslant 0$*, the jump* $\Delta e(t)$ *satisfies*

$$\Delta e(t) = \mathbb{P}\big(X_{t-} + \alpha\Delta e(t) \geqslant 1\big). \tag{2.3}$$

Proof. Given some time $t \geqslant 0$, a necessary and sufficient condition for registering a spike (that is to have $M_t - M_{t-} = 1$), is $X_{t-} + \alpha\Delta e(t) \geqslant 1$. Therefore, the probability of observing a spike is $\mathbb{P}(X_{t-} + \alpha\Delta e(t) \geqslant 1)$, which proves that $\Delta e(t) = \mathbb{P}(\Delta M_t = 1) = \mathbb{P}(X_{t-} + \alpha\Delta e(t) \geqslant 1)$. □

Unfortunately, equation (2.3) is not sufficient to characterize the size of the jumps. Indeed one can guess simple examples of distributions for the law of X_{t-} such that the equation (in η)

$$\eta = \mathbb{P}\big(X_{t-} + \alpha\eta \geqslant 1\big) \tag{2.4}$$

has several solutions. For instance, if X_{t-} has a uniform distribution on $[1-\alpha, 1]$, then the equation is satisfied for every $\eta \in [0, 1]$. In order to determine which solution to (2.4) characterizes the size of the jump, we refer again to what a physical solution to (1.1) must be. In (2.4), $\alpha\eta$ is intended to stand for the magnitude of the kick felt by the particle. The idea behind this is that we consider all the $\omega \in \Omega$ for which the kick is large enough to make the particle cross the barrier. To put it differently there must be enough mass near 1 in the distribution of X_{t-} to 'absorb' the particle from $1-\alpha\eta$ to 1. Implicitly, this requires that there is no gap in the mass. If, for some $\eta' < \eta$, the probability $\mathbb{P}(X_{t-} + \alpha\eta' \geqslant 1)$ is (strictly) less than η', then the kick is not strong enough to absorb the particle when at distance $\alpha\eta'$ from 1. This suggests that, physically, the magnitude of the kick must be given as the largest magnitude for which 'absorption' can occur. Therefore, a reasonable characterization for $\Delta e(t)$ is

$$\begin{aligned}\Delta e(t) &= \sup\big\{\eta \geqslant 0 : \forall \eta' \leqslant \eta, \mathbb{P}\big(X_{t-} + \alpha\eta' \geqslant 1\big) \geqslant \eta'\big\} \\ &= \inf\big\{\eta \geqslant 0 : \mathbb{P}\big(X_{t-} + \alpha\eta \geqslant 1\big) < \eta\big\}.\end{aligned}$$

At this stage of the paper, we will keep this characterization as a necessary condition for a 'physical' solution to (1.1). Again, we will justify this choice in a more detailed way below. With this in mind, we thus make the following precise definition.

Definition 2.2. *We call a (physical) solution to* (1.1) *a pair* $(X_t, M_t)_{t \geqslant 0}$ *of càdlàg adapted processes such that*

(1) $(M_t)_{t \geqslant 0}$ *has integrable marginal distributions;*
(2) *for all* $t \geqslant 0$, $\mathbb{P}(\Delta M_t \leqslant 1) = 1$*;*
(3) $\mathbb{P}$*-almost surely,* (1.1), (2.1) *and* (2.2) *hold true;*

(4) *the discontinuity points of the function* $e : [0, +\infty) \ni t \mapsto \mathbb{E}(M_t)$ *satisfy*

$$\Delta e(t) = \inf\{\eta \geqslant 0 : \mathbb{P}(X_{t-} + \alpha\eta \geqslant 1) < \eta\}.$$

We underline that a physical solution satisfies (2.3), but that we need (4) to characterize the size of the jump $\Delta e(t)$ (and hence avoid non-physical phenomenon as discussed in the above examples – see also the paragraph 'Non-physical solutions' on page 10). A sufficient condition for a physical solution is given in Proposition 2.7 below.

2.2. **Standing assumptions and related literature.** We will make the following two assumptions throughout the article.

Assumption 1 (Globally Lipschitz drift)**.** *The drift* $b : (-\infty, 1] \to \mathbb{R}$ *is Lipschitz continuous such that* $|b(x) - b(y)| \leqslant K|x - y|$*, for all* $x, y \in (-\infty, 1]$.

Assumption 2 (Initial condition)**.** *The initial condition* $X_0 \in (-\infty, 1 - \varepsilon_0]$ *almost surely for some* $\varepsilon_0 > 0$ *and* $X_0 \in L^p(\Omega)$ *for any* $p \geqslant 1$.

The assumption that the distribution of the initial condition has support in $(-\infty, 1-\varepsilon_0]$, rather than in $(-\infty, 1)$, is a slight simplification. It is motivated by technical reasons that will be specified in the core of the proofs.

As mentioned in the Introduction, the existence and uniqueness of a solution to (1.1) is a nontrivial problem. It is addressed in [3] and [8], as well as [5], but in the smaller class of pairs $(X_t, M_t)_{t \geqslant 0}$ for which the mapping $e : [0, +\infty) \ni t \mapsto \mathbb{E}(M_t)$ is continuous (which renders the conditions (2) and (4) in Definition 2.2 useless). The following two theorems summarize the results in [3] and [8] relevant for the present study.

Theorem 2.3 ([3])**.** *For every* $\alpha \in (0, 1)$*, it is possible to find an initial condition* X_0*, such that there is no global solution (global meaning defined on the entire interval* $[0, +\infty)$*) where the mapping e is continuously differentiable. Equivalently, it is possible to find an initial condition* X_0 *such that any solution to* (1.1) *experiences a blow-up, in the sense that* $e'(t) = +\infty$ *for some* $t \geqslant 0$.

Theorem 2.4 ([8])**.** *For all initial conditions* $X_0 = x_0 < 1$*, it is possible to find an* $\alpha_0(x_0) \in (0, 1)$ *such that, whenever* $\alpha \in (0, \alpha_0)$*, equation* (1.1) *possesses a unique (pathwise and thus in law) global solution such that the mapping e is continuously differentiable.*

So far existence and uniqueness within the framework of Definition 2.2 are completely open problems. As mentioned in the Introduction, the purpose of this paper is to provide a general compactness method for approximating solutions to (1.1), and as a by-product prove the existence of a solution according to Definition 2.2, for which the map e may be discontinuous. Inspired by the earlier paper [8], we will make use of the following reformulation of equation (1.1).

Remark 2.5 (Reformulation)**.** *It will be very convenient throughout the article to sometimes work instead with a reformulated version of* (1.1)*, which describes the evolution of the process* $Z = (Z_t)_{t \geqslant 0}$*, defined simply by*

$$Z_t := X_t + M_t, \quad t \geqslant 0.$$

It is then plain to see that M_t *can be completely expressed in terms of* $(Z_s)_{0 \leqslant s \leqslant t}$ *as*

$$M_t = \big\lfloor \big(\sup_{0 \leqslant s \leqslant t} Z_s \big)_+ \big\rfloor = \sup_{0 \leqslant s \leqslant t} \lfloor (Z_s)_+ \rfloor, \quad t \geqslant 0, \tag{2.5}$$

where $\lfloor x \rfloor$ *and* $(x)_+$ *indicate the integer part of* x *and* $\max\{x, 0\}$ *respectively, for any* $x \in \mathbb{R}$*. Indeed, as* $X_t < 1$ *and* $(M_s)_{s \geqslant 0}$ *is non-descreasing, it is clear that* $M_t \geqslant \lfloor (\sup_{0 \leqslant s \leqslant t} Z_s)_+ \rfloor$.

Conversely, for a given $k \geqslant 0$ such that $\tau_k \leqslant t < \tau_{k+1}$, $X_{\tau_k} \geqslant 0$, so that $M_{\tau_k} = k \leqslant Z_{\tau_k}$, which completes the proof of the equality. The reformulated version of (1.1) *is then given by*

$$Z_t = Z_0 + \int_0^t b(Z_s - M_s)ds + \alpha\mathbb{E}(M_t) + W_t, \qquad t \geqslant 0, \tag{2.6}$$

where $Z_0(= X_0) < 1$, and $(M_t)_{t \geqslant 0}$ is defined by (2.5). *One big advantage of any solution $Z = (Z_t)_{t \geqslant 0}$ to* (2.6) *over a solution $X = (X_t)_{t \geqslant 0}$ to* (1.1) *is that discontinuity points of Z are dictated by those of the deterministic mapping $e : [0, +\infty) \ni t \mapsto \mathbb{E}(M_t)$ only.*

Conversely, given a solution $(Z'_t, M'_t)_{t \geqslant 0}$ to (2.6) *and* (2.5), *we recover a (possibly non-physical) solution to the original equation* (1.1) *by setting $X'_t = Z'_t - M'_t$.*

2.3. **A criterion for a physical solution.** The following lemma is an adaptation of [8, Prop. 3.1]. The proof is left to the reader. It relies on Gronwall's lemma and (2.5)–(2.6).

Lemma 2.6. *Consider a pair $(X_t, M_t)_{t \geqslant 0}$ of càdlàg adapted processes such that (1) and (3) hold in Definition 2.2. Then it holds that $\mathbb{E}[\sup_{t\in[0,T]} |Z_t|^p] < +\infty$ for any $p \geqslant 1, T > 0$.*

Next we present a useful application. The reader may skip the proof on a first reading.

Proposition 2.7. *Assume that the pair $(X_t, M_t)_{t \geqslant 0}$ of càdlàg processes is such that (1), (2) and (3) hold in Definition 2.2. Assume also that, at any discontinuity time $t \geqslant 0$ of the mapping $e : [0, +\infty) \ni s \mapsto \mathbb{E}(M_s)$, it holds that*

$$\forall \eta \leqslant \Delta e(t), \quad \mathbb{P}\big(X_{t-} \geqslant 1 - \alpha\eta\big) \geqslant \eta. \tag{2.7}$$

Then the pair $(X_t, M_t)_{t \geqslant 0}$ is a physical solution.

Proof. In order to prove that $(X_t, M_t)_{t \geqslant 0}$ is a physical solution, we must check that, for any $t \geqslant 0$, there exists a decreasing sequence $(\eta_n)_{n \geqslant 1}$, with $\Delta e(t)$ as its limit, such that

$$\mathbb{P}\big(X_{t-} + \alpha\eta_n \geqslant 1\big) < \eta_n, \quad n \geqslant 1.$$

Together with (2.7), this indeed implies (4) in Definition 2.2.

We argue by contradiction. Fix $t \geqslant 0$ and assume that there exists $\eta_0 > \Delta e(t)$ such that

$$\forall \eta \in (\Delta e(t), \eta_0], \quad \mathbb{P}\big(X_{t-} \geqslant 1 - \alpha\eta\big) \geqslant \eta.$$

Then, recalling from (2.3) that $\Delta e(t) = \mathbb{P}(X_{t-} + \alpha\Delta e(t) \geqslant 1)$, we deduce that

$$\begin{aligned} \forall \eta \in (0, \eta_0], \quad &\mathbb{P}\big(1 - \alpha\eta \leqslant X_{t-} < 1 - \alpha\Delta e(t)\big) \\ &= \mathbb{P}\big(X_{t-} \geqslant 1 - \alpha\eta\big) - \mathbb{P}\big(X_{t-} + \alpha\Delta e(t) \geqslant 1\big) \geqslant \eta - \Delta e(t). \end{aligned}$$

Notice that, on the event $\{1 - \alpha\eta \leqslant X_{t-} < 1 - \alpha\Delta e(t)\}$, $\Delta M_t = 0$, so that $X_t = X_{t-} + \alpha\Delta e(t)$. Therefore, with $\eta' = \eta - \Delta e(t)$, we obtain, $\forall \eta' \in (0, \eta_0 - \Delta e(t)]$,

$$\begin{aligned} \mathbb{P}\big(X_t \geqslant 1 - \alpha\eta'\big) &\geqslant \mathbb{P}\big(1 > X_t \geqslant 1 - \alpha\eta', X_{t-} + \alpha\Delta e(t) = X_t\big) \\ &= \mathbb{P}\big(1 - \alpha\eta \leqslant X_{t-} < 1 - \alpha\Delta e(t), X_{t-} + \alpha\Delta e(t) = X_t\big) \\ &= \mathbb{P}\big(1 - \alpha\eta \leqslant X_{t-} < 1 - \alpha\Delta e(t)\big) \geqslant \eta'. \end{aligned} \tag{2.8}$$

To simplify, we let $\eta'_0 := \eta_0 - \Delta e(t) > 0$.

The strategy is then to prove that $\liminf_{h\downarrow 0}[e(t + h) - e(t)] > 0$, which will contradict the right-continuity of e. To do so, we use a stochastic comparison argument. For some small $h \in (0, 1)$, we indeed have

$$e(t + h) - e(t) \geqslant \mathbb{P}\big(\exists s \in (t, t + h] : Y_{s-} \geqslant 1\big),$$

where $(Y_s)_{s\in[t,t+h]}$ solves the equation

$$Y_s = X_t + \int_t^s b(Y_u)du + \alpha(e(s)-e(t)) + W_s - W_t, \quad t \leqslant s \leqslant t+h.$$

Indeed, as long as $(X_s)_{s\in[t,t+h]}$ does not spike, it coincides with $(Y_s)_{s\in[t,t+h]}$. In particular, if $M_{s-} - M_t = 0$ and $Y_{s-} \geqslant 1$, then $X_{s-} \geqslant 1$ and thus $M_s - M_t = 1$. Therefore, $\{\exists s \in (t,t+h] : Y_{s-} \geqslant 1\} \subset \{M_{t+h} - M_t \geqslant 1\}$. We then get, for some constant C (the value of which is allowed to increase from line to line, but will remain independent of h and α),

$$e(t+h)-e(t) \geqslant \mathbb{P}\Big(X_t - Ch\big(1+\sup_{s\in[t,t+h]}|Y_s|\big)+\sup_{s\in[t,t+h]}\Big[\alpha(e(s)-e(t))+W_s-W_t\Big] \geqslant 1\Big). \quad (2.9)$$

By a standard application of Gronwall's lemma (recalling that h can be chosen small enough so that $e(s)-e(t) \leqslant 1$ for all $s \in [t,t+h]$ as e is right continuous),

$$|Y_s| \leqslant C\big(1+|X_t| + \sup_{s\in[t,t+h]}|W_s - W_t|\big), \quad t \leqslant s \leqslant t+h,$$

so that $\mathbb{P}(\sup_{s\in[t,t+h]}|Y_s| \geqslant 3C, |X_t| \leqslant 1) \leqslant Ch$. Since $X_t \geqslant 0$ implies $|X_t| \leqslant 1$, we obtain $\mathbb{P}(\sup_{s\in[t,t+h]}|Y_s| \geqslant 3C, X_t \geqslant 0) \leqslant Ch$, so that, by (2.9),

$$\begin{aligned}
&e(t+h)-e(t)\\
&\geqslant \mathbb{P}\Big(X_t - Ch(1+3C) + \sup_{s\in[t,t+h]}\Big[\alpha(e(s)-e(t))+W_s-W_t\Big] \geqslant 1,\ \sup_{s\in[t,t+h]}|Y_s| \leqslant 3C\Big)\\
&\geqslant \mathbb{P}\Big(X_t - Ch + \sup_{s\in[t,t+h]}\Big[\alpha(e(s)-e(t))+W_s-W_t\Big] \geqslant 1,\ X_t \geqslant 0\Big) - Ch,
\end{aligned}$$

where we have adjusted C.

Assume now that there exists $c \geqslant 0$ such that $\alpha(e(r)-e(t)) \geqslant c\sqrt{r-t}$ for all $r \in [t,t+h]$, which is (at least) true with $c=0$. Then, by the above bound, we get

$$e(t+h)-e(t) \geqslant \int_0^{+\infty} \mathbb{P}\big(X_t - Ch + u \geqslant 1,\ X_t \geqslant 0\big)d\nu(u) - Ch, \quad (2.10)$$

where ν denotes the law of the supremum of $c\sqrt{s}+W_s$ over $s \in [0,h]$. Notice that $u \leqslant 1$ and $X_t - Ch + u \geqslant 1$ implies $X_t \geqslant 0$. Assuming without any loss of generality that $\alpha\eta_0' = \alpha(\eta_0 - \Delta e(t)) \leqslant 1$, we deduce from (2.8) that

$$e(t+h)-e(t) \geqslant \int_{Ch}^{\alpha\eta_0'} \frac{u-Ch}{\alpha}d\nu(u) - Ch \geqslant \frac{1}{\alpha}\int_{Ch}^{\alpha\eta_0'} ud\nu(u) - 2\frac{Ch}{\alpha}, \quad (2.11)$$

the constant C being independent of c. Recall now that $c\sqrt{h} \leqslant \alpha(e(t+h)-e(t)) \leqslant \alpha\eta_0'/2$ for h small enough as e is right continuous. Therefore, using the fact that the tail of $\sup_{s\in[0,h]} W_s$ is Gaussian, we obtain

$$\begin{aligned}
\int_{\alpha\eta_0'}^{+\infty} ud\nu(u) &= \mathbb{E}\Big[\sup_{s\in[0,h]}(c\sqrt{s}+W_s)\mathbf{1}_{\{\sup_{s\in[0,h]}(c\sqrt{s}+W_s) \geqslant \alpha\eta_0'\}}\Big]\\
&\leqslant \mathbb{E}\Big[\sup_{s\in[0,h]}\Big(\frac{\alpha\eta_0'}{2}+W_s\Big)\mathbf{1}_{\{\sup_{s\in[0,h]}W_s \geqslant \alpha\eta_0'/2\}}\Big] \leqslant Ch.
\end{aligned}$$

Moreover, quite obviously, $\int_0^{Ch} u d\nu(u) \leqslant Ch$. Finally, by (2.11), with C independent of c,

$$\begin{aligned}\alpha(e(t+h)-e(t)) \geqslant \int_0^{+\infty} u d\nu(u) - Ch &= \mathbb{E}\Big[\sup_{s\in[0,h]}\big(c\sqrt{s}+W_s\big)\Big] - Ch \\ &= h^{1/2}\bigg(\mathbb{E}\Big[\sup_{s\in[0,1]}\big(c\sqrt{s}+W_s\big)\Big] - Ch^{1/2}\bigg),\end{aligned}$$

the last equality following from Brownian scaling. A similar inequality can be proved for any $r\in[t,t+h]$, that is $\alpha(e(r)-e(t)) \geqslant f(c)\sqrt{r-t}$, where

$$f(c) = \mathbb{E}\Big[\sup_{s\in[0,1]}\big(c\sqrt{s}+W_s\big)\Big] - C\sqrt{h}.$$

We deduce that, if the inequality $\alpha(e(r)-e(t)) \geqslant c\sqrt{r-t}$ holds for all $r\in[t,t+h]$, then $\alpha(e(r)-e(t)) \geqslant f(c)\sqrt{r-t}$ for all $r\in[t,t+h]$. Letting $c_0=0$ and $c_{n+1}=f(c_n)$ for all $n\geqslant 0$, we deduce that $\alpha(e(r)-e(t)) \geqslant c_n\sqrt{r-t}$ for all $r\in[t,t+h]$ and all $n\geqslant 0$.

Clearly, we can choose h small enough so that $c_1>0=c_0$. Since f is non-decreasing, we deduce that the sequence $(c_n)_{n\geqslant 0}$ is non-decreasing. As e is locally bounded, the sequence has a finite limit c^*. Then, as f is obviously Lipschitz continuous, we have $c^*=f(c^*)$, that is,

$$c^* = \mathbb{E}\Big[\sup_{s\in[0,1]}(c^*\sqrt{s}+W_s)\Big] - C\sqrt{h} = \mathbb{E}\Big[\sup_{s\in[0,1]}(c^*\sqrt{s}+W_s) - (c^*+W_1)\Big] + c^* - C\sqrt{h}.$$

Therefore, by time reversal,

$$\begin{aligned}C\sqrt{h} &= \mathbb{E}\Big[\sup_{s\in[0,1]}(c^*\sqrt{s}+W_s) - (c^*+W_1)\Big] \\ &\geqslant \mathbb{E}\Big[\sup_{s\in[0,1]}\big(c^*(s-1)+W_s-W_1\big)\Big] \geqslant \mathbb{E}\Big[\sup_{s\in[0,1]}(-c^*s+W_s)\Big],\end{aligned}$$

which says that c^* must be large when h is small. In particular, we can assume h small enough so that $c^*\geqslant 1$. Then,

$$C\sqrt{h} \geqslant \mathbb{E}\Big[\sup_{s\in[0,(c^*)^{-2}]}(-c^*s+W_s)\Big] = \frac{1}{c^*}\mathbb{E}\Big[\sup_{s\in[0,1]}(-s+W_s)\Big],$$

which proves that, for h small enough, $c^*\sqrt{h}\geqslant\beta$, for some constant $\beta>0$. This implies $\liminf_{h\downarrow 0}[e(t+h)-e(t)] \geqslant \beta/\alpha$, which is a contradiction. □

3. Two candidates for approximate solutions

In this section we present two alternative systems, which are candidates to be approximations of the nonlinear equation (1.1).

3.1. **The particle system approximation.**

As noted above, one of the main motivations for studying (1.1) is the idea that it describes the behavior of a very large number of interacting spiking neurons in a fully connected network, each evolving according to the classical noisy integrate-and-fire model. More precisely, this idea translates into the fact that we would like (1.1) to describe the behavior of the particle system

$$\begin{cases} X_t^{i,N} = X_0^{i,N} + \int_0^t b\big(X_s^{i,N}\big)ds + \dfrac{\alpha}{N}\displaystyle\sum_{j=1}^N M_t^{j,N} + W_t^i - M_t^{i,N} \\ X_0^{i,N} \overset{d}{=} X_0 \text{ independent and identically distributed,} \end{cases} \tag{3.1}$$

for $i \in \{1, \dots, N\}$ and $t \geqslant 0$ when N is large. Here $(X_t^{i,N})_{t\,\geqslant\,0}$ represents the electrical potential of the ith neuron, X_0 satisfies standing Assumption 2, $(W_t^i)_{t\,\geqslant\,0}$ are independent standard Brownian motions, and now $(M_t^{i,N})_{t\,\geqslant\,0}$ is the process that counts the number of times the ith neuron has 'spiked' up until time t. Precisely, we define for $i \in \{1, \dots, N\}$ and $t \geqslant 0$

$$M_t^{i,N} := \sum_{k\,\geqslant\,1} \mathbf{1}_{[0,t]}(\tau_k^{i,N}),$$

where $\tau_0^{i,N} = 0$ and

$$\tau_k^{i,N} := \inf\Big\{ t > \tau_{k-1}^{i,N} : X_{t-}^{i,N} + \frac{\alpha}{N}\sum_{j=1}^N \big(M_t^{j,N} - M_{t-}^{j,N}\big) \geqslant 1 \Big\}, \quad k \geqslant 1, \tag{3.2}$$

which should be compared with (2.2) (and which is as involved as (2.2) since the definitions of $\tau_k^{i,N}$ and $M_t^{i,N}$ are fully coupled). The idea is that the system spikes if one of the particles reaches the threshold 1, but this can cause other particles to instantaneously spike through the empirical mean type interaction. However, exactly as in the previous section, where we defined a solution to (1.1), we must be careful about what we mean by a 'physical' solution to the particle system (3.1). This is because there may in fact exist multiple solutions to (3.1) and (3.2) (see section on 'non-physical' solutions below). The 'physical' solution we will identify is in fact the one in which we require the instantaneous spikes induced at a spike time to be ordered in a natural way, the first spike occuring when one of the particles hits the barrier. See below for a precise description. At time $t = \tau_k^{i,N}$, $X_t^{i,N} = X_{t-}^{i,N} - 1 + (\alpha/N)\sum_{j=1}^N (M_t^{j,N} - M_{t-}^{j,N})$. Again, it should also be noted that the presence of the '$-$' in $X_{t-}^{i,N}$ in (3.2) ensures that $M^{i,N}$ and $X^{i,N}$ are càdlàg.

Anyway, the point is that the system (3.1) is mathematically equivalent to the one used by Ostojic, Brunel and Hakim in [14] to describe the behavior of a finite network of neurons, and that (1.1) is a good guess as to what happens in the limit as $N \to \infty$. Indeed, the extremely well developed theory of mean-field/McKean-Vlasov equations provides many rigorous results about when an individual particle in a system that interacts through an empirical mean becomes independent in the limit as $N \to \infty$, and then behaves according to a distribution dependent (McKean-Vlasov) limit equation. However, despite the use of such a result in [14], we argue that the current situation is quite different to any that has been previously studied in the literature due to the nature of the nonlinearity. Thus, one of the aims of this paper is to provide a complete rigorous proof of this convergence.

Remark 3.1. *The reader may argue that it makes more sense physically to replace the interaction term* $N^{-1}\sum_{j=1}^N M_t^{j,N}$ *in* (3.1) *by* $(N-1)^{-1}\sum_{j\neq i} M_t^{j,N}$*, so that if a single neuron spikes at time* t*, it is reset from the threshold* 1 *to* 0 *(rather than to* α/N*). However, we choose to keep the stated interaction term since it renders the analysis notationally simpler, while remaining mathematically equivalent in the limit* $N \to \infty$*.*

Remark 3.2 (Reformulation)**.** *Following Remark 2.5, it will be convenient to reformulate the particle system* (3.1) *in terms of the processes* $(Z_t^{i,N})_{t\geqslant 0}$*, defined by*

$$Z_t^{i,N} := X_t^{i,N} + M_t^{i,N}, \qquad t \geqslant 0.$$

Then, similarly to (2.6)*, the reformulated system is given by*

$$\begin{aligned} Z_t^{i,N} &= Z_0^{i,N} + \int_0^t b\left(Z_s^{i,N} - M_s^{i,N}\right) ds + \frac{\alpha}{N}\sum_{j=1}^N M_t^{j,N} + W_t^i, \\ M_t^{i,N} &= \lfloor \big(\sup_{s\in[0,t]} Z_s^{i,N}\big)_+ \rfloor = \sup_{s\in[0,t]} \lfloor \big(Z_s^{i,N}\big)_+ \rfloor, \end{aligned} \tag{3.3}$$

for all $t \geqslant 0$ *and* $i \in \{1,\dots,N\}$*, where* $Z_0^{i,N} = X_0^{i,N} \overset{d}{=} X_0$ *are i.i.d and* X_0 *satisfies Assumption 2. We will refer to* (3.3) *as the* Z*-particle system (and the original system* (3.1) *as the* X*-particle system).*

Notation. In the sequel, we will use the convenient notation

$$\bar{e}^N(t) = \frac{1}{N}\sum_{i=1}^N M_t^{i,N}, \quad t \geqslant 0. \tag{3.4}$$

We will also often omit the superscript N in the notations $X^{i,N}$, $Z^{i,N}$, $M^{i,N}$ and $\tau_k^{i,N}$ for simplicity. When no confusion is possible, we thus write X^i, Z^i, M^i and τ_k^i instead.

Non-physical solutions. As mentioned already, the particle system defined above is not well-posed, as it may admit a large number of solutions when α is close to 1.

Actually, uniqueness may fail for several reasons. A first way for constructing different solutions is to allow one particle to admit several spikes at the same time. Indeed, consider the Z-system (3.3) with $b \equiv 0$ and suppose that α has the form $\alpha = 1 - 1/(2m)$, for some integer $m \geqslant 1$. Suppose moreover that, at some time t, it holds that (the system being initialized at $Z_0^i = 0$, $i \in \{1,\dots,N\}$)

$$\forall i \in \{1,\dots,N\}, \quad Z_{t-}^i = 1 - \delta_i, \ M_{t-}^i = 0,$$

with $\delta_1 = 0$ and $\delta_i \in ((i-2)/(4N),(i-1)/(4N))$ for $i = 2,\dots,N$, which, by the support theorem for Brownian motion, happens with positive probability. Then, the system is to spike at time t since the first particle reaches the barrier, but the spike procedure may be arbitrarily chosen. Indeed, setting arbitrarily $M_t^i = \ell \geqslant 1$, for all $i \in \{1,\dots,N\}$, the equation for Z^i gives

$$Z_t^i = 1 - \delta_i + \frac{\alpha}{N}\sum_{i=1}^N \ell = \ell + 1 - \frac{\ell}{2m} - \delta_i \in \left[\ell + \frac{3}{4} - \frac{\ell}{2m}, \ell + 1\right),$$

for all $i \in \{1,\dots,N\}$. Then, if $\ell/m \leqslant 1$, $\lfloor(\sup_{0\leqslant s\leqslant t} Z_s^i)_+\rfloor = \ell = M_t^i$, so that the second relationship in (3.3) is satisfied. Since $\ell \leqslant m$ is arbitrary, the system (3.3) clearly does not possess a unique solution. According to the discussion below, cases where $\ell \geqslant 2$ will be considered as *non-physical*.

We give here a second example where uniqueness fails even if the property $\mathbb{P}(\Delta M_t^i \leqslant 1) = 1$ is fulfilled. We present it with $N = 3$ particles but it could be generalized in an obvious way. Suppose that at some (random) time t, $M_{t-}^1 = M_{t-}^2 = M_{t-}^3 = 0$, $Z_{t-}^1 = 1$, $Z_{t-}^2, Z_{t-}^3 \in (1 - 2\alpha/3, 1 - \alpha/3)$. We can then make explicit two solutions to (3.3): a first one where only particle 1 spikes, that is $M_t^1 = 1$ and $M_t^2 = M_t^3 = 0$, and a second one where all the particles spike at time t, that is $M_t^1 = M_t^2 = M_t^3 = 1$. In this example, the second case will be said *non-physical*. Intuitively, particle 1 is indeed intended to spike 'first'. After particle 1 has spiked, particles 2 and 3 are both strictly below 1, which should prevent them from spiking immediately.

Physical solutions. In view of the above discussion, the problem is that the ordering of the spike cascade is not determined i.e. how spiking neurons instantaneously cause others to spike. We now argue that in fact there is a natural way of ordering this cascade, which then leads to unique 'physical' solutions.

To this end, consider the X-particle system (3.1) and define the set

$$\Gamma_0 := \{i \in \{1,\dots,N\} : X^i_{t-} = 1\}.$$

We say t is a spike time when $\Gamma_0 \neq \emptyset$. At a spike time t, it is certain that all the neurons in Γ_0 spike. It then makes sense to introduce a second *time axis*, called the *cascade time axis at spike time* t, and to say that, along this axis, neurons in Γ_0 are the first ones to spike.

Then it is natural to determine exactly which other neurons spike *given that those in* Γ_0 *have already spiked.* Since the system says that all the other neurons should feel the effect of the ones in Γ_0 spiking by receiving a kick to their potential of size $\alpha|\Gamma_0|/N$, this in turn means that all the neurons in the set

$$\Gamma_1 := \left\{ i \in \{1,\dots,N\}\backslash\Gamma_0 : X^i_{t-} + \alpha\frac{|\Gamma_0|}{N} \geqslant 1 \right\},$$

now have potentials that are instantaneously above the threshold, and so should also spike. Thus we are now sure that all the neurons in $\Gamma_0 \cup \Gamma_1$ spike at t. Along the *cascade time axis at time* t, the neurons in Γ_1 are said to spike after the neurons in Γ_0. Similarly, it is then natural to determine which other neurons spike, *given that those in* $\Gamma_0 \cup \Gamma_1$ *have already spiked.* According to the definition of the system, this is exactly those in the set

$$\Gamma_2 := \left\{ i \in \{1,\dots,N\}\backslash\Gamma_0\cup\Gamma_1 : X^i_{t-} + \alpha\frac{|\Gamma_0\cup\Gamma_1|}{N} \geqslant 1 \right\}.$$

By defining sequentially for general $k \in \mathbb{N}_0$

$$\Gamma_{k+1} := \left\{ i \in \{1,\dots,N\}\backslash\Gamma_0\cup\cdots\cup\Gamma_k : X^i_{t-} + \alpha\frac{|\Gamma_0\cup\cdots\cup\Gamma_k|}{N} \geqslant 1 \right\}, \tag{3.5}$$

the natural cascade is continued in this way until $\Gamma_l = \emptyset$ for some $l \in \{1,\dots,N\}$. Note that this must happen, since by definition $\Gamma_N = \emptyset$ (if $\Gamma_N \neq \emptyset$, all the sets $\Gamma_0,\dots,\Gamma_{N-1}$ contain at least one element; since all of them are disjoint, we obtain a contradiction). Along the cascade time axis at time t, neurons in Γ_{k+1} ($k+1 < \ell$) spike after neurons in $\Gamma_0 \cup \cdots \cup \Gamma_k$. We can then define $\Gamma := \bigcup_{0\leqslant k\leqslant N-1} \Gamma_k$, which is exactly the set of all neurons that spike at time t, according the natural ordering of the spike cascade (see also [6]). Having determined this, it is then straightforward to perform the final update of all the neurons in the network by setting

$$X^i_t = X^i_{t-} + \frac{\alpha|\Gamma|}{N} \text{ if } i \notin \Gamma, \quad X^i_t = X^i_{t-} + \frac{\alpha|\Gamma|}{N} - 1 \text{ if } i \in \Gamma. \tag{3.6}$$

Note that now $X^i_t < 1$ for all $i \in \{1,\dots,N\}$. Indeed, if $i \notin \Gamma$, then i must be such that

$$X^i_{t-} + \frac{\alpha|\Gamma|}{N} < 1 \quad \Rightarrow \quad X^i_t < 1.$$

On the other hand, if $i \in \Gamma$ then, since $|\Gamma| \leqslant N$ and $\alpha < 1$,

$$X_t^i = X_{t-}^i + \frac{\alpha|\Gamma|}{N} - 1 \leqslant X_{t-}^i + \alpha - 1 < X_{t-}^i \leqslant 1.$$

The above idea is completed by the following lemma.

Lemma 3.3. *There exists a unique solution to the particle system* (3.1) *such that, whenever t is a spike time, the entire system jumps according to*

$$X_t^i = X_{t-}^i + \frac{\alpha|\Gamma|}{N} \text{ if } i \notin \Gamma, \quad X_t^i = X_{t-}^i + \frac{\alpha|\Gamma|}{N} - 1 \text{ if } i \in \Gamma.$$

where $\Gamma \subset \{1, \dots, N\}$ is as above. Such a solution will be known as a 'physical' solution.

Proof. It is clear that in between spike times of the system there is no problem of uniqueness (since the particles only interact at spike times). Therefore, since we have specified a unique jumping procedure, any solution must be unique.

The proof of existence is more challenging. The issue is to prove that spike times of the system do not accumulate. We feel it is more convenient to give it at this stage of the paper, but the reader may skip ahead on a first reading.

We in fact prove the existence of a solution to the associated Z-system (3.3) (this is completely equivalent to the existence of a solution to the original system (3.1)) with the given spike cascade. For any $1 \leqslant i \leqslant N$, we define $(Y_t^{1,i})_{t \geqslant 0}$ as the solution of the SDE

$$Y_t^{1,i} = Z_0^i + \int_0^t b(Y_s^{1,i}) ds + W_t^i, \quad t \geqslant 0.$$

We set $\tau^{1,i} = \inf\{t \geqslant 0 : Y_t^{1,i} \geqslant 1\}$, $1 \leqslant i \leqslant N$. Clearly, we have $0 < \tau^{1,i} < \infty$ (a.s.), so that $0 < \inf_{1 \leqslant i \leqslant N}(\tau^{1,i}) < \infty$ (a.s.). For $t \in [0, \tau^1)$, with $\tau^1 = \inf_{1 \leqslant i \leqslant N}(\tau^{1,i})$, we set

$$Z_t^i = Y_t^{1,i}, \quad M_t^i = 0, \quad 0 \leqslant t < \tau^1, \quad 1 \leqslant i \leqslant N.$$

At time τ^1, there exists $i^1 \in \{1, \dots, N\}$ such that $\tau^1 = \tau^{1,i^1}$. We then denote by $\Gamma^{(1)}$ the set of particles that spike at τ^1 according to the physical procedure summarized in (3.6) (pay attention that $\Gamma^{(1)}$ stands for the Γ in (3.6) and not for Γ_1: the positions of the indices are different). Then, according to the cascade, we know that the kick that the particle Z^i receives at time τ^1 is $\alpha|\Gamma^{(1)}|/N$, so that

$$Z_{\tau^1}^i = Y_{\tau^1}^{1,i} + \alpha\frac{|\Gamma^{(1)}|}{N}, \quad 1 \leqslant i \leqslant N.$$

For a coordinate $i \in \Gamma^{(1)}$, it holds $M_{\tau^1}^i = 1$. Since $Z_{\tau^1-}^i \leqslant 1$ and the kick received by i is less than α, it holds $Z_{\tau^1}^i \leqslant 1 + \alpha < M_{\tau^1}^i + 1$. Moreover, we must also have $Z_{\tau^1}^i \geqslant 1$ so that $M_{\tau^1}^i \leqslant Z_{\tau^1}^i < M_{\tau^1}^i + 1$, that is $\lfloor Z_{\tau^1}^i \rfloor = M_{\tau^1}^i$. Since $Z_{\tau^1}^i = \sup_{s \in [0,\tau^1]} Z_s^i = (\sup_{s \in [0,\tau^1]} Z_s^i)_+$, we deduce $M_{\tau^1}^i = \lfloor (\sup_{s \in [0,\tau^1]} Z_s^i)_+ \rfloor$. On the other hand, for a coordinate $i \notin \Gamma^{(1)}$, it holds that $M_{\tau^1}^i = M_{\tau^1-}^i = 0$, and $\sup_{s \in [0,\tau^1]} Z_s^i < 1$, so that $M_{\tau^1}^i = \lfloor (\sup_{s \in [0,\tau^1]} Z_s^i)_+ \rfloor$ as well.

For any $1 \leqslant i \leqslant N$, we then define $(Y_t^{2,i})_{t \geqslant 0}$ as the solution of the SDE

$$Y_t^{2,i} = Z_{\tau^1}^i + \int_{\tau^1}^t b(Y_s^{2,i} - M_{\tau^1}^i) ds + (W_t^i - W_{\tau^1}^i), \quad t \geqslant \tau^1.$$

Define then $\tau^{2,i} = \inf\{t \geqslant \tau^1 : Y_t^{2,i} \geqslant M_{\tau^1}^i + 1\}$, $1 \leqslant i \leqslant N$. Since $Z_{\tau^1}^i < M_{\tau^1}^i + 1$, we have $\tau^{2,i} > \tau^1$. Then, with $\tau^2 = \inf_{1 \leqslant i \leqslant N}(\tau^{2,i})$, we set

$$Z_t^i = Y_t^{2,i}, \quad \tau^1 < t < \tau^2.$$

The spike procedure at time τ^2 is defined according to the process summarized in (3.6), the set of particles jumping at τ^2 being denoted by $\Gamma^{(2)}$. By iteration, we build an increasing sequence of stopping times $(\tau^k)_{k\geqslant 0}$ (with $\tau^0 = 0$) such that

$$Z^i_t = Z^i_0 + \int_0^t b\big(Z^i_s - M^i_{\tau^k}\big)ds + \frac{\alpha}{N}\sum_{j=1}^N M^j_{\tau^k} + W^i_t, \tag{3.7}$$

for $1 \leqslant i \leqslant N$ and $\tau^k < t < \tau^{k+1}$, $k \geqslant 0$, with $\tau^{k+1} = \inf_{1\leqslant i\leqslant N} \tau^{k+1,i}$, where $\tau^{k+1,i} = \inf\{t > \tau^k : Z^i_{t-} \geqslant M^i_{\tau^k} + 1\}$, $1 \leqslant i \leqslant N$. The set of particles that jump at τ^k is then denoted by $\Gamma^{(k)}$. With such a construction, we notice that

$$M^i_t = \big\lfloor \big(\sup_{s\in[0,t]} Z^i_s\big)_+ \big\rfloor, \tag{3.8}$$

for $1 \leqslant i \leqslant N$ and $t \leqslant \tau^k$, for any $k \geqslant 1$. Indeed, at time τ^k, the proof is the same as at time τ^1. At any time $t \in (\tau^k, \tau^{k+1})$, the equality follows from the fact that $Z^i_t < M^i_{\tau^k} + 1$.

To finish with the proof of existence, we prove that $\tau^k \to +\infty$ as $k \to +\infty$. Noting from (3.8) that the drift part in (3.7) can be bounded by $|b(Z^i_s - M^i_s)| \leqslant C(1 + \sup_{0\leqslant r\leqslant s} |Z^i_r|)$, for some constant $C \geqslant 0$, and taking the empirical mean over $i \in \{1, \dots, N\}$, we deduce that, for $t \leqslant \tau^k$, for $k \geqslant 1$,

$$\frac{1-\alpha}{N}\sum_{i=1}^N \sup_{s\in[0,t]} |Z^i_s| \leqslant \frac{1}{N}\sum_{i=1}^N \big(|X^i_0| + \sup_{s\in[0,t]} |W^i_s|\big) + C\int_0^t \big(1 + \frac{1}{N}\sum_{i=1}^N \sup_{r\in[0,s]} |Z^i_r|\big)ds.$$

We deduce from Gronwall's lemma that, for any $1 \leqslant i \leqslant N$ and $t \leqslant \tau^k$, for $k \geqslant 1$,

$$\frac{1}{N}\sum_{i=1}^N \sup_{s\in[0,t]} |Z^i_s| \leqslant C\exp(Ct)\left[t + \frac{1}{N}\sum_{i=1}^N \big(|X^i_0| + \sup_{s\in[0,t]} |W^i_s|\big)\right],$$

for a possibly new value of C. By (3.8), the same bound holds for $N^{-1}\sum_{i=1}^N M^i_t$. Going back to (3.7) and using Gronwall's lemma again, we deduce, that for any $T > 0$, there exists a constant $C_T \geqslant 0$ such that, for any $1 \leqslant i \leqslant N$ and $t \leqslant \tau^k \wedge T$, for $k \geqslant 1$,

$$\sup_{s\in[0,t]} |Z^i_s| \leqslant C_T\left[1 + |X^i_0| + \sup_{s\in[0,t]} |W^i_s| + \frac{1}{N}\sum_{j=1}^N \big(|X^j_0| + \sup_{s\in[0,t]} |W^j_s|\big)\right]. \tag{3.9}$$

Again, by (3.8), the same bound holds for M^i_t. In particular, if $\tau^k \leqslant T$,

$$\sup_{i\in\{1,\dots,N\}} M^i_{\tau^k} \leqslant C_T\Big(1 + \sup_{i\in\{1,\dots,N\}} |X^i_0| + \sup_{i\in\{1,\dots,N\}} \sup_{s\in[0,T]} |W^i_s|\Big).$$

Applying the above inequality with Nk instead of k, we notice that $\sup_{i\in\{1,\dots,N\}} M^i_{\tau^{Nk}}$ is larger than k (as, at time τ^{Nk}, there have been Nk spikes in the system, so that at least one of the particles has spiked at least k times). Therefore, if $\tau^{Nk} \leqslant T$, then

$$k \leqslant C_T\Big(1 + \sup_{i\in\{1,\dots,N\}} |X^i_0| + \sup_{i\in\{1,\dots,N\}} \sup_{s\in[0,T]} |W^i_s|\Big).$$

In particular, the sequence $(\tau^k)_{k\geqslant 1}$ cannot have a finite limit T, as otherwise, passing to the limit, we would get $\sup_{i\in\{1,\dots,N\}} |X^i_0| + \sup_{i\in\{1,\dots,N\}} \sup_{s\in[0,T]} |W^i_s| = +\infty$. □

Physical solutions to the particle system satisfy a discrete version of (4) in Definition 2.2, which motivates the notion of physical solutions to the original equation (1.1):

Proposition 3.4. *For a physical solution as in Lemma 3.3, it holds that, for all $t \geqslant 0$,*

$$N\Delta\bar{e}^N(t) = \sum_{i=1}^{N}\big(M_t^i - M_{t-}^i\big) = \inf\Big\{k \in \{0,\dots,N\} : \sum_{i=1}^{N} \mathbf{1}_{\{X_{t-}^i \geqslant 1-\frac{\alpha k}{N}\}} \leqslant k\Big\}. \qquad (3.10)$$

Proof. We use the description of a physical solution. We know that the left-hand side in (3.10) is equal to $|\Gamma|$ by definition of Γ in page 11. Clearly $|\Gamma| = \sum_{k=0}^{l} |\Gamma_k|$ where l is the largest integer such that $\Gamma_l \neq \emptyset$. Then, for $k \in \{|\Gamma_0 \cup \Gamma_1 \cup \cdots \cup \Gamma_{j-1}|, \dots, |\Gamma_0 \cup \Gamma_1 \cup \cdots \cup \Gamma_j| - 1\}$ and $j \in \{1,\dots,l\}$ (or $k \in \{0,\dots,|\Gamma_0|-1\}$ if $j=0$),

$$\sum_{i=1}^{N} \mathbf{1}_{\{X_{t-}^i \geqslant 1-\frac{\alpha k}{N}\}} \geqslant \sum_{i=1}^{N} \mathbf{1}_{\{X_{t-}^i \geqslant 1-\frac{\alpha|\Gamma_0\cup\Gamma_1\cup\cdots\cup\Gamma_{j-1}|}{N}\}} = \sum_{\ell=0}^{j} |\Gamma_\ell| > k,$$

the equality following from (3.5), proving that the right-hand side in (3.10) is greater than or equal to $|\Gamma|$. On the other hand, by construction, $\sum_{i=1}^{N} \mathbf{1}_{\{X_{t-}^i \geqslant 1-\frac{\alpha}{N}|\Gamma|\}} \leqslant |\Gamma| + \sum_{i\notin\Gamma} \mathbf{1}_{\{X_{t-}^i \geqslant 1-\frac{\alpha}{N}|\Gamma|\}}$, but the last term in the right hand-side is zero. This shows that the right-hand side in (3.10) is less than or equal to $|\Gamma|$. □

3.2. **The system with delays.** In this section we introduce a second approximation of the nonlinear system (1.1), by introducing delays. As we will see below (Proposition 3.5), the advantage of doing this is that the resulting system has a global in time solution, for which the mean-firing rate e' remains finite for any value of the parameter α and initial condition (recall that this is in contrast to the system without delays (1.1) which may 'blow-up' in finite time for some parameter values: see Theorem 2.3).

The point is that the introduction of a delay prevents a macroscopic proportion of the neurons all spiking at the same time. Intuitively, this is because, even if other neurons are close enough to the threshold to be induced to spike as a result of the first neuron spiking, this will occur only after a positive amount of time.

Given that the delayed system does not experience a blow-up and has a global solution (see below), part of our work is dedicated to the analysis of the solutions when the delay converges to zero (see Subsection 4.2). However, the purpose of this current subsection is simply to introduce the system with delays and to check well-posedness. To this end, let $\delta > 0$ and consider the equation

$$X_t^\delta = X_0 + \int_0^t b(X_s^\delta)ds + \alpha e_\delta(t) + W_t - M_t^\delta, \quad t \geqslant 0. \qquad (3.11)$$

Here, similarly to above, M_t^δ counts the number of time $(X_s^\delta)_{s\in[0,t]}$ reaches the threshold. Precisely, we write

$$M_t^\delta = \sum_{k \geqslant 1} \mathbf{1}_{[0,t]}(\tau_k^\delta), \qquad (3.12)$$

where $\tau_0^\delta = 0$ and, for $k \geqslant 1$,

$$\tau_k^\delta = \inf\left\{t > \tau_{k-1}^\delta : X_{t-}^\delta + \alpha\Delta e_\delta(t) \geqslant 1\right\}, \quad \text{with } e_\delta(t) := \begin{cases} 0 & \text{if } t \leqslant \delta \\ \mathbb{E}(M_{t-\delta}^\delta) & \text{if } t > \delta. \end{cases} \qquad (3.13)$$

We write the equation in this way, even though, as the following proposition shows, the delay guarantees that there is a unique solution to (3.11) such that e_δ is always continuously differentiable (so that $\Delta e_\delta \equiv 0$). This makes any notion of 'physical' solutions irrelevant for the delayed equation. As in Section 2, we take $\alpha \in (0,1)$ and $(W_t)_{t\geqslant 0}$ a standard real-valued Brownian motion, and we assume that Assumptions 1 and 2 are in force.

Proposition 3.5. *Let $T > 0$ and $\alpha \in (0,1)$. Then there exists a unique càdlàg process $(X^\delta_t, M^\delta_t)_{t\in[0,T]}$, such that $(M^\delta_t)_{t\geqslant 0}$ has integrable marginal distributions, satisfying* (3.11) *and* (3.12)*. The resulting map e_δ is continuously differentiable.*

Proof. Step 1: Solution on $[0,\delta]$. For $t \leqslant \delta$ (3.11) reads

$$X^\delta_t = X_0 + \int_0^t b(X^\delta_s)ds + W_t - M^\delta_t. \tag{3.14}$$

Clearly, this has a unique strong solution for $t \leqslant \delta$ (there is no difficult nonlinear term). Moreover, by [8, Proposition 4.5], we have that $[0,\delta] \ni t \mapsto \mathbb{E}(M^\delta_t)$ is continuously differentiable, and moreover

$$\frac{d}{dt}\mathbb{E}(M^\delta_t) = -\frac{1}{2}\int_0^t \partial_y p^\delta_0(t-s,1)\frac{d}{ds}\mathbb{E}(M^\delta_s)ds - \frac{1}{2}\partial_y p^\delta_{X_0}(t,1), \qquad t \in [0,\delta], \tag{3.15}$$

where, for a random variable χ_0, $p^\delta_{\chi_0}$ represents the density of the process X^δ killed at 1 with $X^\delta_0 = \chi_0$, namely $p^\delta_{\chi_0}(t,y)dy := \mathbb{P}(X^\delta_t \in dy, \sup_{s\in[0,t]} X^\delta_s < 1 | X^\delta_0 = \chi_0)$. Note that the shift by s that is required in [8, Proposition 4.5] is not necessary here, as (3.14) is time homogeneous. By continuous differentiability, $0 \leqslant \sup_{t\in[0,\delta]}(d/dt)\mathbb{E}(M^\delta_t) < +\infty$.

Step 2: Solution on $[0,2\delta]$. For $t \leqslant 2\delta$ (3.11) reads

$$X^\delta_t = X_0 + \int_0^t b(X^\delta_s)ds + \alpha e_\delta(t) + W_t - M^\delta_t, \tag{3.16}$$

where $e_\delta(t) = 0$ on $[0,\delta]$ and $\mathbb{E}(M^\delta_{t-\delta})$ on $[\delta, 2\delta]$.

We now claim that $[0,2\delta] \ni t \mapsto e_\delta(t)$ is continuously differentiable. This is clearly the case on $[0,\delta]$, and on $[\delta,2\delta]$ by Step 1. It remains to check that it is also true at $t = \delta$, i.e. $\lim_{t\downarrow\delta} e'_\delta(t) = \lim_{t\downarrow 0}[d/dt]\mathbb{E}(M^\delta_t) = 0$. This follows from (3.15), since $\lim_{t\downarrow 0} \partial_y p^\delta_{\chi_0}(t,1) = 0$ for χ_0 satisfying Assumption 2, see [8, Lemma 4.2]. In particular, $0 \leqslant \sup_{t\in[0,2\delta]}(d/dt)\mathbb{E}(M^\delta_t) < +\infty$. By [8, Section 3], equation (3.16) has a unique strong solution.

Moreover, since $[0,2\delta] \ni t \mapsto e_\delta(t)$ is continuously differentiable, we can apply [8, Proposition 4.5] on this new interval to see that $[0,2\delta] \ni t \mapsto \mathbb{E}(M^\delta_t)$ is continuously differentiable.

Step 3: Solution on $[0,3\delta]$*:* We replicate Step 2 by proving that $[0,3\delta] \ni t \mapsto e_\delta(t)$ is continuously differentiable. Indeed, $[\delta,3\delta] \ni t \mapsto e_\delta(t)$ is continuously differentiable by Step 2, and $e_\delta(t)$ is equal to 0 on $[0,\delta]$, but, as already noted, $[d/dt]|_{t=\delta^+} e_\delta(t) = 0$, so that the 'join' is continuously differentiable.

Conclusion: Let $T > 0$. One may iterate this procedure up until Step $\lceil T/\delta \rceil$. This will yield the fact that there exists a unique strong solution to the system given by (3.11) and (3.12) up until time T, such that $[0,T] \ni t \mapsto \mathbb{E}(M^\delta_t)$ is continuously differentiable. □

4. Results

Given the setup described in the previous sections, we are now in a position to present our main results. The objective is to pass to the limit as $N \to \infty$ in the particle system described in Subsection 3.1 and as $\delta \to 0$ in the delayed equation described in Subsection 3.2, deriving as a by-product a new global in time solvability result for the original model (1.1) including solutions that blow up.

With this in mind, the thrust of the paper is to identify a very convenient topology for tackling both problems. Basically, the strategy is to make use of the so-called M1 Skorohod topology, which is different from the more famous J1 topology and which turns out to be much more adapted to the problem at hand. The reason is that relative compactness for the M1 topology is indeed easily checked for sets of monotone càdlàg functions, which exactly fits the nature of the process $(M_t)_{t\geqslant 0}$ in (1.1).

4.1. **The M1 topology.** We first supply the reader with some reminders about the M1 topology. For a complete overview, we refer to the original paper by Skorohod [15] and to the monograph by Whitt [19]. We denote by $\hat{\mathcal{D}}([0,T],\mathbb{R})$ the space of càdlàg functions from $[0,T]$ to $\mathbb{R}$ that are left-continuous at time T [1]. For a function $f\in\hat{\mathcal{D}}([0,T],\mathbb{R})$, we denote by $\mathcal{G}_f$ the completed graph of f i.e.

$$\mathcal{G}_f := \{(x,t)\in\mathbb{R}\times[0,T] : x\in[f(t-),f(t)]\},$$

where $[f(t-),f(t)]$ stands for the non-ordered segment between $f(t-)$ and $f(t)$ ($f(t-)$ could be bigger than $f(t)$). We define an order on $\mathcal{G}_f$ in the following way: for $(x_1,t_1),(x_2,t_2)\in\mathcal{G}_f$, we say that $(x_1,t_1)\leqslant(x_2,t_2)$ if either $t_1<t_2$, or $t_1=t_2$ and $|f(t_1-)-x_1|\leqslant|f(t_1-)-x_2|$. In other words this is the natural order when the graph $\mathcal{G}_f$ is traced out from left to right. We then define a *parametric representation* of $\mathcal{G}_f$ as being a continuous function (u,r) that maps $[0,T]$ onto $\mathcal{G}_f$ that is non-decreasing with respect to the order on $\mathcal{G}_f$ defined above i.e.

$$(u,r):[0,T]\ni t\mapsto(u(t),r(t))\in\mathcal{G}_f,$$

where $u\in\mathcal{C}([0,T],\mathbb{R}), r\in\mathcal{C}([0,T],\mathbb{R})$. We define $\mathcal{R}_f$ as the set of all parametric representations of $\mathcal{G}_f$. A parametric representation of $\mathcal{G}_f$ is thus a way of tracing it out 'without going back on oneself' with respect to the natural order of the graph.

For $f_1,f_2\in\hat{\mathcal{D}}([0,T],\mathbb{R})$ we finally define the M1 distance between them as

$$d_{M_1}(f_1,f_2) := \inf_{\substack{(u_j,r_j)\in\mathcal{R}_{f_j}\\ j=1,2}}\{\|u_1-u_2\|\vee\|r_1-r_2\|\},$$

where $\|\cdot\|$ is the usual supremum norm on $\mathcal{C}([0,T],\mathbb{R})$. In order to characterize the convergence in M1, we define for $f\in\hat{\mathcal{D}}([0,T],\mathbb{R}), t\in[0,T]$ and $\delta>0$,

$$w_T(f,t,\delta) := \sup_{0\vee(t-\delta)\leqslant t_1<t_2<t_3\leqslant T\wedge(t+\delta)}\Big\|f(t_2)-[f(t_1),f(t_3)]\Big\| \tag{4.1}$$

where

$$\Big\|f(t_2)-[f(t_1),f(t_3)]\Big\| = \inf_{\theta\in[0,1]}\Big|\theta f(t_1)+(1-\theta)f(t_3)-f(t_2)\Big|$$

is the distance between $f(t_2)$ and the set $[f(t_1),f(t_3)]$. In particular, if a function $f\in\hat{\mathcal{D}}([0,T],\mathbb{R})$ is monotone (non-increasing or non-decreasing), then $w_T(f,t,\delta)=0$.

From [19, Theorems 12.5.1, 12.4.1, 12.12.2], we have:

Theorem 4.1. *A sequence of functions* $(f_n)_{n\geqslant 1}\subset\hat{\mathcal{D}}([0,T],\mathbb{R})$ *converges to some* $f\in\hat{\mathcal{D}}([0,T],\mathbb{R})$ *in the M1 topology if and only if* $f_n(t)\to f(t)$ *for each* t *in a dense subset of full Lebesgue measure of* $[0,T]$ *that includes* 0 *and* T, *and*

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{t\in[0,T]} w_T(f_n,t,\delta)=0.$$

[1] The condition forcing elements in $\hat{\mathcal{D}}([0,T],\mathbb{R})$ to be left-continuous at the terminal time is implicitly done in Whitt [19]: in Theorem 12.2.2 therein, the piecewise constant functions used for approximating càdlàg functions on $[0,T]$ are precisely assumed to be continuous at terminal time T.

Theorem 4.2. *Suppose that the sequence* $(f_n)_{n\geqslant 1} \subset \hat{\mathcal{D}}([0,T],\mathbb{R})$ *converges to some* $f \in \hat{\mathcal{D}}([0,T],\mathbb{R})$ *in the M1 topology. Then for all points* $t \in [0,T]$ *at which* f *is continuous it holds that*

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{s\in[0\vee(t-\delta),T\wedge(t+\delta)]}|f_n(s)-f(s)| = 0.$$

In particular, if f *is continuous on the entire interval* $[0,T]$*, then the convergence of* $(f_n)_{n\geqslant 1}$ *to* f *is uniform. Moreover, if* f_n *is monotone for each* n*, then* $f_n \to f$ *in M1 if and only if* $f_n(t) \to f(t)$ *for all* t *in a dense subset of full Lebesgue measure of* $[0,T]$ *including* 0 *and* T*.*

Theorem 4.3. *A subset* A *of* $\hat{\mathcal{D}}([0,T],\mathbb{R})$ *has compact closure in the M1 topology if and only if* $\sup_{f\in A}\|f\| < \infty$ *and*

$$\lim_{\delta\to 0}\sup_{f\in A}\Big\{\Big(\sup_{t\in[0,T]} w_T(f,t,\delta)\Big)\vee v_T(f,0,\delta)\vee v_T(f,T,\delta)\Big\} = 0$$

where $v_T(f,t,\delta) := \sup_{0\vee(t-\delta)\,\leqslant\, t_1\,\leqslant\, t_2\,\leqslant\, T\wedge(t+\delta)} \big|f(t_1)-f(t_2)\big|$.

Finally, we mention that $\hat{\mathcal{D}}([0,T],\mathbb{R})$, endowed with M1, is Polish, and that the Borel σ-field coincides with the σ-field generated by the evaluation mappings (see [18, Page 8]). This guarantees that the law of a process over $\hat{\mathcal{D}}([0,T],\mathbb{R})$, endowed with M1, is characterized by its finite-dimensional distributions. The Polish property renders the Skorohod representation theorem licit, both on $\hat{\mathcal{D}}([0,T],\mathbb{R})$ and on $\mathcal{P}(\hat{\mathcal{D}}([0,T],\mathbb{R}))$, which is defined as the set of probability measures on $\hat{\mathcal{D}}([0,T],\mathbb{R})$ (see [2, Theorem 6.7] and [9, Chapter III, Theorem 1.7]). It also renders the Prohorov theorem licit (see [2, Chapter 1, Section 5]): we let the reader derive the tightness criterion from Theorem 4.3 (see [19, Theorem 12.12.3]).

4.2. **Existence of weak solutions with simultaneous spikes.** The purpose of this Section is to state two results showing that the existence of a solution to (1.1) can be deduced by extracting weakly convergent subsequences either along the distributions of the particle systems (as the number of particles N tends to $+\infty$), or along the distributions of the delayed systems (as the delay tends to 0).

Theorem 4.4. *Given* $T > 0$ *and the (physical) solution* $((Z_t^{i,N})_{t\in[0,T]})_{i=1,\dots,N}$ *to the particle system* (3.3)*, consider the extended system*

$$\widetilde{Z}_t^{i,N} := \begin{cases} Z_t^{i,N}, & \text{if } t \leqslant T, \\ W_t^i - W_T^i + Z_T^{i,N}, & \text{if } t \in (T, T+1], \end{cases} \tag{4.2}$$

for $i \in \{i,\dots,N\}$*. Define the empirical measure* $\bar{\mu}_N := \frac{1}{N}\sum_{i=1}^N \mathrm{Dirac}(\widetilde{Z}^{i,N})$*, which reads as a random variable with values in* $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$*. Then, denoting by* Π_N *the law of* $\bar{\mu}_N$*, the family* $(\Pi_N)_{N\geqslant 1}$ *is tight in* $\mathcal{P}(\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R})))$ *endowed with the topology of weak convergence inherited from the M1 topology.*

Moreover, for any weak limit Π_∞*, for* Π_∞*-almost every measure* $\mu \in \mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$*, the canonical process* $(z_t)_{t\in[0,T+1]}$ *on* $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$ *generates, under* μ*, a physical solution to* (2.6)*, and hence to* (1.1)*, up until time* T *i.e.*

(1) *under* μ*,* z_0 *is distributed according to the law of* X_0*;*

(2) *under* μ, $(z_t - z_0 - \int_0^t b(z_s - m_s)ds - \alpha\langle \mu, m_t\rangle)_{t\in[0,T)}$ *is a Brownian motion, where* $m_t = \lfloor (\sup_{0 \leqslant s \leqslant t} z_s)_+ \rfloor$ *and* $\langle \mu, m_t\rangle$ *denotes the expectation of* m_t *under* μ *(*$\langle \cdot, \cdot \rangle$ *is the duality bracket between a probability measure and a measurable function);*

(3) *under* μ*, (1), (2) and (4) in Definition 2.2 are fulfilled.*

Remark 4.5. *In the usual terminology of SDEs, the solution to* (2.6) *as given by Theorem 4.4 is weak as the Brownian motion is part of the solution.*

The extension of the Z*-processes in* (4.2) *to the interval* $(T, T+1]$ *permits to get for free uniform bounds on the modulus of continuity of the particles at the final time* $T+1$*, which is a requirement for tightness for the M1 topology (see Lemma 5.4 below). As recalled in Footnote 1 on Page 16, elements of* $\hat{\mathcal{D}}([0,S],\mathbb{R})$*, for* $S > 0$*, are assumed to be left-continuous at* S*, which requires bounds on the modulus of continuity at* S *when addressing questions of convergence or compactness. With* $S = T$*, we see that* $\widetilde{Z}^{i,N}$ *(or equivalently* $Z^{i,N}$*) may not be continuous at* S*, but with* $S = T+1$*,* $\widetilde{Z}^{i,N}$ *is obviously continuous at* S*, with the modulus of continuity at* S *being controlled by the Brownian part only. Although rather arbitrary, the reason why we include Brownian oscillations in the definition of* $\widetilde{Z}^{i,N}$ *on* $[T, T+1]$ *will be made clear in Lemma 5.6 below. Basically, noise is needed to guarantee that the counting process* $[0,+\infty) \ni t \mapsto M_t$ *in* (2.6) *is stable in law under perturbation of the dynamics. Moreover, this avoids introducing any distinction between the dynamics on* $[0,T]$ *and on* $[T, T+1]$ *in the application of the stability property.*

By Cantor's diagonal argument, it is possible to construct a solution on a sequence of intervals $([0, T_n))_{n \geqslant 1}$*, with* $T_n \to +\infty$*, and thus to prove existence in infinite time.*

We have a similar result for the delayed equation:

Theorem 4.6. *Given* $T > 0$ *and the family of solutions* $((X_t^\delta)_{t\in[0,T]})_{\delta\in(0,1)}$ *to the delayed equation* (3.11)*, with* $X_0^\delta = X_0$ *satisfying Assumption 2, consider the family of extended paths*

$$\widetilde{Z}_t^\delta := \begin{cases} X_t^\delta + M_t^\delta, & \text{if } t \leqslant T, \\ W_t - W_T + \widetilde{Z}_T^\delta, & \text{if } t \in (T, T+1]. \end{cases} \tag{4.3}$$

Define by μ^δ *the law of* $(\widetilde{Z}_t^\delta)_{t\in[0,T+1]}$ *on* $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$*. Then, the family* $(\mu^\delta)_{\delta\in(0,1)}$ *is tight in* $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$ *endowed with the topology of weak convergence inherited from the M1 topology. Moreover, under any weak limit* μ *as* δ *tends to* 0*, the canonical process* $(z_t)_{t\in[0,T+1]}$ *on* $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$ *generates a physical solution to* (2.6)*, and hence to* (1.1)*, until time* T*, in the sense that (1), (2) and (3) in Theorem 4.4 hold true.*

4.3. **Convergence of the particle system and propagation of chaos.** An important corollary to Theorem 4.4 is that when we have uniqueness for equation (1.1), we also have propagation of chaos for the particle system $((Z_t^{i,N}, M_t^{i,N})_{t\in[0,T]})_{i=1,\dots,N}$ given by (3.3):

Theorem 4.7. *Assume that there exists a unique physical solution* $(X_t, M_t)_{t \geqslant 0}$ *to* (1.1) *and denote by* $(Z_t, M_t)_{t \geqslant 0}$ *the reformulated solution (as defined in Remark 2.5). Denote also by* J *the (at most countable) set of discontinuity points of the function* $[0,+\infty) \ni t \mapsto \mathbb{E}(M_t)$*. Then, for any* $S \in [0,+\infty) \setminus J$ *and any* $k \in \{1,\dots,N\}$*,*

$$\big((\hat{Z}_s^{1,N}, \hat{M}_s^{1,N}), \dots, (\hat{Z}_s^{k,N}, \hat{M}_s^{k,N})\big)_{s\in[0,S]} \Rightarrow \mathbb{P}^{\otimes k}_{(Z_s, \hat{M}_s)_{s\in[0,S]}} \quad \text{as } N \to +\infty, \tag{4.4}$$

on the space $[\hat{\mathcal{D}}([0,S],\mathbb{R}) \times \hat{\mathcal{D}}([0,S],\mathbb{R})]^k$ *equipped with the product topology induced by the M1 topology, where*

$$\big(\hat{Z}_s^{i,N}, \hat{M}_s^{i,N}\big) = \begin{cases} \big(Z_s^{i,N}, M_s^{i,N}\big) & \text{if } s < S, \\ \big(Z_{S-}^{i,N}, M_{S-}^{i,N}\big) & \text{if } s = S. \end{cases}$$

Here, for a random variable X, $\mathbb{P}_X$ *stands for the law of* X, *and* $\Rightarrow$ *indicates weak convergence. Moreover, as* $N \to +\infty$, *on* $\hat{\mathcal{D}}([0,S],\mathbb{R})$ *equipped with the M1 topology,*

$$\left(\frac{1}{N}\sum_{i=1}^N \hat{M}_s^{i,N}\right)_{s\in[0,S]} \to \big(\mathbb{E}(M_s)\big)_{s\in[0,S]} \quad \textit{in probability.} \tag{4.5}$$

Remark 4.8. (i) *In the case when the unique solution* $(Z_t, M_t)_{t\geqslant 0}$ *has a continuous firing function* $e : [0,+\infty) \ni t \mapsto \mathbb{E}(M_t)$, *then the process* $(Z_t)_{t\geqslant 0}$ *has continuous paths. Such a situation is guaranteed for some initial conditions and values of* α *by Theorem 2.4. Then, by Theorem 4.2, the weak convergence of the law of the particles* $\hat{Z}^{1,N}, \dots, \hat{Z}^{k,N}$ *in* (4.4) *holds on the space* $[\hat{\mathcal{D}}([0,S],\mathbb{R})]^k$ *equipped with the product uniform topology. Similarly, in such a case, the convergence in* (4.5) *holds on* $\hat{\mathcal{D}}([0,S],\mathbb{R})$ *equipped with the uniform topology.*

(ii) *In* (4.4), *we could replace* $(Z_s)_{s\in[0,S]}$ *by* $(\hat{Z}_s)_{s\in[0,S]}$ *but this would be useless as* Z *is continuous at point* S *for any realization of the randomness: Since* $S \in J$, *the (deterministic) jump function* $[0,+\infty) \ni t \mapsto \mathbb{E}(M_t)$ *is continuous at point* S. *Similarly, we could replace* $(\hat{M}_s)_{s\in[0,S]}$ *by* $(M_s)_{s\in[0,S]}$, *by noticing that, with probability* 1 *under* $\mathbb{P}$, $M_S = M_{S-}$, *but this would be slightly abusive as the paths of* $(M_s)_{s\in[0,S]}$ *are in* $\hat{\mathcal{D}}([0,S],\mathbb{R})$ *with probability* 1 *only (and not for all realizations of the randomness).*

(iii) *Convergence of the* X*-particles in* (4.4) *follows from the relationship* $X_t^{i,N} = Z_t^{i,N} - M_t^{i,N}$, *for* $i \in \{1,\dots,N\}$. *However, since addition may not be continuous for the M1 topology (see Chapter 12 in* [19]*), we cannot deduce the convergence of the* X*-particles on* $\hat{\mathcal{D}}([0,S],\mathbb{R})$. *By Theorem 4.2, the best we can say is that, for any* $k \in \{1,\dots,N\}$, *any* $\ell \geqslant 1$ *and any* $t_1,\dots,t_\ell \notin J$, *the law of the random vector* $((X_{t_j}^{i,N}, M_{t_j}^{i,N})_{i\in\{1,\dots,k\}})_{j\in\{1,\dots,\ell\}}$ *converges towards the finite-dimensional marginals, at times* $t_1,\dots,t_\ell$, *of* k *independent copies of* $(X_t, M_t)_{t\geqslant 0}$.

(iv) *Finally, we emphasize that* (4.5) *is the keystone to switch from the finite system of particles to the dynamics of the McKean-Vlasov type.*

4.4. **Convergence of the delayed system.** Here is the analogue of the previous result for the delayed system $((Z_t^\delta, M_t^\delta)_{t\geqslant 0})_{\delta\in(0,1)}$, where $Z_t^\delta := X_t^\delta + M_t^\delta$ and $(X_t^\delta, M_t^\delta)_{t\geqslant 0}$ is a solution to (3.11).

Theorem 4.9. *Assume that there exists a unique physical solution* $(X_t, M_t)_{t\geqslant 0}$ *to* (1.1) *and denote by* $(Z_t, M_t)_{t\geqslant 0}$ *the reformulated solution (as defined in Remark 2.5). Denote also by* J *the (at most countable) set of discontinuity points of the function* $[0,+\infty) \ni t \mapsto \mathbb{E}(M_t)$. *Then, for any* $S \in [0,+\infty) \setminus J$,

$$\big(Z_s^\delta, \hat{M}_s^\delta\big)_{s\in[0,S]} \Rightarrow \mathbb{P}_{(Z_s,\hat{M}_s)_{s\in[0,S]}} \quad \textit{as } \delta \to 0, \tag{4.6}$$

on the space $\hat{\mathcal{D}}([0,S],\mathbb{R}) \times \hat{\mathcal{D}}([0,S],\mathbb{R})$ *equipped with the product topology induced by the M1 topology, where* $\hat{M}_s^\delta = M_s^\delta$ *if* $s < S$ *and* $\hat{M}_s^\delta = M_{S-}^\delta$ *if* $s = S$. *Moreover,*

$$\big(\mathbb{E}\big(\hat{M}_s^\delta\big)\big)_{s\in[0,S]} \to \big(\mathbb{E}(M_s)\big)_{s\in[0,S]} \tag{4.7}$$

as $\delta \to 0$*, on* $\hat{\mathcal{D}}([0,S],\mathbb{R})$ *equipped with the M1 topology.*

5. Proofs

5.1. **Preliminary estimates for the particle system.** This first subsection is devoted to the proof of two preliminary technical lemmas. The first one will be used for establishing suitable tightness properties of the particle system, while the second one will be needed to show that in the limit the solution does indeed satisfy the required properties to be physical in the sense defined above.

Throughout the section $((Z^i_t, M^i_t)_{t\,\geqslant\,0})_{i=1,\dots,N}$ will denote the physical solution to (3.3). We start with a moment estimate:

Lemma 5.1. *For any* $p \geqslant 1$ *and* $T \geqslant 0$*, there exists* $C^{(p)}_T \geqslant 0$*, independent of* N*, such that*

$$\forall i \in \{1,\dots,N\}, \quad \mathbb{E}\big[\sup_{t\in[0,T]} |Z^i_t|^p + \big(M^i_T\big)^p\big] \leqslant C^{(p)}_T.$$

Proof. The proof is a consequence of (3.8), (3.9) and Assumptions 1 and 2. □

Lemma 5.2. *For all* $\eta > 0$ *there exists a constant* $\lambda(\eta) > 0$ *that is independent of* N *(but depends on the constant* ε_0 *in Assumption 2), such that*

$$\mathbb{P}\big\{\forall t \in [0,\lambda(\eta)], \quad \bar{e}^N(t) \geqslant \big(\lambda(\eta)\big)^{-1} t^{1/4}\big\} \leqslant \eta,$$

where $\bar{e}^N(t)$ *is defined by* (3.4)*.*

Proof. Given $T \in (0,1)$, define τ by

$$\tau = \inf\bigg\{t \geqslant 0 : \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{M^i_t\,\geqslant\,1\}} \geqslant T\bigg\} \quad (\inf\emptyset = +\infty),$$

which is the first time the proportion of particles that have spiked at least once is bigger than T. For $t < \tau \wedge T$, the Cauchy-Schwarz inequality yields

$$\frac{1}{N}\sum_{i=1}^N M^i_t \leqslant \bigg(\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{M^i_t\,\geqslant\,1\}}\bigg)^{1/2}\bigg(\frac{1}{N}\sum_{i=1}^N \big(M^i_t\big)^2\bigg)^{1/2} \leqslant T^{1/2}\bigg(\frac{1}{N}\sum_{i=1}^N \big(M^i_T\big)^2\bigg)^{1/2}. \quad (5.1)$$

The point is now to investigate $\mathbb{P}\{\tau \leqslant T\}$. To this end, we define the events

$$\begin{aligned} &A = \bigg\{\frac{1}{N}\sum_{i=1}^N \big(M^i_T\big)^2 \leqslant T^{-1/2}\bigg\}, \quad A' = \bigg\{\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{A^i} \leqslant T^2\bigg\}, \quad \text{where} \\ &A^i = \bigg\{C\int_0^T (1 + \sup_{r\in[0,s]} |Z^i_r|)ds + \alpha T^{1/4} + \sup_{s\in[0,T]} |W^i_s| \geqslant \varepsilon_0/2\bigg\}, \end{aligned} \quad (5.2)$$

and $C \geqslant 0$ (independent of N and T) is chosen to be the constant such that on A, for $t < \tau \wedge T$ and any $1 \leqslant i \leqslant N$,

$$\sup_{s\in[0,t]} \big(Z^i_s\big)_+ \leqslant \big(Z^i_0\big)_+ + C\int_0^t \big(1 + \sup_{r\in[0,s]} |Z^i_r|\big)ds + \alpha T^{1/4} + \sup_{s\in[0,t]} |W^i_s|.$$

The existence of such a constant follows from the definition of the reformulated particle system (3.3) and (5.1). As $(Z^i_0)_+ \leqslant 1 - \varepsilon_0$ by Assumption 2, we have, on A, for $t < \tau \wedge T$,

$$\sup_{s\in[0,t]} \big(Z^i_s\big)_+ \geqslant 1 - \frac{\varepsilon_0}{2} \Rightarrow C\int_0^T \big(1 + \sup_{r\in[0,s]} |Z^i_r|\big)ds + \alpha T^{1/4} + \sup_{s\in[0,T]} |W^i_s| \geqslant \frac{\varepsilon_0}{2}.$$

If $\tau \leqslant T$, the above is true on A for all $t < \tau$. We deduce that, on $A \cap \{\tau \leqslant T\}$

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{\sup_{s\in[0,\tau)}(Z_s^i)_+ \,\geqslant\, 1-\varepsilon_0/2\}} \leqslant \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{A^i},$$

with $(A^i)_{i=1,\dots,N}$ as in (5.2). On $A \cap A' \cap \{\tau \leqslant T\}$, it thus holds that

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{\sup_{s\in[0,\tau)}(Z_s^i)_+ \,\geqslant\, 1-\varepsilon_0/2\}} \leqslant T^2.$$

Assume that $T^2 \leqslant \varepsilon_0/4$. Then the number of particles such that $\sup_{s\in[0,\tau)}(Z_s^i)_+ \geqslant 1-\varepsilon_0/2$ is at most $N\varepsilon_0/4$. The other particles cannot cross 1 at time τ, since the size of the kick they receive due to those such that $\sup_{s\in[0,\tau)}(Z_s^i)_+ \geqslant 1-\varepsilon_0/2$ is bounded by $\varepsilon_0/4$. Therefore, the number that have crossed 1 up to and including τ must also be less than NT^2, i.e. $(1/N)\sum_{i=1}^N \mathbf{1}_{\{M_\tau^i \,\geqslant\, 1\}} \leqslant T^2$.

Since $T \leqslant \sqrt{\varepsilon_0}/2 \leqslant 1/2$, we have $T^2 < T$. This yields a contradiction since, by definition of τ and by right-continuity, $(1/N)\sum_{i=1}^N \mathbf{1}_{\{M_\tau^i \,\geqslant\, 1\}} \geqslant T$. In other words $A \cap A' \cap \{\tau \leqslant T\} = \emptyset$, so that $\{\tau \leqslant T\} \subset (A\cap A')^{\complement}$. Hence

$$\mathbb{P}(\tau \leqslant T) \leqslant \mathbb{P}\big(A^{\complement}\big) + \mathbb{P}\big((A')^{\complement}\big). \tag{5.3}$$

By Markov's inequality, $\mathbb{P}((A')^{\complement}) = \mathbb{P}((1/N)\sum_{i=1}^N \mathbf{1}_{A^i} > T^2) \leqslant (1/NT^2)\sum_{i=1}^N \mathbb{P}(A^i)$. Thus, by (5.2) and using the fact that $T \leqslant 1$,

$$\mathbb{P}\big((A')^{\complement}\big) \leqslant \frac{1}{NT^2}\sum_{i=1}^{N}\Big[\mathbb{P}\Big((\alpha+C)T^{1/4} + \sup_{s\in[0,T]}|W_s^i| \geqslant \frac{\varepsilon_0}{4}\Big) + \mathbb{P}\Big(\sup_{s\in[0,T]}|Z_s^i| \geqslant \frac{\varepsilon_0}{4CT}\Big)\Big]. \tag{5.4}$$

By Lemma 5.1 (with $p = 3$) and Markov's inequality again, we see that (since $T \leqslant 1$),

$$\mathbb{P}\Big(\sup_{s\in[0,T]}|Z_s^i| \geqslant \frac{\varepsilon_0}{4CT}\Big) \leqslant 4^3 C^3 C_1^{(3)} \varepsilon_0^{-3} T^3 \leqslant C' T^3, \tag{5.5}$$

for another constant C' depending upon ε_0. Under the additional assumption that $(\alpha + C)T^{1/4} \leqslant \varepsilon_0/8$, the first term in the right-hand side of (5.4) can be bounded by

$$\begin{aligned}\frac{1}{NT^2}\sum_{i=1}^{N}\mathbb{P}\bigg((\alpha+C)T^{1/4} + \sup_{s\in[0,T]}|W_s^i| \geqslant \frac{\varepsilon_0}{4}\bigg) &\leqslant \frac{1}{NT^2}\sum_{i=1}^{N}\mathbb{P}\bigg(\sup_{s\in[0,T]}|W_s^i| \geqslant \frac{\varepsilon_0}{8}\bigg)\\ &\leqslant cT^{-2}\exp\big(-c^{-1}T^{-1}\big),\end{aligned}$$

for some constant $c > 0$, independent of N and T (but depending upon ε_0). Here we have used the reflection principle and an elementary bound on the Gaussian distribution function.

In the end, by (5.4), (5.5) and the above inequality, we obtain

$$\mathbb{P}\big((A')^{\complement}\big) \leqslant C'T + cT^{-2}\exp\big(-c^{-1}T^{-1}\big) \leqslant C'T,$$

for some C', independent of N and T and the value of which is allowed to increase from one inequality to another. In a similar way to the proof of (5.5), we also have that $\mathbb{P}(A^{\complement}) \leqslant C'T$. Therefore, (5.3) yields

$$\mathbb{P}(\tau \leqslant T) \leqslant C'T,$$

for $T^2 \leqslant \varepsilon_0/4$ and $(\alpha+C)T^{1/4} \leqslant \varepsilon_0/8$. The point is that this probability is small in T. Finally, by (5.1) and by Lemma 5.1 again,

$$\begin{aligned}\mathbb{P}\big(\bar{e}^N(T) \geqslant T^{1/4}\big) &\leqslant \mathbb{P}\big(\bar{e}^N(T) \geqslant T^{1/4}, T < \tau\big) + \mathbb{P}\big(\tau \leqslant T\big)\\ &\leqslant \mathbb{P}\bigg(\frac{1}{N}\sum_{i=1}^N (M_T^i)^2 \geqslant T^{-1/2}\bigg) + \mathbb{P}\big(\tau \leqslant T\big)\\ &\leqslant \mathbb{P}(A^{\complement}) + \mathbb{P}\big(\tau \leqslant T\big) \leqslant C'T.\end{aligned}$$

Choose now $T = T_k$ and $T_k = \lambda^k$ with $\lambda < 1$ such that $(\alpha+C)\lambda^{1/4} \leqslant \varepsilon_0/8$ and $\lambda^2 \leqslant \varepsilon_0/4$. Then by above,

$$\mathbb{P}\big(\bar{e}^N(T_k) \geqslant T_k^{1/4}\big) \leqslant C'\lambda^k,$$

so that, for any $k_0 \geqslant 1$,

$$\mathbb{P}\bigg(\bigcup_{k\,\geqslant\,k_0} \big\{\bar{e}^N(T_k) \geqslant T_k^{1/4}\big\}\bigg) \leqslant \sum_{k\,\geqslant\,k_0} C'\lambda^k =: \eta(k_0),$$

where η is finite since the sum converges, is independent of N and satisfies $\lim_{x\to+\infty}\eta(x) = 0$. Observe now that, for any $t \in (T_{k+1}, T_k]$, $M_t^i \leqslant M_{T_k}^i$ and $\lambda^{-1/4}t^{1/4} \geqslant T_k^{1/4}$, so that $\bar{e}^N(t) \geqslant \lambda^{-1/4}t^{1/4}$ implies $\bar{e}^N(T_k) \geqslant T_k^{1/4}$ (recall that $\bar{e}^N$ is non-decreasing). Therefore,

$$\mathbb{P}\big(\exists t \in [0, T_{k_0}] : \bar{e}^N(t) \geqslant \lambda^{-1/4}t^{1/4}\big) \leqslant \eta(k_0),$$

thus completing the proof. □

The next proposition shows that there is a very small chance of observing a macroscopic proportion of particles spiking twice or more in a small interval and extends Proposition 3.4 to intervals of non-zero length.

Proposition 5.3. *For a given $T > 0$, consider $0 \leqslant t < t+h \leqslant T$, $h \in (0,1)$. Then, we can find $C \geqslant 0$, independent of h, and an integer $N_0 := N_0(h)$, such that, for $N \geqslant N_0$,*

$$\mathbb{P}\big(\bar{e}^N(t+h) - \bar{e}^N(t-) > 1 + Ch^{1/16}\big) \leqslant Ch$$

and

$$\mathbb{P}\bigg(\forall\lambda \leqslant \big(\bar{e}^N(t+h) - \bar{e}^N(t-) - Ch^{1/16}\big)_+,\ \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{X_{t-}^i \,\geqslant\, 1-\alpha\lambda - Ch^{1/16}\}} \geqslant \lambda\bigg) \geqslant 1 - Ch.$$

Proof. The first step of the proof is to show that the proportion of particles that spike twice in a small interval tends to 0 with the length of the interval, uniformly in $N \geqslant 1$. More precisely, given an interval $[t, t+h]$ and $\beta \in (0,1)$, define

$$\tau(\beta) = \inf\bigg\{s \in [t, t+h] : \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{M_s^i - M_{t-}^i \,\geqslant\, 2\}} \geqslant \beta\bigg\}, \quad \inf\emptyset = +\infty.$$

Then we want to show that, for $\beta = h^{1/4}$, $N > h^{-1/2}$ and $p \geqslant 1$, there exists a constant C_p (independent of h), such that

$$\mathbb{P}\big(\tau(\beta) \leqslant t+h\big) \leqslant C_p h^p. \tag{5.6}$$

In order to do this, we will have to enter into the spike cascade, which will require the *cascade time axis* defined on page 11. Indeed, we will say that particle $i \in \{1,\dots,N\}$ spikes *twice before* j if $\inf\{s \geqslant t, M_s^i \geqslant M_{t-}^i + 2\} < \inf\{s \geqslant t, M_s^j \geqslant M_{t-}^j + 2\}$, or

$\inf\{s \geqslant t, M_s^i \geqslant M_{t-}^i + 2\} = \inf\{s \geqslant t, M_s^j \geqslant M_{t-}^j + 2\} =: \rho$ and $X_{\rho-}^i > X_{\rho-}^j$. This precisely means that particle i will spike twice before j either in (usual) time, or before j along the *cascade time axis*.

Define the set $I = \{i \in \{1,\dots,N\} : M_{t+h}^i - M_{t-}^i \geqslant 2\}$ of particles that have spiked at least twice in the interval $[t-, t+h]$. We prove the following claim:

Claim: Suppose $\tau(\beta) \leqslant t+h$. Then there exists a set $I(\beta) \subset \{1,\dots,N\}$, such that $\beta N \leqslant |I(\beta)| \leqslant \beta N + 1$ and, for all $i \in I(\beta)$, it holds that

$$1 \leqslant \alpha + \beta^{1/2}\Big(\frac{1}{N}\sum_{j=1}^{N}\big(M_{t+h}^j\big)^2\Big)^{1/2} + Ch\big(1 + \sup_{s\in[0,t+h]}|Z_s^i|\big) + \sup_{s\in[t,t+h]}\big|W_s^i - W_t^i\big|. \quad (5.7)$$

To prove (5.7), suppose $\tau(\beta) \leqslant t+h$. By right-continuity of each $(M_s^i)_{s\geqslant 0}$, $|I| \geqslant N\beta$. For $i_0 \in I$, let $I^{(i_0)}$ be the set of particles that have spiked twice before i_0 in the above sense. Whenever $|I^{(i_0)}| < \beta N$, the sum of the kicks received by particle i_0 due to the effect of the particles in $I^{(i_0)}$ spiking before it (again in the previous sense) is bounded by

$$\alpha\Big[1 + \frac{1}{N}\sum_{i\in I^{(i_0)}}\big(M_{t+h}^i - M_{t-}^i\big)\Big] \leqslant \alpha\Big[1 + \frac{1}{N}\sum_{i\in I^{(i_0)}} M_{t+h}^i\Big] \leqslant \alpha + \beta^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\big(M_{t+h}^i\big)^2\right)^{1/2}. \quad (5.8)$$

The first α stands for the kick generated by particles that have spiked once only. The other part corresponds to the particles that have spiked twice or more. At the time when the particle i_0 spikes for the second time, X^{i_0} has to cross 1, or equivalently, the Z-particle i_0 crosses a new integer. Since it is its second spike in the interval $[t, t+h]$, the Z-particle i_0 has run more than 1 since $t-$, i.e. $1 \leqslant \sup_{t\leqslant s\leqslant t+h}|Z_s - Z_{t-}|$, so that

$$1 \leqslant \int_t^{t+h}|b(X_s^{i_0})|ds + \alpha + \beta^{1/2}\Big(\frac{1}{N}\sum_{j=1}^{N}\big(M_{t+h}^j\big)^2\Big)^{1/2} + \sup_{s\in[t,t+h]}\big|W_s^{i_0} - W_t^{i_0}\big|.$$

Using the bound for the growth of b,

$$1 \leqslant \alpha + \beta^{1/2}\Big(\frac{1}{N}\sum_{j=1}^{N}\big(M_{t+h}^j\big)^2\Big)^{1/2} + Ch\big(1 + \sup_{s\in[0,t+h]}|Z_s^{i_0}|\big) + \sup_{s\in[t,t+h]}\big|W_s^{i_0} - W_t^{i_0}\big|.$$

Iterating the argument up until the number of particles that have spiked more than twice is greater than $N\beta$, we can find an index i_1 such that $I^{(i_1)} \subset \{1,\dots,N\}$ and $\beta N \leqslant |I^{(i_1)}| \leqslant \beta N + 1$. This proves the claim.

To proceed, we can take the mean in (5.7) over the particles in $I(\beta)$. Using the bound $|I(\beta)| \leqslant N\beta + 1$ and the Cauchy-Schwarz inequality, we see that

$$\begin{aligned}\beta \leqslant \alpha\beta &+ \beta^{1/2}\big(\beta + \frac{1}{N}\big)\left(\frac{1}{N}\sum_{j=1}^{N}\big(M_{t+h}^j\big)^2\right)^{1/2} \\ &+ \big(\beta + \frac{1}{N}\big)^{1/2}\left[Ch\Big(1 + \frac{1}{N}\sum_{j=1}^{N}\sup_{s\in[0,t+h]}|Z_s^j|^2\Big)^{1/2} + \Big(\frac{1}{N}\sum_{j=1}^{N}\sup_{s\in[t,t+h]}|W_s^j - W_t^j|^2\Big)^{1/2}\right].\end{aligned}$$

Since $\beta = h^{1/4} \geqslant 1/\sqrt{N}$, we have $1/(\beta N) \leqslant 1/\sqrt{N}$. Dividing both sides of the above inequality by β, we deduce that $\tau(\beta) \leqslant t+h$ implies

$$1 \leqslant \alpha + 2\beta^{1/2}\Bigl(\frac{1}{N}\sum_{j=1}^{N}\bigl(M^j_{t+h}\bigr)^2\Bigr)^{1/2}$$
$$+ Ch\beta^{-1/2}\Bigl(1+\frac{1}{N}\sum_{j=1}^{N}\sup_{s\in[0,t+h]}|Z^j_s|^2\Bigr)^{1/2} + 2\beta^{-1/2}\Bigl(\frac{1}{N}\sum_{j=1}^{N}\sup_{s\in[t,t+h]}|W^j_s - W^j_t|^2\Bigr)^{1/2}.$$

We can now apply Markov's inequality with any exponent $p \geqslant 1$. By Lemma 5.1, we get that there exists a constant C_p such that

$$\mathbb{P}\bigl(\tau(\beta) \leqslant t+h\bigr) \leqslant C_p\bigl(\beta^{p/2} + h^p\beta^{-p/2} + h^{p/2}\beta^{-p/2}\bigr) = C_p\bigl(h^{p/8} + h^{7p/8} + h^{3p/8}\bigr). \quad (5.9)$$

On the event $\{\tau(\beta) > t+h\} \cap \{N^{-1}\sum_{i=1}^{N}(M^i_{t+h})^2 \leqslant h^{-1/8}\}$, we have, as in (5.8),

$$\bar{e}^N(t+h) - \bar{e}^N(t-) \leqslant 1 + \beta^{1/2}\Bigl(\frac{1}{N}\sum_{i=1}^{N}\bigl(M^i_{t+h}\bigr)^2\Bigr)^{1/2} \leqslant 1 + \beta^{1/2}h^{-1/16}. \quad (5.10)$$

Since $\beta = h^{1/4}$, we deduce from Lemma 5.1 and (5.9) that the first bound in the statement holds with N large enough.

Now, we can focus on the particles that spike no more than once between t and $t+h$. The set of such particles coincides with $I^{\complement}$. In $I^{\complement}$, there are two kind of particles: The set $I^{\complement,0}$ denotes the set of particles that do not spike and the set $I^{\complement,1}$ the set of particles that spike once. In order to characterize the sets $I^{\complement,0}$ and $I^{\complement,1}$, we can make use of the ordering of the spikes again, as defined on page 11.

A particle $i_1 \in I^{\complement,1}$ is to spike at some time $s \in [t, t+h]$, if, at some moment along the *time cascade axis* at s, the kick it receives from the particles that spike before in the cascade is larger than $1 - X^{i_1}_{s-}$. Now, as i_1 doesn't spike between t and $s-$, $1 - X^{i_1}_{s-}$ is equal to

$$1 - X^{i_1}_{s-} = 1 - X^{i_1}_{t-} - \int_t^s b(X^{i_1}_r)dr - (W^{i_1}_s - W^{i_1}_t) - \alpha\bigl(\bar{e}^N(s-) - \bar{e}^N(t-)\bigr).$$

We observe that $\alpha(\bar{e}^N(s-) - \bar{e}^N(t-))$ represents the kick i_1 receives from the other neurons between $t-$ and $s-$. Therefore, $i_1 \in I^{\complement,1}$ is to spike at some time $s \in [t, t+h]$, if the kick it receives between $t-$ and $s-$ plus the kick it receives along the *time axis cascade* at s before it spikes is greater than

$$1 - X^{i_1}_{t-} - \int_t^s b(X^{i_1}_r)dr - (W^{i_1}_s - W^{i_1}_t).$$

The sum of the two kicks is called the *kick received by* i_1 *before it spikes*. Following (5.8), it can be bounded by

$$\alpha\frac{k}{N} + \Bigl(\frac{|I|}{N} \times \frac{1}{N}\sum_{j=1}^{N}\bigl(M^j_{t+h}\bigr)^2\Bigr)^{1/2},$$

where k stands for the number of particles in $I^{\complement}$ that spike once before i_1. Therefore, for i_1 to be in $I^{\complement,1}$, it must hold that

$$X^{i_1}_{t-} + \alpha \frac{k}{N} + \Big(\frac{|I|}{N^2} \sum_{j=1}^N \big(M^j_{t+h} \big)^2 \Big)^{1/2} + Ch\big(1 + \sup_{s\in[0,t+h]} |Z^{i_1}_s|\big) + \sup_{s\in[t,t+h]} |W^{i_1}_s - W^{i_1}_t| \geqslant 1, \tag{5.11}$$

the two last terms in the left-hand side standing for bounds on the drift and Brownian parts in the dynamics of X^{i_1}. Obviously, the number of particles for which the above inequality holds must be larger than $k+1$ (it must be true for the k particles that spiked before i_1 and for i_1 as well). On the model of Proposition 3.4, this must be true for any $k < |I^{\complement,1}|$.

Now, following (5.10) for estimating the overall kick on $[t, t+h]$, we deduce

$$\frac{|I^{\complement,1}|}{N} \leqslant \bar{e}^N(t+h) - \bar{e}^N(t-) \leqslant \frac{|I^{\complement,1}|}{N} + \Big(\frac{|I|}{N^2} \sum_{j=1}^N \big(M^j_{t+h} \big)^2 \Big)^{1/2}.$$

Therefore, for an integer $0 \leqslant k < N(\bar{e}^N(t+h) - \bar{e}^N(t-)) - [|I| \sum_{j=1}^N (M^j_{t+h})^2]^{1/2}$, we have $k < |I^{\complement,1}|$. By (5.11), we can deduce that

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{B^{i,1}(k)} \geqslant \frac{k}{N}, \quad \text{with } B^{i,1}(k) = \Big\{ X^i_{t-} + \alpha \frac{k}{N} + \Big(\frac{|I|}{N^2} \sum_{j=1}^N \big(M^j_{t+h} \big)^2 \Big)^{1/2} \\ + Ch\big(1 + \sup_{s\in[0,t+h]} |Z^i_s|\big) + \sup_{s\in[t,t+h]} |W^i_s - W^i_t| \geqslant 1 \Big\}. \end{aligned} \tag{5.12}$$

Define now the events

$$B^0 = \{\tau(\beta) > t+h\} \cap \Big\{ \frac{1}{N} \sum_{j=1}^N \big(M^j_{t+h} \big)^2 \leqslant h^{-1/8} \Big\}$$

$$B^2 = \Big\{ \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{B^{i,2}} \leqslant h \Big\}, \quad B^{i,2} = \Big\{ Ch\big(1 + \sup_{s\in[0,t+h]} |Z^i_s|\big) + \sup_{s\in[t,t+h]} |W^i_s - W^i_t| \geqslant h^{1/16} \Big\}.$$

On B^0, the term $((|I|/N^2) \sum_{j=1}^N (M^j_{t+h})^2)^{1/2}$ is less than $h^{1/16}$. Therefore, on B^0, for any $\lambda \leqslant \bar{e}^N(t+h) - \bar{e}^N(t-) - 3h^{1/16}$, we have

$$\lfloor \lambda N + 2N h^{1/16} \rfloor \leqslant N\big(\bar{e}^N(t+h) - \bar{e}^N(t-)\big) - \Big[|I| \sum_{j=1}^N (M^j_{t+h})^2 \Big]^{1/2}.$$

For such a λ, we choose $k = \lfloor \lambda N + 2N h^{1/16} \rfloor$, so that k satisfies the required condition to apply (5.12). On $B^0 \cap B^{i,1}(k) \cap (B^{i,2})^{\complement}$, $X^i_{t-} + \alpha k/N \geqslant 1 - 2h^{1/16}$, so that

$$X^i_{t-} + \alpha \lambda \geqslant X^i_{t-} + \alpha \frac{k}{N} - 2h^{1/16} \geqslant 1 - 4h^{1/16}.$$

Therefore, on $B^0 \cap B^2$,

$$\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{X^i_{t-} \geqslant 1 - \alpha\lambda - 4h^{1/16}\}} \geqslant \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{B^{i,1}(k)} - \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{B^{i,2}} \geqslant \frac{k}{N} - h \geqslant \lambda,$$

the last inequality following from the fact that, since $N > h^{-1/2}$ and $h \leqslant 1$, $k/N - h \geqslant \lambda + 2h^{1/16} - 1/N - h \geqslant \lambda$. To complete the proof, notice from (5.9) and Lemma 5.1 that

$$\mathbb{P}\Big(\big(B^0 \cap B^2\big)^{\complement}\Big) \leqslant C'h + \mathbb{P}\Big(\big(B^2\big)^{\complement}\Big) \leqslant C'h + h^{-1}\frac{1}{N}\sum_{i=1}^{N}\mathbb{P}\big(B^{i,2}\big) \leqslant C'h,$$

the value of C' being allowed to increase from one inequality to another. □

5.2. **Tightness properties and convergent subsequences.** This second section is now devoted to the proof of the tightness property of the family of measures $(\Pi_N)_{N \geqslant 1}$ defined in Theorem 4.4. It is for this result that the M1 Skorohod topology plays a key role. Recall that $\widetilde{Z}^{i,N} \in \hat{\mathcal{D}}([0,T+1],\mathbb{R})$ satisfies

$$\widetilde{Z}^{i,N}_t := \begin{cases} Z^{i,N}_t, & \text{if } t \leqslant T \\ W^i_t - W^i_T + Z^{i,N}_T, & \text{if } t \in (T, T+1] \end{cases} \tag{5.13}$$

for $i \in \{i, \dots, N\}$, where $(Z^{i,N}_t)_{t\in[0,T]}$ is the (physical) solution to the particle system (3.3),

$$\bar{\mu}_N := \frac{1}{N}\sum_{i=1}^{N}\mathrm{Dirac}(\widetilde{Z}^{i,N}),$$

so that $\bar{\mu}_N$ is a random variable taking values in $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$ and $\Pi_N = \mathrm{Law}(\bar{\mu}_N)$.

Lemma 5.4. *For any $T > 0$, the family of laws of $\widetilde{Z}^{1,N}$, $N \geqslant 1$, is tight in $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$ endowed with the weak topology inherited from the M1 topology on $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$.*

Proof. By definition of tightness, we must show that for any $\varepsilon \geqslant 0$, there exists $K \subset \hat{\mathcal{D}}([0,T+1],\mathbb{R})$ compact for the M1 topology, such that $\inf_{N \geqslant 1}\mathbb{P}(\widetilde{Z}^{1,N} \in K) \geqslant 1-\varepsilon$. By Theorem 4.3, K is compact for M1 if and only if $\lim_{\delta\to 0}\sup_{f\in K} u_{T+1}(f,\delta) = 0$, where

$$u_{T+1}(f,\delta) := \Big(\sup_{t\in[0,T+1]} w_{T+1}(f,t,\delta)\Big) \vee v_{T+1}(f,0,\delta) \vee v_{T+1}(f,T+1,\delta) \tag{5.14}$$

is the modulus of continuity appearing in that result, and the functions w_{T+1} and v_{T+1} are given by (4.1) and in Theorem 4.3 respectively. With this in mind, for any $R \geqslant 1$, define the set

$$K_R := \big\{f \in \hat{\mathcal{D}}([0,T+1],\mathbb{R}) : u_{T+1}(f,\delta) \leqslant R\delta^{\frac{1}{4}},\ \delta \in (0, 1/R)\big\}.$$

It is thus clear that K_R is compact for every $R \geqslant 1$. It therefore suffices to show that

$$\lim_{R\to\infty}\inf_{N \geqslant 1}\mathbb{P}\Big(\forall\delta \in (0,1/R),\ u_{T+1}(\widetilde{Z}^{1,N},\delta) \leqslant R\delta^{\frac{1}{4}}\Big) = 1. \tag{5.15}$$

Since u_{T+1} is a maximum of three terms, the above certainly holds if it also holds when u_{T+1} is replaced by each of the three terms appearing in the maximum in (5.14) individually. This is what we aim to show now, starting with the first term.

To this end define the new process $(\widetilde{U}^N_t)_{t\in[0,T+1]}$ as

$$\widetilde{U}^N_t = \widetilde{Z}^{1,N}_t - \frac{1}{N}\sum_{i=1}^{N} M^{i,N}_{t\wedge T}, \qquad t \in [0,T+1].$$

Then $\widetilde{U}^N$ is the continuous part of $\widetilde{Z}^{1,N}$. We use the easily verified fact that

$$w_{T+1}(f+g,t,\delta) \leqslant v_{T+1}(g,t,\delta) \quad t \in [0,T], \delta > 0,$$

whenever $f, g \in \hat{\mathcal{D}}([0, T+1], \mathbb{R})$, and f is monotone. Thus, since $\widetilde{Z}^{1,N} - \widetilde{U}^N$ is non-decreasing,

$$\sup_{t\in[0,T+1]} w_{T+1}(\widetilde{Z}^{1,N}, t, \delta) \leqslant \sup_{t\in[0,T+1]} v_{T+1}(\widetilde{U}^N, t, \delta), \quad \delta > 0,$$

almost surely. Hence, in order to show that (5.15) holds in the first case, it is sufficient to prove that $\inf_{N\geqslant 1} \mathbb{P}(\forall\delta \in (0, 1/R),\ \sup_{t\in[0,T+1]} v_{T+1}(\widetilde{U}^N, t, \delta) \leqslant R\delta^{1/4})$ converges to 1 as $R \to \infty$. Since $(\widetilde{U}_t^N)_{t\in[0,T+1]}$ is a continuous process driven by a Lipschitz drift and a Brownian motion, this directly follows from Lemma 5.1.

To handle the second case, when $u_{T+1}(f, \delta) = v_{T+1}(f, 0, \delta)$, by Lemma 5.2, we know that, for any $\eta > 0$ there exists $\lambda(\eta) > 0$ independent of N (depending only on the ε_0 appearing in Assumption 2) such that

$$\mathbb{P}\big(\forall t \in [0, \lambda(\eta)], \quad \bar{e}^N(t) \geqslant \big(\lambda(\eta)\big)^{-1} t^{1/4}\big) \leqslant \eta.$$

In particular, by taking $R = \big(\lambda(\eta)\big)^{-1}$, it follows that

$$\lim_{R\to+\infty} \inf_{N\geqslant 1} \mathbb{P}\bigg(\forall\delta \in (0, 1/R),\ v_{T+1}\bigg(\frac{1}{N}\sum_{i=1}^{N} M^{i,N}, 0, \delta\bigg) \leqslant R\delta^{1/4}\bigg) = 1,$$

where we have also used the definition of v_{T+1}, given in Theorem 4.3. As the continuous part $(\widetilde{U}^N)$ of the dynamics of $\widetilde{Z}^{1,N}$ can be handled in the standard way, we deduce that

$$\lim_{R\to+\infty} \inf_{N\geqslant 1} \mathbb{P}\Big(\forall\delta \in (0, 1/R),\ v_{T+1}\big(\widetilde{Z}^{1,N}, 0, \delta\big) \leqslant R\delta^{1/4}\Big) = 1.$$

For the final term in the maximum in (5.14), by definition, $\widetilde{Z}^{1,N}$ behaves as W^1 in neighborhoods of $T+1$, so that (again in the standard way)

$$\lim_{R\to+\infty} \inf_{N\geqslant 1} \mathbb{P}\Big(\forall\delta \in (0, 1/R),\ v_{T+1}\big(\widetilde{Z}^{1,N}, T+1, \delta\big) \leqslant R\delta^{1/4}\Big) = 1.$$

This completes the proof. □

By [17, Proposition 2.2], we deduce:

Lemma 5.5. *For any $T > 0$ the family $(\Pi_N)_{N\geqslant 1} \subset \mathcal{P}(\mathcal{P}(\hat{\mathcal{D}}([0, T+1], \mathbb{R})))$ is tight, where $\mathcal{P}(\hat{\mathcal{D}}([0, T+1], \mathbb{R}))$ is endowed with the weak topology deriving from the M1 topology on $\hat{\mathcal{D}}([0, T+1], \mathbb{R})$.*

5.3. **Proof of Theorem 4.4.** We now give the proof of Theorem 4.4. As we will see, the proof will rely on some key convergence results that will be proved afterwards. Throughout the proof, as in the statement of the result, $(z_t)_{t\in[0,T+1]}$ will denote the canonical process on $\hat{\mathcal{D}}([0, T+1], \mathbb{R})$ and $m_t = \lfloor(\sup_{s\in[0,t]} z_s)_+\rfloor$.

The first part of the theorem is contained in Lemma 5.5, namely that $(\Pi_N)_{N\geqslant 1}$ is tight. We can therefore extract a convergent subsequence as N tends to $+\infty$, which we denote in the same way. We set Π_∞ to be the limit point of such a sequence.

Consider the function

$$[0, T+1] \ni t \mapsto \int \mu\big\{z_{t-} = z_t\big\} d\Pi_\infty(\mu).$$

Since the application $A \in \mathcal{B}(\hat{\mathcal{D}}([0, T+1], \mathbb{R})) \mapsto \int \mu(A) d\Pi_\infty(\mu)$ defines a probability measure on $\hat{\mathcal{D}}([0, T+1], \mathbb{R})$, the function matches 1 for any t in $[0, T+1]$ but in some countable subset $J \subset [0, T+1]$ (see Lemma 7.7, p. 131, chap. 3 in Ethier and Kurtz [9]). Therefore, for $t \notin J$, for a.e. μ under the probability Π_∞, $\mu\{z_{t-} = z_t\} = 1$. Similarly, the function

$$[0, T+1] \ni t \mapsto \int \langle \mu, m_t \rangle d\Pi_\infty(\mu)$$

is non-decreasing and has at most a countable number of jumps. Up to a modification of J, we can assume that the jumps are all included in J. Then, for any $t \notin J$,

$$\int \langle \mu, m_t \rangle d\Pi_\infty(\mu) = \int \langle \mu, m_{t-} \rangle d\Pi_\infty(\mu).$$

Therefore, for $t \notin J$, for a.e. μ under the probability measure Π_∞, $\langle \mu, m_{t-} \rangle = \langle \mu, m_t \rangle$.

For $p \geqslant 1$, $S_1, \dots, S_p \notin J$, $0 = S_0 \leqslant S_1 < \dots < S_p < T$ and $f_0, \dots, f_p$ bounded and uniformly continuous functions from $\mathbb{R}$ into itself, put

$$F(z) = \prod_{i=0}^{p} f_i\big(z_{S_i}\big), \quad z \in \hat{\mathcal{D}}([0, T+1], \mathbb{R}). \tag{5.16}$$

For another bounded and uniformly continuous function G from $\mathbb{R}$ into itself, consider

$$Q_N := \mathbb{E}\bigg[G\bigg(\bigg\langle \bar{\mu}_N, F\bigg(z - z_0 - \int_0^{\cdot} b(z_s - m_s)ds - \alpha \langle \bar{\mu}_N, m_\cdot \rangle\bigg)\bigg\rangle\bigg)\bigg].$$

The point is that we may write

$$\widetilde{Z}_t^{i,N} = Z_0^{i,N} + \int_0^{t \wedge T} b(\widetilde{Z}_s^{i,N} - M_s^{i,N})ds + \alpha \langle \bar{\mu}_N, m_{t \wedge (T-)} \rangle + W_t^i$$

for all $t \in [0, T+1]$ and $i \in \{1, \dots, N\}$. It then follows that $Q_N = \mathbb{E}[G(N^{-1} \sum_{i=1}^N F(W^i))]$. By the law of large numbers, we deduce that

$$\lim_{N \to +\infty} Q_N = G\Big(\mathbb{E}\big(F(W)\big)\Big). \tag{5.17}$$

The key result in order to proceed is contained in Lemma 5.10 below, where it is shown that under Π_∞, the functional

$$\mu \in \mathcal{P}\big(\hat{\mathcal{D}}([0, T+1], \mathbb{R})\big) \mapsto \bigg\langle \mu, F\bigg(z - z_0 - \int_0^{\cdot} b(z_s - m_s)ds - \alpha \langle \mu, m_\cdot \rangle\bigg)\bigg\rangle, \tag{5.18}$$

is a.e. continuous. Indeed, with this in hand, by the continuous mapping theorem, we have

$$\lim_{N \to +\infty} Q_N = \int G\bigg(\bigg\langle \mu, F\bigg(z - z_0 - \int_0^{\cdot} b(z_s - m_s)ds - \alpha \langle \mu, m_\cdot \rangle\bigg)\bigg\rangle\bigg) d\Pi_\infty(\mu),$$

so that, from (5.17),

$$\int G\bigg(\bigg\langle \mu, F\bigg(z - z_0 - \int_0^{\cdot} b(z_s - m_s)ds - \alpha \langle \mu, m_\cdot \rangle\bigg)\bigg\rangle\bigg) d\Pi_\infty(\mu) = G\Big(\mathbb{E}\big(F(W)\big)\Big).$$

Applying the above equality with $\widetilde{G}(\cdot) = [G(\cdot) - G(\mathbb{E}(F(W)))]^2$ instead of G itself, we deduce that, Π_∞ a.s.,

$$G\bigg(\bigg\langle \mu, F\bigg(z - z_0 - \int_0^{\cdot} b(z_s - m_s)ds - \alpha \langle \mu, m_\cdot \rangle\bigg)\bigg\rangle\bigg) = G\Big(\mathbb{E}\big(F(W)\big)\Big),$$

so that, for a.a. probability measures μ under Π_∞, under μ, the process

$$\bigg(\Upsilon_t = z_t - z_0 - \int_0^t b(z_s - m_s)ds - \alpha \langle \mu, m \rangle_t\bigg)_{t \in [0, T+1]}$$

has the same finite-dimensional distributions as a Brownian motion at any points $0 \leqslant S_1 < S_2 < \cdots < S_p < T$ which are not in J. Since $(\Upsilon_t)_{t\in[0,T)}$ has right-continuous paths under μ (and $[0,T) \cap J^{\complement}$ is dense in $[0,T)$), it has the same finite-dimensional distributions as a Brownian motion. Moreover, since the Borel σ-field generated by M1 is also generated by the evaluation mappings, we deduce that the distribution of $(\hat{\Upsilon}_t)_{t\in[0,T]}$ on $\hat{\mathcal{D}}([0,T],\mathbb{R})$ is the Wiener distribution, where $\hat{\Upsilon}_t = \Upsilon_t$ if $t < T$ and $\hat{\Upsilon}_T = \Upsilon_{T-}$. This says that, for Π_∞ a.e. μ, the canonical process solves the reformulated equation (2.6) up until time T, with z_0 as initial condition, which proves (2) in Theorem 4.4.

We now check that the law of z_0 under μ is the distribution of X_0 under $\mathbb{P}$. To this end, let $\pi_0 : \hat{\mathcal{D}}([0,T+1],\mathbb{R}) \to \mathbb{R}$ be given by $\pi_0(z) = z_0$, and $\pi_0\sharp\mu$ be the push-forward measure of μ by π_0. Then the mapping $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R})) \ni \mu \mapsto \pi_0\sharp\mu$ is continuous (see Theorem 4.1). Using the Skorohod representation theorem, this allows the joint application of Theorem 4.2 at $t = 0$. By the law of large numbers, $\pi_0\sharp\bar{\mu}_N$ converges towards the law of X_0. Therefore, for a.e. μ under Π_∞, $\pi_0\sharp\mu$ matches the law of X_0, which proves (1) in Theorem 4.4.

We finally prove that, for almost all μ under Π_∞, the canonical process under μ satisfies the required conditions for defining a physical condition up until T. This requires showing the conditions (2) and (4) of Definition 2.2 are satisfied for the canonical process under μ. We make use of Proposition 5.3, which says that, for $0 \leqslant t < t+h < T$ and for N large enough,

$$\begin{aligned}
&\mathbb{P}\big(\bar{e}^N(t+h) - \bar{e}^N(t-) > 1 + Ch^{1/16}\big) \leqslant Ch,\\
&\mathbb{P}\Big(\forall \lambda \leqslant \bar{e}^N(t+h) - \bar{e}^N(t-) - Ch^{1/16}, \bar{\mu}_N\big(z_{t-} - m_{t-} \geqslant 1 - \alpha\lambda - Ch^{1/16}\big) \geqslant \lambda\Big) \geqslant 1 - Ch,
\end{aligned}$$

where $\bar{e}^N(t+h) - \bar{e}^N(t-) = \langle \bar{\mu}_N, m_{t+h} - m_{t-}\rangle$.

By the Skorohod representation theorem, we can assume that $\bar{\mu}_N$ converges almost surely to μ. Choosing t and $t+h$ in $J^{\complement}$, we make use of Lemma 5.9 below. It says that $\lim_{N\to+\infty}\langle \bar{\mu}_N, m_{t+h} - m_{t-}\rangle = \langle \mu, m_{t+h} - m_t\rangle$ (a.e. under Π_∞). Moreover, as $t \notin J$, Theorem 4.2 says that the law of $z_{t-} - m_{t-}$ under $\bar{\mu}_N$ converges to the law of $z_{t-} - m_{t-}$ under μ (for almost all μ under Π_∞). Therefore, following the proof of the Portmanteau theorem and modifying the constant C if necessary, we get

$$\begin{aligned}
&\Pi_\infty\big(\langle \mu, m_{t+h} - m_t\rangle > 1 + Ch^{1/16}\big) \leqslant Ch,\\
&\Pi_\infty\Big(\forall \lambda \leqslant \langle \mu, m_{t+h} - m_t\rangle - Ch^{1/16},\ \mu\big(z_{t-} - m_{t-} \geqslant 1 - \alpha\lambda - Ch^{1/16}\big) \geqslant \lambda\Big) \geqslant 1 - Ch.
\end{aligned} \tag{5.19}$$

The above inequalities are true for any $t, t+h$ that are not in J. Assume now that t is some point in $[0,T) \cap J$. Then, we can find sequences $(t_p)_{p\geqslant 1}$ and $(h_p)_{p\geqslant 1}$ such that $0 \leqslant t_p < t < t_p + h_p < T$, t_p and $t_p + h_p \notin J$, and $t_p \uparrow t$, $h_p \downarrow 0$. Then, applying (5.19) to any $(t_p, t_p + h_p)$ and letting p tend to $+\infty$, we deduce that

$$\begin{aligned}
&\Pi_\infty\big(\langle \mu, m_t - m_{t-}\rangle > 1\big) = 0,\\
&\Pi_\infty\Big(\forall \lambda < \langle \mu, m_t - m_{t-}\rangle, \quad \mu\big(z_{t-} - m_{t-} \geqslant 1 - \alpha\lambda\big) \geqslant \lambda\Big) = 1.
\end{aligned} \tag{5.20}$$

The first equality shows that under μ the canonical process satisfies condition (2) of Definition 2.2 (a.e. under Π_∞). Moreover, since J is countable, we deduce that

$$\Pi_\infty\Big(\forall t \in [0,T)\cap J, \quad \forall \lambda < \langle \mu, m_t - m_{t-}\rangle, \quad \mu\big(z_{t-} - m_{t-} \geqslant 1-\alpha\lambda\big) \geqslant \lambda\Big) = 1,$$

which is enough to conclude that condition (4) of Definition 2.2 is also satisfied (a.e. under Π_∞), by invoking Proposition 2.7.

5.4. **Proof of Theorem 4.7.** Assume now that (2.6) admits a unique solution $(Z_t)_{t\in[0,T]}$ on $[0,T]$. By identification, we deduce from the proof of Theorem 4.4 that, for Π_∞ a.e. μ, the pair $(z_t, m_t)_{t\in[0,T)}$ has the same law, under μ, as the pair $(Z_t, M_t)_{t\in[0,T)}$ under $\mathbb{P}$. In particular, for some fixed $S \in (0,T)$ such that $\mathbb{E}(M_S) = \mathbb{E}(M_{S-})$, we have $\langle \mu, m_S\rangle = \langle \mu, m_{S-}\rangle$ and $z_S = z_{S-}$ for Π_∞ a.e. μ. Define the map

$$\gamma_S : \hat{\mathcal{D}}([0,T+1],\mathbb{R}) \to \hat{\mathcal{D}}([0,S],\mathbb{R}), \quad \gamma_S(z) = \hat{z}, \tag{5.21}$$

where $\hat{z}$ is the element of $\hat{\mathcal{D}}([0,S],\mathbb{R})$ given by z_t if $t < S$ and $\hat{z}_S = z_{S-}$. The map γ_S is measurable for the σ-fields on $\hat{\mathcal{D}}([0,T+1])$ and $\hat{\mathcal{D}}([0,S])$ generated by the evaluation mappings and thus for the Borel σ-fields generated by M1. Moreover, the push forward of μ by γ_S, denoted by $\gamma_S\sharp\mu$, coincides, for Π_∞ a.e. μ, with the law of $(Z_t)_{t\in[0,S]}$ under $\mathbb{P}$.

Thus, defining $\Gamma_S(\mu) = \gamma_S\sharp\mu$ for an arbitrary $\mu \in \mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$, we have that $\Gamma_S\sharp\Pi_\infty = \mathrm{Dirac}(\mu_S)$, where $\mu_S := \mathrm{Law}((Z_s)_{t\in[0,S]})$. Notice that Γ_S is indeed measurable when both $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$ and $\mathcal{P}(\hat{\mathcal{D}}([0,S],\mathbb{R}))$ are equipped with the Borel σ-fields generated by the topology of weak convergence.

Moreover, as a consequence of Theorem 4.2, γ_S is continuous at any $z \in \hat{\mathcal{D}}([0,T+1],\mathbb{R})$ such that $z_{S-} = z_S$. In particular, if μ is a probability measure on $\hat{\mathcal{D}}([0,T+1])$ such that $\mu\{z_{S-} = z_S\} = 1$, then μ is a point of continuity of Γ_S. Therefore, under the probability Π_∞, a.e. μ is a continuity point of Γ_S. Since $(\Pi_N)_{N\geqslant 1}$ converges towards Π_∞ in the weak sense, we deduce that

$$\lim_{N\to+\infty} \Gamma_S\sharp\Pi_N = \Gamma_S\sharp\Pi_\infty = \mathrm{Dirac}(\mu_S).$$

Since (with $\hat{Z}^{i,N} = \gamma_S(Z^{i,N})$ given by (5.21))

$$\Gamma_S\sharp\Pi_N = \mathbb{P}_{\gamma_S\sharp(N^{-1}\sum_{i=1}^N \mathrm{Dirac}(\widetilde{Z}^{i,N}))} = \mathbb{P}_{N^{-1}\sum_{i=1}^N \mathrm{Dirac}(\hat{Z}^{i,N})},$$

we obtain that, in the weak sense,

$$\lim_{N\to+\infty} \mathbb{P}_{N^{-1}\sum_{i=1}^N \mathrm{Dirac}(\hat{Z}^{i,N})} = \mathrm{Dirac}(\mu_S), \tag{5.22}$$

where $N^{-1}\sum_{i=1}^N \mathrm{Dirac}(\hat{Z}^{i,N})$ and μ_S are seen as probability measures on $\hat{\mathcal{D}}([0,S],\mathbb{R})$ equipped with M1. Since the law of the N-tuple $(\hat{Z}^{1,N},\dots,\hat{Z}^{N,N})$ is invariant by permutation, we deduce from [17, Proposition 2.2] that the family $(\hat{Z}^{i,N})_{i=1,\dots,N}$ is chaotic on $\hat{\mathcal{D}}([0,S],\mathbb{R})$ endowed with M1, that is, for any integer $k \geqslant 1$,

$$\big(\hat{Z}^{1,N},\dots,\hat{Z}^{k,N}\big) \Rightarrow \big(\mu_S\big)^{\otimes k} = \mathbb{P}^{\otimes k}_{(Z_s)_{s\in[0,S]}} \quad \text{as } N \to +\infty, \tag{5.23}$$

in the weak sense, on $[\hat{\mathcal{D}}([0,S],\mathbb{R})]^k$ equipped with the product topology induced by M1. By Lemma 5.6 below, this also proves that

$$\big((\hat{Z}^{1,N},\hat{M}^{1,N}),\dots,(\hat{Z}^{k,N},\hat{M}^{k,N})\big) \Rightarrow \mathbb{P}^{\otimes k}_{(Z_s,\hat{M}_s)_{s\in[0,S]}} \quad \text{as } N \to +\infty, \tag{5.24}$$

on $[\hat{\mathcal{D}}([0,S],\mathbb{R})\times\hat{\mathcal{D}}([0,S],\mathbb{R})]^k$ equipped with the product topology induced by M1. Indeed, assuming without loss of generality that the sequence representing the convergence in (5.23) in the almost-sure sense is $(\hat{Z}^{1,N},\dots,\hat{Z}^{k,N})$ itself and denoting the a.s. limit by $(Z^{1,\infty},\dots,Z^{k,\infty})$, Lemma 5.6 says that, for any $\ell=1,\dots,k$, a.s., for t in a dense subset of $[0,S]$

$$\lfloor\big(\sup_{s\in[0,t]}\hat{Z}^{\ell,N}_s\big)_+\rfloor \to \lfloor\big(\sup_{s\in[0,t]}Z^{\ell,\infty}_s\big)_+\rfloor. \tag{5.25}$$

Obviously, (5.25) holds at $t=0$ since both sides are zero. At $t=S$, we know that $\mathbb{E}(M^\ell_S)=\mathbb{E}(M^\ell_{S-})$ since $t\mapsto\mathbb{E}(M_t)$ is continuous at $t=S$, so that $\mathbb{P}(M_S=M_{S-}=\hat{M}_S)=1$. Therefore, by Lemma 5.6, (5.25) holds at $t=S$. By Theorem 4.2, we deduce that, for any $\ell=1,\dots,k$, $((\hat{M}^{\ell,N}_s)_{s\in[0,S]})_{N\geqslant 1}$ converges a.s. to $(M^{\ell,\infty}_s)_{s\in[0,S]}$ in $\hat{\mathcal{D}}([0,S],\mathbb{R})$, where $(M^{\ell,\infty}_s)_{s\in[0,S]}$ is the counting process associated with $(Z^{\ell,\infty}_s)_{s\in[0,S]}$. Actually, the Skorohod representation theorem says that the a.s. convergence holds for a representation sequence only, but, in any case, (5.24) holds.

To complete the proof of Theorem 4.7, we use the Skorohod representation theorem once again. In fact we can assume without loss of generality that the convergence in (5.22) holds almost-surely, namely $N^{-1}\sum_{i=1}^N \mathrm{Dirac}(\hat{Z}^{i,N})$ converges to μ_S almost surely. Lemma 5.9 below then guarantees that a.s.

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{i=1}^N M^{i,N}_s=\mathbb{E}(M_s).$$

for all $s\in[0,S)$, except at any points of discontinuity of $\cdot\mapsto\mathbb{E}(M_\cdot)$ (of which there are countably many). Actually, the Skorohod representation theorem says that convergence holds almost-surely for a representation sequence only. Anyhow, the convergence always holds in probability (using for instance the fact that M1 convergence is metrizable). Since S can be chosen as close as needed to T, we can apply the above result to some $S'\in(S,T)$, where S' is a continuity point of $\cdot\mapsto\mathbb{E}(M_\cdot)$. This says that the limit holds for all continuity points $s\in[0,S]$ since $S\in[0,S')$, and S is also a continuity point. By Theorem 4.2, we deduce that the mapping $[0,S]\ni s\mapsto N^{-1}\sum_{i=1}^N\hat{M}^{i,N}_s$ converges to the mapping $[0,S]\ni s\mapsto\mathbb{E}(M_s)$ in probability on $\hat{\mathcal{D}}([0,S],\mathbb{R})$ equipped with the M1 topology, where $\hat{M}^{i,N}_s=M^{i,N}_s$ for $s\in[0,S)$ and $\hat{M}^{i,N}_S=M^{i,N}_{S-}$.

5.5. **Continuity of related mappings.** The aim of this section is to complete the proof of Theorems 4.4 and 4.7 by providing the technical results needed therein. In particular, we prove the key continuity result used in the proof of Theorem 4.4 i.e. the continuity of the functional given in (5.18) (see Lemma 5.10).

In the whole section, S is a given positive real, and we make use of the notation and definitions of Section 4.1 for the M1 topology. Moreover, as above, $(z_t)_{t\in[0,S]}$ will be the canonical process on $\hat{\mathcal{D}}([0,S],\mathbb{R})$ and $m_t:=\lfloor\big(\sup_{s\in[0,t]}z_s\big)_+\rfloor,\quad t\in[0,S]$.

We begin with the following continuity property.

Lemma 5.6. *Consider a sequence $(z^n)_{n\geqslant 1}$ of functions in $\hat{\mathcal{D}}([0,S],\mathbb{R})$, converging towards some $z\in\hat{\mathcal{D}}([0,S],\mathbb{R})$ for M1. Assume that z has the following* crossing property*:*

$$\forall k\in\mathbb{N}^*,\ \forall h>0,\quad \tau^k<S\Rightarrow\sup_{t\in[\tau^k,\min(S,\tau^k+h)]}\big[z_t-z_{\tau^k}\big]>0, \tag{5.26}$$

where $\tau^k = \inf\{t \in [0,S] : z_t \geqslant k\}$ *(*$\inf \emptyset = S$*). Then, there exists an at most countable subset* $J \subset [0,S]$*, such that*

$$\forall t \in [0,S] \setminus J, \quad \lim_{n \to +\infty} m_t^n = m_t,$$

where $m_t^n = \lfloor (\sup_{s \in [0,t]} z_s^n)_+ \rfloor$, $m_t = \lfloor (\sup_{s \in [0,t]} z_s)_+ \rfloor$*. The set* $[0,S] \setminus J$ *contains all the points* t *of continuity of* z *such that* $(\sup_{s \in [0,t]} z_s)_+$ *is not in* $\mathbb{N} \setminus \{0\}$*.*

Proof. Since $z^n \to z$ for the M1 topology, we can find a sequence of parametric representations $((u^n, r^n))_{n \geqslant 1}$ of $(z^n)_{n \geqslant 1}$ that converges towards a parametric representation (u, r) of z uniformly on $[0,S]$ (see Theorem 12.5.1 in [19]). For any $t \in [0,S]$, we thus have

$$\lim_{n \to +\infty} \sup_{s \leqslant t} u_s^n = \sup_{s \leqslant t} u_s. \tag{5.27}$$

Fix $t \in [0,S]$. Since $\mathcal{G}_{z^n} \subset (u^n, r^n)([0,S])$, we know that, for every $s \in [0,S]$, there exists $s' \in [0,S]$ such that $(u_{s'}^n, r_{s'}^n) = (z_s^n, s)$. If $s < r_t^n$, it therefore must hold that $s' < t$, as $s' \geqslant t$ would imply $s = r_{s'}^n \geqslant r_t^n$ (r^n is non-decreasing). Therefore, $\sup_{s \leqslant t} u_s^n \geqslant \sup_{s < r_t^n} z_s^n$.

Moreover, for any $s \in [0,S]$, $u_s^n \in [z_{r_s^n -}^n, z_{r_s^n}^n]$, so that $u_s^n \leqslant \max(z_{r_s^n -}^n, z_{r_s^n}^n)$. Therefore, for $s \leqslant t$,

$$u_s^n \leqslant \max\big(\sup_{s' \leqslant r_t^n} z_{s'-}^n, \sup_{s' \leqslant r_t^n} z_{s'}^n\big) = \sup_{s' \leqslant r_t^n} z_{s'}^n.$$

In the end we see that $\sup_{s < r_t^n} z_s^n \leqslant \sup_{s \leqslant t} u_s^n \leqslant \sup_{s \leqslant r_t^n} z_s^n$. Similarly, we have $\sup_{s < r_t} z_s \leqslant \sup_{s \leqslant t} u_s \leqslant \sup_{s \leqslant r_t} z_s$. Thus, by (5.27),

$$\liminf_{n \to +\infty} \sup_{s \leqslant r_t^n} z_s^n \geqslant \sup_{s < r_t} z_s, \quad \limsup_{n \to +\infty} \sup_{s < r_t^n} z_s^n \leqslant \sup_{s \leqslant r_t} z_s. \tag{5.28}$$

If r_t is a point of continuity of z, the two right-hand sides above coincide, and moreover, by Theorem 4.2,

$$\lim_{\delta \to 0} \limsup_{n \to +\infty} \sup_{\max(0, r_t - \delta) \leqslant s \leqslant \min(S, r_t + \delta)} |z_s^n - z_{r_t}| = 0.$$

That is, for any $\varepsilon > 0$, we can find $\delta > 0$, such that, for n large enough,

$$\sup_{\max(0, r_t - \delta) \leqslant s \leqslant \min(S, r_t + \delta)} |z_s^n - z_{r_t}| \leqslant \varepsilon. \tag{5.29}$$

In particular, for n large enough that $|r_t^n - r_t| < \delta$,

$$\begin{aligned} \sup_{s \leqslant r_t} z_s^n &= \max\Big(\sup_{s \leqslant \max(0, r_t - \delta)} z_s^n, \sup_{\max(0, r_t - \delta) \leqslant s \leqslant r_t} z_s^n\Big) \\ &\leqslant \max\Big(\sup_{s \leqslant \max(0, r_t - \delta)} z_s^n, z_{r_t}\Big) + \varepsilon \leqslant \max\Big(\sup_{s < r_t^n} z_s^n, z_{r_t}\Big) + \varepsilon. \end{aligned} \tag{5.30}$$

Therefore, by (5.28),

$$\limsup_{n \to +\infty} \sup_{s \leqslant r_t} z_s^n \leqslant \sup_{s \leqslant r_t} z_s. \tag{5.31}$$

Similarly, for n large enough, we have

$$\begin{aligned} \sup_{s \leqslant r_t} z_s^n &\geqslant \max\big(\sup_{s \leqslant \max(0, r_t - \delta)} z_s^n, z_{r_t}\big) - \varepsilon \\ &\geqslant \max\big(\sup_{s \leqslant \max(0, r_t - \delta)} z_s^n, \sup_{\max(0, r_t - \delta) \leqslant s \leqslant r_t^n} z_s^n\big) - 2\varepsilon \geqslant \sup_{s \leqslant r_t^n} z_s^n - 2\varepsilon. \end{aligned}$$

By (5.28) again, $\liminf_{n \to +\infty} \sup_{s \leqslant r_t} z_s^n \geqslant \sup_{s < r_t} z_s$. By (5.31), we deduce that

$$\lim_{n \to +\infty} \sup_{s \leqslant r_t} z_s^n = \sup_{s \leqslant r_t} z_s,$$

whenever r_t is a continuity point of z. Since r maps $[0,S]$ onto itself, we deduce that for any continuity point $t \in [0,S]$ of z,

$$\lim_{n\to+\infty} \sup_{s \leqslant t} z^n_s = \sup_{s \leqslant t} z_s.$$

Now, if $(\sup_{s \leqslant t} z_s)_+$ is not an integer, we deduce that

$$\lim_{n\to+\infty} \lfloor \big(\sup_{s \leqslant t} z^n_s\big)_+ \rfloor = \lfloor \big(\sup_{s \leqslant t} z_s\big)_+ \rfloor.$$

If $(\sup_{s \leqslant t} z_s)_+$ is an integer, there are two cases. If the integer is zero, it means that $\sup_{s \leqslant t} z_s \leqslant 0$, so that $\lim_{n\to+\infty} \sup_{s \leqslant t} z^n_s \leqslant 0$ and thus

$$\lim_{n\to+\infty} (\sup_{s \leqslant t} z^n_s)_+ = 0.$$

In this case, $\lim_{n\to+\infty} \lfloor(\sup_{s \leqslant t} z^n_s)_+\rfloor = 0$. If the integer is not zero and $t < S$, the crossing property (5.26) says that t must be the first time when this integer is crossed (rather than just touched), so that the number of points of continuity of z for which the convergence of the integer parts can fail is finite. ☐

The following corollary is simply a result of Lebesgue's dominated convergence theorem.

Corollary 5.7. *Under the assumptions of Lemma 5.6, the functions* $[0,S] \ni t \mapsto \int_0^t b\big(z^n_s - m^n_s\big)ds$ *converge in* $\mathcal{C}([0,S],\mathbb{R})$ *towards* $[0,S] \ni t \mapsto \int_0^t b\big(z_s - m_s\big)ds$.

The next result guarantees the convergence of the expectation of m_t whenever $\mu^n \to \mu$ and the canonical process z satisfies the crossing property of Lemma 5.6 under μ, under appropriate assumptions.

Proposition 5.8. *Assume that* $(\mu^n)_{n \geqslant 1}$ *is a sequence of probability measures on the space* $\hat{\mathcal{D}}([0,S],\mathbb{R})$*, converging towards some* μ *for the weak topology deriving from M1. If* $\sup_{n \geqslant 1} \langle \mu^n, \sup_{t\in[0,S]} |z_t|^2 \rangle < +\infty$*, then, at any point of continuity of the mapping* $(0,S) \ni t \mapsto \langle \mu, m_t\rangle$*, it holds that*

$$\lim_{n\to+\infty} \langle \mu^n, m_t\rangle = \langle \mu, m_t\rangle,$$

provided that, for any integer $k \geqslant 1$ *and for* $\tau^k := \inf\{t \in [0,S] : z_t \geqslant k\}$,

$$\forall h > 0, \quad \mu\Big\{\tau^k < S, \sup_{s\in[\tau^k,\min(S,\tau^k+h)]} (z_s - z_{\tau^k}) = 0\Big\} = 0, \tag{5.32}$$

Proof. Let $\delta > 0$. Then, for every $n \geqslant 1$, the function

$$[0,S] \ni t \mapsto \bar{m}^n_t = \begin{cases} \langle \mu^n, m_\delta\rangle, & t \in [0,\delta], \\ \langle \mu^n, m_t\rangle, & t \in [\delta, S-\delta], \\ \langle \mu^n, m_{S-\delta}\rangle, & t \in [S-\delta, S] \end{cases}$$

is (deterministic) non-decreasing nonnegative and càdlàg such that $\sup_{n \geqslant 1} \bar{m}^n_S < +\infty$.

Since each $\bar{m}^n$ is constant on both $[0,\delta]$ and $[S-\delta,S]$, the sequence $(\bar{m}^n)_{n \geqslant 1}$ is relatively compact for the M1 topology by Theorem 4.3. Extracting a convergent subsequence, still indexed by n, and denoting the limit by $\bar{m}$, we have, for any bounded and measurable function f from $[0,S]$ into $\mathbb{R}$ vanishing outside $(\delta, S-\delta)$, that

$$\lim_{n\to+\infty} \int_0^S f_t \bar{m}^n_t dt = \int_0^S f_t \bar{m}_t dt.$$

This follows from the dominated convergence theorem and the fact that $(\bar{m}^n)_{n \geqslant 1}$ converges pointwise towards $\bar{m}$ on a subset of $[0, S]$ of full Lebesgue measure, see Theorem 4.1. Since f vanishes outside $(\delta, S-\delta)$, we can forget the definition of $\bar{m}^n$ outside $(\delta, S-\delta)$. Therefore, by the Skorohod representation theorem

$$\int_0^S f_t \bar{m}_t^n dt = \mathbb{E} \int_0^S f_t \big\lfloor \big(\sup_{s \in [0,t]} \zeta_s^n \big)_+ \big\rfloor dt,$$

where $(\zeta^n)_{n \geqslant 1}$ is a sequence of processes distributed according to $(\mu^n)_{n \geqslant 1}$ and converging a.s. for M1 towards some process ζ with law μ. By assumption, we know that, a.s.,

$$\forall k \geqslant 1, \quad \forall h > 0, \qquad \tau^k < S \Rightarrow \sup_{s \in [\tau^k, \min(S, \tau^k + h)]} (\zeta_s - \zeta_{\tau^k}) > 0,$$

so that almost all paths of ζ satisfy the assumption of Lemma 5.6. We deduce that, a.s., the functions $[0, S] \ni t \mapsto \lfloor (\sup_{s \in [0,t]} \zeta_s^n)_+ \rfloor$ converge pointwise towards $[0, S] \ni t \mapsto \lfloor (\sup_{s \in [0,t]} \zeta_s)_+ \rfloor$ on the complement of an at most countable set. Therefore, a.s.

$$\lim_{n \to +\infty} \int_0^S f_t \big\lfloor \big(\sup_{s \in [0,t]} \zeta_s^n \big)_+ \big\rfloor dt = \int_0^S f_t \big\lfloor \big(\sup_{s \in [0,t]} \zeta_s \big)_+ \big\rfloor dt,$$

by the Lebesgue dominated convergence theorem. Using this result again, we deduce that

$$\lim_{n \to +\infty} \mathbb{E} \left[\int_0^S f_t \big\lfloor \big(\sup_{s \in [0,t]} \zeta_s^n \big)_+ \big\rfloor dt \right] = \mathbb{E} \left[\int_0^S f_t \big\lfloor \big(\sup_{s \in [0,t]} \zeta_s \big)_+ \big\rfloor dt \right] = \int_0^S f_t \langle \mu, m_t \rangle dt,$$

the last equality following from the fact that the law of $(\zeta_s)_{0 \leqslant s \leqslant S}$ is μ. By right-continuity, this proves that $\bar{m}_t = \langle \mu, m_t \rangle$ for any $t \in (\delta, S - \delta)$. Therefore, at any point of continuity of $(\delta, S - \delta) \ni t \mapsto \langle \mu, m_t \rangle$, Theorem 4.2 says that

$$\lim_{n \to +\infty} \langle \mu^n, m_t \rangle = \langle \mu, m_t \rangle.$$

Letting δ tend to 0, we complete the proof, as the set of discontinuity points of $[0, S] \ni t \mapsto \langle \mu, m_t \rangle$ is at most countable. □

The previous result is general. In order to be able to apply it, we need to check that whenever Π_∞ is a limit point of the family $(\Pi_N)_{N \geqslant 1}$ (see Theorem 4.4), any measure μ in the support of Π_∞ must satisfy the crossing property.

Lemma 5.9. *For a.e. μ under Π_∞, for any integer $k \geqslant 1$ and any real $h > 0$,*

$$\mu \Big\{ \tau^k < T + 1, \sup_{s \in [\tau^k, \min(T+1, \tau^k + h)]} (z_s - z_{\tau^k}) = 0 \Big\} = 0, \quad \text{with } \tau^k = \inf\{t \geqslant 0 : z_t \geqslant k\}.$$

In particular, if $(\mu^n)_{n \geqslant 1}$ is a sequence of probability measures on the space $\hat{\mathcal{D}}([0, S], \mathbb{R})$, converging towards μ for the weak topology deriving from M1, then, at any point of continuity of the mapping $(0, S) \ni t \mapsto \langle \mu, m_t \rangle$, it holds that $\lim_{n \to +\infty} \langle \mu^n, m_t \rangle = \langle \mu, m_t \rangle$.

Proof. In order to prove this result, we actually will introduce a second empirical measure

$$\bar{\theta}_N := \frac{1}{N} \sum_{i=1}^N \mathrm{Dirac}(\widetilde{Z}^{i,N}, W^i).$$

We will consider $\bar{\theta}_N$ as a random probability measure on $\hat{\mathcal{D}}([0, T+1], \mathbb{R}) \times \mathcal{C}([0, T+1], \mathbb{R})$, endowed with the product of the M1 topology and of the standard topology of uniform convergence. Note that the marginal distribution of $\bar{\theta}_N$ on the first coordinate is $\bar{\mu}_N$. We also define $\Xi_N := \mathrm{Law}(\bar\theta_N)$. Following Lemma 5.5, we have that the family $(\Xi_N)_{N\geqslant 1}$ is tight on $\mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R})\times\mathcal{C}([0,T+1],\mathbb{R}))$ (endowed with the topology of weak convergence), so that we can extract a convergent subsequence of $(\Xi_N)_{N\geqslant 1}$, still indexed by N. We denote its limit by Ξ_∞.

Returning to the particle system, we first claim that there exists a constant $c\geqslant 0$ such that, for any $N\geqslant 1$ and $i=1,\dots,N$,

$$\widetilde{Z}^{i,N}_t-\widetilde{Z}^{i,N}_s\geqslant W^i_t-W^i_s-c\big(1+\sup_{v\in[0,T]}|Z^{i,N}_v|\big)(t-s),$$

for $s,t\in[0,T+1]$ with $s\leqslant t$. Indeed this follows from the fact that b is Lipschitz and that $M^{i,N}_t\leqslant\sup_{v\in[0,T]}|Z^{i,N}_v|$, for all $t\in[0,T]$ and every $i=1,\dots,N$.

Denoting by $(z_t,w_t)_{t\in[0,T+1]}$ the canonical process on $\hat{\mathcal{D}}([0,T+1],\mathbb{R})\times\mathcal{C}([0,T+1],\mathbb{R})$, we get that, $\mathbb{P}$-a.s.,

$$z_t-z_s\geqslant w_t-w_s-c\big(1+\sup_{v\in[0,T+1]}|z_v|\big)(t-s),\quad 0\leqslant s\leqslant t\leqslant T+1. \tag{5.33}$$

holds under the empirical measure $\bar\theta_N$, or that (5.33) holds a.s. under a.e. probability measure θ under Ξ_N. Since $(\Xi_N)_{N\geqslant 1}$ converges towards Ξ_∞ in the weak sense, we know from the Skorohod representation theorem that, on some probability space (still denoted by $(\Omega,\mathcal{A},\mathbb{P})$), there exists a sequence of random probability measures $(\vartheta_N)_{N\geqslant 1}$ on the space $\hat{\mathcal{D}}([0,T+1],\mathbb{R})\times\mathcal{C}([0,T+1],\mathbb{R})$, such that $(\vartheta_N)_{N\geqslant 1}$ converges towards some random probability measure ϑ_∞ a.s., with ϑ_N being distributed according to Ξ_N and ϑ_∞ according to Ξ_∞. Hence, a.s., under ϑ_N the canonical process (z,w) has the property (5.33).

Step 1: The first step is to show that under ϑ_∞ the canonical process (z,w) also satisfies (5.33), simply using the facts that it is true under ϑ_N for each N, and that ϑ_N converges to ϑ_∞ a.s. Again, we can use the Skorohod representation theorem (still with $(\Omega,\mathcal{A},\mathbb{P})$ as the underlying probability space). Indeed, we can find a sequence of processes (ζ^N,ξ^N) with law ϑ_N under $\mathbb{P}$, converging a.s. towards some process (ζ,ξ) distributed according to ϑ_∞. Since (5.33) holds under ϑ_N for each N, we have that $\mathbb{P}$ a.s., for any $0\leqslant s\leqslant t\leqslant T+1$,

$$\zeta^N_t-\zeta^N_s\geqslant\xi^N_t-\xi^N_s-c\big(1+\sup_{v\in[0,T+1]}|\zeta^N_v|\big)(t-s). \tag{5.34}$$

We want to prove that, $\mathbb{P}$ a.s., for any $s\leqslant t$,

$$\zeta_t-\zeta_s\geqslant\xi_t-\xi_s-c\big(1+\sup_{v\in[0,T+1]}|\zeta_v|\big)(t-s). \tag{5.35}$$

It is sufficient to prove that, for an arbitrary sequence $(\zeta^N,\xi^N)_{N\geqslant 1}$ satisfying (5.34) and converging towards (ζ,ξ) in $\hat{\mathcal{D}}([0,T+1],\mathbb{R})\times\mathcal{C}([0,T+1],\mathbb{R})$ equipped with the product topology derived from the M1 and uniform topologies, the limit (ζ,ξ) satisfies (5.35).

To this end, if $\zeta^N\to\zeta$ in $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$ with respect to the M1 topology, then by [19, Theorem 12.5.1], there exist parametric representations (u^N,r^N) of ζ^N and (u,r) of ζ (see Section 4.1 for definition) such that $\|u^N-u\|\vee\|r^N-r\|\to 0$, where

$$\forall t\in[0,T+1],\quad u^N_t\in\big[\zeta^N_{r^N_t-},\zeta^N_{r^N_t}\big],\quad u_t\in\big[\zeta_{r_t-},\zeta_{r_t}\big].$$

Noting that $\sup_{0\leqslant v\leqslant T+1}|\zeta^N_v|=\sup_{0\leqslant v\leqslant T+1}|u^N_v|$, we have by (5.34), for any $s\leqslant t$,

$$\max\big(\zeta^N_{r^N_t-},\zeta^N_{r^N_t}\big)-\min\big(\zeta^N_{r^N_s-},\zeta^N_{r^N_s}\big)\geqslant\xi^N_{r^N_t}-\xi^N_{r^N_s}-c\big(1+\sup_{v\in[0,T+1]}|u^N_v|\big)\big(r^N_t-r^N_s\big),$$

since the right-hand side is continuous in the subscript parameter. Thus

$$u_t^N - u_s^N \geqslant \xi^N_{r_t^N} - \xi^N_{r_s^N} - c\big(1 + \sup_{v\in[0,T+1]} |u_v^N|\big)\big(r_t^N - r_s^N\big).$$

Passing to the limit as $N \to \infty$ and using the continuity of ξ, we obtain, for any $s \leqslant t$,

$$u_t - u_s \geqslant \xi_{r_t} - \xi_{r_s} - c\big(1 + \sup_{v\in[0,T+1]} |u_v|\big)(r_t - r_s).$$

For any $t', s' \in [0, T+1]$, $s' \leqslant t'$, we can find $t, s \in [0, T+1]$ with $s \leqslant t$ such that $\zeta_{t'} = u_t, \zeta_{s'} = u_s$ and $t' = r_t, s' = r_s$. This proves that

$$\zeta_{t'} - \zeta_{s'} \geqslant \xi_{t'} - \xi_{s'} - c\big(1 + \sup_{v\in[0,T+1]} |\zeta_v|\big)(t' - s'),$$

and thus that (ζ, ξ) satisfies (5.35).

Step 2: We now use the previous step to prove the lemma. Indeed, by Step 1 we have that (5.33) holds a.s. under θ, for almost all θ under Ξ_∞. Thus for almost all θ under Ξ_∞, we have for any real R, integer k and $h > 0$

$$\theta\big\{\tau^k < T+1, \sup_{s\in[\tau^k,\min(T,\tau^k+h)]} (z_s - z_{\tau^k}) = 0\big\} \leqslant \theta\big\{\sup_{v\in[0,T+1]} |z_v| \geqslant R\big\}$$
$$+\, \theta\big\{\tau^k < T+1, \sup_{s\in[\tau^k,\min(T,\tau^k+h)]} [w_s - w_{\tau^k} - c(1+R)(s - \tau^k)] = 0\big\}.$$

Clearly, τ^k is a stopping time for the filtration generated by (z, w). Assume that w is a Brownian motion w.r.t. this filtration under θ (this is proved below). Then the strong Markov property says that the second term in the right-hand side of the above is zero for any $h > 0$. Letting R tend to $+\infty$, we get

$$\theta\big\{\tau^k < T+1, \sup_{s\in[\tau^k,\min(T,\tau^k+h)]} (z_s - z_{\tau^k}) = 0\big\} = 0 \quad \text{for a.e. } \theta \text{ under } \Xi_\infty. \tag{5.36}$$

We finally claim that (5.36) then also holds for all μ under Π_∞. For $\theta \in \mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}) \times \mathcal{C}([0,T+1],\mathbb{R}))$ let $\Psi(\theta) \in \mathcal{P}(\hat{\mathcal{D}}([0,T+1],\mathbb{R}))$ be the marginal of θ on the first coordinate. Since $\theta \mapsto \Psi(\theta)$ is obviously continuous and $\Pi_N = \Psi\sharp\Xi_N$, where $\sharp$ indicates the push forward map, we have $\Pi_\infty = \Psi\sharp\Xi_\infty$. Then, for any Borel subset $A \subset \hat{\mathcal{D}}([0,T+1],\mathbb{R})$, $\int \theta\{(z_t)_{t\in[0,T+1]} \in A\} d\Xi_\infty(\theta) = \int \mu\{(z_t)_{t\in[0,T+1]} \in A\} d\Pi_\infty(\mu)$. Choosing $A = \{\tau^k < T+1, \sup_{s\in[\tau^k,\min(T,\tau^k+h)]}(z_s - z_{\tau^k}) = 0\}$, we complete the proof.

It thus remains to prove that w is a Brownian motion under the filtration generated by (z, w), for which it is sufficient to prove that

$$\int f(w_t - w_s) \prod_{i=1}^n g_i(z_{s_i}, w_{s_i}) d\theta(z,w) = \int f(w_t - w_s) d\theta(z,w) \int \prod_{i=1}^n g_i(z_{s_i}, w_{s_i}) d\theta(z,w),$$

for bounded and continuous functions g_i and f, and for $0 \leqslant s_1 < \cdots < s_n \leqslant s < t$ i.e. w is a continuous martingale. This follows from the convergence of the finite-dimensional distributions up to a countable subset under the weak convergence for the M1 Skorokhod topology (see [15, Theorem 3.2.2]) and from the right-continuity of the paths. □

We then finally arrive at the required continuity lemma:

Lemma 5.10. *Under* Π_∞*, the functional* (5.18) *is a.e. continuous.*

Proof. Let μ be in the support of Π_∞, and let $(\mu^n)_{n \geqslant 1}$ be a sequence of measures on $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$ converging towards μ (in the weak topology induced by the M1 topology). By choice of $S_1 < \cdots < S_p = S$ in the definition (5.16) of F, both the canonical process and $\langle \mu, m \rangle$ are a.s. continuous at $S_1, \dots, S_p$. We want to prove that

$$\lim_{n \to +\infty} \Big\langle \mu^n, F\Big(z_\cdot - z_0 - \int_0^\cdot b(z_s - m_s) ds - \alpha \langle \mu^n, m_\cdot \rangle \Big) \Big\rangle$$
$$= \Big\langle \mu, F\Big(z_\cdot - z_0 - \int_0^\cdot b(z_s - m_s) ds - \alpha \langle \mu, m_\cdot \rangle \Big) \Big\rangle.$$

Again, we make use of the Skorohod representation theorem. We consider a sequence of processes $(\zeta^n)_{n \geqslant 1}$, with $(\mu^n)_{n \geqslant 1}$ as distributions on $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$, converging a.s. towards some process ζ, admitting μ as distribution. We set, for any $t \in [0,T]$,

$$\eta^n_t = \big\lfloor \big(\sup_{s \in [0,t]} \zeta^n_s \big)_+ \big\rfloor, \quad \eta_t = \big\lfloor \big(\sup_{s \in [0,t]} \zeta_s \big)_+ \big\rfloor.$$

Since, a.s., S is a point of continuity of ζ, we deduce from Theorems 4.1 and 4.2 that, a.s., $(\hat{\zeta}^n_s)_{s \in [0,S]}$ converges towards $(\zeta_s)_{s \in [0,S]}$ in $\hat{\mathcal{D}}([0,S],\mathbb{R})$, where $\hat{\zeta}^n_s = \zeta^n_s$ if $s < S$ and $\hat{\zeta}^n_S = \zeta^n_{S-}$ as usual.

By Corollary 5.7, a.s.,

$$\lim_{n \to +\infty} \sup_{v \in [0,S]} \left| \int_0^v b(\zeta^n_s - \eta^n_s) ds - \int_0^v b(\zeta_s - \eta_s) ds \right| = 0.$$

Moreover, since $S_0, S_1, \dots, S_p$ are almost-surely points of continuity of ζ, we deduce from Theorem 4.2 that $\zeta^n_{S_i} \to \zeta_{S_i}$ as $n \to +\infty$, for any $i \in \{0, \dots, p\}$.

Similarly, since $S_1, \dots, S_p$ are points of continuity of $[0,T] \ni t \mapsto \langle \mu, m_t \rangle$ (which coincides with the expectation of η_t under μ), we deduce from Proposition 5.8 that $\langle \mu^n, m_{S_i} \rangle \to \langle \mu, m_{S_i} \rangle$ as $n \to +\infty$, for any $i \in \{1, \dots, p\}$. This proves the lemma.

□

5.6. **Proof of Theorems 4.6 and 4.9.** Theorems 4.6 and 4.9 are proved in a completely similar way to Theorems 4.4 and 4.7.

We first discuss the proof of Theorem 4.6, using the same notation as in the statement. Following the proof Lemma 5.1, we can prove that $\sup_{\delta \in (0,1)} \mathbb{E}[\sup_{t \in [0,T]} |Z^\delta_t|^p + (M^\delta_T)^p]$ is finite for any $p \geqslant 1$. Following the proof of [8, Lemma 5.2], we know that $e_\delta(t) \leqslant C t^{1/2}$, for a constant C independent of δ. Following the proof of Lemma 5.4, we deduce that the laws of $(\mu^\delta)_{\delta \in (0,1)}$ are tight in $\hat{\mathcal{D}}([0,T+1],\mathbb{R})$, where μ^δ is defined in the statement of Theorem 4.6. In order to pass to the limit along convergent subsequences, we consider, for a given weak limit μ, the countable set of points J in $[0,T+1]$ at which the function

$$[0,T+1] \ni t \mapsto \mu\big\{ z_{t-} = z_t \big\}$$

differs from one. Checking that μ satisfies the 'crossing' property in Lemma 5.9, it is then quite straightforward to pass to the limit in the identity

$$Z^\delta_{S_{i+1}} = Z^\delta_{S_i} + \int_{S_i}^{S_{i+1}} b\big(Z^\delta_s - M^\delta_s\big) ds + \alpha \langle \mu^\delta, m_{S_{i+1}} - m_{S_i} \rangle + W_{S_{i+1}} - W_{S_i},$$

for points $0 = S_0 < S_1 < \cdots < S_p < T$ that are not in J. This permits to prove that, under μ, the canonical process $(z_t)_{t \in [0,T]}$ satisfies (2) in Theorem 4.4.

The most difficult point is to check (3). It follows from Lemma 5.11 below, which is the counterpart of Proposition 5.3 for the particle system. The end of the proof is then similar. The proof of Theorem 4.9 works on the same model as the one of Theorem 4.7.

Lemma 5.11. *For a given $T > 0$, consider $0 \leqslant t < t+h \leqslant T$, $h \in (0,1)$. Then, we can find $C \geqslant 0$, independent of h, such that, for any $\delta \in (0,1)$,*

$$e_\delta(t+h) - e_\delta(t-) = e_\delta(t+h) - e_\delta(t) \leqslant 1 + Ch^{1/8},$$
$$\forall \lambda \leqslant e_\delta(t+h) - e_\delta(t), \quad \mathbb{P}\big(\, X^\delta_{t-} \geqslant 1 - \alpha\lambda - Ch^{1/8}\big) \geqslant \lambda + e_\delta(t) - e_\delta(t+\delta) - Ch^{1/8}.$$

Notice that the second statement in Lemma 5.11 slightly differs from the second statement in Proposition 5.3. However, the application for passing to the limit is the same. With the same notation as in (5.19), it says that, for $t, t+h, t+\delta' \in J^{\complement}$ and $\lambda < \langle \mu, m_{t+h} - m_t \rangle$, it holds $\mu(z_{t-} - m_{t-} \geqslant 1 - \alpha\lambda - Ch^{1/8}) \geqslant \lambda + \langle \mu, m_t - m_{t+\delta'} \rangle - Ch^{1/8}$. Letting $\delta' \downarrow 0$, it permits to recover (5.19).

Proof. Given some $\delta > 0$ and $0 \leqslant t < t+h \leqslant T$, consider $\sigma^\delta_t := \inf\{s \geqslant t : M^\delta_s - M^\delta_{t-} = 1\}$ and $\tau^\delta_t := \inf\{s > t : M^\delta_s - M^\delta_{t-} = 2\}$. As e_δ is continuously differentiable, we know from [8, Lemma 4.2] that σ^δ_t and τ^δ_t have differentiable cumulative distribution functions. Recalling the definition of (3.11) and setting $Z^\delta = X^\delta + M^\delta$ as usual, we see that

$$\sup_{s \in [t,t+h]} |Z^\delta_s - Z^\delta_t| \leqslant \alpha\big(e_\delta(t+h) - e_\delta(t)\big) + \sup_{s \in [t,t+h]} |W^\delta_s - W^\delta_t| + Ch\big(1 + \sup_{s \in [0,T]} |Z^\delta_s|\big).$$

On $\{\tau^\delta_t \leqslant t+h\}$, $(Z^\delta_s)_{s \in [t,t+h]}$ crosses at least two new integers, i.e. $\sup_{s \in [t,t+h]} |Z^\delta_s - Z^\delta_t| \geqslant 1$:

$$1 \leqslant \alpha\big(e_\delta(t+h) - e_\delta(t)\big) + \sup_{s \in [t,t+h]} |W^\delta_s - W^\delta_t| + Ch\big(1 + \sup_{s \in [0,T]} |Z^\delta_s|\big), \tag{5.37}$$

where (with the convention that $M^\delta_r = M^\delta_0$, $\sigma^\delta_r = \sigma^\delta_0$ and $\tau^\delta_r = \tau^\delta_0$ if $r \leqslant 0$),

$$\begin{aligned} e_\delta(t+h) - e_\delta(t) &\leqslant \mathbb{P}\big(M^\delta_{t+h-\delta} - M^\delta_{t-\delta} \geqslant 1\big) + \mathbb{E}\big[\big(M^\delta_{t+h-\delta} - M^\delta_{t-\delta}\big)\mathbf{1}_{\{\tau^\delta_{t-\delta} \leqslant t+h-\delta\}}\big] \\ &\leqslant \mathbb{P}\big(\sigma^\delta_{t-\delta} \leqslant t+h-\delta\big) + C\big(\mathbb{P}(\tau^\delta_{t-\delta} \leqslant t+h-\delta)\big)^{1/2} \\ &\leqslant 1 + C\big(\mathbb{P}(\tau^\delta_{t-\delta} \leqslant t+h-\delta)\big)^{1/2}. \end{aligned} \tag{5.38}$$

Therefore, taking the expectation in (5.37) on the event $\{\tau^\delta_t \leqslant t+h\}$, applying the Cauchy-Schwarz inequality and noticing that $\mathbb{P}(\tau^\delta_t \leqslant t+h) \neq 0$, we deduce:

$$1 \leqslant \alpha + C\big(\mathbb{P}(\tau^\delta_{t-\delta} \leqslant t-\delta+h)\big)^{1/2} + Ch^{1/2}\big[\mathbb{P}(\tau^\delta_t \leqslant t+h)\big]^{-1/2}.$$

For $t = 0$, this says that

$$1 \leqslant \alpha + C\big(\mathbb{P}(\tau^\delta_0 \leqslant h)\big)^{1/2} + Ch^{1/2}\big[\mathbb{P}(\tau^\delta_0 \leqslant h)\big]^{-1/2}.$$

We claim that (with C the constant found in the above equation)

$$\mathbb{P}\big(\tau^\delta_0 \leqslant h\big) \leqslant h^{1/4}, \quad \text{for } h \in [0, h_0), \quad \text{where } h_0 = \frac{1-\alpha}{2C}\,. \tag{5.39}$$

We argue by contradiction. If there exists $h^* \in (0, h_0)$ such that $\mathbb{P}(\tau^\delta_0 \leqslant h^*) > (h^*)^{1/4}$, we can take, by differentiability of the cumulative distribution function of τ^δ_0, h^* so that in fact equality is satisfied i.e. $\mathbb{P}(\tau^\delta_0 \leqslant h^*) = (h^*)^{1/4}$. We deduce that $1 \leqslant \alpha + 2C(h^*)^{1/8} < \alpha + 2Ch_0^{1/8}$, which is a contradiction.

For $t \in [0,\delta]$, we deduce that $\mathbb{P}(\tau^{\delta}_{t-\delta} \leqslant t-\delta+h) = \mathbb{P}(\tau^{\delta}_0 \leqslant h) \leqslant h^{1/4}$. Assume then that, for some integer $1 \leqslant k \leqslant \lceil T/\delta \rceil$ and for all $t \in [0,k\delta]$, we have $\mathbb{P}(\tau^{\delta}_{t-\delta} \leqslant t-\delta+h) \leqslant h^{1/4}$, for $h \in [0,h_0)$. We then claim that $\mathbb{P}(\tau^{\delta}_t \leqslant t+h) \leqslant h^{1/4}$ for all $t \in [0,(k\delta)\wedge T]$ and $h \in [0,h_0)$. This can be proved by contradiction again by considering h^* such that $\mathbb{P}(\tau^{\delta}_t \leqslant t+h^*) = (h^*)^{1/4}$. We finally deduce that

$$\mathbb{P}(\tau^{\delta}_t \leqslant t+h) \leqslant Ch^{1/4} \quad \text{so that} \quad e_\delta(t+h) - e_\delta(t) \leqslant 1 + Ch^{1/8}, \tag{5.40}$$

the second claim following from (5.38) and proving the first statement.

For the second statement, notice that, on the event $\{\sigma^{\delta}_t \leqslant s\}$, for $t \leqslant s \leqslant t+h$,

$$X^{\delta}_{\sigma^{\delta}_t -} - X^{\delta}_{t-} \leqslant \alpha\big(e_\delta(s) - e_\delta(t)\big) + \sup_{s\in[t,t+h]} |W_s - W_t| + C'h\big(1 + \sup_{s\in[t,t+h]} |Z^{\delta}_s|\big).$$

Since $X^{\delta}_{\sigma^{\delta}_t -} = 1$ (e_δ being continuous, see (3.13)), we have for $e_\delta(s) - e_\delta(t) \leqslant \lambda$:

$$\begin{aligned}
&\mathbb{P}\big(X^{\delta}_{t-} \geqslant 1 - \alpha\lambda - C'h^{1/8}\big) \\
&\geqslant \mathbb{P}\big(\sigma^{\delta}_t \leqslant s\big) - \mathbb{P}\big(\sup_{s\in[t,t+h]} |W_s - W_t| + C'h\big(1 + \sup_{s\in[t,t+h]} |Z^{\delta}_s|\big) \geqslant C'h^{1/8}\big) \\
&\geqslant \mathbb{P}\big(\sigma^{\delta}_t \leqslant s\big) - C'h \geqslant e_\delta(s+\delta) - e_\delta(t+\delta) - 2C'h^{1/8},
\end{aligned}$$

the last inequality following from the second line in (5.38) and from (5.40). Now, for $\lambda \leqslant e_\delta(t+h) - e_\delta(t)$, we can find $s^* \in [t,t+h]$ such that $\lambda = e_\delta(s^*) - e_\delta(t)$, which completes the proof since $e_\delta(s^*+\delta) - e_\delta(t+\delta) \geqslant \lambda + e_\delta(t) - e_\delta(t+\delta)$. □

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*, J. Math. Neurosci., 2 (2012), pp. Art. 10, 50.

[2] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[3] M. J. Cáceres, J. A. Carrillo, and B. Perthame, *Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states*, J. Math. Neurosci., 1 (2011), pp. Art. 7, 33.

[4] M. J. Cáceres and B. Perthame, *Beyond blow-up in excitatory integrate and fire neuronal networks: Refractory period and spontaneous activity*, J. Theor. Biol., 350 (2014), pp. 81 – 89.

[5] J. A. Carrillo, M. D. M. González, M. P. Gualdani, and M. E. Schonbek, *Classical solutions for a nonlinear Fokker-Planck equation arising in computational neuroscience*, Comm. Partial Differential Equations, 38 (2013), pp. 385–409.

[6] E. Catsigeras and P. Guiraud, *Integrate and fire neural networks, piecewise contractive maps and limit cycles*, J. Math. Biol., 67 (2013), pp. 609–655.

[7] A. De Masi, A. Galves, E. Löcherbach, and E. Presutti, *Hydrodynamic limit for interacting neurons*, J. Stat. Phys., (2014), pp. 1–37.

[8] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré, *Global solvability of a networked integrate-and-fire model of McKean-Vlasov type*. `arxiv:1211.0299v4`, to appear in Ann. Appl. Probab., 2014.

[9] S. N. Ethier and T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[10] O. Faugeras, J. Touboul, and B. Cessac, *A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs*, Front. Comput. Neurosci., (2009).

[11] K. Giesecke, K. Spiliopoulos, and R. B. Sowers, *Default clustering in large portfolios: typical events*, Ann. Appl. Probab., 23 (2013), pp. 348–385.

[12] E. Luçon and W. Stannat, *Mean field limit for disordered diffusions with singular interactions*, Ann. Appl. Probab., 24 (2014), pp. 1946–1993.

[13] S. Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, in Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), vol. 1627 of Lecture Notes in Math., Springer, Berlin, 1996, pp. 42–95.
[14] S. Ostojic, N. Brunel, and V. Hakim, *Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities*, J. Comput. Neurosci., 26 (2009), pp. 369–392.
[15] A. V. Skorohod, *Limit theorems for stochastic processes*, Teor. Veroyatnost. i Primenen., 1 (1956), pp. 289–319.
[16] K. Spiliopoulos, J. A. Sirignano, and K. Giesecke, *Fluctuation analysis for the loss from default*, Stochastic Process. Appl., 124 (2014), pp. 2322–2362.
[17] A.-S. Sznitman, *Topics in propagation of chaos*, in École d'Été de Probabilités de Saint-Flour XIX—1989, vol. 1464 of Lecture Notes in Math., Springer, Berlin, 1991, pp. 165–251.
[18] W. Whitt, *The reflection map with discontinuities*, Math. Oper. Res., 26 (2001), pp. 447–484.
[19] ———, *Stochastic-process limits*, Springer Series in Operations Research, Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.

François Delarue, Sylvain Rubenthaler, Laboratoire J.-A. Dieudonné, Université de Nice Sophia-Antipolis, Parc Valrose, 06108 Nice Cedex 02, France.
E-mail address: delarue@unice.fr,rubentha@unice.fr

James Inglis, Etienne Tanré, Equipe Tosca, INRIA Sophia-Antipolis Méditerranée, 2004 route des lucioles, BP 93, 06902 Sophia Antipolis Cedex, France
E-mail address: James.Inglis@inria.fr,Etienne.Tanre@inria.fr