MEAN FIELD LIMIT FOR DISORDERED DIFFUSIONS WITH SINGULAR INTERACTIONS

By Eric Luçon*,†,‡ and Wilhelm Stannat*,†,‡

Technische Universität, Berlin† *and Bernstein Center for Computational Neuroscience*‡

Motivated by considerations from neuroscience (macroscopic behavior of large ensembles of interacting neurons), we consider a population of mean field interacting diffusions in $\mathbf{R}^m$ in the presence of a random environment and with spatial extension: each diffusion is attached to one site of the lattice $\mathbf{Z}^d$ and the interaction between two diffusions is attenuated by a spatial weight that depends on their positions. For a general class of singular weights (including the case already considered in the physical literature when interactions obey to a power-law of parameter $0 < \alpha < d$), we address the convergence as $N \to \infty$ of the empirical measure of the diffusions to the solution of a deterministic McKean-Vlasov equation and prove well-posedness of this equation, even in the degenerate case without noise. We provide also precise estimates of the speed of this convergence, in terms of an appropriate weighted Wasserstein distance, exhibiting in particular nontrivial fluctuations in the power-law case when $\frac{d}{2} \leqslant \alpha < d$. Our framework covers the case of polynomially bounded monotone dynamics that are especially encountered in the main models of neural oscillators.

1. Introduction. The purpose of this paper is to provide a general convergence result for the empirical distribution of spatially extended networks of mean field coupled diffusions in a random environment. The main novelty of the paper is to consider a family of interacting diffusions indexed by the box $\Lambda_N := [\![-N, \ldots, N]\!]^d$ of volume $|\Lambda_N| := (2N+1)^d$ in the d-dimensional lattice $\mathbf{Z}^d$ $(d \geqslant 1)$ where the interaction between two diffusions in Λ_N depends on their relative positions. We are in particular interested in diffusions modeling the spiking activity of neurons in a noisy environment. To motivate the mathematical model we want to work with, let us consider, as a particular example, a family of stochastic FitzHugh-Nagumo neurons (see [2, 14] and references therein for further neurophysiological insights on the model)

$$(1.1) \quad \begin{cases} \mathrm{d}V_i(t) &= \left(V_i(t) - \frac{V_i(t)^3}{3} - w_i(t) + I\right) \mathrm{d}t + \sigma_V \, \mathrm{d}B_i^V(t), \\ \mathrm{d}w_i(t) &= \left(a_i(b_i V_i(t) - w_i(t))\right) \mathrm{d}t + \sigma_w \, \mathrm{d}B_i^w(t), \end{cases}$$

for $i \in \Lambda_N$, with exterior input current I. The variable $V_i(t)$ denotes the voltage activity of the neuron and $w_i(t)$ plays the role of a recovery variable. $\left(B_i^V(t), B_i^w(t)\right)$ are independent Brownian motions modeling exterior stochastic forces. Depending on the parameters $(a_i, b_i) \in \mathbf{R}^2$, the neurons exhibit an oscillatory, excitable or inhibitory behavior. Suppose

that the precise values of $\omega_i = (a_i, b_i)$ are unknown, which will always be the case in a real-world applications, but rather are given as independent and identically distributed random variables. From a point of view from statistical physics, this additional randomness in (1.1) may be considered as a *disorder*. For simplicity we suppose that the ω_i are independent of the time t. Equation (1.1) can be written as

$$(1.2) \qquad \mathrm{d}\theta_i(t) = c(\theta_i, \omega_i)\,\mathrm{d}t + \sigma \cdot \mathrm{d}B_i(t), \quad t \geqslant 0, \; i \in \Lambda_N,$$

using the shorthand notations $\theta = (V, w)$, $\omega = (a, b)$, $c(\theta, \omega) = \left(V - \frac{V^3}{3} - w + I, a(bV - w) \right)$, $B = \left(B^V, B^w \right)$ and $\sigma = \begin{pmatrix} \sigma_V & 0 \\ 0 & \sigma_w \end{pmatrix}$. We suppose that the individual neurons are coupled with the help of a possibly nonlinear and random coupling term $\Gamma\left(\theta_i, \omega_i, \theta_j, \omega_j\right)$, $(i, j \in \Lambda_N)$ modeling electrical synapses between the neurons. The coupling intensity between neurons i and j will depend in addition on some weight $\Psi_N(i, j)$ (Ψ_N may be thought as a function of the distance, but not necessarily), so that the resulting system gets the following type:

$$(1.3) \quad \mathrm{d}\theta_i(t) = c(\theta_i(t), \omega_i)\,\mathrm{d}t$$
$$+ \frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} \Gamma\left(\theta_i(t), \omega_i, \theta_j(t), \omega_j\right) \Psi_N(i, j)\,\mathrm{d}t + \sigma \cdot \mathrm{d}B_i(t), \; t \geqslant 0, \; i \in \Lambda_N.$$

The purpose of the paper is to address the behavior of the system (1.3) in large populations $(N \to \infty)$, under general assumptions on the dynamics c, the coupling Γ and the spatial constraint Ψ_N.

1.1. *Empirical measure and mean-field limit.* All the statistical information of the neural ensemble is contained in its empirical distribution of the diffusions θ_j (with disorder ω_j and with renormalized position $x_j := \frac{1}{2N} \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$):

$$(1.4) \qquad \nu_t^{(N)}(\,\mathrm{d}\theta, \,\mathrm{d}\omega, \,\mathrm{d}x) := \frac{1}{|\Lambda_N|} \sum_{j \in \Lambda_N} \delta_{(\theta_i(t), \omega_i, x_j)}(\,\mathrm{d}\theta, \,\mathrm{d}\omega, \,\mathrm{d}x), \; t \geqslant 0$$

that can be seen as a random probability measure.

REMARK 1.1. *The renormalization of the positions by $\frac{1}{2N}$ maps $\Lambda_N = [\![-N, \ldots, N]\!]^d$ to a discrete subset of $\left[-\frac{1}{2}, \frac{1}{2} \right]^d$. The necessity of this renormalization will become clear in the discussion on the spatial constraints below in this introduction.*

Since we are interested in the collective behavior of a large numbers of neurons, as it is the case for neural ensembles in the brain, understanding the asymptotic behavior of $\nu_t^{(N)}$ as $N \to \infty$ is important.

Under the assumption that

$$(1.5) \qquad \Psi_N(i, j) = \Psi\left(\frac{i}{2N}, \frac{j}{2N} \right)$$

for a general class of functions Ψ defined on $\left[-\frac{1}{2}, \frac{1}{2} \right]^d \times \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, we prove, as part of our main results in this paper (see Theorems 2.13 and 2.18), that $\nu_t^{(N)}$ converges to a

deterministic measure $\nu_t(\,\mathrm{d}\theta,\,\mathrm{d}\omega,\,\mathrm{d}x) = q_t(\theta,\omega,x)\,\mathrm{d}\theta\mu(\,\mathrm{d}\omega)\,\mathrm{d}x$ where q_t is a weak solution of the McKean-Vlasov equation

(1.6)
$$\partial_t q_t = \frac{1}{2}\mathrm{div}_\theta\big(\sigma\sigma^T\nabla_\theta q_t\big) - \mathrm{div}_\theta\left(q_t\left\{c(\theta,\omega) + \int \Gamma(\theta,\omega,\bar\theta,\bar\omega)\Psi(x,\bar x)q_t(\bar\theta,\bar\omega,\bar x)\,\mathrm{d}\bar\theta\,\mathrm{d}\mu(\bar\omega)\,\mathrm{d}\bar x\right\}\right).$$

For a formal derivation of this equation, we refer to the end of § 2.4 below. The measure ν_t is called the mean field limit of the system (1.3). Through Theorems 2.13 and 2.18, we not only prove the convergence $\nu_t^{(N)}$ towards ν_t, but we also provide some explicit estimates on the speed of convergence in terms of an appropriate weighted Wasserstein distance.

1.2. *Existing literature and motivations.*

1.2.1. *The non-spatial case:* $\Psi_N \equiv 1$. Of course, since there is no spatial interaction in this case, indexing the diffusions by a subset of $\mathbf{Z}^d$ is not relevant. Systems of type (1.3) are called *mean field models* (or *weakly interacting diffusions*) in statistical physics and have attracted much attention in the past years (see e.g. [26, 15, 28, 35, 10]), since they are capable of modeling complex dynamical behavior of various types of real-world models from physics to biology, like e.g. synchronization of large populations of individuals, collective behavior of social insects, emergence of synchrony in neural networks ([2, 37, 11]), and providing particle approximations for various nonlinear PDEs appearing in physics ([7, 6, 4, 24, 5]).

The most prominent example of such models is the Kuramoto model which has been widely considered in the literature as the main prototype for synchronization phenomena (see e.g. [1, 23, 3, 18, 34]):

$$(1.7) \qquad \mathrm{d}\theta_i(t) = \omega_i\,\mathrm{d}t + \frac{K}{N}\sum_{j=1}^{N}\sin\left(\theta_j - \theta_i\right)\,\mathrm{d}t + \sigma\,\mathrm{d}B_i(t),\ t \geqslant 0,\ i = 1,\ldots,N.$$

where $K \geqslant 0$ is the intensity of interaction and $\theta_i \in \mathbf{S} := \mathbf{R}/2\pi$.

In the context of weighted interactions, a notable attempt to go beyond pure mean field interactions has been to consider *moderately interacting diffusions* (see [29, 27, 21]).

1.2.2. *The spatial case.* The motivation of going beyond pure mean-field interaction comes from the biological observation that neurons do not interact in a mean-field way (see e.g. [39] and references therein) and a vast literature exists in physics about synchronization on general networks. In particular, several papers have already considered the model (1.3) (in dimension $d = 1$) for different choices of spatial weight Ψ defined in (1.5). In this paper, we will be more particularly interested in two classes of spatial weights:

1. **The P-nearest neighbor model:** this model (see [30, 31]) concerns the case where each diffusion $\theta_i \in \Lambda_N$ only interacts with its neighbors within a box $\Lambda_P \subseteq \Lambda_N$, where P is smaller than N:

$$(1.8) \qquad \mathrm{d}\theta_i(t) = c(\theta_i,\omega_i)\,\mathrm{d}t + \frac{1}{|\Lambda_P|}\sum_{\substack{j\in\Lambda_P \\ j\neq i}} \Gamma\left(\theta_i,\omega_i,\theta_j,\omega_j\right)\,\mathrm{d}t + \sigma\cdot\mathrm{d}B_i(t),\ i\in\Lambda_N.$$

We are concerned in this work with the case where P is proportional to N, that is

$$(1.9) \qquad\qquad\qquad\qquad\qquad P = RN,$$

for a fixed proportion $R \in (0, 1]$.

REMARK 1.2. *The case of $R = 1$ corresponds to the mean field case. Understanding the behavior of the system* (1.8) *in the case of a pure local interaction (that is when $P \ll N$) does not enter into the scope of this work. In particular, we will not address the question of P of order smaller than N (e.g. $P = RN^\alpha$ for some $\alpha < 1$), whose behavior as $N \to \infty$ seems to be quite different.*

Under the assumption (1.9), the P-nearest-neighbor model (1.8) enters into the framework of (1.3) for the following choice of Ψ in (1.5):

$$(1.10) \qquad \forall x, y \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d, \ \Psi(x, y) := \chi_R(x - y) := \frac{1}{(2R)^d} \mathbb{1}_{[-R,R]^d} (x - y).$$

2. **The power-law model:** this model also considered in the physical literature (see [9, 19, 25, 33]) corresponds to the case where Ψ in (1.5) is given by:

$$(1.11) \qquad\qquad \forall x, y \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d, \ \Psi(x, y) := \frac{1}{\| x - y \|^\alpha},$$

for some parameter $\alpha \geqslant 0$, that is
$$(1.12)$$
$$\mathrm{d}\theta_i(t) = c(\theta_i, \omega_i)\, \mathrm{d}t + \frac{1}{|\Lambda_N|} \sum_{\substack{j \in \Lambda_N \\ j \neq i}} \Gamma\left(\theta_i, \omega_i, \theta_j, \omega_j\right) \left\| \frac{i-j}{2N} \right\|^{-\alpha} \mathrm{d}t + \sigma \cdot \mathrm{d}B_i(t), i \in \Lambda_N.$$

Note that the pure mean field case corresponds again to $\alpha = 0$. As observed in the articles mentioned above on the basis of numerical simulations, it appears that the behavior of the system is strongly dependent on the value of the parameter α. The situation which is considered in this paper corresponds to the subcritical case where the parameter is smaller than the dimension:

$$(1.13) \qquad\qquad\qquad\qquad\qquad \alpha < d.$$

The case of $\alpha \geqslant d$ is much more delicate and will be the object of future work. We refer to Remark 2.7 below for further explanations on this case.

It is easy to see that in the case of (1.13) the renormalization of the positions by a factor $\frac{1}{2N}$ in (1.12) is necessary: by standard arguments, the diverging series $\sum_{j \in \Lambda_N, j \neq i} \| i - j \|^{-\alpha}$ is of order $N^{d-\alpha}$. Consequently, $\frac{1}{|\Lambda_N|} \sum_{\substack{j \in \Lambda_N \\ j \neq i}} \left\| \frac{i-j}{2N} \right\|^{-\alpha}$ is of order $\frac{N^\alpha}{|\Lambda_N|} N^{d-\alpha} = O(1)$, so that we should expect a nontrivial limit in (1.12), as $N \to \infty$.

1.3. *Main lines of proof and organization of the paper.* The strategy usually used in the literature on mean-field models (see [15, 21, 23, 28]) for the convergence of the empirical measure (1.4) is the following: first prove tightness of $(\nu^{(N)})_{N \geqslant 1}$ in the set of measure-valued continuous processes and second, prove uniqueness of any possible limit points, that is, uniqueness in the McKean-Vlasov equation (1.6).

In our context, *a priori* uniqueness in (1.6) appears unclear, due the fact that our model includes *singular spatial weights* (discontinuous in (1.10) and singular in (1.11)) and also a class of dynamics with *no global-Lipschitz continuity* and *polynomial growth* (recall the FitzHugh-Nagumo case (1.1)). Note that we are also concerned with the case where σ is degenerate (even equally zero) for which uniqueness in (1.6) is also not clear.

To bypass this difficulty, we adopt a converse strategy: we first prove *existence* of a solution to the mean-field limit (1.6) (through an *ad-hoc* fixed point argument, using ideas from Sznitman [36]). Secondly, via a propagator method (see [12] for related ideas) we prove the convergence (with respect to a Wasserstein-like distance adapted to the singularities of the interaction) of the empirical measure to *any* solution to (1.6). In particular, easy byproducts of this method are *uniqueness* of any solution to (1.6) as well as *explicit rates of convergence* to the McKean-Vlasov limit. In that sense, one of the main conclusions of the paper is to exhibit a *phase transition* in the size of the fluctuations in the power-law case (see Theorem 2.18). An actual Central Limit Theorem in this case is of course a natural perspective and is currently under investigation.

The paper is organized as follows: we give in Section 2 the main assumptions on the model and we state the main results (Theorems 2.13 and 2.18). Section 3 contains the proof of Proposition 2.9 concerning the existence of a solution to the McKean-Vlasov equation (1.6). Section 4 summarizes the main ideas and results concerning the propagator method. The proofs of the laws of large numbers are provided in Section 5 for the P-nearest case and in Section 6 for the power-law case. An additional assumption of regularity is made from Section 4 to 6, with is is discarded in Section 7.

2. Mathematical set-up and main results.

2.1. *The model.* Fix $N \geqslant 1$, $T > 0$ and let Λ_N be the hypercube $[\![-N, \ldots, N]\!]^d \subset \mathbf{Z}^d$ and $|\Lambda_N| = (2N + 1)^d$ be its volume. We consider $|\Lambda_N|$ diffusions on $[0, T]$ with values in the state space[1] $\mathcal{X} := \mathbf{R}^m$ for a certain $m \geqslant 1$.

Each diffusion θ_i is attached to the site i of Λ_N. The local dynamics of θ_i is governed by the following stochastic differential equation which is perturbed by a random environment represented by a vector $\omega_i \in \mathcal{E} := \mathbf{R}^n$ ($n \geqslant 1$).

$$(2.1) \qquad \mathrm{d}\theta_i(t) = c(\theta_i, \omega_i)\,\mathrm{d}t + \sigma \cdot \mathrm{d}B_i(t), \ 0 \leqslant t \leqslant T, i \in \Lambda_N.$$

where $\sigma \in \mathbf{R}^{m \times m}$ is the covariance matrix, $c(\cdot, \cdot)$ is a function from $\mathcal{X} \times \mathcal{E}$ to $\mathcal{X}$ and (B_i) is a given sequence of independent Brownian motions in $\mathcal{X}$. The vectors $(\omega_i)_{i \in \Lambda_N}$ are supposed to be i.i.d. realizations of a law μ and are hence seen as a random environment for the diffusions.

---

[1]Note that it is also possible to choose $\mathcal{X}$ as the circle $\mathbf{S} := \mathbf{R}/2\pi\mathbf{Z}$ in the case of the Kuramoto model, but we will stick to $\mathcal{X} := \mathbf{R}^m$ for simplicity.

When connected to the others, the diffusions interact in a mean field way with spatial extension:

$$(2.2) \qquad \mathrm{d}\theta_i(t) = c(\theta_i, \omega_i) \, \mathrm{d}t + \frac{1}{|\Lambda_N|} \sum_{\substack{j \in \Lambda_N \\ j \neq i}} \Gamma(\theta_i, \omega_i, \theta_j, \omega_j) \, \Psi\left(\frac{i}{2N}, \frac{j}{2N}\right) \mathrm{d}t + \sigma \cdot \mathrm{d}B_i(t),$$

$$0 \leqslant t \leqslant T, i \in \Lambda_N,$$

where Γ is a function from $(\mathcal{X} \times \mathcal{E})^2$ to $\mathcal{X}$, and $(x, y) \mapsto \Psi(x, y)$ is a function from $\left[-\frac{1}{2}, \frac{1}{2}\right]^d \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d$ to $[0, \infty)$. The required assumptions for the function Ψ will be made precise in Assumption 2.5 below. One should notice at this point that $\Psi(x, y)$ does not need to depend on the difference $x - y$.

We suppose that, at time $t = 0$, the variables $(\theta_i(0))_{1 \leqslant i \leqslant N}$ are independent and identically distributed according to a probability distribution $\zeta(\mathrm{d}\theta)$ on $\mathcal{X}$.

REMARK 2.1. *Instead of considering diffusions on Λ_N, we can also suppose periodic boundary conditions, i.e. when Λ_N is replaced by $\Lambda_{N,per} := \mathbb{T}_N^d$, where $\mathbb{T}_N$ is the discrete N-torus, that is $[\![-N, \dots, N]\!]$ with $-N$ and N identified. The only thing that changes in what follows in the continuous model is that one should replace $\left[-\frac{1}{2}, \frac{1}{2}\right]^d$ by $\mathbb{T}^d$ where $\mathbb{T} := \left[-\frac{1}{2}, \frac{1}{2}\right]/_{-\frac{1}{2} \sim \frac{1}{2}}$. Since the corresponding changes in the proofs of this paper remain marginal, we will restrict to the non periodic case and let the interested reader make the appropriate modifications in the periodic case.*

2.2. *Notations and assumptions.* From now on, we will suppose that the following assumptions (Assumptions 2.2, 2.4 and 2.5) are satisfied throughout the paper. In particular, saying that Assumption 2.5 is true means that we are either in the P-nearest neighbor case or in the power-law case (see Hypotheses (H1) and (H2) below).

ASSUMPTION 2.2 (Hypothesis on Γ and c). *We make the following assumptions:*

- *The function $(\theta, \omega) \mapsto c(\theta, \omega)$ is supposed to be locally Lipschitz-continuous in θ (for fixed ω) and satisfy a one-sided Lipschitz condition w.r.t. the two variables (θ, ω):*

$$(2.3) \qquad \forall (\theta, \omega), (\bar{\theta}, \bar{\omega}), \ \langle \theta - \bar{\theta}, c(\theta, \omega) - c(\bar{\theta}, \bar{\omega}) \rangle \leqslant L \left(\|\theta - \bar{\theta}\|^2 + \|\omega - \bar{\omega}\|^2 \right),$$

for some constant L (not necessarily positive). We suppose also some polynomial bound about the function c:

$$(2.4) \qquad \forall (\theta, \omega), \|c(\theta, \omega)\| \leqslant \|\!|c|\!\| \left(1 + \|\theta\|^\kappa + \|\omega\|^\iota \right),$$

for some constant $\|\!|c|\!\| > 0$ and where $\kappa \geqslant 2$ and $\iota \geqslant 1$.
- *The interaction term Γ is supposed to be bounded by $\|\Gamma\|_\infty$ and globally Lipschitz-continuous on $(\mathcal{X} \times \mathcal{E})^2$, with a Lipschitz constant $\|\Gamma\|_{Lip}$.*

We also assume that for fixed $\bar{\theta}, \omega, \bar{\omega}$, the functions $\theta \mapsto c(\theta, \omega)$ and $\theta \mapsto \Gamma(\theta, \omega, \bar{\theta}, \bar{\omega})$ are twice differentiable with continuous derivatives.

REMARK 2.3. *Assumption 2.2 is in particular satisfied for the FitzHugh-Nagumo case. One technical difficulty is the dynamics is not globally Lispchitz continuous. This will entail some technical complications in the following. Note also that the constant $\|\!\| c \|\!\|$ mentioned in (2.3) does not take part in the estimates of Sections 4 to 6. It only enters into account in Section 3.*

ASSUMPTION 2.4 (Assumptions on μ and ζ). *We suppose that the initial distribution ζ of θ satisfies the following moment condition:*

$$(2.5) \qquad \int_{\mathcal{X}} \|\theta\|^{\kappa} \zeta(\mathrm{d}\theta) < \infty,$$

and that the law of the disorder μ satisfies the moment condition:

$$(2.6) \qquad \int_{\mathcal{E}} \|\omega\|^{\iota} \mu(\mathrm{d}\omega) < \infty,$$

where the constants κ and ι are given by (2.4) in Assumption 2.2.

ASSUMPTION 2.5 (Assumptions on the weight Ψ). *In order to cover the case of both the P-nearest model and the power-law interaction introduced in § 1.2.2, we suppose that either Hypothesis (H1) or Hypothesis (H2) is true:*

(H1) P-nearest-neighbor:

$$(2.7) \qquad \forall x, y \in \left[-\tfrac{1}{2}, \tfrac{1}{2}\right]^d, \ \Psi(x,y) := \chi_R(x,y)$$

where χ_R is defined in (1.10).

(H2) Power-law: the function Ψ is supposed to be a nonnegative function on $\left[-\tfrac{1}{2}, \tfrac{1}{2}\right]^d \times \left[-\tfrac{1}{2}, \tfrac{1}{2}\right]^d$ such that the following properties are satisfied:

$$(2.8) \qquad \mathcal{I}_1(\Psi) := \sup_{a,\, x \in \left[-\frac{1}{2}, \frac{1}{2}\right]^d} \|x-a\|^{\alpha} \Psi(x,a) < \infty,$$

$$(2.9) \qquad \mathcal{I}_2(\Psi) := \sup_{x,\, y \in \left[-\frac{1}{2}, \frac{1}{2}\right]^d} \frac{\int |\Psi(x,\bar{x}) - \Psi(y,\bar{x})| \, \mathrm{d}\bar{x}}{\|x-y\|^{(d-\alpha)\wedge 1}} < \infty,$$

$$(2.10) \qquad \mathcal{I}_3(\Psi) := \sup_{a,\, x,\, y \in \left[-\frac{1}{2}, \frac{1}{2}\right]^d} \frac{\left| \|x-a\|^{2\gamma} \Psi(x,a) - \|y-a\|^{2\gamma} \Psi(y,a) \right|}{\|x-y\|^{(2\gamma-\alpha)\wedge 1}} < \infty,$$

for some parameters $\alpha \in [0,d)$ and γ chosen to be

$$(2.11) \qquad \begin{cases} \gamma \in \left[\alpha, \frac{d}{2}\right) & \text{if } \alpha \in \left[0, \frac{d}{2}\right) \\ \gamma = \frac{d}{2} & \text{otherwise.} \end{cases}$$

REMARK 2.6. *Note that we could have chosen simply $\gamma = \frac{d}{2}$ in any case. But this would have led to worse convergence rates than the ones that we obtain below in Theorem 2.18.*

Of course, the main prototype for Hypothesis (H2) is when $\Psi(x,y) = \| x - y \|^{-\alpha}$, for $\alpha < d$ (recall (1.11)). But, the assumptions made in (H2) cover a larger class of examples: the reader may think of the general case of $\Psi(x,y) := \psi(x,y) \| x - y \|^{-\alpha}$, for a bounded Lipschitz-continuous function ψ. Note also that the case of bounded Lipschitz interactions is also captured (take $\alpha = 0$).

REMARK 2.7 (About the supercritical case).   *The case of a power-law interaction with $\alpha \geqslant d$ is more delicate and requires more attention. Note that, to our knowledge, no proposition for any continuous limit has been made in the literature in this case. We are only aware of [9], where the system (2.12) below is considered for finite N.*

One trivial observation is that the series $\sum_{\substack{j \in \Lambda_N \\ j \neq i}} \| i - j \|^{-\alpha}$ is in this case already convergent. Consequently, an interaction term of the form $\frac{1}{|\Lambda_N|} \sum_{\substack{j \in \Lambda_N \\ j \neq i}} \Gamma(\theta_i, \omega_i, \theta_j, \omega_j) \| i - j \|^{-\alpha}$ simply vanishes to 0 as $N \to \infty$. Hence, the correct model in this case is where the factor $\frac{1}{|\Lambda_N|}$ is absent

$$(2.12) \qquad \mathrm{d}\theta_i(t) = c(\theta_i, \omega_i)\,\mathrm{d}t + \sum_{\substack{j \in \Lambda_N \\ j \neq i}} \Gamma(\theta_i, \omega_i, \theta_j, \omega_j) \| i - j \|^{-\alpha}\,\mathrm{d}t + \sigma.\,\mathrm{d}B_i(t),\ i \in \Lambda_N.$$

The main difficulty for the derivation of the correct continuous limit in the case of (2.12) lies in the fact that the interaction term $\sum_{\substack{j \in \Lambda_N \\ j \neq i}} \Gamma(\theta_i, \omega_i, \theta_j, \omega_j) \| i - j \|^{-\alpha}$ is not sufficiently mixing: if it exists, the McKean-Vlasov limit in this case should be random. We believe that the correct continuous limit should be governed by a stochastic partial differential equation instead of a deterministic PDE. This case is currently under investigation and will be the object of a future work.

2.3. *The empirical measure.*   Let us consider for fixed horizon T and time $t \in [0, T]$, the empirical measure $\nu_t^{(N)}$ (introduced in (1.4)):

$$(2.13) \qquad \nu_t^{(N)}(\,\mathrm{d}\theta,\,\mathrm{d}\omega,\,\mathrm{d}x) := \frac{1}{|\Lambda_N|} \sum_j \delta_{(\theta_j(t),\,\omega_j,\,x_j)}(\,\mathrm{d}\theta,\,\mathrm{d}\omega,\,\mathrm{d}x),$$

as a probability measure on $\mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2} \right]^d$. Here

$$(2.14) \qquad x_j := \frac{j}{2N} \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d,\ j \in \Lambda_N.$$

2.4. *The McKean-Vlasov equation.*   The convergence of the empirical measure at $t = 0$ is clear: since $(\theta_i(0), \omega_i)_{1 \leqslant i \leqslant N}$ are i.i.d. random variables sampled according to $\zeta \otimes \mu$, the initial empirical measure $\nu_0^{(N)}$ converges, as $N \to \infty$, to

$$(2.15) \qquad \nu_0(\,\mathrm{d}\theta,\,\mathrm{d}\omega,\,\mathrm{d}x) := \zeta(\,\mathrm{d}\theta)\mu(\,\mathrm{d}\omega)\,\mathrm{d}x.$$

An application of Ito's formula to (2.2) (for any $(\theta, \omega, x) \mapsto f(\theta, \omega, x)$ bounded function of class $\mathcal{C}^2$ w.r.t. θ with bounded derivatives) leads to the following martingale representation

for $\nu^{(N)}$:

$$(2.16) \quad \left\langle \nu_t^{(N)}, f \right\rangle = \left\langle \nu_0^{(N)}, f \right\rangle + \int_0^t \left\langle \nu_s^{(N)}, \frac{1}{2}\text{div}_\theta \left(\sigma\sigma^T \nabla_\theta f\right) + \nabla_\theta f \cdot c(\cdot, \cdot) \right\rangle \, \mathrm{d}s$$

$$+ \int_0^t \left\langle \nu_s^{(N)}, \nabla_\theta f \cdot \int \Gamma(\cdot, \cdot, \bar\theta, \bar\omega)\Psi(\cdot, \bar x)\nu_s^{(N)}(\mathrm{d}\bar\theta, \mathrm{d}\bar\omega, \mathrm{d}\bar x) \right\rangle \, \mathrm{d}s + M_t^{(N)}(f),$$

where $M_t^{(N)}(f) := \frac{1}{|\Lambda_N|} \sum_j \int_0^t \nabla_\theta f(\theta_j(s), \omega_j, x_j) \cdot \sigma \, \mathrm{d}B_j(s)$ is a martingale. Note that we use here the usual duality notation $\langle \nu, f \rangle = \int f \, \mathrm{d}\nu$ for the integral of a test function f against a measure ν.

Taking formally $N \to \infty$ in (2.16) shows that any limit point of $\nu^{(N)}$ should satisfy the following nonlinear McKean-Vlasov equation

$$(2.17) \quad \partial_t \langle \nu_t, f \rangle = \left\langle \nu_t, \frac{1}{2}\text{div}_\theta \left(\sigma\sigma^T \nabla_\theta f\right) + \nabla_\theta f \cdot c(\cdot, \cdot) \right\rangle$$

$$+ \left\langle \nu_t, \nabla_\theta f \cdot \int \Gamma(\cdot, \cdot, \bar\theta, \bar\omega)\Psi(\cdot, \bar x)\nu_t(\mathrm{d}\bar\theta, \mathrm{d}\bar\omega, \mathrm{d}\bar x) \right\rangle,$$

where $\Psi(\cdot, \cdot)$ is the weight function introduced either in Hypothesis (H1) or in Hypothesis (H2).

REMARK 2.8. *An important remark about* a priori *properties of* (2.17) *is the following: taking a test function f in* (2.17) *that does not depend on θ implies*

$$\langle \nu_0, f \rangle = \langle \nu_t, f \rangle, \ \forall t \in [0, T].$$

In particular, the marginal distribution of (ω, x) w.r.t. the measure ν_t is independent of t and equal to $\mathrm{d}\mu \otimes \mathrm{d}x$. This implies that, for the class of singular weight we consider here, Ψ is always integrable against ν_t, for all t, since the function $y \mapsto \|x - y\|^{-\alpha}$ is integrable w.r.t. to the Lebesgue measure on $\left[-\frac{1}{2}, \frac{1}{2}\right]^d$.

Moreover, since the function c is supposed to have a polynomial growth (recall (2.4)), one has to justify in particular the term $\langle \nu_t, \nabla_\theta f \cdot c(\cdot, \cdot) \rangle$ in (2.17) *(the others are easily integrable). Thus, one should look for solutions $t \mapsto \nu_t$ having finite moment: for all $t \in [0, T]$, $\int_{\mathcal{X} \times \mathcal{E}} \|\theta\|^\kappa \|\omega\|^\iota \nu_t(\mathrm{d}\theta, \mathrm{d}\omega, \mathrm{d}x) < \infty$.*

In particular, well-posedness in (2.17) *will be addressed within the class of all measure-valued processes satisfying the properties mentioned above.*

Formally integrating by parts in equation (2.17), assuming the existence of a density $\nu_t(\mathrm{d}\theta, \mathrm{d}\omega, \mathrm{d}x) = q_t(\theta, \omega, x) \, \mathrm{d}\theta\mu(\mathrm{d}\omega) \, \mathrm{d}x$, q_t satisfies

$$(2.18) \quad \partial_t q_t = \frac{1}{2}\text{div}_\theta \left(\sigma\sigma^T \nabla_\theta q_t\right) - \text{div}_\theta \left(q_t(\theta, \omega, x)c(\theta, \omega)\right)$$

$$- \text{div}_\theta \left(q_t(\theta, \omega, x) \int \Gamma(\theta, \omega, \bar\theta, \bar\omega)\Psi(x, \bar x)q_t(\bar\theta, \bar\omega, \bar x) \, \mathrm{d}\bar\theta\mu(\mathrm{d}\bar\omega) \, \mathrm{d}\bar x\right), \ t > 0,$$

In the case where σ is non degenerate, one can make this integration by parts rigorous: using the same arguments as in [17, Appendix A], one can show that for any measure-valued initial condition in (2.17), by the regularizing properties of the heat kernel, the

solution of (2.17) has a regular density q_t for all positive time that solves (2.18). We refer to [17, Prop. A.1] for further details. But of course, if σ is degenerate, the strong formulation (2.18) does not necessarily make sense and one has to restrict to the weak formulation (2.17) in that case.

2.5. *Results.* The first result of this paper, whose proof is given in Section 3, concerns the existence of a weak solution to the McKean-Vlasov equation (2.17):

PROPOSITION 2.9. *Under Assumptions 2.2, 2.4 and 2.5, for any initial condition* $\nu_0(\,\mathrm{d}\theta, \mathrm{d}\omega, \mathrm{d}x) = \zeta(\,\mathrm{d}\theta)\mu(\,\mathrm{d}\omega)\,\mathrm{d}x$, *there exists a solution* $t \mapsto \nu_t$ *to (2.17).*

Having proven the existence of at least one such solution in the general case, we turn to the issue of the convergence of the empirical measure to *any* of such solution. From now on, we specify the problem to the case of Hypothesis (H1) (§ 2.5.1) and of Hypothesis (H2) (§ 2.5.2). For each case, in order to state the convergence result, one needs to define an appropriate distance between two random measures that is basically the supremum over evaluations against a set of test functions. Such a space of test functions must incorporate the kind of singularities that are present either in Hypothesis (H1) or (H2).

2.5.1. *The P-nearest-neighbor case.* Suppose that the weight function Ψ satisfies Hypothesis (H1) of Assumption 2.5.

DEFINITION 2.10 (Test functions for *P*-nearest-neighbor). *For fixed* $R \in (0,1]$ *and* $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, *let* $\mathcal{C}_{R,a}$ *be the set of functions* f *on* $\mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2} \right]^d$ *of the form:*

$$f : (\theta, \omega, x) \mapsto g(\theta, \omega) \cdot \chi_R(x - a),$$

where χ_R *is given in (1.10) and* g *is globally Lipschitz-continuous w.r.t.* (θ, ω):

$$(2.19) \qquad \exists C > 0, \forall (\theta, \omega, \bar{\theta}, \bar{\omega}), \; \left\| g(\theta, \omega) - g(\bar{\theta}, \bar{\omega}) \right\| \leqslant C \left(\left\| \theta - \bar{\theta} \right\| + \left\| \omega - \bar{\omega} \right\| \right).$$

Let

$$\| f \|_{R,a} := \sup_{\theta, \bar{\theta}, \omega, \bar{\omega}} \frac{\left\| g(\theta, \omega) - g(\bar{\theta}, \bar{\omega}) \right\|}{\left\| \theta - \bar{\theta} \right\| + \left\| \omega - \bar{\omega} \right\|}$$

be the corresponding seminorm.

REMARK 2.11. *Note that for any* $f \in \mathcal{C}_{R,a}$ *that is* $\mathcal{C}^1$ *in the variable* θ, *the following estimate holds:*

$$(2.20) \qquad \forall \theta, \omega, x, \; \left\| \nabla_\theta f(\theta, \omega, x) \right\| \leqslant \| f \|_{R,a} \, \chi_R(x - a).$$

We now turn to the appropriate distance between two random measures:

DEFINITION 2.12 (Distance for *P*-nearest neighbor). *For random probability measures* λ *and* ν *on* $\mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, *let*

$$d_R(\lambda, \nu) := \sup_f \left(\mathbf{E} \left\| \langle f , \lambda \rangle - \langle f , \nu \rangle \right\|^2 \right)^{1/2}$$

where the supremum is taken over all functions $f \in \bigcup_{a \in [-1,1]^d} \mathcal{C}_{R,a}$, *such that* $\| f \|_{R,a} \leqslant 1$, $\| f \|_\infty \leqslant 1$.

Our convergence result is given in the following

THEOREM 2.13 (Law of Large Numbers). *Under Assumptions 2.2, 2.4 and Hypothesis (H1) of Assumption 2.5, for all $R \in (0, 1]$, for any arbitrary solution ν to the mean-field equation (2.17), we have:*

$$(2.21) \qquad \sup_{0 \leqslant t \leqslant T} d_R(\nu_t^{(N)}, \nu_t) \leqslant \frac{C}{N^{1 \wedge \frac{d}{2}}}$$

where the constant $C > 0$ only depends on T, Γ, R and c.

2.5.2. *The case of the power-law interaction.* Assume that the weight function Ψ satisfies Hypothesis (H2). In view of the form of Ψ in this case (recall Assumption 2.5), the main idea is to consider test functions $(\theta, \omega, x) \mapsto f(\theta, \omega, x)$ that become regular when renormalized by $\| x - a \|^\alpha$. The seminorm $\| \cdot \|_a$ introduced in (2.25) below should therefore be thought of as a *weighted Hölder seminorm*.

DEFINITION 2.14 (Test functions for power-law interaction). *For fixed α and γ as in Assumption 2.5 and for fixed $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, let $\mathcal{C}_a$ be the set of functions $(\theta, \omega, x) \mapsto f(\theta, \omega, x)$ on $\mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2} \right]^d$ satisfying:*

- *Regularity w.r.t. (θ, ω): $(\theta, \omega) \mapsto \| x - a \|^\alpha f(\theta, \omega, x)$ is globally Lipschitz-continuous on $\mathcal{X} \times \mathcal{E}$, uniformly in x, that is*
  $$(2.22) \quad \exists C > 0, \forall(\theta, \omega, \bar{\theta}, \bar{\omega}), \ \| x - a \|^\alpha \left\| f(\theta, \omega, x) - f(\bar{\theta}, \bar{\omega}, x) \right\| \leqslant C \left(\left\| \theta - \bar{\theta} \right\| + \left\| \omega - \bar{\omega} \right\| \right)$$

- *Regularity w.r.t. x: $x \mapsto \| x - a \|^\alpha f(\theta, \omega, x)$ is uniformly bounded*
  $$(2.23) \qquad \exists C > 0, \| x - a \|^\alpha \left\| f(\theta, \omega, x) \right\| \leqslant C,$$
  and $x \mapsto |x - a|^{2\gamma} f(\theta, \omega, x)$ is globally $(2\gamma - \alpha) \wedge 1$-Hölder, uniformly in (θ, ω):
  $$(2.24) \quad \exists C > 0, \left\| \| x - a \|^{2\gamma} f(\theta, \omega, x) - \| y - a \|^{2\gamma} f(\theta, \omega, y) \right\| \leqslant C \| x - y \|^{(2\gamma - \alpha) \wedge 1}.$$

Denote by

$$(2.25) \quad \| f \|_a := \sup_{\theta, \bar{\theta}, \omega, \bar{\omega}, x} \frac{\| x - a \|^\alpha \left\| f(\theta, \omega, x) - f(\bar{\theta}, \bar{\omega}, x) \right\|}{\left\| \theta - \bar{\theta} \right\| + \left\| \omega - \bar{\omega} \right\|} + \sup_{\theta, \omega, x} \| x - a \|^\alpha \left\| f(\theta, \omega, x) \right\|$$
$$+ \sup_{\theta, \omega, x, y} \frac{\left\| \| x - a \|^{2\gamma} f(\theta, \omega, x) - \| y - a \|^{2\gamma} f(\theta, \omega, y) \right\|}{\| x - y \|^{(2\gamma - \alpha) \wedge 1}}$$

the corresponding seminorm.

REMARK 2.15. *Note that for any $f \in \mathcal{C}_a$ that is $\mathcal{C}^1$ in the variable θ, the following holds:*

$$(2.26) \qquad \forall \theta, \omega, x, \ \| \nabla_\theta f(\theta, \omega, x) \| \leqslant \frac{\| f \|_a}{\| x - a \|^\alpha}.$$

The corresponding definition of the distance between two random measures is similar to Definition 2.12 given in the P-nearest neighbor case. The main difference here is that one needs to take care of test functions with singularities. Since those singularities happen at points of the form $\frac{i}{2N}$ (for some i and N) that are regularly distributed on $\left[-\frac{1}{2}, \frac{1}{2}\right]^d$, we first need to introduce some further notations: for all integer $K \geqslant 1$, we denote by $\mathcal{D}_K$ the regular discretization of $\left[-\frac{1}{2}, \frac{1}{2}\right]^d$ with mesh of length $\frac{1}{2K}$:

$$(2.27) \qquad \mathcal{D}_K := \left\{ \left(\frac{j_1}{2K}, \ldots, \frac{j_d}{2K} \right); -K \leqslant j_1 \leqslant K, \ldots, -K \leqslant j_d \leqslant K \right\} \subset \left[-\frac{1}{2}, \frac{1}{2} \right]^d.$$

The appropriate distance between two random measures is then:

DEFINITION 2.16 (Distance for power-law interaction).   *Let $\alpha < d$ and $p \geqslant 2$ be defined by:*

$$(2.28) \qquad\qquad p := \begin{cases} 2 & \text{if } \alpha \in [0, \frac{d}{2}), \\ \lceil \frac{d}{d-\alpha} \rceil & \text{if } \alpha \in [\frac{d}{2}, d), \end{cases}$$

where $\lceil x \rceil$ stands for the smallest integer strictly larger than x. On the set of random probability measures on $\mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, let us define a sequence of distances $\left(d_K^{(p)}(\cdot, \cdot) \right)_{K \geqslant 1}$ indexed by $K \geqslant 1$, between two elements λ and ν by

$$d_K^{(p)}(\lambda, \nu) = \sup_f \left(\mathbf{E} \left\| \langle f, \lambda \rangle - \langle f, \nu \rangle \right\|^p \right)^{1/p}$$

where the supremum is taken over all the functions $f \in \bigcup_{\substack{a \in \mathcal{D}_{K'} \\ 1 \leqslant K' \leqslant K}} \mathcal{C}_a$, such that $\| f \|_a \leqslant 1$. Let us then define the distance $d_\infty^{(p)}(\cdot, \cdot)$ by

$$(2.29) \qquad\qquad d_\infty^{(p)}(\lambda, \nu) := \sum_{K \geqslant 1} \frac{1}{2^K} \frac{e^{-CK^{\frac{dp}{q}}}}{K^{2d}} \left(d_K^{(p)}(\lambda, \nu) \wedge 1 \right),$$

for a sufficiently large constant C (that depends on the parameters of our model) and where q is the conjugate of p: $\frac{1}{p} + \frac{1}{q} = 1$. For a precise estimate on C, we refer to Proposition 6.5 below.

Apart from the weight $\frac{e^{-CK^{\frac{dp}{q}}}}{K^{2d}}$ (which is precisely here to compensate the estimate that we find in Proposition 6.5 below), the definition of $d_\infty^{(p)}(\cdot, \cdot)$ exactly follows the usual Fréchet construction (see e.g. [16]).

REMARK 2.17.   *The choice of the integer p in (2.28) is made for integrability reasons that will become clear in the proof of Theorem 2.18. One only has to notice here that p has been precisely defined so that its conjugate q always satisfies $q\alpha < d$.*

The main result of this work is the following:

THEOREM 2.18 (Law of Large Numbers in the power-law case). *Under Assumptions 2.2, 2.4 and Hypothesis (H2) of Assumption 2.5, for any arbitrary solution ν to the mean-field equation (2.17), we have:*

$$(2.30) \qquad \sup_{0 \leqslant t \leqslant T} d_\infty^{(p)}(\nu_t^{(N)}, \nu_t) \leqslant C \begin{cases} \dfrac{1}{N^{\gamma \wedge 1}}, & \text{if } \alpha \in \left[0, \frac{d}{2}\right), \\[2mm] \dfrac{\ln N}{N^{\frac{d}{2} \wedge 1}}, & \text{if } \alpha = \frac{d}{2}, \\[2mm] \dfrac{\ln N}{N^{(d-\alpha) \wedge 1}}, & \text{if } \alpha \in \left(\frac{d}{2}, d\right), \end{cases}$$

where the constant $C > 0$ only depends on T, Γ, Ψ, α and c.

Note that the speed of convergence found in Theorem 2.18 is never smaller that $N^{-\frac{d}{2}}$ which is the optimal speed for the case without spatial extension (recall the CLT results in the mean field case in [23]). Note also that, in the case where $0 \leqslant \alpha < \frac{d}{2}$, we have obtained a speed of convergence which is arbitrarily close to $N^{-(\frac{d}{2} \wedge 1)}$ (since in that case γ is arbitrarily close to $\frac{d}{2}$). We believe that the optimal speed in this case should be *exactly* $N^{-(\frac{d}{2} \wedge 1)}$, but the proof we propose in this work does not seem to reach this optimal result.

Nevertheless, in the case where we only consider a *bounded Lispchitz-continuous* weight function Ψ (i.e. with no singularity at all), the proof of Theorem 2.18 can be considerably simplified and one obtains a speed that is $N^{-\frac{d}{2}}$.

Note also that the fluctuations when $\alpha \in \left[\frac{d}{2}, d\right)$ appear to be nontrivial. A natural perspective of this work would be to prove a precise Central Limit Theorem in this case and to study the limiting fluctuation process in details.

2.6. *Well-posedness of the McKean-Vlasov equation.* A straightforward corollary of Theorems 2.13 and 2.18 is that uniqueness holds for the McKean-Vlasov equation (2.17):

PROPOSITION 2.19 (Well-posedness of the McKean-Vlasov equation). *Under Assumptions 2.2, 2.4 and 2.5, for every initial condition $\nu_0(\,\mathrm{d}\theta,\,\mathrm{d}\omega,\,\mathrm{d}x) = \zeta(\,\mathrm{d}\theta)\mu(\,\mathrm{d}\omega)\,\mathrm{d}x$, there exists a unique solution $t \mapsto \nu_t \in \mathcal{M}_1\left(\mathcal{C}([0,T], \mathcal{X}) \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d\right)$ to the McKean-Vlasov equation (2.17).*

3. The non-linear process and the existence of a continuous-limit. The purpose of this paragraph is to prove Proposition 2.9 concerning the existence of a solution to the McKean-Vlasov equation (2.17). This part is reminiscent of the techniques used by Sznitman ([36]) in order to prove propagation of chaos for non disordered models.

3.1. *Distance on probability measures.* Let us first consider the set $\mathcal{M}_\mathcal{X}$ of probability measures on $\mathcal{C}([0,T], \mathcal{X})$ with finite moments of order κ (where $\kappa \geqslant 2$ is given in (2.4)) and endow this set with the Wasserstein metric

$$(3.1) \qquad \delta_\mathcal{X}^{(T)}(p_1, p_2) := \inf \left\{ \mathbf{E}\left(\sup_{s \leqslant T} \left\| \vartheta_s^{(1)} - \vartheta_s^{(2)} \right\|^\kappa \right)^{\frac{1}{\kappa}} \right\},$$

where the infimum in (3.1) is considered over all couplings $\left(\vartheta^{(1)}, \vartheta^{(2)}\right)$ with respective marginals p_1 and p_2. Here, the $\vartheta^{(i)}$ are understood as random variables on a certain

probability space $(\Omega, \mathbf{P})$. Note however that the definition of (3.1) does not depend on its particular choice. (3.1) defines a complete metric on $\mathcal{M}_{\mathcal{X}}$ encoding the topology of convergence in law with convergence of moments up to order κ (see [38, Th. 6.9, p. 96]). We endow $\mathcal{M}_{\mathcal{X}}$ with the corresponding Borel σ-field.

Fix some probability measure m on $\mathcal{C}([0,T], \mathcal{X}) \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d$ (endowed with its Borel σ-field) such that its marginal on $\mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d$ is absolutely continuous w.r.t. $\mu(\mathrm{d}\omega) \otimes \mathrm{d}x$. Thanks to a usual disintegration result (see e.g. [13, Th. 10.2.2]) one can write m as

$$m(\mathrm{d}\theta, \mathrm{d}\omega, \mathrm{d}x) = m^{\omega,x}(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\,\mathrm{d}x,$$

where $(\omega, x) \mapsto m^{\omega,x}(\mathrm{d}\theta)$ is a measurable map from $\mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d$ (endowed with its Borel σ-field) into $\mathcal{M}_{\mathcal{X}}$. We consider the set $\mathcal{M}$ of such measures m such that for all (ω, x), $m^{\omega,x}$ belongs to $\mathcal{M}_{\mathcal{X}}$, endowed with the following metric:

DEFINITION 3.1. *Fix p to be equal to 2 in the case of Hypothesis (H1) or as in (2.28) in the case of Hypothesis (H2). Then define*

$$(3.2) \qquad \forall m_1, m_2 \in \mathcal{M}, \;\; \delta_T(m_1, m_2) := \left[\int_{\mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d} \left(\delta_{\mathcal{X}}^{(T)}\left(m_1^{\omega,x}, m_2^{\omega,x}\right)\right)^p \mu(\mathrm{d}\omega)\,\mathrm{d}x\right]^{\frac{1}{p}}.$$

The space $\mathcal{M}$ endowed with δ_T is a complete metric space (see [36, p.173]).

Note that, by construction (see (2.15)), the initial condition $\mathrm{d}\nu_0(\theta, \omega, x) = \zeta(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\,\mathrm{d}x$ belongs to $\mathcal{M}$.

3.2. *The nonlinear process.* The proof of Proposition 2.9 is based on a Picard iteration in the space $\mathcal{M}$ endowed with the metric introduced in Definition 3.1. For fixed $\omega \in \mathcal{E}$ and Brownian motion B in $\mathcal{X}$, independent of the sequence $(B_k)_{k \geqslant 1}$, and for a fixed $m \in \mathcal{M}$, consider the following stochastic differential equation in $\mathcal{X}$:

$$(3.3) \qquad \mathrm{d}\theta(t) = c(\theta(t), \omega)\,\mathrm{d}t + \int \Gamma(\theta(t), \omega, \bar{\theta}, \bar{\omega})\Psi(x, \bar{x})m_t(\mathrm{d}\bar{\theta}, \mathrm{d}\bar{\omega}, \mathrm{d}\bar{x})\,\mathrm{d}t + \sigma \cdot \mathrm{d}B(t),$$

with initial condition $\theta(0) \sim \zeta$. Note here that for all $t \geqslant 0$, $m_t(\mathrm{d}\theta, \mathrm{d}\omega, \mathrm{d}x)$, probability measure on $\mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2}, \frac{1}{2}\right]^d$, stands for the projection of m at time t. The integral term in (3.3) is well-defined since

$$\int \|\Gamma(\theta(t), \omega, \bar{\theta}, \bar{\omega})\|\,\Psi(x, \bar{x})m_t(\mathrm{d}\bar{\theta}, \mathrm{d}\bar{\omega}, \mathrm{d}\bar{x})$$

$$\leqslant \|\Gamma\|_\infty \int_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d} \Psi(x, \bar{x}) \underbrace{\int_{\mathcal{X} \times \mathcal{E}} m_t^{\bar{\omega},\bar{x}}(\mathrm{d}\bar{\theta})\mu(\mathrm{d}\bar{\omega})}_{=1}\,\mathrm{d}\bar{x} \leqslant \|\Gamma\|_\infty\, S(\Psi),$$

where the quantity

$$(3.4) \qquad\qquad\qquad S(\Psi) := \sup_x \int_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d} \Psi(x, \bar{x})\,\mathrm{d}\bar{x}$$

is smaller than 1 in case of Hypothesis (H1) and smaller that $\mathcal{I}_1(\Psi)$ (using (2.8)) in the case of Hypothesis (H2). Moreover, thanks to the regularity properties of Γ and c, equation (3.3) has a unique (strong) solution.

Let us denote by $\Theta : \mathcal{M} \to \mathcal{M}$ the functional which maps any measure $m(\, \mathrm{d}\theta, \, \mathrm{d}\omega, \, \mathrm{d}x) \in \mathcal{M}$ to the law $\Theta(m)$ of (θ, ω, x) where $(\theta_t)_{0 \leqslant t \leqslant T}$ is the unique solution to (3.3). Note that the functional Θ effectively preserves the set $\mathcal{M}$. Proposition 2.9 is a direct consequence of the following lemma:

LEMMA 3.2.   *The functional Θ admits a fixed point $\bar{\nu}$ in $\mathcal{M}$.*

PROOF OF LEMMA 3.2. As in [36], we prove the following

$$(3.5) \qquad \forall m_1, m_2 \in \mathcal{M}, \ \forall t \leqslant T, \ \delta_t(\Theta(m_1), \Theta(m_2))^\kappa \leqslant C_T \int_0^t \delta_u(m_1, m_2)^\kappa \, \mathrm{d}u.$$

If (3.5) is proved, the proof of Proposition 2.9 will be finished since in that case, one can iterate this inequality and find

$$\forall k \geqslant 1, \ \delta_T(\Theta^{k+1}(\nu_0), \Theta^k(\nu_0))^\kappa \leqslant C_T^k \frac{T^k}{k!} \delta_T(\Theta(\nu_0), \nu_0)^\kappa,$$

which gives that $\left(\Theta^k(\nu_0) \right)_{k \geqslant 1}$ is a Cauchy sequence, and thus converges to some fixed-point $\bar{\nu}$ of Θ. Let us now prove (3.5). The key calculation is the following: there exists a constant $C > 0$ such that for all $\theta_1, \theta_2 \in \mathcal{X}$, $\omega \in \mathcal{E}$, $x \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, for all $m_1, m_2 \in \mathcal{M}$,

$$(3.6) \quad \delta\Gamma := \left\| \int \Gamma(\theta_1, \omega, \cdot, \cdot) \, \Psi(x, \cdot) \, \mathrm{d}m_{1,t} - \int \Gamma(\theta_2, \omega, \cdot, \cdot) \, \Psi(x, \cdot) \, \mathrm{d}m_{2,t} \right\|$$

$$\leqslant C \left(\| \theta_2 - \theta_1 \| \wedge 1 + \delta_t(m_1, m_2) \right).$$

Indeed,

$$(3.7) \quad \begin{aligned} \delta\Gamma &\leqslant \left\| \int \Gamma(\theta_1, \omega, \cdot, \cdot) \, \Psi(x, \cdot) \, \mathrm{d}m_{1,t} - \int \Gamma(\theta_2, \omega, \cdot, \cdot) \, \Psi(x, \cdot) \, \mathrm{d}m_{1,t} \right\| \\ &+ \left\| \int \Gamma(\theta_2, \omega, \cdot, \cdot) \, \Psi(x, \cdot) \, \mathrm{d}m_{1,t} - \int \Gamma(\theta_2, \omega, \cdot, \cdot) \, \Psi(x, \cdot) \, \mathrm{d}m_{2,t} \right\| := \delta\Gamma_1 + \delta\Gamma_2. \end{aligned}$$

The first term $\delta\Gamma_1$ in (3.7) is easily bounded by $\| \Gamma \|_{Lip} S(\Psi) \| \theta_2 - \theta_1 \|$, where $S(\Psi)$ is defined by (3.4). The second term $\delta\Gamma_2$ in (3.7) can be successively bounded by

$$\delta\Gamma_2 = \left\| \iint_{\left[-\frac{1}{2}, \frac{1}{2} \right]^d \times \mathcal{E}} \Psi(x, \bar{x}) \left(\int \Gamma(\theta_2, \omega, \bar{\theta}, \bar{\omega}) \, m_{1,t}^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta}) - \int \Gamma(\theta_2, \omega, \bar{\theta}, \bar{\omega}) \, m_{2,t}^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta}) \right) \, \mathrm{d}\bar{x}\mu(\mathrm{d}\bar{\omega}) \right\|$$

$$\leqslant \left(\int_{\left[-\frac{1}{2}, \frac{1}{2} \right]^d} \Psi(x, \bar{x})^q \, \mathrm{d}\bar{x} \right)^{\frac{1}{q}} \cdot$$

$$\left(\iint_{\left[-\frac{1}{2}, \frac{1}{2} \right]^d \times \mathcal{E}} \left\| \int \Gamma(\theta_2, \omega, \bar{\theta}, \bar{\omega}) \, m_{1,t}^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta}) - \int \Gamma(\theta_2, \omega, \bar{\theta}, \bar{\omega}) \, m_{2,t}^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta}) \right\|^p \, \mathrm{d}\bar{x}\mu(\mathrm{d}\bar{\omega}) \right)^{\frac{1}{p}}.$$

Note that the first term in the last inequality is always bounded: it is straightforward in the P-nearest neighbor case and comes from Remark 2.17 in the power-law case. Indeed, q has been precisely chosen so that $q\alpha < d$, so that $\Psi(x, \cdot)^q$ is integrable.

Using the Lipschitz-continuity of Γ, we see that, for *any* coupling $m^{\omega,x}(\mathrm{d}\vartheta_1, \mathrm{d}\vartheta_2)$ of $m_1^{\omega,x}$ and $m_2^{\omega,x}$

$$\delta\Gamma_2 \leqslant C \parallel \Gamma \parallel_{Lip} \left(\int_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d \times \mathcal{E}} \left(\mathbf{E}_{m^{\omega,x}} \parallel \vartheta_1(t) - \vartheta_2(t) \parallel \right)^p \mathrm{d}\bar{x}\mu(\mathrm{d}\bar{\omega}) \right)^{\frac{1}{p}},$$

$$\leqslant C \parallel \Gamma \parallel_{Lip} \left(\int_{\left[-\frac{1}{2}, \frac{1}{2}\right]^d \times \mathcal{E}} \left([\mathbf{E}_{m^{\omega,x}} \parallel \vartheta_1(t) - \vartheta_2(t) \parallel^\kappa]^{\frac{1}{\kappa}} \right)^p \mathrm{d}\bar{x}\mu(\mathrm{d}\bar{\omega}) \right)^{\frac{1}{p}}.$$

By Definition 3.1, this gives $\delta\Gamma_2 \leqslant C \parallel \Gamma \parallel_{Lip} \delta_t(m_1, m_2)$, which proves (3.6). We are now in position to prove (3.5). Let us consider (θ_1, ω, x) and (θ_2, ω, x) solutions to (3.3) for two different measures m_1 and m_2 in $\mathcal{M}$ driven by the same Brownian motion, with the same initial condition. We have for all $0 \leqslant t \leqslant T$,

$$\parallel \theta_1(t) - \theta_2(t) \parallel^2 = 2 \int_0^t \langle \theta_1(s) - \theta_2(s), c(\theta_1(s), \omega) - c(\theta_2(s), \omega) \rangle \mathrm{d}s$$

$$+ 2 \int_0^t \left\langle \theta_1(s) - \theta_2(s), \int \Gamma(\theta_1(s), \omega, \cdot, \cdot) \Psi(x, \cdot) \mathrm{d}m_1 - \int \Gamma(\theta_2(s), \omega, \cdot, \cdot) \Psi(x, \cdot) \mathrm{d}m_2 \right\rangle \mathrm{d}s.$$

Using the one-sided Lipschitz condition (2.3) and (3.6), we obtain

$$\parallel \theta_1(t) - \theta_2(t) \parallel^2 \leqslant C \int_0^t \parallel \theta_1(s) - \theta_2(s) \parallel^2 \mathrm{d}s + C \int_0^t \parallel \theta_1(s) - \theta_2(s) \parallel \delta_s(m_1, m_2) \mathrm{d}s,$$

$$\leqslant C \int_0^t \parallel \theta_1(s) - \theta_2(s) \parallel^2 \mathrm{d}s + C \int_0^t \delta_s(m_1, m_2)^2 \mathrm{d}s.$$

Consequently, using Gronwall's Lemma

$$\sup_{s \leqslant t} \parallel \theta_1(s) - \theta_2(s) \parallel^2 \leqslant C e^{CT} \int_0^t \delta_s(m_1, m_2)^2 \mathrm{d}s.$$

Elevating this inequality to the power $\frac{\kappa}{2} \geqslant 1$ gives

$$\sup_{s \leqslant t} \parallel \theta_1(s) - \theta_2(s) \parallel^\kappa \leqslant \left(C e^{CT} \right)^{\frac{\kappa}{2}} \left(\int_0^t \delta_s(m_1, m_2)^2 \mathrm{d}s \right)^{\frac{\kappa}{2}}$$

$$\leqslant \left(C e^{CT} \right)^{\frac{\kappa}{2}} T^{\frac{\kappa-2}{2}} \int_0^t \delta_s(m_1, m_2)^\kappa \mathrm{d}s,$$

which gives

$$\delta_{\mathcal{X}}^{(t)}(\Theta(m_1)^{\omega,x}, \Theta(m_2)^{\omega,x}) \leqslant \left(C e^{CT} \right)^{\frac{1}{2}} T^{\frac{\kappa-2}{2\kappa}} \left(\int_0^t \delta_s(m_1, m_2)^\kappa \mathrm{d}s \right)^{\frac{1}{\kappa}}.$$

Elevating this inequality to the power p and integrating over ω and x leads to the desired result (3.5). Lemma 3.2 is proved. $\qquad\square$

We are now in position to prove Proposition 2.9.

PROOF OF PROPOSITION 2.9. It remains to prove that if $\bar{\nu}$ is a fixed point of Θ then $\bar{\nu}$ is a solution to the weak formulation of the continuous limit (2.17). Indeed if $\bar{\nu} = \Theta(\bar{\nu})$, one can write $\bar{\nu}(\,\mathrm{d}\theta,\,\mathrm{d}\omega,\,\mathrm{d}x) = \bar{\nu}^{\omega,x}(\,\mathrm{d}\theta)\mu(\,\mathrm{d}\omega)\,\mathrm{d}x$ where, for fixed ω, x, $\bar{\nu}^{\omega,x}(\,\mathrm{d}\theta)$ is the law of the process solution to (3.3). Applying Ito's formula, one obtains for all $f(\theta,\omega,x)$, $\mathcal{C}^2$ w.r.t. θ with bounded derivatives,

(3.8)

$$
f(\theta(t),\omega,x) = f(\theta_0,\omega,x) + \frac{1}{2}\int_0^t \mathrm{div}_\theta\left(\sigma\sigma^T\nabla_\theta f\right)(\theta(s),\omega,x)\,\mathrm{d}s + \int_0^t \nabla_\theta f\cdot c(\theta(s),\omega)\,\mathrm{d}s
$$
$$
+ \int_0^t \nabla_\theta f\cdot\int\Gamma(\theta(t),\omega,\bar\theta,\bar\omega)\Psi(x,\bar x)\bar\nu_t^{\bar\omega,\bar x}(\,\mathrm{d}\bar\theta)\mu(\,\mathrm{d}\bar\omega)\,\mathrm{d}\bar x\,\mathrm{d}s + \int_0^t \nabla_\theta f(\theta(s),\omega,x)\cdot(\sigma\,\mathrm{d}B_s).
$$

Taking the expectation in (3.8) leads to (2.17). But in order to do so, we need to know that the term $\nabla_\theta f(\theta,\omega,x)\cdot c(\theta,\omega)$ is integrable w.r.t. the measure $\bar\nu^{\omega,x}(\,\mathrm{d}\theta)\mu(\,\mathrm{d}\omega)\,\mathrm{d}x$ (the other terms are integrable, by assumptions on f). This is ensured by (2.5), the fact that (by construction) $\bar\nu^{\omega,x}(\,\mathrm{d}\theta)$ has finite moments up to order κ, and the fact that μ has finite moment of order ι (recall (2.6)). $\square$

The rest of the document is devoted to provide a proof for Theorems 2.13 and 2.18.

4. Definition and properties of the propagator. For reasons that will be made clear in Remark 4.2 below, we make in this section, as well as in Sections 5 and 6 some supplementary assumption on the regularity on the dynamics c:

ASSUMPTION 4.1 (Additional regularity on c). *We assume that for all ω, the function $\theta\mapsto c(\theta,\omega)$ is globally Lispchitz continuous.*

Of course, the FitzHugh-Nagumo case does not enter into the framework of Assumption 4.1. Assumption 4.1 is made in order to ensure the existence of a backward Kolmogorov equation (see Remark 4.2). The purpose of Section 7 will be to discard this assumption.

In this section, the function Ψ is either defined as in Hypothesis (H1) or as in Hypothesis (H2). We know from Proposition 2.9 that there exists at least one measure-valued solution $t\mapsto\nu_t$ to the continuous equation (2.17). We fix once and for all one such solution. We can then consider the stochastic differential equation:

(4.1)
$$
\mathrm{d}\theta(t) = c(\theta(t),\omega)\,\mathrm{d}t + \int\Gamma(\theta(t),\omega,\bar\theta,\bar\omega)\Psi(x,\bar x)\nu_t(\,\mathrm{d}\bar\theta,\,\mathrm{d}\bar\omega,\,\mathrm{d}\bar x)\,\mathrm{d}t + \sigma\cdot\mathrm{d}B(t)
$$
$$
=: c(\theta(t),\omega)\,\mathrm{d}t + v(t,\theta(t),\omega,x)\,\mathrm{d}t + \sigma\cdot\mathrm{d}B(t),
$$

where $\theta(0)\sim\zeta$. Thanks to the regularity properties of Γ and c and to the integrability of Ψ, (4.1) has a unique solution. Define the propagator corresponding to (4.1):

(4.2)
$$
\forall s,t\in[0,T],\ P_{s,t}f(\theta,\omega,x) := \mathbf{E}_B f(\Phi_s^t(\theta;\omega,x),\omega,x),
$$

where $\mathbf{E}_B$ is the expectation w.r.t. the Brownian motion B, f is a bounded measurable function on $\mathcal{X}\times\mathcal{E}\times\left[-\frac{1}{2},\frac{1}{2}\right]^d$, $0\leqslant s\leqslant t$ and $t\mapsto\Phi_s^t(\theta;\omega,x)$ is the unique solution to (4.1) such that $\Phi_s^s(\theta;\omega,x) = \theta$.

REMARK 4.2. *If f is $\mathcal{C}^2$ w.r.t. the variable θ, under Assumptions 2.2 and 4.1 made about c and Γ, it is standard to see that the function $P_{s,t}f$ is of class $\mathcal{C}^2$ in θ and $\mathcal{C}^1$ in s and satisfies the Backward Kolmogorov equation (see for example [32, Remark 2.3]):*

$$(4.3) \quad \forall(\theta,\omega,x,s,t), \ \partial_s P_{s,t}f(\theta,\omega,x) + \frac{1}{2}\mathrm{div}_\theta\left(\sigma\sigma^T\nabla_\theta P_{s,t}\right)(\theta,\omega,x)$$
$$+\left([c(\theta,\omega) + v(t,\theta,\omega,x)]\cdot\nabla_\theta\right)P_{s,t}f(\theta,\omega,x) = 0.$$

The main problem which motivates the work of Section 7 at the end of this paper is that proving similar Kolmogorov when Assumption 4.1 is discarded appears to be difficult (see in particular the recent work in this direction [20]). Nevertheless, we work in this section under this additional hypothesis and we provide in Section 7 a way to bypass this technical difficulty.

The key calculation of this work is the object of Lemma 4.3:

LEMMA 4.3. *Let $f : \mathcal{X} \times \mathcal{E} \times \left[-\frac{1}{2},\frac{1}{2}\right]^d \to \mathbf{R}$, be $\mathcal{C}^2$ w.r.t. the variable θ. Then,*

(4.4)
$$\left\langle f,\, \nu_T^{(N)} - \nu_T\right\rangle = \left\langle P_{0,T}f,\, \nu_0^{(N)} - \nu_0\right\rangle + \frac{1}{|\Lambda_N|}\sum_k\int_0^T \nabla_\theta\left(P_{t,T}f\right)(\theta_k(t),\omega_k,x_k)\cdot\sigma\,\mathrm{d}B_k(t)$$
$$+\frac{1}{|\Lambda_N|}\sum_k\int_0^T \nabla_\theta\left(P_{t,T}f\right)(\theta_k(t),\omega_k,x_k)\cdot\left[\left\langle\Gamma\left(\theta_k,\omega_k,\cdot,\cdot\right)\Psi\left(x_k,\cdot\right),\, \nu_t^{(N)} - \nu_t\right\rangle\right]\mathrm{d}t.$$

PROOF OF LEMMA 4.3. An application of Ito's formula gives: for all k and $0 < t < T$,

$$P_{t,T}f\left(\theta_k(t),\omega_k,x_k\right) = P_{0,T}f\left(\theta_k(0),\omega_k,x_k\right) + \int_0^t \partial_s P_{s,T}f\left(\theta_k(s),\omega_k,x_k\right)\mathrm{d}s$$
$$+\int_0^t\nabla_\theta P_{s,T}f\left(\theta_k(s),\omega_k,x_k\right)\cdot\mathrm{d}\theta_k(s) + \frac{1}{2}\int_0^t\mathrm{div}_\theta\left(\sigma\sigma^T\nabla_\theta P_{s,T}f\right)(\theta_k(s),\omega_k,x_k)\,\mathrm{d}s.$$

Using the definition of θ_k (recall (2.2)) and (4.3) we obtain:

$$P_{t,T}f\left(\theta_k(t),\omega_k,x_k\right) = P_{0,T}f\left(\theta_k(0),\omega_k,x_k\right) - \int_0^t v(s,\theta_k(s),\omega_k,x_k)\cdot\nabla_\theta P_{s,T}f\left(\theta_k(s),\omega_k,x_k\right)\mathrm{d}s$$
$$+\int_0^t\nabla_\theta P_{s,T}f\left(\theta_k(s),\omega_k,x_k\right)\cdot\left\langle\Gamma\left(\theta_k,\omega_k,\cdot,\cdot\right)\Psi\left(x_k,\cdot\right),\, \nu_s^{(N)}\right\rangle\mathrm{d}s$$
$$+\int_0^t\nabla_\theta P_{s,t}f(\theta_k(s),\omega_k,x_k)\cdot(\sigma\,\mathrm{d}B_k(s)).$$

Then, using the definition of $v(\cdot)$ (recall (4.1)) and summing over k lead to:

$$\left\langle P_{t,T}f,\, \nu_t^{(N)}\right\rangle = \left\langle P_{0,T}f,\, \nu_0^{(N)}\right\rangle + \frac{1}{|\Lambda_N|}\sum_k\int_0^t\nabla_\theta P_{s,t}f(\theta_k(s),\omega_k,x_k)\cdot(\sigma\,\mathrm{d}B_k(s))$$
$$+\frac{1}{|\Lambda_N|}\sum_k\int_0^t\nabla_\theta P_{s,T}f\left(\theta_k(s),\omega_k,x_k\right)\cdot\left\langle\Gamma\left(\theta_k,\omega_k,\cdot,\cdot\right)\Psi\left(x_k,\cdot\right),\, \nu_s^{(N)} - \nu_s\right\rangle\mathrm{d}s.$$

A straightforward calculation using (4.3) shows that $\partial_t \langle P_{t,T} f, \nu_t \rangle = 0$. Using this and the previous equality, one obtains the desired result (choose $t = T$ and recall that $P_{T,T} f = f$). Lemma 4.3 is proved. $\qquad\square$

The purpose of the following lemma is to establish regularity properties of the propagator $P_{t,T}$:

LEMMA 4.4 (Estimates on the propagator $P_{t,T}$). *Fix $T > 0$, $0 < t < T$ and $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$.*

1. *Assume Ψ satisfies Hypothesis (H1). For any $R \in (0,1]$ and any f in $\mathcal{C}_{R,a}$, $P_{t,T} f$ is also in $\mathcal{C}_{R,a}$ and one has the following estimate*

$$(4.5) \qquad \| P_{t,T} f \|_{R,a} \leqslant \sqrt{2} e^{\|P\|(T-t)} \| f \|_{R,a},$$

for some constant $\|P\|$ (that can be chosen equal to $L + 3/2 \| \Gamma \|_{Lip}$, recall (2.3)).

2. *Assume Ψ satisfies Hypothesis (H2). For every $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, for any f in $\mathcal{C}_a$, $P_{t,T} f$ is also in $\mathcal{C}_a$ and one has the following estimate*

$$(4.6) \qquad \| P_{t,T} f \|_a \leqslant \|\!|P|\!\| e^{\|\!|P|\!\|(T-t)} \| f \|_a,$$

for some constant $\|\!|P|\!\|$ (that only depends on Γ, Ψ and c).

PROOF. Note that, by a usual density argument, one only needs to prove (4.5) and (4.6) for test functions f that are $\mathcal{C}^2$ w.r.t. θ. Fix $T > 0$, $0 < t < T$, $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$ and consider two different flows for (4.1) $\Phi_s^t(\theta_i; \omega_i, x)$, for $i = 1, 2$, with different initial condition and parameter but at the same site x, with the same Brownian motion. For simplicity, we write $\Phi_s^t(i)$ instead of $\Phi_s^t(\theta_i; \omega_i, x)$. Then, using the one-sided Lipschitz condition (2.3) on c, we obtain

$$\left\| \Phi_s^t(2) - \Phi_s^t(1) \right\|^2 = \| \theta_2 - \theta_1 \|^2 + 2 \int_s^t \left\langle \Phi_s^u(2) - \Phi_s^u(1), c(\Phi_s^u(2), \omega_2) - c(\Phi_s^u(1), \omega_1) \right\rangle \, \mathrm{d}u$$

$$+ 2 \int_s^t \left\langle \Phi_s^u(2) - \Phi_s^u(1), v(u, \Phi_s^u(2), \omega_2, x) - v(u, \Phi_s^u(1), \omega_1, x) \right\rangle \, \mathrm{d}u,$$

$$\leqslant \| \theta_2 - \theta_1 \|^2 + 2L \int_s^t \left(\| \Phi_s^u(2) - \Phi_s^u(1) \|^2 + \| \omega_2 - \omega_1 \|^2 \right) \, \mathrm{d}u$$

$$+ 2 \int_s^t \| \Phi_s^u(2) - \Phi_s^u(1) \| \underbrace{\| v(u, \Phi_s^u(2), \omega_2, x) - v(u, \Phi_s^u(1), \omega_1, x) \|}_{:=\delta v(u)} \, \mathrm{d}u,$$

where the definition of $v(\cdot)$ is given in (4.1). The Lipschitz-continuity of Γ implies

$$\delta v(u) \leqslant \int \| \Gamma(\Phi_s^u(2), \omega_2, \bar{\theta}, \bar{\omega}) - \Gamma(\Phi_s^u(1), \omega_1, \bar{\theta}, \bar{\omega}) \| \Psi(x, \bar{x}) \nu_u^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta}) \mu(\mathrm{d}\bar{\omega}) \, \mathrm{d}\bar{x}$$

$$\leqslant \| \Gamma \|_{Lip} S(\Psi) \left(\| \Phi_s^u(2) - \Phi_s^u(1) \| + \| \omega_2 - \omega_1 \| \right),$$

where $S(\Psi)$ has already been defined in (3.4). Putting things together we see that, for $C = 2L + 3 \| \Gamma \|_{Lip} S(\Psi)$,

$$(4.7) \quad \left\| \Phi_s^t(2) - \Phi_s^t(1) \right\|^2 \leqslant \| \theta_2 - \theta_1 \|^2 + C \int_s^t \left(\| \Phi_s^u(2) - \Phi_s^u(1) \|^2 + \| \omega_2 - \omega_1 \|^2 \right) \, \mathrm{d}u.$$

An application of Gronwall's lemma leads to

(4.8)
$$\left\| \Phi_s^t(\theta_2, \omega_2, x) - \Phi_s^t(\theta_1, \omega_1, x) \right\|^2 + \left\| \omega_2 - \omega_1 \right\|^2 \leqslant e^{C(t-s)}\left(\left\| \theta_2 - \theta_1 \right\|^2 + \left\| \omega_2 - \omega_1 \right\|^2 \right).$$

Then, in the case where Ψ satisfies Hypothesis (H1), we have $P_{t,T}f(\theta, \omega, x) = \chi_R(x - a)g(\Phi_t^T(\theta; \omega, x), \omega)$, when $f(\theta, \omega, x) = \chi_R(x - a)g(\theta, \omega)$. But then,

$$
\begin{aligned}
\left\| g(\Phi_t^T(\theta_2; \omega_2, x), \omega_2) - g(\Phi_t^T(\theta_1; \omega_1, x), \omega_1) \right\|^2 &\leqslant \left\| f \right\|_{R,a}^2 \left(\left\| \Phi_t^T(2) - \Phi_t^T(1) \right\| + \left\| \omega_2 - \omega_1 \right\| \right)^2, \\
&\leqslant 2 \left\| f \right\|_{R,a}^2 \left(\left\| \Phi_t^T(2) - \Phi_t^T(1) \right\|^2 + \left\| \omega_2 - \omega_1 \right\|^2 \right), \\
&\leqslant 2 \left\| f \right\|_{R,a}^2 e^{C(T-t)}\left(\left\| \theta_2 - \theta_1 \right\|^2 + \left\| \omega_2 - \omega_1 \right\|^2 \right),
\end{aligned}
$$

so that

$$\left\| g(\Phi_t^T(\theta_2; \omega_2, x), \omega_2) - g(\Phi_t^T(\theta_1; \omega_1, x), \omega_1) \right\| \leqslant \sqrt{2} \left\| f \right\|_{R,a} e^{\frac{C}{2}(T-t)}\left(\left\| \theta_2 - \theta_1 \right\| + \left\| \omega_2 - \omega_1 \right\| \right),$$

which is the desired estimate (2.19) and gives (4.5). The same kind of calculation in the case of Hypothesis (H2) leads to the estimate (2.22) for $P_{t,T}f$.

Thus, it remains to prove estimates (2.23) and (2.24) for $P_{t,T}f$ in the case of Hypothesis (H2). The case of (2.23) is straightforward. As far as (2.24) is concerned, the same kind of calculation with two different flows $\Phi_s^t(x) := \Phi_s^t(\theta; \omega, x)$ and $\Phi_s^t(y) := \Phi_s^t(\theta; \omega, y)$, with the same θ and ω but at different sites x and y leads to

$$
\begin{aligned}
\left\| \Phi_s^t(x) - \Phi_s^t(y) \right\|^2 \leqslant{}& 2L \int_s^t \left\| \Phi_s^u(x) - \Phi_s^u(y) \right\|^2 \mathrm{d}u \\
&+ 2\int_s^t \left\| \Phi_s^u(x) - \Phi_s^u(y) \right\| \underbrace{\left\| v(u, \Phi_s^u(x), \omega, x) - v(u, \Phi_s^u(y), \omega, y) \right\|}_{:= \delta v(u, x, y)} \mathrm{d}u,
\end{aligned}
$$

with,

$$
\begin{aligned}
\delta v(u, x, y) \leqslant{}& \int \left\| \Gamma(\Phi_s^u(x), \omega, \bar{\theta}, \bar{\omega})\Psi(x, \bar{x}) - \Gamma(\Phi_s^u(y), \omega, \bar{\theta}, \bar{\omega})\Psi(y, \bar{x}) \right\| \nu_u^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta})\mu(\mathrm{d}\bar{\omega})\mathrm{d}\bar{x} \\
\leqslant{}& \int \left\| \Gamma(\Phi_s^u(x), \omega, \bar{\theta}, \bar{\omega}) - \Gamma(\Phi_s^u(y), \omega, \bar{\theta}, \bar{\omega}) \right\| \Psi(x, \bar{x})\nu_u^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta})\mu(\mathrm{d}\bar{\omega})\mathrm{d}\bar{x} \\
&+ \int \left\| \Gamma(\Phi_s^u(y), \omega, \bar{\theta}, \bar{\omega}) \right\| |\Psi(x, \bar{x}) - \Psi(y, \bar{x})|\nu_u^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta})\mu(\mathrm{d}\bar{\omega})\mathrm{d}\bar{x} \\
\leqslant{}& \left\| \Gamma \right\|_{Lip} S(\Psi)\left(\left\| \Phi_s^u(x) - \Phi_s^u(y) \right\| \right) \\
&+ \left\| \Gamma \right\|_\infty \int_{[-1,1]^d} |\Psi(x, \bar{x}) - \Psi(y, \bar{x})| \underbrace{\int_{\mathcal{X} \times \mathcal{E}} \nu_u^{\bar{\omega}, \bar{x}}(\mathrm{d}\bar{\theta})\mu(\mathrm{d}\bar{\omega})}_{=1} \mathrm{d}\bar{x} \\
\leqslant{}& \left\| \Gamma \right\|_{Lip} S(\Psi)\left\| \Phi_s^u(x) - \Phi_s^u(y) \right\| + \mathcal{I}_2(\Psi)\left\| \Gamma \right\|_\infty \left\| x - y \right\|^{(d-\alpha)\wedge 1},
\end{aligned}
$$

where $S(\Psi)$ is defined in (3.4) and where we used assumption (2.9). This gives, for $C = 2L + 2\left\| \Gamma \right\|_{Lip} S(\Psi) + \mathcal{I}_2(\Psi)\left\| \Gamma \right\|_\infty$,

$$\left\| \Phi_s^t(x) - \Phi_s^t(y) \right\|^2 \leqslant C\int_s^t \left\| \Phi_s^u(x) - \Phi_s^u(y) \right\|^2 \mathrm{d}u + \mathcal{I}_2(\Psi)\left\| \Gamma \right\|_\infty (t-s)\left\| x - y \right\|^{2((d-\alpha)\wedge 1)},$$

Consequently, by Gronwall's lemma,

$$(4.9) \qquad \left\| \Phi_s^t(\theta; \omega, x) - \Phi_s^t(\theta; \omega, y) \right\|^2 \leqslant \mathcal{I}_2(\Psi) \left\| \Gamma \right\|_\infty (t-s) e^{C(t-s)} \left\| x - y \right\|^{2((d-\alpha)\wedge 1)}.$$

Then, for any $0 < t \leqslant T$, we have

$$\left\| \delta P_{t,T} f \right\|^2 := \left\| \left\| x - a \right\|^{2\gamma} P_{t,T} f(\theta, \omega, x) - \left\| y - a \right\|^{2\gamma} P_{t,T} f(\theta, \omega, y) \right\|^2$$

$$= \left\| \left\| x - a \right\|^{2\gamma} f(\Phi_t^T(\theta; \omega, x), \omega, x) - \left\| y - a \right\|^{2\gamma} f(\Phi_t^T(\theta; \omega, y), \omega, y) \right\|^2,$$

$$\leqslant \left(\left\| x - a \right\|^{2\gamma} \left\| f(\Phi_t^T(\theta; \omega, x), \omega, x) - f(\Phi_t^T(\theta; \omega, y), \omega, x) \right\| \right.$$

$$\left. + \left\| \left\| x - a \right\|^{2\gamma} f(\Phi_t^T(\theta; \omega, y), \omega, x) - \left\| y - a \right\|^{2\gamma} f(\Phi_t^T(\theta; \omega, y), \omega, y) \right\| \right)^2,$$

$$(4.10)$$

$$\leqslant \left\| f \right\|_a^2 \left(\left\| \Phi_t^T(x) - \Phi_t^T(y) \right\| + \left\| x - y \right\|^{(2\gamma-\alpha)\wedge 1} \right)^2,$$

$$\leqslant 2 \left\| f \right\|_a^2 \left(\left\| \Phi_t^T(x) - \Phi_t^T(y) \right\|^2 + \left\| x - y \right\|^{2((2\gamma-\alpha)\wedge 1)} \right),$$

$$(4.11)$$

$$\leqslant 2 \left\| f \right\|_a^2 \left(\mathcal{I}_2(\Psi) \left\| \Gamma \right\|_\infty (T-t) \vee 1 \right) e^{C(T-t)} \left(\left\| x - y \right\|^{2((d-\alpha)\wedge 1)} + \left\| x - y \right\|^{2((2\gamma-\alpha)\wedge 1)} \right),$$

where we used assumptions (2.23) and (2.24) in (4.10) and the estimation (4.9) in (4.11). Using the definition of γ (recall (2.11)), it is always true that $d - \alpha \geqslant 2\gamma - \alpha$. Consequently,

$$\left\| \left\| x - a \right\|^{2\gamma} P_{t,T} f(\theta, \omega, x) - \left\| y - a \right\|^{2\gamma} P_{t,T} f(\theta, \omega, y) \right\|$$

$$\leqslant 2 \left(T \mathcal{I}_2(\Psi) \left\| \Gamma \right\|_\infty \vee 1 \right)^{\frac{1}{2}} e^{\frac{C}{2}(T-t)} \left\| f \right\|_a \left\| x - y \right\|^{(2\gamma-\alpha)\wedge 1},$$

which leads to (2.24). Lemma 4.4 is proved. $\qquad\square$

REMARK 4.5. *One could wonder why we have not simply used in the calculation above the global Lipschitz assumption about c (recall Assumption 4.1), instead of the more involved one-sided Lipschitz inequality used here. The crucial reason for this is that in order to be able to discard Assumption 4.1 in Section 7 below, we need to ensure that the estimates of Lemma 4.4 do not depend on the modulus of continuity of c, but only on its one-sided Lipschitz constant L.*

Using (4.5) (respectively (4.6)) in (4.4), we easily see that for every $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$, for any given $f \in \mathcal{C}_{R,a}$ with $\left\| f \right\|_{R,a} \leqslant 1$ (respectively $f \in \mathcal{C}_a$ with $\left\| f \right\|_a \leqslant 1$), we have

$$(4.12) \quad \left\| \left\langle f, \nu_T^{(N)} \right\rangle - \left\langle f, \nu_T \right\rangle \right\| \leqslant \left\| \left\langle P_{0,T} f, \nu_0^{(N)} \right\rangle - \left\langle P_{0,T} f, \nu_0 \right\rangle \right\|$$

$$+ \left\| \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f \right)(\theta_k(t), \omega_k, x_k) \cdot (\sigma \, \mathrm{d}B_k(t)) \right\|$$

$$+ \frac{1}{|\Lambda_N|} \sum_k \int_0^T \left\| \nabla_\theta P_{t,T} f \right\| \left\| \left\langle \Gamma(\theta_k, \omega_k, \cdot, \cdot) \Psi(x_k, \cdot), \nu_t^{(N)} - \nu_t \right\rangle \right\| \mathrm{d}t.$$

Using (2.20) and (4.5) (resp. (2.26) and (4.6)), the term $\|\nabla_\theta P_{t,T} f\|(\theta_k(t), \omega_k, x_k)$ in the third summand of (4.12) can be bounded by $\sqrt{2} e^{\|P\|(T-t)} \|\chi_R\|_\infty$ in case of Hypothesis (H1) and by $\|x_k - a\|^{-\alpha} \|P\| e^{\|P\|(T-t)}$ in case of Hypothesis (H2). In both cases, the bound that we find can be written in the form

$$(4.13) \qquad \|\nabla_\theta P_{t,T} f\|(\theta_k(t), \omega_k, x_k) \leqslant e^{\|P\|(T-t)} \rho(x_k)$$

(ρ is a constant in the first case and proportional to $\|x_k - a\|^{-\alpha}$ in the second). In particular, it is uniform in f and (θ_k, ω_k). Let us now fix the integer p equal to 2 in the case of Hypothesis (H1) or defined as in (2.28) in the case of Hypothesis (H2). Elevating inequality (4.12) to the power p and taking the expectation lead to

$$
\begin{aligned}
(4.14) \quad & \frac{1}{3^{p-1}} \mathbf{E} \left\| \left\langle f, \nu_T^{(N)} - \nu_T \right\rangle \right\|^p \leqslant \mathbf{E} \left\| \left\langle P_{0,T} f, \nu_0^{(N)} - \nu_0 \right\rangle \right\|^p \\
& + \mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f\right)(\theta_k(t), \omega_k, x_k) \cdot (\sigma \, \mathrm{d}B_k(t)) \right\|^p \\
& + \mathbf{E} \left| \frac{1}{|\Lambda_N|} \sum_k \int_0^T e^{\|P\|(T-t)} \rho(x_k) \left\| \left\langle \Gamma(\theta_k, \omega_k, \cdot, \cdot) \Psi(x_k, \cdot), \nu_t^{(N)} - \nu_t \right\rangle \right\| \, \mathrm{d}t \right|^p.
\end{aligned}
$$

Let us concentrate on the third term of the last inequality, that we denote by D_N. By successive use of Hölder's inequality (recall that $\frac{1}{p} + \frac{1}{q} = 1$), one has:

$$
\begin{aligned}
(4.15) \quad D_N & \leqslant \left(\int_0^T e^{q\|P\|(T-t)} \, \mathrm{d}t \right)^{\frac{p}{q}} \mathbf{E} \int_0^T \left| \frac{1}{|\Lambda_N|} \sum_k \rho(x_k) \left\| \left\langle \Gamma(\theta_k, \omega_k, \cdot, \cdot) \Psi(x_k, \cdot), \nu_t^{(N)} - \nu_t \right\rangle \right\| \right|^p \, \mathrm{d}t \\
& \leqslant \left(\frac{e^{q\|P\|T} - 1}{q\|P\|} \right)^{\frac{p}{q}} \left(\frac{1}{|\Lambda_N|} \sum_k \rho(x_k)^q \right)^{\frac{p}{q}} \int_0^T \frac{1}{|\Lambda_N|} \sum_k \mathbf{E} \left\| \left\langle \Gamma(\theta_k, \omega_k, \cdot, \cdot) \Psi(x_k, \cdot), \nu_t^{(N)} - \nu_t \right\rangle \right\|^p \, \mathrm{d}t.
\end{aligned}
$$

At this point, here are the main steps of proof that we will follow in the remaining of this paper: we have built the spaces of test functions (recall Definitions 2.10 and 2.14) in such a way that they precisely include the functions $(\theta, \omega, x) \mapsto \Gamma(\theta_k, \omega_k, \theta, \omega) \Psi(x_k, x)$ for all k (in this case, a is equal to x_k). Since the distances between two random measures introduced in Definitions 2.12 and 2.16 are exactly the suprema of evaluations over all such test functions, we are thus able to bound the term within the integral in (4.15) in terms of the distance between $\nu^{(N)}$ and ν.

The second point of the proof is to obtain an estimate (uniform in f) of the speed of convergence to 0 of the two first terms in (4.14). Taking the supremum over all test functions f and applying Gronwall's Lemma lead to the conclusion.

Those steps are somehow easy to follow in the P-nearest neighbor case (see Section 5) but are more technically demanding in the power-law case (see Section 6).

5. Law of Large Numbers in the P-nearest neighbor case. The purpose of this section is to prove Theorem 2.13. Thus, throughout this section, we suppose that Ψ satisfies Hypothesis (H1) for some $R \in (0, 1]$. In this case, the integer p introduced in (4.14) is equal to 2 and the function ρ in (4.13) is bounded (equal to $\sqrt{2} \|\chi_R\|_\infty$). In particular,

the two terms in front of the integral in (4.15) are trivially bounded by a constant, equal to $\frac{e^{2\|P\|T}-1}{2\|P\|}\,\|\chi_R\|_\infty^2$.

The following proposition proves the convergence to 0 of the first term in (4.14) together with explicit rates:

PROPOSITION 5.1 (Convergence of the initial condition). *There exists a numerical constant $C_1 > 0$ (independent of R) such that for all $f \in \bigcup_{a\in\left[-\frac{1}{2},\frac{1}{2}\right]^d} \mathcal{C}_{R,a}$ with $\|f\|_{R,a} \leqslant 1$ and $\|f\|_\infty \leqslant 1$*

$$(5.1) \qquad \mathbf{E}\left\|\left\langle P_{0,T}f,\, \nu_0^{(N)}\right\rangle - \left\langle P_{0,T}f,\, \nu_0\right\rangle\right\|^2 \leqslant \frac{C_1}{N^{d\wedge 2}}.$$

PROOF OF PROPOSITION 5.1. Recall that the couples $(\theta_i(0),\omega_i)_{1\leqslant i\leqslant N}$ are supposed to be i.i.d. samples of the law $\zeta(\mathrm{d}\theta)\otimes\mu(\mathrm{d}\omega)$ on $\mathcal{X}\times\mathcal{E}$. Let $f\in\mathcal{C}_{R,a}$: by definition, $f(\theta,\omega,x) = g(\theta,\omega)\chi_R(x-a)$ so that $P_{0,T}f = \chi(x-a)P_{0,T}g$. Let write $\varphi := P_{0,T}g$ for simplicity. Then:

$$\delta_N(f) := \mathbf{E}\left\|\left\langle P_{0,T}f,\, \nu_0^{(N)}\right\rangle - \left\langle P_{0,T}f,\, \nu_0\right\rangle\right\|^2$$

$$= \mathbf{E}\left\|\frac{1}{|\Lambda_N|}\sum_j \varphi(\theta_j,\omega_j)\chi_R(x_j-a) - \int\varphi(\theta,\omega)\chi_R(x-a)\zeta(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\,\mathrm{d}x\right\|^2$$

$$\leqslant 2\mathbf{E}\left\|\chi_R(x_j-a)\frac{1}{|\Lambda_N|}\sum_j\left(\varphi(\theta_j,\omega_j) - \int\varphi(\theta,\omega)\zeta(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\right)\right\|^2$$

$$+ 2\left\|\int\varphi(\theta,\omega)\zeta(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\left(\frac{1}{|\Lambda_N|}\sum_j\chi_R(x_j-a) - \int\chi_R(x-a)\,\mathrm{d}x\right)\right\|^2$$

$$\leqslant \frac{2}{(2R)^{2d}}\mathbf{E}\left\|\frac{1}{|\Lambda_N|}\sum_j\left(\varphi(\theta_j,\omega_j) - \int\varphi(\theta,\omega)\zeta(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\right)\right\|^2$$

$$+ 2\|\varphi\|_\infty^2\left|\frac{1}{|\Lambda_N|}\sum_j\chi_R(x_j-a) - \int\chi_R(x-a)\,\mathrm{d}x\right|^2$$

$$(5.2)\qquad := A_N + B_N.$$

Since the (θ_i,ω_i) are i.i.d. random variables (with law $\zeta\otimes\mu$), a standard calculation shows

$$A_N = \frac{2}{|\Lambda_N|^2\,(2R)^{2d}}\sum_j\mathbf{E}\left\|\varphi(\theta_j,\omega_j) - \int\varphi(\theta,\omega)\zeta(\mathrm{d}\theta)\mu(\mathrm{d}\omega)\right\|^2 \leqslant \frac{8\|f\|_\infty^2}{2^dN^d},$$

since $\|\varphi\|_\infty = \|P_{0,T}g\|_\infty = (2R)^d\|f\|_\infty$ and $|\Lambda_N| = (2N+1)^d \geqslant (2N)^d$.

Let us now turn to the case of the term B_N in (5.2). We place ourselves in the case of non-periodic boundary condition (recall Remark 2.1). The periodic case is simpler and left to the reader. Let $a = (a_1,\ldots,a_d)$. One has

$$(5.3)\qquad \int_{\left[-\frac{1}{2},\frac{1}{2}\right]^d}\chi_R(x-a)\,\mathrm{d}x = \prod_{l=1}^d\left(\frac{1}{2R}\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathbb{1}_{|x-a_l|\leqslant R}\,\mathrm{d}x\right) := \prod_{l=1}^d\mathcal{I}(a_l).$$

In the same way,

$$\frac{1}{|\Lambda_N|} \sum_j \chi_R(x_j - a) = \prod_{l=1}^{d} \left(\frac{1}{2R(2N+1)} \sum_{j=-N}^{N} \mathbb{1}_{|x_j - a_l| \leqslant R} \right) := \prod_{l=1}^{d} \mathcal{I}_N(a_l).$$

Then, from the obvious equality

$$\prod_{l=1}^{d} \mathcal{I}_N(a_l) - \prod_{l=1}^{d} \mathcal{I}(a_l) = \sum_{k=1}^{d} \mathcal{I}_N(a_1) \ldots \mathcal{I}_N(a_{k-1}) \left(\mathcal{I}_N(a_k) - \mathcal{I}(a_k) \right) \mathcal{I}(a_{k+1}) \ldots \mathcal{I}(a_d).$$

and a recursion argument, one only needs to consider the case $d = 1$ in order to prove (5.1). An easy calculation shows the following: for all $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]$, for all $R \in (0, 1]$,

$$\mathcal{I}(a) = \frac{1}{2R} \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathbb{1}_{|x-a| \leqslant R} \, dx = \begin{cases} \frac{1}{2R} \left(R + \frac{1}{2} + a \right), & \text{if } -\frac{1}{2} \leqslant a \leqslant -\frac{1}{2} + R, \\ 1 & \text{if } -\frac{1}{2} + R \leqslant a \leqslant \frac{1}{2} - R, \\ \frac{1}{2R} \left(R + \frac{1}{2} - a \right), & \text{if } \frac{1}{2} - R \leqslant a \leqslant \frac{1}{2}. \end{cases}$$

Thus, in the one-dimensional case, we need to distinguish three cases, depending on the position of $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]$ w.r.t. R; we only treat the case $-\frac{1}{2} \leqslant a \leqslant -\frac{1}{2} + R$, the two others being similar and left to the reader. In this case, one has successively,

$$|\mathcal{I}_N(a) - \mathcal{I}(a)|^2 = \frac{1}{4R^2} \left| \frac{1}{2N+1} \sum_{j=-N}^{N} \mathbb{1}_{|j-2aN| \leqslant 2RN} - \left(R + \frac{1}{2} + a \right) \right|^2$$

$$= \frac{1}{4R^2} \left| \frac{1}{2N+1} \left(\lfloor 2N(R+a) \rfloor + N \right) - \left(R + \frac{1}{2} + a \right) \right|^2$$

$$\leqslant \frac{(R+a)^2}{4R^2(1+2N)^2} \leqslant \frac{(2R - 1/2)^2}{16R^2 N^2} \leqslant \frac{1}{4N^2}.$$

Proposition 5.1 is proved. $\square$

We are now in position to prove Theorem 2.13:

PROOF OF THEOREM 2.13. Fix some $a \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d$ and some $f \in \mathcal{C}_{R,a}$ such that $\| f \|_{R,a} \leqslant 1$ and $\| f \|_\infty \leqslant 1$. Let us first give an estimate of the second term in (4.12). Recall that B_k is a Brownian motion in $\mathcal{X} = \mathbf{R}^m$ so that B_k may be written as m i.i.d. Brownian motions $\left(B_k^{(1)}, \ldots, B_k^{(m)} \right)$. Then, using (2.20) (recall Remark 2.11) in (5.4) and using (4.5) (recall Lemma 4.4) in (5.5)

$$\mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f \right) (\theta_k(t), \omega_k, x_k) \cdot dB_k(t) \right\|^2 = \frac{1}{|\Lambda_N|^2} \sum_k \sum_{l=1}^{m} \mathbf{E} \int_0^T \partial_{\theta^{(l)}} \left(P_{t,T} f \right)^2 \, dt$$

$$(5.4) \hspace{6cm} \leqslant \frac{m \| \chi_R \|_\infty^2}{|\Lambda_N|} \int_0^T \| P_{t,T} f \|_{R,a}^2 \, dt$$

$$(5.5) \hspace{6cm} \leqslant \frac{m \| \chi_R \|_\infty^2}{|\Lambda_N|} 2 \int_0^T e^{2\|P\|(T-t)} \, dt$$

$$(5.6) \hspace{6cm} = \frac{m(e^{2\|P\|T} - 1)}{(2R)^{2d} |\Lambda_N|} \leqslant \frac{C_2}{N^d}.$$

where $C_2 = \frac{m(e^{2\|P\|T}-1)}{8^d R^{2d}}$ and where $\|P\|$ is defined by (4.5).

Let us now give an estimate of the term D_N in (4.15): by Definition 2.10, due to the assumptions made on Γ, it is easy to see that for fixed k the function $f_k := \Gamma(\theta_k, \omega_k, \cdot, \cdot) \Psi(x_k, \cdot)$ belongs to $\mathcal{C}_{R,x_k}$ with norm $\| f_k \|_{R,x_k} = \| \Gamma \|_{Lip}$. Consequently, by construction of the distance d_R (recall Definition 2.12), one has the following:

$$\forall t > 0, \ \mathbf{E} \left\| \left\langle \Gamma(\theta_k, \omega_k, \cdot, \cdot) \Psi(x_k, \cdot) \, , \, \nu_t^{(N)} - \nu_t \right\rangle \right\|^2 \leqslant \| \Gamma \|_{Lip}^2 \, d_R(\nu_t^{(N)}, \nu_t)^2.$$

Putting together (4.14), (5.1) and (5.6), we obtain finally

$$\mathbf{E} \left\| \left\langle f \, , \, \nu_T^{(N)} - \nu_T \right\rangle \right\|^2 \leqslant 3\frac{C_1}{N^{2 \wedge d}} + 3\frac{C_2}{N^d} + 3\frac{e^{2\|P\|T}-1}{(2R)^{2d}\|P\|} \| \Gamma \|_{Lip}^2 \int_0^T d_R(\nu_t^{(N)}, \nu_t)^2 \, \mathrm{d}t.$$

Taking the supremum over all functions f in $\bigcup_{a \in [-1,1]^d} \mathcal{C}_{R,a}$ and applying Gronwall's lemma leads to the result. Theorem 2.13 is proved. $\square$

6. Law of Large Numbers in the power-law case. We suppose in this section that the weight Ψ satisfies Hypothesis (H2).

Let us begin with a technical lemma that will be of constant use throughout this part:

LEMMA 6.1. *There exists a constant $C_0 > 0$ (that only depends on β), such that for all $N, K \geqslant 1$, for all $a \in \mathcal{D}_K$,*

1. for all $0 < \beta < d$, one has

$$(6.1) \qquad \sum_{j; \ j/N \neq a} \left\| \frac{j}{2N} - a \right\|^{-\beta} \leqslant C_0 \begin{cases} N^d K^d & \text{if } a \notin \mathcal{D}_N \\ N^d & \text{if } a \in \mathcal{D}_N. \end{cases}$$

2. for $\beta = d$, one has

$$(6.2) \qquad \sum_{j; \ j/N \neq a} \left\| \frac{j}{2N} - a \right\|^{-d} \leqslant C_0 \begin{cases} K^d N^d \ln N & \text{if } a \notin \mathcal{D}_N \\ N^d \ln N & \text{if } a \in \mathcal{D}_N. \end{cases}$$

3. for all $\beta > d$, one has

$$(6.3) \qquad \sum_{j; \ j/N \neq a} \left\| \frac{j}{2N} - a \right\|^{-\beta} \leqslant C_0 \begin{cases} N^\beta K^\beta & \text{if } a \notin \mathcal{D}_N \\ N^\beta & \text{if } a \in \mathcal{D}_N. \end{cases}$$

REMARK 6.2. *The estimates given in Lemma 6.1 in the case $a \in \mathcal{D}_N$ are standard and optimal. The main technical problem of Lemma 6.1 lies in the case of $a \notin \mathcal{D}_N$: in this case, the point a of the discretization $\mathcal{D}_K$ can be arbitrarily close to one point $\frac{j}{2N}$ in the above sum. Those points belong to the discretization $\mathcal{D}_N$. The minimal distance between a and the discretization $\mathcal{D}_N$ depends on K (actually it depends on the greatest common divisor of K and N, see the proof of Lemma 6.1). This explains the dependence in K of the estimations of Lemma 6.1.*

The proof of Lemma 6.1 is postponed to the appendix. Lemma 6.1 will be at the basis of most of the estimations in this section.

Theorem 2.18 is a consequence of the two following propositions:

PROPOSITION 6.3. *Let fix $\alpha \in [0, d)$, γ and p defined in (2.11) and (2.28) respectively. There exists a constant $C_1 > 0$ (that only depends on p and C_0 defined in Lemma 6.1) such that for all $K \geqslant 1$, $N \geqslant 1$, $a \in \mathcal{D}_K$ and $f \in \mathcal{C}_a$ with $\| f \|_a \leqslant 1$,*

$$(6.4) \qquad \mathbf{E} \left\| \left\langle P_{0,T} f, \nu_0^{(N)} \right\rangle - \left\langle P_{0,T} f, \nu_0 \right\rangle \right\|^p \leqslant C_1 \begin{cases} \left(\frac{K^d}{N^{\gamma \wedge 1}} \right)^p, & \text{if } \alpha \in \left[0, \frac{d}{2} \right), \\ \left(\frac{K^d \ln N}{N^{\frac{d}{2} \wedge 1}} \right)^p, & \text{if } \alpha = \frac{d}{2}, \\ \left(\frac{K^{\frac{3d}{2}} \ln N}{N^{(d-\alpha) \wedge 1}} \right)^p, & \text{if } \alpha \in \left(\frac{d}{2}, d \right). \end{cases}$$

Moreover, in the case where $a \in \mathcal{D}_N$, the previous estimates are true for $K = 1$.

PROPOSITION 6.4. *Let fix $\alpha \in [0, d)$, γ and p defined in (2.11) and (2.28) respectively. There exists a constant $C_2 > 0$ such that for all $K \geqslant 1$, for all $a \in \mathcal{D}_K$, for all $f \in \mathcal{C}_a$ such that $\| f \|_a \leqslant 1$*
(6.5)

$$\mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f \right) (\theta_k(t), \omega_k, x_k) \cdot \mathrm{d}B_k(t) \right\|^p \leqslant C_2 \begin{cases} \left(\frac{K^d}{N^d} \right)^{\frac{p}{2}}, & \text{if } \alpha \in \left[0, \frac{d}{2} \right), \\ \left(\frac{K^d \ln N}{N^d} \right)^{\frac{p}{2}}, & \text{if } \alpha = \frac{d}{2}, \\ \left(\frac{K^d}{N^{d-\alpha}} \right)^p, & \text{if } \alpha \in \left(\frac{d}{2}, d \right). \end{cases}$$

Moreover, in the particular case where $a \in \mathcal{D}_N$, the previous estimates are true for $K = 1$.

Let us admit for a moment Propositions 6.3 and 6.4. Then the result of Theorem 2.18 is a straightforward consequence of the following proposition:

PROPOSITION 6.5. *Under the assumptions made above, there exist constants C_3 and C_4 such that for all $K, N \geqslant 1$, one has:*

$$(6.6) \qquad \sup_{0 \leqslant t \leqslant T} d_K(\nu_t^{(N)}, \nu_t) \leqslant C_3 \begin{cases} \frac{1}{N^{\gamma \wedge 1}} K^d e^{C_4 K^d}, & \text{if } \alpha \in \left[0, \frac{d}{2} \right), \\ \frac{\ln N}{N^{\frac{d}{2} \wedge 1}} K^d e^{C_4 K^{2d}}, & \text{if } \alpha = \frac{d}{2}, \\ \frac{\ln N}{N^{(d-\alpha) \wedge 1}} K^{\frac{3d}{2}} e^{C_4 K^{\frac{dp}{q}}}, & \text{if } \alpha \in \left(\frac{d}{2}, d \right), \end{cases}$$

where q in (6.6) is the conjugate of p and where C_3 and C_4 are large enough constants that depend only on p, T, Γ, Ψ, c and on the constants C_1 and C_2 defined in Propositions 6.3 and 6.4.

PROOF OF PROPOSITION 6.5. Let us fix $K \geqslant 1$, $a \in \mathcal{D}_K$ and $f \in \mathcal{C}_a$ with $\| f \|_a \leqslant 1$.

Let us recall the estimate obtained in (4.14) and (4.15):

$$(6.7) \quad \mathbf{E} \left\| \left\langle f, \nu_T^{(N)} - \nu_T \right\rangle \right\|^p \leqslant$$

$$3^{p-1} \mathbf{E} \left\| \left\langle P_{0,T} f, \nu_0^{(N)} - \nu_0 \right\rangle \right\|^p + 3^{p-1} \mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f \right) (\theta_k(t), \omega_k, x_k) \cdot (\sigma \, \mathrm{d}B_k(t)) \right\|^p$$

$$+ \, 3^{p-1} \left(\frac{e^{q\|P\|T} - 1}{q\|P\|} \right)^{\frac{p}{q}} \left(\frac{1}{|\Lambda_N|} \sum_k \frac{1}{|x_k - a|^{q\alpha}} \right)^{\frac{p}{q}}.$$

$$\int_0^T \frac{1}{|\Lambda_N|} \sum_k \mathbf{E} \left\| \left\langle \Gamma (\theta_k, \omega_k, \cdot, \cdot) \, \Psi (x_k, \cdot), \, \nu_t^{(N)} - \nu_t \right\rangle \right\|^p \mathrm{d}t.$$

We understand here the necessity of choosing p (and its conjugate q) different from 2. Indeed, the integer q (recall Remark 2.17) has been precisely chosen such that $q\alpha < d$ which ensures that the term $\left(\frac{1}{|\Lambda_N|} \sum_k \frac{1}{\|x_k - a\|^{q\alpha}} \right)^{\frac{p}{q}}$ is finite: more precisely, an application of Lemma 6.1, (6.1) shows that this quantity is smaller than $K^{\frac{dp}{q}}$ whenever $a \in \mathcal{D}_K$ and smaller than 1 in the particular case where $a \in \mathcal{D}_N$.

Let us now prove (6.6) in the case where $K > N$. Notice first that, thanks to the assumptions made on Ψ and Γ in § 2.2, for all k the function $f_k : (\theta, \omega, x) \mapsto \Gamma (\theta_k, \omega_k, \theta, \omega) \, \Psi (x_k, x)$ belongs to the space $\mathcal{C}_{x_k}$ where $x_k \in \mathcal{D}_N$. Indeed (recall the definition of $\mathcal{I}_1(\Psi)$ (2.8)), for all k and $(\theta, \omega, \bar{\theta}, \bar{\omega}, x)$,

$$\| x - x_k \|^\alpha \, \Psi (x_k, x) \left\| \Gamma(\theta_k, \omega_k, \theta, \omega) - \Gamma(\theta_k, \omega_k, \bar{\theta}, \bar{\omega}) \right\| \leqslant \mathcal{I}_1(\Psi) \, \| \Gamma \|_{Lip} \left(\| \bar{\theta} - \theta \| + \| \bar{\omega} - \omega \| \right),$$

and

$$\| x - x_k \|^\alpha \, \Psi (x_k, x) \, \| \Gamma(\theta_k, \omega_k, \theta, \omega) \| \leqslant \mathcal{I}_1(\Psi) \, \| \Gamma \|_\infty.$$

As far as condition (2.24) is concerned, we have (using (2.10)):

$$\left\| \| x - x_k \|^{2\gamma} f_k(\theta, \omega, x) - \| y - x_k \|^{2\gamma} f_k(\theta, \omega, y) \right\|$$

$$\leqslant \| \Gamma \|_\infty \left| \| x - x_k \|^{2\gamma} \Psi(x_k, x) - \| y - x_k \|^{2\gamma} \Psi(x_k, y) \right| \leqslant \mathcal{I}_3(\Psi) \| \Gamma \|_\infty |x - y|^{(2\gamma - \alpha) \wedge 1}.$$

Therefore, since $K > N$, by definition of the distance $d_K^{(p)}(\cdot, \cdot)$ (recall Definition 2.16), for all k, the following holds

$$\mathbf{E} \left\| \left\langle \Gamma (\theta_k, \omega_k, \cdot, \cdot) \, \Psi (x_k, \cdot), \, \nu_t^{(N)} - \nu_t \right\rangle \right\|^p \leqslant \eta_1 d_K^{(p)}(\nu_t^{(N)}, \nu_t)^p,$$

for the constant $\eta_1 := \max \left(\mathcal{I}_1(\Psi) \| \Gamma \|_{Lip}, \, \mathcal{I}_1(\Psi), \, \mathcal{I}_3(\Psi) \| \Gamma \|_\infty \right)^p$. Using this estimate in (6.7) and taking the supremum over all functions f in $\bigcup_{\substack{a \in \mathcal{D}_L \\ 1 \leqslant L \leqslant K}} \mathcal{C}_a$, one obtains

$$d_K^{(p)}(\nu_T^{(N)}, \nu_T)^p \leqslant 3^{p-1} \sup_f \mathbf{E} \left\| \left\langle P_{0,T} f, \nu_0^{(N)} \right\rangle - \left\langle P_{0,T} f, \nu_0 \right\rangle \right\|^p$$

$$+ \, 3^{p-1} \sup_f \mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f \right) (\theta_k(t), \omega_k, x_k) \cdot (\sigma \, \mathrm{d}B_k(t)) \right\|^p$$

$$+ \, 3^{p-1} \eta_2 K^{\frac{dp}{q}} \int_0^T d_K^{(p)}(\nu_t^{(N)}, \nu_t)^p \, \mathrm{d}t,$$

for $\eta_2 := \eta_1 \left(\frac{e^{q\|\|P\|\|T}-1}{q\|\|P\|\|} \right)^{\frac{p}{q}}$. The results of Propositions 6.3 and 6.4 together with an application of Gronwall's lemma leads to the estimate (6.6) in the case where $K > N$. Note that one can choose in this case the constants $C_3 := 3^{\frac{p-1}{p}} \left(2 \max \left(C_1, C_2 \right) \right)^{\frac{1}{p}}$ (where C_1 and C_2 come from Propositions 6.3 and 6.4) and $C_4 := \frac{3^{p-1}}{p} T \eta_2$.

Let us now turn to the case where $K \leqslant N$. In this situation, we cannot use Gronwall's inequality in order to obtain an analogous estimate on $d_K^{(p)}(\nu^{(N)}, \nu)$, since the function f_k ($k \in \Lambda_N$) defined at the beginning of this proof has not the sufficient regularity (f_k belongs to $\mathcal{C}_{x_k}$ where $x_k \in \mathcal{D}_N$ and hence may not belong to $\bigcup_{\substack{a \in \mathcal{D}_{K'} \\ 1 \leqslant K' \leqslant K}} \mathcal{C}_a$ for $K < N$). Nonetheless, one can bound the term $\frac{1}{\eta_1} \mathbf{E} \left\| \left\langle \Gamma \left(\theta_k, \omega_k, \cdot, \cdot \right) \Psi \left(x_k, \cdot \right) , \nu_t^{(N)} - \nu_t \right\rangle \right\|^p$ by $\sup_f \mathbf{E} \left\| \left\langle f , \nu_t^{(N)} \right\rangle - \left\langle f , \nu_t \right\rangle \right\|^p$, where the supremum is taken over functions f in $\bigcup_{a \in \mathcal{D}_N} \mathcal{C}_a$ with $\| f \|_a \leqslant 1$. Using this estimate in (6.7) and a calculation similar to the previous one gives the following estimate:
(6.8)
$$
\sup_{0 \leqslant t \leqslant T} \sup_{f \in \bigcup_{a \in \mathcal{D}_N} \mathcal{C}_a} \left(\mathbf{E} \left\| \left\langle f , \nu_t^{(N)} \right\rangle - \left\langle f , \nu_t \right\rangle \right\|^p \right) \leqslant \left(C_3 e^{C_4} \right)^p
\begin{cases}
\left(\frac{1}{N^{\gamma \wedge 1}} \right)^p, & \text{if } \alpha \in \left[0, \frac{d}{2} \right), \\
\left(\frac{\ln N}{N^{\frac{d}{2} \wedge 1}} \right)^p, & \text{if } \alpha = \frac{d}{2}, \\
\left(\frac{1}{N^{(d-\alpha) \wedge 1}} \right)^p, & \text{if } \alpha \in \left(\frac{d}{2}, d \right).
\end{cases}
$$

But then, for instance in the case $\alpha \in \left[0, \frac{d}{2} \right)$ (we let the two other cases to the reader), for all $K \leqslant N$, for all $f \in \bigcup_{a \in \mathcal{D}_{K'}} \mathcal{C}_a$ for $K' \leqslant K$, inserting directly (6.8) into (6.7) and using again Propositions 6.3 and 6.4 leads to:

$$
\mathbf{E} \left\| \left\langle f , \nu_t^{(N)} \right\rangle - \left\langle f , \nu_T \right\rangle \right\|^p \leqslant 3^{p-1} C_1 \left(\frac{K^d}{N^{\gamma \wedge 1}} \right)^p + 3^{p-1} C_2 \left(\frac{K^{d/2}}{N^{d/2}} \right)^p
$$
$$
+ 3^{p-1} \left(\frac{e^{q\|\|P\|\|T}-1}{q\|\|P\|\|} \right)^{\frac{p}{q}} T \left(C_3 e^{C_4} \right)^p \left(\frac{K^{\frac{d}{q}}}{N^{\gamma \wedge 1}} \right)^p .
$$

Up to a change in the constant C_3, this term is anyway smaller than $\left(\frac{C_3}{N^{\gamma \wedge 1}} K^d e^{C_4 K^d} \right)^p$. Taking the supremum over all f in $\bigcup_{\substack{a \in \mathcal{D}_{K'} \\ K' \leqslant K}} \mathcal{C}_a$, one obtains the result. $\qquad \square$

The rest of this part is devoted to prove Propositions 6.3 and 6.4:

PROOF OF PROPOSITION 6.3. Recall that the couples $(\theta_i(0), \omega_i)_{1 \leqslant i \leqslant N}$ are supposed to be chosen i.i.d. according to the law $\zeta(\mathrm{d}\theta) \otimes \mu(\mathrm{d}\omega)$ on $\mathcal{X} \times \mathcal{E}$. Fix $a = \frac{l}{K} \in \mathcal{D}_K$, $f \in \mathcal{C}_a$ with $\| f \|_a \leqslant 1$ as well as $\alpha \in (0, d)$ and the integer $p \geqslant 2$ defined in (2.28). Write again

$\varphi := P_{0,T} f$ for simplicity. Then,

$$\delta_N(f) := \mathbf{E} \left\| \left\langle P_{0,T} f \,, \nu_0^{(N)} \right\rangle - \left\langle P_{0,T} f \,, \nu_0 \right\rangle \right\|^p$$

$$= \mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_j \varphi(\theta_j, \omega_j, x_j) - \int \varphi(\theta, \omega, x) \zeta(\mathrm{d}\theta) \mu(\mathrm{d}\omega) \, \mathrm{d}x \right\|^p,$$

$$\leqslant 2^{p-1} \mathbf{E} \left\| \frac{1}{|\Lambda_N|} \sum_j \varphi(\theta_j, \omega_j, x_j) - \frac{1}{|\Lambda_N|} \sum_j \int \varphi(\theta, \omega, x_j) \zeta(\mathrm{d}\theta) \mu(\mathrm{d}\omega) \right\|^p$$

$$+ 2^{p-1} \left\| \frac{1}{|\Lambda_N|} \sum_j \int \varphi(\theta, \omega, x_j) \zeta(\mathrm{d}\theta) \mu(\mathrm{d}\omega) - \int \varphi(\theta, \omega, x) \zeta(\mathrm{d}\theta) \mu(\mathrm{d}\omega) \, \mathrm{d}x \right\|^p,$$

$$:= A_N + B_N.$$

For simplicity, let us write $X_j := \varphi(\theta_j, \omega_j, x_j) - \int \varphi(\theta, \omega, x_j) \zeta(\mathrm{d}\theta) \mu(\mathrm{d}\omega)$; note that $\mathbf{E} X_j = 0$ for all j. Since the (θ_i, ω_i) are i.i.d. random variables with law $\zeta \otimes \mu$, the first term A_N becomes

$$A_N = \frac{1}{|\Lambda_N|^p} \sum_{l=1}^{\lfloor p/2 \rfloor} \sum_{(k_1 + \ldots + k_l = \lfloor p/2 \rfloor)} \sum_{j_1, \ldots, j_l} \mathbf{E} \left(X_{j_1}^{2k_1} \cdots X_{j_l}^{2k_l} \right),$$

$$(6.9) \qquad \leqslant \frac{2^{2\lfloor p/2 \rfloor}}{|\Lambda_N|^p} \sum_{l=1}^{\lfloor p/2 \rfloor} \sum_{(k_1 + \ldots + k_l = \lfloor p/2 \rfloor)} \sum_{j_1, \ldots, j_l} \frac{1}{\| x_{j_1} - a \|^{2\alpha k_1}} \cdots \frac{1}{\| x_{j_l} - a \|^{2\alpha k_l}},$$

where we used $\| f \|_a \leqslant 1$ and assumption (2.23) in (6.9). Let us concentrate on the contribution of $l = 1$ to the sum in (6.9), that we call $\tilde{A}_N$ (where $\tilde{p} = 2\lfloor p/2 \rfloor$):

$$\tilde{A}_N = \frac{2^{\tilde{p}}}{|\Lambda_N|^p} \sum_j \frac{1}{\| x_j - a \|^{2\tilde{p}\alpha}}.$$

Here, one has to distinguish two cases, depending on the value of $\alpha \in [0, d)$:

1. If $0 \leqslant \alpha < \frac{d}{2}$ then by definition $p = 2$ and $p\alpha < d$ so that an application of Lemma 6.1, (6.1) leads to

$$(6.10) \qquad \qquad \tilde{A}_N \leqslant \frac{1}{N^{2d}} C_0 \cdot K^d N^d = C_0 \frac{K^d}{N^d}.$$

2. If $\alpha \geqslant \frac{d}{2}$, then p is chosen such that $p > \frac{d}{d-\alpha}$ so that $p\alpha > d$. Then Lemma 6.1, (6.3) leads to:

$$(6.11) \qquad \qquad \tilde{A}_N \leqslant \frac{1}{N^{pd}} C_0 \cdot K^{p\alpha} N^{p\alpha} = C_0 \frac{K^{p\alpha}}{N^{p(d-\alpha)}}.$$

It is also easy to see that the other terms in (6.9) are negligible w.r.t. $\tilde{A}_N$ as $N \to \infty$.

Let us now turn to the second term B_N: $(B_N)^{\frac{1}{p}}$ is the difference between the Riemann sum of the function $\Phi := x \mapsto \int \varphi(\theta, \omega, x) \zeta(\mathrm{d}\theta) \mu(\mathrm{d}\omega)$ and its integral, so that it should

be small with N. But one has to be careful since φ as a discontinuity (φ belongs to some $\mathcal{C}_a$ for some a) and since we want to have a result uniform in the function φ:

$$(6.12) \qquad \frac{1}{2^{p-1}} B_N = \left\| \frac{1}{|\Lambda_N|} \sum_j \Phi(x_j) - \int \Phi(x)\, \mathrm{d}x \right\|^p \leqslant \left| \sum_j \int_{\Delta_j} \| \Phi(x_j) - \Phi(x) \|\, \mathrm{d}x \right|^p,$$

where $\Delta_j := \left\{ z \in \left[-\frac{1}{2}, \frac{1}{2} \right]^d; \ \forall k = 1, \ldots, d, \ j_k \leqslant z_k < j_k + \frac{1}{2N} \right\}$ is the infinitesimal sub-domain of Λ_N of size $\frac{1}{2N}$ of corner j. Let us begin with the following straightforward inequality:

$$\| \Phi(x) - \Phi(y) \| \leqslant \left\| \| x - a \|^{-\gamma} - \| y - a \|^{-\gamma} \right\| \left\| \| x - a \|^{\gamma} \Phi(x) + \| y - a \|^{\gamma} \Phi(y) \right\|$$
$$(6.13) \qquad \qquad + \frac{1}{\| x - a \|^{\gamma} \| y - a \|^{\gamma}} \left\| \Phi(x) \| x - a \|^{2\gamma} - \Phi(y) \| y - a \|^{2\gamma} \right\|.$$

Using the assumptions made on f, we deduce in particular from (2.23) and $\| f \|_a \leqslant 1$ that $\| x - a \|^{\gamma} \Phi(x)$ is bounded by $\| x - a \|^{\gamma - \alpha}$. Using also (2.24), it is then immediate to see that

$$\| \Phi(x) - \Phi(y) \| \leqslant \frac{\| x - y \|^{\gamma}}{\| x - a \|^{\gamma} \| y - a \|^{\gamma}} \left(\| x - a \|^{\gamma - \alpha} + \| y - a \|^{\gamma - \alpha} \right) + \frac{\| x - y \|^{(2\gamma - \alpha) \wedge 1}}{\| x - a \|^{\gamma} \| y - a \|^{\gamma}},$$
$$(6.14) \qquad = \frac{\| x - y \|^{\gamma}}{\| x - a \|^{\alpha} \| y - a \|^{\gamma}} + \frac{\| x - y \|^{\gamma}}{\| x - a \|^{\gamma} \| y - a \|^{\alpha}} + \frac{\| x - y \|^{(2\gamma - \alpha) \wedge 1}}{\| x - a \|^{\gamma} \| y - a \|^{\gamma}}.$$

Using (6.14) in (6.12), one obtains that

$$B_N \leqslant 2^{p-1} \left(\sum_j \int_{\Delta_j} \frac{\| x - x_j \|^{\gamma}}{\| x - a \|^{\alpha} \| x_j - a \|^{\gamma}}\, \mathrm{d}x + \sum_j \int_{\Delta_j} \frac{\| x - x_j \|^{\gamma}}{\| x - a \|^{\gamma} \| x_j - a \|^{\alpha}}\, \mathrm{d}x \right.$$
$$(6.15) \qquad \left. + \sum_j \int_{\Delta_j} \frac{\| x - x_j \|^{(2\gamma - \alpha) \wedge 1}}{\| x - a \|^{\gamma} \| x_j - a \|^{\gamma}}\, \mathrm{d}x \right)^p := 2^{p-1} \left(S_N^{(1)} + S_N^{(2)} + S_N^{(3)} \right)^p.$$

The first of the three sums in (6.15) can be bounded by the following quantity:

$$S_N^{(1)} \leqslant \sum_j \frac{1}{\min\left(\| x_{j-1} - a \|^{\alpha}, \| x_j - a \|^{\alpha} \right) \| x_j - a \|^{\gamma}} \int_{\Delta_j} \| x - x_j \|^{\gamma}\, \mathrm{d}x$$
$$= \frac{1}{N^{d+\gamma}} \sum_j \frac{1}{\min\left(\| x_{j-1} - a \|^{\alpha}, \| x_j - a \|^{\alpha} \right) \| x_j - a \|^{\gamma}}.$$

Let us once again distinguish three cases, depending on the value of α:

1. if $\alpha \in \left[0, \frac{d}{2} \right)$, then $\alpha + \gamma < d$ (recall (2.11)), so that an application of Lemma 6.1, (6.1) leads to

$$(6.16) \qquad \qquad \qquad S_N^{(1)} \leqslant C_0 \frac{K^d}{N^{\gamma}}.$$

2. if $\alpha = \frac{d}{2}$, then $\alpha + \gamma = d$ (recall (2.11)), so that Lemma 6.1, (6.2) gives

$$(6.17) \qquad\qquad S_N^{(1)} \leqslant C_0 \frac{K^d \ln N}{N^{\frac{d}{2}}}.$$

3. if $\alpha \in \left(\frac{d}{2}, d\right)$ then $\alpha + \gamma > d$, so that Lemma 6.1, (6.3) gives

$$(6.18) \qquad\qquad S_N^{(1)} \leqslant C_0 \frac{K^{\alpha+\gamma}}{N^{d-\alpha}} \leqslant C_0 \frac{K^{\frac{3d}{2}}}{N^{d-\alpha}}.$$

The same calculation leads to the same estimates for the second term $S_N^{(2)}$ in (6.15). A very similar calculation also leads to the following estimate for the last term $S_N^{(3)}$:

$$(6.19) \qquad\qquad S_N^{(3)} \leqslant C_0 \begin{cases} \frac{K^d}{N^{(2\gamma-\alpha)\wedge 1}} & \text{if } \alpha \in \left[0, \frac{d}{2}\right), \\[2mm] \frac{K^d \ln N}{N^{(d-\alpha)\wedge 1}} & \text{if } \alpha \in \left[\frac{d}{2}, d\right). \end{cases}$$

Combining estimations (6.19) and (6.10) (resp. (6.11)) and (6.16), (resp. (6.17) or (6.18)) leads to the desired estimation (6.4). The proof of the case where $a \in \mathcal{D}_N$ is analogous and uses the estimates for $a \in \mathcal{D}_N$ in Lemma 6.1. Proposition 6.3 is proved. $\qquad\square$

It remains to prove Proposition 6.4, whose purpose is to control the martingale term in (4.12):

PROOF OF PROPOSITION 6.4. Fix some $K \geqslant 1$, $a \in \mathcal{D}_K$ and $f \in \mathcal{C}_a$ such that $\| f \|_a \leqslant 1$. The martingale $M_t^N := \frac{1}{|\Lambda_N|} \sum_k \int_0^T \nabla_\theta \left(P_{t,T} f\right) (\theta_k(t), \omega_k, x_k) \cdot \mathrm{d}B_k(t)$ may be written as $M_t^N = \frac{1}{|\Lambda_N|} \sum_k \sum_{l=1}^m \int_0^T \partial_{\theta^{(l)}} \left(P_{t,T} f\right) (\theta_k(t), \omega_k, x_k) \, \mathrm{d}B_k^{(l)}(t)$, where for all k, $B_k = (B_k^{(1)}, \ldots, B_k^{(m)})$. Consequently, its quadratic variation process is given by

$$\langle M^N \rangle_t = \frac{1}{|\Lambda_N|^2} \sum_k \sum_{l=1}^m \int_0^T \| \partial_{\theta^{(l)}} P_{t,T} f (\theta_k(t), \omega_k, x_k) \|^2 \, \mathrm{d}t.$$

Applying Remark 2.15 and Lemma 4.4, we have almost surely that

$$\langle M^N \rangle_t \leqslant \frac{m \|\!\| P \|\!\|^2}{|\Lambda_N|^2} \sum_k \frac{1}{|x_k - a|^{2\alpha}} \int_0^T e^{2\|\!\| P \|\!\|(T-t)} \, \mathrm{d}t$$

An argument repeatedly used in this work shows that one can bound the quadratic variation by $C \frac{K^d}{N^d}$ (respectively $C \frac{K^d \ln N}{N^d}$ and $C \frac{K^{2\alpha}}{N^{2(d-\alpha)}}$) when $\alpha < \frac{d}{2}$ (respectively $\alpha = \frac{d}{2}$ and $\alpha > \frac{d}{2}$), for some constant $C > 0$. Then, Burkholder-Davis-Gundy inequality $\mathbf{E}\left(\| M_t^N \|^p\right) \leqslant C_p \mathbf{E}\left(\langle M^N \rangle_t^{\frac{p}{2}}\right)$ gives the result. Proposition 6.4 is proved. $\qquad\square$

7. The case of a locally Lipschitz dynamics $c(\cdot)$. One of the key arguments of the proofs of Theorems 2.13 and 2.18 is the fact that one can derive a Kolmogorov equation (recall (4.3)) for the propagator $P_{s,t} f$ defined in (4.2). Under Assumption 2.2 on the dynamics $c(\cdot)$ (one-sided Lipschitz condition and absence of *global* Lispchitz continuity), deriving such a Kolmogorov equation appears to be problematic (see in particular [22, 20]). Even if such a result existed, we could not find a proper reference in the literature.

One can bypass this technical difficulty and prove nevertheless Theorem 2.13 and 2.18 by an approximation argument. We will suppose throughout this section that c satisfies only Assumption 2.2.

7.1. *Yosida approximation.* Let us denote for all (θ, ω), $\tilde{c}(\theta, \omega) := c(\theta, \omega) - L\theta$, where we recall that L is the constant appearing in the one-sided Lipschitz continuity assumption (2.3). In terms of $\tilde{c}$, (2.3) reads:

$$(7.1) \qquad \forall (\theta, \omega), (\bar{\theta}, \bar{\omega}), \ \left\langle \theta - \bar{\theta}, \ \tilde{c}(\theta, \omega) - \tilde{c}(\bar{\theta}, \bar{\omega}) \right\rangle \leqslant 0,$$

and, for example, the mean field evolution (4.1) reads:

$$(7.2) \qquad \mathrm{d}\theta(t) = \tilde{c}(\theta(t), \omega) \, \mathrm{d}t + \tilde{v}(t, \theta(t), \omega, x) \, \mathrm{d}t + \sigma \cdot \mathrm{d}B(t),$$

where $\tilde{v}(t, \theta(t), \omega, x) := v(t, \theta(t), \omega, x) + L\theta(t)$.

For all $\lambda > 0$, consider $\tilde{c}_\lambda$ the Yosida approximation of $\tilde{c}$ (see [8, Appendix A] for a review of the basic properties of Yosida approximations):

$$(7.3) \qquad \forall (\theta, \omega), \ \tilde{c}_\lambda(\theta, \omega) := \tilde{c}(R_\lambda(\lambda\theta), \omega),$$

for

$$(7.4) \qquad \forall (\theta, \omega), \ R_\lambda(\theta, \omega) := (\lambda - \tilde{c}(\cdot, \omega))^{-1}(\theta).$$

Consider now the solution θ_λ of the following SDE (with the same initial condition and driven by the same Brownian motion B as in (7.2))

$$(7.5) \qquad \mathrm{d}\theta_\lambda(t) = \tilde{c}_\lambda(\theta_\lambda(t), \omega) \, \mathrm{d}t + \tilde{v}(t, \theta_\lambda(t), \omega, x) \, \mathrm{d}t + \sigma \cdot \mathrm{d}B(t),$$

that is, the analog of (7.2) where $\tilde{c}$ has been replaced by its Yosida approximation. Note that one can proceed exactly in the same way for the microscopic system (2.2). From now on, whatever X may be, the subscript notation X_λ will refer to the analog of X when the dynamics has been replaced by its Yosida approximation. Note that we will most of the time drop the dependencies of the functions in ω, for simplicity of notations.

It is easy to see that $\tilde{c}$ and $\tilde{c}_\lambda$ have the same regularity in θ (see e.g. [8, p.304]). Moreover, $\tilde{c}_\lambda$ has the supplementary property to be uniformly Lipschitz continuous. In other words, $\tilde{c}_\lambda$ satisfies Assumption 2.2 as well as Assumption 4.1, so that everything that has been done before is applicable: Theorems 2.13 and 2.18 are true in the case of an interaction ruled by $\tilde{c}_\lambda$:

$$(7.6) \qquad \sup_{t \in [0,T]} d\left(\nu_{t,\lambda}^{(N)}, \nu_{t,\lambda} \right) \leqslant C N^{-\beta},$$

for d either equal to $d_R(\cdot, \cdot)$ or $d_\infty^{(p)}(\cdot, \cdot)$ and β one of the appropriate exponent appearing in the formulation of Theorems 2.13 and 2.18. Note that the constant C in (7.6) *does not depend on* λ. Indeed, the assumption made in Section 4 about the global Lipschitz continuity of c was made only to ensure the existence of the Kolmogorov equation. In particular, the modulus of continuity of c did not enter into the calculation made in Section 4: the only dependence in the dynamics c was in its one sided-Lipschitz constant L (recall Lemma 4.4), which is conserved by the Yosida approximation. In other words, every constant estimates made upon evolution (7.5) is independent on λ.

Now, Theorems 2.13 and 2.18 in our general framework are an easy consequence of the triangular inequality and the proposition:

PROPOSITION 7.1. *For all $N \geqslant 1$,*

$$(7.7) \qquad \sup_{t \in [0,T]} d\left(\nu_{t,\lambda}^{(N)}, \nu_t^{(N)}\right) \to_{\lambda \to \infty} 0$$

$$(7.8) \qquad \sup_{t \in [0,T]} d\left(\nu_{t,\lambda}, \nu_t\right) \to_{\lambda \to \infty} 0.$$

The rest of this section is devoted to prove Proposition 7.1. Let us begin with some a priori estimate:

LEMMA 7.2. *We have the following a priori estimates*

$$(7.9) \qquad \sup_{\lambda > 0} \mathbf{E}\left(\sup_{t \in [0,T]} \|\theta_\lambda(t)\|^2\right) < \infty.$$

and,

$$(7.10) \qquad \mathbf{P}\left(\sup_{\lambda > 0} \int_0^T \|\tilde{c}_\lambda(\theta_\lambda(s))\|^2 \, \mathrm{d}s < \infty\right) = 1.$$

PROOF OF LEMMA 7.2. Let us first prove the first estimate (7.9): applying Ito formula,

$$\|\theta_\lambda(t)\|^2 = \|\theta_\lambda(0)\|^2 + 2\int_0^t \langle \theta_\lambda(s), \, \tilde{c}_\lambda(\theta_\lambda(s)) + \tilde{v}(s, \theta_\lambda(s), \omega, x)\rangle \, \mathrm{d}s$$

$$+ 2\int_0^t \langle \theta_\lambda(s), \, \mathrm{d}B(s)\rangle + tr(\sigma\sigma^T)t,$$

$$\leqslant \|\theta_\lambda(0)\|^2 + 2\left(\|\tilde{c}(0)\| + L + \|\Gamma\|_\infty S(\Psi)\right) \int_0^t \|\theta_\lambda(s)\|^2 \, \mathrm{d}s$$

$$+ 2\int_0^t \langle \theta_\lambda(s), \, \mathrm{d}B(s)\rangle + tr(\sigma\sigma^T)T.$$

Taking expectations and using Burkholder-Davis-Gundy inequality, we obtain that for some constant $C > 0$ (independent of λ),

$$\mathbf{E}\left(\sup_{s \leqslant t} \|\theta_\lambda(s)\|^2\right) \leqslant \mathbf{E}\left(\|\theta(0)\|^2\right) + tr(\sigma\sigma^T)T + 2C \int_0^t \mathbf{E}\left(\sup_{u \leqslant s} \|\theta_\lambda(u)\|^2\right) \, \mathrm{d}s$$

$$+ 6tr(\sigma\sigma^T)^{1/2} \mathbf{E}\left(\left(\int_0^t \|\theta_\lambda(u)\|^2 \, \mathrm{d}u\right)^{\frac{1}{2}}\right)$$

$$\leqslant \mathbf{E}\left(\|\theta(0)\|^2\right) + tr(\sigma\sigma^T)T + 2C \int_0^t \mathbf{E}\left(\sup_{u \leqslant s} \|\theta_\lambda(u)\|^2\right) \, \mathrm{d}s$$

$$+ 18tr(\sigma\sigma^T)T + \frac{1}{2}\mathbf{E}\left(\sup_{u \leqslant t} \|\theta_\lambda(u)\|^2\right)$$

which implies

$$\mathbf{E}\left(\sup_{s \leqslant t} \|\theta_\lambda(s)\|^2\right) \leqslant 2\left(\mathbf{E}\left(\|\theta(0)\|^2\right) + 19tr(\sigma\sigma^T)T\right) + 4C \int_0^t \mathbf{E}\left(\sup_{u \leqslant s} \|\theta_\lambda(u)\|^2\right) \, \mathrm{d}s$$

and Gronwall lemma leads to the result.

Let us now turn to the second estimate (7.10): define $Y_\lambda(t) := \theta_\lambda(t) - \sigma \cdot B(t)$. Then, Y_λ satisfies:

$$(7.11) \qquad \mathrm{d}Y_\lambda(t) = (\tilde{c}_\lambda(Y_\lambda(t) + B(t), \omega) + \tilde{v}(t, Y_\lambda(t) + B(t), \omega, x)) \, \mathrm{d}t.$$

Clearly,

$$\| Y_\lambda(t) \|^2 = \| Y_\lambda(0) \|^2 + 2 \int_0^t \langle Y_\lambda(s) \,,\, \tilde{c}_\lambda(Y_\lambda(s) + \sigma \cdot B(s)) \rangle \, \mathrm{d}s$$

$$+ 2 \int_0^t \langle Y_\lambda(s) \,,\, \tilde{v}(s, Y_\lambda(s) + \sigma \cdot B(s)), \omega, x \rangle \, \mathrm{d}s$$

$$\leqslant \| Y_\lambda(0) \|^2 + 2 \left(\| \tilde{c}(0) \| + L + \| \Gamma \|_\infty S(\Psi) \right) \int_0^t \| Y_\lambda(s) \|^2 \, \mathrm{d}s$$

$$+ 2 \int_0^t \langle Y_\lambda(s) \,,\, \tilde{c}_\lambda(\sigma \cdot B(s)) \rangle \, \mathrm{d}s$$

$$\leqslant \| Y_\lambda(0) \|^2 + 2 \left(\| \tilde{c}(0) \| + L + \| \Gamma \|_\infty S(\Psi) + \int_0^t \| \tilde{c}_\lambda(\sigma \cdot B(s)) \|^2 \, \mathrm{d}s \right) \int_0^t \| Y_\lambda(s) \|^2 \, \mathrm{d}s$$

taking the supremum in λ and using $Y_\lambda(0) = \theta_\lambda(0) = \theta(0)$, we have

$$\sup_\lambda \| Y_\lambda(t) \|^2 \leqslant \| \theta(0) \|^2 + 2 \left(C + \int_0^t \| \tilde{c}_\lambda(\sigma \cdot B(s)) \|^2 \, \mathrm{d}s \right) \int_0^t \sup_\lambda \| Y_\lambda(s) \|^2 \, \mathrm{d}s$$

$$\leqslant \| \theta(0) \|^2 + 2 \left(C + \int_0^t \| \tilde{c}(\sigma \cdot B(s)) \|^2 \, \mathrm{d}s \right) \int_0^t \sup_\lambda \| Y_\lambda(s) \|^2 \, \mathrm{d}s$$

where we used the pointwise estimate $\| \tilde{c}_\lambda(\theta) \| \leqslant \| \tilde{c}(\theta) \|$. Gronwall lemma gives

$$\sup_\lambda \| Y_\lambda(t) \|^2 \leqslant \| \theta(0) \|^2 \exp \left(2 \left(C + \int_0^T \| \tilde{c}(\sigma \cdot B(s)) \|^2 \, \mathrm{d}s \right) T \right)$$

that is almost surely finite, since $\tilde{c}$ is locally bounded and the trajectories of B are almost surely bounded. Consequently

$$\sup_\lambda \sup_{t \leqslant T} \| \theta_\lambda(t) \|^2 \leqslant \sup_\lambda \sup_{t \leqslant T} \| Y_\lambda(t) \|^2 + \sup_{t \leqslant T} \| B(t) \|^2 < \infty, \ a.s.$$

Since $\tilde{c}$ is polynomially bounded, this implies now that

$$\sup_\lambda \int_0^T \| \tilde{c}_\lambda(\theta_\lambda(t)) \|^2 \, \mathrm{d}t < \infty, \ a.s.$$

which is the result.   □

The key estimate of this section is the following

PROPOSITION 7.3. *Almost surely, the following holds*

$$(7.12) \qquad\qquad\qquad \limsup_{\lambda \to \infty} \sup_{t \in [0,T]} \| \theta(t) - \theta_\lambda(t) \| = 0.$$

PROOF OF PROPOSITION 7.3. Let us fix $\lambda < \mu$. Since the Brownian motion is the same, one has successively (for a constant $C = L + \|\Gamma\|_{Lip} S(\Psi)$)

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} e^{-2Ct} \|\theta_\mu(t) - \theta_\lambda(t)\|^2 &= -2C e^{-2Ct} \|\theta_\mu(t) - \theta_\lambda(t)\|^2 \\
&\quad + 2 e^{-2Ct} \langle \theta_\mu(t) - \theta_\lambda(t), \, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \rangle \\
&\quad + 2 e^{-2Ct} \langle \theta_\mu(t) - \theta_\lambda(t), \, \tilde{v}(t, \theta_\mu(t), \omega, x) - \tilde{v}(t, \theta_\lambda(t), \omega, x) \rangle \\
&\leqslant -2C e^{-2Ct} \|\theta_\mu(t) - \theta_\lambda(t)\|^2 \\
&\quad + 2 e^{-2Ct} \langle \theta_\mu(t) - \theta_\lambda(t), \, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \rangle \\
&\quad + 2 e^{-2Ct} \left(L + \|\Gamma\|_{Lip} S(\Psi) \right) \|\theta_\mu(t) - \theta_\lambda(t)\|^2 \\
&\leqslant 2 e^{-2Ct} \langle \theta_\mu(t) - \theta_\lambda(t), \, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \rangle \\
&= 2 e^{-2Ct} \Big\langle \left(R_\mu(\mu \theta_\mu(t)) - \frac{1}{\mu} \tilde{c}(R_\mu(\mu \theta_\mu(t))) \right) \\
&\quad - \left(R_\lambda(\lambda \theta_\lambda(t)) - \frac{1}{\lambda} \tilde{c}(R_\lambda(\lambda \theta_\lambda(t))) \right), \tilde{c}(R_\mu(\mu \theta_\mu(t))) - \tilde{c}(R_\lambda(\lambda \theta_\lambda(t))) \Big\rangle \\
&\leqslant -2 e^{-2Ct} \Big\langle \frac{1}{\mu} \tilde{c}_\mu(\theta_\mu(t)) - \frac{1}{\lambda} \tilde{c}_\lambda(\theta_\lambda(t)), \, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \Big\rangle .
\end{aligned}
$$

Integrating this inequality gives (since the initial condition is the same)

$$
\frac{1}{2} e^{-2CT} \|(\theta_\mu - \theta_\lambda)(T)\|^2 \leqslant -\int_0^T e^{-2Ct} \Big\langle \frac{1}{\mu} \tilde{c}_\mu(\theta_\mu(t)) - \frac{1}{\lambda} \tilde{c}_\lambda(\theta_\lambda(t)), \, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \Big\rangle \, \mathrm{d}t.
$$

This gives in particular that

$$
\int_0^T e^{-2Ct} \Big\langle \frac{1}{\mu} \tilde{c}_\mu(\theta_\mu(t)) - \frac{1}{\lambda} \tilde{c}_\lambda(\theta_\lambda(t)), \, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \Big\rangle \, \mathrm{d}t \leqslant 0.
$$

Let us denote as $\|\cdot\|_H$ the Hilbert norm in $H := L^2([0,T], e^{-2Cs}\,\mathrm{d}s; \mathcal{X})$. Then, from the identity

$$
\begin{aligned}
2 \Big\langle \tilde{c}_\mu(\theta_\mu) - \tilde{c}_\lambda(\theta_\lambda), \, \frac{1}{\mu} \tilde{c}_\mu(\theta_\mu) - \frac{1}{\lambda} \tilde{c}_\lambda(\theta_\lambda) \Big\rangle_H &= \left(\frac{1}{\mu} + \frac{1}{\lambda} \right) \|\tilde{c}_\mu(\theta_\mu) - \tilde{c}_\lambda(\theta_\lambda)\|_H^2 \\
&\quad + \left(\frac{1}{\mu} - \frac{1}{\lambda} \right) \left(\|\tilde{c}_\mu(\theta_\mu)\|_H^2 - \|\tilde{c}_\lambda(\theta_\lambda)\|_H^2 \right)
\end{aligned}
$$

one obtains that

$$
(7.13) \qquad \left(\frac{1}{\mu} + \frac{1}{\lambda} \right) \|\tilde{c}_\mu(\theta_\mu) - \tilde{c}_\lambda(\theta_\lambda)\|_H^2 \leqslant \left(\frac{1}{\lambda} - \frac{1}{\mu} \right) \left(\|\tilde{c}_\mu(\theta_\mu)\|_H^2 - \|\tilde{c}_\lambda(\theta_\lambda)\|_H^2 \right)
$$

which gives in particular that $\lambda \mapsto \|\tilde{c}_\lambda(\theta_\lambda)\|_H^2$ is increasing and by (7.10) bounded and thus, convergent. The same inequality (7.13) shows also that $\|\tilde{c}_\mu(\theta_\mu) - \tilde{c}_\lambda(\theta_\lambda)\|_H^2 \to_{\lambda,\mu\to\infty} 0$, so that $(\tilde{c}_\lambda(\theta_\lambda)(t))$ converges in H to some $c_\infty(t)$.

Going back to the first inequality of the proof, one has

$$\frac{1}{2} \sup_{t \in [0,T]} e^{-2Ct} \| \theta_\mu(t) - \theta_\lambda(t) \|^2 \leqslant \int_0^T e^{-2Ct} \langle \theta_\mu(t) - \theta_\lambda(t) \,,\, \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \rangle \,\mathrm{d}t$$

$$\leqslant \frac{1}{4T} \int_0^T e^{-2Ct} \| \theta_\mu(t) - \theta_\lambda(t) \|^2 \,\mathrm{d}t$$

$$+ T \int_0^T e^{-2Ct} \| \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \|^2 \,\mathrm{d}t$$

$$\leqslant \frac{1}{4} \sup_{t \in [0,T]} e^{-2Ct} \| \theta_\mu(t) - \theta_\lambda(t) \|^2$$

$$+ T \int_0^T e^{-2Ct} \| \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \|^2 \,\mathrm{d}t.$$

Hence

$$\sup_{t \in [0,T]} e^{-2Ct} \| \theta_\mu(t) - \theta_\lambda(t) \|^2 \leqslant 4T \int_0^T e^{-2Ct} \| \tilde{c}_\mu(\theta_\mu(t)) - \tilde{c}_\lambda(\theta_\lambda(t)) \|^2 \,\mathrm{d}t,$$

which goes to 0 as $\lambda, \mu \to \infty$. This implies that there exists an adapted process $\bar{\theta}$ with continuous trajectories such that $\lim_{\lambda \to \infty} \theta_\lambda = \bar{\theta}$, uniformly and almost surely. Clearly, for all t, the strong continuity $\lim_{\lambda \to \infty} R_\lambda(\lambda \bar{\theta}(t)) = \bar{\theta}(t)$ of the resolvent and the uniform Lipschitz continuity $\| R_\lambda(\lambda \theta_\lambda(t)) - R_\lambda(\lambda \bar{\theta}(t)) \| \leqslant \| \theta_\lambda(t) - \theta(t) \|$ implies that $\lim_\lambda R_\lambda(\lambda \theta_\lambda(t)) = \bar{\theta}(t)$. Finally, continuity of $\tilde{c}$ gives $\lim_{\lambda \to \infty} \tilde{c}_\lambda(\theta_\lambda(t)) = \tilde{c}(R_\lambda(\lambda \theta_\lambda(t))) = \tilde{c}(\bar{\theta}(t))$. Consequently, we have that, almost surely $\tilde{c}(\bar{\theta}_t) = c_\infty(t)$, so that $\bar{\theta}$ solves equation (7.2), so that by uniqueness $\bar{\theta} = \theta$ almost surely. □

We are now in position to prove Proposition 7.1:

PROOF OF PROPOSITION 7.1. We only prove (7.8), the proof of (7.7) follows from analogous estimates with the microscopic equation (2.2). We only treat the (more complicated) case of the power-law interaction. Fix any f in $\mathcal{C}_a$ for some a with $\| f \|_a \leqslant 1$. Then, by Lispchitz continuity of f in the variable θ

$$|\langle f \,,\, \nu_{t,\lambda} \rangle - \langle f \,,\, \nu_t \rangle| \leqslant S(\Psi) \mathbf{E}_B \| \theta_\lambda(t) - \theta(t) \|.$$

Taking the supremum in f and in t leads to

$$\sup_{t \in [0,T]} d(\nu_{t,\lambda}, \nu_t) \leqslant S(\Psi) \mathbf{E}_B \sup_{t \in [0,T]} \| \theta_\lambda(t) - \theta(t) \|.$$

By (7.12) we have the almost sure convergence to 0 of $\sup_{t \in [0,T]} \| \theta_\lambda(t) - \theta(t) \|$ and (7.9) gives the boundedness in L^2 implying uniform integrability. The result follows. □

APPENDIX A: PROOF OF A TECHNICAL LEMMA

PROOF OF LEMMA 6.1. Let us proceed by induction on the dimension d. Let fix $d = 1$:

- Let us begin with the case where $a \notin \mathcal{D}_N$: let J be the integer such that $\frac{J}{2N} < a < \frac{J+1}{2N}$. Then, an easy comparison with integrals shows the following

$$
\sum_j \left| \frac{j}{2N} - a \right|^{-\beta} \leqslant 2^\beta N^\beta \Big(\int_0^J |2aN - t|^{-\beta} \, \mathrm{d}t + |2aN - J|^{-\beta} + |2aN - (J+1)|^{-\beta}
$$
$$
+ \int_{J+1}^N |t - 2aN|^{-\beta} \, \mathrm{d}t \Big)
$$
$$
= 2^\beta N^\beta \int_0^J |2aN - t|^{-\beta} \, \mathrm{d}t + 2^\beta N^\beta \int_{J+1}^N |t - aN|^{-\beta} \, \mathrm{d}t
$$
$$
+ \left| a - \frac{J}{2N} \right|^{-\beta} + \left| a - \frac{J+1}{2N} \right|^{-\beta}.
$$

It is straightforward to see that the two first integral terms are smaller than $\frac{N}{d-\beta}$ whereas each of the two remaining terms is smaller than $\rho(N,K)^{-\beta}$, where $\rho(N,K) := \inf_{\substack{|j| \leqslant N \\ |l| \leqslant K \\ j/N \neq l/K}} \left| \frac{j}{2N} - \frac{l}{2K} \right| = \frac{\gcd(K,N)}{2KN} \geqslant \frac{1}{2KN}$. Consequently, since $K \geqslant 1$ and $\beta < 1$,

$$
\sum_j \left| \frac{j}{N} - a \right|^{-\beta} \leqslant \frac{2N}{d-\beta} + 2K^\beta N^\beta \leqslant C_0 N K.
$$

- The case where $a \in \mathcal{D}_N$ is easier: in that case, $a = \frac{k}{2N}$ for some k. Then, once again by comparison with integrals,

$$
\sum_{j;\ j/N \neq a} \left| \frac{j}{2N} - a \right|^{-\beta} = 2^\beta N^\beta \sum_{j \neq k} |j - k|^{-\beta} \leqslant \frac{N^\beta}{1-\beta} \left((N+k)^{1-\beta} + (N-k)^{1-\beta} \right) \leqslant \frac{2^{2-\beta} N}{1-\beta}.
$$

The other cases ($\beta = 1$ and $\beta > 1$) are similar and left to the reader. Lemma 6.1 is proved in the particular case of $d = 1$.

The case of higher dimension is nothing but a technical complication of the previous case $d = 1$. Let us fix $d > 1$, $a = (a_1, \ldots, a_d) \in \mathcal{D}_K$ and denote by $j = (j_1, \ldots, j_d)$ any element of $\mathbf{Z}^d$.

Let us begin with the case where $a \notin \mathcal{D}_K$. Let $(J_1, \ldots, J_d)$ the d integers between $-N$ and N such that for all $l = 1, \ldots, d$, $J_l \leqslant 2a_l N \leqslant J_l + 1$, with at least one inequality that is strict. The coordinates J_l and $J_l + 1$ are by construction the closest integers to $2a_l N$ in $-N, \ldots, N$. For the rest of this proof, we will refer to them as *critical* coordinates. Then, one can decompose the sum $\sum_j \left\| \frac{j}{2N} - a \right\|^{-\beta}$ according to the number p of critical coordinates among $(j_1, \ldots, j_d) = j$, where j s a typical index:

(A.1) $$ \sum_j \left\| \frac{j}{2N} - a \right\|^{-\beta} = \sum_{p=0}^d \sum_{(i_1, \ldots, i_p)} \sum_{j \in \mathcal{J}_{(i_1, \ldots, i_p)}} \left\| \frac{j}{2N} - a \right\|^{-\beta}, $$

where the second sum is taken over all the vectors $(i_1, \ldots, i_p)$ with strictly increasing indices taken among $1, \ldots, d$ and where $\mathcal{J}(i_1, \ldots, i_p)$ is a notation for the set of vectors $j = (j_1, \ldots, j_d)$ such that j_{i_l} is critical for every $l = 1, \ldots, p$.

In the sum (A.1), let us treat the cases $p = 0$ and $p > 0$ separately. Let us first focus on the case $p = 0$: it corresponds to vectors j without critical coordinates, which means that we restrict ourselves to j such that for every $k = 1, \ldots, d$, either $j_k < J_k$ (in such case $|j_k - 2a_k N| = 2a_k N - j_k$) or either $j_k > J_k + 1$ (in such case $|j_k - 2a_k N| = j_k - 2a_k N$). In particular, this sum can be divided into 2^d sums $\sum_{j \in D} \left\| \frac{j}{2N} - a \right\|^{-\beta}$ where D is a connected subdomain of $[-1/2, 1/2]^d$, which is defined by this binary choice for each j_k. For simplicity, we only treat the case of $D_0 := \{j = (j_1, \ldots, j_d); \forall k = 1, \ldots, d, \ j_k < J_k\}$. The case of the other $2^d - 1$ subdomains can be treated in a similar way.

We have successively,

$$
(A.2) \qquad \sum_{j \in D_0} \left\| \frac{j}{2N} - a \right\|^{-\beta} = 2^\beta N^\beta \sum_{j_k < J_k - 1} \left| \sum_{l=1}^d (2a_l N - j_l)^2 \right|^{-\beta/2}
$$

$$
(A.3) \qquad \leqslant 2^\beta N^\beta \int_{-N}^{J_1} \cdots \int_{-N}^{J_d} \left| \sum_{l=1}^d (2a_l N - t_l)^2 \right|^{-\beta/2} \mathrm{d}t_1 \ldots \mathrm{d}t_d
$$

$$
(A.4) \qquad = 2^\beta N^\beta \int_{2a_1 N - J_1}^{N + 2a_1 N} \cdots \int_{2a_d N - J_d}^{N + 2a_d N} \left| \sum_{l=1}^d u_l^2 \right|^{-\beta/2} \mathrm{d}u_1 \ldots \mathrm{d}u_d
$$

$$
(A.5) \qquad \leqslant C N^\beta \int_{w_N}^{2N} \frac{1}{r^\beta} r^{d-1} \, \mathrm{d}r,
$$

where $w_N > 0$ is the distance to 0 of the point of coordinates $(2a_1 N - J_1, \ldots, 2a_d - J_d)$. The estimates found in Lemma 6.1 are then straightforward: for example in the case $\beta < d$, an upper bound for the last quantity is $C N^\beta N^{d-\beta} = C N^d$. The other cases are treated in the same manner and lead to the same desired estimate.

As far as the case $0 < p \leqslant d$ is concerned, the particular case $p = d$ is a bit special: it corresponds to vectors j with only critical coordinates. Since in that case, each $\left| \frac{j_k}{2N} - a_k \right|$ is either equal to $\left| \frac{J_k}{2N} - a_k \right|$ or $\left| \frac{J_k + 1}{2N} - a_k \right|$ and is anyway larger than $\rho_{N,K} \geqslant \frac{1}{2NK}$ (where the quantity $\rho_{N,K}$ has been defined in the beginning of this proof), the contribution of this case to the whole sum can be bounded by $2^d \cdot \frac{1}{(d\rho_{N,K}^2)^{\beta/2}} \leqslant \frac{2^d 2^\beta}{d^{\beta/2}} N^\beta K^\beta = C N^\beta K^\beta$.

Let us now concentrate on the case $0 < p < d$: Then for a fixed choice of indices $(i_1, \ldots, i_p)$, we have

$$
\sum_{j \in \mathcal{J}_{(i_1, \ldots, i_p)}} \left\| \frac{j}{2N} - a \right\|^{-\beta} = \sum_{j \in \mathcal{J}_{(i_1, \ldots, i_p)}} \left| \sum_{i = i_1, \ldots, i_p} \left(\frac{j_i}{2N} - a_i \right)^2 + \sum_{i \neq i_1, \ldots, i_p} \left(\frac{j_i}{2N} - a_i \right)^2 \right|^{-\beta/2}
$$

$$
\leqslant \sum_{j \in \mathcal{J}_{(i_1, \ldots, i_p)}} \left| \sum_{i \neq i_1, \ldots, i_p} \left(\frac{j_i}{2N} - a_i \right)^2 \right|^{-\beta/2}.
$$

But this last sum is nothing else than $\sum_{\bar{j}} \left\| \frac{\bar{j}}{2N} - \bar{a} \right\|^{-\beta}$, where $\bar{a}$ (resp. $\bar{j}$) is the vector in $[-1, 1]^{d-p}$, built upon the vector a (resp. j) with all its coordinates of index in $\{i_1, \ldots, i_p\}$

removed. Since $p > 0$, we see that, by induction hypothesis, that the previous sum can be bounded by

$$\begin{cases} CN^{d-p}K^{d-p}\ln N & \text{if } \beta \leqslant d - p \\ CN^{\beta} & \text{if } \beta > d - p. \end{cases}$$

In particular, if $\beta \geqslant d$, then the contribution to (A.1) of the sum over $0 < p < d$ can be bounded by $CN^{d-p}K^{d-p}\ln N \leqslant \min\left(CK^d N^d \ln N, CN^\beta\right)$. If $\beta < d$, it is also straightforward to see that this contribution is also smaller than $CN^d K^d$. The proof of Lemma 6.1 follows, by induction. $\qquad\square$

ACKNOWLEDGEMENTS

REFERENCES

[1] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.*, 77:137–185, Apr 2005.

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

[3] L. Bertini, G. Giacomin, and C. Poquet. Synchronization and random long time dynamics for mean-field plane rotators, 2012. `arxiv.org/abs/1209.4537`.

[4] F. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.*, 44(5):867–884, 2010.

[5] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields*, 137(3-4):541–593, 2007.

[6] M. Bossy, J.-F. Jabir, and D. Talay. On conditional McKean Lagrangian stochastic models. *Probab. Theory Related Fields*, 151(1-2):319–351, 2011.

[7] M. Bossy and D. Talay. Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation. *Ann. Appl. Probab.*, 6(3):818–861, 1996.

[8] S. Cerrai. *Second order PDE's in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.

[9] D. Chowdhury and M. C. Cross. Synchronization of oscillators with long-range power law interactions. *Phys. Rev. E*, 82:016205, Jul 2010.

[10] P. Dai Pra and F. den Hollander. McKean-Vlasov limit for interacting random processes in random media. *J. Statist. Phys.*, 84(3-4):735–772, 1996.

[11] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. 68 pages (2012), 2012. `arxiv.org/abs/1211.0299`.

[12] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Process. Appl.*, 86(2):193–216, 2000.

[13] 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.

[14] G. B. Ermentrout and D. H. Terman. *Mathematical foundations of neuroscience*, volume 35 of *Interdisciplinary Applied Mathematics*. Springer, New York, 2010.

[15] J. Gärtner. On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, 137:197–248, 1988.

[16] I. M. Gel'fand and N. Y. Vilenkin. *Generalized functions. Vol. 4*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1964 [1977].

[17] G. Giacomin, E. Luçon, and C. Poquet. Coherence stability and effect of random natural frequencies in population of coupled oscillators, 2011. `arxiv.org/abs/1111.3581`.

[18] G. Giacomin, K. Pakdaman, and X. Pellegrin. Global attractor and asymptotic dynamics in the Kuramoto model for coupled noisy phase oscillators. *Nonlinearity*, 25(5):1247, 2012.

[19] 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.

[20] M. Hairer, M. Hutzenthaler, and A. Jentzen. Loss of regularity for Kolmogorov equations, 2012. `arxiv.org/abs/1209.6035`

[21] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(6):727–766, 1998.

[22] N. V. Krylov. *Introduction to the theory of diffusion processes*, volume 142 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1995.

[23] E. Luçon. Quenched limits and fluctuations of the empirical measure for plane rotators in random media. *Electron. J. Probab.*, 16:792–829, 2011.

[24] F. Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003.

[25] M. Maródi, F. d'Ovidio, and T. Vicsek. Synchronization of oscillators with long range interaction: Phase transition and anomalous finite size effects. *Phys. Rev. E*, 66:011109, Jul 2002.

[26] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.

[27] S. Méléard and S. Roelly-Coppoletta. A propagation of chaos result for a system of particles with moderate interaction. *Stoch. Proc. Appl.*, 26(2):317–332, 1987.

[28] K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 1984.

[29] K. Oelschläger. A law of large numbers for moderately interacting diffusion processes. *Z. Wahrsch. Verw. Gebiete*, 69(2):279–322, 1985.

[30] 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.

[31] I. Omelchenko, B. Riemenschneider, P. Hövel, Y. Maistrenko, and E. Schöll. Transition from spatial coherence to incoherence in coupled chaotic systems. *Phys. Rev. E*, 85:026212, Feb 2012.

[32] G. D. Prato. Some remarks about backward Itô formula and applications. *Stoch. Anal. Appl.*, 16(6):993–1003, 1998.

[33] J. L. Rogers and L. T. Wille. Phase transitions in nonlinear oscillator chains. *Phys. Rev. E*, 54:R2193–R2196, Sep 1996.

[34] S. H. Strogatz and R. E. Mirollo. Stability of incoherence in a population of coupled oscillators. *J. Statist. Phys.*, 63(3-4):613–635, 1991.

[35] A.-S. Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.*, 56(3):311–336, 1984.

[36] 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.

[37] J. Touboul. Propagation of chaos in neural fields, 2011. `arxiv.org/abs/1108.2414`

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

[39] G. Wainrib and J. Touboul. Topological and dynamical complexity of random neural networks. *Phys. Rev. Lett.*, 110:118101, Mar 2013.

Institut für Mathematik
Technische Universität Berlin
Strasse des 17. Juni 136
D-10623 Berlin
Germany
E-mail: lucon@math.tu-berlin.de
          stannat@math.tu-berlin.de

Bernstein Center for Computational Neuroscience
Philippstr. 13
D-10115 Berlin
Germany