HAL
open science

Long time dynamics for interacting oscillators on graphs

Fabio Coppini

To cite this version:

Fabio Coppini. Long time dynamics for interacting oscillators on graphs. The Annals of Applied Probability, 2022, 32, pp.360-391. 10.1214/21-AAP1680 . hal-02263485v2

HAL Id: hal-02263485

https://hal.science/hal-02263485v2

Submitted on 18 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LONG TIME DYNAMICS FOR INTERACTING OSCILLATORS ON GRAPHS

FABIO COPPINI

Abstract. The stochastic Kuramoto model defined on a sequence of graphs is analyzed: the emphasis is posed on the relationship between the mean field limit, the connectivity of the underlying graph and the long time behavior. We give an explicit deterministic condition on the sequence of graphs such that, for any finite time and any initial condition, even dependent on the network, the empirical measure of the system stays close to the solution of the McKean-Vlasov equation associated to the classical mean field limit. Under this condition, we study the long time behavior in the subcritical and in the supercritical regime: in both regimes, the empirical measure stays close to the (possibly degenerate) manifold of stable stationary solutions, up to times which can diverge as fast as the exponential of the size of the system, before Large Deviation phenomena take over. The condition on the sequence of graphs is derived by means of Grothendieck's Inequality and expressed through a concentration in $\ell_\infty \to \ell_1$ norm. It is shown to be satisfied by a large class of graphs, random and deterministic, provided that the average number of neighbors per site diverges, as the size of the system tends to infinity.

2010 Mathematics Subject Classification: 60K35, 82C20, 82C31, 82C44.

Keywords and phrases: Interacting oscillators, Long time dynamics, Kuramoto model, Random Graphs, Stochastic partial differential equations, Cut-norm, Grothendieck's Inequality, Self-normalized processes

1. Introduction

1.1. **Synchronization of mean field systems on graphs.** In recent years, synchronization of complex networks has become a very important topic for explaining real world phenomena. While in the physics literature the analysis has been pushed quite far and several extended reviews are available (e.g. [15, 33]), from a mathematical point of view these studies and the associated numerical simulations, can be regarded more as heuristic arguments than conclusive proofs.

The mathematical community has started working on particle systems on (random) graphs from the statistical mechanics point of view in the equilibrium regime and, with respect to the graph setting, assuming a locally tree-like structure (e.g [14]). Only in the last few years the attention has been focused on the dynamics of weakly interacting particles, tackling mean field systems on graphs, and their relationship with the corresponding thermodynamical limit (e.g. [5, 13]). These results, and the one presented here, are obtained for graphs in an intermediate regime between the sparse and the dense case, i.e. if G_n has n vertices and np_n represents the average number of edges, then $1 \ll np_n \leqslant n$. In the case of sparse graphs, i.e. $np_n = O(1)$, the limiting system seems to show a different phenomenology ([23, 29]).

Today, many results on the behavior of the empirical measure of such systems are available ([5, 10, 13, 25, 32]), but there is no agreement on the *weakest* hypothesis the class of graphs should satisfy in order to obtain the classical mean field limit. It turns

out that, depending on the setting one is considering, i.e. the normalization chosen in the interaction and/or the hypothesis on the initial data, different requirements on the graph may be asked.

To the author's knowledge, there exists no result on the longtime dynamics of a system defined on a sequence of graphs and the question whether the network is influencing the dynamics on long time scales, is still open and very much awaited with regards to applications.

In this work, we attack these issues by considering a well known model of synchronization defined on a sequence of graphs: we consider the Kuramoto model (e.g. [1]) for which an extensive literature is available and many tools have now been developed ([3, 4, 17]). For the sake of clarity, we study the model without the natural frequencies but our techniques apply as well in the quenched setting. We look for a result of mean field type with the minimal hypothesis on the initial conditions, i.e. the weak convergence of the empirical measure only, and by proposing a (deterministic) condition on the sequence of graphs which is shown to be satisfied by a large class of *homogeneous* graphs, including Erdős-Rényi random graphs with diverging average degree.

Finally, we show that the condition on the graph is not only sufficient for the system to converge to the mean field limit on bounded time intervals, but also that it is enough to study it on longer time scales. Namely, we push our analysis to the Large Deviation barrier of exponential time scales showing that, if the system synchronizes, then it keeps synchronized for long times.

1.2. **The model.** For each $n \in \mathbb{N}$, let $\xi^{(n)}$ be the adjacency matrix of a graph $(V^{(n)}, E^{(n)})$ with n vertices:

$$V^{(n)} = \{1, \dots, n\}, \quad E^{(n)} = \left\{(i,j) \in V^{(n)} \times V^{(n)} : \xi^{(n)}_{ij} \geqslant 1\right\}. \tag{1.1}$$

We consider both directed and undirected graphs as well as multigraphs so that $\xi^{(n)}_{ij}$ can take values in $\{1, \dots, n\}$ and not need to be equal to $\xi^{(n)}_{ji}$. We denote the corresponding (multi)graph by $\xi^{(n)}$ itself. Together with $\xi^{(n)}$, we consider a *dilution* parameter $p_n \in (0,1]$ representing the average density of neighbors per site. The two quantities will be coupled so that it is useful to think of them as one single object, we refer to Subsection 1.4 for the precise condition we required on it.

Given $\left(\xi^{(n)}, p_n\right)$, let $\{\theta^{i,n}_\cdot\}_{i=1,\dots,n}$ be the family of oscillators on $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$, which satisfy:

$$\begin{cases} \mathrm{d}\theta^{i,n}_t = \frac{1}{np_n} \sum_{j=1}^n \xi^{(n)}_{ij}\, J(\theta^{i,n}_t - \theta^{j,n}_t)\, \mathrm{d}t + \mathrm{d}B^i_t, \quad \text{for } t > 0, \\ \theta^{i,n}_0 = \theta^i_0, \quad \text{for } i \in \{1, \dots, n\}, \end{cases} \tag{1.2}$$

where $J(\cdot) = -K \sin(\cdot)$ with $K \geqslant 0$. Denote by $\mathbf{P}$ the law induced by $\left\{B^i_\cdot\right\}_{i\in\mathbb{N}}$ which are independent and identically distributed (IID) Brownian motions on $\mathbb{T}$ and by $\left\{\theta^i_0\right\}_{i\in\mathbb{N}}$ the initial conditions. We consider both deterministic and random initial data and, whenever they are random, they have to be independent of the Brownian motions.

If $\{\xi^{(n)}_{ij}\}_{i,j}$ are symmetric, i.e. $\xi^{(n)}_{ij} = \xi^{(n)}_{ji}$ for $1 \leqslant i < j \leqslant n$, then the model is reversible (e.g. [3]) with respect to the probability measure on $\mathbb{T}^n$ given by

$$\pi^{(n)}(\,\mathrm{d}\theta) = \frac{1}{Z^{(n)}} \exp\left(-\frac{K}{n} \sum_{i,j=1}^n \xi^{(n)}_{ij} \cos(\theta^i - \theta^j)\right) \lambda_n(\,\mathrm{d}\theta), \tag{1.3}$$

where $Z^{(n)}$ is the normalizing constant and λ_n the uniform probability measure on $\mathbb{T}^n$.

The main quantity of interest in system (1.2) is the empirical measure μ_t^n associated to $\{\theta_t^{i,n}\}_{i=1,\dots,n}$ and it is defined for all $t \geqslant 0$ by

$$\mu_t^n := \frac{1}{n}\sum_{j=1}^n \delta_{\theta_t^{j,n}} \in \mathcal{P}(\mathbb{T}), \tag{1.4}$$

the space of probability measure on the torus being denoted by $\mathcal{P}(\mathbb{T})$.

1.3. **The reversible Kuramoto model and its mean field limit.** When $\xi_{ij}^{(n)} = 1$ for $1 \leqslant i, j \leqslant n$ and $p_n \equiv 1$ for all $n \in \mathbb{N}$, i.e. $\xi^{(n)}$ is the complete graph, system (1.2) becomes:

$$\begin{cases} \mathrm{d}\bar{\theta}_t^{i,n} = (J * \bar{\mu}_t^n)(\bar{\theta}_t^{i,n})\,\mathrm{d}t + \mathrm{d}B_t^i, & \text{for } t > 0, \\ \bar{\theta}_0^{i,n} = \theta_0^i, & \text{for } i \in \{1,\dots,n\}, \end{cases} \tag{1.5}$$

where $\bar{\mu}_t^n := \frac{1}{n}\sum_{j=1}^n \delta_{\bar{\theta}_n^{j,n}}$ is the associated empirical measure and $*$ stands for the convolution. We refer to (1.5) as the reversible Kuramoto model (e.g. [3]).

It is well known (e.g. [3, Proposition 3.1]) that for all fixed time T, $\bar{\mu}_{t\in[0,T]}^n$ seen as a continuous function over $\mathcal{P}(\mathbb{T})$, weakly converges in $\mathcal{C}^0([0,T],\mathcal{P}(\mathbb{T}))$ to a deterministic limit $\mu_\cdot \in \mathcal{C}^0([0,T],\mathcal{P}(\mathbb{T}))$ that is solution to the following partial differential equation (PDE):

$$\begin{cases} \partial_t \mu_t(\theta) = \frac{1}{2}\partial_\theta^2 \mu_t(\theta) - \partial_\theta[\mu_t(\theta)(J*\mu_t)(\theta)], & \text{for } \theta \in \mathbb{T},\ 0 < t \leqslant T, \\ \mu_t\,|_{t=0} = \mu_0, \end{cases} \tag{1.6}$$

provided that μ_0^n weakly converges to μ_0 in $\mathcal{P}(\mathbb{T})$. If μ_0 does not have a density, than (1.6) has to be intended in the weak sense; however the regularity properties of the Laplacian operator make μ_t smooth for all $t > 0$ (see again [3, Proposition 3.1]). Equation (1.6) is often called McKean-Vlasov or Fokker-Planck equation and we refer to its solution $\mu_\cdot$ as to the mean field limit of the diffusions solving (1.5).

We recall here the most important results on (1.6), without giving any proof but referring to [17] (and references therein) where a complete analysis of the global dynamics is presented.

As for the mean field limit of the classical Kuramoto model, (1.6) is known to admit a phase transition depending on the coupling strength K: in the subcritical regime, for $0 \leqslant K < K_c := 1$, the particles behave as they were independently distributed on the circle; in the supercritical regime, for $K > 1$, they tend to synchronize around the same phase. We do not consider the critical case $K = 1$, since it does not add anything to the purpose of this work.

More precisely, in the subcritical regime there is a unique stationary solution which corresponds to the incoherent state $\frac{1}{2\pi}$, the uniform measure on the torus (see [17, Proposition 4.1]). It is globally attractive and the linear operator around it has negative spectrum bounded away from zero: we will make use of this property showing that the fluctuations given by the graph structure are controlled for all times, whereas the random fluctuations given by the Brownian motions are not and will make the system escape from $1/2\pi$ after some (very long) time, i.e. a Large Deviation phenomenon.

In the supercritical regime, when $K > 1$, there is a manifold of stable stationary solutions corresponding to the synchronous states of the oscillators $\{\theta_\cdot^{i,n}\}_{i=1,\dots,n}$ (see [17,

Subsection 4.3] and [3]). Up to a rotation, all stable stationary solutions of (1.6) are given by

$$q(\theta) = \frac{\exp\{2Kr\cos(\theta)\}}{Z}, \tag{1.7}$$

where Z is the normalizing constant and $r = r(K)$ is the unique solution in $(0,1)$ of a fixed point equation $r = \Psi(2Kr)$, see [3] for a explicit formula of Ψ. The parameter r is often referred to as the degree of synchronization of the system: r close to 0 indicates that the particles are scattered around the circle, r close to 1 that they are almost fully synchronized. We just recall that whenever $K < 1$, the fixed point equation has a unique solution $r = 0$, which in (1.7) boils down to the uniform measure $1/2\pi$, and whenever $K > 1$ the value $r = 0$ is still a solution but the corresponding measures solving (1.6) are unstable so that we will not consider them.

Let $K > 1$ and $0 < r < 1$. Observe that system (1.5) (and also (1.2)) is invariant under rotations, this property is maintained in the limit (1.6) and the manifold of stationary solutions M can be described as

$$M = \{q_\psi \,:\, q_\psi(\cdot) = q(\cdot - \psi),\, \psi \in \mathbb{T}\}\,. \tag{1.8}$$

It is possible to show that, unless one starts from the unstable manifold

$$U = \left\{\mu \in \mathcal{P}(\mathbb{T}) \,:\, \int_{\mathbb{T}} \exp(i\theta)\mu(\,\mathrm{d}\theta) = 0\right\}, \tag{1.9}$$

the measure μ_t solution to (1.6) converges to some $q_\psi \in M$ as t tends to infinity, the phase $\psi \in \mathbb{T}$ depending only on μ_0. Since each $q \in M$ is a stationary solution, the dynamics of μ_t is fully characterized for all times t.

1.4. **The graph's perspective.** The aim of this work is to investigate the weakest assumptions on the sequences $\xi = \left\{\xi^{(n)}\right\}_{n\in\mathbb{N}}$ and $\{p_n\}_{n\in\mathbb{N}}$, such that the long time behavior of (1.2) is well understood: in other words, whenever system (1.2) is comparable to (1.5) or to the mean field limit (1.6), under a proper scale between size of the system n and some horizon time T_n.

The normalization sequence p_n has to be chosen such that the interaction term in (1.2) makes sense. At least, this requires the assumption that the quantity

$$\frac{1}{np_n}\sum_{j=1}^{n} \xi_{ij}^{(n)} \tag{1.10}$$

is of order one, for almost each vertex i in the graph.

Remark 1.1. *Observe that whenever* (1.10) *converges to zero or diverges, one should look for a different normalization in order to obtain a proper limit. A control on* (1.10) *is thus required to exclude degenerate cases, yet it cannot be sufficient for our purpose: whenever one considers a graph composed of two (or more) highly connected components, the degree of each vertex can be correctly defined, but one cannot expect the convergence of the empirical measure since the behavior on each component may differ, depending on the initial conditions! We refer to* [10, Remark 1.2] *and* [13, Remark 1.4] *for concrete examples and a precise analysis from this perspective, see also Remark 2.2 in the next section.*

For $n \in \mathbb{N}$, define the normalized adjacency matrix $P^{(n)} = \{P^{(n)}_{ij}\}_{i,j=1,\dots,n}$ by

$$P^{(n)}_{ij} := \frac{\xi^{(n)}_{ij}}{p_n}, \qquad \text{for } i,j = 1,\dots,n. \tag{1.11}$$

Recall that we do not assume any symmetry on $\xi^{(n)}$ and that it can also represent a multigraph. Define $\mathbf{1}^{(n)}$ as the adjacency matrix associated to the classical mean field model, i.e. $\mathbf{1}^{(n)}_{ij} = 1$ for $i,j = 1,\dots,n$. One would like to compare $P^{(n)}$ to $\mathbf{1}^{(n)}$.

It turns out that a sufficient condition for what we aim at, is given by a control on the difference between $P^{(n)}$ and $\mathbf{1}^{(n)}$ through the $\ell_\infty \to \ell_1$ norm. This norm is defined for a matrix $G = \{G_{ij}\}_{i,j=1,\dots,n}$ as

$$\|G\|_{\infty\to 1} := \sup_{\|s\|_\infty \leqslant 1} \|Gs\|_1 = \sup_{s,t\in\{-1,1\}^n} Gst^\top = \sup_{s_i,t_j\in\{-1,1\}} \sum_{i,j=1}^n G_{ij}s_it_j. \tag{1.12}$$

It has received a lot of attention in the last years: it appears in many applications in computer science (e.g. [20]) and it has been shown to be very useful in graphs concentration (e.g. [19, 24, 28]). Part of this success is because of the equivalence to the cut-norm (e.g. [2]) and, as already remarked in [19, 32], of Grothendieck's Inequality, which is recalled hereafter.

Theorem 1.2 (Grothendieck's inequality, [31, Theorem 2.4])**.** *Let $\{a_{ij}\}_{i,j=1,\dots,n}$ be a $n\times n$ real matrix such that for all $s_i, t_j \in \{-1,1\}$*

$$\sum_{i,j=1}^n a_{ij}s_it_j \leqslant 1. \tag{1.13}$$

Then, there exists a constant $K_R > 0$, such that for every Hilbert space $(H, \langle\cdot,\cdot\rangle_H)$ and for all S_i and T_j in the unit ball of H

$$\sum_{i,j=1}^n a_{ij}\langle S_i, T_j\rangle_H \leqslant K_R. \tag{1.14}$$

It is indeed thanks to this inequality that $\ell_\infty \to \ell_1$ norm turns out to be the natural choice for our setting: an important part of the proof (Lemmas 3.2 and 4.3) consists in showing that the fluctuations due to the graph structure can be described by expressions like (1.14), and thus controlled by $\|\cdot\|_{\infty\to 1}$.

From now on, the only condition we require on $\left(\xi^{(n)}, p_n\right)_{n\in\mathbb{N}}$ is to satisfy:

$$\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1} = o(n^2), \tag{1.15}$$

or, in other words,

$$\lim_{n\to\infty} \sup_{s_i,t_j\in\{-1,1\}} \frac{1}{n^2}\sum_{i,j=1}^n \left(\frac{\xi^{(n)}_{ij}}{p_n} - 1\right) s_it_j = 0. \tag{1.16}$$

In Proposition A.3 it is shown that Erdős-Rényi random graphs with parameter p_n satisfy condition (1.16) almost surely, provided that $np_n \uparrow \infty$. We also provide a class of deterministic graphs, Ramanujan graphs, that satisfies (1.16) (see Proposition A.6) and give some link with the theory of graphons.

Appendix A presents such results and includes remarks on the relationship between condition (1.16), the degree condition (1.10) and the connectivity of $\{\xi^{(n)}\}_{n\in\mathbb{N}}$.

1.5. **Set-up and notations.** The closeness between μ_t^n and μ_t is studied through a norm which controls the bounded Lipschitz (or 1-Wasserstein) distance between probability measures, in an appropriate class of weighted Hilbert spaces $H_{-1,w}$. This class is defined as follows.

Denote by $\mathcal{C}_0^1(\mathbb{T})$ the space of $\mathcal{C}^1$ functions on the torus with zero mean and consider

$$\mathcal{L}_0^2 = \left\{ f \in \mathcal{L}^2(\mathbb{T}) : \int_{\mathbb{T}} f = 0 \right\}, \tag{1.17}$$

with canonical scalar product $(u,v) := \int_{\mathbb{T}} uv$, for $u,v \in \mathcal{L}_0^2$. Let $w \in \mathcal{C}^1(\mathbb{T},(0,\infty))$ and V be the closure of $\mathcal{C}_0^1(\mathbb{T})$ with respect to the norm $\|\varphi\|_{H_{1,1/w}} = \sqrt{\int_{\mathbb{T}} \frac{(\varphi')^2}{w}}$ for $\varphi \in \mathcal{C}_0^1(\mathbb{T})$. It is easy to see that V is continuously and densely injected in $\mathcal{L}_0^2$ (thanks to the compactness of $\mathbb{T}$ and Poincaré inequality). Moreover, one can define an inner product on V which makes it an Hilbert space $H_{1,1/w} := (V, \langle\cdot,\cdot\rangle_{H_{1,1/w}})$ where $\langle\varphi,\psi\rangle_{H_{1,1/w}} = \int_{\mathbb{T}} \frac{\varphi'\psi'}{w}$ for all $\varphi,\psi \in \mathcal{C}_0^1(\mathbb{T})$. The dual space of $H_{1,1/w}$ is denoted by $H_{-1,w}$. Observe that if $u,v \in \mathcal{L}_0^2$ and $v \in H_{1,1/w}$, then $u \in H_{-1,w}$ and

$$u(v) := \langle u,v\rangle_{-1,1} = (u,v), \tag{1.18}$$

where $\langle\cdot,\cdot\rangle_{-1,1}$ denotes the action of $H_{-1,w}$ on $H_{1,1/w}$, we omit the weight w.

The action of a probability measure μ on a test function h is denoted by $\langle\mu,h\rangle = \int h \,\mathrm{d}\mu$: of course whenever u and v are regular enough, one has $u(v) = \langle u,v\rangle_{-1,1} = \langle u,v\rangle = (u,v)$, where we have abused of notation, denoting the density of a probability measure by the probability measure itself.

Finally, observe that different weights w give equivalent norms so that whenever the geometry of the space is not important, we consider the case $w \equiv 1$ and simply note $\|\cdot\|_{-1}$. More information about the construction of $H_{-1,\omega}$ are given in Appendix B.

Hereafter we drop the dependency on $\mathbb{T}$, i.e. we write $\mathcal{C}_0^1$ instead of $\mathcal{C}_0^1(\mathbb{T})$ and so on for the other spaces and integrals.

2. Main results

We present the results in three consecutive subsections: we start by the finite time behavior, then pass to the supercritical regime and, finally, the subcritical case.

In all results, the convergence of empirical measures is stated in the norm $\|\cdot\|_{-1}$. It is not difficult to see that the difference of two probability measures belongs to H_{-1} and that the distance induced on $\mathcal{P}(\mathbb{T})$ controls the bounded Lipschitz distance (or, equivalently, the 1-Wasserstein distance). These details are covered in Appendix B.

Recall that throughout the paper, we only require $\left(\xi^{(n)}, p_n\right)_{n\in\mathbb{N}}$ to satisfy condition (1.16) and $\mu_0 \in \mathcal{P}(\mathbb{T})$, no independence between μ_0^n and $\xi^{(n)}$ is demanded.

2.1. **The finite time behavior.** We give the result and then comment it.

Theorem 2.1. *Let $K \geqslant 0$. Suppose that for all $\varepsilon_0 > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left(\|\mu_0^n - \mu_0\|_{-1} \leqslant \varepsilon_0\right) = 1. \tag{2.1}$$

Then, for every fixed time $T > 0$ and for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T]} \|\mu_t^n - \mu_t\|_{-1} \leqslant \varepsilon\right) = 1. \tag{2.2}$$

The finite time behavior of weakly interacting particle systems on graphs is already known under suitable hypothesis on the initial conditions and on the graph sequence, we refer to Subsection 2.5 for a comprehensive literature on the subject. We decide to present Theorem 2.1 because, contrary to all the previous results, it does not require any independence between initial conditions and (the realization of) the sequence of graphs. In particular, even if one accurately assigns the initial conditions for each vertex, the mixing properties of the graph will shuffle all the information and make the empirical measure converge, loosing any memory of the initial coupling. This property is crucial for studying the longtime behavior as pointed out in the next subsections.

Observe that Theorem 2.1 implies the existence of a unique giant component in $\{\xi^{(n)}\}_{n\in\mathbb{N}}$, as pointed out in the next remark.

Remark 2.2. *The result is independent of K. First observe that this implies the uniqueness of a giant component: if there are two, then one can accurately prepare the initial conditions so to obtain different behaviors on the twos and loose the proximity to* (1.6)*. Secondly, with the same argument one deduces that the size of the giant component is asymptotically n, i.e. all but $o(n)$ vertices are connected. Finally, the existence comes from the fact that the system cannot synchronize on components of size $o(n)$, no matter the value of K. Lemma A.2 shows that condition* (1.10) *indeed implies the existence of a giant component of size asymptotically n.*

2.2. **Long time behavior in the supercritical regime.** In the supercritical regime, we suppose to be already close to the manifold M at time 0. However, since we do not assume any independence between graph and initial data, this hypothesis can be weakened by requiring the initial condition μ_0 to be in the domain of attraction of M, i.e. $\mu_0 \in \mathcal{P}(\mathbb{T}) \setminus U$, and using Theorem 2.1. One can then start after some time T with initial condition given now by μ_T^n (and dependent on the graph!): if T is big enough, than μ_T^n will be close to M. Observe that the choice of T depends only on how close to M μ_t has to be, it thus depends only on μ_0.

Before stating the theorem, we define the distance of a probability measure from M. For $\mu \in \mathcal{P}(\mathbb{T})$, let

$$\operatorname{dist}(\mu, M) := \inf_{\nu\in M} \|\mu - \nu\|_{-1}. \tag{2.3}$$

We are ready for the main result of this section.

Theorem 2.3. *Let $K > 1$. Suppose there exists $\psi \in \mathbb{T}$ such that for every $\varepsilon_0 > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left(\|\mu_0^n - q_\psi\|_{-1} \leqslant \varepsilon_0\right) = 1. \tag{2.4}$$

Then, for every positive sequence $\{T_n\}_{n\in\mathbb{N}}$ *such that* $T_n = \exp(o(n))$*, and for all* $\varepsilon > 0$ *small enough*

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T_n]} \operatorname{dist}(\mu_t^n, M) \leqslant \varepsilon\right) = 1. \tag{2.5}$$

Theorem 2.3 implies the proximity of the empirical measure to the manifold of solutions of the McKean-Vlasov equation (1.6) for almost exponential times. On this time scale, Large Deviation phenomena take control of the finite system (e.g. [11, 16, 30]) making it escape from the stationary solutions.

Observe that we do not prove the closeness to the mean field limit $\mu_\cdot$. Indeed, it is by now well known that, on longtime scales, the mean field limit is not a faithful description of the finite system of n diffusions. In other words, the behavior of $\mu_{T_n}^n$ highly depends on the scale of time T_n under consideration, whereas the dynamics of μ_t is deterministic and completely known for large t, i.e. it sticks to q_ψ.

In [4], a deep analysis of the longtime dynamics for the classical mean field system (1.5) is presented. Namely, it is shown that μ_t solution to the PDE (1.6) is a reasonable approximation of $\bar{\mu}_t^n$ for times scales of order $o(\log n)$. On times proportional to n, the dynamics of the empirical measure can be coupled to a Brownian motion on M with a non trivial diffusion coefficient that can be explicitly computed (see [4, Theorem 1.1]). Whereas the PDE prescribes the system to stay synchronized on a fixed phase, the noise induced by the Brownian motions makes this phase oscillate and it turns out that the oscillations become significant on times proportional to the size of the system n.

We do not show this property, yet extend the closeness to M for exponential times, whereas in [4] this is shown up to polynomial times.

Theorem 2.3, as Theorem 2.1, does not depend on the speed of convergence of the condition on the graph (1.16). The escaping time is indeed only due to the stochastic nature of the system, given by the Brownian motions, and it cannot be improved as explained above. The reason why one can control the perturbation induced by the graph structure for long times is somehow hidden in the martingale properties of $\mu_\cdot^n$ and in the fact that we do not really analyze the dynamics near M (which can, a priori, depend on the graph). We refer to the proof of the subcritical regime for a clear control on the perturbations given by the graph, through the exponential stability of the stationary solution.

2.3. **Longtime behavior in the subcritical regime.** The subcritical regime is somehow easier than the supercritical regime since there is an unique stable stationary solution. We decide to include this case firstly because, to the author's knowledge, it is missing in the literature and, secondly, because the proof enlightens some aspects hidden in the supercritical regime. As a byproduct, we obtain the equivalent of maximal inequalities for Ornstein-Uhlenbeck processes in infinite dimensional Hilbert spaces, see Corollary 2.5.

Theorem 2.4. *Let* $0 \leqslant K < 1$*. Suppose that condition* (2.1) *holds, i.e. for all* $\varepsilon_0 > 0$

$$\lim_{n\to\infty} \mathbf{P}\left(\|\mu_0^n - \mu_0\|_{-1} \leqslant \varepsilon_0\right) = 1. \tag{2.6}$$

Then, for every positive sequence $\{T_n\}_{n\in\mathbb{N}}$ *such that* $T_n = \exp(o(n))$*, and for all* $\varepsilon > 0$ *small enough*

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T_n]} \|\mu_t^n - \mu_t\|_{-1} \leqslant \varepsilon\right) = 1. \tag{2.7}$$

Observe that, since $\mu_\cdot$ converges as t tends to infinity to $1/2\pi$ for all initial conditions μ_0, then Theorem 2.4 implies the proximity of μ_t^n to the stable solution up to exponential times.

Of independent interest, we present a corollary of Theorem 2.4 in the limit case $K = 0$. This result seems to be well known, yet the author was unable to find it elsewhere.

Corollary 2.5. *Let $\mu_\cdot^n$ be the empirical measure of n independent Brownian motions $\{B_\cdot^{j,n}\}_{j=1,\dots,n}$ on $\mathbb{T}$ with initial conditions $\{\theta_0^i\}_{1\leqslant i\leqslant n}$ satisfying (2.1). Then, there exist $C > 0$ and $T_0 > 0$ such that for all $T > T_0$, the following maximal inequality holds:*

$$\mathbf{E}\left[\sup_{t\in[T_0,T]}\left\|\mu_t^n - \frac{1}{2\pi}\right\|_{-1}^2\right] \leqslant C\,\frac{\log(1+T-T_0)}{n}. \tag{2.8}$$

Corollary 2.5 shows a maximal inequality for the empirical measure of n independent Brownian motions on the torus, establishing the SPDE version of the result for Ornstein-Uhlenbeck processes presented in [18] for stochastic ordinary differential equations.

Observe that if the initial conditions and the graph are exchangeable (not necessarily independent), then Corollary 2.5 and a classical result by Sznitman ([34, Proposition 2.2]) implies the creation of chaos for all times $T_n = o(\exp(n))$.

2.4. **Organization of the paper.** This section ends presenting the existing literature and giving an outline of the proof for the three theorems.

Sections 3, 4 and 5 concern the proofs of the three results. In particular, Section 3 is devoted to the long time dynamics close to M, it starts from the derivation of a mild formulation for the empirical measure, then proceeds with the control on the graph and the noise, and it ends with the proof of Theorem 2.3. Section 4 concerns the subcritical regime where a different control on the perturbations is given. Finally, Section 5 proves Theorem 2.1 by using slight variations of the previous techniques.

Appendix A gives a few examples of graph sequences that satisfy condition (1.16), together with remarks on the degrees and connectivity of such sequences. Appendix B contains information about the Hilbert spaces $H_{-1,\omega}$ and the linear operator L_ψ.

2.5. **A glance at the existing literature.** The results presented here are at a crossroads of two different research areas: the long time dynamics of stochastic differential equations and the role of a network in a mean field model.

Concerning the long time behavior of weakly interacting particle systems, Theorems 2.3 and 2.4 can be seen as a complement to the previous results presented in [4], filling the gap of the exponential time scale which has not been addressed so far. To the author's knowledge, they also represent the first equivalent, in infinite dimensional Hilbert spaces, to the famous result for stochastic ordinary differential equations in $\mathbb{R}^d$ by Friedlin and Wentzell ([16]).

Looking at variations on the same model, the behavior of the classical Kuramoto model with intrinsic frequencies has been studied in [26], showing that the longtime dynamics is indeed dependent on the quenched setting given by the frequencies. A macroscopic constant speed in the phase appears on time scales of order $O(\sqrt{n})$, making the effects of the noise vanish. The results and the techniques presented here should be easily adaptable to this case showing the proximity to the manifold of solutions for long times, yet loosing the precise characterization of the motion on M for which a deeper analysis is needed.

The Kuramoto model is an example of system which admits more than one stable stationary solution, a continuous manifold as already precised, and that's one of the reasons why it shows a rich phenomenology depending on the time scale under consideration. For similar results on different models, one has to dip in the context of stochastic partial differential equations (SPDEs) with vanishing noise. Since the aim of this work is more oriented on the effects of the network rather than the longtime dynamics of SPDEs, the author refers to the bibliography in [4, 26] for a more comprehensive discussion.

Turning to interacting particle systems on graphs, the subject has become an interesting topic in the mathematical community given the several applications to complex systems, in particular regarding the Kuramoto model and synchronization phenomena (e.g. [1, 33]), yet it has always been addressed on a finite time scale or up to times slowly diverging on n, i.e. $T_n = O(\log n)$.

The first articles [5, 13] attack the problem under a propagation of chaos viewpoint, requiring independent and identically distributed initial conditions and also independent of the realization of the graph. In this setting, the condition on the graph boils down to a condition on the degrees only so that very general graphs are allowed (see again [13, Remark 1.4] and [10, Remark 1.2]). Regarding more inhomogeneous settings, [25] extends [13] to graphons and [32] presents a Large Deviation result again in the graphon setting. Observe that [32] already makes use of Grothendieck's inequality and the norm $\|\cdot\|_{\infty\to 1}$ to control the graphs fluctuations. Up to now, the only result not assuming independence in the initial data is given by [10], where general systems of interacting particles are defined on Erdősh-Rényi random graphs and the empirical measure is shown to satisfy a Law of Large Number and a Large Deviation principle, implying the convergence to the respective mean field limit.

In the deterministic setting, the Kuramoto model has been studied on different networks with various hypothesis on the initial conditions, we refer to [9, 27] and references therein.

In all the cited works, the condition on the normalization p_n is slightly stronger, or equivalent, to the one required in (1.16), see in particular Propositions A.4 and A.6 in Appendix A. If np_n is not diverging as n tends to infinity, i.e. the case of sparse graphs, the limiting behavior of the empirical measure seems to be rather different from the mean field limit, see [23, 29].

2.6. **Outline of the proofs.** The three theorems are proven in a similar way and the main ingredients are given by

(1) A mild formulation satisfied by $\mu^n_\cdot$ for each $n \in \mathbb{N}$;
(2) The control on the perturbations given by the graph structure through Grothendieck's Inequality;
(3) The control on the random perturbations given by the Brownian motions through maximal inequalities for self-normalized processes.

The mild formulation will be different in all the three cases and will depend on the linear dynamics around M or $1/2\pi$ or on the properties of μ_t. In Section 3, we give a full derivation of the stochastic partial differential equation satisfied by $\mu^n_t - q_\psi$ in a neighborhood of M.

The control on the graph will also depend whether there is a strong contraction given by the dynamics, or not. Whenever the evolution is contracting in all direction, as in the subcritical case around $1/2\pi$, these perturbations can be controlled uniformly in time, we refer to Lemma 4.3 for a precise statement.

A fine control on the random perturbations turns out to be rather delicate and one has to exploit all the properties associated to the Hilbert structure as well as the ones associated to the linear dynamics around the stationary solutions to get the job done. We give two independent explicit proofs:

- Around M, we study the noise using of a strong result on self-normalized martingales ([12]);
- Around $1/2\pi$, we extend a result on maximal inequalities for Ornstein-Uhlenbeck processes ([18]), to an infinite dimensional setting.

Once the perturbations are controlled, a Gronwall-like lemma is used to bound the difference between $\mu^n_\cdot$ and the relative target. In the supercritical case, we need to set up an (easy) iterative scheme in order to estimate the distance between $\mu^n_\cdot$ and M on bounded time intervals, and then make use of the martingale property of system (1.2) to extend the result up to almost exponential times; in the subcritical case, the result is directly obtained by the bound on the noise, the graph perturbations are indeed controlled for all times.

3. Longtime dynamics close to M

In a neighborhood of M, one can exploit the properties of the linear dynamics around q_ψ. For $u \in \mathcal{L}^2_0$, let

$$L_\psi u \;:=\; \tfrac{1}{2}\partial_\theta^2 u - \partial_\theta \left[u(J * q_\psi) + q_\psi (J * u)\right], \tag{3.1}$$

be the linear operator at q_ψ, its domain is given by $D(L_\psi) = \{u \in \mathcal{C}^2(\mathbb{T}),\ \int_{\mathbb{T}} u(\theta)\,\mathrm{d}\theta = 0\}$. The operator L_ψ is self-adjoint in $H_{-1,1/q_\psi}$ and its adjoint L^*_ψ in $\mathcal{L}^2_0$ has the following expression

$$L^*_\psi u \;=\; \tfrac{1}{2}\partial_\theta^2 u + (J * q_\psi)\partial_\theta u - (J * q_\psi\, \partial_\theta u) - \int_{\mathbb{T}} (J * q_\psi\, \partial_\theta u), \tag{3.2}$$

and domain $D(L^*_\psi) = D(L_\psi)$.

We recall here the most important properties of L_ψ, referring to Appendix B for more informations. The linear operator $-L_\psi$ has compact resolvent and its spectrum lies in $[0,\infty)$: the smallest eigenvalue $\lambda_0 := 0$, associated to the eigenfunction $\partial_\theta q_\psi$, is isolated from the rest of the spectrum. In particular, this implies that H_{-1} can be decomposed into a direct sum $T_\psi \oplus N_\psi$, where $T_\psi = \mathrm{Span}(\partial_\theta q_\psi)$; we denote by P^0_ψ the projection on T_ψ along N_ψ and $P^s_\psi = 1 - P^0_\psi$. Observe that both P^0_ψ and P^s_ψ commute with L_ψ (e.g. [21]).

Let $\{\lambda_0 < \lambda_1 \leqslant \dots\} \subset [0,\infty)$ denote the set of eigenvalues and let $\{e^\psi_l\}_{l=0,1,\dots}$ be the correspondent set of eigenfunctions, normalized in $H_{-1,1/q_\psi}$, i.e.

$$-L_\psi e^\psi_l = \lambda_l e^\psi_l, \qquad \text{for } l = 0,1,\dots. \tag{3.3}$$

Observe that $e^\psi_l \in \mathcal{C}^\infty(\mathbb{T})$. Moreover, the eigenvalues do not depend on the phase ψ, whereas the eigenfunctions do in a rather simple way given by the rotation symmetry of the system, i.e.

$$e^\psi_l(\cdot) = e^\theta_l(\cdot + \psi - \theta), \qquad \text{for } l = 0,1,\dots. \tag{3.4}$$

As a matter of fact, we will study the system only around some q_ψ. The dual eigenfunctions f^ψ_l associated to L^*_ψ will play an important role for studying the noise perturbation, their properties are studied in Proposition B.3.

From the previous properties one deduces that L_ψ (respectively L^*_ψ) generates a strong continuous semigroup on H_{-1} (resp. H_1) that we denote by e^{tL_ψ} (resp. $e^{tL^*_\psi}$). These semigroups have many properties that we will recall and use throughout this section, see Proposition B.2 for a general statement.

A final remark: we use the letter C for all the constants even if they are possibly different, the value of C can change from one line to another if the constant is replaced by another constant and the context is clear.

3.1. **The mild formulation around $q_\psi \in M$.** As shown in [4], $\bar{\mu}^n_\cdot$ satisfies a SPDE written in mild form once it is close to M; in this subsection we extend this formulation to μ^n_t. Let $\nu^n_\cdot := \mu^n_\cdot - q_\psi$, then

Proposition 3.1. *The process $\nu^n_t \in H_{-1}$ satisfies the following stochastic partial differential equation in $\mathcal{C}^0([0,T], H_{-1})$:*

$$\nu^n_t = e^{tL_\psi}\nu^n_0 - \int_0^t e^{(t-s)L_\psi}\partial_\theta\left[\nu^n_s(J*\nu^n_s)\right]\mathrm{d}s - g^n_t + z^n_t, \tag{3.5}$$

where

$$g^n_t = \frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) e^{(t-s)L_\psi}\partial_\theta\left[\delta_{\theta^{i,n}_s}(J*\delta_{\theta^{j,n}_s})\right]\mathrm{d}s, \tag{3.6}$$

and $z^n_t \in H_{-1}$ is defined for $h \in H_1$ by

$$\langle z^n_t, h\rangle_{-1,1} = \frac{1}{n}\sum_{j=1}^n \int_0^t \left[\partial_\theta e^{(t-s)L^*_\psi}h\right](\theta^{j,n}_s)\,\mathrm{d}B^j_s. \tag{3.7}$$

Proof. Let $F = F_t(\theta) \in \mathcal{C}^{1,2}([0,\infty)\times\mathbb{T})$, with $\int F_t = 0$ for all $t \geqslant 0$. For some $t \geqslant 0$, an application of Ito formula, together with the definition of L^*_ψ gives

$$\begin{aligned}\langle \mu^n_t - q_\psi, F_t\rangle &= \langle \mu^n_0 - q_\psi, F_0\rangle + \int_0^t \langle \mu^n_s - q_\psi, \partial_s F_s + L^*_\psi F_s\rangle\,\mathrm{d}s + \\ &\quad + \int_0^t \langle (\mu^n_s - q_\psi)(J*(\mu^n_s - q_\psi)), \partial_\theta F_s\rangle\,\mathrm{d}s + G^n_t(F) + Z^n_t(F),\end{aligned} \tag{3.8}$$

where

$$\begin{aligned}G^n_t(F) &= \frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) J(\theta^{i,n}_s - \theta^{j,n}_s)\partial_\theta F_s(\theta^{i,n}_s)\,\mathrm{d}s, \\ Z^n_t(F) &= \frac{1}{n}\sum_{j=1}^n \int_0^t \partial_\theta F_s(\theta^{j,n}_s)\,\mathrm{d}B^j_s.\end{aligned} \tag{3.9}$$

The properties of $e^{tL^*_\psi}$, see Proposition B.2, assure that the function

$$F = F_s(\theta) = e^{(t-s)L^*_\psi}h(\theta), \qquad \text{for some } h \in \mathcal{C}^2(\mathbb{T}),\ \int h = 0, \tag{3.10}$$

is $\mathcal{C}^{1,2}([0,t]\times\mathbb{T})$. But then $\partial_s F_s = -L^*_\psi F_s$ and one obtains

$$\langle \nu^n_t, F_t\rangle = \langle \nu^n_0, e^{tL^*_\psi}h\rangle + \int_0^t \langle \nu^n_s(J*\nu^n_s), \partial_\theta e^{(t-s)L^*_\psi}h\rangle\,\mathrm{d}s + g^n_t(h) + z^n_t(h), \tag{3.11}$$

where we have used the definition of ν_t^n and the notations

$$g_t^n(h) = \frac{1}{n^2}\sum_{i,j=1}^{n}\int_0^t \left(\frac{\xi_{ij}}{p_n}-1\right) J(\theta_s^{i,n}-\theta_s^{j,n})\left[\partial_\theta e^{(t-s)L_\psi^*}h\right](\theta_s^{i,n})\,\mathrm{d}s, \tag{3.12}$$

$$z_t^n(h) = \frac{1}{n}\sum_{j=1}^{n}\int_0^t \left[\partial_\theta e^{(t-s)L_\psi^*}h\right](\theta_s^{j,n})\,\mathrm{d}B_s^j. \tag{3.13}$$

We aim at proving that (3.11) is the weak formulation of the mild equation (3.5) in H_{-1}.

Let $\{\nu_l\}_{l\geqslant 1}\subset\mathcal{L}_0^2$ such that $\nu_l \xrightarrow{l\uparrow\infty} \nu_0^n$ in H_{-1}. Then, for $h\in\mathcal{C}^2$

$$\langle \nu_l, e^{tL_\psi^*}h\rangle_{-1,1} = \left(\nu_l, e^{tL_\psi^*}h\right) = \left(e^{tL_\psi}\nu_l, h\right) = \langle e^{tL_\psi}\nu_l, h\rangle_{-1,1}. \tag{3.14}$$

By continuity of the operators, $e^{tL_\psi}\nu_l$ converges in H_{-1} to $e^{tL_\psi}\nu_0^n$ as $l\uparrow\infty$. Taking the limit for $l\uparrow\infty$ in both sides of (3.14), we deduce

$$\langle \nu_0^n, e^{tL_\psi^*}h\rangle_{-1,1} = \langle e^{tL_\psi}\nu_0^n, h\rangle_{-1,1}. \tag{3.15}$$

We now focus on

$$\omega_s^n := \nu_s^n(J*\nu_s^n). \tag{3.16}$$

Consider $\{\nu_{s,l}\}_{l\geqslant 1}\subset\mathcal{L}_0^2$ which converges to ν_s^n in H_{-1} as $l\uparrow\infty$, and define

$$\omega_{s,l} := \nu_{s,l}(J*\nu_s^n). \tag{3.17}$$

For any $l\geqslant 1$, it holds

$$\begin{aligned}\langle \omega_{s,l}, \partial_\theta e^{(t-s)L_\psi^*}h\rangle_{-1,1} &= \left(\omega_{s,l}, \partial_\theta e^{(t-s)L_\psi^*}h\right) = \\ &= -\left(e^{(t-s)L_\psi}\partial_\theta\,\omega_{s,l}, h\right) = -\langle e^{(t-s)L_\psi}\partial_\theta\,\omega_{s,l}, h\rangle_{-1,1}.\end{aligned} \tag{3.18}$$

Using the properties of the semigroup one obtains

$$\begin{aligned}\left|\langle e^{(t-s)L_\psi}\partial_\theta(\omega_{s,l}-\omega_s^n), h\rangle_{-1,1}\right| &\leqslant \|h\|_1 \left\|e^{(t-s)L_\psi}\partial_\theta(\omega_{s,l}-\omega_s^n)\right\|_{-1} \leqslant \\ \leqslant \|h\|_1 \frac{C}{\sqrt{t-s}}\|\partial_\theta(\omega_{s,l}-\omega_s^n)\|_{-2} &= \|h\|_1\frac{C}{\sqrt{t-s}}\|\omega_{s,l}-\omega_s^n\|_{-1},\end{aligned} \tag{3.19}$$

which implies

$$\left\|e^{(t-s)L_\psi}\partial_\theta(\omega_{s,l}-\omega_s^n)\right\|_{-1} \leqslant \frac{C}{\sqrt{t-s}}\|\omega_{s,l}-\omega_s^n\|_{-1}. \tag{3.20}$$

Since h is regular and $\omega_{s,l}\xrightarrow{l\uparrow\infty}\omega_s^n$ in H_{-1}, this implies

$$\langle \omega_s^n, \partial_\theta e^{(t-s)L_\psi^*}h\rangle_{-1,1} = -\langle e^{(t-s)L_\psi}\partial_\theta\,\omega_s^n, h\rangle_{-1,1}. \tag{3.21}$$

We now observe from (3.20) that

$$\left\|e^{(t-s)L_\psi}\partial_\theta\omega_s^n\right\|_{-1} \leqslant \frac{C}{\sqrt{t-s}} \tag{3.22}$$

thus the integral in (3.5)

$$\int_0^t e^{(t-s)L_\psi}\partial_\theta\left[\nu_s^n(J*\nu_s^n)\right]\mathrm{d}s \tag{3.23}$$

is almost surely finite. Using [35, Theorem 1, p.133], we deduce that (3.23) makes sense as a Bochner integral in H_{-1}. The continuity is a direct consequence of the continuity of e^{tL_ψ}.

Assume that $g_t^n(h) = \langle g_t^n, h\rangle_{-1,1}$ and $z_t^n(h) = \langle z_t^n, h\rangle_{-1,1}$ are well defined and continuous with respect to t for all $h \in H_1$; we have shown that

$$\begin{aligned}\langle \nu_t^n, h\rangle_{-1,1} &= \langle e^{tL_\psi}\nu_0^n, h\rangle_{-1,1} + \\ &- \langle \int_0^t e^{(t-s)L_\psi}\partial_\theta\left[\nu_s^n(J*\nu_s^n)\right]\,\mathrm{d}s, h\rangle_{-1,1} - \langle g_t^n, h\rangle_{-1,1} + \langle z_t^n, h\rangle_{-1,1}.\end{aligned} \tag{3.24}$$

Since (3.24) holds for all $h \in H_1$, the identity (3.5) follows. All elements in (3.5) take values in $\mathcal{C}^0([0,T], H_{-1})$ and the proof is then concluded modulo regularity and wellposedness of $g_\cdot^n$ and $z_\cdot^n$. We refer to Lemma 3.2 and Lemma 3.3 which are presented in the next subsection. □

3.2. **Control on the perturbations.** Two kinds of perturbations are present in the SPDE (3.5): $z_\cdot^n$ given by the stochastic nature of the system and $g_\cdot^n$ given by the presence of a network structure. In this subsection, we exhibit the control over the two perturbations. Observe that all the estimates are independent of ψ.

We start with the control on the graph structure, which uses Grothendieck's Inequality presented in Theorem 1.2.

Lemma 3.2 (Wellposedness and bounds on g_t^n)**.** *For $n \in \mathbb{N}$ and $t \geqslant 0$, let g_t^n be given by*

$$g_t^n = \frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) e^{(t-s)L_\psi}\partial_\theta\left[\delta_{\theta_s^{i,n}}(J*\delta_{\theta_s^{j,n}})\right]\,\mathrm{d}s. \tag{3.25}$$

Then

(1) *$g^n \in \mathcal{C}^0([0,\infty), H_{-1})$. In particular, for every $h \in H_1$ and $t \geqslant 0$*

$$g_t^n(h) = -\frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) J(\theta_s^{i,n} - \theta_s^{j,n})\left[\partial_\theta e^{(t-s)L_\psi^*}h\right](\theta_s^{i,n})\,\mathrm{d}s. \tag{3.26}$$

(2) *There exists $D > 0$ such that*

$$\|g_t^n\|_{-1} \leqslant D\sqrt{t}\,\frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}, \qquad \text{for all } t \geqslant 0. \tag{3.27}$$

Proof. Fix n large. Consider $\{\phi_l\}_{l\geqslant 1} \subset \mathcal{C}^\infty$ such that $\phi_l \geqslant 0$, $\phi_l(\theta) = 0$ for $\theta \in [1/l, 2\pi - 1/l]$, $\int \phi_l = 1$ for every $l \geqslant 1$ and $\lim_{l\to\infty}\int F\phi_l = F(0)$ for every $F \in \mathcal{C}^0$. For $i = 1, \dots, n$, define

$$\phi_{s,l}^i := \phi_l * \delta_{\theta_s^{i,n}}. \tag{3.28}$$

We start by establishing (3.26). For each $h \in \mathcal{C}^2$

$$\begin{aligned}
&\langle \frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)\phi^i_{s,l}(J*\delta_{\theta^{j,n}_s}), \partial_\theta e^{(t-s)L^*_\psi}h\rangle_{-1,1} = \\
&= \left(\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)\phi^i_{s,l}(J*\delta_{\theta^{j,n}_s}), \partial_\theta e^{(t-s)L^*_\psi}h\right) = \\
&= -\left(\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\phi^i_{s,l}(J*\delta_{\theta^{j,n}_s})\right], h\right) = \\
&= -\langle \frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\phi^i_{s,l}(J*\delta_{\theta^{j,n}_s})\right], h\rangle_{-1,1}
\end{aligned} \tag{3.29}$$

But $\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)\phi^i_{s,l}(J*\delta_{\theta^{j,n}_s})$ converges to $\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)\delta_{\theta^{i,n}_s}(J*\delta_{\theta^{j,n}_s})$ since

$$\left\|\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)\left(\phi^i_{s,l}-\delta_{\theta^{i,n}_s}\right)(J*\delta_{\theta^{j,n}_s})\right\|_{-1} \leqslant \frac{1}{p_n}\sup_{i=1,\dots,n}\left\|\phi^i_{s,l}-\delta_{\theta^{i,n}_s}\right\|_{-1}, \tag{3.30}$$

which tends to zero as l tends to infinity.

Thanks to the properties of the semigroup, the same holds true for

$$\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\left(\phi^i_{s,l}-\delta_{\theta^{i,n}_s}\right)(J*\delta_{\theta^{j,n}_s})\right]; \tag{3.31}$$

indeed, by Proposition B.2,

$$\begin{aligned}
&\left\|\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\left(\phi^i_{s,l}-\delta_{\theta^{i,n}_s}\right)(J*\delta_{\theta^{j,n}_s})\right]\right\|_{-1} \leqslant \\
&\qquad\qquad \leqslant \frac{C}{p_n\sqrt{t-s}}\sup_{i=1,\dots,n}\left\|\phi^i_{s,l}-\delta_{\theta^{i,n}_s}\right\|_{-1}.
\end{aligned} \tag{3.32}$$

A similar argument shows that

$$\left\|\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\delta_{\theta^{i,n}_s}(J*\delta_{\theta^{j,n}_s})\right]\right\|_{-1} \leqslant \frac{C}{p_n\sqrt{t-s}}, \tag{3.33}$$

which, in turn, implies that

$$\frac{1}{n^2}\sum_{i,j=1}^{n}\int_0^t\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\delta_{\theta^{i,n}_s}(J*\delta_{\theta^{j,n}_s})\right]\mathrm{d}s \tag{3.34}$$

is almost surely finite and continuous with respect to t. We deduce (3.26).

For the second part (3.27), observe that

$$\begin{aligned}\langle \frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)e^{(t-s)L_\psi}\partial_\theta\left[\delta_{\theta_s^{i,n}}(J*\delta_{\theta_s^{j,n}})\right],h\rangle_{-1,1} = \\ = -\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\frac{\xi_{ij}}{p_n}-1\right)\langle \delta_{\theta_s^{i,n}},(J*\delta_{\theta_s^{j,n}})\partial_\theta e^{(t-s)L_\psi^*}h\rangle_{-1,1}.\end{aligned} \tag{3.35}$$

We claim that this last term can be controlled by $\left\|P^{(n)}-\mathbf{1}^{(n)}\right\|_{\infty\to 1}$ through Grothendieck's inequality. By choosing $H=H_{-1}$ and

$$\begin{aligned} a_{ij} &= \left(\tfrac{\xi_{ij}}{p_n}-1\right), \\ S_i &= \delta_{\theta_s^{i,n}}, \\ T_j &= \frac{\sqrt{t-s}}{C}\left(J*\delta_{\theta_s^{j,n}}\right)\partial_\theta e^{(t-s)L_\psi^*}\frac{h}{\|h\|_1}, \end{aligned} \tag{3.36}$$

Theorem 1.2 allows us to bound the expression in (3.35) by

$$K_R\frac{C}{\sqrt{t-s}}\|h\|_1\ \frac{\left\|P^{(n)}-\mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}. \tag{3.37}$$

This shows that

$$\|g_t^n\|_{-1}\leqslant K_RC\frac{\left\|P^{(n)}-\mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}\int_0^t\frac{1}{\sqrt{t-s}}\,\mathrm{d}s\ = D\sqrt{t}\frac{\left\|P^{(n)}-\mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}\ , \tag{3.38}$$

where $D:=K_RC/2>0$. The proof is concluded. □

We now turn to the stochastic term $z_\cdot^n$ in (3.5). Recall that L_ψ^* is diagonal in the basis $\{f_l^\psi\}_{l\geqslant 0}$ of H_{1,q_ψ}, with eigenvalues denoted by $\{\lambda_l\}_{\lambda\geqslant 0}$, see Proposition B.3. We precisely analyze $z_\cdot^n$ through its coefficients in the orthonormal basis given by L_ψ^*.

Lemma 3.3 (Wellposedness and bounds on z_t^n)**.** *For $n\in\mathbb{N}$ and $t>0$, let z_t^n be defined by*

$$z_t^n=\sum_{l\geqslant 1}\langle z_t^n,e_l^\psi\rangle_{H_{-1,1/q_\psi}}\ e_l^\psi. \tag{3.39}$$

Then

(1) *$z_\cdot^n\in\mathcal{C}^0([0,\infty),H_{-1})$ almost surely. In particular, for every $h\in H_1$*

$$z_t^n(h)=\frac{1}{n}\sum_{j=1}^{n}\int_0^t\left[\partial_\theta e^{(t-s)L_\psi^*}h\right](\theta_s^{j,n})\,\mathrm{d}B_s^j. \tag{3.40}$$

(2) *For every $T>0$, there exists a constant $Z=Z(T)>0$, such that for n large enough it holds that*

$$\forall\eta>0,\quad \mathbf{P}\left(\sup_{t\in[0,T]}\|z_t^n\|_{-1}>\eta\right)\leqslant\exp\left\{-Zn\eta^2\right\}. \tag{3.41}$$

Proof. We start by observing that the definition of z_t^n in the basis of $H_{-1,1/q_\psi}$, coincides with the one give in the mild formulation. For $h = \sum_{l \geqslant 0} \langle h, f_l^\psi \rangle_{H_{1,q_\psi}} f_l^\psi$, one obtains

$$\begin{aligned} \langle z_t^n, h \rangle_{-1,1} &= \sum_{l \geqslant 0} \langle z_t^n, e_l^\psi \rangle_{H_{-1,1/q_\psi}} \langle h, f_l^\psi \rangle_{H_{1,q_\psi}} = \\ &= \sum_{l \geqslant 0} z_t^n(f_l^\psi) \langle h, f_l^\psi \rangle_{H_{1,q_\psi}} = z_t^n(h), \end{aligned} \tag{3.42}$$

where we have used the properties of e_l^ψ and f_l^ψ, see Proposition B.3.

Before proving (1), we prove (2) and this will imply the existence of a continuous version of $z_\cdot^n$ almost surely.

Concerning (2), we start by observing that

$$\|z_t^n\|_{-1,1/q_\psi} = \sum_{l \geqslant 0} \left| z_t^n(f_l^\psi) \right|^2. \tag{3.43}$$

Let $l \geqslant 1$ and consider $z_t^n(f_l^\psi)$, by the definition of f_l^ψ and the properties of the semigroup, one gets

$$z_t^n(f_l^\psi) = \frac{1}{n} \sum_{j=1}^n \int_0^t \left[\partial_\theta e^{(t-s)L_\psi^*} f_l^\psi \right] (\theta_s^{j,n}) \, \mathrm{d}B_s^j = \frac{1}{n} \sum_{j=1}^n \int_0^t e^{-(t-s)\lambda_l} \left[\partial_\theta f_l^\psi \right] (\theta_s^{j,n}) \, \mathrm{d}B_s^j. \tag{3.44}$$

Set $c = \sup_{l \geqslant 0} \left| \partial_\theta f_l^\psi \right| < \infty$, see Proposition B.3. We rewrite the last expression as

$$z_t^n(f_l^\psi) = \frac{c e^{-t\lambda_l}}{\sqrt{2\lambda_l n}} A_t, \tag{3.45}$$

where A_t is a continuous martingale given by

$$A_t = \sqrt{\frac{2\lambda_l}{c^2 n}} \sum_{j=1}^n \int_0^t e^{s\lambda_l} \left[\partial_\theta f_l^\psi \right] (\theta_s^{j,n}) \, \mathrm{d}B_s^j \tag{3.46}$$

and quadratic variation bounded by

$$\langle A \rangle_t = \frac{2\lambda_l}{c^2 n} \sum_{j=1}^n \int_0^t e^{2s\lambda_l} \left[\partial_\theta f_l^\psi \right]^2 (\theta_s^{j,n}) \, \mathrm{d}s \leqslant 2\lambda_l \int_0^t e^{2\lambda_l s} \, \mathrm{d}s \leqslant e^{2t\lambda_l} - 1. \tag{3.47}$$

From (3.45) and (3.47), one deduces that $z_t^n(f_l^\psi)$ is a self normalized process. For estimating $\mathbf{P}\left(\sup_{t \in [0,T]} \left| z_t^n(f_l^\psi) \right|^2 > \eta \right)$, we can this use the following result

Theorem 3.4 ([12, Theorem 4.1 and the following remark])**.** *Let $T > 0$, $\alpha \in (0, \frac{1}{2})$ and $(A_t)_{t \in [0,T]}$ be a martingale with $A_0 = 0$. There exists $C > 0$, depending only on α, such that*

$$\mathbf{E}\left[\sup_{t \in [0,T]} \exp\left\{ \frac{\alpha A_t^2}{\langle A \rangle_t \log\log\left(\langle A \rangle_t \vee e^2 \right)} \right\} \right] \leqslant C. \tag{3.48}$$

By standard computations, we obtain

$$\begin{aligned}
&\mathbf{P}\left(\sup_{t\in[0,T]}\left|z_t^n(f_l^\psi)\right|^2>\eta\right)=\mathbf{P}\left(\sup_{t\in[0,T]}e^{-2t\lambda_l}A_t^2>\frac{2\lambda_l n}{c^2}\eta\right)=\\
&=\mathbf{P}\left(\sup_{t\in[0,T]}\frac{A_t^2}{4e^{2t\lambda_l}\log(2+2T\lambda_l)}>\frac{\lambda_l n}{2c^2\log(2+2T\lambda_l)}\eta\right)\leqslant\\
&\leqslant\mathbf{E}\left[\sup_{t\in[0,T]}\exp\left\{\frac{A_t^2}{4e^{2t\lambda_l}\log(2+2T\lambda_l)}\right\}\right]\exp\left\{-\frac{\lambda_l n}{2c^2\log(2+2T\lambda_l)}\eta\right\}.
\end{aligned}\tag{3.49}$$

We now use the fact that

$$\alpha:=\sup_{t\in[0,T]}\frac{\langle A\rangle_t\log\log(\langle A\rangle_t\vee e^2)}{4e^{2t\lambda_l}\log(2+2T\lambda_l)}\leqslant\frac{1}{4},\quad\text{for all } l\geqslant 1,\tag{3.50}$$

and, by Theorem 3.4, we obtain that there exists $C>0$, independent of T,n and l, such that

$$\mathbf{P}\left(\sup_{t\in[0,T]}\left|z_t^n(f_l^\psi)\right|^2>\eta\right)\leqslant C\exp\left\{-\frac{\lambda_l n}{c^2\log(2+2T\lambda_l)}\eta\right\}.\tag{3.51}$$

The case $l=0$ is somehow easier since

$$z_t^n(f_0^\psi)=\frac{1}{n}\sum_{j=1}^n\int_0^t\left[\partial_\theta f_0^\psi\right](\theta_s^{j,n})\,\mathrm{d}B_s^j\tag{3.52}$$

is a standard martingale with bounded quadratic variation: one can use Theorem 3.4 or, more simply, exponential estimates and Doob's inequality to obtain that there exists $c_0>0$ (depending on T) such that for all $\eta>0$

$$\mathbf{P}\left(\sup_{t\in[0,T]}\left|z_t^n(f_0^\psi)\right|^2>\eta\right)\leqslant c_0\exp\{-c_0 n\eta\}.\tag{3.53}$$

The last part of the proof consists in exploiting the exponential inequalities (3.51) and (3.53), and to transfer them to $\sup_{t\in[0,T]}\|z_t^n\|_{-1}^2$. For this purpose, let $S>0$ be defined by

$$S:=\left(\sum_{l\geqslant 0}\frac{1}{(1+l)^{4/3}}\right)^{-1}.\tag{3.54}$$

For some $\eta>0$, it holds that

$$\begin{aligned}
&\mathbf{P}\left(\sup_{t\in[0,T]}\|z_t^n\|_{-1,1/q_\psi}>\eta\right)\leqslant\mathbf{P}\left(\sum_{l\geqslant 0}\sup_{t\in[0,T]}\left|z_t^n(f_l^\psi)\right|^2>\eta^2\right)\leqslant\\
&\leqslant\sum_{l\geqslant 0}\mathbf{P}\left(\sup_{t\in[0,T]}\left|z_t^n(f_l^\psi)\right|^2>\frac{S}{(1+l)^{4/3}}\eta^2\right)\leqslant\\
&\leqslant c_0\exp\{-c_0 S\,n\eta^2\}+C\sum_{l\geqslant 1}\exp\left\{-\frac{S}{c^2}\,\frac{\lambda_l}{\log(2+2T\lambda_l)(1+l)^{4/3}}\,n\eta^2\right\}.
\end{aligned}\tag{3.55}$$

Now we use the fact that $\lambda_l = \Theta(l^2)$ as l tends to infinity, see Proposition B.3. In particular, there exists $L > 0$, depending on T, such that

$$\mathbf{P}\left(\sup_{t\in[0,T]} \|z_t^n\|_{-1,1/q_\psi} > \eta\right) \leqslant c_0 \exp\{-c_0 S\, n\eta^2\} + \\ + CL \exp\left\{-\frac{S}{c^2 \log(2+2T\lambda_L)} n\eta^2\right\} + C\sum_{l>L} \exp\left\{-\frac{S}{c^2}\sqrt{l}\, n\eta^2\right\}. \tag{3.56}$$

Observing that $\int_1^\infty e^{-\sqrt{x}n}\,\mathrm{d}x = \frac{2(n+1)}{n^2}e^{-n}$, taking n large enough and Z an suitable constant depending on c_0, S, L and C, the proof of (2) is concluded.

Back to (1), observe that for $s, t \in [0,T]$ and for some $k \geqslant 1$

$$\|z_t^n - z_s^n\|_{-1}^2 \leqslant \sum_{l=0}^{k} \left|z_t^n(f_l^\psi) - z_s^n(f_l^\psi)\right|^2 + 2\sum_{l>k} \sup_{t\in[0,T]} \left|z_t^n(f_l^\psi)\right|^2. \tag{3.57}$$

The first term can be make small by using the continuity of $z_t^n(e_l)$; for the second one, observe that we have just proven that $\mathbf{E}\left[\sum_{l\,\geqslant\,1} \sup_{t\in[0,T]} |z_t^n(e_l)|^2\right] < \infty$. This implies that there exists a subsequence $\{k_m\}_{m\in\mathbb{N}}$ such that $\sum_{l>k_m} \sup_{t\in[0,T]} |z_t^n(e_l)|^2$ tends to 0 almost surely as m tends to infinity. The almost sure continuity in (3.57) is then established by choosing s and t close enough and k large enough.

□

3.3. **Proof of Theorem 2.3.** Recall that μ_0^n converges in H_{-1} to $q_\psi \in M$. Next lemma assures that the projection of μ_0^n on M is well defined for n big enough.

Lemma 3.5. [26, Lemma 2.8] *There exists $\sigma > 0$ such that for all $\mu \in H_{-1}$ such that $\mathrm{dist}(\mu, M) \leqslant \sigma$, there exists a unique phase $\psi := \mathrm{proj}_M(\mu) \in \mathbb{T}$ such that $P_\psi^0(\mu - q_\psi) = 0$ and the mapping $\mu \mapsto \mathrm{proj}_M(\mu)$ is $\mathcal{C}^\infty$.*

Let $\psi_n = \mathrm{proj}_M(\mu_0^n)$. Fix $\varepsilon > 0$, we place ourselves in

$$A_1^n = \{\|\mu_0^n - q_{\psi_n}\|_{-1} \leqslant \varepsilon/2\}. \tag{3.58}$$

Denote by d_t^n the distance between μ_t^n and M, i.e.

$$d_t^n := \mathrm{dist}(\mu_t^n, M). \tag{3.59}$$

We want to prove that μ_t^n stays close to M for long times. Let $T > 0$, $N \in \mathbb{N}$ and define

$$T_i = iT, \qquad \text{for } i = 0, \dots, N. \tag{3.60}$$

If, for $N = N_n = o(\exp(n))$, we show that

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T_{N_n}]} d_t^n \leqslant \varepsilon\right) = 1, \tag{3.61}$$

then we are done. For sake of notation, we just employ N.

For $0 \leqslant a < b < \infty$, define the events

$$E^n(a,b) = \left\{\max\left\{2d_a^n,\, 2d_b^n,\, \sup_{t\in(a,b)} d_t^n\right\} \leqslant \varepsilon\right\}, \tag{3.62}$$

clearly

$$\mathbf{P}\left(\sup_{t\in[0,T_N]} d_t^n \leqslant \varepsilon\right) \geqslant \mathbf{P}\left(E^n(0,T_N)\right). \tag{3.63}$$

The Markov property of system (1.2) implies that

$$\begin{aligned}\mathbf{P}\left(E^n(0,T_N)\right) &\geqslant \mathbf{P}\left(E^n(0,T_N)\middle|E^n(0,T_{N-1})\right)\mathbf{P}\left(E^n(0,T_{N-1})\right) = \\ &= \mathbf{P}\left(E^n(0,T)\right)\mathbf{P}\left(E^n(0,T_{N-1})\right) \geqslant \mathbf{P}\left(E^n(0,T)\right)^N.\end{aligned} \tag{3.64}$$

Let's then focus on the bounded interval of time $[0,T]$ and consider $\nu_\cdot^n = \mu_\cdot^n - q_{\psi_n}$, which satisfies the stochastic partial differential equation (3.5). Taking the norm in H_{-1} on both sides of (3.5) and using the properties of the semigroup together with the fact that (e.g. [4, Lemma 7.3])

$$\|\partial_\theta(\mu(J*\nu))\|_{-2} \leqslant C\|\mu\|_{-1}\|\nu\|_{-1}, \qquad \text{for all } \mu,\nu\in H_{-1}, \tag{3.65}$$

one is left with (with a new constant C)

$$\|\nu_t^n\|_{-1} \leqslant \left\|e^{tL_{\psi^n}}\nu_0^n\right\|_{-1} + \int_0^t \frac{C}{\sqrt{t-s}}\|\nu_s^n\|_{-1}^2\,\mathrm{d}s + \|g_t^n\|_{-1} + \|z_t^n\|_{-1}. \tag{3.66}$$

By taking ε small enough, one can apply a Gronwall-type inequality (similar to Lemma B.4) that leads to (recall (3.58) and the fact that the semigroup is continuous)

$$\sup_{t\in[0,T]}\|\nu_t^n\|_{-1} \leqslant \frac{2}{3}\varepsilon + \sup_{t\in[0,T]}\|g_t^n\|_{-1} + \sup_{t\in[0,T]}\|z_t^n\|_{-1}. \tag{3.67}$$

For $\eta>0$, define $A_2^n(\eta) = \{\sup_{t\in[0,T]}\|z_t^n\|_{-1} \leqslant \eta\}$. If n is large enough, Lemma 3.2 assures that $\sup_{t\in[0,T]}\|g_t^n\|_{-1}$ is arbitrarily small a.s., and, placing ourselves in $A_2^n(\varepsilon/10)$, one obtains

$$\sup_{t\in[0,T]}\|\nu_t^n\|_{-1} \leqslant \frac{2}{3}\varepsilon + \frac{1}{5}\varepsilon \leqslant \varepsilon. \tag{3.68}$$

Plugging this estimate in (3.66) for $t=T$, observing that $P^0_{\psi^n}\nu_0^n = 0$ by construction so that $\left\|e^{tL_{\psi^n}}\nu_0^n\right\|_{-1} \leqslant Ce^{-\lambda_1 t/2}\|\nu_0^n\|_{-1}$ (e.g. [26, Proposition B.6]), one obtains

$$\|\nu_T^n\|_{-1} \leqslant e^{-\lambda_1 T/2}\frac{\varepsilon}{2} + \varepsilon\left(\varepsilon\int_0^T \frac{C}{\sqrt{T-s}}\,\mathrm{d}s\right) + \frac{\varepsilon}{5}, \tag{3.69}$$

choosing T and ε such that

$$\begin{aligned} T &\geqslant \frac{2}{\lambda_1}\log(5), \\ \varepsilon &\leqslant \frac{1}{20C\sqrt{T}}, \end{aligned} \tag{3.70}$$

one finally gets

$$\|\nu_T^n\|_{-1} \leqslant \frac{\varepsilon}{2}. \tag{3.71}$$

Since $d_t^n \leqslant \|\nu_t^n\|_{-1}$ for all $t\geqslant 0$, we have then proven that $A_2^n(\varepsilon/10)\subset E^n(0,T)$. In particular,

$$\mathbf{P}\left(\sup_{t\in[0,T_N]} d_t^n \leqslant \varepsilon\right) \geqslant \mathbf{P}\left(E^n(0,T)\right)^N \geqslant \left(1-\mathbf{P}\left(A_2^n(\epsilon/10)^{\complement}\right)\right)^N. \tag{3.72}$$

We can than use the estimate (3.41) in Lemma 3.3 to get

$$\mathbf{P}\left(A_2^n(\epsilon/10)^{\complement}\right) = \mathbf{P}\left(\sup_{t\in[0,T]} \|z_t^n\|_{-1} > \frac{\varepsilon}{10}\right) \leqslant \exp\left\{-\frac{Z}{100}n\varepsilon^2\right\}. \tag{3.73}$$

Putting all together, one is left with

$$\begin{aligned}\mathbf{P}\left(\sup_{t\in[0,T_N]} d_t^n \leqslant \varepsilon\right) &\geqslant \left(1-\mathbf{P}\left(A_2^n(\epsilon/10)^{\complement}\right)\right)^N \geqslant \left(1-\exp\left\{-\frac{Z}{100}n\varepsilon^2\right\}\right)^N \\ &= \exp\left\{N\log\left[1-\exp\left(-\frac{Z}{100}n\varepsilon^2\right)\right]\right\} \geqslant \exp\left\{-\frac{3}{2}N\exp\left(-\frac{Z}{100}n\varepsilon^2\right)\right\},\end{aligned} \tag{3.74}$$

where we have used that $\log(1-x) \geqslant -3/2x$ for $0 \leqslant x \leqslant 1/2$. But the right hand side of (3.74) tends to 1 for all $N = N_n = o(\exp(n))$ and the proof is concluded.

4. Longtime behavior around $1/2\pi$

In this section we will suppose that the finite time behavior is already known, so that for n large enough, μ_t^n is very close to μ_t; thus, for a large T_0, μ_{T_0} will be very close to $1/2\pi$ and so will be for $\mu_{T_0}^n$. At the end of the day, we may suppose that we are starting close to $1/2\pi$. Since we are not assuming any independence between initial conditions and graph, instead of proving Theorem 2.4, we rather prove the following Proposition.

Proposition 4.1. *If for every $\varepsilon_0 > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left(\left\|\mu_0^n - \tfrac{1}{2\pi}\right\|_{-1} \leqslant \varepsilon_0\right) = 1. \tag{4.1}$$

Then, there exists $A > 0$ such that for every positive increasing sequence $\{T_n\}_{n\in\mathbb{N}}$ such that $T_n = \exp(o(n))$ and for all $0 < \varepsilon < A$, it holds

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T_n]} \left\|\mu_t^n - \tfrac{1}{2\pi}\right\|_{-1} \leqslant \varepsilon\right) = 1. \tag{4.2}$$

The end of the section is thus devoted to prove Proposition 4.1.

4.1. **A mild formulation around $1/2\pi$.** We place ourselves aroud the stationary solution $\frac{1}{2\pi}$. The system evolution is captured by the linear dynamics around $\frac{1}{2\pi}$ and the corresponding linear operator $L_{2\pi}$ is given by

$$L_{2\pi}u \;:=\; \tfrac{1}{2}\partial_\theta^2 u - \tfrac{1}{2\pi}(\partial_\theta J) * u, \qquad \text{for } u \in \mathcal{C}^2(\mathbb{T}),\ \int_{\mathbb{T}} u(\theta)\,\mathrm{d}\theta = 0. \tag{4.3}$$

The adjoint $L_{2\pi}^*$ of $L_{2\pi}$ in $\mathcal{L}_0^2$ has the following expression

$$L_{2\pi}^* u \;=\; \tfrac{1}{2}\partial_\theta^2 u - \tfrac{1}{2\pi} J * (\partial_\theta u), \tag{4.4}$$

and domain $D(L_{2\pi}^*) = D(L_{2\pi})$. These operators are diagonal in the Fourier basis $\{e_l\}_{l\geqslant 1}$, with eigenvalues denoted by $\{\lambda_l^{2\pi}\}_{l\geqslant 1}$. The spectrum is negative and bounded away from 0, let $\gamma_K = \lambda_1^{2\pi} = \frac{1-K}{2} > 0$ denote the spectral gap. The operator $L_{2\pi}$ (resp. $L_{2\pi}^*$) defines an analytic semigroup $e^{tL_{2\pi}}$ (resp. $e^{tL_{2\pi}^*}$) with the following property:

$$\left\|e^{tL_{2\pi}}u\right\|_{-1} \leqslant C\frac{e^{-\gamma t/2}}{\sqrt{t}}\|u\|_{-2}, \quad \text{for some } C > 0, \tag{4.5}$$

for all $\gamma \in [0, \gamma_K)$, all $t > 0$ and $u \in H_{-1}$. We will not prove (4.5) but refer to Appendix B for similar estimates.

Define $\nu_t^n := \mu_t^n - \frac{1}{2\pi}$. As done in Section 3, we derive a mild formulation for $\nu_\cdot^n$. We omit the proof.

Proposition 4.2. *The process $\nu_t^n \in H_{-1}$ satisfies the following stochastic partial differential equation in $C([0,T], H_{-1})$:*

$$\nu_t^n = e^{tL_{2\pi}}\nu_0^n - \int_0^t e^{(t-s)L_{2\pi}}\partial_\theta \left[\nu_s^n (J * \nu_s^n)\right] \mathrm{d}s - g_t^n + z_t^n, \tag{4.6}$$

where

$$g_t^n = \frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) e^{(t-s)L_{2\pi}}\partial_\theta \left[\delta_{\theta_s^{i,n}}(J * \delta_{\theta_s^{j,n}})\right] \mathrm{d}s, \tag{4.7}$$

and $z_t^n \in H_{-1}$ is defined for $h \in H_1$ by

$$\langle z_t^n, h\rangle_{-1,1} = \frac{1}{n}\sum_{j=1}^n \int_0^t \partial_\theta e^{(t-s)L_{2\pi}^*} h(\theta_s^{j,n})\, \mathrm{d}B_s^j. \tag{4.8}$$

4.2. **Control on the perturbations.** Contrary to the supercritical case, the operator $L_{2\pi}$ is contracting along all direction or, in other words, all its eigenvalues are negative. This property gives a stronger control on $g_\cdot^n$ and $z_\cdot^n$, as shown in the next Lemmas.

Lemma 4.3 (Wellposedness and bounds on g_t^n)**.** *For $n \in \mathbb{N}$ and $t \geqslant 0$, let g_t^n be given by*

$$g_t^n = \frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) e^{(t-s)L_{2\pi}}\partial_\theta \left[\delta_{\theta_s^{i,n}}(J * \delta_{\theta_s^{j,n}})\right] \mathrm{d}s. \tag{4.9}$$

Then

(1) *$g^n \in \mathcal{C}^0([0,\infty), H_{-1})$. In particular, for every $h \in H_1$ and $t \geqslant 0$*

$$g_t^n(h) = -\frac{1}{n^2}\sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) J(\theta_s^{i,n} - \theta_s^{j,n})\partial_\theta e^{(t-s)L_{2\pi}^*} h(\theta_s^{i,n})\, \mathrm{d}s. \tag{4.10}$$

(2) *There exists $D > 0$, independent of t, such that*

$$\|g_t^n\|_{-1} \leqslant D\, \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}, \qquad \text{for all } t \geqslant 0. \tag{4.11}$$

Proof. We only prove (2). Observe that, as in (3.35),

$$\begin{aligned} &\langle \frac{1}{n^2}\sum_{i,j=1}^n \left(\frac{\xi_{ij}}{p_n} - 1\right) e^{(t-s)L_{2\pi}}\partial_\theta \left[\delta_{\theta_s^{i,n}}(J * \delta_{\theta_s^{j,n}})\right], h\rangle_{-1,1} = \\ &= -\frac{1}{n^2}\sum_{i,j=1}^n \left(\frac{\xi_{ij}}{p_n} - 1\right) \langle \delta_{\theta_s^{i,n}}, (J * \delta_{\theta_s^{j,n}})\partial_\theta e^{(t-s)L_{2\pi}^*} h\rangle_{-1,1}. \end{aligned} \tag{4.12}$$

Applying Theorem 1.2, this time with

$$\begin{aligned} a_{ij} &= \left(\tfrac{\xi_{ij}}{p_n} - 1\right), \\ S_i &= \delta_{\theta_s^{i,n}}, \\ T_j &= \frac{\sqrt{t-s}}{Ce^{-\gamma(t-s)}}\left(J * \delta_{\theta_s^{j,n}}\right)\partial_\theta e^{(t-s)L_{2\pi}^*}\, \frac{h}{\|h\|_1}, \end{aligned} \tag{4.13}$$

allows us to bound the expression in (4.12) by

$$K_R \frac{Ce^{-\gamma(t-s)}}{\sqrt{t-s}} \|h\|_1 \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}. \tag{4.14}$$

This shows that

$$\|g_t^n\|_{-1} \leqslant K_R C \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2} \int_0^t \frac{e^{-\gamma(t-s)}}{\sqrt{t-s}} \,\mathrm{d}s \leqslant D \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}, \tag{4.15}$$

where $D := K_R C \int_0^\infty \frac{e^{-\gamma s}}{\sqrt{s}} \,\mathrm{d}s > 0$ since the integral converges. The proof is concluded. □

We now turn to the stochastic term z_t^n in (4.6). Recall that $L_{2\pi}$ is diagonal in the Fourier basis $\{e_l\}_{l \geqslant 1}$ of H_{-1}, with eigenvalues denoted by $\lambda_l^{2\pi}$. Then

Lemma 4.4 (Wellposedness and bounds on z_t^n)**.** *For $n \in \mathbb{N}$ and $t > 0$, let z_t^n be defined by*

$$z_t^n = \sum_{l \geqslant 1} \langle z_t^n, e_l \rangle_{H_{-1}} \, e_l, \tag{4.16}$$

where

$$\langle z_t^n, e_l \rangle_{H_{-1}} = z_t^n \left(\frac{e^{il\cdot}}{l} \right) = \frac{i}{n} \sum_{j=1}^n \int_0^t e^{(t-s)\lambda_l^{2\pi}} e^{il\theta_s^{j,n}} \,\mathrm{d}B_s^j. \tag{4.17}$$

Then

(1) *$z^n \in \mathcal{C}^0([0,\infty), H_{-1})$ almost surely.*
(2) *There exists $C > 0$ independent of n, such that for all $T > 0$*

$$\mathbf{E}\left[\sup_{t\in[0,T]} \|z_t^n\|_{-1}^2 \right] \leqslant C \, \frac{\log(1 + 2\gamma_K T)}{n}. \tag{4.18}$$

(3) *For every positive increasing sequence $\{T_n\}_{n\in\mathbb{N}}$ such that $T_n = \exp(o(n))$ and for all $\eta > 0$, it holds*

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T_n]} \|z_t^n\|_{-1} \leqslant \eta \right) = 1. \tag{4.19}$$

Proof. We only prove (2). For $l \geqslant 1$, let $x_t^l := \sqrt{2\lambda_l^{2\pi} n} \, e^{\lambda_l^{2\pi} t} \left| z_t^n(e^{il\cdot}) \right|$. In particular

$$x_t^l = \left| \frac{\sqrt{2\lambda_l^{2\pi}}}{\sqrt{n}} \sum_{j=1}^n \int_0^t e^{s\lambda_l^{2\pi}} e^{il\theta_s^{j,n}} \,\mathrm{d}B_s^j \right| = \left| a_t^l + i\, b_t^l \right|, \tag{4.20}$$

where a^l and b^l are two continuous real valued martingales. Let $\langle x^l \rangle_t = \langle a^l \rangle_t + \langle b^l \rangle_t$ where $\langle a^l \rangle_t$ and $\langle b^l \rangle_t$ are the quadratic variations of a_t^l and b_t^l respectively, then

$$\langle x^l \rangle_t = \frac{2\lambda_l^{2\pi}}{n} \sum_{j=1}^n \int_0^t e^{2s\lambda_l^{2\pi}} (\cos^2 + \sin^2)(l\theta_s^{j,n}) \,\mathrm{d}s = e^{2\lambda_l^{2\pi} t} - 1. \tag{4.21}$$

We now use

Lemma 4.5. *Let $Y_t = A_t + i\,B_t$, where A_t and B_t are continuous real valued martingales. Define $X_t = |Y_t|$ and $\langle X\rangle_t = \langle A\rangle_t + \langle B\rangle_t$, where $\langle A\rangle_t$ and $\langle B\rangle_t$ are the quadratic variations of A and B respectively. Then, there exists $C > 0$ such that, for all $T > 0$,*

$$\mathbf{E}\left[\sup_{t\in[0,T]} \frac{X_t^2}{1+\langle X\rangle_t}\right] \leqslant C\log(1+\log(1+\langle X\rangle_t)). \tag{4.22}$$

The proof of Lemma 4.5 is presented at the end of the section. By choosing $X_t = x_t^l$, $A_t = a_t^l$ and $B_t = b_t^l$, one obtains that, for $T > 0$,

$$\mathbf{E}\left[\sup_{t\in[0,T]} |z_t^n(e_l)|^2\right] = \frac{1}{2\lambda_l^{2\pi} n}\,\mathbf{E}\left[\sup_{t\in[0,T]} \frac{(x_t^l)^2}{1+\langle x^l\rangle_t}\right] \leqslant \frac{C}{2\lambda_l^{2\pi} n}\log(1+2\lambda_l^{2\pi}T). \tag{4.23}$$

It remains to observe that

$$\mathbf{E}\left[\sup_{t\in[0,T]} \|z_t^n\|_{-1}^2\right] \leqslant \mathbf{E}\left[\sum_{l\geqslant 1}\sup_{t\in[0,T]} |z_t^n(e_l)|^2\right] \leqslant C\sum_{l\geqslant 1}\frac{1}{2\lambda_l^{2\pi} n}\log(1+2\lambda_l^{2\pi}T). \tag{4.24}$$

The conclusion holds by factorizing the first term of the sum and modifying the constant C accordingly: observe that $\sum_{l\geqslant 1}\sup_{T\geqslant 1}\frac{\log(1+2\lambda_l^{2\pi}T)}{\lambda_l^{2\pi}\log(1+2\lambda_1^{2\pi}T)} < \infty$.

The proof is concluded modulo Lemma 4.5, proven hereafter. □

Proof of Lemma 4.5. Recall that A_t is a martingale, in particular a slight variation of [18, Corollary 2.8] implies that there exists $D > 0$ such that

$$\mathbf{E}\left[\sup_{t\in[0,T]} \frac{A_t^2}{1+\langle A\rangle_t}\right] \leqslant D\log(1+\log(1+\langle A\rangle_t)). \tag{4.25}$$

Thus, one can develop

$$\begin{aligned}\mathbf{E}\left[\sup_{t\in[0,T]} \frac{X_t^2}{1+\langle X\rangle_t}\right] &\leqslant \mathbf{E}\left[\sup_{t\in[0,T]} \frac{A_t^2}{1+\langle A\rangle_t}\right] + \mathbf{E}\left[\sup_{t\in[0,T]} \frac{B_t^2}{1+\langle B\rangle_t}\right] \leqslant \\ &\leqslant D\log(1+\log(1+\langle A\rangle_t)) + D\log(1+\log(1+\langle B\rangle_t)) \leqslant \\ &\leqslant 2D\log(1+\log(1+\langle X\rangle_t)),\end{aligned} \tag{4.26}$$

and the proof is done by taking $C = 2D$. □

4.3. **Proof of Proposition 4.1.** Fix $\varepsilon > 0$. From Proposition 4.2 we know that $\nu_t^n := \mu_t^n - \frac{1}{2\pi}$ satisfies

$$\nu_t^n = e^{tL_{2\pi}}\nu_0^n - \int_0^t e^{(t-s)L_{2\pi}}\partial_\theta\left[\nu_s^n(J*\nu_t^n)\right]\,\mathrm{d}s - g_t^n + z_t^n. \tag{4.27}$$

Taking the norm and using the properties of $e^{tL_{2\pi}}$, together with the estimate (3.65), for all $0 < \gamma < \gamma_K$ one obtains

$$\|\nu_t^n\|_{-1} \leqslant \|\nu_0^n\|_{-1} + C\int_0^t \frac{e^{-\gamma(t-s)}}{\sqrt{t-s}}\,\|\nu_s^n\|_{-1}^2\,\mathrm{d}s + \|g_t^n\|_{-1} + \|z_t^n\|_{-1}. \tag{4.28}$$

Thanks to the contractive properties of $L_{2\pi}$, there exists $D > 0$ (Lemma 4.3) such that

$$\sup_{t\geqslant 0}\|g_t^n\|_{-1} < D\,\frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}. \tag{4.29}$$

Define now $B_1^n(\varepsilon_0) = \{\|\nu_0^n\| \leqslant \varepsilon_0\}$ and $B_2^n(\eta) = \{\sup_{t\in[0,T_n]} \|z_t^n\|_{-1} \leqslant \eta\}$. On $B_1^n(\varepsilon/3) \cap B_2^n(\varepsilon/4)$ and for n large enough, we can apply Lemma B.4 with

$$\begin{aligned} &\delta = \frac{\varepsilon}{3}, \qquad T = T_n, \\ &f(t) = \|\nu_t^n\|_{-1}, \\ &g(t) = \|g_t^n\|_{-1} + \|z_t^n\|_{-1}, \end{aligned} \tag{4.30}$$

and obtain

$$\sup_{t\in[0,T_n]} \|\nu_t^n\|_{-1} \leqslant \varepsilon. \tag{4.31}$$

The proof is concluded since by hypothesis $\mathbf{P}(B_1^n) \to 1$ and Lemma 4.4 implies that $\mathbf{P}(B_2^n) \to 1$ as n tends to infinity.

5. Finite time behavior

The aim of this section is to study the closeness of $\mu_\cdot^n$ to $\mu_\cdot$ on bounded time interval.

Proof of Theorem 2.1. Fix $\varepsilon > 0$ and $T > 0$. It is not difficult to see that $\mu_\cdot^n - \mu_\cdot$ satisfies again a mild equation in $\mathcal{C}^0([0,T], H_{-1})$, which is given by

$$\mu_t^n - \mu_t = e^{t\frac{\Delta}{2}}(\mu_0^n - \mu_0) - \int_0^t e^{(t-s)\frac{\Delta}{2}} \partial_\theta \left[\mu_s^n (J * \mu_s^n) - \mu_s (J * \mu_s)\right] \mathrm{d}s - g_t^n + z_t^n, \tag{5.1}$$

where

$$g_t^n = \frac{1}{n^2} \sum_{i,j=1}^n \int_0^t \left(\frac{\xi_{ij}}{p_n} - 1\right) e^{(t-s)\frac{\Delta}{2}} \partial_\theta \left[\delta_{\theta_s^{i,n}} (J * \delta_{\theta_s^{j,n}})\right] \mathrm{d}s, \tag{5.2}$$

and z_t^n is denoted for $h \in H_1$ by

$$z_t^n(h) = \frac{1}{n} \sum_{j=1}^n \int_0^t \partial_\theta e^{(t-s)\frac{\Delta}{2}} h(\theta_s^{j,n}) \,\mathrm{d}B_s^j. \tag{5.3}$$

Observe that we are using the Laplacian operator which is very similar to $L_{2\pi}$ except for the first eigenvalue that is now given by $-(1-K)/2$. We will thus use all the results about $L_{2\pi}$ and its semigroup to control $z_\cdot^n$ and $g_\cdot^n$.

Taking the H_{-1} norm in (5.1) and applying (3.65), one is left with

$$\|\mu_t^n - \mu_t\|_{-1} \leqslant \|\mu_0^n - \mu_0\|_{-1} + \int_0^t \tfrac{C}{\sqrt{t-s}} \|\mu_s^n - \mu_s\|_{-1} \,\mathrm{d}s + \|g_t^n\|_{-1} + \|z_t^n\|_{-1}. \tag{5.4}$$

The term involving the graph g_t^n can be controlled again by $\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}$: minor modifications to Lemma 4.3 show that there exists $D > 0$ such that

$$\sup_{t\in[0,T]} \|g_t^n\|_{-1} \leqslant D \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2}. \tag{5.5}$$

For the initial conditions and the stochastic part z_t^n, define the two sets:

$$\begin{aligned} C_1^n = C_1^n(\varepsilon_0) &= \left\{\|\mu_0^n - \mu_0\|_{-1} \leqslant \varepsilon_0\right\}; \\ C_2^n = C_2^n(T,\eta) &= \left\{\sup_{t\in[0,T]} \|z_t^n\|_{-1} \leqslant \eta\right\}. \end{aligned} \tag{5.6}$$

On $C_1^n \cap C_2^n$, one obtains

$$\|\mu_t^n - \mu_t\|_{-1} \leqslant \varepsilon_0 + \int_0^t \tfrac{C}{\sqrt{t-s}} \|\mu_s^n - \mu_s\|_{-1} \, \mathrm{d}s + D \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2} + \eta. \tag{5.7}$$

Gronwall-Henry's inequality ([21, Lemma 7.1.1 and Exercice 1]) leads to

$$\sup_{t\in[0,T]} \|\mu_t^n - \mu_t\|_{-1} \leqslant 2\left(\varepsilon_0 + D \frac{\left\|P^{(n)} - \mathbf{1}^{(n)}\right\|_{\infty\to 1}}{n^2} + \eta\right) e^{aT}, \tag{5.8}$$

where a is independent of n, ε_0 and η. Considering ε_0 and η small enough and n large enough, the proof is concluded modulo showing that

$$\lim_{n\to\infty} \mathbf{P}\left(C_1^n \cap C_2^n\right) = 1. \tag{5.9}$$

From the hypothesis on the intial condition (2.1), it is clear that for all ε_0 one has $\mathbf{P}\left(C_1^n(\varepsilon_0)\right) \to 1$ as n tends to infinity. The same conclusion holds for C_2^n by slightly modifying the proof of Lemma 4.4. The proof is concluded. □

Appendix A. Graphs

A.1. **General properties of the graphs under consideration.** We observe that condition (1.16) implies a weak form of degree homogeneity (recall (1.10)):

Lemma A.1. *Suppose that* (1.16) *holds. Let* $\delta > 0$*, define*

$$I_n^\delta := \left\{ i \in \{1,\dots,n\} : \lim_{n\to\infty} \left| \frac{1}{n} \sum_{j=1}^n \frac{\xi_{i,j}^{(n)}}{p_n} - 1 \right| \geqslant \delta \right\}. \tag{A.1}$$

Then $|I_n^\delta| = o(n)$.

Proof. Suppose that $\lim_{n\to\infty} \frac{|I_n|}{n} = c$ for some $c > 0$. Then

$$\begin{aligned} \sup_{s_i,t_j\in\{\pm1\}} \frac{1}{n^2} \sum_{i,j=1}^n \left(\frac{\xi_{ij}^{(n)}}{p_n} - 1\right) s_i t_j &\geqslant \sup_{s_i\in\{\pm1\}} \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{n} \sum_{j=1}^n \left(\frac{\xi_{ij}^{(n)}}{p_n} - 1\right)\right] s_i \geqslant \\ &\geqslant \frac{1}{n} \sum_{i\in I_n} \left| \frac{1}{n} \sum_{j=1}^n \left(\frac{\xi_{ij}^{(n)}}{p_n} - 1\right) \right| \geqslant \frac{|I_n|}{n} \inf_{i\in I_n} \left| \frac{1}{n} \sum_{j=1}^n \left(\frac{\xi_{ij}^{(n)}}{p_n} - 1\right) \right|. \end{aligned} \tag{A.2}$$

This last term does not go to zero as n tends to infinity, against (1.16). □

It also implies the existence of an unique giant component.

Lemma A.2. *Suppose that* (1.16) *holds. Then, there exists a unique sequence of connected components* $\{\mathcal{C}^{(n)}\}$ *in* $\{\xi^{(n)}\}$ *and* $\lim_{n\to\infty} \left|\mathcal{C}^{(n)}\right|/n = 1$.

Proof. We prove the uniqueness first. Suppose that for every n there exist $\mathcal{C}_1^{(n)}$ and $\mathcal{C}_2^{(n)}$ distinct connected components of $\xi^{(n)}$ such that $\left|\mathcal{C}_i^{(n)}\right| = n_i = \Theta(n)$ for $i = 1, 2$. Without loss of generality, one can suppose $\mathcal{C}_1^{(n)}$ consisting in the first n_1 vertices of $\xi^{(n)}$ and $\mathcal{C}_2^{(n)}$ in the following n_2.

Using the equivalence of $\ell_\infty \to \ell_1$ norm with the cut-norm (e.g. [2]), one obtains

$$\|P_n - \mathbf{1}_n\|_{\infty\to 1} \geqslant \sup_{x_i,y_j\in\{0,1\}} \left| \sum_{i,j=1}^n \left(\frac{\xi_{ij}}{p_n} - 1\right) x_i y_j \right| \geqslant \sum_{\substack{1\leqslant i\leqslant n_1 \\ n_1\leqslant j\leqslant n_2-n_1}} 1 \,=\, n_1 n_2 = \Theta(n^2). \tag{A.3}$$

For the existence, suppose the connected components of $\xi^{(n)}$ are ordered from the biggest one in size (the first n_1 vertices) to the smallest one (the last vertices). Take the first m components such that $|\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m| \geqslant n/4$. One easily sees that $|\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m| \leqslant n/2$. Applying the same reasoning of before with $1 \leqslant i \leqslant n/4$ and $n/2 \leqslant j \leqslant n$, the proof is concluded. □

A.2. **Examples of graph sequences.** We exhibit two classes of graphs, a random and a deterministic one, that satisfy assumption (1.16). The only hypothesis required on p_n is equivalent to asking that the mean degree per site diverges as n tends to infinity, i.e. $np_n \uparrow \infty$.

Erdős-Rényi random graphs. As mentioned in the introduction, $\|\cdot\|_{\infty\to 1}$ has been found very useful for random graph concentration and this is indeed the case of ER graphs (e.g. [19]). We recall the definition and give the result.

For every $n \in \mathbb{N}$, let $\{\xi_{ij}^{(n)}\}_{1\leqslant i\neq j\leqslant n}$ be IID Bernoulli random variables with parameter p_n, $\mathbb{P}$ denoting the associated probability. For every i, $\xi_{ii}^{(n)}$ is set equal to 0, i.e. self loop are not admitted.

Lemma A.3. *Assume that*

$$\lim_{n\to\infty} np_n = \infty. \tag{A.4}$$

There exists $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}\left(\sup_{s_i,t_j} \frac{1}{n^2} \sum_{i,j=1}^n \left(\tfrac{\xi_{ij}}{p_n} - 1\right) s_i t_j \geqslant \frac{2}{\sqrt{np_n}} \right) \leqslant e^{-2n}, \quad \textit{for all } n \geqslant n_0. \tag{A.5}$$

Proof. The proof is just an union bound and an application of Bernstein's inequality. Indeed,

$$\mathbb{P}\left(\sup_{s_i,t_j} \frac{1}{n^2} \sum_{i,j=1}^n \left(\frac{\xi_{ij}}{p_n} - 1\right) s_i t_j \geqslant \frac{\delta}{\sqrt{np_n}} \right) \leqslant \sum_{s_i,t_j} \mathbb{P}\left(\frac{1}{n^2} \sum_{i,j=1}^n \left(\frac{\xi_{ij}}{p_n} - 1\right) s_i t_j \geqslant \frac{\delta}{\sqrt{np_n}} \right). \tag{A.6}$$

Bernstein's inequality ([6, Corollary 2.11]) says that if $X_1, \ldots, X_n$ are independent zero-mean random variables such that $|X_j| \leq M$ a.s. for all j, then for all $t \geq 0$

$$\mathbb{P}\left(\sum_{j=1}^n X_j > t \right) \,\leq\, \exp\left\{ -\frac{t^2}{2\sum_{j=1}^n \mathbb{E}[X_j^2] + \frac{2}{3}Mt} \right\}.$$

Let $X_{k(i,j)} = \frac{s_i t_j}{n^2 p_n}(\xi_{ij} - p_n)$ with k some bijection from $\{1,\dots,n\}^2$ to $\{1,\dots,n^2\}$. Then $|X_k| \leqslant \frac{1}{n^2 p_n}$ and $\mathbb{E}\left[X_k^2\right] \leqslant \frac{2}{n^4}$. For n large enough, we thus obtain

$$\mathbb{P}\left(\sum_{k=1}^{n^2} X_k \geqslant \frac{\delta}{\sqrt{np_n}}\right) \leq \exp\left\{-\frac{n\delta^2}{4p_n + \frac{2}{3}\frac{\delta}{\sqrt{np_n}}}\right\} \leqslant \exp\left\{-n\delta^2\right\}. \tag{A.7}$$

The proof is concluded observing that the sum in (A.6) consists in 4^n elements and choosing $\delta = 2$. □

We thus have

Proposition A.4. *Given* (A.4)*, ER graphs satisfy condition* (1.16) $\mathbb{P}$*-almost surely.*

Proof. It suffices to apply Borel-Cantelli lemma to (A.5). □

Similarly one can prove that symmetric ER random graphs satisfy (1.16) a.s..

Ramanujan graphs. Let $d = 2, 3, \dots$, consider a d-regular graph, i.e. graph where each vertex has exactly d neighbors. We start recalling a well-known result

Lemma A.5 (Expander mixing lemma)**.** *Let G be a d-regular random graph (G denoting the adjacency matrix itself), it holds*

$$\frac{1}{n^2}\left\|\tfrac{n}{d}G - \mathbf{1}^{(n)}\right\|_{\infty\to 1} \leqslant 4\,\frac{\lambda(d)}{d}, \tag{A.8}$$

where $\lambda(d)$ is the second biggest eigenvalue (in absolute value) associated to G.

Proof. The proof is classical but it is in general formulated in terms of the cut-norm (e.g. [22]). One easily sees that the cut-norm is equivalent (paying a factor 4, e.g. [2]) to the $\ell_\infty \to \ell_1$ norm. □

Ramanujan graphs are d-regular graphs such that $\lambda(d) \leqslant 2\sqrt{d-1}$, they are very well known for their expander properties (e.g. [22]). Condition (1.16) holds whenever d_n diverges; indeed

Proposition A.6. *Let $d_n = np_n$. Suppose that* (A.4) *holds, i.e.*

$$\lim_{n\to\infty} d_n = \infty. \tag{A.9}$$

Then, every sequence of Ramanujan graphs satisfies condition (1.16)*.*

Proof. Rewriting (A.8) in terms of p_n, it becomes

$$\frac{1}{n^2}\left\|\tfrac{G}{p_n} - \mathbf{1}^{(n)}\right\|_{\infty\to 1} \leqslant \frac{8}{\sqrt{np_n}}. \tag{A.10}$$

The proof is concluded taking the limit for n which tends to infinity. □

A.3. **Links with graphons.** The norm $\|\cdot\|_{\infty\to 1}$ is strictly related to the canonical distance $d_{\mathcal{W}}$ on the space of (sparse) graphons $\mathcal{W}$ (e.g. [8]). In fact, whenever $\xi^{(n)}/p_n$ is (a realization of) a graphon $W^{(n)}$, condition (1.16) is implied by the convergence of $W^{(n)}$ to the constant graphon $W \equiv 1$ in $\mathcal{W}$. One can then consider system (1.2) on a sequence of (sparse) graphons and require, instead of condition (1.16), the convergence in $\mathcal{W}$ to the constant graphon.

We have decided not to add another level of complexity in order to keep the results as clear as possible, but everything could be reformulated within this more general framework and the proofs would basically not change.

Appendix B. H_{-1} and Semigroups

B.1. **On the relationship between H_{-1} and $\mathcal{P}(\mathbb{T})$.** Consider $H_1 := H_{1,1}$, its dual space, denoted by H_{-1}, can be described through the Fourier orthonormal basis $\{e_l\}_{l\geqslant 1}$, where $e_l(\theta) = le^{il\theta}$. With this characterization one easily obtains that $\mathcal{P}(\mathbb{T}) - \frac{1}{2\pi} \subset H_{-1}$. Indeed, for $\mu \in \mathcal{P}(\mathbb{T})$,

$$\left\|\mu - \frac{1}{2\pi}\right\|_{-1} = \sqrt{\sum_{l\geqslant 1} \left|\langle \mu, le^{il\cdot}\rangle_{H_{-1}}\right|^2} = \sqrt{\sum_{l\geqslant 1} \frac{1}{l^2} \left|\langle \mu, e^{il\cdot}\rangle\right|^2} \leqslant \sqrt{\sum_{l\geqslant 1} \frac{1}{l^2}} < \infty. \tag{B.1}$$

In particular, the difference between two probability measures belongs to H_{-1}.

Observe now that H_{-1} induces a distance on $\mathcal{P}(\mathbb{T})$ which controls the bounded-Lipschitz distance d_{bL}, i.e. for all $\mu, \nu \in \mathcal{P}(\mathbb{T})$

$$\begin{aligned} d_{\mathrm{bL}}(\mu,\nu) &= \sup_{\|f\|_{\mathrm{bL}}=1} \int f\,(\,\mathrm{d}\mu - \,\mathrm{d}\nu) \leqslant \sup_{h\in\mathcal{C}_0^1, \|h\|_1=1} \int h\,(\,\mathrm{d}\mu - \,\mathrm{d}\nu) = \\ &= \sup_{h\in\mathcal{C}_0^1, \|h\|_1=1} \int h'\,(\mathcal{U} - \mathcal{V}) = \sup_{\|h\|_1=1} \langle \mu - \nu, h\rangle_{-1,1} = \\ &= \|\mu - \nu\|_{-1}\,. \end{aligned} \tag{B.2}$$

Where we have used the density of $\mathcal{C}_0^1$ in H_1, and denoted by $\mathcal{U}$ and $\mathcal{V}$ the primitives of μ and ν respectively.

B.2. **On the weighted Hilbert space $H_{-1,\omega}$.** Recall that, one has this sequence of continuous and dense inclusions:

$$H_{1,1/\omega} \subset \mathcal{L}_0^2 = \mathcal{L}_0^{2^*} \subset H_1^* =: H_{-1,\omega}, \tag{B.3}$$

where we have chosen the canonical identification for $\mathcal{L}_0^2$. We can explicit the isometry between $H_{1,1/\omega}$ to $H_{-1,\omega}$. Consider the operator

$$\begin{aligned} A_\omega : \mathcal{C}^\infty(\mathbb{T}) &\to \mathcal{C}^\infty(\mathbb{T}) \\ f &\mapsto -\partial_\theta\left(\omega^{-1}\,\partial_\theta f\right) \end{aligned} \tag{B.4}$$

It is known [7, pag. 82] that $A_\omega(H_{1,1/\omega})$ is dense in $H_{-1,\omega}$ and the injection is continuous. This allows considering $H_{1,1/\omega}$ as a subset of $H_{-1,\omega}$ by identifying u and $A_\omega u$.

The inner product in $H_{-1,\omega}$, dual to the one in $H_{1,1/w}$, is given by

$$\langle u, v\rangle_{H_{-1,w}} = \int w\,\mathcal{U}\mathcal{V}, \tag{B.5}$$

where $\mathcal{U}$ and $\mathcal{V}$ are primitive of u and v respectively, such that $\int w\,\mathcal{U} = 0 = \int w\,\mathcal{V}$ (e.g. [3, Subsection 2.2]). Then, for $f, g \in \mathcal{C}^\infty$, it holds

$$\langle A_\omega f, A_\omega g\rangle_{-1,\omega} = \int \omega^{-1} f' g' = \langle f, g\rangle_{1,1/\omega}. \tag{B.6}$$

B.3. **The linear operators L_ψ and L_ψ^* and their semigroups.** This subsection recalls the known results on L_ψ, its dual L_ψ^* and the associated semigroups e^{tL_ψ} and $e^{tL_\psi^*}$.

We start with the spectral properties of L_ψ.

Proposition B.1. *The operator L_ψ (resp. L_ψ^*) is essentially self-adjoint with compact resolvent in $H_{-1,1/q}$ (resp. $H_{1,q}$). Its spectrum is pure point and lies in $(-\infty, -\lambda_1] \cup \{0\}$, where $\lambda_1 > 0$ and 0 is a simple eigenvalue of L_ψ with eigenvector $\partial_\theta q_\psi$.*

Moreover, both $L_{2\pi}$ and $L_{2\pi}^*$ generate a $\mathcal{C}^0$ semigroup $t \mapsto e^{tL_\psi}$ (resp. $t \mapsto e^{tL_\psi^*}$) in $\mathcal{L}_0^2$ and $e^{tL_\psi^*} = \left(e^{tL_\psi}\right)^*$.

Proof. The result about L_ψ is given in [3]. Observe that, due to the isometry (B.4) between $H_{-1,1/q_\psi}$ and H_{1,q_ψ}, $L_\psi^* = A_{1/q_\psi}^{-1} L_\psi A_{1/q_\psi}$ and it has thus the same spectral properties of L_ψ.

From the spectral properties of L_ψ and L_ψ^*, one deduces that the two operators are sectorial (and with dense domain in H_{-1}), standard techniques assure the existence of the analytic semigroup (e.g. [21]). □

An accurate analysis of the semigroup has already been established in [4] by means of interpolating norms and Fourier decomposition. We recall here the most important properties. We will use the space H_{-2}, defined in an analogous way of H_{-1}.

Proposition B.2 ([4, Lemma 7.2])**.** *For all $t > 0$, the operator e^{tL_ψ} extends to a bounded operator from H_{-2} to H_{-1} and there exists $C > 0$ such that for all $u \in H_{-2}$*

$$\left\|e^{tL_\psi} u\right\|_{-1} \leqslant C\left(1 + \frac{1}{\sqrt{t}}\right)\|u\|_{-2}. \tag{B.7}$$

Moreover, for all $\epsilon \in (0, 1/2)$, $\delta \geqslant 0$ and all $u \in H_{-1}$

$$\left\|e^{(t+\delta)L_\psi} u - e^{tL_\psi} u\right\|_{-1} \leqslant C\delta^\epsilon\left(1 + \frac{1}{t^{1/2+\epsilon}}\right)\|u\|_{-2}. \tag{B.8}$$

By duality, observe that for all $h \in H_1$

$$\left\|e^{tL_\psi^*} h\right\|_2 \leqslant C\left(1 + \frac{1}{\sqrt{t}}\right)\|h\|_1. \tag{B.9}$$

We end this subsection with an useful result on the eigenvalues and eigenfunctions associated to L_ψ, recall (3.3).

Proposition B.3. *There exists $C > 1$ such that for all $l \in \mathbb{N}$*

$$\frac{l^2}{C} \leqslant \lambda_l \leqslant C l^2. \tag{B.10}$$

Let $f_l^\psi = A^{-1}_{1/q_\psi} e_l^\psi$*, then* f_l^ψ *is an eigenfunction of* L^*_ψ *associated to* $-\lambda_l$ *and*

$$\sup_{l\in\mathbb{N}} \left\| \partial_\theta f_l^\psi \right\|_\infty < \infty. \tag{B.11}$$

Proof. The first part is covered in [4, Remark 8.3] and the second one in [4, Corollary 8.6]. □

B.4. **Analytical estimate.** A variation on Gronwall Lemma.

Lemma B.4. *Let* $T > 0$*,* $\gamma \geqslant 0$*. Let* $f : [0,T] \to [0,\infty)$ *be a continuous function and* $g : [0,T] \to [0,\infty)$ *be such that for all* $0 \leqslant t \leqslant T$

$$f(t) \leqslant f(0) + \int_0^t \frac{e^{-\gamma(t-s)}}{\sqrt{t-s}} f^2(s)\,\mathrm{d}s + g(t). \tag{B.12}$$

There exists $A > 0$*, which depends on* T *only if* $\gamma = 0$*, such that for all* $0 < \delta < A$ *and if* $f(0) < \delta$*,* $\sup_{t\in[0,T]} g(t) < \delta$*, then*

$$\sup_{t\in[0,T]} f(t) \leqslant 3\delta. \tag{B.13}$$

Proof. Consider the set $O = \{t \,:\, f(t) \leqslant 3\delta\} \subset [0,T]$. Since f is continuous and $f(0) \leqslant \delta$, O is a non-empty open set in $[0,T]$. Suppose that $\sup(O) = u < T$; we show that $u \in O$, which implies $O = [0,T]$.

Consider

$$\begin{aligned} f(u) =& f(0) + \int_0^u \frac{e^{-\gamma(u-s)}}{\sqrt{u-s}} f^2(s)\,\mathrm{d}s + g(u) \leqslant \\ \leqslant& 2\delta + \delta \left(9\delta \int_0^u \frac{e^{-\gamma(u-s)}}{\sqrt{u-s}}\,\mathrm{d}s \right) \leqslant 3\delta, \end{aligned} \tag{B.14}$$

where the last inequality holds for all $\delta \leqslant A := \left(9\int_0^\infty \frac{e^{-\gamma s}}{\sqrt{s}}\,\mathrm{d}s\right)^{-1}$ whenever $\gamma > 0$ or for all $\delta \leqslant \frac{1}{18\sqrt{T}}$ in case $\gamma = 0$. Thus $u \in O$ and the proof is concluded. □

Acknowledgments

The author is thankful to Giambattista Giacomin for proposing this subject and his insightful advises and to Helge Dietert for the help on the analytic part. He would also thank Florian Bechtold, Simon Coste, Christophe Poquet and Assaf Shapira for discussions and comments on the previous drafts of the work.

The author acknowledges the support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665850.

References

[1] J.A. Acebron, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort and R. Spigler, *The Kuramoto Model: A Simple Paradigm for Synchronization Phenomena*, Reviews of Modern Physics **77**, no. 1 (2005): 137–85.

[2] N. Alon and A Naor, *Approximating the Cut-Norm via Grothendieck's Inequality*, SIAM Journal on Computing **35** (2006): 787–803.

[3] L. Bertini, G. Giacomin and K. Pakdaman, *Dynamical aspects of mean field plane rotators and the Kuramoto model*, Journal of Statistical Physics **138**, n. 1-3 (2010): 270-90.

[4] L. Bertini, G. Giacomin and C. Poquet, *Synchronization and Random Long Time Dynamics for Mean-Field Plane Rotators*, Probability Theory and Related Fields **160** (2014), n. 3.4: 593-653.

[5] S. Bhamidi, A. Budhiraja and R. Wu, *Weakly Interacting Particle Systems on Inhomogeneous Random Graphs*, Stochastic Processes and their Applications **129**, n. 6 (2019): 2174–2206.

[6] S. Boucheron, G. Lugosi and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*, OUP Oxford (2013).

[7] H. Brezis, *Analyse fonctionnelle: Théorie et applications*, Dunod, Paris (1999).

[8] C. Borgs, J. Chayes, H. Cohn and Y. Zhao, *An $\mathcal{L}^p$ Theory of Sparse Graph Convergence I: Limits, Sparse Random Graph Models, and Power Law Distributions*, Transactions of the American Mathematical Society **372**, no. 5 (2019): 3019-62.

[9] H. Chiba and G. S. Medvedev, *The Mean Field Analysis for the Kuramoto Model on Graphs I. The Mean Field Equation and Transition Point Formulas*, Discrete & Continuous Dynamical Systems - A 39, n. **1** (2019): 131.

[10] F. Coppini, H. Dietert and Giambattista Giacomin, *A Law of Large Numbers and Large Deviations for interacting diffusions on Erdős-Rényi graphs*, Stochastics and Dynamics, appeared online on July 2019, arXiv:1807.10921 (2019).

[11] D. A. Dawson and J. Gärtner, *Large deviations from the mckean-vlasov limit for weakly interacting diffusions*, Stochastics **20** (1987): 247–308.

[12] V. H. de la Pena, M. J. Klass and T. L. Lai, *Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws*, The Annals of Probability **32**, n. 3A (2004): 1902–33.

[13] S. Delattre, G. Giacomin and E. Luçon, *A note on dynamical models on random graphs and Fokker-Planck equations*, Journal of Statistical Physics **165** (2016), no. 4: 785-798.

[14] A, Dembo and A. Montanari, *Ising Models on Locally Tree-like Graphs*, The Annals of Applied Probability 20, n. **2** (2010): 565–92.

[15] F. Dörfler and F. Bullo, *Synchronization in complex networks of phase oscillators: A survey*, Automatica 50, n. **6** (2014): 1539–64.

[16] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren der mathematischen Wissenschaften Vol. **260**, New York, Springer US (1984).

[17] G. Giacomin, K. Pakdaman and X. Pellegrin, *Global attractor and asymptotic dynamics in the Kuramoto model for coupled noisy phase oscillators*, Nonlinearity **25** (2012), n. 5: 1247-73.

[18] S. E. Graversen and G. Peskir. *Maximal Inequalities for the Ornstein-Uhlenbeck Process*, Proceedings of the American Mathematical Society **128**, no. 10 (2000): 3035–42.

[19] O. Guédon and R. Vershynin, *Community Detection in Sparse Networks via Grothendieck's Inequality*, Probability Theory and Related Fields **165** (2016), n. 3: 1025–49.

[20] E. Heiman, G. Schechtman and A. Shraibman, *Deterministic Algorithms for Matrix Completion*, Random Structures & Algorithms **45** (2014), n. 2: 306–17.

[21] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics **840**, Springer Berlin Heidelberg (1981).

[22] S. Hoory, N. Linial and A. Wigderson, *Expander Graphs and Their Applications*, Bulletin of the American Mathematical Society **43**, no. 04 (2006): 439–562.

[23] D. Lacker, K. Ramanan and R. Wu, *Large sparse networks of interacting diffusions*, arXiv:1904.02585 (2019).

[24] C. M. Le, E. Levina and R. Vershynin, *Concentration and Regularization of Random Graphs*, Random Structures & Algorithms **51** (2017), n. 3: 538–61.

[25] E. Luçon, *Quenched asymptotics for interacting diffusions on inhomogeneous random graphs*, arXiv:1811.09229 (2018).

[26] E. Luçon and C. Poquet, *Long time dynamics and disorder-induced traveling waves in the stochastic Kuramoto model*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **53** (2017), n. 3: 1196–1240.

[27] G. S. Medvedev, *The continuum limit of the Kuramoto model on sparse random graphs*, Communications in Mathematical Sciences, vol. 17, no. **4** (2019): 883-898.

[28] R. I. Oliveira, *Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges*, arXiv:0911.0600 [math] (2009).

[29] R. I. Oliveira, G. H. Reis and L. M. Stolerman *Interacting diffusions on sparse graphs: hydrodynamics from local weak limits*, arXiv:1812.11924 (2018). http://arxiv.org/abs/1812.11924.

[30] E. Olivieri and M. E. Vares, *Large Deviations and Metastability*, Cambridge: Cambridge University Press (2005).

[31] G. Pisier, *Grothendieck's Theorem, Past and Present*, Bulletin of the American Mathematical Society **49** (2012), n. 2: 237–323.

[32] G. H. Reis and R. I. Oliveira, *Interacting diffusions on random graphs with diverging degrees: hydrodynamics and large deviations*, Journal of Statistical Physics, (2019). https://doi.org/10.1007/s10955-019-02332-1.

[33] F. A. Rodrigues, T. K. DM. Peron, P. Ji and J. Kurths. *The Kuramoto Model in Complex Networks*, Physics Reports, **610** (2016): 1-98.

[34] A.-S. Sznitman, *Topics in Propagation of Chaos*, Ecole d'Eté de Probabilités de Saint-Flour XIX, Springer Berlin Heidelberg (1991), 1464:165-251.

[35] K. Yosida, *Functional analysis*, Grundlehren der Mathematischen Wissenschaften vol. 123, Springer-Verlag, Berlin-New York, VI ed. (1980).

Université Paris Diderot, Sorbonne Paris Cité, Laboratoire de Probabilités Statistique et Modélisation, UMR 8001, F- 75205 Paris,France