Quenched asymptotics for interacting diffusions on inhomogeneous random graphs

Eric Luçon

Quenched asymptotics for interacting diffusions on inhomogeneous random graphs

Eric Luçon

MAP5 (UMR CNRS 8145), Université de Paris, F-75006 Paris, France,
eric.lucon@paridescartes.fr.

Abstract

The aim of the paper is to address the behavior in large population of diffusions interacting on a random, possibly diluted and inhomogeneous graph. This is the natural continuation of a previous work, where the homogeneous Erdős-Rényi case was considered. The class of graphs we consider includes disordered W-random graphs, with possibly unbounded graphons. The main result concerns a quenched convergence (that is true for almost every realization of the random graph) of the empirical measure of the system towards the solution of a nonlinear Fokker-Planck PDE with spatial extension, also appearing in different contexts, especially in neuroscience. The convergence of the spatial profile associated to the diffusions is also considered, and one proves that the limit is described in terms of a nonlinear integro-differential equation which matches the neural field equation in certain particular cases.

Keywords: mean-field system, interacting diffusions, nonlinear Fokker-Planck equation, spatially-extended systems, nonlinear heat equation, neural field equation, random graphs, graph convergence
2010 MSC: 60F15, 82C20, 35K55, 35Q84, 35Q92, 92B20

1. The model

1.1. Interacting diffusions on a graph

For all $n \geqslant 1$, consider the system of coupled diffusions $(\theta_{1,t}^{(n)}, \ldots, \theta_{n,t}^{(n)})$ in $\mathbb{R}^d$ $(d \geqslant 1)$

$$\mathrm{d}\theta_{i,t}^{(n)} = c(\theta_{i,t}^{(n)})\mathrm{d}t + \frac{\kappa_i^{(n)}}{n} \sum_{j=1}^{n} \xi_{i,j}^{(n)} \Gamma\left(\theta_{i,t}^{(n)}, \theta_{j,t}^{(n)}\right) \mathrm{d}t + \sigma \mathrm{d}B_{i,t},\ 0 \leqslant t \leqslant T,\ i = 1, \ldots, n\ . \quad (1.1)$$

The dynamics in (1.1) is decomposed into three terms: a local dynamics, represented by $c(\cdot) : \mathbb{R}^d \to \mathbb{R}^d$, a mean-field coupling (governed by the binary kernel $\Gamma(\cdot,\cdot) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$) and a noise term, in the presence of i.i.d. standard Brownian motions in $\mathbb{R}^d$, $B_1, \ldots, B_n$. Here σ is a constant, but possibly degenerate (even equally 0) diffusion matrix. The time horizon T is fixed (but arbitrary).

In (1.1), the diffusions $(\theta_1^{(n)}, \ldots, \theta_n^{(n)})$ no longer interact on the complete graph (as it is a common framework for a mean-field analysis) but through a nontrivial graph of interaction, encoded by the matrix $(\xi_{i,j}^{(n)})_{i,j=1,\ldots,n}$ in $\{0,1\}^{n^2}$. More precisely, we define the graph of interaction of (1.1) as $\mathcal{G}^{(n)} := (\mathcal{V}^{(n)}, \mathcal{E}^{(n)})$ with set of vertices $\mathcal{V}^{(n)} = [n] := \{1, \ldots, n\}$ and set of oriented edges $\mathcal{E}^{(n)} = \{(i,j) \in \mathcal{V}^{(n)} \times \mathcal{V}^{(n)},\ \xi_{i,j}^{(n)} = 1\}$. The aim of the paper is to analyse the large population behavior of (1.1) for situations where the graph of interaction $\mathcal{G}^{(n)}$ is possibly inhomogeneous. This paper is the natural continuation of [22] where the homogeneous Erdős-Rényi case is considered.

Remark 1.1. *Note that it would also be possible to include disordered coefficients $c(\theta_i, \omega_i)$ and $\Gamma(\theta_i, \omega_i, \theta_j, \omega_j)$ in (1.1), where $(\omega_i)_{i\in[n]}$ is some i.i.d sequence independent of everything, as it is customary for Kuramoto-type models (see Section 2.8 below). Everything below works with this additional random environment up to an additional expectation w.r.t. this disorder, under appropriate moment conditions.*

1.2. Construction of the interaction graph

The construction of the graph $\mathcal{G}^{(n)}$ goes back to the formalism of W-random graphs developed in [36, 9, 10, 8, 7], which has been used in particular in a series of papers [42, 32, 33, 16, 43] on macroscopic limits for Kuramoto-type models (see Section 2.8), in the deterministic case $\sigma = 0$. In addition to the fact that we consider here more general dynamics, the crucial point is the presence of noise in (1.1) that changes considerably the analysis (in particular, the techniques used in [33, 16] for the convergence of the empirical measure when $\sigma = 0$ do not seem to be directly applicable to the case $\sigma \neq 0$). The work that is closest to the present analysis is the recent work [48] where annealed large deviations estimates are given in the case of bounded graphons.

Let I be a closed subset of $\mathbb{R}^p$ ($p \geqslant 1$), endowed with a probability measure ℓ with support I. We associate to each vertex $i \in [n]$ a position variable $x_i^{(n)} \in I$ encoding some local inhomogeneity for the vertex i in the graph $\mathcal{G}^{(n)}$. In many situations (see e.g. [36, 9]), the set of positions I is taken to be $[0, 1]$, but most of the results presented below remain valid in more general cases, closer to situations where $x_i^{(n)}$ actually encodes some real spatial position of the particle $\theta_i^{(n)}$. Spatial extensions of mean-field dynamics are particularly relevant in a context of neuroscience where one accounts for the spatial organization of neurons in the cortex (see [38, 46, 14, 15, 44] and references therein for further details). The way positions $(x_1^{(n)}, \ldots, x_n^{(n)})$ are chosen in I will be made precise later (see Assumptions 3.1 and 3.3 below). For now, we suppose that these positions are deterministic. In the rest of the paper, we denote by $\ell_n(\mathrm{d}x)$ the empirical measure of the positions:

$$\ell_n(\mathrm{d}x) := \frac{1}{n} \sum_{k=1}^{n} \delta_{x_k^{(n)}}(\mathrm{d}x). \tag{1.2}$$

Then, we introduce a kernel $W_n : I^2 \to [0, 1]$ such that $W_n\left(x_i^{(n)}, x_j^{(n)}\right) \in [0, 1]$ represents the probability of the presence of the edge $\xi_{i,j}^{(n)}$ in the graph $\mathcal{G}^{(n)}$:

Definition 1.2. *On a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we give ourselves a family of random variables $\left(\xi_{i,j}^{(n)}\right)_{i,j\in[n]; n \geqslant 1}$ on Ω, such that, under $\mathbb{P}$, for each $n \geqslant 1$, $(\xi_{i,j}^{(n)})_{i,j\in[n]}$ is a collection of independent Bernoulli random variables with parameter $W_n\left(x_i^{(n)}, x_j^{(n)}\right)$.*

In (1.1), the parameter $\kappa_i^{(n)} > 0$ is a dilution parameter that compensates for the possible local sparsity of the graph $\mathcal{G}^{(n)}$ around vertex i: vertices with fewer neighbors will have a larger dilution parameter. Note that each $\kappa_i^{(n)}$ may actually depend on the whole sequence of positions in the graph $\mathcal{G}^{(n)}$: $\kappa_i^{(n)} = \kappa_i^{(n)}(x_1^{(n)}, \ldots, x_n^{(n)})$. In the following $\Xi := \left(\xi_{i,j}^{(n)}\right)_{i,j\in[n];n\geqslant 1}$ and $\mathcal{X} := \left(x_i^{(n)}\right)_{i\in[n];n\geqslant 1}$ stand for the whole sequence of connections and positions. For fixed $n \geqslant 1$, we also write $\underline{x} := \left(x_i^{(n)}\right)_{i\in[n]}$ and $\underline{\xi} := \left(\xi_{i,j}^{(n)}\right)_{i,j\in[n]}$. In absence of ambiguity, we write x_i instead of $x_i^{(n)}$, $\xi_{i,j}$ instead of $\xi_{i,j}^{(n)}$ and θ_i instead of $\theta_i^{(n)}$. The notations $\mathbf{P}(\cdot)$ and $\mathbf{E}[\cdot]$ stand for the probability and expectation w.r.t. the randomness in the Brownian motions and initial condition in (1.1). We use both notations $x \cdot y$ or $\langle x\, ,\, y\rangle$ for the scalar product of $x, y \in \mathbb{R}^d$, and $|x|$ denotes the Euclidean norm of x. $\sigma^\dagger$ stands for the transpose of the matrix σ. The notation $\langle \mu\, ,\, f\rangle := \int f \mathrm{d}\mu$ is also used for the usual duality between a measure and some test function.

1.3. The macroscopic kernel

In order to obtain a macroscopic limit as $n \to \infty$ for (1.1), we require some averaging for the probability field $W_n(\cdot, \cdot)$: we assume the existence of a nonnegative measurable function $W : I^2 \to [0, +\infty)$ so that the probability field $(W_n)_{n\geqslant 1}$, correctly renormalized by the dilution parameters $\kappa_i^{(n)}$, converges along the sequence $\mathcal{X}$ as $n \to \infty$ to the macroscopic kernel W (anticipating on the definitions of the Section 2.3 below, what we rigorously mean is that $\delta_n(\underline{x})$ in (2.19) goes to 0 as $n \to \infty$). This assumption encodes some notion of graph convergence (in the sense of [36, 8]) that is discussed in Section 2.7 below.

Remark 1.3. *Without loss of generality, we suppose that one particle does not interact with itself, that is $\xi_{i,i}^{(n)} = 0$ for all $i \in [n]$. In the limit as $n \to \infty$, this boils down to the assumption that the macroscopic kernel W is zero on the diagonal. We make these assumptions throughout this work without further notice.*

1.4. The McKean-Vlasov process and the nonlinear Fokker-Planck equation

The natural limit of the particle system (1.1) is then described by the nonlinear process $\bar{\theta}^x$ (at position $x \in I$) solution to

$$\mathrm{d}\bar{\theta}_t^x = c(\bar{\theta}_t^x)\mathrm{d}t + \int W(x,y)\Gamma\left(\bar{\theta}_t^x, \tilde{\theta}\right)\nu_t^y(\mathrm{d}\tilde{\theta})\ell(\mathrm{d}y)\mathrm{d}t + \sigma\mathrm{d}B_t,\ 0 \leqslant t \leqslant T. \tag{1.3}$$

where, for fixed (x,t), $\nu_t^x(\mathrm{d}\theta)$ is the law of $\bar{\theta}_t^x$. It is standard to see that the joint law

$$\nu(\mathrm{d}\theta, \mathrm{d}x) = \nu^x(\mathrm{d}\theta)\ell(\mathrm{d}x)$$

of $(\bar{\theta}^x, x)$ solves the nonlinear Fokker-Planck equation

$$\begin{aligned}\langle \nu_t\, ,\, \varphi\rangle = \langle \nu_0\, ,\, \varphi\rangle &+ \int_0^t \left\langle \nu_s\, ,\, \frac{1}{2}\nabla_\theta\left(\sigma\sigma^\dagger\nabla_\theta\varphi\right) + \nabla_\theta\varphi(\cdot)\cdot c(\cdot)\right\rangle \mathrm{d}s \\ &+ \int_0^t \left\langle \nu_s(\mathrm{d}\theta, \mathrm{d}x)\, ,\, \nabla_\theta\varphi(\theta, x) \cdot \int W(x,y)\Gamma(\theta,\tilde{\theta})\nu_s(\mathrm{d}\tilde{\theta}, \mathrm{d}y)\right\rangle \mathrm{d}s, \end{aligned} \tag{1.4}$$

where φ is a regular test function. Writing formally $\nu_t(\mathrm{d}\theta, \mathrm{d}x) = q_t(\theta, x)\mathrm{d}\theta\ell(\mathrm{d}x)$, (1.4) is the weak formulation of

$$\partial_t q_t = \frac{1}{2}\nabla_\theta \left(\sigma\sigma^\dagger \nabla_\theta q_t\right) - \nabla_\theta \left(q_t \left(c(\cdot) + \int \Gamma(\cdot, \theta') W(\cdot, y) q_t(\theta', y) \mathrm{d}\theta' \ell(\mathrm{d}y)\right)\right). \tag{1.5}$$

The precise meaning we give to (1.3) and (1.4) is given in Section 2.2 below. (1.3) and (1.5) are spatially-extended versions of standard McKean-Vlasov models that are natural large population limits of mean-field particle systems such as (1.1). A recent interest in models with spatial extension similar to (1.5) has been shown in a neuroscience context (see e.g. [38, 14, 48, 46, 44] and references therein).

2. Main assumptions and results

2.1. General assumptions

For any $r \geqslant 1$, for any $x \in I$ such that $W(x, \cdot) \in L^r(I, \ell)$, denote by

$$\mathcal{W}_r(x) := \int W(x, y)^r \ell(\mathrm{d}y). \tag{2.1}$$

Assumption 2.1 (Assumption on the kernel W)**.** *We require the minimal assumption that*

$$\left\|\mathcal{W}_2\right\|_\infty := \sup_{z\in I} \mathcal{W}_2(z) < \infty. \tag{2.2}$$

Suppose also that

$$\inf_{x\in I} \mathcal{W}_1(x) > 0. \tag{2.3}$$

Condition (2.2) of Assumption 2.1 implies in particular that $\left\|\mathcal{W}_1\right\|_\infty < \infty$: in the limit $n \to \infty$, the degree of each node $x \in I$ in the macroscopic graph W remains uniformly bounded ([23]).

Remark 2.2. *A closer look to the proofs below shows that* (2.3) *can be discarded if one assumes more integrability on W (for example,* (2.3) *is not needed if W is bounded).*

Remark 2.3. *An important remark is that we do not suppose any symmetry of the kernels W and W_n, nor that we suppose that W and W_n are simple functions of the distance $x - y$ (this is a natural hypothesis if one thinks of applications in neuroscience, as the mutual influence between neuron i on neuron j need not be symmetric). There are also some interesting examples where W is not symmetric, even if $\mathcal{G}^{(n)}$ might be (see Section 3). Note that the proof of Theorem 2.20 below requires to consider asymmetric kernels (see* (B.4)*).*

Assumption 2.4 (Assumptions on the coefficients Γ and c)**.** *We suppose that $(\theta, \tilde{\theta}) \mapsto \Gamma(\theta, \tilde{\theta})$ and $\theta \mapsto c(\theta)$ are twice differentiable on $\mathbb{R}^d$ with continuous derivatives and that Γ is Lipschitz continuous with sublinear behavior: there exists a constant $L_\Gamma > 0$ such that*

$$\left|\Gamma(\theta_1, \theta_2) - \Gamma(\tilde{\theta}_1, \tilde{\theta}_2)\right| \leqslant L_\Gamma \left(\left|\theta_1 - \tilde{\theta}_1\right| + \left|\theta_2 - \tilde{\theta}_2\right|\right), \ \theta_1, \tilde{\theta}_1, \theta_2, \tilde{\theta}_2 \in \mathbb{R}^d, \tag{2.4}$$

$$\left|\Gamma(\theta, \tilde{\theta})\right| \leqslant L_\Gamma \left(1 + |\theta| + \left|\tilde{\theta}\right|\right), \ \theta, \tilde{\theta} \in \mathbb{R}^d. \tag{2.5}$$

We require that $c(\cdot)$ is one-sided Lipschitz: there exists a constant $L_c > 0$ such that

$$\left\langle \theta - \tilde{\theta}\, , \, c(\theta) - c(\tilde{\theta}) \right\rangle \leqslant L_c \left|\theta - \tilde{\theta}\right|^2, \ \theta, \tilde{\theta} \in \mathbb{R}^d. \tag{2.6}$$

We also suppose some polynomial control on $c(\cdot)$: there exists $k \geqslant 2$ such that

$$\sup_{\theta \in \mathbb{R}^d} \frac{|c(\theta)|}{1 + |\theta|^k} < \infty. \tag{2.7}$$

Unless specified otherwise, we only assume (2.4), (2.5), (2.6) and (2.7). Nonetheless, for some of the results of the paper, we may restrict for simplicity to a generic subset of these assumptions:

Model 2.5 (Polynomial interactions)**.** *A particular case of the previous assumptions is to require that c is polynomial of degree smaller than k satisfying* (2.6) *and that Γ is either bounded or linear: $\Gamma(\theta, \tilde{\theta}) = \Gamma \cdot \left(\theta - \tilde{\theta}\right)$ for some (possibly degenerate) matrix Γ.*

We give in Section 2.8 below several dynamics satisfying the present hypotheses, a significant application to have in mind being FitzHugh-Nagumo oscillators with electrical synapses (see Section 2.8, item 2).

We now turn to the assumptions related to the initial condition in (1.4): if (S, d) is a Polish space, let $w_1(\cdot, \cdot)$ be the usual Wasserstein distance [53] on S: for any probability measures ν_1, ν_2 on S,

$$w_1(\nu_1, \nu_2) := \inf_{\pi} \left\{ \int d\left(\theta_1, \theta_2\right) \pi(\mathrm{d}\theta_1, \mathrm{d}\theta_2) \right\}, \tag{2.8}$$

where the infimum is taken on all couplings π on $S \times S$ with marginals ν_1 and ν_2. In (2.11) below, we take $S = \mathbb{R}^d$, but we will also use the same definition later for $S = \mathbb{R}^d \times I$ when necessary.

Assumption 2.6 (Assumption on the initial condition)**.** *We assume that* (1.4) *is endowed with an initial condition $\nu_0(\mathrm{d}\theta, \mathrm{d}x)$ of the form*

$$\nu_0(\mathrm{d}\theta, \mathrm{d}x) = \nu_0^x(\mathrm{d}\theta) \ell(\mathrm{d}x) \tag{2.9}$$

with some uniform a priori control on its moments: for k given by (2.7)*, suppose that*

$$\sup_{x \in I} \int |\theta|^{2k} \nu_0^x(\mathrm{d}\theta) < +\infty. \tag{2.10}$$

We assume here that there exist $L_0 > 0$ and $\iota_1 \in (0,1]$ such that

$$w_1\left(\nu_0^x, \nu_0^y\right) \leqslant L_0 \left|x-y\right|^{\iota_1}, \ x, y \in I. \tag{2.11}$$

We suppose finally that the initial condition of the particle system (1.1) *is such that $(\theta_{1,0}^{(n)}, \ldots, \theta_{n,0}^{(n)})$ are independent, with respective law $\theta_{k,0}^{(n)} \sim \nu_0^{x_k}(\mathrm{d}\theta)$, for $k \in [n]$. Note here that we allow a priori the initial law of the particles to depend on their positions $x_k = x_k^{(n)}$.*

2.2. Well-posedness on the nonlinear Fokker-Planck equation and a priori estimates

Consider $\mathcal{M}$ the set of probability measures ν on $\mathcal{C}([0,T], \mathbb{R}^d) \times I$ with marginals on I equal to ℓ. Since $\mathcal{C}([0,T], \mathbb{R}^d)$ is Polish, it follows from the disintegration Theorem [25] that any $\nu \in \mathcal{M}$ may be written as $\nu(\mathrm{d}\theta, \mathrm{d}x) = \nu^x(\mathrm{d}\theta)\ell(\mathrm{d}x)$. We endow $\mathcal{M}$ with the following Wasserstein-type metric [51]

$$\delta_T(\nu, \mu) := \sup_{x \in I} \inf_{\pi} \left\{ \sup_{s \leqslant T} \int \left|\vartheta_{1,s}^x - \vartheta_{2,s}^x\right|^{2k} \pi(\mathrm{d}\vartheta_1, \mathrm{d}\vartheta_2) \right\}^{\frac{1}{2k}}, \tag{2.12}$$

where the infimum is taken over all couplings π under which $\vartheta_1^x \sim \nu^x$ and $\vartheta_2^x \sim \mu^x$, for ℓ-almost every $x \in I$.

Proposition 2.7. *Under Assumptions 2.1, 2.4 and 2.6, there exists a unique weak solution to* (1.4) *in $\mathcal{M}$ with initial condition $\nu_0(\mathrm{d}\theta, \mathrm{d}x) = \nu_0^x(\mathrm{d}\theta)\ell(\mathrm{d}x)$. This solution ν is such that for ℓ-almost every $x \in I$, $\nu^x(\mathrm{d}\theta)$ is the law of the nonlinear process $(\bar{\theta}_t^x)_{t\in[0,T]}$ given by* (1.3)*.*

Remark 2.8. *A byproduct of Proposition 2.7 is the following: under the hypotheses of Proposition 2.7, there exists a constant C_0, only depending on Γ, c, W, σ, T and ν_0, such that*

$$\sup_{x \in I} \mathbf{E}\left[\sup_{s \leqslant T} \left|\bar{\theta}_s^x\right|^{2k}\right] \leqslant C_0. \tag{2.13}$$

Similarly, it is standard to prove, under the same hypotheses, a similar estimate for the particle system (1.1)*:*

$$\sup_{i \in [n]} \mathbf{E}\left[\sup_{s \leqslant T} \left|\theta_{i,s}^{(n)}\right|^{2k}\right] \leqslant C_0. \tag{2.14}$$

The proof of Proposition 2.7 is standard and relies on a fixed-point argument [51] on the McKean-Vlasov diffusion (1.3). Existence and uniqueness in (1.3) provides existence of a solution to (1.4). Uniqueness in (1.4) comes from a propagator method. Similar well-posedness results for spatially-extended McKean-Vlasov processes may be found in [38, 46, 48]. Proposition 2.7, as well as some further regularity estimates concerning ν, is proven in Appendix A.

2.3. A general propagation of chaos estimate

In this paragraph, we fix a sequence of positions $(x_i)_{i\in[n]}$ (that is supposed to be deterministic in Theorem 2.13 below), a probability field $W_n(x_i,x_j)$ and a kernel W. We are interested in the approximation of the microscopic system (1.1) by its mean-field limit (1.3): let $(\bar{\theta}_1^{x_1},\dots,\bar{\theta}_n^{x_n})$, n independent copies of the nonlinear process driven by the same Brownian motions $(B_1,\dots,B_n)$, with the same positions x_i and initial conditions as in (1.1). For simplicity, we write $\bar{\theta}_{i,s}$ in place of $\bar{\theta}_{i,s}^{x_i}$.

In order to state the result, we need some hypotheses on the graph $\mathcal{G}^{(n)}$:

Assumption 2.9 (Convergence of $(\mathcal{G}^{(n)},\kappa^{(n)})$ to W)**.** *We assume the following:*

1. *Uniform control on $\kappa^{(n)}$: we suppose some uniformity in the dilution parameters $(\kappa_i^{(n)})_{i\in[n]}$, namely the existence of $\kappa_n\geqslant 1$ and $w_n\in(0,1]$ such that*

$$\kappa_\infty^{(n)}(\underline{x}) := \max_{i\in[n]}\left(\kappa_i^{(n)}(\underline{x})\right)\leqslant \kappa_n, \tag{2.15}$$

$$\max_{i,j\in[n]}\left(W_n(x_i,x_j)\right)\leqslant w_n, \tag{2.16}$$

satisfying, as $n\to\infty$,

$$\frac{1}{\kappa_n}\leqslant w_n\leqslant 1, \tag{2.17}$$

$$\kappa_n^2 w_n = o\left(\frac{n}{\log(n)}\right),\ \text{ as } n\to\infty. \tag{2.18}$$

2. *Convergence of the weighted graph $(\mathcal{G}^{(n)},\kappa^{(n)})$: defining*

$$\delta_n(\underline{x}) := \sup_{i\in[n]}\left(\frac{1}{n}\sum_{k=1}^n\left|\kappa_i^{(n)}W_n(x_i,x_k)-W(x_i,x_k)\right|\right), \tag{2.19}$$

we assume that

$$\delta_n(\underline{x})\xrightarrow[n\to\infty]{}0. \tag{2.20}$$

Remark 2.10. *A few comments are in order here:*

1. *Assumptions (2.16) and (2.17) are mostly technical, since it is always possible to take $w_n=1$. Nonetheless, there are simple cases where it is natural to take $w_n\to 0$ as $n\to\infty$: e.g. consider $W_n(x_i,x_j)\equiv w_n=\rho_n$ for some $\rho_n\xrightarrow[n\to\infty]{}0$ (this corresponds to a uniform diluted Erdős-Rényi graph $\mathcal{G}^{(n)}$, see [22]). In this case, it is natural to renormalize the sum in (1.1) by the mean degree of each vertex (equal to $n\rho_n$), so that we take $\kappa_i^{(n)}\equiv\kappa_n=\frac{1}{\rho_n}$ for all $i\in[n]$. Then, (2.18) boils down to the condition $\kappa_n=o\left(\frac{n}{\log(n)}\right)$, which is exactly the condition found in [22], Eq. (1.12) in the Erdős-Rényi case. A general extension of this simple case is considered in Section 3.3.*

2. *The convergence* (2.20) *of the microscopic probability field* W_n *(properly renormalized by* $\kappa^{(n)}$*) to the macroscopic kernel* W *encodes some notion of convergence of the underlying graph* $\mathcal{G}^{(n)}$ *as* $n \to \infty$*. Further comments on this point are made in Section 2.7.*
   In the remaining of the paper, we will say that $(\mathcal{G}^{(n)}, \kappa^{(n)})$ *converges to* W *as* $n \to \infty$ *if Assumption 2.9 holds. General examples of converging graphs are given in Section 3.*

The second set of assumptions concerns regularity estimates on the limiting kernel W. The following notations are used throughout the paper:

$$[\Gamma]_u(\theta, x) := \int \Gamma(\theta, \tilde{\theta}) \nu_u^x(\mathrm{d}\tilde{\theta}), \ \theta \in \mathbb{R}^d, x \in I, u \geqslant 0. \tag{2.21}$$

$$\Upsilon_t(x, y, z) := \int_0^t \int \langle [\Gamma]_u(\theta, y)\, , \, [\Gamma]_u(\theta, z) \rangle \, \nu_u^x(\mathrm{d}\theta)\mathrm{d}u, \ x, y, z \in I, t \geqslant 0. \tag{2.22}$$

A priori controls on $[\Gamma]$ and Υ are given in Lemma A4 below. Define (recall the definition of $\ell_n(\mathrm{d}x)$ in (1.2)), for $i \in [n]$

$$\begin{aligned} \epsilon_{n,T}^{(1,i)}(\underline{x}) &:= \int W(x_i, y)W(x_i, z)\Upsilon_T(x_i, y, z) \left\{ \ell_n(\mathrm{d}y)\ell_n(\mathrm{d}z) - \ell(\mathrm{d}y)\ell(\mathrm{d}z) \right\}, \\ \epsilon_{n,T}^{(2,i)}(\underline{x}) &:= \int W(x_i, y)W(x_i, z)\Upsilon_T(x_i, y, z) \left\{ \ell_n(\mathrm{d}y) - \ell(\mathrm{d}y) \right\} \ell(\mathrm{d}z), \\ \epsilon_{n,T}^{(3,i)}(\underline{x}) &:= \int W(x_i, y)W(x_i, z)\Upsilon_T(x_i, y, z)\ell(\mathrm{d}y) \left\{ \ell_n(\mathrm{d}z) - \ell(\mathrm{d}z) \right\}. \end{aligned} \tag{2.23}$$

Assumption 2.11 (Regularity of W along $\mathcal{X}$)**.** *We require that*

$$\epsilon_{n,T}^{(m)}(\underline{x}) := \sup_{i \in [n]} \left| \epsilon_{n,T}^{(m,i)}(\underline{x}) \right| \xrightarrow[n \to \infty]{} 0, \ \textit{for } m = 1, 2, 3, \tag{2.24}$$

as well as

$$\sup_{n \geqslant 1} \sup_{i \in [n]} \int W(x_i, y)\ell_n(\mathrm{d}y) = \sup_{n \geqslant 1} \sup_{i \in [n]} \frac{1}{n} \sum_{k=1}^n W(x_i, x_k) < +\infty. \tag{2.25}$$

Remark 2.12. *Assumption 2.11 captures a notion of regularity of the macroscopic kernel* W*:* (2.25) *is the discrete counterpart of* (2.2) *(when* $r = 1$*) and (forgetting about the factor* Υ_T *in* (2.23)*)* (2.24) *essentially says that various empirical means in* W *converge to their expectation. Hence, we will say that the kernel* W *is regular along the sequence of positions* $\mathcal{X}$ *if Assumption 2.11 is satisfied. We give in Section 3.1 some sufficient conditions for Assumption 2.11 to hold.*

We are now in position to state the first main result of the paper:

Theorem 2.13. *Fix $T \geqslant 0$. Suppose that Assumptions 2.1, 2.4, 2.6, 2.9 and 2.11 hold. In this case,*

$$\sup_{i \in [n]} \mathbf{E}\left[\sup_{s \leqslant T} \left|\theta_{i,s}^{(n)} - \bar{\theta}_{i,s}\right|^2\right] \to 0 \text{ as } n \to \infty, \tag{2.26}$$

for almost every realization of the connectivity sequence Ξ given by Definition 1.2.

Theorem 2.13 is proven in Section 4.1. We detail in Section 3 generic examples of graphs that are regular and convergent in the sense of Assumptions 2.9 and 2.11. The rest of the present section is organized as follows: a byproduct of Theorem 2.13 concerns the convergence of the empirical measure of (1.1) (Section 2.4). A different approach to the macroscopic description of (1.1) is given in Section 2.5 and links between the two approaches are provided in Section 2.6. A discussion on the notion of graph convergence encoded by assumption (2.20) is given in Section 2.7. Comments on applications and links with existing literature are given in Section 2.8.

2.4. Convergence of the empirical measure

It is standard (see e.g. [22] for a similar result in the Erdős-Rényi case) to derive from Theorem 2.13 the convergence of the empirical measure of the system (1.1)

$$\nu_{n,t}(\mathrm{d}\theta, \mathrm{d}x) = \frac{1}{n} \sum_{i=1}^{n} \delta_{(\theta_{i,t}^{(n)}, x_i^{(n)})}(\mathrm{d}\theta, \mathrm{d}x),\ t \geqslant 0 \tag{2.27}$$

to the solution ν to the nonlinear Fokker-Planck equation (1.4).

Assumption 2.14 (Further regularity on W and convergence of ℓ_n)**.** *We assume the following:*

1. *there exist $L_W > 0$ and $\iota_2 \in (0, 1]$ such that*

$$\delta\mathcal{W}(x, y) := \int |W(x, z) - W(y, z)|\, \mathrm{d}z \leqslant L_W\, |x - y|^{\iota_2},\ x, y \in I. \tag{2.28}$$

2. *Convergence of ℓ_n (recall (1.2)): when I is not bounded, we suppose $\iota_1 = \iota_2$ (recall (2.11)) and denote in any case $\iota := \iota_1 \wedge \iota_2 \in (0, 1]$. Let $\mathcal{H}_\iota$ be the set of ι-Hölder functions on I*

$$\mathcal{H}_\iota := \left\{\varphi : I \to \mathbb{R}, \|\varphi\|_{\mathcal{H}_\iota} < \infty\right\} \tag{2.29}$$

where $\|\varphi\|_{\mathcal{H}_\iota} := \sup_{x \in I} |\varphi(x)| + \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{|x-y|^\iota}$. Denote by $d_{\mathcal{H}_\iota}(\cdot, \cdot)$ the distance given by, for every probability measures μ and ν on I,

$$d_{\mathcal{H}_\iota}(\mu, \nu) = \sup_{\varphi \in \mathcal{H}_\iota, \|\varphi\|_{\mathcal{H}_\iota} \leqslant 1} |\langle \mu - \nu\, ,\, \varphi\rangle|\, . \tag{2.30}$$

We suppose that the empirical measure of the positions ℓ_n satisfies

$$d_{\mathcal{H}_\iota}(\ell_n, \ell) \xrightarrow[n \to \infty]{} 0. \tag{2.31}$$

Under these assumptions, we state the convergence of the empirical measure (2.27) in the following simple way:

Theorem 2.15. *Suppose that Assumptions 2.1, 2.4, 2.6, 2.9, 2.11 and 2.14 hold. Then for any continuous function $\varphi : \mathbb{R}^d \times I \to \mathbb{R}$ such that for some $C_\varphi > 0$, for any $\theta, \tilde{\theta} \in \mathbb{R}^d$, $x, y \in I$, $\left|\varphi(\theta, x) - \varphi(\tilde{\theta}, y)\right| \leqslant C_\varphi \left(\left|\theta - \tilde{\theta}\right| + |x-y|^\iota\right)$ and $\sup_{x\in I} |\varphi(\theta, x)| \leqslant C_\varphi(1 + |\theta|)$, the following convergence holds:*

$$\sup_{t\in[0,T]} \mathbf{E}\left[|\langle \nu_{n,t} - \nu_t \, , \, \varphi\rangle|^2\right] \xrightarrow[n\to\infty]{} 0, \tag{2.32}$$

for almost every realization of the connectivity sequence Ξ given by Definition 1.2.

Theorem 2.15 is proven in Section 4.2.

2.5. The nonlinear spatial profile

In [42, 43, 32, 33], a different approach to the large population behavior of (1.1) is considered. The point of view here is to consider the deterministic macroscopic spatial profile $(\psi(\cdot, t))_{t\in[0,T]}$ that, in our context, solves the following nonlinear integro-differential equation

$$\partial_t \psi(x,t) = c(\psi(x,t)) + \int_I \Gamma(\psi(x,t), \psi(y,t)) W(x,y) \ell(\mathrm{d}y). \tag{2.33}$$

In the context of [43], (2.33) is referred to as the nonlinear heat equation on the graph W. For FitzHugh-Nagumo dynamics with linear interaction, (2.33) corresponds to the reaction-diffusion equation addressed in the recent work [20]. We consider here weak solutions to (2.33) in the sense of the following definition: if $\mathcal{C}([0,T], L^k(I,\ell))$ is the set of continuous functions with values in $L^k(I,\ell)$ where $k \geqslant 2$ is given in (2.10),

Definition 2.16. *We say that $\psi(\cdot,t)_{t\in[0,T]} \in \mathcal{C}([0,T], L^k(I,\ell))$ is a weak solution to (2.33) if for all regular test functions $J : I \to \mathbb{R}^d$, for all $t \in [0,T]$, we have*

$$\begin{aligned}\int_I \langle \psi(x,t)\, , \, J(x)\rangle \, \ell(\mathrm{d}x) = \int_I \langle \psi(x,0)\, , \, J(x)\rangle \, \ell(\mathrm{d}x) + \int_0^t \int_I \langle c(\psi(x,s))\, , \, J(x)\rangle \, \ell(\mathrm{d}x)\mathrm{d}s \\ + \int_0^t \int_{I^2} \langle \Gamma(\psi(x,s), \psi(y,s))\, , \, J(x)\rangle \, W(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x)\mathrm{d}s.\end{aligned} \tag{2.34}$$

We first state a uniqueness result for (2.34):

Proposition 2.17. *Under the Assumptions 2.1, 2.4 and 2.6, for any $\psi_0 \in L^k(I,\ell)$, there is at most one weak solution in $\mathcal{C}([0,T], L^k(I,\ell))$ to (2.34) with initial condition ψ_0.*

For the rest of Section 2.5, we restrict ourselves to the case $I := [0,1]$, endowed with its Lebesgue measure $\ell(\mathrm{d}x) := \mathrm{d}x$ and where $x_i^{(n)} := \frac{i}{n}$ for $i \in [n]$ (see Section 3.1 for further details). Following the approach of [43], it is possible to consider the spatial field:

$$\theta_n(x,t) := \theta^{(n)}_{\lfloor nx+1 \rfloor, t} = \sum_{i=1}^n \theta^{(n)}_{i,t} \ \mathbf{1}_{[x^{(n)}_{i-1}, x^{(n)}_i)}(x), \ x \in I, \ t \geqslant 0. \tag{2.35}$$

We restrict here for simplicity to Model 2.5. We suppose that the hypotheses of Section 2.1 hold and that $(\mathcal{G}^{(n)}, \kappa^{(n)})$ converges to W in the sense of Assumption 2.9. We require the regularity of $(W, \mathcal{X})$ in the sense of Assumption 2.11 together with the following supplementary condition: suppose that

$$\sum_{i,j=1}^{n} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \int_{\frac{j-1}{n}}^{\frac{j}{n}} \left| W\left(\frac{i}{n}, \frac{j}{n}\right) - W(x,y) \right|^2 \mathrm{d}x\mathrm{d}y \to 0, \text{as } n \to \infty. \tag{2.36}$$

The convergence is the following:

Theorem 2.18. *Restrict to Model 2.5 and suppose that the Assumptions 2.1, 2.6, 2.9, 2.11 and condition (2.36) hold. Then, for almost every realization of the connectivity sequence Ξ given in Definition 1.2, the spatial field (θ_n) given in (2.35) converges weakly in $\mathcal{C}([0,T], L^k)$ to $\psi(\cdot,t)_{t\in[0,T]}$, unique solution in $\mathcal{C}([0,T], L^k)$ to (2.34) with initial condition*

$$\psi_0(x) := \int_{\mathbb{R}^d} \theta \nu_0^x(\mathrm{d}\theta). \tag{2.37}$$

The present result can be seen as a generalization of [42, 43, 32, 33], where the case of Kuramoto-type interaction (namely Γ of the form $\Gamma(\theta - \tilde{\theta})$ with $\Gamma(\cdot)$ and $c(\cdot)$ Lipschitz and bounded) in absence of noise ($\sigma = 0$) is considered. Theorem 2.18 is proven in Section 5.2.

2.6. Identification

A direct consequence of Theorems 2.15 and 2.18 is the identification between the spatial profile $\psi(\cdot,t)$, weak solution to (2.34) in terms of the expected value of the solution ν_t of (1.4) (see (2.38) and (2.39) below). This identification is straightforward in the case $I = [0,1]$, $\ell(\mathrm{d}x) = \mathrm{d}x$, as it based on both convergence of processes $(\nu_{n,t})_{n\geqslant 1}$ and $(\theta_n(\cdot,t))_{n\geqslant 1}$:

Theorem 2.19 (Identification in the compact case)**.** *Suppose $I = [0,1]$, $\ell(\mathrm{d}x) = \mathrm{d}x$ and restrict to Model 2.5. Under the Assumptions 2.1, 2.6, 2.9, 2.11 and 2.14 and condition (2.36), let $(\nu_t)_{t\in[0,T]}$ be the unique solution to (1.4) with initial condition ν_0 and $\psi(\cdot,t)_{t\in[0,T]}$ be the unique solution to (2.34) with initial condition $\psi_0(x) := \int \theta\nu_0^x(\mathrm{d}\theta)$. Then, for all $t \in [0,T]$, all regular test functions J on $[0,1]$*

$$\int_{[0,1]} \langle \psi(x,t)\, ,\, J(x) \rangle \,\mathrm{d}x = \int_{[0,1]} \left\langle \int_{\mathbb{R}^d} \theta\nu_t^x(\mathrm{d}\theta)\, ,\, J(x) \right\rangle \mathrm{d}x. \tag{2.38}$$

Theorem 2.19 is proven in Section 5.2. When $I \subset \mathbb{R}^p$ is not compact, there is no natural construction of the spatial profile $(\theta_n(\cdot,t))$ as in (2.35). However, by a simple truncation argument, it is still possible to get the following identification result (which may have an interest of its own, independently of the context of random graphs):

Theorem 2.20 (Identification when $I = \mathbb{R}^p$)**.** *Suppose that $I = \mathbb{R}^p$ is endowed with a probability measure $\ell(\mathrm{d}x) = \ell(x)\mathrm{d}x$ that is absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}^p$. Suppose that ℓ is $\mathcal{C}^1$ on $\mathbb{R}^p$ and fix a kernel $W(x,y)$ that is $\mathcal{C}^1$ on $\mathbb{R}^p \times \mathbb{R}^p$. Restrict to Model 2.5 and suppose that Assumptions 2.1 and 2.6 hold.*

Then, (2.34) has a unique weak solution $\psi(\cdot,t)_{t\in[0,T]}$ with initial condition $\psi_0(x) := \int \theta\nu_0^x(\mathrm{d}\theta)$. If $(\nu_t)_{t\in[0,T]}$ is the unique solution to (1.4) with initial condition ν_0, then, for all $t \in [0,T]$, all test regular functions J with compact support on $\mathbb{R}^p$,

$$\int_{\mathbb{R}^p} \langle \psi(x,t)\, ,\, J(x) \rangle \,\ell(\mathrm{d}x) = \int_{\mathbb{R}^p} \left\langle \int_{\mathbb{R}^d} \theta\nu_t^x(\mathrm{d}\theta)\, ,\, J(x) \right\rangle \ell(\mathrm{d}x). \tag{2.39}$$

Theorem 2.20 is proven in Appendix B, page 52. In the case of FitzHugh-Nagumo oscillators, (2.39) (i.e. the identification of the expected value of (1.5) as a solution of (2.33)) can be seen as a weak formulation of a recent work [20] where a similar issue is addressed, using PDE techniques. Although we consider here a more general class of model, the present identification is weaker, as it is only valid in the sense of distributions (in particular, the spatial regularity of $(\psi(\cdot, t))$ is not addressed here). Another significant difference with [20] is that we crucially need here to have a probability measure $\ell(\mathrm{d}x)$ on the spatial variable x, whereas [20] adresses directly the case where ℓ is Lebesgue on $\mathbb{R}^p$.

2.7. A comment on graph convergence

The aim of this paragraph is to question the notion of graph convergence given by Assumption 2.9. The point we want to raise here is that this notion of convergence does not really concern so much the unlabeled and (possibly undirected) graph $\mathcal{G}^{(n)}$ constructed in Definition 1.2, but is rather a notion of convergence of a directed and weighted graph $\bar{\mathcal{G}}^{(n)}$ that is coupled to $\mathcal{G}^{(n)}$, with weights that depend on the dilution sequence $\kappa^{(n)}$. The reason is that, even though the original graph $\mathcal{G}^{(n)}$ might be symmetric, the renormalization $\kappa_i^{(n)}$ for each node i in (1.1) induces an asymmetry in the interaction between i and any of its neighbor j. To be more specific:

Definition 2.21. *Let $\bar{\mathcal{G}}^{(n)}$ be the directed and weighted graph (with vertex set $[n]$) constructed from $\mathcal{G}^{(n)}$ in the following way: for any $i \neq j \in [n]$, both edges $i \to j$ and $j \to i$ are present in $\bar{\mathcal{G}}^{(n)}$ if and only if the undirected edge $\{i, j\}$ is present in $\mathcal{G}^{(n)}$. Then, attribute the weight $\kappa_i^{(n)}$ (resp. $\kappa_j^{(n)}$) to $i \to j$ (resp. $j \to i$) in $\bar{\mathcal{G}}^{(n)}$.*

We suppose here again for simplicity that $I := [0, 1]$, endowed with its Lebesgue measure $\ell(\mathrm{d}x) := \mathrm{d}x$ and where $x_i^{(n)} := \frac{i}{n}$ for $i \in [n]$. We assume in this paragraph that Assumption 2.9 holds as well as:

$$\sum_{i,j=1}^{n} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \int_{\frac{j-1}{n}}^{\frac{j}{n}} \left| W(x_i^{(n)}, x_j^{(n)}) - W(x, y) \right| \mathrm{d}x \mathrm{d}y \xrightarrow[n\to\infty]{} 0. \tag{2.40}$$

Proposition 2.22. *Let $\bar{\mathcal{G}}^{(n)}$ be given by Definition 2.21. Under the above hypotheses, we have the following convergence result:*

$$d_\square \left(\bar{\mathcal{G}}^{(n)}, W \right) \xrightarrow[n\to\infty]{} 0, \tag{2.41}$$

where $d_\square(\cdot, \cdot)$ is the cut-off distance.

We use here the formalism of graph convergence developed in [36, 9, 6, 7, 8] (and references therein). The precise definition of the cut-off distance together with the proof of Proposition 2.22 are given in Appendix C. Note that one needs to slightly generalize the formalism of [7, 8] to the case of directed graphs and asymmetric kernels W (but this is of minor difficulty).

An important point concerning Proposition 2.22 is that a lot of the structure of the microscopic graph $\mathcal{G}^{(n)}$ is lost in the limit $n \to \infty$: the macroscopic limit (1.5) essentially captures a dynamics that lives on the *renormalized* graph $\bar{\mathcal{G}}^{(n)}$, which may be significantly different to $\mathcal{G}^{(n)}$. More precisely, one of the main contributions of the theory developed in

[36, 9, 6, 7, 8]) is to show the existence of a large class of generic models of microscopic graphs $\mathcal{G}^{(n)}$ that converge to some graphon $\mathcal{P}(x,y)$. The point we want to stress is that the limit $\mathcal{P}$ of $\mathcal{G}^{(n)}$ is in general different from the limit W of $\bar{\mathcal{G}}^{(n)}$ provided by (2.41). This is due to the presence of the renormalizing coefficients $(\kappa_i^{(n)})_{i\in[n]}$ in (1.1) and typically true when $\mathcal{G}^{(n)}$ has vertices with diverging degree as $n\to\infty$, see e.g. Section 3, (Examples 3.7 and 3.11) and Remark 3.13, where we have two different $\mathcal{G}^{(n)}$, converging to different $\mathcal{P}$, such that their renormalized graphs $\bar{\mathcal{G}}^{(n)}$ converge to the same W.

What is more, even though the graph of interaction $\mathcal{G}^{(n)}$ might be of power-law type, the renormalized graph $\bar{\mathcal{G}}^{(n)}$ and its macroscopic counterpart W that we consider in this paper are *never* of power-law type: a crucial assumption that is constantly used in this work is (2.2), i.e. the degree of each macroscopic node remains uniformly of order 1 (note that this uniformity in degrees crucially depends on the choice of the dilution coefficients $(\kappa_i^{(n)})$). The situation where this uniform control on macroscopic degrees is discarded is unclear: to illustrate this, consider the graph $\mathcal{G}^{(n)}$ with diverging degrees defined in Example 3.11 where, instead of (3.17), we choose now

$$\kappa_i^{(n)} = \frac{1}{\rho_n},\ i\in[n]\,, \tag{2.42}$$

that is, the same uniform dilution as for bounded kernels (3.9) (example already considered in [32], § 6.2). The graph $(\mathcal{G}^{(n)},\kappa^{(n)})$ remains convergent in the sense of Assumption 2.9 to

$$W(x,y) := \mathcal{P}(x,y) = (1-\alpha)^2 x^{-\alpha} y^{-\alpha}. \tag{2.43}$$

Indeed, choosing for simplicity $0<\alpha<\frac{1}{6}$ and $\rho_n = n^{-\delta}$ with $2\alpha<\delta<\frac{1}{2}-\alpha$ in (3.6),

$$\delta_n(\underline{x}) = \sup_{i\in[n]} \frac{1}{n}\sum_{j=1}^{n} \left|\min\left(n^{\delta}, (1-\alpha)^2(x_ix_j)^{-\alpha}\right) - (1-\alpha)^2(x_ix_j)^{-\alpha}\right|,$$

is equally 0 for all n, since $\delta>2\alpha$. But now, the uniform renormalization (2.42) (well adapted to vertices with low degree, with position away from 0) is no longer sufficient to compensate for vertices with high degree with position close to 0 and the boundedness assumption (2.2) is no longer satisfied for (2.43): macroscopic nodes $x\in[0,1]$ have diverging degrees as $x\to 0$. At the level of generality considered in this work (but even for Kuramoto-type interaction), it is unclear if the convergence results (Theorem 2.13 or Theorem 2.18) remain true when assumption (2.2) is discarded.

2.8. Applications and links with the existing literature

Applications

The kind of applications we have in mind are:

1. The Kuramoto model and its variants: take $d=1$ and $\Gamma(\theta,\tilde\theta)=\sin\left(\tilde\theta-\theta\right)$. Examples of local dynamics are usually $c(\cdot)\equiv\omega$ or $c(\theta)=1+a\sin(\theta)$ [28]. In this context, (1.4) gives rise to a family of Kuramoto models with spatial extension already considered in the literature (P-nearest neighbor model [49], long-range interactions [30]). The question of characterizing the synchronized states, as for the original Kuramoto model, is still an ongoing question (see [1, 29] and references therein). In the case of

Kuramoto type ODEs (that is when $\sigma = 0$), a series of papers (see [42, 32, 33, 16, 43] and references therein) have addressed similar issues to the ones addressed here. In addition to the fact that we consider more general hypotheses (e.g. possibly unbounded and asymmetric Γ and c being non-Lipschitz), the main difficulty of the present analysis is that noise is present in (1.1). In particular, the fixed-point argument [47] used in [33, 16] for the convergence of the empirical measure in the deterministic case does not seem to generalize easily to the case $\sigma \neq 0$.

2. FitzHugh-Nagumo oscillators: this corresponds to $d = 2$, $\theta = (V, w)$, $c(V, w) := (V - \frac{V^3}{3} - w, \frac{1}{\tau}(V + a - bw))$ (for appropriate parameters a, b, τ) and $\Gamma(\theta, \tilde{\theta}) := (V - \tilde{V}, 0)$. Here, V stands for the potential of one neuron and w its recovery variable. We refer to [11] and references therein for more details on this particular model and its applications to neuroscience. Once again, (1.4) gives a spatially-extended version of a Fokker-Planck PDE, already analyzed in the context of neurons interacting through a deterministic spatial kernel ([52, 38, 39]). The present work gives a new interpretation of such spatially-extended PDEs in terms of the mean-field limit of diffusions on random graphs.

Longtime behavior

Theorems 2.13 and 2.15 are the natural extensions of [22] that concerns the case of homogeneous Erdős-Rényi graphs. One should also mention at this point the recent work [19] which addresses quenched propagation of chaos and large deviations results on homogeneous graphs. The result that is closest to this work is the recent [48] where a similar result of convergence is addressed in the case of bounded and Lipschitz coefficients c, Γ. The analysis in [48] restricts to bounded kernels W and random positions (that is a particular case of Section 3.3 below). Note also that the convergence of [48] is annealed in both disorder (connections and positions), whereas the present analysis is quenched. Contrary to [48], we do not address large deviation estimates here.

This work comes with all the comments and restrictions raised in [22]: the convergence results are only valid on bounded time intervals $[0, T]$, where T is independent of n (although that, with a little more work, it would certainly be possible to extend the result up to times T which grows logarithmically in n, as in [22], Corollary 1.2). In any case, the behavior of the empirical measure (2.27) on larger time scales is beyond the scope of this paper. Even for interactions on the complete graph, the longtime analysis of (1.1) relies heavily on the dynamical structure of the macroscopic limit (1.5) (e.g. existence of stable fixed-point or periodic manifolds, see [4, 37, 40] for results in this direction). The difficulty is here even more present for general graphs, since the interaction in (1.1) cannot be written as a closed expression of the empirical measure (2.27). One should mention the recent work [18] addressing longtime dynamics of Kuramoto oscillators on homogeneous Erdös-Rényi graphs. A trajectorial Central Limit Theorem associated to Theorem 2.15 remains open (in this direction, see [5] for an annealed fluctuation theorem in the Erdős-Rényi case).

One point raised in [22] concerned the necessity of the independence of the initial condition of (1.1) $(\theta_{1,0}, \dots, \theta_{n,0})$ with respect to the graph $\mathcal{G}^{(n)}$ (in [22], the $(\theta_{1,0}, \dots, \theta_{n,0})$ are identically distributed; see also [19] where a convergence result is proven in the Erdős-Rényi case assuming only the convergence of the empirical measure of the initial condition). Here, we note that the present framework allows for a slight connection between the initial condition and the graph: the law $\nu_0^{x_k}$ of $\theta_{k,0}$ depends on its position x_k which encodes for the way the graph $\mathcal{G}^{(n)}$ is built. A simple illustration is when the graph is made of two complete disconnected components (one concerning the particles with positions in $[0, \frac{1}{2}]$, with initial law $\nu_0^x = \mu_1$ and one concerning particles with positions in $[\frac{1}{2}, 1]$, with initial law $\nu_0^x = \mu_2$). In this case, the behavior of the system is governed by (1.5), with macroscopic kernel $W = \mathbf{1}_{[0,\frac{1}{2}]^2} + \mathbf{1}_{[\frac{1}{2},1]^2}$.

Neural field equation and traveling waves

In the particular case where $d = 1$, $c(\theta) = -\alpha\theta$ (for some $\alpha > 0$) and $\Gamma(\theta, \tilde{\theta}) = f(\tilde{\theta})$ (typically f is a sigmoid function), (2.33) becomes:

$$\partial_t \psi(x,t) = -\alpha\psi(x,t) + \int_I f(\psi(y,t))W(x,y)\ell(\mathrm{d}y). \tag{2.44}$$

Equation (2.44) is nothing else than the neural field equation, introduced by Wilson and Cowan [54] and Amari [3] in order to describe the macroscopic activity of a population of neurons with spatial extension. Eq. (2.44) has been the subject of an extensive literature (see [13, 12] and references therein; see in particular the recent work [15] showing that (2.44) is a proper limit for spatially-extended Hawkes processes). An important issue here is the existence and stability of traveling waves [27, 50]. The point we want to raise here is the possibility of studying such traveling waves through the analysis of the corresponding McKean-Vlasov PDE (1.4) (whose dynamics is, to our knowledge, much less studied than (2.44), see [41]), through the identification (2.38). An interesting and open question concerns the possibility of extending this identification beyond finite time scales (as for the Kuramoto model). In this context, it is reasonable to expect that the effect of thermal noise will persist on larger time intervals, resulting in stochastic neural field equations [26, 31, 34, 35].

3. Examples

The point of this section is twofold: to describe generic models $(\mathcal{X}, W)$ that are regular in the sense of Assumption 2.11 (Section 3.1 below) and to give examples of graphs $(\mathcal{G}^{(n)}, \kappa^{(n)})$ that are convergent in the sense of Assumption 2.9 (Sections 3.2, 3.3 and 3.4 below). These examples are directly inspired by the formalism of *W-random graphs*, introduced in [36, 9, 10, 8, 7] and used in [42, 32, 33, 16, 43] (and references therein) in the context of Section 2.5. In this framework, a usual setting is to consider the compact $I = [0, 1]$. This set-up is particularly well adapted to the choice of deterministic regular positions (see Assumption 3.1 below). Nonetheless, as already mentioned, it also makes sense to consider a general state space I endowed with a general probability measure ℓ, where each x_i actually encodes for a real position [14, 15].

3.1. Two classes of regular models

We describe in this paragraph two generic classes of positions $\mathcal{X}$ and macroscopic kernels W and provide simple conditions in both models ensuring that W is regular in the sense of Assumption 2.11.

Deterministic positions

A first set of hypotheses corresponds to deterministic positions [42, 14, 15, 46]:

Assumption 3.1 (Deterministic positions)**.** *We suppose that $I := [0,1]$, endowed with its Lebesgue measure $\ell(\mathrm{d}x) := \mathrm{d}x$. For all $n \geqslant 1$, the sequence $\underline{x} = (x_1^{(n)}, \ldots, x_n^{(n)})$ is deterministic, regularly positioned on I:*

$$x_i^{(n)} := \frac{i}{n},\ n \geqslant 1, i \in [n]\,. \tag{3.1}$$

We set $x_0^{(n)} := 0$ for notational convenience.

A sufficient condition for the regularity of W is to require (2.28) and

$$s_n(W)(\underline{x}) := \sup_{i\in[n]} \left(\sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} |W(x_i, x_k) - W(x_i, y)|\, \mathrm{d}y \right) \xrightarrow[n\to\infty]{} 0. \tag{3.2}$$

Note that (3.2) is in particular true when W is uniformly Lipschitz in (x, y). In the following, we give examples of discontinuous W still satisfying (3.2).

Proposition 3.2. *In the framework of Assumption 3.1, suppose that Assumptions 2.1, 2.4 and 2.6 and conditions (2.28) and (3.2) hold. Then, Assumption 2.11 is verified: the kernel W is regular along the sequence $\mathcal{X}$.*

Proposition 3.2 is proven in Section 4.3.

Random positions

Another general framework concerns the case of random positions [48, 43]:

Assumption 3.3 (Random positions)**.** *Let I be a closed subset of $\mathbb{R}^p$ that is the support of a probability measure $\ell(\mathrm{d}x)$. The sequence*

$$\mathcal{X} = (x_1, x_2, \ldots) \tag{3.3}$$

is the realization of i.i.d. random variables with law $\ell(\mathrm{d}x)$ on I.

For any $p \geqslant 1$ denote by $\|W\|_{L^p} = \|W\|_{L^p(I^2, \ell\otimes\ell)}$ the usual L^p-norm of W on I^2:

$$\|W\|_{L^p} := \left(\int_{I^2} W(x, y)^p \ell(\mathrm{d}x)\ell(\mathrm{d}y) \right)^{\frac{1}{p}}. \tag{3.4}$$

Proposition 3.4. *In the framework of Assumption 3.3, suppose that Assumptions 2.1, 2.4 and 2.6 hold, as well as the following moment condition: there exists some $\chi > 9$ such that*

$$\|W\|_{L^\chi} < \infty. \tag{3.5}$$

Then, for ℓ-almost every realization of the sequence $\mathcal{X}$, Assumption 2.11 is verified: the kernel W is regular along the sequence $\mathcal{X}$.

Proposition 3.4 is proven in Section 4.4.

3.2. A class of convergent graphs

We give now examples of microscopic graphs $\mathcal{G}^{(n)}$ constructed as in Definition 1.2, that fulfill the requirements of Section 2, in both situations of deterministic (Assumption 3.1) and random positions (Assumption 3.3). The present examples fall into the framework of W-random graphs (with possibly unbounded graphons), that is, when the probability field $W_n(x, y)$ is directly constructed upon a predetermined determinisitic kernel $\mathcal{P}(x, y)$. A general framework may be given by (see [8, 32] for similar definitions)

Definition 3.5 (Generic random graph with graphon $\mathcal{P}$)**.** *For fixed (I, ℓ), $n \geqslant 1$ and a given positive measurable kernel $(x, y) \mapsto \mathcal{P}(x, y)$ on I^2, we define*

$$W_n(x, y) := \rho_n \min\left(\frac{1}{\rho_n}, \mathcal{P}(x, y)\right), \tag{3.6}$$

where $\rho_n \in [0, 1]$.

One important aim of [36, 9, 10, 8, 7] (and references therein) is precisely to prove that, under various hypotheses, $\mathcal{G}^{(n)}$ converges to $\mathcal{P}$. In this context, one generally distinguishes between *bounded graphons* $\mathcal{P}$ [36, 9, 10] and *unbounded graphons* (a typical hypothesis being that $\mathcal{P} \in L^p(I^2)$ for some $p \geqslant 1$ see [8, 7] and references therein). Note that when the graphon $\mathcal{P}$ is bounded (and up to the change $\rho_n \leftrightarrow \frac{\rho_n}{\|\mathcal{P}\|_\infty}$, one can always suppose that $\|\mathcal{P}\|_\infty = 1$), (3.6) boils down to

$$W_n(x, y) = \rho_n \mathcal{P}(x, y),\ x, y \in I. \tag{3.7}$$

When $\rho_n = 1$, we are dealing with dense graphs, whereas in the case $\rho_n \xrightarrow[n\to\infty]{} 0$, we consider diluted graphs. A very simple particular case of (3.7) corresponds to $\mathcal{P}(x, y) \equiv 1$ which boils to a (possibly diluted) homogeneous Erdős-Rényi random graph, already studied in [22, 19]. Thus, one has to think of (3.7) as an inhomogeneous version of the Erdős-Rényi case. When $\mathcal{P}$ is not bounded, one usually assumes in (3.6) that $\rho_n \to 0$ and $n\rho_n \to \infty$, as $n \to \infty$.

Here the distinction is not really on the boundedness of $\mathcal{P}$ in (3.6), but rather between $\sup_x \int \mathcal{P}(x, y)\ell(\mathrm{d}y) < +\infty$ and $\sup_x \int \mathcal{P}(x, y)\ell(\mathrm{d}y) = +\infty$. In the first case, there is uniform control on the asymptotic degree of each node in the graph $\mathcal{G}^{(n)}$ whereas in the second, $\mathcal{G}^{(n)}$ has nodes with diverging degree as $n \to \infty$. We treat these two cases in Section 3.3 and 3.4 below.

3.3. Convergent graphs: the case of graphons with uniformly bounded degrees

We assume in this paragraph that W_n is given by (3.6) for

$$\sup_{x\in I} \int \mathcal{P}(x, y)\ell(\mathrm{d}y) < +\infty. \tag{3.8}$$

Here, we adopt a renormalization that is uniform on the nodes $i \in [n]$: set

$$\kappa_i^{(n)} = \kappa_n := \frac{1}{\rho_n},\ i = 1, \ldots, n \text{ and } w_n = \rho_n, \tag{3.9}$$

satisfying

$$\kappa_n = \frac{1}{\rho_n} = o\left(\frac{n}{\log(n)}\right), \text{ as } n \to \infty. \tag{3.10}$$

In this case, the appropriate limit for $(\mathcal{G}^{(n)}, \kappa_n)$ is simply given by $W := \mathcal{P}$ itself. The verification of the hypotheses of Section 2.1 require that

$$\sup_{z\in I} \int \mathcal{P}(z,y)^2 \ell(\mathrm{d}y) < +\infty \text{ and } \inf_{z\in I} \int \mathcal{P}(z,y)\ell(\mathrm{d}y) > 0. \tag{3.11}$$

We address now the question of the convergence of $(\mathcal{G}^{(n)}, \kappa_n)$ to $\mathcal{P}$ (Assumption 2.9) as well as the regularity of the model (Assumption 2.11) in both cases of deterministic (Assumption 3.1) and random positions (Assumption 3.3).

Convergence of the graph $(\mathcal{G}^{(n)}, \kappa^{(n)})$:

Note that (2.15), (2.16) and (2.17) are trivially verified here, as well as (2.18), by (3.10). This dilution condition was already noticed in [22] in the Erdős-Rényi case $\mathcal{P} \equiv 1$: the microscopic graphs that are relevant for the present work have an averaged degree larger than $\log(n)$. The only point that we need to check is (2.20). Note first that when $\mathcal{P}$ is bounded, (2.20) is immediately verified, since we have $\delta_n(\underline{x}) \equiv 0$ for all n. This can be slightly generalized into the following sufficient condition:

Proposition 3.6. *Suppose that the sequence of positions $\mathcal{X}$ and the kernel $\mathcal{P}$ are such that for some $\epsilon \in \left[0, \frac{1}{2}\right)$ and $C > 0$, for all $n \geqslant 1$ sufficiently large,*

$$\sup_{i,j\in[n]} \mathcal{P}(x_i, x_j) \leqslant Cn^{\epsilon}. \tag{3.12}$$

Choose δ such that $\epsilon < \delta < \frac{1}{2}$ and define

$$\rho_n := n^{-\delta}. \tag{3.13}$$

Then, $(\mathcal{G}^{(n)}, \kappa^{(n)})$ given by (3.6) and (3.9) (with ρ_n given by (3.13)) converges to $W := \mathcal{P}$ in the sense of Assumption 2.9.

Proof of Proposition 3.6. It suffices to note that by (3.12) and (3.13), we have, for n sufficiently large, $\min\left(\frac{1}{\rho_n}, \mathcal{P}(x_i, x_k)\right) = \mathcal{P}(x_i, x_k)$ for all $i, k \in [n]$. Hence, $\delta_n(\underline{x}) = 0$ for such n. □

Regular kernels for deterministic positions:

In this paragraph, we suppose that the positions are deterministic $x_i = \frac{i}{n}$ (Assumption 3.1). If $\mathcal{P}$ is bounded, having in mind Proposition 3.2, a simple sufficient condition is $W = \mathcal{P}$ β-Hölder for some $\beta \in (0, 1]$: for such $\mathcal{P}$, (2.28) holds for $\iota_2 = \beta$ and (3.2) is straightforward. The supplementary regularity condition (2.36) required for Theorem 2.18 is also valid. Interesting examples include $\mathcal{P}(x, y) = 1 - \max(x, y)$ or $\mathcal{P}(x, y) = 1 - xy$ which are encountered in the context of dense inhomogeneous graphs (see [7, 6] for many interesting examples). However, note that the hypotheses (2.28), (3.2) and (2.36) are sufficiently general to capture some interesting cases where $\mathcal{P}$ is not continuous: the P-nearest neighbor model [38] corresponds to $\rho_n = 1$ and $\mathcal{P}(x, y) = \mathbf{1}_{|x-y| \leqslant R}$, $x, y \in I$, for some

$R \in (0,1]$. It is immediate to see that (2.28) is true for $\iota_2 = 1$, that $s_n(W) = O\left(n^{-1}\right)$ in (3.2) and that (2.36) holds. Another interesting case of unbounded kernel $\mathcal{P}$ (which still satisfies (3.8)) is $\mathcal{P}(x,y) := \frac{1}{|x-y|^\alpha}$ on $I = [0,1]$ (already considered in [38]). This enters into the present framework for $0 < \alpha < \frac{1}{2}$.

In order to compare with Section 3.4, we end this paragraph with the following example:

Example 3.7 ([32])**.** *Let $(\mathcal{G}^{(n)}, \kappa^{(n)})$ be given by* (3.6) *where*

$$\mathcal{P}(x,y) := (1-\alpha)y^{-\alpha},\ \alpha \in \left[0, \frac{1}{2}\right),\ x,y \in I = [0,1], \tag{3.14}$$

and $\kappa_i^{(n)}$ is given by (3.9) *and $\rho_n = n^{-\delta}$ for some $\alpha < \delta < \frac{1}{2}$. Then, Assumptions 2.9 and 2.11 are satisfied: $(\mathcal{G}^{(n)}, \kappa_n)$ is regular and convergent to $\mathcal{P}$.*

Indeed a rough bound gives that $\sup_{i,j\in[n]} \mathcal{P}(x_i,x_j) \leqslant (1-\alpha)n^\alpha$, so that (3.12) holds for $\epsilon = \alpha$, so that one can conclude about Assumption 2.9 from Proposition 3.6. Moreover,

$$s_n(W)(\underline{x}) = (1-\alpha)\left(\int_0^1 y^{-\alpha}\mathrm{d}y - \frac{1}{n^{1-\alpha}}\sum_{k=1}^n k^{-\alpha}\right) = O\left(\frac{1}{n^{1-\alpha}}\right),$$

which proves (3.2) since $\alpha < 1/2$. In a same way, (2.36) is true. Inequality (2.28) (for $\iota_2 = 1$) is trivial, so that Assumption 2.11 is verified.

Regular kernels for random positions:

In this paragraph, we suppose that the positions are random (Assumption 3.3). In the case of random positions, in addition to (3.11), we need to verify Proposition 3.4 and Proposition 3.6. It turns out that condition (3.5) is sufficient for both: indeed, fix $\mathcal{P} \in L^\chi(I^2)$ with $\chi > 9$ and let $\epsilon \in \left(\frac{3}{\chi}, \frac{1}{2}\right)$. For any $i,j \in [n]$,

$$\mathbb{P}\left(|\mathcal{P}(x_i,x_j)| > n^\epsilon\right) \leqslant \frac{\mathbb{E}\left[|\mathcal{P}(x_i,x_j)|^\chi\right]}{n^{\epsilon\chi}} = \frac{\|\mathcal{P}\|_\chi^\chi}{n^{\epsilon\chi}}.$$

A rough union bound on $i,j \in [n]$ gives

$$\mathbb{P}\left(\sup_{i,j\in[n]} |\mathcal{P}(x_i,x_j)| > n^\epsilon\right) \leqslant \frac{\|\mathcal{P}\|_\chi^\chi}{n^{\epsilon\chi-2}}.$$

Since $\epsilon\chi - 2 > 1$, by Borel-Cantelli Lemma, we have almost surely that $\sup_{i,j\in[n]} |\mathcal{P}(x_i,x_j)| \leqslant n^\epsilon$ for n sufficiently large. In particular, any bounded measurable weights $\mathcal{P}$ in $L^\infty([0,1]^2)$ satisfy the hypotheses. Among interesting examples which have not been addressed so far, one can highlight the case of kernels with values in $\{0,1\}$ (with $\rho_n = 1$). This case corresponds to deterministic graphs (see [42], § 4).

3.4. Convergent graphs: the case of graphons with diverging degrees

We consider here W_n given by (3.6) for $\mathcal{P}$ satisfying

$$\forall x \in I, \int \mathcal{P}(x,y)^2 \ell(\mathrm{d}y) < +\infty \text{ and } \mathcal{P}_* := \inf_{z \in I} \int \mathcal{P}(z,y)\ell(\mathrm{d}y) > 0. \tag{3.15}$$

The point of the paragraph is to discuss the consequences of having possibly

$$\sup_{x \in I} \int \mathcal{P}(x,y)\ell(\mathrm{d}y) = +\infty. \tag{3.16}$$

Here, the uniform renormalization (3.9) is no longer adapted: we consider instead

$$\kappa_i^{(n)} = \frac{n}{\rho_n \sum_{j=1}^n \min\left(\frac{1}{\rho_n}, \mathcal{P}(x_i, x_j)\right)},\ i \in [n]\,. \tag{3.17}$$

This corresponds to renormalizing the interaction in (1.1) by the averaged degree $\sum_j W_n(x_i, x_j)$ of each vertex. Here, the correct choice for the macroscopic interaction kernel W is ([32])

$$W(x,y) := \frac{\mathcal{P}(x,y)}{\int \mathcal{P}(x,z)\ell(\mathrm{d}z)},\ x, y \in I. \tag{3.18}$$

Remark 3.8. *By construction,* $\sup_{x\in I} \int W(x,y)\ell(\mathrm{d}y) = 1$*. Verifying the technical second moment* (2.2) *requires to have* $\int \frac{\mathcal{P}(x,y)^2}{\left(\int \mathcal{P}(x,z)\ell(\mathrm{d}z)\right)^2}\ell(\mathrm{d}y) < \infty$*, which is in particular immediate for Example 3.11 below. Of course, it could be possible to apply the present renormalization* (3.17) *to the previous case of Section 3.3. However, when* $\mathcal{P}$ *is bounded, the two renormalizations* (3.9) *and* (3.17) *lead to slightly different macroscopic models: in* (3.18)*,* W *is renormalized by the factor* $\int \mathcal{P}(x,z)\ell(\mathrm{d}z)$*, which is not natural in the bounded case.*

On the convergence of the graph $(\mathcal{G}^{(n)}, \kappa^{(n)})$*:*

The following result is the counterpart of Proposition 3.6:

Proposition 3.9. *Suppose that the sequence of positions* $\mathcal{X}$ *and the kernel* $\mathcal{P}$ *are such that for some* $\epsilon \in \left[0, \frac{1}{2}\right)$ *and* $C > 0$*, for all* $n \geqslant 1$ *sufficiently large, estimate* (3.12) *holds. Choose also* δ *such that* $\epsilon < \delta < \frac{1}{2}$ *and define* $\rho_n := n^{-\delta}$ *as in* (3.13)*. Suppose*

$$\sup_{i \in [n]} \left| \frac{1}{n} \sum_{j=1}^n \mathcal{P}(x_i, x_j) - \int \mathcal{P}(x_i, z)\ell(\mathrm{d}z) \right| \to 0,\ \text{as } n \to \infty. \tag{3.19}$$

Then, $(\mathcal{G}^{(n)}, \kappa^{(n)})$ *given by* (3.6) *and* (3.17) *converges to* W *given by* (3.18) *in the sense of Assumption 2.9, for the choice of* $\kappa_n := \frac{2n^\delta}{\mathcal{P}_*}$ *and* $w_n := 1$.

Proof of Proposition 3.9. By (3.12) and (3.13), we have, for n sufficiently large, $\min\left(\frac{1}{\rho_n}, \mathcal{P}(x_i, x_k)\right) = \mathcal{P}(x_i, x_k)$ for all $i, k \in [n]$. Hence, for such n, for $i \in [n]$, $\kappa_i^{(n)} = \frac{n^\delta}{\frac{1}{n}\sum_{j=1}^n \mathcal{P}(x_i, x_j)}$. Using (3.19), we have $\inf_{i\in[n]} \frac{1}{n}\sum_{j=1}^n \mathcal{P}(x_i, x_j) \geqslant \frac{1}{2}\int \mathcal{P}(x_i, y)\ell(\mathrm{d}y) \geqslant \frac{\mathcal{P}_*}{2}$, for n sufficiently large.

This proves (2.15) for $\kappa_n := \frac{2n^\delta}{\mathcal{P}_*}$. Condition (2.17) is trivial for $w_n = 1$ and the dilution condition (2.18) holds since $\delta < \frac{1}{2}$. It remains to check (2.20): denote by

$$S_n(x) := \frac{1}{n}\sum_{j=1}^n \mathcal{P}(x, x_j), \quad \text{and } S(x) := \int \mathcal{P}(x,z)\ell(\mathrm{d}z), \; x \in I.$$

With these notations, $\frac{1}{n}\sum_{k=1}^n \left|\kappa_i^{(n)} W_n(x_i, x_k) - W(x_i, x_k)\right| = \frac{|S_n(x_i) - S(x_i)|}{S(x_i)}$ and the result follows immediately from (3.15) and (3.19). □

Regular kernels for random positions (Assumption 3.3):

As for Section 3.3, regularity and convergence holds under sufficient integrability of the kernel $\mathcal{P}$:

Proposition 3.10. *Suppose that Assumption 3.3 holds. For any $\mathcal{P} \in L^\chi(I^2)$ with $\chi > 9$ which verifies* (3.15)*, the following is true: for almost every realization of the sequence $\mathcal{X}$, $(\mathcal{G}^{(n)}, \kappa^{(n)})$ defined by* (3.6) *and* (3.17) *(with $\rho_n = n^{-\delta}$, $\frac{3}{\chi} < \delta < \frac{1}{2}$) converges to W given by* (3.18) *in the sense of Assumption 2.9, (for the choice of $\kappa_n = \frac{2n^\delta}{\mathcal{P}_*}$ and $w_n = 1$) and Assumption 2.11 is verified.*

Proof of Proposition 3.10. Once again, we apply Proposition 3.4 and Proposition 3.9 together with a Borel-Cantelli argument. Let $\epsilon \in \left(\frac{3}{\chi}, \frac{1}{2}\right)$. The same reasoning as before shows that, since $\epsilon\chi - 2 > 1$, we have almost surely that $\sup_{i,j\in[n]} |\mathcal{P}(x_i, x_j)| \leqslant n^\epsilon$ for n sufficiently large. Secondly, let

$$\bar{\mathcal{P}}(x,y) := \mathcal{P}(x,y) - \int \mathcal{P}(x,z)\ell(\mathrm{d}z). \tag{3.20}$$

Compute

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{j=1}^n \bar{\mathcal{P}}(x_i, x_j)\right)^6\right] = \frac{1}{n^6}\sum_{j_1,\ldots,j_6=1}^n \mathbb{E}\left[\prod_{l=1}^6 \bar{\mathcal{P}}(x_i, x_{j_l})\right]. \tag{3.21}$$

Among the sum above, consider the case where one index (for example j_1) is such that $j_1 \notin \{j_2, \ldots, j_6\}$. In this case, conditioning w.r.t. $(x_i, x_j, j \neq j_1)$ within the previous expectation gives $\mathbb{E}\left[\prod_{l=1}^6 \bar{\mathcal{P}}(x_i, x_{j_l})\right] = 0$, by independence of the $(x_k)_{k\in[n]}$ and by definition of $\bar{\mathcal{P}}$. Hence, the nontrivial contributions to (3.21) are necessarily of the form $\mathbb{E}\left[\prod_{l=1}^3 \bar{\mathcal{P}}(x_i, x_{u_l})^2\right]$ for any $(u_1, u_2, u_3) \in [n]^3$. This means that there exists a constant $C > 0$ (independent of $i \in [n]$) such that

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{j=1}^n \bar{\mathcal{P}}(x_i, x_j)\right)^6\right] \leqslant \frac{C}{n^3}. \tag{3.22}$$

By a union bound on $i \in [n]$ and Markov inequality, we obtain, for all $p \geqslant 1$

$$\mathbb{P}\left(\sup_{i\in[n]}\left|\frac{1}{n}\sum_{j=1}^n \bar{\mathcal{P}}(x_i, x_j)\right| \geqslant \frac{1}{p}\right) \leqslant \sum_{i=1}^n \mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^n \bar{\mathcal{P}}(x_i, x_j)\right| \geqslant \frac{1}{p}\right) \leqslant \frac{Cp^6}{n^2}.$$

An application of Borel-Cantelli Lemma shows that (3.19) holds almost surely. □

We finish this section with an example in the deterministic case:

Example 3.11 ([32], Ex. 2.1). *Consider here the model* $(\mathcal{G}^{(n)}, \kappa^{(n)})$ *given by* (3.6) *and* (3.17) *where*

$$\mathcal{P}(x,y) := (1-\alpha)^2 x^{-\alpha} y^{-\alpha},\ \alpha \in \left[0, \frac{1}{2}\right),\ x, y \in I = [0,1]. \tag{3.23}$$

Proposition 3.12. *Suppose that Assumption 3.1 holds. Let* $\alpha \in \left[0, \frac{1}{2}\right)$ *and* $\mathcal{P}$ *defined by* (3.23)*. There exists* $\delta(\alpha) < 1/2$ *such that for all* $\delta(\alpha) < \delta < \frac{1}{2}$*, the renormalized graph* $(\mathcal{G}^{(n)}, \kappa^{(n)})$ *given by* (3.6) *and* (3.17) *with* $\rho_n := n^{-\delta}$ *converges to* $W(x,y) = (1-\alpha)y^{-\alpha}$*, for the choice of* $\kappa_n := \frac{n^\delta}{(1-\alpha)^2}$ *and* $w_n := 1$ *and* W *is regular in the sense of Assumption 2.11.*

Remark 3.13. *The macroscopic limits in Example 3.7 and Example 3.11 are the same, although the underlying graphs* $\mathcal{G}^{(n)}$ *have really different structures. In Example 3.7,* $\mathcal{G}^{(n)}$ *is more or less homogeneous whereas Example 3.11 is much more hub-like: nodes with positions close to* 0 *are connected to the whole population with probability close to* 1*. We see here the effect of the renormalization* (3.17)*: it compensates for the hubs in the graph* $\mathcal{G}^{(n)}$ *so that, even though the graphs* $\mathcal{G}^{(n)}$ *might be different, the renormalized graphs* $\bar{\mathcal{G}}^{(n)}$ *are actually quite similar.*

Proof of Proposition 3.12. Recall the following simple asymptotics: for $\alpha \in (0,1)$, there exist some sequence $\epsilon_n \to_{n\to\infty} 0$ and some constant $C(\alpha) \neq 0$ such that

$$\sum_{k=1}^{n} k^{-\alpha} = \frac{n^{1-\alpha}}{1-\alpha} + C(\alpha) + \epsilon_n,\ n \geqslant 1. \tag{3.24}$$

We first verify Assumption 2.9. First, the easy case where $\alpha \in \left[0, \frac{1}{4}\right)$ can be treated via Proposition 3.9: the following rough bound $\sup_{i,j\in[n]} \mathcal{P}(x_i, x_j) \leqslant (1-\alpha)^2 n^{2\alpha}$ holds so that (3.12) is true for $\epsilon = 2\alpha < \frac{1}{2}$. Morever,

$$\left| \frac{1}{n}\sum_{j=1}^{n} \mathcal{P}(x_i, x_j) - \int \mathcal{P}(x_i, z)\mathrm{d}z \right| = (1-\alpha)x_i^{-\alpha} \left| \frac{1-\alpha}{n^{1-\alpha}} \sum_{j=1}^{n} j^{-\alpha} - 1 \right|,$$

which, since $x_i^{-\alpha} \leqslant n^\alpha$, is of order $n^{-(1-2\alpha)}$, uniformly in $i \in [n]$. Thus, Proposition 3.9 is true, for any δ such that $2\alpha < \delta < \frac{1}{2}$.

The case $\alpha \in \left[\frac{1}{4}, \frac{1}{2}\right)$ is more technical and cannot be dealt via Proposition 3.9 directly. Choose any δ such that $\alpha < \delta < \frac{1}{2} \leqslant 2\alpha < 1$. Since $x_i \leqslant 1$, we have

$$\kappa_i^{(n)} = \frac{n^{1+\delta}}{\sum_{j=1}^{n} \min\left(n^\delta, (1-\alpha)^2 \frac{1}{x_i^\alpha x_j^\alpha}\right)} \leqslant \frac{n^\delta}{(1-\alpha)^2},$$

so that (2.15) and (2.17) are true for $w_n = 1$. The choice of $\delta < \frac{1}{2}$ ensures that (2.18) is verified. We have $\delta_n(\underline{x}) = \sup_{i\in[n]} \delta_{n,i}(\underline{x})$ with

$$\delta_{n,i}(\underline{x}) := \frac{1}{n}\sum_{k=1}^{n} \left| \kappa_i^{(n)} W_n(x_i, x_k) - W(x_i, x_k) \right|,\ i \in [n]\,.$$

Using the notations

$$j_n^i := \left\lfloor (1-\alpha)^{2/\alpha} \frac{n^{2-\delta/\alpha}}{i} \right\rfloor \tag{3.25}$$

and

$$D_{n,i} := \sum_{j=1}^{n} \min\left(n^\delta, (1-\alpha)^2 \frac{1}{x_i^\alpha x_j^\alpha}\right) = j_n^i n^\delta + (1-\alpha)^2 \frac{n^\alpha}{x_i^\alpha} \sum_{j=j_n^i+1}^{n} \frac{1}{j^\alpha},$$

we obtain

$$\begin{aligned}
\delta_{n,i}(\underline{x}) &= \frac{1}{n} \sum_{k=1}^{j_n^i} \left| \frac{n^{1+\delta}}{D_{n,i}} - (1-\alpha) \frac{n^\alpha}{k^\alpha} \right| + \left| \frac{(1-\alpha)n}{D_{n,i} x_i^\alpha} - 1 \right| \left(\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=j_n^i+1}^{n} \frac{1}{k^\alpha} \right), \\
&\leqslant \frac{j_n^i n^\delta}{D_{n,i}} + \frac{1-\alpha}{n^{1-\alpha}} \sum_{k=1}^{j_n^i} \frac{1}{k^\alpha} + \left| \frac{(1-\alpha)n}{D_{n,i} x_i^\alpha} - 1 \right| \left(\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=j_n^i+1}^{n} \frac{1}{k^\alpha} \right).
\end{aligned} \tag{3.26}$$

By (3.24), we have for any $\beta \in (0,1]$

$$\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=1}^{\lfloor n^\beta \rfloor} \frac{1}{k^\alpha} \begin{cases} \to_{n\to\infty} 1 & \text{if } \beta = 1, \\ O\left(\frac{1}{n^{(1-\alpha)(1-\beta)}}\right) \to_{n\to\infty} 0 & \text{if } \beta \in (0,1). \end{cases} \tag{3.27}$$

First observe that for all $i = 1, \ldots, n$ $j_n^i \leqslant \lfloor n^{2-\delta/\alpha} \rfloor$. Since $\delta > \alpha$, $2 - \delta/\alpha < 1$ and by (3.27), the second term in (3.26) is such that

$$\sup_{i\in[n]} \left(\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=1}^{j_n^i} \frac{1}{k^\alpha} \right) \to 0, \text{ as } n \to \infty. \tag{3.28}$$

Moreover we have the existence of $n_0 \geqslant 1$ such that for all $n \geqslant n_0$

$$\inf_{i\in[n]} \left(\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=j_n^i+1}^{n} \frac{1}{k^\alpha} \right) \geqslant \frac{1}{2}. \tag{3.29}$$

Let us now concentrate on the first term of (3.26): for all $i \in [n]$,

$$\begin{aligned}
\frac{j_n^i n^\delta}{D_{n,i}} &= \frac{1}{1 + (1-\alpha)^2 \frac{n^{2\alpha}}{i^\alpha j_n^i n^\delta} \sum_{j=j_n^i+1}^{n} \frac{1}{j^\alpha}}, \\
&\leqslant \frac{1}{1 + (1-\alpha)^{1-2/\alpha} n^{\alpha-\delta-1+\delta/\alpha} \left(\frac{1-\alpha}{n^{1-\alpha}} \sum_{j=j_n^i+1}^{n} \frac{1}{j^\alpha} \right)}, \text{ (by (3.25) and since } i \geqslant 1), \\
&\leqslant \frac{1}{1 + \frac{(1-\alpha)^{1-2/\alpha}}{2} n^{\alpha-\delta-1+\delta/\alpha}} \to 0, \quad \text{by (3.29)},
\end{aligned}$$

since $\alpha - \delta - 1 + \delta/\alpha = \frac{\delta-\alpha}{\alpha}(1-\alpha) > 0$. This proves that the first term of (3.26) converges to 0 uniformly in $i \in [n]$. The third term of (3.26) $\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=j_n^i+1}^{n} \frac{1}{k^\alpha}$ is smaller than $\frac{1-\alpha}{n^{1-\alpha}} \sum_{k=1}^{n} \frac{1}{k^\alpha}$ which converges (to 1) and hence, bounded. It remains to control $\left|\frac{(1-\alpha)n}{D_{n,i}x_i^\alpha} - 1\right|$. We can write

$$\begin{aligned}
\left|\frac{(1-\alpha)n}{D_{n,i}x_i^\alpha} - 1\right| &= \frac{(1-\alpha)n}{D_{n,i}x_i^\alpha}\left|-\frac{j_n^i i^\alpha}{(1-\alpha)n^{1+\alpha-\delta}} + 1 - \frac{1-\alpha}{n^{1-\alpha}}\sum_{j=j_n^i+1}^{n}\frac{1}{j^\alpha}\right|, \\
&\leqslant \frac{(1-\alpha)n}{D_{n,i}x_i^\alpha}\left(\frac{(1-\alpha)^{2/\alpha-1}}{n^{-1+\alpha-\delta+\delta/\alpha}} + \frac{1-\alpha}{n^{1-\alpha}}\sum_{j=1}^{j_n^i}\frac{1}{j^\alpha} + \left|1 - \frac{1-\alpha}{n^{1-\alpha}}\sum_{j=1}^{n}\frac{1}{j^\alpha}\right|\right).
\end{aligned}$$

The terms within the brackets converge to 0, uniformly in $i \in [n]$, (recall (3.27) and (3.28)) and we have

$$\frac{(1-\alpha)n}{D_{n,i}x_i^\alpha} = \frac{1-\alpha}{j_n^i n^{\delta-1} x_i^\alpha + (1-\alpha)^2 n^{\alpha-1}\sum_{j=j_n^i+1}^{n}\frac{1}{j^\alpha}} \leqslant \frac{1}{\frac{1-\alpha}{n^{1-\alpha}}\sum_{j=j_n^i+1}^{n}\frac{1}{j^\alpha}} \leqslant 2,$$

at least for large n, once again by (3.29). This proves (2.20) in the case $\alpha \in \left[\frac{1}{4}, \frac{1}{2}\right)$. The proof of the regularity of W has already been done in Section 3.3. This concludes the proof of Proposition 3.12. □

4. Proofs for the propagation of chaos results

4.1. Proof of Theorem 2.13

For the moment, the sequence of positions $\mathcal{X}$ and the associated connectivity sequence Ξ are fixed. For all $n \geqslant 1$, for fixed $\underline{x}$ and $\underline{\xi}$, we introduce the following quantities

$$b_n(\underline{\xi}) := \sup_{i\in[n]}\left|\frac{1}{n}\sum_{k=1}^{n}\kappa_i^{(n)}\xi_{i,k}^{(n)}\right|, \tag{4.1}$$

$$d_{n,t}(\underline{\xi},\underline{x}) := \sup_{i\in[n]}\int_0^t \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^{n}\kappa_i^{(n)}\left(\xi_{i,k}^{(n)} - W_n(x_i,x_k)\right)[\Gamma]_u(\bar\theta_{i,u},x_k)\right|^2\right]\mathrm{d}u. \tag{4.2}$$

Proposition 4.1. *Suppose that Assumptions 2.1, 2.4 and 2.6 are true. For $n \geqslant 1$ and $T > 0$, let $(\theta_{i,t})_{i\in[n]} := (\theta_{i,t}^{(n)})_{i\in[n]}$ be the solution of (1.1) and $(\bar\theta_{i,t})_{i\in[n]} := (\bar\theta_t^{x_i})_{i\in[n]}$ independent copies of (1.3) with the same initial conditions and Brownian motions as (1.1). There exists a constant $C > 0$ (independent of n and T), such that for any fixed choice of connectivities and positions $(\underline{\xi},\underline{x})$,*

$$\begin{aligned}
\sup_{i\in[n]}\mathbf{E}\left[\sup_{s\leqslant T}\left|\theta_{i,s} - \bar\theta_{i,s}\right|^2\right] &\leqslant C\left(T + e^{CT}\right)e^{Cb_n(\underline{\xi})^2 T} \\
&\times\left(b_n(\underline{\xi})\frac{\kappa_\infty^{(n)}(\underline{x})}{n} + \delta_n(\underline{x})^2 + d_{n,T}(\underline{\xi},\underline{x}) + \sum_{m=1}^{3}\epsilon_{n,T}^{(m)}(\underline{x})\right),
\end{aligned} \tag{4.3}$$

where we recall the definitions of $\kappa_\infty^{(n)}(\underline{x})$ in (2.15), of $\delta_n(\underline{x})$ in (2.19) and of $\epsilon_{n,T}^{(m)}(\underline{x})$ in (2.24).

Proof of Proposition 4.1. For all $s \leqslant t$, $i \in [n]$, by the one-sided Lipschitz-continuity (2.6) of c,

$$
\begin{aligned}
&\left|\theta_{i,s} - \bar{\theta}_{i,s}\right|^2 = 2\int_0^s \left\langle \theta_{i,u} - \bar{\theta}_{i,u}\,,\, c(\theta_{i,u}) - c(\bar{\theta}_{i,u})\right\rangle \mathrm{d}u \\
&+ 2\int_0^s \left\langle \theta_{i,u} - \bar{\theta}_{i,u}\,,\, \left(\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k}\Gamma(\theta_{i,u},\theta_{k,u}) - \int W(x_i,\tilde{x})\Gamma(\bar{\theta}_{i,u},\tilde{\theta})\nu_u(\mathrm{d}\tilde{\theta},\mathrm{d}\tilde{x})\right)\right\rangle \mathrm{d}u, \\
&\leqslant (2L_c+1)\int_0^s \left|\theta_{i,u} - \bar{\theta}_{i,u}\right|^2 \mathrm{d}u + \int_0^s \left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k}\Gamma(\theta_{i,u},\theta_{k,u}) - \int W(x_i,\tilde{x})\Gamma(\bar{\theta}_{i,u},\tilde{\theta})\nu_u(\mathrm{d}\tilde{\theta},\mathrm{d}\tilde{x})\right|^2 \mathrm{d}u.
\end{aligned}
$$

Taking the supremum in $s \leqslant t$ and the expectation w.r.t. Brownian motions and initial conditions,

$$
\begin{aligned}
\mathbf{E}\left[\sup_{s\leqslant t}\left|\theta_{i,s} - \bar{\theta}_{i,s}\right|^2\right] &\leqslant (2L_c+1)\int_0^t \mathbf{E}\left[\sup_{v\leqslant u}\left|\theta_{i,v} - \bar{\theta}_{i,v}\right|^2\right]\mathrm{d}u \\
&+ \int_0^t \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k}\Gamma(\theta_{i,u},\theta_{k,u}) - \int W(x_i,\tilde{x})\Gamma(\bar{\theta}_{i,u},\tilde{\theta})\nu_u(\mathrm{d}\tilde{\theta},\mathrm{d}\tilde{x})\right|^2\right]\mathrm{d}u. \quad (4.4)
\end{aligned}
$$

It remains to control the last term in (4.4), which can be bounded from above by $\int_0^t \left(6\sum_{k=1}^6 A_{n,i,u}^{(k)}\right)\mathrm{d}u$, where, for $u \leqslant t$,

$$A_{n,i,u}^{(1)} := \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k}\left(\Gamma(\theta_{i,u},\theta_{k,u}) - \Gamma(\bar{\theta}_{i,u},\theta_{k,u})\right)\right|^2\right], \quad (4.5)$$

$$A_{n,i,u}^{(2)} := \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k}\left(\Gamma(\bar{\theta}_{i,u},\theta_{k,u}) - \Gamma(\bar{\theta}_{i,u},\bar{\theta}_{k,u})\right)\right|^2\right], \quad (4.6)$$

$$A_{n,i,u}^{(3)} := \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k}\left(\Gamma(\bar{\theta}_{i,u},\bar{\theta}_{k,u}) - [\Gamma]_u(\bar{\theta}_{i,u},x_k)\right)\right|^2\right], \quad (4.7)$$

$$A_{n,i,u}^{(4)} := \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \left(\xi_{i,k} - W_n(x_i,x_k)\right)[\Gamma]_u(\bar{\theta}_{i,u},x_k)\right|^2\right], \quad (4.8)$$

$$A_{n,i,u}^{(5)} := \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \left(\kappa_i^{(n)} W_n(x_i,x_k) - W(x_i,x_k)\right)[\Gamma]_u(\bar{\theta}_{i,u},x_k)\right|^2\right], \quad (4.9)$$

$$A_{n,i,u}^{(6)} := \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n W(x_i,x_k)[\Gamma]_u(\bar{\theta}_{i,u},x_k) - \int W(x_i,\tilde{x})[\Gamma]_u(\bar{\theta}_{i,u},\tilde{x})\ell(\mathrm{d}\tilde{x})\right|^2\right], \quad (4.10)$$

where we recall the definition of $[\Gamma]_u$ in (2.21). Note as this point that, $d_{n,t}$ defined in (4.2) is such that $d_{n,t} = \sup_{i\in[n]}\int_0^t A_{n,i,u}^{(4)}\mathrm{d}u$. Among the other terms, $A_n^{(1)}$ and $A_n^{(2)}$ capture the approximation of the particle system $(\theta_i)_i$ by its mean-field limit $(\bar{\theta}_i)_i$ and $A_n^{(3)}$ relates the empirical measure of the mean-field particle system to its deterministic limit ν. The convergence of $(W_n(x_i, x_j))_{i,j}$ to the macroscopic kernel $(W(x_i, x_j))_{i,j}$ is controlled by $A_n^{(5)}$. By the Lipschitz continuity of Γ (2.4), we have (recall the definition of b_n in (4.1)),

$$A_{n,i,u}^{(1)} \leqslant L_\Gamma^2 b_n(\underline{\xi})^2 \sup_{r\in[n]} \mathbf{E}\left[\sup_{v\in[0,u]} \left|\theta_{r,v} - \bar{\theta}_{r,v}\right|^2\right].$$

In a same way, by (2.4),

$$\begin{aligned}
A_{n,i,u}^{(2)} &\leqslant L_\Gamma^2 \mathbf{E}\left[\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)} \xi_{i,k} \left|\theta_{k,u} - \bar{\theta}_{k,u}\right|\right]^2 = \frac{L_\Gamma^2}{n^2}\sum_{k,l=1}^n \left(\kappa_i^{(n)}\right)^2 \xi_{i,k}\xi_{i,l} \mathbf{E}\left[\left|\theta_{k,u} - \bar{\theta}_{k,u}\right| \left|\theta_{l,u} - \bar{\theta}_{l,u}\right|\right],\\
&\leqslant \frac{L_\Gamma^2}{2n^2}\sum_{k,l=1}^n \left(\kappa_i^{(n)}\right)^2 \xi_{i,k}\xi_{i,l} \mathbf{E}\left[\left|\theta_{k,u} - \bar{\theta}_{k,u}\right|^2 + \left|\theta_{l,u} - \bar{\theta}_{l,u}\right|^2\right] \leqslant L_\Gamma^2 b_n(\underline{\xi})^2 \sup_{r\in[n]} \mathbf{E}\left[\sup_{v\in[0,u]} \left|\theta_{r,v} - \bar{\theta}_{r,v}\right|^2\right].
\end{aligned}$$

Concerning the term $A^{(3)}$, denote by (recall (2.21))

$$\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u}) := \left(\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u}) - [\Gamma]_u(\bar{\theta}_{i,u}, x_k)\right) = \left(\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u}) - \int \Gamma(\bar{\theta}_{i,u}, \tilde{\theta})\nu_u^{x_k}(\mathrm{d}\tilde{\theta})\right). \tag{4.11}$$

One has in particular, using (2.5) and Remark 2.8,

$$\begin{aligned}
\mathbf{E}\left[\left|\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u})\right|^2\right] &\leqslant 2\mathbf{E}\left[\left|\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u})\right|^2\right] + 2\mathbf{E}\left[\int \left|\Gamma(\bar{\theta}_{i,u}, \tilde{\theta})\right| \nu_u^{x_k}(\mathrm{d}\tilde{\theta})\right]^2,\\
&\leqslant 2L_\Gamma^2 \mathbf{E}\left[1 + \left|\bar{\theta}_{i,u}\right| + \left|\bar{\theta}_{k,u}\right|\right]^2 + 2L_\Gamma^2 \mathbf{E}\left[1 + \left|\bar{\theta}_{i,u}\right| + \mathbf{E}\left|\bar{\theta}_{k,u}\right|\right]^2,\\
&\leqslant 12L_\Gamma^2\left(1 + \mathbf{E}\left[\left|\bar{\theta}_{i,u}\right|^2\right] + \mathbf{E}\left[\left|\bar{\theta}_{k,u}\right|^2\right]\right),\\
&\leqslant 12L_\Gamma^2\left(1 + 2\sup_{r\in[n]}\mathbf{E}\left[\sup_{v\leqslant u}\left|\bar{\theta}_{r,v}\right|^2\right]\right) \leqslant 12L_\Gamma^2(1 + 2C_0),
\end{aligned} \tag{4.12}$$

for C_0 given by (2.13). Thus,

$$\begin{aligned}
A_{n,i,u}^{(3)} &\leqslant \mathbf{E}\left[\left|\frac{1}{n}\sum_{k=1}^n \kappa_i^{(n)}\xi_{i,k}\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u})\right|^2\right],\\
&= \frac{\left(\kappa_i^{(n)}\right)^2}{n^2}\sum_{k=1}^n \xi_{i,k}\mathbf{E}\left[\left|\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u})\right|^2\right] + \frac{\left(\kappa_i^{(n)}\right)^2}{n^2}\sum_{k\neq l}^n \xi_{i,k}\xi_{i,l}\mathbf{E}\left[\left\langle\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u}),\, \delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{l,u})\right\rangle\right].
\end{aligned}$$

Since we have supposed that $\xi_{i,i} = 0$, one can suppose that $k \neq i$ and $l \neq i$ in the second sum and conditioning by $\mathcal{F}_k := \sigma(\bar{\theta}_{r,u}, r \neq k)$ gives

$$\begin{aligned}
\mathbf{E}\left[\left\langle\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u}),\, \delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{l,u})\right\rangle\right] &= \mathbf{E}\left[\mathbf{E}\left[\left\langle\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u}),\, \delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{l,u})\right\rangle|\mathcal{F}_k\right]\right],\\
&= \mathbf{E}\left[\left\langle\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{l,u}),\, \mathbf{E}\left[\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u})|\mathcal{F}_k\right]\right\rangle\right] = 0,
\end{aligned}$$

by definition of $\delta\Gamma(\bar{\theta}_{i,u}, \bar{\theta}_{k,u})$. Consequently, by (4.12) (recall the definition of $\kappa_\infty^{(n)}(\underline{x})$ in (2.15)),

$$A^{(3)}_{n,i,u} \leqslant 12L_\Gamma^2 \left(1 + 2C_0\right) \frac{\left(\kappa_i^{(n)}\right)^2}{n^2} \sum_{k=1}^n \xi_{i,k} \leqslant 12L_\Gamma^2 \left(1 + 2C_0\right) b_n(\underline{\xi}) \frac{\kappa_\infty^{(n)}(\underline{x})}{n}.$$

Concerning the term $A^{(5)}_{n,i,u}$, using (A.23) and (2.13) (recall the definition of $\delta_n(\underline{x})$ in (2.19)), we have

$$A^{(5)}_{n,i,u} \leqslant 3L_\Gamma^2 \left(1 + 2C_0\right) \left(\frac{1}{n} \sum_{k=1}^n \left|\kappa_i^{(n)} W_n(x_i, x_k) - W(x_i, x_k)\right|\right)^2 \leqslant 3L_\Gamma^2 \left(1 + 2C_0\right) \delta_n(\underline{x})^2.$$

Finally, let us control the last term $A^{(6)}_{n,i,u}$: using the shortcut

$$\overline{[\Gamma W]}_u(\theta, x, y) = W(x, y)[\Gamma]_u(\theta, y) - \int W(x, z)[\Gamma]_u(\theta, z)\ell(\mathrm{d}z) \tag{4.13}$$

we have (recall the definition of Υ_t in (2.22))

$$\begin{aligned}
\int_0^t A^{(6)}_{n,i,u}\mathrm{d}u &= \int_0^t \int \mathbf{E}\left[\left\langle \overline{[\Gamma W]}_u(\bar{\theta}_{i,u}, x_i, y)\,,\, \overline{[\Gamma W]}_u(\bar{\theta}_{i,u}, x_i, z)\right\rangle\right] \ell_n(\mathrm{d}y)\ell_n(\mathrm{d}z)\mathrm{d}u \\
&= \int W(x_i, y)W(x_i, z)\Upsilon_t(x_i, y, z)\ell_n(\mathrm{d}y)\ell_n(\mathrm{d}z) + \int W(x_i, y)W(x_i, z)\Upsilon_t(x_i, y, z)\ell(\mathrm{d}y)\ell(\mathrm{d}z) \\
&- \int W(x_i, y)W(x_i, z)\Upsilon_t(x_i, y, z)\ell_n(\mathrm{d}y)\ell(\mathrm{d}z) - \int W(x_i, z)W(x_i, y)\Upsilon_t(x_i, y, z)\ell(\mathrm{d}y)\ell_n(\mathrm{d}z).
\end{aligned}$$

So that (recall the definition of the $\epsilon^{(m,i)}_{n,T}(\underline{x})$ in (2.23)),

$$\sup_{i \in [n]} \int_0^t A^{(6)}_{n,i,u}\mathrm{d}u \leqslant \sum_{m=1}^3 \epsilon^{(m)}_{n,t}(\underline{x}).$$

Define now

$$f_t := \sup_{i \in [n]} \mathbf{E}\left[\sup_{s \leqslant t} \left|\theta_{i,s} - \bar{\theta}_{i,s}\right|^2\right], t \leqslant T. \tag{4.14}$$

Taking the supremum on $i \in [n]$ and gathering all the previous estimates in (4.4) gives, for some constant $C = C(\Gamma, c, W, \sigma, \nu_0) > 0$,

$$f_t \leqslant C b_n(\underline{\xi})^2 \int_0^t f_u \mathrm{d}u + C\left(t + e^{Ct}\right)\left(b_n(\underline{\xi})\frac{\kappa_\infty^{(n)}(\underline{x})}{n} + \delta_n(\underline{x})^2\right) + d_{n,t}(\underline{\xi}, \underline{x}) + \sum_{m=1}^3 \epsilon^{(m)}_{n,t}(\underline{x}),\ t \in [0, T].$$

An application of Grönwall's Lemma gives the conclusion. This proves Proposition 4.1. □

At this point of the proof, in (4.3), $\delta_n(\underline{x})$ and $\sum_{m=1}^3 \epsilon^{(m)}_{n,T}(\underline{x})$ go to 0 as $n \to \infty$, by hypothesis. The point now is to prove that the two remaining terms in (4.3) (that depend on the realization of the connectivity sequence $\underline{\xi}$) are such that, first, $b_n(\underline{\xi})$ is bounded and second, that $d_{n,T}(\underline{\xi}, \underline{x})$ goes to 0 as $n \to \infty$, almost surely. This is the purpose of Proposition 4.4 below. The following concentration estimate may be found in [24], Corollary 2.4.7:

Lemma 4.2. *Fix $n \geqslant 1$ and $(Y_l)_{l=1,\dots,n}$ real valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that there exists $v > 0$ such that, almost surely, for all $l = 1, \dots, n-1$, $Y_l \leqslant 1$, $\mathbb{E}[Y_{l+1}|Y_l] = 0$ and $\mathbb{E}\left[Y_{l+1}^2|Y_l\right] \leqslant v$. Then for all $x \geqslant 0$,*

$$\mathbb{P}\left(n^{-1}(Y_1 + \ldots + Y_n) \geqslant x\right) \leqslant \exp\left(-nH\left(\frac{x+v}{1+v}\Big|\frac{v}{1+v}\right)\right), \tag{4.15}$$

where $H(p|q) = p\log(p/q) + (1-p)\log((1-p)/(1-q))$, for $p, q \in [0,1]$.

Using this result, one can prove the following (below $X \sim \mathcal{B}(p)$ stands for a Bernoulli variable with parameter $p \in [0,1]$):

Lemma 4.3. *Fix $n \geqslant 1$, $(p_1, \dots, p_n)$ in $[0,1]$ and a sequence $(v_1, \dots, v_n)$ such that $|v_l| \leqslant 1$ for all $l \in [n]$. Suppose that there exist $\kappa_n > 0$ and $w_n \in (0,1]$, $n \geqslant 1$ satisfying (2.17) and (2.18) such that $p_l \leqslant w_n$ for all $l \in [n]$. Define*

$$\varepsilon_n^2 := 32\frac{\kappa_n^2 w_n}{n}\log(n), \tag{4.16}$$

which goes to 0 as $n \to \infty$, by (2.18). If $(U_1, \dots, U_n)$ are independent random variable with $U_l \sim \mathcal{B}(p_l)$ for all $l \in [n]$, we have the following estimate

$$\mathbb{P}\left(\left|\frac{\kappa_n}{n}\sum_{l=1}^n (U_l - p_l)v_l\right| > \varepsilon_n\right) \leqslant 2\exp\left(-16\log(n)B\left(4\sqrt{2}\left(\frac{\log(n)}{nw_n}\right)^{1/2}\right)\right), \tag{4.17}$$

where

$$B(u) := u^{-2}\left[(1+u)\log(1+u) - u\right]. \tag{4.18}$$

Proof of Lemma 4.3. Fix $n \geqslant 1$, $\kappa_n > 0$, $(p_l)_{l\in[n]}$, $(v_l)_{l\in[n]}$ and $(U_l)_{l\in[n]}$ previously defined. Let $Y_l := (U_l - p_l)v_l$. $(Y_1, \dots, Y_n)$ are independent random variables such that for all $l \in [n]$, $\mathbb{E}[Y_l] = 0$ and $\mathbb{E}\left[Y_l^2\right] = v_l^2\mathbb{V}(U_l) = v_l^2 p_l(1-p_l) \leqslant w_n$, since $|v_l| \leqslant 1$ and $p_l \leqslant w_n$. Then, for all $\varepsilon > 0$, by Lemma 4.2,

$$\mathbb{P}\left(\frac{\kappa_n}{n}\sum_{l=1}^n (U_l - p_l)v_l > \varepsilon\right) \leqslant \exp\left(-nH\left(\frac{\varepsilon\kappa_n^{-1} + w_n}{1+w_n}\Big|\frac{w_n}{1+w_n}\right)\right) \leqslant \exp\left(-\frac{n\varepsilon^2}{2w_n\kappa_n^2}B\left(\frac{\varepsilon}{w_n\kappa_n}\right)\right),$$

where we used the inequality ([24], Exercise 2.4.21)

$$H\left(\frac{x+v}{1+v}\Big|\frac{v}{1+v}\right) \geqslant \frac{x^2}{2v}B\left(\frac{x}{v}\right),\ x, v > 0, \tag{4.19}$$

so that, for the choice of $\varepsilon = \varepsilon_n$ defined by (4.16), we have

$$\mathbb{P}\left(\frac{\kappa_n}{n}\sum_{l=1}^n (U_l - p_l)v_l > \varepsilon_n\right) \leqslant \exp\left(-16\log(n)B\left(4\sqrt{2}\left(\frac{\log(n)}{nw_n}\right)^{1/2}\right)\right)$$

Doing the same for the sequence $(-v_l)_{l\in[n]}$, one obtains (4.17). This proves Lemma 4.3. □

Theorem 2.13 is an immediate consequence of the following result:

Proposition 4.4. *Suppose that Assumptions 2.1, 2.4, 2.6 and 2.9 are true. There exist a deterministic sequence $(\varepsilon_n)_{n\geqslant 1}$ with $\varepsilon_n \to 0$ as $n\to\infty$, a constant $C>0$ such that for all $T>0$, there is an event $\mathcal{O}\in\mathcal{F}$ with $\mathbb{P}(\mathcal{O})=1$ such that the following is true: for every $\omega\in\mathcal{O}$, there exists $n_0<+\infty$, such that for all $n\geqslant n_0$,*

$$b_n(\underline{\xi}(\omega)) \leqslant 2 + \sup_{i\in[n]} \frac{1}{n}\sum_{k=1}^n W(x_i,x_k), \tag{4.20}$$

$$d_{n,T}(\underline{\xi}(\omega),\underline{x}) \leqslant \varepsilon_n C\left(T+e^{CT}\right)\left(1+\sup_{i\in[n]}\frac{1}{n}\sum_{k=1}^n W(x_i,x_k)\right). \tag{4.21}$$

Proof of Proposition 4.4. Introduce the following notations

$$\bar{\xi}_{i,k} := \xi_{i,k} - W_n(x_i,x_k), \tag{4.22}$$

$$\gamma^{(i)}_{k,l,u} = \mathbf{E}\left[[\Gamma]_u(\bar{\theta}_{i,u},x_k)[\Gamma]_u(\bar{\theta}_{i,u},x_l)\right],\ k,l\in[n], \tag{4.23}$$

so that we can rewrite (4.8) as

$$A^{(4)}_{n,i,u} = \frac{\left(\kappa_i^{(n)}\right)^2}{n^2}\sum_{k,l=1}^n \bar{\xi}_{i,k}\bar{\xi}_{i,l}\gamma^{(i)}_{k,l,u}. \tag{4.24}$$

Using (A.23), for some constant $C>0$ independent of i,k,l,u,

$$\begin{aligned}\left|\gamma^{(i)}_{k,l,u}\right| &\leqslant \mathbf{E}\left[\left|[\Gamma]_u(\bar{\theta}_{i,u},x_k)[\Gamma]_u(\bar{\theta}_{i,u},x_l)\right|\right] \leqslant \frac{1}{2}\mathbf{E}\left[\left|[\Gamma]_u(\bar{\theta}_{i,u},x_k)\right|^2\right]+\frac{1}{2}\mathbf{E}\left[\left|[\Gamma]_u(\bar{\theta}_{i,u},x_l)\right|^2\right],\\ &\leqslant C\left(1+\mathbf{E}\left[\left|\bar{\theta}_{i,u}\right|^2\right]\right)\leqslant C\left(1+\sup_{r\in[n]}\mathbf{E}\left[\left|\bar{\theta}_{r,u}\right|^2\right]\right)\leqslant C\left(1+C_0\right),\end{aligned} \tag{4.25}$$

using (2.13). Setting

$$G^{(i)}_{k,l,t} := \int_0^t \gamma^{(i)}_{k,l,u}\mathrm{d}u, \text{ and } \left\|G_t\right\|_\infty := \sup_{k,l,i}\left|G^{(i)}_{k,l,t}\right|, \tag{4.26}$$

one obtains from (4.25) that

$$\left\|G_t\right\|_\infty \leqslant Ct \tag{4.27}$$

for some appropriate constant $C>0$. Let us define now

$$X^{(n)}_{i,k,t} := \frac{\kappa_i^{(n)}}{n}\sum_{l=1}^n \bar{\xi}_{i,l}\frac{G^{(i)}_{k,l,t}}{\left\|G_t\right\|_\infty} \tag{4.28}$$

we can write

$$\begin{aligned}\int_0^t A^{(4)}_{n,i,u}\mathrm{d}u &= \frac{\kappa_i^{(n)}\left\|G_t\right\|_\infty}{n}\sum_{k=1}^n(\xi_{i,k}-W_n(x_i,x_k))X^{(n)}_{i,k,t},\\ &= \frac{\kappa_i^{(n)}\left\|G_t\right\|_\infty}{n}\sum_{k=1}^n(1-W_n(x_i,x_k))\xi_{i,k}X^{(n)}_{i,k,t}+\frac{\kappa_i^{(n)}\left\|G_t\right\|_\infty}{n}\sum_{k=1}^n(-W_n(x_i,x_k))(1-\xi_{i,k})X^{(n)}_{i,k,t}.\end{aligned}$$

Consequently, almost surely, for all $t \geqslant 0$, the following inequality holds

$$d_{n,t} \leqslant \|G_t\|_\infty \left(\sup_{i,k\in[n]} \left| X^{(n)}_{i,k,t} \right| \right) \sup_{i\in[n]} \left(\frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n \xi_{i,k} + \frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n W_n(x_i,x_k) \right). \tag{4.29}$$

We first derive a uniform bound on $\left(X^{(n)}_{i,k,t}\right)_{i,k\in[n]}$. For fixed $i,k \in [n]$, apply Lemma 4.3 for the choice of $U_l = \xi_{i,l}$, $\kappa_n = \kappa_i^{(n)}$, $p_l = W_n(x_i,x_l)$ and $v_l = \frac{G^{(i)}_{k,l,t}}{\|G_t\|_\infty}$: inequality (4.17) together with a simple union bound gives

$$\mathbb{P}\left(\sup_{i,k\in[n]} \left| X^{(n)}_{i,k,t} \right| > \varepsilon_n \right) \leqslant 2n^2 \exp\left(-16\log(n) B\left(4\sqrt{2}\left(\frac{\log(n)}{nw_n} \right)^{1/2} \right) \right).$$

Note that under the assumptions (2.17) and (2.18) on κ_n and w_n, we have

$$\frac{\log(n)}{nw_n} \leqslant \frac{\log(n)\kappa_n^2 w_n}{n} \to 0.$$

Since $B(u) \to \frac{1}{2}$ as $u \to 0$, choose a deterministic $p \geqslant 1$ such that for all $n \geqslant p$, $B\left(4\sqrt{2}\left(\frac{\log(n)}{nw_n}\right)^{1/2}\right) \geqslant \frac{1}{4}$. For such an n,

$$\mathbb{P}\left(\sup_{i,k\in[n]} \left| X^{(n)}_{i,k,t} \right| > \varepsilon_n \right) \leqslant 2n^2 \exp\left(-4\log(n)\right) = \frac{2}{n^2}.$$

Hence, by Borel-Cantelli Lemma, there exists $\mathcal{O}_1 \in \mathcal{F}$ with $\mathbb{P}(\mathcal{O}_1) = 1$ such that, on $\mathcal{O}_1$, there exists $n_1 < +\infty$ such that for all $n \geqslant n_1$,

$$\sup_{i,k\in[n]} \left| X^{(n)}_{i,k,t} \right| \leqslant \varepsilon_n. \tag{4.30}$$

Secondly, apply once again Lemma 4.3 for the choice of $U_l = \xi_{i,l}$, $\kappa_n = \kappa_i^{(n)}$, $p_l = W_n(x_i,x_l)$ and $v_l \equiv 1$. The same reasoning as above gives for $n \geqslant p$,

$$\mathbb{P}\left(\sup_{i\in[n]} \left| \frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n \bar{\xi}_{i,k} \right| > \varepsilon_n \right) \leqslant 2n \exp\left(-4\log(n)\right) = \frac{2}{n^3},$$

so that, there exists $\mathcal{O}_2 \in \mathcal{F}$ such that $\mathbb{P}(\mathcal{O}_2) = 1$, such that on $\mathcal{O}_2$, there exists $n_2 < +\infty$, such that for all $n \geqslant n_2$,

$$\sup_{i\in[n]} \left(\frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n \xi_{i,k} \right) \leqslant \varepsilon_n + \sup_{i\in[n]} \frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n W_n(x_i,x_k) \leqslant 1 + \sup_{i\in[n]} \frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n W_n(x_i,x_k). \tag{4.31}$$

On the event $\mathcal{O} := \mathcal{O}_1 \cap \mathcal{O}_2$ (of probability 1), the inequality (4.29) together with (4.30) and (4.31) gives, for $n \geqslant \max(n_1,n_2)$

$$d_{n,T} \leqslant C\varepsilon_n(T + e^{CT}) \left(1 + 2\sup_{i\in[n]} \frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n W_n(x_i,x_k) \right). \tag{4.32}$$

Using now the fact that $\delta_n(\underline{x}) \to 0$ as $n \to \infty$, $\sup_{i\in[n]} \frac{\kappa_i^{(n)}}{n} \sum_{k=1}^n W_n(x_i,x_k)$ is smaller than $1 + \sup_{i\in[n]} \frac{1}{n}\sum_{k=1}^n W(x_i,x_k)$ at least for large n. This concludes the proof of Proposition 4.4. $\square$

4.2. Proof of Theorem 2.15

Recall that ℓ_n below is the empirical measure of the positions (1.2). We first show that the convergence (2.31) is true for both deterministic positions (Assumption 3.1) and random positions (Assumption 3.3):

Lemma 4.5. *In the deterministic case (Assumption 3.1) (resp. in the random case, Assumption 3.3), the convergence* (2.31) *holds (resp. almost surely).*

Proof of Lemma 4.5. In the deterministic case, it is immediate to see that for any $\varphi \in \mathcal{H}_\iota$ with $\|\varphi\|_{\mathcal{H}_\iota} \leqslant 1$, $|\langle \ell_n - \ell\,,\, \varphi\rangle| \leqslant \frac{1}{(1+\iota)n^\iota}$ so that (2.31) follows directly. We now focus on the random case. First note that the following holds: for all $\varepsilon > 0$, there exists a finite set $\mathcal{H}_\iota^\varepsilon \subset \mathcal{H}_\iota$ such that, for all $n \geqslant 1$, $d_{\mathcal{H}_\iota}(\ell_n, \ell) \leqslant \varepsilon + \max_{\varphi\in\mathcal{H}_\iota^\varepsilon} |\langle \ell_n - \ell\,,\, \varphi\rangle|$. The proof of this point follows closely the proof of [25], Th.11.3.3: for any $\varepsilon > 0$, take $K^\varepsilon \subset I$ compact such that $\ell(K^\varepsilon) > 1 - \varepsilon$. The set of functions $B := \left\{\varphi,\ \|\varphi\|_{\mathcal{H}_\iota} \leqslant 1\right\}$, restricted to K^ε, is compact, by Ascoli-Arzelà theorem. Thus there exists $k \geqslant 1$ and $\varphi_1, \ldots, \varphi_k \in B$ such that for all $\varphi \in B$, there exists $j \leqslant k$ such that $\sup_{y\in K^\varepsilon} |\varphi(y) - \varphi_j(y)| < \varepsilon$, so that $\sup_{x\in K^\varepsilon} |\varphi(x) - \varphi_j(x)| < 3\varepsilon$. Set $g(x) := \max\left(0, 1 - d(x, K^\varepsilon)/\varepsilon\right)$. Then $g \in \mathcal{H}_\iota$ and for n large enough, $\ell_n(K^\varepsilon) \geqslant \int g \mathrm{d}\ell_n > 1 - 2\varepsilon$. Thus

$$|\langle \ell_n - \ell\,,\, \varphi\rangle| \leqslant \int |\varphi - \varphi_j| \,\mathrm{d}(\ell_n + \ell) + |\langle \ell_n - \ell\,,\, \varphi_j\rangle| \leqslant 12\varepsilon + \max_{j=1,\ldots,k} |\langle \ell_n - \ell\,,\, \varphi_j\rangle|\,.$$

Thus, the result follows from the fact that almost surely, for every bounded continuous function φ (and hence, for all $\varphi \in \mathcal{H}_\iota$) we have $|\langle \ell_n - \ell\,,\, \varphi\rangle| \xrightarrow[n\to\infty]{} 0$ ([25], Th. 11.4.1). $\square$

Proof of Theorem 2.15. For $t \in [0, T]$ and $\varphi : \mathbb{R}^d \times I \to \mathbb{R}$ satisfying, for $\theta, \tilde{\theta} \in \mathbb{R}^d$, $x, y \in I$, $\left|\varphi(\theta, x) - \varphi(\tilde{\theta}, y)\right| \leqslant C_\varphi \left(\left|\theta - \tilde{\theta}\right| + |x - y|\right)$ and $\sup_{x\in I} |\varphi(\theta, x)| \leqslant C_\varphi(1 + |\theta|)$, we have

$$\mathbf{E}\left[|\langle \nu_{n,t} - \nu_t\,,\, \varphi\rangle|^2\right] \leqslant 2\mathbf{E}\left[|\langle \nu_{n,t} - \bar{\nu}_{n,t}\,,\, \varphi\rangle|^2\right] + 2\mathbf{E}\left[|\langle \bar{\nu}_{n,t} - \nu_t\,,\, \varphi\rangle|^2\right].$$

Introducing the empirical measure of the nonlinear processes $(\bar{\theta}_i)_{i\in[n]}$ defined in Section 2.3:

$$\bar{\nu}_{n,t} = \frac{1}{n}\sum_{i=1}^n \delta_{(\bar{\theta}_{i,t}, x_i^{(n)})},\ t \geqslant 0, \tag{4.33}$$

we can estimate the first term above as

$$\begin{aligned}\mathbf{E}\left[|\langle \nu_{n,t} - \bar{\nu}_{n,t}\,,\, \varphi\rangle|^2\right] &\leqslant \frac{1}{n}\sum_{i=1}^n \mathbf{E}\left[\left|\varphi(\theta_{i,t}, x_i) - \varphi(\bar{\theta}_{i,t}, x_i)\right|^2\right] \leqslant \frac{C}{n}\sum_{i=1}^n \mathbf{E}\left[\left|\theta_{i,t} - \bar{\theta}_{i,t}\right|^2\right],\\ &\leqslant C \sup_{i\in[n]} \mathbf{E}\left[\sup_{s\in[0,T]} \left|\theta_{i,s} - \bar{\theta}_{i,s}\right|^2\right] \xrightarrow[n\to\infty]{} 0,\end{aligned}$$

by Theorem 2.13. Concerning the second term

$$\mathbf{E}\left[|\langle \bar{\nu}_{n,t} - \nu_t\,,\, \varphi\rangle|^2\right] \leqslant 2\mathbf{E}\left[\left|\frac{1}{n}\sum_{i=1}^n \left(\varphi(\bar{\theta}_{i,t}, x_i) - \int \varphi(\theta, x_i)\nu_t^{x_i}(\mathrm{d}\theta)\right)\right|^2\right] \tag{4.34}$$

$$+ 2\left|\frac{1}{n}\sum_{i=1}^n \int \varphi(\theta, x_i)\nu_t^{x_i}(\mathrm{d}\theta) - \int \varphi(\theta, x)\nu_t^x(\mathrm{d}\theta)\ell(\mathrm{d}x)\right|^2. \tag{4.35}$$

The first term (4.34) above can be computed as

$$\mathbf{E}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\left(\varphi(\bar{\theta}_{i,t},x_i)-\int\varphi(\theta,x_i)\nu_t^{x_i}(\mathrm{d}\theta)\right)\right|^2\right]=\frac{1}{n^2}\sum_{i=1}^{n}\mathbf{E}\left[\left(\varphi(\bar{\theta}_{i,t},x_i)-\int\varphi(\theta,x_i)\nu_t^{x_i}(\mathrm{d}\theta)\right)^2\right],$$
$$\leqslant\frac{1}{n^2}\sum_{i=1}^{n}\int\varphi(\theta,x_i)^2\nu_t^{x_i}(\mathrm{d}\theta)\leqslant\frac{C_\varphi^2}{n^2}\sum_{i=1}^{n}\int(1+|\theta|)^2\,\nu_t^{x_i}(\mathrm{d}\theta),$$

which goes to 0 as $n\to\infty$, uniformly in $t\in[0,T]$, by (2.13). For the last term (4.35), write for simplicity $u_t(x):=\int\varphi(\theta,x)\nu_t^x(\mathrm{d}\theta)$. Note first that $|u_t(x)|\leqslant C_\varphi\int(1+|\theta|)\,\nu_t^x(\mathrm{d}\theta)$ which, by (2.13), is bounded by some $C_1(\varphi,T)$ uniformly in $t\in[0,T]$ and $x\in I$. Secondly, (recall the definition of $\delta\mathcal{W}$ in (2.28)), for some constant $C_2(\varphi,T)>0$ sufficiently large that changes from one line to the other,

$$|u_t(x)-u_t(y)|\leqslant\int|\varphi(\theta,x)-\varphi(\theta,y)|\,\nu_t^x(\mathrm{d}\theta)+\left|\int\varphi(\theta,y)\nu_t^x(\mathrm{d}\theta)-\int\varphi(\theta,y)\nu_t^y(\mathrm{d}\theta)\right|,$$
$$\leqslant C_\varphi\,|x-y|^\iota+C_2(\varphi,T)\,(\delta\mathcal{W}(x,y)+w_1(\nu_0^x,\nu_0^y))\;\leqslant C_2(\varphi,T)\,|x-y|^\iota\,,$$

where we used (A.16), (2.11) and (2.28). Thus, for $C(\varphi,T)=C_1(\varphi,T)+C_2(\varphi,T)$, for all $t\in[0,T]$, $\left\|\frac{u_t}{C(\varphi,T)}\right\|_{\mathcal{H}_\iota}\leqslant 1$. Hence, the term (4.35) can be estimated as

$$\left|\frac{1}{n}\sum_{i=1}^{n}u_t(x_i)-\int u_t(x)\ell(\mathrm{d}x)\right|^2=|\langle\ell_n-\ell\,,\,u_t\rangle|^2\;\leqslant C(\varphi,T)^2d_{\mathcal{H}}(\ell_n,\ell)^2$$

which goes to 0 as $n\to\infty$ by (2.31). □

4.3. Proof of Proposition 3.2: regularity of the kernel in the deterministic case

We suppose here that Assumption 3.1 holds. Since for all $n\geqslant 1$,

$$\left|\frac{1}{n}\sum_{k=1}^{n}W(x_i,x_k)-\int W(x_i,y)\mathrm{d}y\right|\leqslant\sum_{k=1}^{n}\int_{x_{k-1}}^{x_k}|W(x_i,x_k)-W(x_i,y)|\,\mathrm{d}y\leqslant s_n(W),$$

this inequality together with (3.2) and (2.2) implies (2.25). So we are left with proving (2.24). For $i\in[n]$, we have

$$\begin{aligned}\left|\epsilon_{n,T}^{(1,i)}\right|&\leqslant\sum_{k,l=1}^{n}\int_{x_{l-1}}^{x_l}\int_{x_{k-1}}^{x_k}|W(x_i,x_k)W(x_i,x_l)\Upsilon_T(x_i,x_k,x_l)-W(x_i,y)W(x_i,z)\Upsilon_T(x_i,y,z)|\,\mathrm{d}y\mathrm{d}z\\&\leqslant\sum_{k,l=1}^{n}\int_{x_{l-1}}^{x_l}\int_{x_{k-1}}^{x_k}|W(x_i,x_k)W(x_i,x_l)-W(x_i,y)W(x_i,z)|\,|\Upsilon_T(x_i,x_k,x_l)|\,\mathrm{d}y\mathrm{d}z\\&+\sum_{k,l=1}^{n}\int_{x_{l-1}}^{x_l}\int_{x_{k-1}}^{x_k}|\Upsilon_T(x_i,x_k,x_l)-\Upsilon_T(x_i,y,z)|\,W(x_i,y)W(x_i,z)\mathrm{d}y\mathrm{d}z:=(I)+(II).\end{aligned}\tag{4.36}$$

Concerning the first term above, we have, by Lemma A4:

$$(I) \leqslant C \sum_{k,l=1}^{n} \int_{x_{l-1}}^{x_l} \int_{x_{k-1}}^{x_k} |W(x_i, x_k)W(x_i, x_l) - W(x_i, y)W(x_i, z)| \,\mathrm{d}y\mathrm{d}z,$$

$$\leqslant C \left(\frac{1}{n} \sum_{l=1}^{n} W(x_i, x_l) + \int_0^1 W(x_i, y)\mathrm{d}y \right) s_n(W). \tag{4.37}$$

By (2.2), (2.25) and (3.2), this last quantity converges to 0 as $n \to \infty$, uniformly in $i \in [n]$. Concerning the second term in (4.36):

$$(II) \leqslant \sum_{k,l=1}^{n} \int_{x_{l-1}}^{x_l} \int_{x_{k-1}}^{x_k} |\Upsilon_T(x_i, x_k, x_l) - \Upsilon_T(x_i, y, x_l)| \, W(x_i, y)W(x_i, z)\mathrm{d}y\mathrm{d}z,$$

$$+ \sum_{k,l=1}^{n} \int_{x_{l-1}}^{x_l} \int_{x_{k-1}}^{x_k} |\Upsilon_T(x_i, y, x_l) - \Upsilon_T(x_i, y, z)| \, W(x_i, y)W(x_i, z)\mathrm{d}y\mathrm{d}z.$$

Using Lemma A4 again, we have

$$(II) \leqslant C \sum_{k,l=1}^{n} \int_{x_{l-1}}^{x_l} \int_{x_{k-1}}^{x_k} \left(\delta\mathcal{W}(x_k, y) + w_1(\nu_0^{x_k}, \nu_0^{y})\right) W(x_i, y)W(x_i, z)\mathrm{d}y\mathrm{d}z,$$

$$+ C \sum_{k,l=1}^{n} \int_{x_{l-1}}^{x_l} \int_{x_{k-1}}^{x_k} \left(\delta\mathcal{W}(x_l, z) + w_1(\nu_0^{x_l}, \nu_0^{z})\right) W(x_i, y)W(x_i, z)\mathrm{d}y\mathrm{d}z.$$

Using now (2.28) and (2.11), we obtain for $L = \max(L_0, L_W)$

$$(II) \leqslant 2CL \left(\frac{1}{n^{\iota_2}} + \frac{1}{n^{\iota_1}} \right) \int W(x_i, y)W(x_i, z)\mathrm{d}y\mathrm{d}z \leqslant 2CL \left\| \mathcal{W} \right\|_{\infty}^2 \left(\frac{1}{n^{\iota_2}} + \frac{1}{n^{\iota_1}} \right). \tag{4.38}$$

Taking $\sup_{i \in [n]}$ in (4.36), we conclude by (4.37) and (4.38) that $\epsilon^{(1)}_{n,T}(\underline{x}) \xrightarrow[n\to\infty]{} 0$. The two other terms $\epsilon^{(2)}_{n,T}(\underline{x})$ and $\epsilon^{(3)}_{n,T}(\underline{x})$ can be dealt in a similar way, we leave the proof to the reader. This proves Proposition 3.2.

4.4. Proof of Proposition 3.4: regularity of the kernel in the random case

We assume here that Assumption 3.3 holds. For simplicity of notations, we will write $\epsilon_n^{(m,i)}$ in place of $\epsilon_{n,T}^{(m,i)}(\underline{x})$ and $\epsilon_n^{(m)}$ in place of $\epsilon_{n,T}^{(m)}(\underline{x})$ in (2.24). Introduce the following truncation:

$$W_M : (x, y) \mapsto W(x, y) \wedge M, \; M > 0 \tag{4.39}$$

and define $\epsilon_{n,M}^{(m,i)}$ (resp. $\epsilon_{n,M}^{(m)}$) as the truncated version of $\epsilon_n^{(m,i)}$ (resp. $\epsilon_n^{(m)}$), that is, when W is replaced by W_M.

Claim 1: there exists a constant $C > 0$ such that for any $M > 0$, $i \in [n]$, $m = 1, 2, 3$,

$$\left| \epsilon_n^{(m,i)} - \epsilon_{n,M}^{(m,i)} \right| \leqslant C \left(\frac{1}{n} \sum_{l=1}^{n} W(x_i, x_l) + 1 \right)$$

$$\left(\frac{1}{n} \sum_{k=1}^{n} |W(x_i, x_k) - W_M(x_i, x_k)| + \int |W_M(x_i, y) - W(x_i, y)| \, \ell(\mathrm{d}y) \right). \tag{4.40}$$

To prove Claim 1, we only consider $m = 1$ and leave the other cases to the reader. Fix $M > 0$ and $i \in [n]$. Then, using Lemma A4,

$$\begin{aligned}
\Delta^{(1)}_{n,i} &:= \left| \frac{1}{n^2} \sum_{k,l=1}^{n} \{W(x_i, x_k)W(x_i, x_l) - W_M(x_i, x_k)W_M(x_i, x_l)\} \Upsilon_T(x_i, x_k, x_l) \right|, \\
&\leqslant \frac{C}{n^2} \sum_{k,l=1}^{n} |W(x_i, x_k)W(x_i, x_l) - W_M(x_i, x_k)W_M(x_i, x_l)|, \\
&\leqslant C \left(\frac{1}{n} \sum_{k=1}^{n} W(x_i, x_k) + \frac{1}{n} \sum_{k=1}^{n} W_M(x_i, x_k) \right) \left(\frac{1}{n} \sum_{k=1}^{n} |W(x_i, x_k) - W_M(x_i, x_k)| \right), \\
&\leqslant 2C \left(\frac{1}{n} \sum_{k=1}^{n} W(x_i, x_k) \right) \left(\frac{1}{n} \sum_{k=1}^{n} |W(x_i, x_k) - W_M(x_i, x_k)| \right),
\end{aligned}$$

and, in a similar way (recall (2.2))

$$\begin{aligned}
\Delta^{(2)}_{n,i} &:= \left| \int \{W(x_i, y)W(x_i, z) - W_M(x_i, y)W_M(x_i, z)\} \Upsilon_T(x_i, y, z)\ell(\mathrm{d}y)\ell(\mathrm{d}z) \right|, \\
&\leqslant C \int |W(x_i, y)W(x_i, z) - W_M(x_i, y)W_M(x_i, z)| \, \ell(\mathrm{d}y)\ell(\mathrm{d}z), \\
&\leqslant 2C \, \|\mathcal{W}_1\|_\infty \int |W(x_i, y) - W_M(x_i, y)| \, \ell(\mathrm{d}y).
\end{aligned}$$

Doing the same for $m = 2, 3$, this proves Claim 1. The next point is now to show that the quantities in (4.40) can be almost surely controlled for large n, choosing carefully the truncation parameter M.

Claim 2: Let us fix parameters $\delta_1, \delta_2 > 0$ (to be chosen later) and define

$$M := n^{\delta_1}. \tag{4.41}$$

For this choice of M, we have (recall the hypothesis on the norm of W (3.5)), for some constant $C > 0$ independent of n,

$$\mathbb{P}\left(\sup_{i \in [n]} \frac{1}{n} \sum_{k=1}^{n} |W(x_i, x_k) - W_M(x_i, x_k)| > n^{-\delta_2} \right) \leqslant \frac{3 \, \|W\|^{\chi}_{L^\chi(I^2)}}{n^{\delta_1(\chi-1)-\delta_2-1}}, \tag{4.42}$$

$$\mathbb{P}\left(\sup_{i \in [n]} \int |W_M(x_i, \tilde{x}) - W(x_i, \tilde{x})| \, \ell(\mathrm{d}\tilde{x}) > n^{-\delta_2} \right) \leqslant \frac{3 \, \|W\|^{\chi}_{L^\chi(I^2)}}{n^{\delta_1(\chi-1)-\delta_2-1}}, \tag{4.43}$$

$$\mathbb{P}\left(\sup_{i \in [n]} \left| \frac{1}{n} \sum_{l=1}^{n} W(x_i, x_l) \right| > 1 + \|\mathcal{W}_1\|_\infty \right) \leqslant \frac{C}{n^2}. \tag{4.44}$$

Let us prove Claim 2: we have

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{n} \sum_{k=1}^{n} |W(x_i, x_k) - W_M(x_i, x_k)| \right] &= \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[|W(x_i, x_k) - M| \, \mathbf{1}_{W(x_i, x_k) > M} \right], \\
&\leqslant \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[W(x_i, x_k) \mathbf{1}_{W(x_i, x_k) > M} \right].
\end{aligned}$$

Now, for any independent X, Y on I with law ℓ,

$$\begin{aligned}\mathbb{E}\left[W(X,Y)\mathbf{1}_{W(X,Y)>M}\right] &= \sum_{l=0}^{+\infty}\mathbb{E}\left[W(X,Y)\mathbf{1}_{2^l M<W(X,Y)\leqslant 2^{l+1}M}\right],\\ &\leqslant M\sum_{l=0}^{+\infty}2^{l+1}\left(\mathbb{P}\left(W(X,Y)>2^l M\right)-\mathbb{P}\left(W(X,Y)>2^{l+1}M\right)\right),\\ &= M\left(2\mathbb{P}\left(W(X,Y)>M\right)+\sum_{l=1}^{+\infty}2^l\mathbb{P}\left(W(X,Y)>2^l M\right)\right),\\ &\leqslant \frac{\mathbb{E}\left[W(X,Y)^\chi\right]}{M^{\chi-1}}\left(2+\sum_{l=1}^{+\infty}2^{-(\chi-1)l}\right)\leqslant \frac{3\left\|W\right\|_{L^\chi(I^2)}^\chi}{M^{\chi-1}}.\end{aligned}$$

This gives, for M given by (4.41):

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^n\left|W(x_i,x_k)-W_M(x_i,x_k)\right|\right]\leqslant\frac{3\left\|W\right\|_{L^\chi(I^2)}^\chi}{n^{\delta_1(\chi-1)}}.$$

Hence, (4.42) follows immediately from Markov inequality and a union bound. In a same way, we have

$$\mathbb{E}\left[\int\left|W_M(x_i,\tilde{x})-W(x_i,\tilde{x})\right|\ell(\mathrm{d}\tilde{x})\right]\leqslant\frac{3\left\|W\right\|_{L^\chi(I^2)}^\chi}{n^{\delta_1(\chi-1)}},$$

so that inequality (4.43) holds. Inequality

$$\mathbb{P}\left(\sup_{i\in[n]}\left|\frac{1}{n}\sum_{l=1}^n W(x_i,x_l)-\int W(x_i,z)\ell(\mathrm{d}z)\right|>1\right)\leqslant\frac{C}{n^2}\tag{4.45}$$

follows directly from Markov inequality and the fact that $\chi>9\geqslant 6$ in (3.5) and the independence of the variables $(x_k)_{k\in[k]}$. Then (4.44) is a consequence of the inequality $\sup_{i\in[n]}\int W(x_i,z)\ell(\mathrm{d}z)\leqslant\left\|\mathcal{W}_1\right\|_\infty<+\infty$. This proves Claim 2.

Claim 3: Let $\delta_3>0$ be a last constant to be defined later. There exists a constant $C=C_T>0$ such that, for the choice of $M=n^{\delta_1}$ defined in (4.41),

$$\mathbb{P}\left(\epsilon_{n,M}^{(1)}+\epsilon_{n,M}^{(2)}+\epsilon_{n,M}^{(3)}\geqslant C\left(\frac{1}{n^{\delta_3-2\delta_1}}+\frac{1}{n^{1-2\delta_1}}\right)\right)\leqslant Cn\exp\left(-\frac{n^{1-2\delta_3}}{8}\right).\tag{4.46}$$

Let us prove Claim 3: we only control $\epsilon_{n,M}^{(1)}$ and leave the two other terms to the reader. Since for fixed $x\in I$, $(y,z)\mapsto\Upsilon_T(x,y,z)$ is symmetric, we have, for $i\in[n]$

$$\begin{aligned}\epsilon_{n,M}^{(1,i)} &:= \frac{1}{n^2}\sum_{k=1}^n\left\{W_M(x_i,x_k)^2\Upsilon_T(x_i,x_k,x_k)-\int W_M(x_i,y)W_M(x_i,z)\Upsilon_T(x_i,y,z)\ell(\mathrm{d}y)\ell(\mathrm{d}z)\right\}\\ &+\frac{2}{n^2}\sum_{1\leqslant l<k\leqslant n}\left\{W_M(x_i,x_k)W_M(x_i,x_l)\Upsilon_T(x_i,x_k,x_l)-\int W_M(x_i,y)W_M(x_i,z)\Upsilon_T(x_i,y,z)\ell(\mathrm{d}y)\ell(\mathrm{d}z)\right\},\\ &=(I)+(II).\end{aligned}$$

Using Lemma A4, the first term is easily bounded (almost surely for all t) by $\frac{CM^2}{n} = \frac{C}{n^{1-2\delta_1}}$. We now turn to the control of the second term:

$$\begin{aligned}(II) &= \frac{2}{n^2} \sum_{1 \leqslant l<k \leqslant n}^{n} W_M(x_i, x_l) \left\{ W_M(x_i, x_k)\Upsilon_T(x_i, x_k, x_l) - \int W_M(x_i, y)\Upsilon_T(x_i, y, x_l)\ell(\mathrm{d}y) \right\} \\ &+ \frac{2}{n^2} \sum_{1 \leqslant l<k \leqslant n}^{n} \left\{ W_M(x_i, x_l) \int W_M(x_i, y)\Upsilon_T(x_i, y, x_l)\ell(\mathrm{d}y) - \int W_M(x_i, y)W_M(x_i, z)\Upsilon_T(x_i, y, z)\ell(\mathrm{d}y)\ell(\mathrm{d}z \right. \\ &:= \zeta_{n,M}^{(1,i)} + \zeta_{n,M}^{(2,i)}.\end{aligned}$$

We only make the calculations for the $\zeta_{n,M}^{(1,i)}$ and leave the (easier) term $\zeta_{n,M}^{(2,i)}$ to the reader. Denote by

$$\pi_{k,l}^{i,M} = W_M(x_i, x_l) \left\{ W_M(x_i, x_k)\Upsilon_T(x_i, x_k, x_l) - \int W_M(x_i, y)\Upsilon_T(x_i, y, x_l)\ell(\mathrm{d}y) \right\}.$$

By definition of W_M in (4.39) and using Lemma A4, $\Pi_n^M := \sup_{k,l,i} \left| \pi_{k,l}^{i,M} \right| \leqslant C_T M^2 = C_T n^{2\delta_1}$. Writing differently the summation in the definition of $\zeta_{n,M}^{(1,i)}$ leads to

$$\zeta_{n,M}^{(1,i)} = \frac{2\Pi_n^M}{n} \sum_{r=2}^{n} Y_{i,r}^{(M)} = \frac{2\Pi_n^M}{n} Y_{i,i}^{(M)} + \frac{2\Pi_n^M}{n} \sum_{r\in[n],\ r\neq i} Y_{i,r}^{(M)}, \tag{4.47}$$

where we have defined $Y_{i,r}^{(M)} := \frac{1}{n}\sum_{p=1}^{r-1} \frac{\pi_{r,p}^{i,M}}{\Pi_n^M}$, $r \in [n]$. The first term in (4.47) is easily bounded (almost surely for all $t \leqslant T$) by $\frac{2\Pi_n^M}{n} \leqslant \frac{2C}{n^{1-2\delta_1}}$. We now turn to the second term of (4.47): for each for all $r = 2, \dots, n$, $r \neq i$, denote by $\mathcal{F}_r^{(i)}$ the σ-field $\mathcal{F}_r^{(i)} := \sigma\left(x_i, x_k;\ k = 1, \dots, r\right)$. Then,

$$\begin{aligned}\mathbb{E}\left[Y_{i,r}^{(M)} | \mathcal{F}_{r-1}^{(i)}\right] &= \frac{1}{n\Pi_n^M} \sum_{p=1}^{r-1} \mathbb{E}\left[\pi_{r,p}^{(i,M)} \middle| \mathcal{F}_{r-1}^{(i)}\right], \\ &= \frac{1}{n\Pi_n^M} \sum_{p=1}^{r-1} \mathbb{E}\left[W_M(x_i, x_p) \left\{ W_M(x_i, x_r)\Upsilon_T(x_i, x_r, x_p) - \int W_M(x_i, y)\Upsilon_T(x_i, y, x_p)\ell(\mathrm{d}y) \right\} \middle| \mathcal{F}_{r-1}^{(i)} \right],\end{aligned}$$

so that by independence of the $(x_k)_{k\in[n]}$, $\mathbb{E}\left[Y_{i,r}^{(M)} | \mathcal{F}_{r-1}^{(i)}\right] = 0$. Since $Y_{i,r-1}^{(M)}$ is measurable w.r.t. $\mathcal{F}_{r-1}^{(i)}$ we obtain that $\mathbb{E}\left[Y_{i,r}^{(M)} | Y_{i,r-1}^{(M)}\right] = 0$. Note that this calculation only works for $r \neq i$ (this is why we have treated the term $Y_{i,i}^{(M)}$ apart in (4.47)). Since we have by construction $\left|Y_{i,r}^{(M)}\right| \leqslant 1$, we have obviously $\mathbb{E}\left[\left(Y_{i,r}^{(M)}\right)^2 | Y_{i,r-1}^{(M)}\right] \leqslant 1$ for all $r = 2, \dots, n$. We are now in position to apply Lemma 4.2: for all $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\frac{2\Pi_n^M}{n} \sum_{r\in[n],\ r\neq i} Y_{i,r}^{(M)} \geqslant \varepsilon \right) &= \mathbb{P}\left(\frac{1}{n} \sum_{r\in[n],\ r\neq i} Y_{i,r,t}^{(M)} \geqslant \frac{\varepsilon}{2\Pi_n^M} \right), \\ &\leqslant \exp\left(-nH\left(\frac{\frac{\varepsilon}{2}\left(\Pi_n^M\right)^{-1} + 1}{2} \middle| \frac{1}{2} \right) \right) \leqslant \exp\left(-\frac{n\varepsilon^2}{8\left(\Pi_n^M\right)^2} \right).\end{aligned}$$

Doing the same for $-Y_{i,r}$ and by a union bound, we obtain finally, for the choice $\varepsilon := n^{-\delta_3}\Pi_n^M$,

$$\mathbb{P}\left(\sup_{i\in[n]}\left|\frac{2\Pi_n^M}{n}\sum_{r\in[n],\ r\neq i} Y_{i,r}^{(M)}\right| \geqslant n^{-\delta_3}\Pi_n^M\right) \leqslant 2n\exp\left(-\frac{n^{1-2\delta_3}}{8}\right). \tag{4.48}$$

This proves Claim 3.

Conclusion: let $\varepsilon_0 := \chi - 9 > 0$. Define

$$\delta_1 := \frac{1}{4} - \frac{\varepsilon_0}{8(\varepsilon_0+6)} > 0,\ \delta_2 := \frac{\varepsilon_0}{2(\varepsilon_0+6)} > 0 \text{ and } \delta_3 := \frac{1}{2} - \frac{\varepsilon_0}{8(\varepsilon_0+6)} > 0. \tag{4.49}$$

For this choice of parameters, one has obviously $1-2\delta_1 > 0$, $1-2\delta_3 > 0$ and $\delta_3 - 2\delta_1 > 0$, and it is easy to verify that $\delta_1(\chi-1) - \delta_2 - 1 > 1$. This means that the probabilities in (4.42), (4.43), (4.46) are summable in n and that both $n^{-(\delta_3-2\delta_1)}$ and $n^{-(1-2\delta_1)}$ go to 0 as $n\to\infty$. From (4.40), (4.42), (4.43), (4.44), (4.46), we deduce from Borel Cantelli Lemma that there exists an event of probability 1, such that on this event,

$$\sum_{m=1}^{3} \epsilon_{n,T}^{(m)}(\underline{x}) \leqslant C\left(n^{-\delta_2} + n^{-(\delta_3-2\delta_1)} + n^{-(1-2\delta_1)}\right) \tag{4.50}$$

and

$$\sup_{i\in[n]} \frac{1}{n}\sum_{k=1}^{n} W(x_i, x_k) \leqslant 1 + \|\mathcal{W}_1\|_\infty. \tag{4.51}$$

This concludes the proof of Proposition 3.4.

5. The spatial field and the nonlinear heat equation

5.1. Uniqueness of a solution

Here, we consider the general case where I is a closed subset of $\mathbb{R}^p$ and ℓ a probability measure with support I. We suppose here that the hypotheses of Proposition 2.17 hold.

Proof of Proposition 2.17. Let φ, ψ be two weak solutions in $\mathcal{C}\left([0,T], L^k(I,\ell)\right)$ with the same initial condition and denote by $\rho := \varphi - \psi$ the difference. Since $\varphi \in \mathcal{C}([0,T], L^k(I,\ell))$ (resp. ψ) is a weak solution, then φ (resp. ψ) belongs to $H^1([0,T], L^k(I,\ell))$ and we have, for all test function J, for almost every $t\in[0,T]$,

$$\left\langle \frac{\mathrm{d}\varphi}{\mathrm{d}t}(\cdot,t)\,,\,J\right\rangle_{L^2(I,\ell)} = \left\langle c(\varphi(\cdot,t)) + \int_I \Gamma(\varphi(\cdot,t),\varphi(y,t))W(\cdot,y)\ell(\mathrm{d}y)\,,\,J\right\rangle_{L^2(I,\ell)}. \tag{5.1}$$

By density, this also true for all test function $(t,x)\mapsto J(x,t)$ in $L^2([0,T]\times I)$. Substracting the two equations for φ and ψ and choosing the test function $J(\cdot,t) = \varphi(\cdot,t) - \psi(\cdot,t)$, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|\rho(\cdot,t)\|^2_{L^2(I,\ell)} &= \langle c(\varphi(\cdot,t)) - c(\psi(\cdot,t))\,,\,\rho(\cdot,t)\rangle_{L^2(I,\ell)} \\ &+ \left\langle \int_I \left\{\Gamma(\varphi(\cdot,t),\varphi(y,t)) - \Gamma(\psi(\cdot,t),\psi(y,t))\right\} W(\cdot,y)\ell(\mathrm{d}y)\,,\,\rho(\cdot,t)\right\rangle_{L^2(I,\ell)}.\end{aligned}$$

The first term is bounded by $L_c \|\rho(\cdot,t)\|^2_{L^2(I,\ell)}$. The second can be evaluated as, by the properties of Γ

$$\begin{aligned}&\left|\left\langle \int_I \left\{\Gamma(\varphi(\cdot,t),\varphi(y,t)) - \Gamma(\psi(\cdot,t),\psi(y,t))\right\} W(\cdot,y)\ell(\mathrm{d}y)\,,\, \rho(\cdot,t)\right\rangle_{L^2(I,\ell)}\right|\\ &\leqslant L_\Gamma \int_I\int_I |\rho(x,t)|^2 W(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x) + L_\Gamma \int_I |\rho(x,t)| \int_I |\rho(y,t)|\, W(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x).\end{aligned}$$

The first term is bounded by $L_\Gamma \|\mathcal{W}_1\|_\infty \|\rho(\cdot,t)\|^2_{L^2(I,\ell)}$. Concerning the second, by Cauchy-Schwarz inequality,

$$\begin{aligned}\int_I |\rho(x,t)| \int_I |\rho(y,t)|\, W(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x) &\leqslant \int_I |\rho(x,t)| \left(\int_I |\rho(y,t)|^2\, \ell(\mathrm{d}y)\right)^{\frac{1}{2}} \left(\int_I W(x,y)^2 \ell(\mathrm{d}y)\right)^{\frac{1}{2}} \ell(\mathrm{d}x)\\ &\leqslant \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \|\rho(\cdot,t)\|_{L^2(I,\ell)} \|\rho(\cdot,t)\|_{L^1(I,\ell)} \leqslant \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \|\rho(\cdot,t)\|^2_{L^2(I,\ell)}.\end{aligned}$$

Using that $\rho(\cdot,0)\equiv 0$, we obtain uniqueness by a Grönwall's Lemma. □

5.2. Convergence of the spatial profile

The point of this paragraph is to prove Theorem 2.18. Recall that $I=[0,1]$, $\ell(\mathrm{d}x)=\mathrm{d}x$ with regular deterministic positions (Assumption 3.1) and we suppose that the hypotheses of Theorem 2.18 hold. We use the shortcut $\{x\}_n := \lfloor nx+1\rfloor$ for any $x\in I$. Recall the definition of the spatial field $\theta_n(x,t)=\theta^{(n)}_{\{x\}_n,t}$ in (2.35). Introduce the coupling (where the initial conditions and Brownian motions are the same as for (1.1)):

$$\mathrm{d}\psi^{(n)}_{i,t} = c(\psi^{(n)}_{i,t})\mathrm{d}t + \frac{1}{n}\sum_{j=1}^n W(x_i,x_j)\Gamma\left(\psi^{(n)}_{i,t},\psi^{(n)}_{j,t}\right)\mathrm{d}t + \sigma \mathrm{d}B_{i,t},\ 0\leqslant t\leqslant T,\ i\in[n]\,, \tag{5.2}$$

as well as its corresponding spatial field (recall that $x_0=0$ by definition):

$$\psi_n(x,t) := \sum_{i=1}^n \psi^{(n)}_{i,t}\ \mathbf{1}_{[x_{i-1},x_i)}(x) = \psi^{(n)}_{\{x\}_n,t},\ x\in I. \tag{5.3}$$

With these notations at hand, we directly see that $(\theta_n(x,t),\psi_n(x,t))$ are solutions to

$$\mathrm{d}\theta_n(x,t) = c(\theta_n(x,t))\mathrm{d}t + \int_I \Gamma(\theta_n(x,t),\theta_n(y,t))\Xi^{(n)}_{x,y}\mathrm{d}y\mathrm{d}t + \sigma\mathrm{d}B_{\{x\}_n,t}, \tag{5.4}$$

$$\mathrm{d}\psi_n(x,t) = c(\psi_n(x,t))\mathrm{d}t + \int_I \Gamma(\psi_n(x,t),\psi_n(y,t))\hat{W}_n(x,y)\mathrm{d}y\mathrm{d}t + \sigma\mathrm{d}B_{\{x\}_n,t}, \tag{5.5}$$

where

$$\Xi^{(n)}_{x,y} := \kappa^{(n)}_{\{x\}_n}\xi^{(n)}_{\{x\}_n,\{y\}_n} \text{ and } \hat{W}_n(x,y) := W(\{x\}_n,\{y\}_n). \tag{5.6}$$

Proposition 5.1. *Under the hypotheses of Section 2.5, the process $(\theta_n(\cdot,t))_{n\geqslant 1,t\in[0,T]}$ is tight in $\mathcal{C}([0,T],L^k([0,1]))$.*

Proof. Apply Ito's formula:

$$
\begin{aligned}
|\theta_n(x,u)|^k = |\theta_n(x,0)|^k &+ k\int_0^u |\theta_n(x,s)|^{k-2} \langle \theta_n(x,s)\,,\, c(\theta_n(x,s))\rangle \,\mathrm{d}s \\
&+ k\int_0^u |\theta_n(x,s)|^{k-2} \left\langle \theta_n(x,s)\,,\, \int \Gamma\left(\theta_n(x,s),\theta_n(y,s)\right) \Xi^{(n)}_{x,y}\mathrm{d}y \right\rangle \mathrm{d}s \\
&+ k\int_0^u |\theta_n(x,s)|^{k-2} \left\langle \theta_n(x,s)\,,\, \sigma \mathrm{d}B_{\{x\}_n,s} \right\rangle \\
&+ \int_0^u k\left(\frac{k}{2}-1\right) |\theta_n(x,s)|^{k-4} \left\langle \sigma\sigma^\dagger \theta_n(x,s)\,,\, \theta_n(x,s) \right\rangle \mathrm{d}s + kd\int_0^u |\theta_n(x,s)|^{k-2}\,\mathrm{d}s.
\end{aligned}
$$

Let

$$
\theta_n^*(x,t) := \sup_{s\in[0,t]} |\theta_n(x,s)|\,.
$$

Using the hypothesis on c and Γ in Section 2.1, we obtain

$$
\begin{aligned}
\theta_n^*(x,t)^k \leqslant |\theta_n(x,0)|^k &+ k\int_0^t \theta_n^*(x,s)^{k-2}\left(\left(L_c + \frac{1}{2}\right)\theta_n^*(x,s)^2 + \frac{|c(0)|^2}{2}\right)\mathrm{d}s \\
&+ kL_\Gamma \int_0^t \theta_n^*(x,s)^{k-1}\int_I \left(1+\theta_n^*(x,s)+\theta_n^*(y,s)\right)\Xi^{(n)}_{x,y}\mathrm{d}y\mathrm{d}s \\
+k\left((\frac{k}{2}-1)\left|\sigma\sigma^\dagger\right| + d\right) &\int_0^t \theta_n^*(x,s)^{k-2}\mathrm{d}s + k\sup_{u\,\leqslant\, t}\left|\int_0^u |\theta_n(x,s)|^{k-2}\left\langle \theta_n(x,s)\,,\, \sigma\mathrm{d}B_{\{x\}_n,s}\right\rangle\right|.
\end{aligned}
$$

Taking the square, using Jensen's inequality and taking the expectation, we obtain, for some constant $C>0$ depending on k,T,c,Γ,σ,d:

$$
\begin{aligned}
\mathbf{E}\left[\theta_n^*(x,t)^{2k}\right] &\leqslant C\mathbf{E}\left[|\theta_n(x,0)|^{2k}\right] \\
&+ C\left(1+\left(\sup_{x\in I}\int_I \Xi^{(n)}_{x,y}\mathrm{d}y\right)^2\right)\int_0^t \left(\mathbf{E}\left[\theta_n^*(x,s)^{2k-4}\right] + \mathbf{E}\left[\theta_n^*(x,s)^{2k-2}\right] + \mathbf{E}\left[\theta_n^*(x,s)^{2k}\right]\right)\mathrm{d}s, \\
&+ C\mathbf{E}\left[\int_0^t \theta_n^*(x,s)^{2k-2}\left(\int_I \theta_n^*(y,s)\Xi^{(n)}_{x,y}\mathrm{d}y\right)^2\mathrm{d}s\right] \qquad (5.7) \\
&+ C\mathbf{E}\left[\sup_{u\,\leqslant\, t}\left|\int_0^u |\theta_n(x,s)|^{k-2}\left\langle \theta_n(x,s)\,,\, \sigma\mathrm{d}B_{\{x\}_n,s}\right\rangle\right|^2\right]. \qquad (5.8)
\end{aligned}
$$

Concentrate on the term (5.7): for any $x\in I$, setting $i=\{x\}_n$, note that

$$
\begin{aligned}
\left(\int_I \theta_n^*(y,s)\Xi^{(n)}_{x,y}\mathrm{d}y\right)^2 &= \left(\sum_{j=1}^n \kappa_i^{(n)}\xi_{i,j}\int_{x_{j-1}}^{x_j}\theta_n^*(y,s)\mathrm{d}y\right)^2, \\
&\leqslant \left(\frac{1}{n}\sum_{j=1}^n \kappa_i^{(n)}\xi_{i,j}\right)\sum_{j=1}^n \kappa_i^{(n)}\xi_{i,j}\int_{x_{j-1}}^{x_j}\theta_n^*(y,s)^2\mathrm{d}y,
\end{aligned}
$$

applying Jensen's inequality for the probability measure on I $\rho_n = \frac{\sum_{j=1}^n \kappa_i^{(n)} \xi_{i,j} \mathbf{1}_{[x_{j-1},x_j)}(y)\mathrm{d}y}{\frac{1}{n}\sum_{j=1}^n \kappa_i^{(n)}\xi_{i,j}}$. Note also that by Hölder's inequality,

$$\mathbf{E}\left[\theta_n^*(x,s)^{2k-2}\theta_n^*(y,s)^2\right] \leqslant \mathbf{E}\left[\theta_n^*(x,s)^{2k}\right]^{\frac{2k-2}{2k}} \mathbf{E}\left[\theta_n^*(y,s)^{2k}\right]^{\frac{1}{k}}.$$

Hence, if we define

$$\Theta_n(t) := \sup_{x\in I} \mathbf{E}\left[\theta_n^*(x,t)^{2k}\right] \tag{5.9}$$

the term (5.7) may be bounded by $\left(\int_I \Xi_{x,y}^{(n)}\mathrm{d}y\right)^2 \int_0^t \Theta_n(s)\mathrm{d}s$. Finally, concerning the last term (5.8), an application of Burkholder-Davis-Gundy inequality gives, for some constant $C>0$ that depends on σ,

$$\mathbf{E}\left[\sup_{u\leqslant t}\left|\int_0^u |\theta_n(x,s)|^{k-2}\left\langle \theta_n(x,s)\,,\,\sigma\mathrm{d}B_{\{x\}_n,s}\right\rangle\right|^2\right] \leqslant C\int_0^t \mathbf{E}\left[|\theta_n^*(x,s)|^{2k-2}\right]\mathrm{d}s.$$

Putting everything together and using the fact that there exist a universal constant $C_k>0$ (only depending on k) such that $|x|^{2k-4}+|x|^{2k-2} \leqslant C_k + |x|^{2k}$, we obtain finally that for some constant $C'>0$ depending on T,k,σ,c,Γ,

$$\Theta_n(t) \leqslant C'\Theta_n(0) + C'\left(1+\sup_{x\in I}\left(\int_I \Xi_{x,y}^{(n)}\mathrm{d}y\right)^2\right) + C'\left(1+\sup_{x\in I}\left(\int_I \Xi_{x,y}^{(n)}\mathrm{d}y\right)^2\right)\int_0^t \Theta_n(s)\mathrm{d}s.$$

Since we have almost surely $\sup_{x\in I}\left(\int \Xi_{x,y}^{(n)}\mathrm{d}y\right)^2 = \sup_{i\in[n]}\left(\frac{1}{n}\sum_{j=1}^n \kappa_i^{(n)}\xi_{i,j}\right)^2 < \infty$ (recall (4.1), (4.20) and (2.25)) as well as the hypothesis on the initial condition (2.10), we obtain from Grönwall's Lemma that

$$\sup_{n\geqslant 1}\sup_{x\in I}\mathbf{E}\left[\sup_{t\in[0,T]}|\theta_n(x,t)|^{2k}\right] < \infty. \tag{5.10}$$

This implies in particular that, by two successive applications of Jensen's inequality,

$$\begin{aligned}
\sup_{n\geqslant 1}\mathbf{E}\left[\sup_{t\in[0,T]}\|\theta_n(\cdot,t)\|_{L^k(I)}^2\right] &= \sup_{n\geqslant 1}\mathbf{E}\left[\sup_{t\in[0,T]}\left(\int|\theta_n(x,t)|^k\,\mathrm{d}x\right)^{\frac{2}{k}}\right],\\
&\leqslant \sup_{n\geqslant 1}\mathbf{E}\left[\left(\int\sup_{t\in[0,T]}|\theta_n(x,t)|^{2k}\,\mathrm{d}x\right)^{\frac{1}{k}}\right] \leqslant \sup_{n\geqslant 1}\mathbf{E}\left[\int\sup_{t\in[0,T]}|\theta_n(x,t)|^{2k}\,\mathrm{d}x\right]^{\frac{1}{k}},\\
&= \sup_{n\geqslant 1}\left(\int\mathbf{E}\left[\sup_{t\in[0,T]}|\theta_n(x,t)|^{2k}\right]\mathrm{d}x\right)^{\frac{1}{k}} \leqslant \left(\sup_{n\geqslant 1}\sup_{x\in I}\mathbf{E}\left[\sup_{t\in[0,T]}|\theta_n(x,t)|^{2k}\right]\right)^{\frac{1}{k}} < \infty.
\end{aligned} \tag{5.11}$$

We now turn to a similar estimate concerning the modulus of continuity: set $\delta>0$ and $s<t\in[0,T]$ such that $|t-s|\leqslant\delta$. Set

$$\Delta_{n,s,t}(x) := |\theta_n(x,t)-\theta_n(x,s)| \tag{5.12}$$

and

$$\Delta^*_{n,\delta}(x) := \sup_s \Delta_{n,s,s+\delta}(x) \tag{5.13}$$

Then, by the same calculations as before, it is straightforward to see that, for another constant $C' > 0$,

$$\sup_{x\in I} \mathbf{E}\left[\Delta^*_{n,\delta}(x)^{2k}\right] \leqslant C'\left(1 + \left(\sup_{x\in I}\int_I \Xi^{(n)}_{x,y}\mathrm{d}y\right)^2\right)\int_0^\delta \sup_{x\in I}\mathbf{E}\left[\Delta^*_{n,u}(x)^{2k}\right]\mathrm{d}u$$
$$+ C'\mathbf{E}\left[\sup_{s,t,|t-s|\leqslant\delta}\left|\int_s^t |\theta_n(x,u)-\theta_n(x,s)|^{k-2}\left\langle \theta_n(x,u)-\theta_n(x,s)\,,\,\sigma\mathrm{d}B_{\{x\}_n,u}\right\rangle\right|^2\right].$$

Hence, once again by Hölder and Burkholder-Davis-Gundy inequalities,

$$\sup_x \mathbf{E}\left[\Delta^*_{n,\delta}(x)^{2k}\right] \leqslant C''\delta + C''\left(1+\sup_{x\in I}\left(\int \Xi^{(n)}_{x,y}\mathrm{d}y\right)^2\right)\int_0^\delta \sup_{x\in I}\mathbf{E}\left[\Delta^*_{n,u}(x)^{2k}\right]\mathrm{d}u.$$

Grönwall's Lemma gives directly that

$$\sup_{n\geqslant 1}\sup_{x\in I}\mathbf{E}\left[\sup_{|t-s|\leqslant\delta}|\theta_n(x,t)-\theta_n(x,s)|^{2k}\right]\xrightarrow[\delta\searrow 0]{} 0, \tag{5.14}$$

which finally gives

$$\sup_{n\geqslant 1}\mathbf{E}\left[\sup_{|t-s|\leqslant\delta}\|\theta_n(\cdot,t)-\theta_n(\cdot,s)\|^2_{L^k(I)}\right]\xrightarrow[\delta\searrow 0]{} 0. \tag{5.15}$$

The required tightness result follows directly from (5.11) and (5.15) (see for example [45]). □

Proposition 5.2. *Under the assumption of Section 2.5, any accumulation point of* $(\theta_n(\cdot,t))_{n\geqslant 1, t\in[0,T]}$ *in* $\mathcal{C}([0,T],L^k(I))$ *is a weak solution to* (2.34).

Proof. Let us consider the field ψ_n (5.3) corresponding to the particle system (5.2) driven by the same Brownian motions as (1.1) with the same initial condition. By calculations similar to the previous proof, the difference $\theta_n - \psi_n$ verifies, using the assumptions on c,

$$\mathbf{E}\left[|\theta_n(x,t)-\psi_n(x,t)|^{2k}\right] \leqslant C\int_0^t \mathbf{E}\left[|\theta_n(x,u)-\psi_n(x,u)|^{2k}\right]\mathrm{d}u + C\mathbf{E}\int_0^t |\theta_n(x,u)-\psi_n(x,u)|^{2k-2}$$
$$\left\langle \theta_n(x,u)-\psi_n(x,u)\,,\, \int\left(\Gamma\left(\theta_n(x,u),\theta_n(y,u)\right)\Xi^{(n)}_{x,y} - \Gamma\left(\psi_n(x,u),\psi_n(y,u)\right)\hat{W}_n(x,y)\right)\mathrm{d}y\right\rangle\mathrm{d}u$$

Using the assumptions on Γ, (2.4), the difference in the last term above can be bounded by $C(A_1+A_2)$ where A_1 and A_2 are given below. First,

$$\begin{aligned}
A_1 &:= \int_0^t \mathbf{E}\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k-2} \\
&\quad \int_I \langle \theta_n(x,u)-\psi_n(x,u)\,,\,\Gamma(\theta_n(x,u),\theta_n(y,u))-\Gamma(\psi_n(x,u),\psi_n(y,u))\rangle\, \Xi^{(n)}_{x,y}\mathrm{d}y\mathrm{d}u, \\
&\leqslant C\int_0^t \mathbf{E}\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k-1}\int_0^1 \left(\left|\theta_n(x,u)-\psi_n(x,u)\right|+\left|\theta_n(y,u)-\psi_n(y,u)\right|\right)\Xi^{(n)}_{x,y}\mathrm{d}y\mathrm{d}u, \\
&\leqslant 2C\int_0^1 \Xi^{(n)}_{x,y}\mathrm{d}y\int_0^t \mathbf{E}\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k}\mathrm{d}u \\
&\quad + 2C\int_0^t\int_0^1 \mathbf{E}\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k-1}\left|\theta_n(y,u)-\psi_n(y,u)\right|\Xi^{(n)}_{x,y}\mathrm{d}y\mathrm{d}u.
\end{aligned}$$

By Hölder's inequality followed by Young's inequality,

$$\begin{aligned}
&\mathbf{E}\left[\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k-1}\left|\theta_n(y,u)-\psi_n(y,u)\right|\right] \\
&\quad\leqslant \mathbf{E}\left[\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k}\right]^{\frac{2k-1}{2k}}\mathbf{E}\left[\left|\theta_n(y,u)-\psi_n(y,u)\right|^{2k}\right]^{\frac{1}{2k}}, \\
&\quad\leqslant \frac{2k-1}{2k}\mathbf{E}\left[\left|\theta_n(x,u)-\psi_n(x,u)\right|^{2k}\right]+\frac{1}{2k}\mathbf{E}\left[\left|\theta_n(y,u)-\psi_n(y,u)\right|^{2k}\right], \\
&\quad\leqslant \sup_{z\in[0,1]}\mathbf{E}\left[\left|\theta_n(z,u)-\psi_n(z,u)\right|^{2k}\right].
\end{aligned}$$

Hence,

$$A_1 \leqslant 4C\left(\sup_{x\in I}\int_I \Xi^{(n)}_{x,y}\mathrm{d}y\right)\int_0^t \sup_{z\in I}\mathbf{E}\left[\left|\theta_n(z,u)-\psi_n(z,u)\right|^{2k}\right]\mathrm{d}u.$$

Secondly,

$$A_2 := \int_0^t \mathbf{E}\int_I \langle \theta_n(x,u)-\psi_n(x,u)\,,\,\Gamma(\psi_n(x,u),\psi_n(y,u))\rangle\left(\Xi^{(n)}_{x,y}-\hat{W}_n(x,y)\right)\mathrm{d}y\mathrm{d}u.$$

For $i=\{x\}_n$, we have

$$\begin{aligned}
A_2 &:= 2\int_0^t \frac{1}{n}\sum_{j=1}^n \zeta^{(n)}_{i,j,u}\left(\kappa_i^{(n)}\xi_{i,j}-W(x_i,x_j)\right)\mathrm{d}u, \\
&= 2\int_0^t \frac{1}{n}\sum_{j=1}^n \zeta^{(n)}_{i,j,u}\kappa_i^{(n)}\left(\xi_{i,j}-W_n(x_i,x_j)\right)\mathrm{d}u + 2\int_0^t \frac{1}{n}\sum_{j=1}^n \zeta^{(n)}_{i,j,u}\left(\kappa_i^{(n)}W_n(x_i,x_j)-W(x_i,x_j)\right)\mathrm{d}u,
\end{aligned}$$

where $\zeta^{(n)}_{i,j,s} := \mathbf{E}\left[\left\langle \theta^{(n)}_{i,s}-\psi^{(n)}_{i,s}\,,\,\Gamma(\psi^{(n)}_{i,s},\psi^{(n)}_{j,s})\right\rangle\right]$. The same reasoning (together with the apriori control (2.14)) as in Proposition 4.4 shows that the supremum in $i\in[n]$ of the first term in the sum above goes almost surely to 0 as $n\to\infty$. The second term in the sum is bounded above by $\sup_{i\in[n]}\frac{C}{n}\sum_{j=1}^n\left|\kappa_i^{(n)}W_n(x_i,x_j)-W(x_i,x_j)\right|$, which goes to

0 uniformly in i as $n \to \infty$, by (2.19). From the previous estimates and a Grönwall's Lemma, we conclude that $\sup_{x\in I, t\in[0,T]} \mathbf{E}\left|\theta_n(x,t) - \psi_n(x,t)\right|^{2k} \to_{n\to\infty} 0$. In particular,

$$\sup_{t\in[0,T]} \mathbf{E} \int_I \left|\theta_n(x,t) - \psi_n(x,t)\right|^{2k} \mathrm{d}x \to 0, \text{ as } n \to \infty. \tag{5.16}$$

It now remains to identify the limit. By (5.16), any limit point of $(\theta_n(\cdot,t))_{n\geqslant 1}$ in $\mathcal{C}([0,T], L^k(I))$ is also a limit point of $(\psi_n(\cdot,t))_{n\geqslant 1}$. Recall that, for any $\mathcal{C}^1$-bounded test function $J : I \to \mathbb{R}^d$ with bounded derivative

$$\begin{aligned}\int_I \langle \psi_n(x,t)\,,\, J(x)\rangle \,\mathrm{d}x &= \int_I \langle \psi_n(x,0)\,,\, J(x)\rangle \,\mathrm{d}x + \int_0^t \int_I \langle c(\psi_n(x,s))\,,\, J(x)\rangle \,\mathrm{d}x\mathrm{d}s \\ &+ \int_0^t \int_{I^2} \left\langle \Gamma(\psi_n(x,s), \psi_n(y,s))\hat{W}_n(x,y)\,,\, J(x)\right\rangle \mathrm{d}x\mathrm{d}y\mathrm{d}s + \int_I \left\langle \sigma B_{\{x\}_n, t}\,,\, J(x)\right\rangle \mathrm{d}x.\end{aligned} \tag{5.17}$$

Concerning the initial condition $\theta_n(x,0) = \psi_n(x,0)$ (recall the definition of ψ_0 in (2.37)):

$$\begin{aligned}&\mathbf{E}\left|\int_I \langle \theta_n(x,0)\,,\, J(x)\rangle \,\mathrm{d}x - \int_I \langle \psi_0(x)\,,\, J(x)\rangle \,\mathrm{d}x\right|^2 \\ &\qquad\leqslant 3\mathbf{E}\left|\frac{1}{n}\sum_{k=1}^n \left\langle \theta_{k,0}^{(n)} - \mathbf{E}\left[\theta_{k,0}^{(n)}\right]\,,\, n\int_{x_{k-1}}^{x_k} J(u)\mathrm{d}u\right\rangle\right|^2 \\ &\qquad+ 3\left|\frac{1}{n}\sum_{k=1}^n \left\langle \int \theta\nu_0^{x_k}(\mathrm{d}\theta)\,,\, n\int_{x_{k-1}}^{x_k} J(u)\mathrm{d}u - J(x_k)\right\rangle\right|^2 \\ &\qquad+ 3\left|\frac{1}{n}\sum_{k=1}^n \left\langle \int \theta\nu_0^{x_k}(\mathrm{d}\theta)\,,\, J(x_k)\right\rangle - \int_I \langle \psi_0(x)\,,\, J(x)\rangle \,\mathrm{d}x\right|^2, \\ &\leqslant \frac{3\left\|J\right\|_\infty^2}{n} \sup_{x\in I} \int \left|\theta - \int \theta\nu_0^x(\mathrm{d}\theta)\right|^2 \nu_0^x(\mathrm{d}\theta) + 3\sup_{x\in I}\left|\int \theta\nu_0^x(\mathrm{d}\theta)\right|^2 \frac{\left\|J'\right\|_\infty^2}{4n^2} \\ &\qquad+ 3\left|\frac{1}{n}\sum_{k=1}^n \left\langle \int \theta\nu_0^{x_k}(\mathrm{d}\theta)\,,\, J(x_k)\right\rangle - \int_I \langle \psi_0(x)\,,\, J(x)\rangle \,\mathrm{d}x\right|^2.\end{aligned}$$

The two first terms above converge to 0, by (2.10). It is straightforward to see that the third term is $3\left|\int_I \langle \psi_0(\cdot)\,,\, J(\cdot)\rangle\,(\mathrm{d}\ell_n - \mathrm{d}\ell)\right|$ where $x \mapsto \langle \psi_0(x)\,,\, J(x)\rangle$ is ι_1-Hölder, by (2.11). Hence, this term also goes to 0 as $n \to \infty$. Concerning the noise term in (5.17), we have, for some constant $C > 0$

$$\begin{aligned}\mathbf{E}\left[\left|\int_I \left\langle \sigma B_{\{x\}_n, t}\,,\, J(x)\right\rangle \mathrm{d}x\right|^2\right] &= \mathbf{E}\left[\left|\sum_{j=1}^n \int_{x_{j-1}}^{x_j} \langle \sigma B_{j,t}\,,\, J(x)\rangle \,\mathrm{d}x\right|^2\right], \\ &= \sum_{i,j=1}^n \int_{x_{i-1}}^{x_i}\int_{x_{j-1}}^{x_j} \mathbf{E}\left[\langle \sigma B_{i,t}\,,\, J(x)\rangle \langle \sigma B_{j,t}\,,\, J(y)\rangle\right] \mathrm{d}x\mathrm{d}y, \\ &\leqslant Ct\sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} |J(x)|\,\mathrm{d}x\right)^2 \leqslant \frac{CT}{n}\sum_{i=1}^n \int_{x_{i-1}}^{x_i} |J(x)|^2\,\mathrm{d}x = \frac{C(\sigma)T}{n}\int_I |J(x)|^2\,\mathrm{d}x,\end{aligned}$$

which goes to 0 as $n \to \infty$, uniformly on $t \in [0,T]$.

Let $(\psi(x,t))_{x\in I, t\in[0,T]} \in \mathcal{C}([0,T], L^k(I))$ an accumulation point of (ψ_n): there exists a subsequence $(n_k)_{k\geqslant 1}$ (that we rename n for simplicity of exposition) such that $(\psi_n)_{n\geqslant 1}$ converges in law to ψ in $\mathcal{C}([0,T], L^k(I))$. By the Skhorokhod representation theorem ($\mathcal{C}([0,T], L^k(I))$ is separable, [2], p.125), one can suppose that (ψ_n) converges almost surely in $\mathcal{C}([0,T], L^k(I))$ to ψ. Since we have convergence in $\mathcal{C}([0,T], L^k(I))$, we have that $\int_I \langle \psi_n(x,t)\,,\, J(x)\rangle\, \mathrm{d}x \xrightarrow[n\to\infty]{} \int_I \langle \psi(x,t)\,,\, J(x)\rangle\, \mathrm{d}x$, uniformly in $t \in [0,T]$. For the same reason, $\int_0^t \int_I \langle c(\psi_n(x,s))\,,\, J(x)\rangle\, \mathrm{d}x\mathrm{d}s \xrightarrow[n\to\infty]{} \int_0^t \int_I \langle c(\psi(x,s))\,,\, J(x)\rangle\, \mathrm{d}x\mathrm{d}s$ (by an application of dominated convergence theorem when c is bounded or using the convergence in $\mathcal{C}([0,T], L^k(I))$ when c is polynomial). We now turn the interaction term in (5.17). Assume first that Γ is bounded. Then,

$$\left|\int_0^t \int_{I^2} \left\langle \Gamma(\psi_n(x,s), \psi_n(y,s)) \left(\hat{W}_n(x,y) - W(x,y)\right)\,,\, J(x)\right\rangle \mathrm{d}x\mathrm{d}y\mathrm{d}s\right| \leqslant \\ \left\|\Gamma\right\|_\infty \left\|J\right\|_\infty T \int_{I^2} \left|\hat{W}_n(x,y) - W(x,y)\right| \mathrm{d}x\mathrm{d}y,$$

which goes to 0 as $n \to \infty$, by (2.36). Moreover,

$$\left|\int_0^t \int_{I^2} \langle (\Gamma(\psi_n(x,s), \psi_n(y,s)) - \Gamma(\psi(x,s), \psi(y,s)))\, W(x,y)\,,\, J(x)\rangle\, \mathrm{d}x\mathrm{d}y\mathrm{d}s\right| \\ \leqslant \int_0^t \int_{I^2} \left|\Gamma(\psi_n(x,s), \psi_n(y,s)) - \Gamma(\psi(x,s), \psi(y,s))\right| W(x,y) \left|J(x)\right| \mathrm{d}x\mathrm{d}y\mathrm{d}s,$$

which goes to 0 as $n \to \infty$, by dominated convergence theorem. In the case where $\Gamma(\theta, \theta') = \Gamma \cdot (\theta - \theta')$ is linear, we have firstly

$$\left|\int_0^t \int_{I^2} \left\langle (\psi_n(x,s) - \psi(x,s))\, \hat{W}_n(x,y)\,,\, J(x)\right\rangle \mathrm{d}x\mathrm{d}y\mathrm{d}s\right| \\ \leqslant \sup_{z\in I}\left(\int_I \hat{W}_n(z,y)\mathrm{d}y\right) \left\|J\right\|_\infty \int_0^t \int_I \left|\psi_n(x,s) - \psi(x,s)\right| \mathrm{d}x\mathrm{d}s$$

which goes to 0 as $n \to \infty$, by (2.25) and since we have convergence in $\mathcal{C}([0,T], L^k(I))$. Secondly,

$$\left|\int_0^t \int_{I^2} \langle \psi(x,s)\,,\, J(x)\rangle \left(\hat{W}_n(x,y) - W(x,y)\right) \mathrm{d}x\mathrm{d}y\mathrm{d}s\right|, \\ \leqslant \left|\int_0^t \left(\int_I \left|\langle \psi(x,s)\,,\, J(x)\rangle\right|^2 \mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{I^2} \left(\hat{W}_n(x,y) - W(x,y)\right)^2 \mathrm{d}y\mathrm{d}x\right)^{\frac{1}{2}} \mathrm{d}s\right|,$$

which goes to 0 as $n \to \infty$, by (2.36) and since $\psi \in \mathcal{C}([0,T], L^k(I))$. Thirdly,

$$\left|\int_0^t \int_{I^2} \langle (\psi_n(y,s) - \psi(y,s))\,,\, J(x)\rangle\, \hat{W}_n(x,y)\mathrm{d}x\mathrm{d}y\mathrm{d}s\right| \\ \leqslant \left\|J\right\|_\infty \left|\int_0^t \left(\int_I \left|\psi_n(y,s) - \psi(y,s)\right|^2 \mathrm{d}y\right)^{\frac{1}{2}} \left(\int_{I^2} \hat{W}_n(x,y)^2 \mathrm{d}x\mathrm{d}y\right)^{\frac{1}{2}} \mathrm{d}s\right|,$$

which goes to 0 as $n \to \infty$. Finally,

$$\left|\int_0^t \int_{I^2} \left(\hat{W}_n(x,y) - W(x,y)\right) \langle \psi(y,s)\, ,\, J(x)\rangle \,\mathrm{d}x\mathrm{d}y\mathrm{d}s\right|,$$
$$\leqslant \|J\|_\infty \int_0^t \left(\int_I |\psi(y,s)|^2 \mathrm{d}y\right)^{\frac{1}{2}} \left(\int_{I^2} \left|\hat{W}_n(x,y) - W(x,y)\right|^2 \mathrm{d}x\mathrm{d}y\right)^{\frac{1}{2}} \mathrm{d}s,$$

which also goes to 0 as $n \to \infty$. These estimates altogether give the convergence of the interaction term in (5.17). Putting everything together, we obtain that any accumulation point of $(\theta_n)_{n \geqslant 1}$ in $\mathcal{C}([0,T], L^k(I))$ is a weak solution to (2.34). □

5.3. Identification in the compact case

We prove Theorem 2.19. Let $x \mapsto J(x)$ be a regular ($\mathcal{C}^1$) test function on $[0,1]$. Then,

$$\int_I \langle \theta_n(x,t)\, ,\, J(x)\rangle \,\mathrm{d}x = \frac{1}{n}\sum_{i=1}^n \left\langle \theta_{i,t}^{(n)}\, ,\, \bar{J}\left(\frac{i}{n}\right)\right\rangle, \tag{5.18}$$

where $\bar{J}\left(\frac{i}{n}\right) := n\int_{\frac{i-1}{n}}^{\frac{i}{n}} J(x)\mathrm{d}x$. The expression $\bar{J}\left(\frac{i}{n}\right)$ is not an actual function of $\frac{i}{n}$, but one can replace $\bar{J}\left(\frac{i}{n}\right)$ by $J\left(\frac{i}{n}\right)$:

$$\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^n \left\langle \theta_{i,t}^{(n)}\, ,\, \bar{J}\left(\frac{i}{n}\right) - J\left(\frac{i}{n}\right)\right\rangle\right]^2 \leqslant \frac{1}{n}\sum_{i=1}^n \mathbf{E}\left|\theta_{i,t}^{(n)}\right|^2 \left|\bar{J}\left(\frac{i}{n}\right) - J\left(\frac{i}{n}\right)\right|^2,$$
$$\leqslant \frac{C}{n}\sum_{i=1}^n \left|\bar{J}\left(\frac{i}{n}\right) - J\left(\frac{i}{n}\right)\right|^2 \leqslant \frac{C\,\|J'\|_\infty^2}{4n^2},$$

using (2.14). So the limit of $\int_I \langle \theta_n(x,t)\, ,\, J(x)\rangle \,\mathrm{d}x$ as $n \to \infty$ is the same as

$$U_{n,t}(J) := \frac{1}{n}\sum_{i=1}^n \left\langle \theta_{i,t}^{(n)}\, ,\, J\left(\frac{i}{n}\right)\right\rangle = \int \langle \theta\, ,\, J(x)\rangle\, \nu_{n,t}(\mathrm{d}\theta, \mathrm{d}x). \tag{5.19}$$

Taking the limit as $n \to \infty$ in (5.18), using Theorem 2.15, one obtains that, for all $t \in [0,T]$,

$$\int_I \langle \psi(x,t)\, ,\, J(x)\rangle \,\mathrm{d}x = \int_I \int \langle \theta\, ,\, J(x)\rangle\, \nu_t^x(\mathrm{d}\theta)\mathrm{d}x. \tag{5.20}$$

Theorem 2.19 follows.

Ackowledgements

I would like to thank the referees for their careful reading and useful comments on the paper. This is supported by ANR-19-CE40-0024 (CHAllenges in MAthematical NEuroscience) and ANR–19–CE40–0023 (PERISTOCH). This research has been conducted within the FP2M federation (CNRS FR 2036).

A. Well-posedness and regularity results for the nonlinear Fokker-Planck PDE

The aim of this section is to prove Proposition 2.7, as well as some regularity estimates concerning the solution ν to (1.4).

A.1. Existence of a solution to the nonlinear Fokker-Planck PDE

We prove here the existence part of Proposition 2.7 and the result of Remark 2.8. Recall the definition of the set $\mathcal{M}$ in Section 2.2. For any $m \in \mathcal{M}$, consider the solution to

$$\mathrm{d}\theta_t^x = c(\theta_t^x)\mathrm{d}t + \int \Gamma(\theta_t^x, \tilde{\theta})W(x,y)m_t(\mathrm{d}\tilde{\theta}, \mathrm{d}y)\mathrm{d}t + \sigma \mathrm{d}B_t, \tag{A.1}$$

with initial condition $\theta_0^x \sim \nu_0^x$. Consider $\Theta : \mathcal{M} \to \mathcal{M}$ the functional which maps any measure $m(\mathrm{d}\theta, \mathrm{d}x)$ to the law $\Theta(m)$ of (θ^x, x), where θ^x solves (A.1). The point is to prove that Θ admits a fixed-point in $\mathcal{M}$, which gives the existence of a solution ν to (1.4). Recall the definition of the Wasserstein metric in (2.12) and the definition of k in (2.7). Then, for two driving measures m_1 and m_2 in $\mathcal{M}$, for the coupling (θ_1, θ_2) with same initial conditions and Brownian noise, using the properties on c and Γ, we get

$$\begin{aligned}
|\theta_{1,t} - \theta_{2,t}|^2 \leqslant C &\int_0^t |\theta_{1,s} - \theta_{2,s}|^2 \,\mathrm{d}s \\
&+ C \int_0^t |\theta_{1,s} - \theta_{2,s}| \int \left|\Gamma(\theta_{1,s}, \tilde{\theta}) - \Gamma(\theta_{2,s}, \tilde{\theta})\right| W(x,y) m_{1,s}(\mathrm{d}\tilde{\theta}, \mathrm{d}y)\mathrm{d}s \\
&+ C \int_0^t |\theta_{1,s} - \theta_{2,s}| \left| \int \left\{\Gamma(\theta_{2,s}, u) - \Gamma(\theta_{2,s}, v)\right\} p_s^y(\mathrm{d}u, \mathrm{d}v) W(x,y)\ell(\mathrm{d}y) \right| \mathrm{d}s,
\end{aligned}$$

where $p^y(\mathrm{d}u, \mathrm{d}v)$ is *any* coupling of m_1^y and m_2^y. Hence,

$$\begin{aligned}
|\theta_{1,t} - \theta_{2,t}|^2 \leqslant C \left(1 + \|\mathcal{W}_1\|_\infty\right) &\int_0^t |\theta_{1,s} - \theta_{2,s}|^2 \,\mathrm{d}s \\
&+ C \int_0^t |\theta_{1,s} - \theta_{2,s}| \int |u - v| \, p_s^y(\mathrm{d}u, \mathrm{d}v) W(x,y)\ell(\mathrm{d}y)\mathrm{d}s.
\end{aligned}$$

Elevating everything to the power $k \geqslant 1$, using Jensen's inequality and taking the supremum in $s \leqslant t$ and the expectation, one obtains, for another constant $C = C(T)$

$$\begin{aligned}
\mathbf{E}\left[\sup_{s \leqslant t} |\theta_{1,s} - \theta_{2,s}|^{2k}\right] \leqslant C \left(1 + \|\mathcal{W}_1\|_\infty^k\right) &\int_0^t \mathbf{E}\left[\sup_{u \leqslant s} |\theta_{1,s} - \theta_{2,s}|^{2k}\right] \mathrm{d}s \\
&+ C \int_0^t \mathbf{E}\left[\sup_{u \leqslant s} |\theta_{1,u} - \theta_{2,u}|^k\right] \left\{ \int |u - v| \, p_s^y(\mathrm{d}u, \mathrm{d}v) W(x,y)\ell(\mathrm{d}y) \right\}^k \mathrm{d}s.
\end{aligned}$$

Using $\int_0^t u(s)v(s)\mathrm{d}s \leqslant \frac{1}{2}\int_0^t u(s)^2\mathrm{d}s + \frac{1}{2}\int_0^t v(s)^2\mathrm{d}s$ and Jensen's inequality for the last term, we obtain, for another constant C,

$$\begin{aligned}
\mathbf{E}\left[\sup_{s \leqslant t} |\theta_{1,s} - \theta_{2,s}|^{2k}\right] \leqslant C \left(1 + \|\mathcal{W}_1\|_\infty^k\right) &\int_0^t \mathbf{E}\left[\sup_{u \leqslant s} |\theta_{1,s} - \theta_{2,s}|^{2k}\right] \mathrm{d}s \\
&+ C \int_0^t \left\{ \sup_{u \leqslant s} \int |u - v| \, p_u^y(\mathrm{d}u, \mathrm{d}v) W(x,y)\ell(\mathrm{d}y) \right\}^{2k} \mathrm{d}s.
\end{aligned}$$

Applying Jensen's inequality to the probability measure (for fixed $x \in I$) $\frac{W(x,y)}{\int W(x,z)\ell(\mathrm{d}z)}\ell(\mathrm{d}y)$ (recall (2.3)), we can bound the last term above by

$$C\left\|\mathcal{W}_1\right\|_\infty^{2k-1}\int_0^t \sup_{u\leqslant s}\int |u-v|^{2k}\,p_u^y(\mathrm{d}u,\mathrm{d}v)W(x,y)\ell(\mathrm{d}y)\mathrm{d}s.$$

Since this is true for all coupling, we obtain

$$\mathbf{E}\left[\sup_{s\leqslant t}|\theta_{1,s}-\theta_{2,s}|^{2k}\right]\leqslant C\left(1+\|\mathcal{W}_1\|_\infty\right)^k\int_0^t \mathbf{E}\left[\sup_{u\leqslant s}|\theta_{1,s}-\theta_{2,s}|^{2k}\right]\mathrm{d}s+C\left\|\mathcal{W}_1\right\|_\infty^{2k}\int_0^t \delta_s(m_1,m_2)^{2k}\mathrm{d}s.$$

Hence, by Grönwall's Lemma, we obtain finally, for constant C depending on T and $\|\mathcal{W}_1\|_\infty$,

$$\delta_t(\Theta(m_1),\Theta(m_2))^{2k}\leqslant C\int_0^t \delta_s(m_1,m_2)^{2k}\mathrm{d}s.$$

Iterating this estimate gives, for all $l\geqslant 1$, $\delta_T(\Theta^{l+1}(\nu_0),\Theta^l(\nu_0))^{2k}\leqslant C^l\frac{T^l}{l!}\delta_T(\Theta(\nu_0),\nu_0)^{2k}$, so that $\left(\Theta^l(\nu_0)\right)$ is a Cauchy sequence, and hence, converging to ν, solution to (1.4).

A.2. Uniqueness of a solution to the nonlinear Fokker-Planck PDE

We now turn to the uniqueness part of Proposition 2.7. Let $\mu=\{\mu_t\}_{t\in[0,T]}$ be any other weak solution to (1.5) in $\mathcal{M}$ such that $\nu_0=\mu_0$. The point is to prove that $\mu_t=\nu_t$ for $t\in[0,T]$.

Definition A1. *For $m\in\{\nu,\mu\}$, for ℓ-almost every $x\in I$, for any $0\leqslant s\leqslant t$, any $\theta\in\mathbb{R}^d$, denote by $\left\{\varphi_s^{m,t}(\theta,x)\right\}_{s\leqslant t\leqslant T}$ the unique solution of*

$$\mathrm{d}\vartheta_t^{m,x} \;=\; \left(c(\vartheta_t^{m,x})+\int\Gamma\left(\vartheta_t^{m,x},\tilde{\theta}\right)W(x,y)m_t(\mathrm{d}\tilde{\theta},\mathrm{d}y)\right)\mathrm{d}t+\sigma\mathrm{d}B_t\,, \tag{A.2}$$

with position x and initial condition $\varphi_s^{m,s}=\theta$ at $t=s$. In the case $m=\nu$, define finally, for any test function $f:\theta\mapsto f(\theta)$, $\theta\in\mathbb{R}^d$ and ℓ-almost every $x\in I$,

$$P_{s,t}f(\theta,x):=\mathbf{E}f\left(\varphi_s^{\nu,t}(\theta,x)\right). \tag{A.3}$$

Let us suppose that c is uniformly Lispchitz on $\mathbb{R}^d$ (one can remove this assumption by replacing $c(\cdot)$ by its Yosida approximation, we refer to [38], Section 7, where the same procedure is carried out). Under the assumptions made on the model, the propagator P satisfies the following Backward Kolmogorov equation (see [21], Remark 2.3): for $\theta\in\mathbb{R}^d$ and $x\in I$,

$$\partial_s P_{s,t}f+\frac{1}{2}\nabla_\theta\cdot\left(\sigma\sigma^\dagger\nabla_\theta P_{s,t}f\right)+\left(\left(c(\theta)+\int\Gamma(\theta,\tilde{\theta})W(x,y)\nu_t(\mathrm{d}\tilde{\theta},\mathrm{d}y)\right)\cdot\nabla_\theta\right)P_{s,t}f=0. \tag{A.4}$$

For any regular test function f, applying Ito formula to $t \mapsto P_{t,T}f(\vartheta_t^{\mu,x}, x)$, where $\vartheta^{\mu,x}$ solves (A.2) for the choice of $m = \mu$, gives:

$$P_{s,t}f(\vartheta_s^{\mu,x}, x) = P_{0,t}f(\vartheta_0^{\mu,x}, x) + \int_0^s \partial_v P_{v,t}f(\vartheta_v^{\mu,x}, x)\mathrm{d}v$$
$$+ \int_0^s \nabla_\theta P_{v,t}f(\vartheta_v^{\mu,x}, x) \cdot \left(c(\vartheta_v^{\mu,x}) + \int \Gamma\left(\vartheta_v^{\mu,x}, \tilde{\theta}\right) W(x,y)\mu_v(\mathrm{d}\tilde{\theta}, \mathrm{d}y) \right) \mathrm{d}v$$
$$+ \frac{1}{2}\int_0^s \nabla_\theta \cdot \left(\sigma\sigma^\dagger \nabla_\theta P_{v,t}f\right)(\vartheta_v^{\mu,x}, x)\mathrm{d}v + \int_0^s \nabla_\theta P_{v,t}f(\vartheta_v^{\mu,x}, x) \cdot \sigma \mathrm{d}B_v. \quad \text{(A.5)}$$

Using (A.4), this simplifies into

$$P_{s,t}f(\vartheta_t^{\mu,x}, x) = P_{0,t}f(\vartheta_0^{\mu,x}, x) + \int_0^s \nabla_\theta P_{v,t}f(\vartheta_v^{\mu,x}, x) \cdot \sigma \mathrm{d}B_v$$
$$+ \int_0^s \nabla_\theta P_{v,t}f(\vartheta_v^{\mu,x}, x) \cdot \int \Gamma\left(\vartheta_v^{\mu,x}, \tilde{\theta}\right) W(x,y)(\mu_v(\mathrm{d}\tilde{\theta}, \mathrm{d}y) - \nu_v(\mathrm{d}\tilde{\theta}, \mathrm{d}y))\mathrm{d}v. \quad \text{(A.6)}$$

Since the law of $\vartheta_v^{\mu,x}$ is $\mu_v^x(\mathrm{d}\theta)$, taking the expectation w.r.t. the Brownian motion, we obtain for $s = t$ (recall that $P_{s,s}f = f$),

$$\int f(\theta)\mu_s^x(\mathrm{d}\theta) = \int P_{0,s}f(\theta, x)\mu_0^x(\mathrm{d}\theta)$$
$$+ \int_0^s \int \left\{ \int \nabla_\theta P_{v,s}f(\theta, x) \cdot \Gamma\left(\theta, \tilde{\theta}\right) W(x,y)\mu_v^x(\mathrm{d}\theta) \right\} (\mu_v(\mathrm{d}\tilde{\theta}, \mathrm{d}y) - \nu_v(\mathrm{d}\tilde{\theta}, \mathrm{d}y))\mathrm{d}v. \quad \text{(A.7)}$$

Furthermore, taking in (1.4) test functions of the form $\varphi(\theta, x) = \phi(\theta)\psi(x)$ shows that for every regular test function $\theta \mapsto \phi(\theta)$, for ℓ-almost every x,

$$\int \phi(\theta)\nu_s^x(\mathrm{d}\theta) = \int \phi(\theta)\nu_0^x(\mathrm{d}\theta) + \int_0^s \int \left\{ \frac{1}{2}\nabla_\theta\left(\sigma\sigma^\dagger \nabla_\theta \phi(\theta)\right) + \nabla_\theta\phi(\theta) \cdot c(\theta) \right\} \nu_v^x(\mathrm{d}\theta)\mathrm{d}v$$
$$+ \int_0^s \int \nabla_\theta \phi(\theta) \cdot \int W(x,y)\Gamma(\theta, \tilde{\theta})\nu_v(\mathrm{d}\tilde{\theta}, \mathrm{d}y)\nu_v^x(\mathrm{d}\theta)\mathrm{d}v. \quad \text{(A.8)}$$

From this, we get

$$\partial_v \int P_{v,t}f(\theta, x)\nu_v^x(\mathrm{d}\theta) = \int \partial_v P_{v,t}f(\theta, x)\nu_v^x(\mathrm{d}\theta) + \int P_{v,t}f(\theta, x)\partial_v \nu_v^x(\mathrm{d}\theta) = 0. \quad \text{(A.9)}$$

We obtain finally, for ℓ-almost every x,

$$\int f(\theta)\left\{\mu_s^x(\mathrm{d}\theta) - \nu_s^x(\mathrm{d}\theta)\right\} = \int P_{0,s}f(\theta, x)\left\{\mu_0^x(\mathrm{d}\theta) - \nu_0^x(\mathrm{d}\theta)\right\}$$
$$+ \int_0^s \int \left\{ \int \nabla_\theta P_{v,s}f(\theta, x) \cdot \Gamma\left(\theta, \tilde{\theta}\right) W(x,y)\mu_v^x(\mathrm{d}\theta) \right\} (\mu_v^y(\mathrm{d}\tilde{\theta}) - \nu_v^y(\mathrm{d}\tilde{\theta}))\ell(\mathrm{d}y)\mathrm{d}v. \quad \text{(A.10)}$$

Let us recall the definition of the Wasserstein metric $w_1(\cdot, \cdot)$ in (2.8). By the Kantorovich-Rubinstein duality, an equivalent expression of this distance is

$$w_1(\mu, \nu) = \sup_{\|f\|_{Lip} \leqslant 1} \left| \int f\mathrm{d}\mu - \int f\mathrm{d}\nu \right|. \quad \text{(A.11)}$$

An important point is to note that there exists a constant $C > 0$ such that, uniformly in (θ, x), $|\nabla_\theta P_{v,s} f(\theta, x)| \leqslant C$, where C is uniform on $v, s \in [0, T]$ and f such that $\|f\|_{Lip} \leqslant 1$ (see [38], Lemma 4.4 for more details). Thus, for fixed $x, y \in I$ and $v, s \in [0, T]$, the function $\tilde{\theta} \mapsto I_{x,y}(\tilde{\theta}) := \int \nabla_\theta P_{v,s} f(\theta, x) \cdot \Gamma\left(\theta, \tilde{\theta}\right) W(x, y) \mu_v^x(\mathrm{d}\theta)$ is Lipschitz:

$$
\begin{aligned}
|I_{x,y}(\theta_1) - I_{x,y}(\theta_2)| &= \left| \int \nabla_\theta P_{v,s} f(\theta, x) \cdot \left(\Gamma(\theta, \theta_1) - \Gamma(\theta, \theta_2)\right) W(x, y) \mu_v^x(\mathrm{d}\theta) \right| \\
&\leqslant C L_\Gamma W(x, y) \left|\theta_1 - \theta_2\right|.
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
\left| \int f(\theta) \left\{\mu_s^x(\mathrm{d}\theta) - \nu_s^x(\mathrm{d}\theta)\right\} \right| &\leqslant \left| \int P_{0,s} f(\theta, x) \left\{\mu_0^x(\mathrm{d}\theta) - \nu_0^x(\mathrm{d}\theta)\right\} \right| \\
&\quad + C L_\Gamma \int_0^s \int \sup_{u \leqslant v} w_1(\mu_u^y, \nu_u^y) W(x, y) \ell(\mathrm{d}y) \mathrm{d}v. \qquad \text{(A.12)}
\end{aligned}
$$

Taking the supremum on f with $\|f\|_{Lip} \leqslant 1$ and using the fact that $\mu_0 = \nu_0$, we obtain

$$
w_1\left(\mu_s^x, \nu_s^x\right) \leqslant C L_\Gamma \int_0^s \int \sup_{u \leqslant v} w_1(\mu_u^y, \nu_u^y) W(x, y) \ell(\mathrm{d}y) \mathrm{d}v.
$$

By Cauchy-Schwarz and Jensen inequalities,

$$
\begin{aligned}
w_1\left(\mu_s^x, \nu_s^x\right)^2 &\leqslant C^2 L_\Gamma^2 \left(\int_0^s \left(\int \sup_{u \leqslant v} w_1(\mu_u^y, \nu_u^y)^2 \ell(\mathrm{d}y) \right)^{\frac{1}{2}} \left(\int W(x, y)^2 \ell(\mathrm{d}y) \right)^{\frac{1}{2}} \mathrm{d}v \right)^2 \\
&\leqslant C^2 L_\Gamma^2 T \left\|\mathcal{W}_2\right\|_\infty \int_0^s \int \sup_{u \leqslant v} w_1(\mu_u^y, \nu_u^y)^2 \ell(\mathrm{d}y) \mathrm{d}v.
\end{aligned}
$$

Taking the supremum in $s \leqslant t$ and integrating w.r.t. x gives

$$
\int \sup_{s \leqslant t} w_1\left(\mu_s^x, \nu_s^x\right)^2 \ell(\mathrm{d}x) \leqslant C^2 L_\Gamma^2 T \left\|\mathcal{W}_2\right\|_\infty \int_0^t \int \sup_{u \leqslant v} w_1(\mu_u^y, \nu_u^y)^2 \ell(\mathrm{d}y) \mathrm{d}v
$$

so that Grönwall's Lemma gives uniqueness.

A.3. A priori estimates and spatial regularity

We now gather some estimates concerning the solution ν to (1.4). Recall Definition A1: in the following, for $\theta \in \mathbb{R}^d$, $x \in I$, we set $\varphi_s^t(\theta, x) := \varphi_s^{\nu,t}(\theta, x)$ (now $m = \nu$).

Lemma A2. *For all $\theta_1, \theta_2 \in \mathbb{R}^d$, $x, y \in I$, $t \leqslant T$, consider $\varphi_0^t(\theta_1, x)$ and $\varphi_0^t(\theta_2, y)$ coupled in such a way that they are driven by the same Brownian motion B in* (A.2). *Under the Assumptions 2.1, 2.4 and 2.6, there exists a constant C_1 depending only on $\Gamma, \sigma, c, W, \nu_0$ such that for all $t \geqslant 0$,*

$$
\mathbf{E}\left[\sup_{u \in [0,t]} \left|\varphi_0^u(\theta_1, x) - \varphi_0^u(\theta_2, y)\right|^2 \right] \leqslant C_1 e^{C_1 t} \left(\delta\mathcal{W}(x, y)^2 + |\theta_1 - \theta_2|^2 \right), \qquad \text{(A.13)}
$$

where we recall the definition of $\delta\mathcal{W}$ in (2.28).

Proof of Lemma A2. Set for simplicity $\Phi_u := \varphi_0^u(\theta_1, x)$ and $\Psi_u := \varphi_0^u(\theta_2, y)$. Then, for all $s \leqslant t$,

$$\begin{aligned}|\Phi_s - \Psi_s|^2 &= |\theta_1 - \theta_2|^2 + 2\int_0^s \langle \Phi_u - \Psi_u\,,\, c(\Phi_u) - c(\Psi_u)\rangle \,\mathrm{d}u \\ &+ 2\int_0^s \left\langle \Phi_u - \Psi_u\,,\, \int \left(W(x,z)\Gamma(\Phi_u, \tilde{\theta}) - W(y,z)\Gamma(\Psi_u, \tilde{\theta})\right)\nu_u(\mathrm{d}\tilde{\theta}, \mathrm{d}z)\right\rangle \mathrm{d}u. \quad \text{(A.14)}\end{aligned}$$

Since c is one-sided Lipschitz (2.6), taking the supremum in $s \in [0, t]$

$$\begin{aligned}\sup_{s\in[0,t]} |\Phi_s - \Psi_s|^2 \leqslant\; &|\theta_1 - \theta_2|^2 + (2L_c + 1)\int_0^t \sup_{v\in[0,u]} |\Phi_v - \Psi_v|^2 \,\mathrm{d}u \\ &+ \int_0^t \left|\int \left(W(x,z)\Gamma(\Phi_u, \tilde{\theta}) - W(y,z)\Gamma(\Psi_u, \tilde{\theta})\right)\nu_u(\mathrm{d}\tilde{\theta}, \mathrm{d}z)\right|^2 \mathrm{d}u. \quad \text{(A.15)}\end{aligned}$$

The last term within the last integral can be estimated above by $2(a_u^{(1)} + a_u^{(2)})$ where

$$a_u^{(1)} = \left|\int W(x,z)\left(\Gamma(\Phi_u, \tilde{\theta}) - \Gamma(\Psi_u, \tilde{\theta})\right)\nu_u(\mathrm{d}\tilde{\theta}, \mathrm{d}z)\right|^2 \leqslant L_\Gamma^2 \|\mathcal{W}\|_\infty^2 \sup_{v\in[0,u]} |\Phi_v - \Psi_v|^2\,,$$

and where

$$\begin{aligned}a_u^{(2)} &= \left(\int \left\{\int \Gamma(\Psi_u, \tilde{\theta})\nu_u^z(\mathrm{d}\tilde{\theta})\right\} |W(x,z) - W(y,z)|\, \ell(\mathrm{d}z)\right)^2 \\ &\leqslant L_\Gamma^2 \left(1 + |\Psi_u|^2 + \sup_{z\in I} \mathbf{E}\left[\sup_{v\,\leqslant\, u} \left|\bar{\theta}_v^z\right|^2\right]\right) \delta\mathcal{W}(x,y)^2.\end{aligned}$$

Taking the expectation, using (2.13) and Grönwall's Lemma, one obtains (A.13). □

Lemma A3. *Let $f : \theta \mapsto f(\theta)$ be such that $|f(\theta) - f(\theta')| \leqslant c_f^{(1)} |\theta - \theta'|$ and $|f(\theta)| \leqslant c_f^{(2)} (1 + |\theta|)$. Then, Under the Assumptions 2.1, 2.4 and 2.6, for the same constant C_1 as in* (A.13), *for all $t \geqslant 0$, for $x, y \in [0, 1]$,*

$$\left|\int f(\theta)\nu_t^x(\mathrm{d}\theta) - \int f(\theta)\nu_t^y(\mathrm{d}\theta)\right| \leqslant c_f^{(1)} C_1 e^{C_1 t} \left(\delta\mathcal{W}(x,y) + w_1(\nu_0^x, \nu_0^y)\right). \quad \text{(A.16)}$$

Proof of Lemma A3. Recall Definition A1 (in particular the definition of P in (A.3)) and the calculations made in the proof of uniqueness in Section A.2. Apply (A.10) to the case $\mu = \nu$: for any regular test function f satisfying the hypotheses of Lemma A3,

$$\int f(\theta)\nu_t^x(\mathrm{d}\theta) = \int P_{0,t} f(\theta, x)\nu_0^x(\mathrm{d}\theta). \quad \text{(A.17)}$$

In particular, for $x, y \in I$,

$$\begin{aligned}\left|\int f(\theta)\nu_t^x(\mathrm{d}\theta) - \int f(\theta)\nu_t^y(\mathrm{d}\theta)\right| &\leqslant \int |P_{0,t} f(\theta, x) - P_{0,t} f(\theta, y)|\, \nu_0^x(\mathrm{d}\theta) \quad \text{(A.18)} \\ &+ \left|\int P_{0,t} f(\theta, y)\nu_0^x(\mathrm{d}\theta) - \int P_{0,t} f(\theta, y)\nu_0^y(\mathrm{d}\theta)\right|. \quad \text{(A.19)}\end{aligned}$$

For the first term (A.18), we have

$$\int |P_{0,t}f(\theta,x) - P_{0,t}f(\theta,y)|\, \nu_0^x(\mathrm{d}\theta) = \int \left|\mathbf{E}\left[f(\varphi_0^t(\theta,x))\right] - \mathbf{E}\left[f(\varphi_0^t(\theta,y))\right]\right| \nu_0^x(\mathrm{d}\theta),$$
$$\leqslant c_f^{(1)} \int \mathbf{E}\left|\varphi_0^t(\theta,x) - \varphi_0^t(\theta,y)\right| \nu_0^x(\mathrm{d}\theta), \tag{A.20}$$

where the coupling $(\varphi_0^t(\theta,x), \varphi_0^t(\theta,y))$ is given by Lemma A2. By (A.13), one obtains that

$$\int |P_{0,t}f(\theta,x) - P_{0,t}f(\theta,y)|\, \nu_0^x(\mathrm{d}\theta) \leqslant c_f^{(1)} C_1 e^{C_1 t} \delta\mathcal{W}(x,y). \tag{A.21}$$

As far as the second term (A.19) is concerned, for any fixed (θ_1,θ_2), applying once again Lemma A2 gives $|P_{0,t}f(\theta_1,y) - P_{0,t}f(\theta_2,y)| \leqslant c_f^{(1)} C_1 e^{C_1 t} |\theta_1 - \theta_2|$. In particular, for any coupling $(\bar\theta_0^x, \bar\theta_0^y)$,

$$\left|\int P_{0,t}f(\theta,y)\nu_0^x(\mathrm{d}\theta) - \int P_{0,t}f(\theta,y)\nu_0^y(\mathrm{d}\theta)\right| = \left|\mathbf{E}\left[P_{0,t}f(\bar\theta_0^x,y) - P_{0,t}f(\bar\theta_0^y,y)\right]\right|,$$
$$\leqslant \mathbf{E}\left|P_{0,t}f(\bar\theta_0^x,y) - P_{0,t}f(\bar\theta_0^y,y)\right| \leqslant c_f^{(1)} C_1 e^{C_1 t} \mathbf{E}\left|\bar\theta_0^x - \bar\theta_0^y\right|.$$

Since this true for all coupling of the initial conditions $(\bar\theta_0^x, \bar\theta_0^y)$, one obtains finally that

$$\left|\int P_{0,t}f(\theta,y)\nu_0^x(\mathrm{d}\theta) - \int P_{0,t}f(\theta,y)\nu_0^y(\mathrm{d}\theta)\right| \leqslant c_f^{(1)} C_1 e^{C_1 t} w_1(\nu_0^x, \nu_0^y) \tag{A.22}$$

This concludes the proof of Lemma A3. □

Lemma A4. *Recall the definition of* $[\Gamma]_u$ *in* (2.21) *and of* Υ_u *in* (2.22) *(*$u \geqslant 0$*). Under the Assumptions 2.1, 2.4 and 2.6, there exists a constant* $C > 0$ *(that depends only on* T*,* c*,* Γ *and* σ*) such that for all* $u \in [0,T]$*,* $\theta \in \mathbb{R}^d$*,* $x, y, y', z, z' \in I$*,*

$$\sup_{u\in[0,T]} \sup_{x\in I} |[\Gamma]_u(\theta,x)| \leqslant C(1+|\theta|), \tag{A.23}$$

$$\sup_{u\in[0,T]} \sup_{x,y,z\in I} |\Upsilon_u(x,y,z)| \leqslant C \tag{A.24}$$

as well as

$$\sup_{u\in[0,T]} |[\Gamma]_u(\theta,x) - [\Gamma]_u(\theta,y)| \leqslant C\left(\delta\mathcal{W}(x,y) + w_1(\nu_0^x,\nu_0^y)\right), \tag{A.25}$$

$$\sup_{u\in[0,T]} \sup_{x,z\in I} \left|\Upsilon_u(x,y,z) - \Upsilon_u(x,y',z)\right| \leqslant C\left(\delta\mathcal{W}(y,y') + w_1(\nu_0^y,\nu_0^{y'})\right), \tag{A.26}$$

$$\sup_{u\in[0,T]} \sup_{x,y\in I} \left|\Upsilon_u(x,y,z) - \Upsilon_u(x,y,z')\right| \leqslant C\left(\delta\mathcal{W}(z,z') + w_1(\nu_0^z,\nu_0^{z'})\right). \tag{A.27}$$

Proof of Lemma A4. The estimates on Υ are an easy consequence of the estimates on $[\Gamma]$. The bound (A.23) on $[\Gamma]$ is a direct consequence of (2.5) and the uniform estimates we have on $\bar\theta$ (2.13). Let us now prove (A.25): apply the results of Lemma A3 to the test function $f := \tilde\theta \mapsto \Gamma(\theta,\tilde\theta)$ (for fixed $\theta \in \mathbb{R}^d$). The test function f satisfies the hypothesis of Lemma A3 for $c_f^{(1)} := L_\Gamma$ (and some $c_f^{(2)}$ that depends on θ, but note that $c_f^{(2)}$ does not enter into account in the estimates of Lemma A3). In particular, uniformly in θ,

$$|[\Gamma]_u(\theta, x) - [\Gamma]_u(\theta, y)| = \left| \int \Gamma(\theta, \tilde{\theta})\nu_u^x(\mathrm{d}\tilde{\theta}) - \int \Gamma(\theta, \tilde{\theta})\nu_u^y(\mathrm{d}\tilde{\theta}) \right| \leqslant C \left(\delta\mathcal{W}(x, y) + w_1(\nu_0^x, \nu_0^y)\right),$$

which gives the result. $\square$

B. Identification in the general case: proof of Theorem 2.20

The point of this section is to prove Theorem 2.20. Recall that $I = \mathbb{R}^p$ endowed with a probability measure $\ell(\mathrm{d}x) := \ell(x)\mathrm{d}x$ with a $\mathcal{C}^1$ density ℓ. We fix a $\mathcal{C}^1$-kernel $W(x, y)$ on $\mathbb{R}^p \times \mathbb{R}^p$. We proceed by truncation from the compact case (Theorem 2.19): fix $M > 0$, define

$$B_M := [-M, M]^p \text{ and } \Lambda_M := B_M \times B_M, \tag{B.1}$$

and introduce the following probability measure, whose support is B_M:

$$\ell^{(M)}(\mathrm{d}x) := \ell^{(M)}(x)\mathrm{d}x := \ell(x)\frac{\mathbf{1}_{B_M}(x)}{\ell(B_M)}\mathrm{d}x. \tag{B.2}$$

In what follows, we choose M sufficiently large so that

$$\ell(B_M) \geqslant \frac{1}{2}. \tag{B.3}$$

What has been done in Section 5.3 for $I = [0, 1]$ with deterministic regular positions can be transposed without difficulties (up to obvious notational changes) to $I = B_M$, endowed with its renormalized Lebesgue measure $\frac{\mathrm{d}x}{|B_M|}$, where $|B_M| = (2M)^p$. Hence, we can apply the result of Section 5.3 for $I = B_M$ and deterministic positions, for the choice of kernel on B_M^2

$$\tilde{W}_M(x, y) := W(x, y)\ell^{(M)}(y) \left|B_M\right|, \; x, y \in B_M. \tag{B.4}$$

Indeed, the kernel $(x, y) \mapsto \tilde{W}_M(x, y)$ is bounded and $\mathcal{C}^1$ on B_M^2 so that $\tilde{W}_M$ satisfies the assumptions of Section 2.1. Moreover, we see from Section 3.3 that $\tilde{W}_M$ can be realized as the macroscopic limit of a graph $(\mathcal{G}^{(n)}, \kappa^{(n)})$ constructed as in (3.7) and (3.9). Since W and ℓ are regular, Assumptions 2.9 and 2.11 are satisfied. Hence, both solutions $\nu^{(M)}$ to (1.4) (with initial condition $\nu_0^{(M)}(\mathrm{d}\theta, \mathrm{d}x) = \nu_0^x(\mathrm{d}\theta)\ell^{(M)}(\mathrm{d}x)$) and $\psi^{(M)}$ to (2.34) (with initial condition $\psi_0^{(M)}(x) = \int \theta\nu_0^{x,(M)}(\mathrm{d}\theta)$) in the case $I = B_M$ endowed with $\frac{\mathrm{d}x}{|B_M|}$, for the kernel $\tilde{W}_M$ are well posed and satisfy the identification

$$\int_{B_M} \left\langle \psi^{(M)}(x, t) \, , \, J(x) \right\rangle \frac{\mathrm{d}x}{|B_M|} = \int_{B_M} \int \left\langle \theta \, , \, J(x) \right\rangle \nu_t^{(M),x}(\mathrm{d}\theta) \frac{\mathrm{d}x}{|B_M|}, \tag{B.5}$$

for all regular test functions J on B_M. The point of the remaining is to make $M \to \infty$ in (B.5). We treat the two sides of (B.5) separately. Concerning the lefthand, $\psi^{(M)}$ is the unique weak solution to

$$\int_{B_M} \left\langle \psi^{(M)}(x,t)\,,\, J(x) \right\rangle \frac{\mathrm{d}x}{|B_M|} = \int_{B_M} \left\langle \psi_0^{(M)}(x)\,,\, J(x) \right\rangle \frac{\mathrm{d}x}{|B_M|} + \int_0^t \int_{B_M} \left\langle c(\psi^{(M)}(x,s))\,,\, J(x) \right\rangle \frac{\mathrm{d}x}{|B_M|}\mathrm{d}s + \int_0^t \int_{B_M^2} \left\langle \Gamma(\psi^{(M)}(x,s), \psi^{(M)}(y,s))\,,\, J(x) \right\rangle \tilde{W}_M(x,y) \frac{\mathrm{d}y}{|B_M|}\frac{\mathrm{d}x}{|B_M|}\mathrm{d}s. \quad \text{(B.6)}$$

Multiplying everything by $|B_M|$ and choosing test functions of the form $J(x)\ell^{(M)}(x)$ gives, by definition of $\tilde{W}_M$,

$$\int_{\mathbb{R}^p} \frac{\mathrm{d}}{\mathrm{d}t} \left\langle \psi^{(M)}(x,t)\,,\, J(x) \right\rangle \ell^{(M)}(\mathrm{d}x) = \int_{\mathbb{R}^p} \left\langle c(\psi^{(M)}(x,t))\,,\, J(x) \right\rangle \ell^{(M)}(\mathrm{d}x) + \int_{\mathbb{R}^p}\int_{\mathbb{R}^p} \left\langle \Gamma(\psi^{(M)}(x,t), \psi^{(M)}(y,t))\,,\, J(x) \right\rangle W(x,y)\ell^{(M)}(\mathrm{d}y)\ell^{(M)}(\mathrm{d}x). \quad \text{(B.7)}$$

We first give some a priori bound on $\psi^{(M)}$: by density, (B.7) is also true for all test functions $J(x,t)$ and for $J(x,t) = \left|\psi^{(M)}(x,t)\right|^{2k-2}\psi^{(M)}(x,t)$, we obtain, using the properties on c and Γ

$$\begin{aligned}
\frac{1}{2k}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^p} \left|\psi^{(M)}(x,t)\right|^{2k} \ell^{(M)}(\mathrm{d}x) &= \int_{\mathbb{R}^p} \left\langle c(\psi^{(M)}(x,t))\,,\, \psi^{(M)}(x,t) \right\rangle \left|\psi^{(M)}(x,t)\right|^{2k-2} \ell^{(M)}(\mathrm{d}x) \\
&+ \int_{\mathbb{R}^{2p}} \left\langle \Gamma(\psi^{(M)}(x,t), \psi^{(M)}(y,t))\,,\, \psi^{(M)}(x,t) \right\rangle \left|\psi^{(M)}(x,t)\right|^{2k-2} W(x,y)\ell^{(M)}(\mathrm{d}y)\ell^{(M)}(\mathrm{d}x), \\
&\leqslant \left(L_c + \frac{1}{2}\right)\int_{\mathbb{R}^p} \left|\psi^{(M)}(x,t))\right|^{2k} \ell^{(M)}(\mathrm{d}x) + \frac{|c(0)|^2}{2}\int_{\mathbb{R}^p} \left|\psi^{(M)}(x,t))\right|^{2k-2} \ell^{(M)}(\mathrm{d}x) \\
&+ L_\Gamma \int_{\mathbb{R}^{2p}} \left(\left|\psi^{(M)}(x,t)\right|^{2k-1} + \left|\psi^{(M)}(x,t)\right|^{2k}\right) W(x,y)\ell^{(M)}(\mathrm{d}y)\ell^{(M)}(\mathrm{d}x) \\
&+ L_\Gamma \int_{\mathbb{R}^{2p}} \left|\psi^{(M)}(y,t)\right| \left|\psi^{(M)}(x,t)\right|^{2k-1} W(x,y)\ell^{(M)}(\mathrm{d}y)\ell^{(M)}(\mathrm{d}x). \quad \text{(B.8)}
\end{aligned}$$

The last term in (B.8) is bounded by

$$\begin{aligned}
&\int_{\mathbb{R}^p} \left|\psi^{(M)}(x,t)\right|^{2k-1} \left(\int_{\mathbb{R}^p} \left|\psi^{(M)}(y,t)\right|^2 \ell^{(M)}(\mathrm{d}y)\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^p} W(x,y)^2 \ell^{(M)}(\mathrm{d}y)\right)^{\frac{1}{2}} \ell^{(M)}(\mathrm{d}x), \\
&\leqslant \sqrt{2}\,\|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \left(\int_{\mathbb{R}^p} \left|\psi^{(M)}(y,t)\right|^2 \ell^{(M)}(\mathrm{d}y)\right)^{\frac{1}{2}} \int_{\mathbb{R}^p} \left|\psi^{(M)}(x,t)\right|^{2k-1} \ell^{(M)}(\mathrm{d}x),
\end{aligned}$$

where we used (B.3) and (2.2). Applying Young's inequality for $p = 2k$ and $q = p^* = \frac{2k}{2k-1}$, the last quantity is smaller than

$$\begin{aligned}
&\sqrt{2}\,\|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \left\{ \frac{1}{2k}\left(\int_{\mathbb{R}^p}\left|\psi^{(M)}(y,t)\right|^2 \ell^{(M)}(\mathrm{d}y)\right)^k + \frac{2k-1}{2k}\left(\int_{\mathbb{R}^p}\left|\psi^{(M)}(x,t)\right|^{2k-1}\ell^{(M)}(\mathrm{d}x)\right)^{\frac{2k}{2k-1}}\right\},\\
&\leqslant \sqrt{2}\,\|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \left\{ \frac{1}{2k}\int_{\mathbb{R}^p}\left|\psi^{(M)}(y,t)\right|^{2k}\ell^{(M)}(\mathrm{d}y) + \frac{2k-1}{2k}\int_{\mathbb{R}^p}\left|\psi^{(M)}(x,t)\right|^{2k}\ell^{(M)}(\mathrm{d}x)\right\},\\
&= \sqrt{2}\,\|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \int_{\mathbb{R}^p}\left|\psi^{(M)}(y,t)\right|^{2k}\ell^{(M)}(\mathrm{d}y),
\end{aligned}$$

by Jensen's inequality. Using this bound in (B.8), the fact that there are constants $c_0, c_1 > 0$ such that $|x|^{2k-2} + |x|^{2k-1} + |x|^{2k} \leqslant c_0 + c_1 |x|^{2k}$ and Grönwall's Lemma gives

$$\sup_{t\in[0,T]}\sup_{M\,\geqslant\,1}\int_{\mathbb{R}^p}\left|\psi^{(M)}(x,t)\right|^{2k}\ell^{(M)}(\mathrm{d}x) < +\infty. \tag{B.9}$$

We now prove that $(\psi^{(M)}\mathbf{1}_{B_M})$ is Cauchy: for $N > M$, set

$$\rho(x,t) = \rho_{N,M}(x,t) := \psi^{(N)}(x,t)\mathbf{1}_{B_N}(x) - \psi^{(M)}(x,t)\mathbf{1}_{B_M}(x). \tag{B.10}$$

Multiplying by $\ell(B_M)$ in (B.7), we have,

$$\begin{aligned}
\int_{\mathbb{R}^p}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \psi^{(M)}(x,t)\mathbf{1}_{B_M}(x)\,,\,J(x,t)\right\rangle \ell(\mathrm{d}x) &= \int_{\mathbb{R}^p}\left\langle c(\psi^{(M)}(x,t))\mathbf{1}_{B_M}(x)\,,\,J(x,t)\right\rangle \ell(\mathrm{d}x)\\
&+\frac{1}{\ell(B_M)}\int_{\mathbb{R}^{2p}}\left\langle \Gamma(\psi^{(M)}(x,t),\psi^{(M)}(y,t))\,,\,J(x,t)\right\rangle W(x,y)\mathbf{1}_{\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x).
\end{aligned}$$

Use the notation $\tilde{c}(\theta) = c(\theta) - c(0)$. Note that $\tilde{c}$ also satisfies (2.6) for the same constant L_c. Since $\tilde{c}(0) = 0$, one has $\tilde{c}(\theta)\mathbf{1}_{B_M}(x) = \tilde{c}(\theta\mathbf{1}_{B_M}(x))$. In a same way, for all θ_1, θ_2, $\Gamma(\theta_1,\theta_2)\mathbf{1}_{\Lambda_M}(x,y) = \Gamma(\theta_1\mathbf{1}_{B_M}(x),\theta_2\mathbf{1}_{B_M}(y))\mathbf{1}_{\Lambda_M}(x,y)$. Hence, we can write

$$\begin{aligned}
\int_{\mathbb{R}^p}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \psi^{(M)}(x,t)\mathbf{1}_{B_M}(x)\,,\,J(x,t)\right\rangle \ell(\mathrm{d}x) &= \int_{\mathbb{R}^p}\left\langle \tilde{c}(\psi^{(M)}(x,t)\mathbf{1}_{B_M}(x))\,,\,J(x,t)\right\rangle \ell(\mathrm{d}x)\\
&+\int_{\mathbb{R}^p}\left\langle c(0)\,,\,J(x,t)\right\rangle \mathbf{1}_{B_M}(x)\ell(\mathrm{d}x)\\
+\frac{1}{\ell(B_M)}\int_{\mathbb{R}^{2p}}\Big\langle \Gamma(\psi^{(M)}(x,t)\mathbf{1}_{B_M}(x),\psi^{(M)}(y,t)&\mathbf{1}_{B_M}(y))\,,\,J(x,t)\Big\rangle W(x,y)\mathbf{1}_{\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x).
\end{aligned}$$

Taking the difference between $\psi^{(M)}$ and $\psi^{(N)}$ gives for $J(x,t) = \rho(x,t)\,|\rho(x,t)|^{k-2}$:

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^p} |\rho(x,t)|^k\,\ell(\mathrm{d}x) = \int_{\mathbb{R}^p} \Big\langle \tilde{c}(\psi^{(N)}(x,t)\mathbf{1}_{B_N}(x)) - \tilde{c}(\psi^{(M)}(x,t)\mathbf{1}_{B_M}(x))\,,\ J(x,t)\Big\rangle\,\ell(\mathrm{d}x)\\
&+\int_{\mathbb{R}^p} \langle c(0)\,,\ J(x,t)\rangle\,\mathbf{1}_{B_N\setminus B_M}(x)\ell(\mathrm{d}x)\\
&+\frac{1}{\ell(B_M)}\int_{\mathbb{R}^{2p}} \Big\langle \Gamma(\psi^{(N)}(x,t)\mathbf{1}_{B_N}(x), \psi^{(N)}(y,t)\mathbf{1}_{B_N}(y))\,,\ J(x,t)\Big\rangle\, W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x)\\
&+\left(\frac{1}{\ell(B_N)} - \frac{1}{\ell(B_M)}\right)\int_{\mathbb{R}^{2p}} \Big\langle \Gamma(\psi^{(N)}(x,t)\mathbf{1}_{B_N}(x), \psi^{(N)}(y,t)\mathbf{1}_{B_N}(y))\,,\ J(x,t)\Big\rangle\, W(x,y)\mathbf{1}_{\Lambda_N}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x)\\
&+\frac{1}{\ell(B_M)}\int_{\mathbb{R}^{2p}} \Big\langle \Gamma(\psi^{(N)}(x,t)\mathbf{1}_{B_N}(x), \psi^{(N)}(y,t)\mathbf{1}_{B_N}(y))\\
&\qquad\qquad - \Gamma(\psi^{(M)}(x,t)\mathbf{1}_{B_M}(x), \psi^{(M)}(y,t)\mathbf{1}_{B_M}(y))\,,\ J(x,t)\Big\rangle W(x,y)\mathbf{1}_{\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x)\\
&:= (I) + (II) + (III) + (IV) + (V).
\end{aligned}
\tag{B.11}
$$

Concerning the first term in (B.11), we have, by the property of c, $|(I)| \leqslant L_c \int_{\mathbb{R}^p} |\rho(x,t)|^k\,\ell(\mathrm{d}x)$. The second term is controlled as

$$
|(II)| \leqslant |c(0)| \left(\int_{\mathbb{R}^p} |\rho(x,t)|^k\,\ell(\mathrm{d}x)\right)^{1-\frac{1}{k}} \ell(B_N\setminus B_M)^{\frac{1}{k}} \leqslant |c(0)| \left\{\frac{k-1}{k}\int_{\mathbb{R}^p} |\rho(x,t)|^k\,\ell(\mathrm{d}x) + \frac{1}{k}\ell(B_N\setminus B_M)\right\}.
$$

We now turn to (III): by (B.3) and (2.5),

$$
\begin{aligned}
|(III)| \leqslant\ & 2L_\Gamma \int_{\mathbb{R}^{2p}} |\rho(x,t)|^{k-1}\, W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x)\\
&+ 2L_\Gamma \int_{\mathbb{R}^{2p}} \left|\psi^{(N)}(x,t)\mathbf{1}_{B_N}(x)\right| |\rho(x,t)|^{k-1}\, W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}x)\ell(\mathrm{d}y)\\
&+ 2L_\Gamma \int_{\mathbb{R}^{2p}} \left|\psi^{(N)}(y,t)\mathbf{1}_{B_N}(y)\right| |\rho(x,t)|^{k-1}\, W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x)\\
&:= (i) + (ii) + (iii).
\end{aligned}
\tag{B.12}
$$

The term (i) is bounded by

$$
\begin{aligned}
(i) \leqslant 2L_\Gamma &\left(\int_{\mathbb{R}^p} |\rho(x,t)|^k\,\ell(\mathrm{d}x)\right)^{1-\frac{1}{k}} \left(\int_{\mathbb{R}^p}\left(\int_{\mathbb{R}^p} W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\right)^k \ell(\mathrm{d}x)\right)^{\frac{1}{k}}\\
&\leqslant \frac{2(k-1)L_\Gamma}{k}\int_{\mathbb{R}^p} |\rho(x,t)|^k\,\ell(\mathrm{d}x) + \frac{2L_\Gamma}{k}\,\|\mathcal{W}_2\|_\infty^{\frac{k}{2}}\,\ell(\Lambda_N\setminus\Lambda_M)^{\frac{k}{2}}.
\end{aligned}
$$

Secondly,

$$
\begin{aligned}
(ii) \leqslant\ & 2^{k-1} L_\Gamma \int_{B_N} \left|\psi^{(N)}(x,t)\right|^k \left(\int_{\mathbb{R}^p} W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\right)\ell(\mathrm{d}x) \\
& + 2^{k-1} L_\Gamma \int_{B_M} \left|\psi^{(N)}(x,t)\right| \left|\psi^{(M)}(x,t)\right|^{k-1} \left(\int_{\mathbb{R}^p} W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\right)\ell(\mathrm{d}x), \\
& \leqslant 2^{k-1} L_\Gamma \left(\int_{B_N} \left|\psi^{(N)}(x,t)\right|^{2k} \ell(\mathrm{d}x)\right)^{\frac{1}{2}} \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \ell(\Lambda_N\setminus\Lambda_M)^{\frac{1}{2}} \\
& + 2^{k-1} L_\Gamma \left(\int_{B_N} \left|\psi^{(N)}(x,t)\right|^{qr} \ell(\mathrm{d}x)\right)^{\frac{1}{qr}} \left(\int_{B_M} \left|\psi^{(M)}(x,t)\right|^{qr^*(k-1)} \ell(\mathrm{d}x)\right)^{\frac{1}{qr^*}} \\
& \left(\int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^p} W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\right)^{q^*} \ell(\mathrm{d}x)\right)^{\frac{1}{q^*}},
\end{aligned}
$$

by two successive applications of Hölder's inequality for the last term. Choosing $q = q^* = 2$ and $(r, r^*) = (k, \frac{k}{k-1})$, using (B.9), we see that the term (ii) is finally bounded by $C\,\|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \ell(\Lambda_N\setminus\Lambda_M)^{\frac{1}{2}}$. Finally, the last term in (B.12) is controlled as

$$
\begin{aligned}
(iii) \leqslant\ & 2^{k-1} L_\Gamma \int_{B_N^2} \left|\psi^{(N)}(y,t)\right| \left|\psi^{(N)}(x,t)\right|^{k-1} W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x) \\
& + 2^{k-1} L_\Gamma \int_{B_M}\int_{B_N} \left|\psi^{(N)}(y,t)\right| \left|\psi^{(M)}(x,t)\right|^{k-1} W(x,y)\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x), \\
& := (a) + (b).
\end{aligned}
$$

By Hölder's inequality, the term (a) is controlled as

$$
\begin{aligned}
(a) \leqslant\ & 2^{k-1} L_\Gamma \left(\int_{B_N} \left|\psi^{(N)}(y,t)\right|^{q} \ell(\mathrm{d}y)\right)^{\frac{1}{q}} \left(\int_{B_N} \left|\psi^{(N)}(x,t)\right|^{r(k-1)} \ell(\mathrm{d}x)\right)^{\frac{1}{r}} \\
& \left(\int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^p} W(x,y)^{q^*}\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\right)^{\frac{r^*}{q^*}} \ell(\mathrm{d}x)\right)^{\frac{1}{r^*}}.
\end{aligned}
$$

For the choice $(r, r^*) = (\frac{2k}{k-1}, \frac{2k}{k+1})$, $(q, q^*) = (\frac{2-\delta}{1-\delta}, 2-\delta)$ for $\delta \in (0,1)$,

$$
\begin{aligned}
(a) \leqslant\ & 2^{k-1} L_\Gamma \left(\int_{B_N} \left|\psi^{(N)}(y,t)\right|^{\frac{2-\delta}{1-\delta}} \ell(\mathrm{d}y)\right)^{\frac{1-\delta}{2-\delta}} \left(\int_{B_N} \left|\psi^{(N)}(x,t)\right|^{2k} \ell(\mathrm{d}x)\right)^{\frac{k-1}{2k}} \\
& \left(\int_{\mathbb{R}^p} \left(\int_{\mathbb{R}^p} W(x,y)^{2-\delta}\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)\ell(\mathrm{d}y)\right)^{\frac{2k}{(2-\delta)(k+1)}} \ell(\mathrm{d}x)\right)^{\frac{k+1}{2k}}.
\end{aligned}
$$

Choosing δ sufficiently small so that $\frac{2-\delta}{1-\delta} < 2k$ (possible since $k > 1$), we obtain

$$
(a) \leqslant 2^{k-1} L_\Gamma \sup_{N\geqslant 1} \sup_{t\in[0,T]} \left(\int_{B_N} \left|\psi^{(N)}(y,t)\right|^{2k} \ell(\mathrm{d}y)\right)^{\frac{1}{2}} \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \ell(\Lambda_N\setminus\Lambda_M)^{\frac{\delta}{2(2-\delta)}}.
$$

The same calculation gives the same estimate for (b). Concerning the term (IV), we have obviously $\left|\frac{1}{\ell(B_N)} - \frac{1}{\ell(B_M)}\right| \leqslant 4\ell(B_N \setminus B_M)$ so that it only suffices to have a uniform bound on the integral term in (IV). This is indeed true, by the same calculations as for (III) (with $\mathbf{1}_{\Lambda_N\setminus\Lambda_M}(x,y)$ replaced by $\mathbf{1}_{\Lambda_N}(x,y) \leqslant 1$). It remains to control (V): by (2.4),

$$
\begin{aligned}
|(V)| \leqslant 2L_\Gamma \int_{(\mathbb{R}^p)^2} |\rho(x,t)|^k W(x,y)\mathbf{1}_{\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x) \\
+ 2L_\Gamma \int_{(\mathbb{R}^p)^2} |\rho(y,t)|\, |\rho(x,t)|^{k-1} W(x,y)\mathbf{1}_{\Lambda_M}(x,y)\ell(\mathrm{d}y)\ell(\mathrm{d}x).
\end{aligned}
$$

The first term in the sum above si easily bounded by $2L_\Gamma \|\mathcal{W}_1\|_\infty \int_{\mathbb{R}^p} |\rho(x,t)|^k \ell(\mathrm{d}x)$. The second is controlled by

$$
\begin{aligned}
2L_\Gamma \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \left(\int_{\mathbb{R}^p} |\rho(y,t)|^2 \ell(\mathrm{d}y)\right)^{\frac{1}{2}} \int_{\mathbb{R}^p} |\rho(x,t)|^{k-1} \ell(\mathrm{d}x) \\
\leqslant 2L_\Gamma \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \left\{ \frac{1}{k} \left(\int_{\mathbb{R}^p} |\rho(y,t)|^2 \ell(\mathrm{d}y)\right)^{\frac{k}{2}} + \frac{k-1}{k} \left(\int_{\mathbb{R}^p} |\rho(x,t)|^{k-1} \ell(\mathrm{d}x)\right)^{\frac{k}{k-1}} \right\} \\
\leqslant 2L_\Gamma \|\mathcal{W}_2\|_\infty^{\frac{1}{2}} \int_{\mathbb{R}^p} |\rho(x,t)|^k \ell(\mathrm{d}x),
\end{aligned}
$$

by Jensen's inequality (since $k \geqslant 2$).

Gathering all these estimates into (B.11), by a Grönwall's Lemma and the fact that $\ell(\Lambda_N \setminus \Lambda_M) \leqslant \ell(\Lambda_M^c) \xrightarrow[M\to\infty]{} 0$, we see that the sequence $\left(\psi^{(M)}\mathbf{1}_{B_M}\right)_{M \geqslant 1}$ is Cauchy in $\mathcal{C}([0,T], L^k(I,\ell))$ and hence converges to some $\psi(\cdot,t)$. By the same argument as before, it is easy to show that ψ is a weak solution (and hence, by Proposition 2.17 the only solution) to (2.34). Moreover, the convergence in $\mathcal{C}([0,T], L^k(I,\ell))$ implies that $\int_{\mathbb{R}^p} \left\langle \psi^{(M)}(x,t)\,,\, J(x)\right\rangle \mathrm{d}x \xrightarrow[M\to\infty]{} \int_{\mathbb{R}^p} \langle \psi(x,t)\,,\, J(x)\rangle \,\mathrm{d}x$ for all $t \in [0,T]$ and every test functions J with compact support.

We now turn to the righthand side of (B.5): we apply the same propagator technique as in Section A.2 to the process $\nu^{(M)}$ in (B.5). By the same procedure as in Section A.2, it is possible to prove that, for all test functions $\theta \mapsto f(\theta)$, for $\ell^{(M)}$-almost every x,

$$
\begin{aligned}
&\int f(\theta) \left\{\nu_s^{(M),x}(\mathrm{d}\theta) - \nu_s^x(\mathrm{d}\theta)\right\} = \int P_{0,s} f(\theta,x) \left\{\nu_0^{(M),x}(\mathrm{d}\theta) - \nu_0^x(\mathrm{d}\theta)\right\} \\
&+ \int_0^s \int_{\mathbb{R}^d\times\mathbb{R}^p} \left\{\int_{\mathbb{R}^d} \nabla_\theta P_{v,t} f(\theta,x) \cdot \Gamma\left(\theta,\tilde{\theta}\right) \nu_v^{(M),x}(\mathrm{d}\theta)\right\} W(x,y) \left\{\nu_v^{(M),y}(\mathrm{d}\tilde{\theta}) - \nu_v^y(\mathrm{d}\tilde{\theta})\right\} \ell^{(M)}(\mathrm{d}y)\mathrm{d}v \\
&+ \int_0^s \int_{\mathbb{R}^d\times\mathbb{R}^p} \left\{\int_{\mathbb{R}^d} \nabla_\theta P_{v,t} f(\theta,x) \cdot \Gamma\left(\theta,\tilde{\theta}\right) \nu_v^{(M),x}(\mathrm{d}\theta)\right\} W(x,y)\nu_v^y(\mathrm{d}\tilde{\theta}) \left\{\frac{\mathbf{1}_{B_M}(x)}{\ell(B_M)} - 1\right\} \ell(\mathrm{d}y)\mathrm{d}v \\
&:= (A) + (B). \quad \text{(B.13)}
\end{aligned}
$$

Concerning the first term: it is bounded by

$$|(A)| \leqslant \frac{CL_\Gamma}{\ell(B_M)} \int_0^s \int \sup_{u \leqslant v} w_1 \left(\nu_u^{(M),y}, \nu_u^y\right) W(x,y)\ell(\mathrm{d}y)\mathrm{d}v$$
$$\leqslant \frac{CL_\Gamma \left\|\mathcal{W}_2\right\|_\infty^{\frac{1}{2}}}{\ell(B_M)} \int_0^s \left(\int \sup_{u \leqslant v} w_1 \left(\nu_u^{(M),y}, \nu_u^y\right)^2 \ell(\mathrm{d}y)\right)^{\frac{1}{2}} \mathrm{d}v.$$

Concerning the second term, using the fact that $\sup_{s \leqslant T, M \geqslant 1, y \in I} \int |\theta| \, \nu_s^{M,y}(\mathrm{d}\theta) < \infty$, for some constant $C > 0$

$$|(B)| \leqslant C \int_0^s \left|\frac{\mathbf{1}_{B_M}(x)}{\ell(B_M)} - 1\right| \int\int \left(1 + |\theta| + \left|\tilde{\theta}\right|\right) \nu_v^{(M),x}(\mathrm{d}\theta)\nu_v^y(\mathrm{d}\tilde{\theta})W(x,y)\ell(\mathrm{d}y)\mathrm{d}v$$
$$\leqslant C \left\|\mathcal{W}_1\right\|_\infty \left|\frac{\mathbf{1}_{B_M}(x)}{\ell(B_M)} - 1\right| s.$$

These estimates and Grönwall Lemma gives that $\sup_x \sup_{s \in [0,T]} \int f(\theta) \left\{\nu_s^{(M),x}(\mathrm{d}\theta) - \nu_s^x(\mathrm{d}\theta)\right\} \xrightarrow[M \to \infty]{} 0$.

C. Convergence of graphs: proof of Proposition 2.22

The point of this section is to prove Proposition 2.22. To do so, define first two other auxiliary (directed and weighted) graphs:

1. $\mathcal{H}_1^{(n)}$, with vertex set $[n]$: in $\mathcal{H}_1^{(n)}$, for all $i \neq j \in [n]$, both directed edges $i \to j$ and $j \to i$ are present and associated with the respective weights $\kappa_i^{(n)} W_n(x_i^{(n)}, x_j^{(n)})$ and $\kappa_j^{(n)} W_n(x_i^{(n)}, x_j^{(n)})$,
2. $\mathcal{H}_2^{(n)}$, with vertex set $[n]$: in $\mathcal{H}_2^{(n)}$, for all $i \neq j \in [n]$, the edge $i \to j$ (resp. $j \to i$) is present and associated with the weight $W(x_i^{(n)}, x_j^{(n)})$ (resp. $W(x_i^{(n)}, x_j^{(n)})$).

C.1. Some distances and norms on graphs and kernels

Before proving Proposition 2.22, we need to introduce the necessary definitions coming from graph convergence theory (see [36, 9, 6] and references therein). Concerning the notion of cut-off distance considered in Proposition 2.22 and other related definitions, we follow here closely [7, 8]. In particular, we generalize here the definitions of [8], § 2.3 to the case of directed graphs and non-symmetric kernels: let $G = (V(G), E(G))$ be a possibly directed weighted graph, where each vertex $i \in V(G)$ is associated to a weight α_i and each edge $i \to j$ is associated to a weight $\beta_{i,j}$ (where possibly $\beta_{i,j} \neq \beta_{j,i}$). We define $\alpha_G = \sum_{i \in V(G)} \alpha_i$ and the kernel W^G on $[0,1]^2$ in the following way: divide $[0,1]$ into intervals $I_1, \ldots, I_{|V(G)|}$ of length $|I_i| = \frac{\alpha_i}{\alpha_G}$ and define

$$W^G(x,y) = \sum_{i,j \in V(G)} \beta_{i,j} \mathbf{1}_{(x,y) \in I_i \times I_j}. \tag{C.1}$$

For one kernel W define

$$\|W\|_{\infty \to 1} := \sup_{\|f\|_\infty, \|g\|_\infty \leqslant 1} \left|\int_{[0,1]^2} W(x,y)f(x)g(y)\mathrm{d}x\mathrm{d}y\right| \tag{C.2}$$

as well as the cut norm

$$\|W\|_{\square} := \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)\mathrm{d}x\mathrm{d}y \right|. \tag{C.3}$$

Note that the norms $\|\cdot\|_{\infty\to 1}$ and $\|\cdot\|_{\square}$ are equivalent (see [8], Eq. (2.3)): for some $C>0$, for any kernel W,

$$\|W\|_{\square} \leqslant \|W\|_{\infty\to 1} \leqslant C\,\|W\|_{\square}. \tag{C.4}$$

For any weighted directed graphs $\mathcal{G}$ and $\mathcal{G}'$ with vertex set $[n]$, the same nodeweights α_i and with respective weights $(\beta_{i,j})_{i,j\in[n]}$ and $(\beta'_{i,j})_{i,j\in[n]}$, define the cut-off distance

$$d_{\square}(G,G') := \left\| W^G - W^{G'} \right\|_{\square} = \max_{S,T\subset[n]} \left| \sum_{i\in S, j\in T} \frac{\alpha_i\alpha_j}{\alpha_G^2}(\beta_{i,j}-\beta'_{i,j}) \right|. \tag{C.5}$$

Finally, we define the following L^1-distance between two kernels W and W' (not necessarily symmetric):

$$d_1(W,W') := \left\|W-W'\right\|_1 := \int_{[0,1]^2} \left|W(x,y)-W'(x,y)\right| \mathrm{d}x\mathrm{d}y. \tag{C.6}$$

C.2. Proof of Proposition 2.22

Proposition 2.22 is a direct consequence of Propositions C1, C3 and C4 below.

Proposition C1. *Under the hypotheses of Section 2.7, we have*

$$d_{\square}\left(\bar{\mathcal{G}}^{(n)}, \mathcal{H}_1^{(n)}\right) \xrightarrow[n\to\infty]{} 0, \quad \textit{with probability } 1. \tag{C.7}$$

Proof of Proposition C1. Let $S,T\subseteq[n]$: we have

$$\mathbb{P}\left(\left| \frac{1}{n^2} \sum_{i\in S, j\in T} \kappa_i^{(n)} \bar{\xi}_{i,j} \right| > \varepsilon \right) \leqslant \mathbb{P}\left(\left| \frac{1}{n^2} \sum_{i\in S, j\in T, i<j} \kappa_i^{(n)} \bar{\xi}_{i,j} \right| > \frac{\varepsilon}{2} \right) + \mathbb{P}\left(\left| \frac{1}{n^2} \sum_{i\in S, j\in T, i>j} \kappa_i^{(n)} \bar{\xi}_{i,j} \right| > \frac{\varepsilon}{2} \right).$$

Let $U_+ := \left\{(i,j)\in[n]^2,\ i\in S,\ j\in T, i<j\right\}$. We have $|U_+| \leqslant |S|\,|T|$. We use the following result ([17], Th 3.3):

Proposition C2. *Let $X_1,\dots,X_n$ be independent random variables with $X_i\sim\mathcal{B}(p_i)$. For $X=\sum_{i=1}^n a_iX_i$ with $a_i>0$, define $\nu=\sum_{i=1}^n a_i^2p_i$. Then, for $a=\max\{a_1,\dots,a_n\}$,*

$$\mathbb{P}\left(X-\mathbb{E}[X] < -\lambda\right) \leqslant e^{-\frac{\lambda^2}{2\nu}}, \tag{C.8}$$

$$\mathbb{P}\left(X-\mathbb{E}[X] > \lambda\right) \leqslant e^{-\frac{\lambda^2}{2(\nu+a\lambda/3)}}, \tag{C.9}$$

Since, by definition the variables $(\xi_{i,j})_{(i,j)\in U_+}$ are independent (recall that $\bar{\xi}_{i,j}=\xi_{i,j}-W_n(x_i,x_j)$, apply (C.8) with $a_i=\kappa_i^{(n)}$, so that

$$\mathbb{P}\left(\frac{1}{n^2} \sum_{(i,j)\in U_+} \kappa_i^{(n)} \bar{\xi}_{i,j} < -\frac{\varepsilon}{2} \right) \leqslant \exp\left(-\frac{\varepsilon^2 n^4}{8\sum_{(i,j)\in U_+} \left(\kappa_i^{(n)}\right)^2 W_n(x_i,x_j)} \right).$$

Since

$$\frac{1}{n^2\kappa_n}\sum_{(i,j)\in U_+}\left(\kappa_i^{(n)}\right)^2 W_n(x_i,x_j)\leqslant\frac{1}{n^2}\sum_{i,j=1}^{n}\kappa_i^{(n)}W_n(x_i,x_j)\leqslant\delta_n(\underline{x})+\frac{1}{n^2}\sum_{i,j=1}^{n}W(x_i,x_j),$$

using the fact that $\frac{1}{n^2}\sum_{i,j=1}^{n}W(x_i,x_j)\xrightarrow[n\to\infty]{}\int W(x,y)\mathrm{d}x\mathrm{d}y=\|W\|_{L^1}>0$, we have for n sufficiently large,

$$\frac{1}{n^2\kappa_n}\sum_{(i,j)\in U_+}\left(\kappa_i^{(n)}\right)^2 W_n(x_i,x_j)\leqslant 1+\frac{3\,\|W\|_{L^1}}{2}.$$

For such n,

$$\mathbb{P}\left(\frac{1}{n^2}\sum_{(i,j)\in U_+}\kappa_i^{(n)}\bar{\xi}_{i,j}<-\frac{\varepsilon}{2}\right)\leqslant\exp\left(-\frac{\varepsilon^2n^2}{8\kappa_n\left(1+\frac{3\|W\|_{L^1}}{2}\right)}\right).$$

Moreover, by similar arguments, for n sufficiently large and $\varepsilon<6$, we have

$$\mathbb{P}\left(\frac{1}{n^2}\sum_{(i,j)\in U_+}\kappa_i^{(n)}\bar{\xi}_{i,j}>\frac{\varepsilon}{2}\right)=\mathbb{P}\left(\sum_{(i,j)\in U_+}\kappa_i^{(n)}\bar{\xi}_{i,j}>\frac{\varepsilon n^2}{2}\right)\leqslant\exp\left(-\frac{\varepsilon^2n^2}{8\kappa_n\left(2+\frac{3\|W\|_{L^1}}{2}\right)}\right).$$

Putting things together, we obtain

$$\mathbb{P}\left(\left|\frac{1}{n^2}\sum_{(i,j)\in U_+}\kappa_i^{(n)}\bar{\xi}_{i,j}\right|>\frac{\varepsilon}{2}\right)\leqslant 2\exp\left(-\frac{\varepsilon^2n^2}{8\kappa_n\left(2+\frac{3\|W\|_{L^1}}{2}\right)}\right).$$

By the same argument on U_-, we have

$$\mathbb{P}\left(\left|\frac{1}{n^2}\sum_{i\in S,j\in T}\kappa_i^{(n)}\bar{\xi}_{i,j}\right|>\varepsilon\right)\leqslant 4\exp\left(-\frac{\varepsilon^2n^2}{8\kappa_n\left(2+\frac{3\|W\|_{L^1}}{2}\right)}\right).$$

Writing $w_n\kappa_n^2=u_n\frac{n}{\log(n)}$ with $u_n\xrightarrow[n\to\infty]{}0$ (recall (2.18)), we have from $\frac{1}{\kappa_n}\leqslant w_n$ (recall (2.17)) that $\kappa_n\leqslant u_n\frac{n}{\log(n)}$. Hence, for $\varepsilon:=\varepsilon_n=u_n^{\frac{1}{2}}\xrightarrow[n\to\infty]{}0$, a union bound on $S,T\subseteq[n]$ gives:

$$\mathbb{P}\left(\max_{S,T\subseteq[n]}\left|\frac{1}{n^2}\sum_{i\in S,j\in T}\kappa_i^{(n)}\bar{\xi}_{i,j}\right|>\varepsilon_n\right)\leqslant 4^{n+1}\exp\left(-\frac{n\log(n)}{8\left(2+\frac{3\|W\|_{L^1}}{2}\right)}\right).$$

Borel-Cantelli Lemma gives the convergence (C.7). □

Proposition C3. *With the previous definitions, assuming that* (2.19) *holds, we have*

$$d_\square\left(\mathcal{H}_1^{(n)},\mathcal{H}_2^{(n)}\right)\xrightarrow[n\to\infty]{}0. \tag{C.10}$$

Proof of Proposition C3. For the norm (C.2), we have

$$\left\| W^{\mathcal{H}_1^{(n)}} - W^{\mathcal{H}_2^{(n)}} \right\|_{\infty \to 1} = \sup_{\|f\|_\infty, \|g\|_\infty \leqslant 1} \left| \int_{[0,1]^2} \left(W^{\mathcal{H}_1^{(n)}}(x,y) - W^{\mathcal{H}_2^{(n)}}(x,y) \right) f(x) g(y) \mathrm{d}x \mathrm{d}y \right|,$$

$$\leqslant \sum_{k,l=1}^{n} \int_{I_k \times I_l} \left| W^{\mathcal{H}_1^{(n)}}(x,y) - W^{\mathcal{H}_2^{(n)}}(x,y) \right| \mathrm{d}x \mathrm{d}y \leqslant \delta_n(\underline{x}) \xrightarrow[n \to \infty]{} 0.$$

Hence, we obtain from (C.4) and (C.5) the required convergence (C.10). □

Proposition C4. *Under* (2.40)*, we have*

$$d_\square \left(\mathcal{H}_2^{(n)}, W \right) \xrightarrow[n \to \infty]{} 0. \tag{C.11}$$

Proof of Proposition C4. Note that by (C.2), we have that $\|W\|_{\infty \to 1} \leqslant \|W\|_1$, so that, by (C.4), the convergence w.r.t. the distance $d_1(\cdot, \cdot)$ implies the convergence w.r.t. the distance $d_\square(\cdot, \cdot)$. This, it suffices to prove that

$$d_1 \left(\mathcal{H}_2^{(n)}, W \right) \xrightarrow[n \to \infty]{} 0. \tag{C.12}$$

But this is exactly equivalent to (2.40). □

References

[1] D. M. Abrams and S. H. Strogatz. Chimera states in a ring of nonlocally coupled oscillators. *International Journal of Bifurcation and Chaos*, 16(01):21–37, 2006.

[2] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[3] S.-I. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.*, 27(2):77–87, June 1977.

[4] L. Bertini, G. Giacomin, and C. Poquet. Synchronization and random longtime dynamics for mean-field plane rotators. *Probability Theory and Related Fields*, 160(3):593–653, Dec 2014.

[5] S. Bhamidi, A. Budhiraja, and R. Wu. Weakly interacting particle systems on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 129(6):2174 – 2206, 2019.

[6] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi. Limits of randomly grown graph sequences. *European J. Combin.*, 32(7):985–999, 2011.

[7] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence. *Ann. Probab.*, 46(1):337–396, 01 2018.

[8] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062, May 2019.

[9] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

[10] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)*, 176(1):151–219, 2012.

[11] M. Bossy, O. Faugeras, and D. Talay. Clarification and complement to "Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons". *J. Math. Neurosci.*, 5:Art. 19, 23, 2015.

[12] P. C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *J. Phys. A*, 45(3):033001, 109, 2012.

[13] P. C. Bressloff. *Waves in neural media*. Lecture Notes on Mathematical Modelling in the Life Sciences. Springer, New York, 2014. From single neurons to neural fields.

[14] T. Cabana and J. Touboul. Large deviations for randomly connected neural networks: I. spatially extended systems. *arXiv:1510.06957*, 2015.

[15] J. Chevallier, A. Duarte, E. Löcherbach, and G. Ost. Mean-field limits for nonlinear spatially extended hawkes processes with exponential memory kernels. *Stochastic Processes and their Applications*, 129(1):1 – 27, 2019.

[16] H. Chiba, , and G. S. M. and. The mean-field analysis of the Kuramoto model on graphs I. The mean-field equation and transition point formulas. *Discrete & Continuous Dynamical Systems - A*, 39(1):131–155, 2019.

[17] F. Chung and L. Lu. Concentration inequalities and martingale inequalities: a survey. *Internet Math.*, 3(1):79–127, 2006.

[18] F. Coppini. Longtime dynamics for interacting oscillators on dense graphs. *arXiv e-prints*, page arXiv:1908.01520, Aug 2019.

[19] F. Coppini, H. Dietert, and G. Giacomin. A Law of Large Numbers and Large Deviations for interacting diffusions on Erdős-Rényi graphs. *ArXiv e-prints*, July 2018.

[20] J. Crevat, G. Faye, and F. Filbet. Rigorous Derivation of the Nonlocal Reaction-Diffusion Fitzhugh–Nagumo System. *SIAM Journal on Mathematical Analysis*, 51(1):346–373, 2019.

[21] G. Da Prato and L. Tubaro. Some remarks about backward Itô formula and applications. *Stochastic Analysis and Applications*, 16(6):993–1003, 1998.

[22] S. Delattre, G. Giacomin, and E. Luçon. A Note on Dynamical Models on Random Graphs and Fokker–Planck Equations. *J. Stat. Phys.*, 165(4):785–798, 2016.

[23] J.-F. Delmas, J.-S. Dhersin, and M. Sciauveau. Asymptotic for the cumulative distribution function of the degrees and homomorphism densities for random graphs sampled from a graphon. *ArXiv e-prints:1807.09989*, July 2018.

[24] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.

[25] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[26] O. Faugeras and J. Inglis. Stochastic neural field equations: a rigorous footing. *J. Math. Biol.*, 71(2):259–300, 2015.

[27] G. Faye and Z. Kilpatrick. Threshold of front propagation in neural fields: An interface dynamics approach. *SIAM Journal on Applied Mathematics*, 78(5):2575–2596, 2018.

[28] G. Giacomin, K. Pakdaman, X. Pellegrin, and C. Poquet. Transitions in active rotator systems: Invariant hyperbolic manifold approach. *SIAM Journal on Mathematical Analysis*, 44(6):4165–4194, 2012.

[29] S. Gupta, A. Campa, and S. Ruffo. Kuramoto model of synchronization: equilibrium and nonequilibrium aspects. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(8):R08001, 2014.

[30] S. Gupta, M. Potters, and S. Ruffo. One-dimensional lattice of oscillators coupled through power-law interactions: Continuum limit and dynamics of spatial Fourier modes. *Phys. Rev. E*, 85:066201, Jun 2012.

[31] J. Inglis and J. MacLaurin. A general framework for stochastic traveling waves and patterns, with application to neural field equations. *SIAM Journal on Applied Dynamical Systems*, 15(1):195–234, 2016.

[32] D. Kaliuzhnyi-Verbovetskyi and G. Medvedev. The semilinear heat equation on sparse random graphs. *SIAM Journal on Mathematical Analysis*, 49(2):1333–1355, 2017.

[33] D. Kaliuzhnyi-Verbovetskyi and G. Medvedev. The mean-field equation for the kuramoto model on graph sequences with non-lipschitz limit. *SIAM Journal on Mathematical Analysis*, 50(3):2441–2465, 2018.

[34] J. Krüger and W. Stannat. Front propagation in stochastic neural fields: A rigorous mathematical framework. *SIAM Journal on Applied Dynamical Systems*, 13(3):1293–1310, 2014.

[35] E. Lang. A multiscale analysis of traveling waves in stochastic neural fields. *SIAM Journal on Applied Dynamical Systems*, 15(3):1581–1614, 2016.

[36] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933 – 957, 2006.

[37] E. Luçon and C. Poquet. Longtime dynamics and disorder-induced traveling waves in the stochastic Kuramoto model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1196–1240, 2017.

[38] E. Luçon and W. Stannat. Mean-field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.*, 24(5):1946–1993, 2014.

[39] E. Luçon and W. Stannat. Transition from Gaussian to non-Gaussian fluctuations for mean-field diffusions in spatial interaction. *Ann. Appl. Probab.*, 26(6):3840–3909, 2016.

[40] E. Luçon and C. Poquet. Emergence of Oscillatory Behaviors for Excitable Systems with Noise and Mean-Field Interaction: A Slow-Fast Dynamics Approach. *Communications in Mathematical Physics*, 373(3):907–969, Dec. 2019.

[41] J. MacLaurin, J. Salhi, and S. Toumi. Mean-field dynamics of a Wilson–Cowan neuronal network with nonlinear coupling term. *Stochastics and Dynamics*, page 1850046, 2017.

[42] G. S. Medvedev. The nonlinear heat equation on dense graphs and graph limits. *SIAM J. Math. Anal.*, 46(4):2743–2766, 2014.

[43] G. S. Medvedev. The nonlinear heat equation on W-random graphs. *Arch. Ration. Mech. Anal.*, 212(3):781–803, 2014.

[44] S. Mehri, M. Scheutzow, W. Stannat, and B. Z. Zangeneh. Propagation of Chaos for Stochastic Spatially Structured Neuronal Networks with Delay driven by Jump Diffusions. *arXiv e-prints:1805.01654, to appear in Ann. Appl. Prob.*, May 2018.

[45] I. Mitoma. Tightness of probabilities on $C([0,1];\mathcal{S}')$ and $D([0,1];\mathcal{S}')$. *Ann. Probab.*, 11(4):989–999, 1983.

[46] P. E. Müller. Path large deviations for interacting diffusions with local mean-field interactions in random environment. *Electron. J. Probab.*, 22:56 pp., 2017.

[47] H. Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic Theories and the Boltzmann Equation*, pages 60–110. Springer Berlin Heidelberg, 1984.

[48] R. I. Oliveira and G. Reis. Interacting diffusions on random graphs with diverging degrees: hydrodynamics and large deviations. *arXiv e-prints*, page arXiv:1807.06898, July 2018.

[49] I. Omelchenko, Y. Maistrenko, P. Hövel, and E. Schöll. Loss of coherence in dynamical networks: Spatial chaos and chimera states. *Phys. Rev. Lett.*, 106:234102, Jun 2011.

[50] J. Rankin, D. Avitabile, J. Baladron, G. Faye, and D. Lloyd. Continuation of localized coherent structures in nonlocal neural field equations. *SIAM Journal on Scientific Computing*, 36(1):B70–B93, 2014.

[51] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[52] J. Touboul. Limits and dynamics of stochastic neuronal networks with random heterogeneous delays. *J. Stat. Phys.*, 149(4):569–597, 2012.

[53] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[54] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24, 01 1972.