Interacting diffusions on sparse graphs: hydrodynamics from local weak limits

Roberto I. Oliveira*   Guilherme H. Reis†   Lucas M. Stolerman‡

Abstract

We prove limit theorems for systems of interacting diffusions on sparse graphs. For example, we deduce a hydrodynamic limit and the propagation of chaos property for the stochastic Kuramoto model with interactions determined by Erdös-Rényi graphs with constant mean degree. The limiting object is related to a potentially infinite system of SDEs defined over a Galton-Watson tree. Our theorems apply more generally, when the sequence of graphs ("decorated" with edge and vertex parameters) converges in the local weak sense. Our main technical result is a locality estimate bounding the influence of far-away diffusions on one another. We also numerically explore the emergence of synchronization phenomena on Galton-Watson random trees, observing rich phase transitions from synchronized to desynchronized activity among nodes at different distances from the root.

1 Introduction

The results of this paper were inspired by a concrete problem. Let $n \in \mathbb{N}$ and $0 < p(n) \leq 1$. Define the Erdős-Rényi random graph $G_n = G(n, p(n))$ as the random graph with vertex set $[n] := \{1, \ldots, n\}$ where two vertices are adjacent with probability $p(n)$, independently of all other pairs. Write $i \sim_{(n)} j$ if $i, j \in [n]$ are adjacent and let $d_i^{(n)}$ denote the degree of i in G_n. We consider the stochastic Kuramoto model [16] over each realization of the graph G_n, which is defined as a system of interacting diffusions indexed

by $i \in [n]$, solutions of the following system of Itô Stochastic Differential Equations (SDEs) in time interval $[0, T]$:

$$d\theta_i^{(n)}(t) = \frac{1}{d_i^{(n)}} \sum_{j \in [n] \,:\, j \sim_{(n)} i} \sin(\theta_j^{(n)}(t) - \theta_i^{(n)}(t))dt + \omega_i \, dt + dB_i(t). \qquad (1.1)$$

Here the B_i are independent Brownian motions, and the initial positions $\theta_i^{(n)}(0)$ and "natural frequencies" ω_i are sampled from some product measure independently from the B_i and G_n. We adopt the convention that the first term in the RHS of (1.1) is zero in the case that $d_i^{(n)} = 0$.

The following question arises.

Problem: *What is the bulk behavior of this system when $n \to +\infty$ for different choices of $p(n)$?*

More precisely, we want to understand the behavior of the empirical measure of particle trajectories over a time interval $[0, T]$:

$$L_n := \frac{1}{n} \sum_{v \in [n]} \delta_{\theta_v^{(n)}(\cdot)},$$

that is a random measure over the space of continuous functions from $[0, T]$ to $\mathbb{R}$, the space $C([0, T]; \mathbb{R})$. Our problem is potentially interesting because the graph G_n can be very different depending on $p(n)$. For instance, when $n \to +\infty$, G_n is typically connected if $p(n) \gg \log n/n$ and typically disconnected if $p(n) \ll \log n/n$.

As it turns out, all that matters for our problem is the behavior of $np(n)$ as $n \to +\infty$, which is the expected degree of a vertex in G_n (up to a small error). In a recent paper [26], we proved that $np(n) \to +\infty$ implies that L_n has the same a.s. limit and obeys the same large deviations principle as in the case $p(n) \equiv 1$ of a complete interaction graph. In particular, the limit of L_n is the law of a McKean-Vlasov diffusion, a Markovian process with trajectories in $C([0, T]; \mathbb{R})$.

In this paper we complement the result for $np(n) \to +\infty$ by describing what happens when $np(n) \to c \in \mathbb{R}_+$. We prove that L_n converges to the law of a non-Markovian process, which is described by a system of the form (1.1) on a potentially infinite Galton-Watson (GW) tree. The mechanism behind this fact is a general theorem relating the *local weak convergence of networks* to the hydrodynamics of systems of diffusions on these networks.

Remark 1.1. To the best of our knowledge, the preprint version of this paper was the first work to explicitly relate systems of interacting diffusions to local weak convergence. However, an earlier paper by Maclaurin [22] obtained stronger large deviation results in the specific setting of diffusions over discrete tori, which converge to a system over $\mathbb{Z}^d$. In that sense, Maclaurin's result anticipates some aspects of our work. A few months after the present paper appeared in the Arxiv, Lacker et al. [17, 18] independently arrived at more systematic (and essentially more general) results of the same kind. More details about this related work are discussed in Section 3.2.

Finally, we numerically investigate synchronization phase transitions for the stochastic Kuramoto Model on GW trees. In particular, we compute synchronization levels among nodes at different distances from the root, by varying the coupling strength between oscillators, as well as their natural frequencies and initial conditions. In contrast with the full interaction case of the complete graph, we generally observe the emergence of desynchronization phenomena at distant nodes in the sparse setting.

In section 2, we give an informal description of our main results. In section 3, we make comments about the proofs, review past results, and give the outline of the remainder of the paper.

2 Informal definitions and main results

2.1 Infinite networks and interacting diffusions

We will need the concept of a *network*. Informally, this is an object of the form

$$N = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta(0)}),$$

where:

1. $G = (V, E)$ is a locally finite graph with countable vertex set V and edge set E.
2. $\boldsymbol{\mu} = (\mu_e)_{e \in E}$ is a vector of positive weights $\mu_e > 0$ for the edges of G.
3. $\boldsymbol{\omega} = (\omega_v)_{v \in V}$ is a vector of "media variables" $\omega_v \in \mathbb{R}$ associated with the vertices.
4. $\boldsymbol{\theta(0)} = (\theta_v(0))_{v \in V}$ is a vector of initial conditions $\theta_v(0) \in \mathbb{R}$ for each vertex.

We wil call $\boldsymbol{\mu}$ the edge marks and $\boldsymbol{\omega}, \boldsymbol{\theta(0)}$ the vertex marks. We will say that a network N is *finite* if the graph G is finite. We will often abuse notation and write "$v \in N$" instead of "$v \in V$". We also write $\mu_{vu} = \mu_{uv} := \mu_e$ for the weights of pairs $e = \{v, u\} \in E$, and set $\mu_{vu} = 0$ if vu is not an edge.

Suppose N is given. Let $\psi : \mathbb{R}^2 \to \mathbb{R}$, $\phi : \mathbb{R}^4 \to \mathbb{R}$ be Lipschitz functions with only ϕ bounded, and define for each $v \in V$, the total weight

$$\mu_v := \sum_{u \in V} \mu_{uv}.$$

Assume that we have a collection $(B_v(\cdot))_{v \in V}$ of i.i.d. standard Brownian motions associated with the vertices of N. A *system of interacting diffusions on the network N* (with this choice of functions ψ, ϕ) is a random vector

$$\boldsymbol{\theta}^N(\cdot) = (\theta_v^N(\cdot))_{v \in V} \in C([0, T]; \mathbb{R})^{|V|}$$

which is a strong solution of the following system of Itô Stochastic Differential Equations (SDEs): for each $v \in N$,

$$\mathrm{d}\theta_v^N(t) = \frac{1}{\mu_v} \sum_{u \in V} \mu_{uv} \phi(\theta_u^N(t), \theta_v^N(t); \omega_v, \omega_u)\mathrm{d}t + \psi(\theta_v^N(t); \omega_v)\mathrm{d}t + \mathrm{d}B_v(t), \qquad (2.1)$$

in the time interval $[0, T]$ and with initial conditions $(\theta_v(0))_{v \in N}$. Heretofore, we adopt the convention that the first term in the RHS of (2.1) is 0 whenever $\mu_v = 0$.

When N is finite, our conditions on ψ and ϕ are more than sufficient to imply existence and uniqueness for this problem. Our first finding is that the same holds for infinite networks with at-most-exponential growth.

Theorem 2.1 (Loose statement of Theorem 6.7)**.** *Suppose N is infinite, but there exists a vertex $o \in N$ such that balls around o grow at most exponentially. Also assume that the weights $\mu_{vu} \neq 0$ are uniformly bounded away from 0 and $+\infty$. Then there exists a system of interacting diffusions over N that is the unique strong solution of (2.1).*

2.2 Local weak limits and hydrodynamics

To state our next result, we need the concepts of *local metric* and *local weak convergence* of networks. Both of these concepts are defined precisely in Section 5; for now, we only give an informal description.

The local metric is defined on *rooted networks*, that is, on pairs (N, o) where N is a network and $o \in N$. According to this metric, two networks (N, o) and (N', o') are close if there are large balls around o and o' where both the graphs and the corresponding marks

can be matched nearly exactly. Note that, for this to make sense, we need to consider these networks up to "rooted isomorphisms"; see Section 5 for details.

Now, given a *sequence* of finite networks

$$\{N_n = (G_n, \boldsymbol{\mu}_n, \boldsymbol{\omega}_n, \boldsymbol{\theta}(\mathbf{0})_n)\}_{n \in \mathbb{N}},$$

let o_n be a node of N_n chosen uniformly at random, for each $n \in \mathbb{N}$. Let $\mathcal{L}(N_n, o_n)$ denote the law of the random rooted network (N_n, o_n); all the randomness comes from the choice of o_n. We say that $\{N_n\}_{n \in \mathbb{N}}$ converges in the local weak sense to a distribution ν over rooted networks if the probability laws $\mathcal{L}(N_n, o_n)$ converge weakly to ν.

Remark 2.2. If we forget about the marks $\boldsymbol{\mu}_n$, $\boldsymbol{\omega}_n$ and $\boldsymbol{\theta}(\mathbf{0})_n$, this is nothing but the better known concept of local weak convergence of graphs. In this case, it is known e.g. that n-cycles converge to the deterministic rooted graph $\delta_{(\mathbb{Z},0)}$; that the Erdős-Rényi graph $G(n, c/n)$ a.s. converges to a Poisson GW tree with parameter c; and that random d-regular graphs on n vertices a.s. converge to the infinite d-regular tree. We show in Section 5.2.1 that if the marks are chosen independently, with each vector $\boldsymbol{\mu}_n, \boldsymbol{\omega}_n, \boldsymbol{\theta}(\mathbf{0})_n$ i.i.d., then the corresponding networks converge a.s. in the local weak sense.

Now note that, for a arbitrary network N, if we can define a system of interacting diffusions over N this gives rise to a *random network*

$$N^\theta = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}^N(\cdot)),$$

where the initial conditions are replaced by the particle trajectories in the time interval $[0, T]$. We are abusing notation by calling by the name network two different classes of objects.

Coming back to the sequence of finite networks $(N_n)_{n \in \mathbb{N}}$, the next theorem relates the local weak convergence of N_n^θ to that of N_n.

Theorem 2.3 (Loose statement of Theorem 6.8 and Corollary 6.10)**. *Assume that ν is a probability measure on rooted networks which is supported over pairs (N, o) satisfying the assumptions of Theorem 2.1. Then for almost all samples $(N, o) \sim \nu$ we can solve the system of interacting diffusions (cf. (2.1)) as in Theorem 2.1 to consider (N^θ, o).*

Now consider a sequence of networks $\{N_n\}_{n \in \mathbb{N}}$, each N_n with n vertices, which converges in the local weak sense to ν. Assume also that the largest vertex degree in N_n is $n^{o(1)}$ for large n. Then, almost surely, the sequence $\{N_n^\theta\}_{n \in \mathbb{N}}$ of networks marked with the diffusions is locally weakly convergent to the law of (N^θ, o) when $(N, o) \sim \nu$. As a consequence, the empirical measures

$$L_n := \frac{1}{n} \sum_{i=1}^n \delta_{\theta_i^{N_n}(\cdot)}$$

converge almost surely to the distribution of $\theta_o^N(\cdot)$ when (N, o) is sampled from ν.

Our results also imply a propagation-of-chaos property (see Corollary 6.12).

2.3 Synchronization phenomena and sparsity

We come back to the particular case of stochastic Kuramoto model. Our results from theorems 2.1 and 2.3 motivate us to explore synchronization phenomena on finite GW trees, since they appear as the limit object from a sequence of random Erdős-Rényi graphs. If we denote $\mathcal{T}$ as the random GW tree with m vertices rooted at vertex 1, we consider the system of SDEs: for each $i \in [m]$,

$$d\theta_i^{\mathcal{T}}(t) = K \sum_{j=1}^m a_{ij} \sin(\theta_j^{\mathcal{T}}(t) - \theta_i^{\mathcal{T}}(t)) dt + \omega_i dt + \varepsilon dB_i(t),$$

where $\theta_j^{\mathcal{T}}(t)$ and ω_j represent the angular phase and natural frequency of the oscillator indexed by $j \in \{1, 2, \ldots, m\}$, respectively. The parameter $K \in \mathbb{R}_+$ represents the coupling strength between nodes, and $a_{ij} = 1$ if nodes i and j are connected in $\mathcal{T}$ or $a_{ij} = 0$ otherwise. In our numerical analysis, we do not divide the summation over neighbours of i by the degree of i. We sample both initial conditions $\theta_j^{\mathcal{T}}(0)$ and natural frequencies ω_i from distinct distributions, and by changing the coupling strength between nodes, we compute synchronization levels between the root and those nodes at different distances. In our simulations, we chose two different models for generating the GW trees:

1. **Binomial model:** The offspring is a binomial random variable with distribution $Bin(n, p)$.

2. **D-regular model:** The root node has $\mathcal{C}$ children, while the other ones have exactly $\mathcal{C} - 1$ children.

In Section 10, we describe our numerical methods and results in details. Interestingly, we observe how desynchronization emerges among distant nodes, depending on the choice of the model parameters. These findings enlighten our understanding of synchronization in complex networks and pave the way for new phase transition studies on Kuramoto dynamics.

3 Comments, references and organization

3.1 Comment on proofs

We now briefly comment on our proofs. The key step is to show that our system (cf. (2.1)) satisfies a locality property, Lemma 7.1 below. Loosely speaking, this property states that information does not propagate too fast over the graph in systems like (2.1). To prove this Lemma, we rely on a linear Gronwall argument, which leads to a matrix exponential. A nice wrinkle in the proof is that this exponential can be related to a heat kernel for a random walk over a network, which we can analyze via the Carne-Varoupoulos bound. With this Lemma in hand, our main results follow easily from general principles, including the definition of weak convergence.

One last comment is that it seems clear that our result is an exemplar of a more general principle. One can gather from our arguments that "local" systems of particles on graphs should have a "local hydrodynamic limit" whenever the sequence of underlying graphs converges. In this sense, our main technical contributions consist of formulating this principle precisely and proving the required locality estimate in our setting.

3.2 Discussion

As stated above, our motivation was to understand what happens to interacting difusions in the simple case of an Erdös-Rényi graph with a constant average degree. Our recent preprint [26] showed that the entire regime of a diverging average degree has the same behavior, even at the level of large deviations, as the complete graph (mean-field interactions). Of course, proving an LDP in the setting of the present paper is an interesting topic for further study.

We continue with a very brief review of the literature, referring to [26, Section 1.2] for more details. The study of our class of systems over complete graphs is a classical topic; see e.g. [28, 25] for early results. Recent papers have obtained hydrodynamic limits in settings with singular interactions [20, 21] or Gaussian couplings and delays [9, 7, 8]. More recently, several authors [11, 23, 14] have explicitly considered the case of relatively sparse random graphs. A recent preprint by Coppini et al. [13] obtains

an LDP under a stronger degree condition than in our paper [26], but with otherwise weaker or incomparable assumptions.

To the best of our knowledge, the present work is the first paper to explore how interacting diffusions behave in the bulk of relatively general graphs of constant average degree. However, we have recently become aware of earlier work [22] by Maclaurin. In that paper, the author proved a large deviation principle for interacting diffusions indexed by subsets of the lattice $\mathbb{Z}^d$. In the present paper, we only obtain a law of large numbers for the diffusion. However, Maclaurin's methods rely on the specific structure of $\mathbb{Z}^d$, whereas our results apply more broadly.

A few months after the first version of our paper appeared in the Arxiv, Lacker, Ramanan and Wu [17, 18] posted two preprints where they consider related systems of interacting diffusions over sparse graphs. Outside of our considering weighted edges, their results greatly generalize ours, by allowing the drift and diffusion coefficients of a vertex to depend nonlinearly on the empirical measures of neighboring vertices. Additionally, they make weaker requirements on the sequences of graphs. One technical difference is that our proofs give more quantitative estimates on correlation decay, whereas they rely on "softer" weak convergence tools. Their preprint [18] obtain closed-form descriptions of the non-Markovian dynamics of a vertex and its neighbors.

From the synchronization viewpoint, our study introduces novel results for the Kuramoto Model on GW trees. Over the past years, many studies have analyzed synchronization phenomena on various network topologies ([2, 19]), yet little attention has been given to sparse random trees. More recently, Chiba et al. [12] studied transitions to synchronization for a large family of random graphs, relating their onset of synchronization and the well-known phase transition for the fully connected network. With a more computational approach, Sokolov and Ermentrout [27] related network structure with global stability of phase-locked solutions. For power-law random networks, Medvedev and Tang [24] studied the effects of scale-free connectivity and compared the synchronization thresholds with dense graphs. In contrast with all those recent findings, our analysis on GW-trees allows to investigate the emergence of *desynchronizaton* among nodes that are distant from the root, which illustrates how full synchronization is not always achievable by increasing the coupling strength beyond a fixed value.

3.3 Organization of the paper

The remainder of the work is organized as follows: Section 4 reviews notation of functions and measures. In Section 4 we also present some preliminaries about graphs. Section 5 reviews networks and local weak convergence. The reader familiar with networks and local weak convergence just need to read this section to know what notation we adopted here. Section 6 states in full details our main results.

We prove the Locality lemma in Section 7, and the other main results are derived from this lemma in subsequent sections. We solve the infinite system of SDEs in Section 8, and we address the hydrodynamic limit in Section 9.

Finally, in Section 10 we present numerical simulations to discuss the synchronization phenomena. Auxiliary results are found in the Appendix, starting at section A.

4 Preliminaries

In this section we fix notation and briefly review some important concepts.

4.1 Numbers

$\mathbb{N}$ is the set of nonnegative integers. For a natural number $n \in \mathbb{N}\backslash\{0\}$, we let $[n] := \{1, \ldots, n\}$. We define the maximum and minimum of two numbers $x, y \in \mathbb{R}$ by $x \vee y$ and $x \wedge y$, respectively. We define $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$.

4.2 Functions and spaces of probability measures

Let (S, d) be a Polish metric space. We define

$$C(S; \mathbb{R}) := \{h : S \to \mathbb{R} \; : \; h \text{ is continuous}\},$$

and for a map $h : S \to \mathbb{R}$, we have the norms:

$$
\begin{aligned}
\|h\|_\infty &:= \sup_{x \in S} |h(x)|; \\
\|h\|_{Lip} &:= \sup_{x,y \in S \,:\, x \neq y} \frac{|h(x) - h(y)|}{\mathrm{d}(x,y)}; \\
\|h\|_{BL} &:= \|h\|_\infty + \|h\|_{Lip}.
\end{aligned}
$$

We let $\mathcal{P}(S)$ denote the set of probability measures over (the Borel sets of) S. If $X \in S$ is a random element we denote $\delta_X \in \mathcal{P}(S)$ the Dirac measure at X which is a random measure.

Given a measure $\mu \in \mathcal{P}(S)$ and a Borel function $h : S \to \mathbb{R}$ we write

$$\mu(h) = \int_S h \, \mathrm{d}\mu.$$

If $X \in S$ is a random element, and $\mathbb{E}\,[\,\cdot\,]$ is the expectation in the probability space that X is defined, we write $\mathcal{L}(X) \in \mathcal{P}(S)$ for its law:

$$\mathcal{L}(X)(h) = \mathbb{E}\,[h(X)].$$

The topology of weak convergence in $\mathcal{P}(S)$ is metrized by the Bounded-Lipschitz metric defined for $\mu, \nu \in \mathcal{P}(S)$:

$$\mathrm{d}_{BL}(\mu, \nu) := \sup \{\mu(h) - \nu(h) \; : \; h : S \to \mathbb{R}, \|h\|_{BL} \leq 1\}.$$

If X and Y are random elements in S defined on the same probability space and $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$ then

$$\mathrm{d}_{BL}(\mu, \nu) \leq \mathbb{E}\,[\mathrm{d}(X,Y) \wedge 2]. \tag{4.1}$$

4.3 Graphs

In this paper, a graph $G = (V, E)$ has vertex set V and unoriented edge set E. The set V is either finite or countably infinite. We write $x \sim y$ to denote that $xy = yx := \{x, y\} \in E$. Notice that we allow $x \sim x$ (i.e. a loop edge). The degree d_x of $x \in V$ is the number of $y \in V$ with $y \sim x$. When we need to specify the dependency on G we write V_G, E_G, $x \sim_G y$, and d_x^G. We always assume G is *locally finite*, i.e. $d_x < +\infty$ for all $x \in V$. We write $|G|$ and $e(G)$ for the number of vertices and edges in G, respectively.

Given a subgraph $H \subset G$ we define

1. $\partial H = \{v \in H : \exists u \in G \setminus H \text{ with } v \sim u\}$,
2. We write $\mathrm{dist}(v, u)$ for the distance between $v, u \in V$, i.e., the size of the shortest path between v and u in G,
3. $\mathrm{dist}(v, \partial H) = \inf_{u \in \partial H} \mathrm{dist}(v, u)$, and
4. For a subset of vertices $H_0 \subset H$ we define

$$\mathrm{dist}(H_0, \partial H) = \inf_{v \in H_0} \mathrm{dist}(v, \partial H).$$

We will also consider weighted graphs. For a graph $G = (V, E)$ the vector $\boldsymbol{\mu} = (\mu_e)_{e \in E}$ is a vector of weights for G if $\mu_e > 0$ for any $e \in E$. To each vector of weights $\boldsymbol{\mu}$ we can associate a matrix $(\mu_{vu})_{v,u \in V}$ such that for $v, u \in V$

- if $e = \{v, u\} \in E$ then $\mu_{vu} = \mu_{uv} = \mu_e$, and
- if $\{v, u\} \notin E$ then $\mu_{vu} = \mu_{uv} = 0$.

We write $\mu_v = \sum_{u \in V} \mu_{vu}$ for the total weight of v. We say that $(G, \boldsymbol{\mu})$ is a weighted graph and we identify the vector $\boldsymbol{\mu}$ with the associated matrix $(\mu_{vu})_{v,u \in V}$.

4.4 Models of random graphs

Some examples of our theory are related to random graph models. Given $n \in \mathbb{N}$, $p \in [0, 1]$, the Erdös-Rényi random graph $G(n, p)$ is the random graph with vertex set $[n]$ with no loops, where any two distinct $x, y \in [n]$ are adjacent with probability p, independently of all other pairs. We consider (as is customary) sequences of random graphs $G(n, p)$ where $p = p(n)$ may depend on n.

Given $n \in \mathbb{N}$ and $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{N}^n$, we let $G(n, \mathbf{d})$ denote a random graph with degree sequence $\mathbf{d}$, i.e. a graph that is chosen uniformly at random from the set of graphs G with $V(G) = [n]$, no loops, and $d_v^G = d_v$ for each $v \in [n]$. This makes sense only for certain sequences $\mathbf{d}$. One important particular case is that of random d-regular graphs, where $\mathbf{d} = (d, d, \ldots, d)$ for some $d \geq 3$ and we only need to assume dn even.

5 Local metrics and weak convergence of networks

In this section, we review the basic aspects of the local topology and local weak convergence of networks. We start with the case of graphs, which is better known. We then discuss the case of networks with more details. Our main references are the survey by Aldous and Steele [1], the lecture notes by Bordenave [5], and the paper [4].

5.1 Rooted graphs and local weak convergence

When we consider sparse random graphs, we will need to consider their *local weak limits*.

A rooted graph (G, o) consists of a (countable, locally finite) graph G with a distinguished vertex $o \in V_G$. Two rooted graphs (G, o), (H, p) are rooted isomorphic ($(G, o) \cong (H, p)$) if there exists a bijection $f : V_G \to V_H$ mapping o to p and preserving edges. The space $\mathcal{G}^*$ of rooted graphs considered up to isomorphisms can be endowed with a metrizable "local topology" that makes it a Polish space. Therefore, we may speak of random elements in this space (we will define a more general metric on networks below).

Given $r \in \mathbb{N}$, $(G, o)_r$ is the rooted graph with root o that contains the vertices $x \in V_G$ within distance r from o, and all the edges between these vertices. $[G, o]_r$ is the equivalence class of $(G, o)_r$. We write $G(v)$ for the connected component of v in G. We write $(G(v), v)$ for the graph $G(v)$ rooted at v and $[G(v), v]$ for its equivalence class up to isomorphism.

Definition 5.1. *For each finite graph $G \in \mathcal{G}$ we define the empirical neighbourhood distribution:*

$$U(G) = \frac{1}{|V_G|} \sum_{v \in V_G} \delta_{[G(v), v]}.$$

We say that a sequence of finite graphs G_n converges locally weakly to the measure $\rho \in \mathcal{P}(\mathcal{G}^*)$ if

$$U(G_n) \to \rho \text{ in the weak topology of } \mathcal{P}(\mathcal{G}^*).$$

If the sequence of finite graphs G_n is random then we say G_n converges almost surely to ρ in the local weak sense if the locally weakly convergence of G_n to ρ holds in a set of probability 1 with respect to the law of the sequence G_n.

Example 5.2. Cycle graphs C_n with n vertices locally weakly converge to $\delta_{(\mathbb{Z},0)}$.

Example 5.3. Suppose that, for each n, G_n has the law of the Erdös-Rényi random graph $G(n, c/n)$ for $c > 0$ constant. Then for almost all realizations of G_n, the sequence G_n locally weakly converges to the rooted GW tree with Poisson offspring distribution with mean c.

Example 5.4. Suppose that for each $n \in \mathbb{N}$ we have a vector

$$\mathbf{d}_n = (d_{n,1}, \ldots, d_{n,n}) \in \mathbb{N}^n.$$

Assume the sequence $\mathbf{d}_n$ has $\max_{1 \leq i \leq n} d_{i,n} \leq n^{\epsilon_n}$ with $\epsilon_n \to 0$ and the measures

$$P_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{d_{i,n}}$$

converge weakly to some P with finite first moment. If we sample G_n from $G(n, \mathbf{d}_n)$, then for almost all realizations of G_n, G_n locally weakly converges to the unimodular rooted GW tree $UGW(P)$, where the root has offspring distribution P, and all other nodes have offspring distribution

$$\widehat{P}(k) := \frac{(k+1)\,P(k+1)}{\sum_{i=1}^{\infty} iP(i)} \quad (k \in \mathbb{N}).$$

In particular, if $\mathbf{d}_n = (d, d, d, \ldots, d)$ for all n, $UGW(P)$ is the infinite (deterministic) d-regular tree rooted at a node.

5.2 Rooted networks and local weak convergence

Roughly speaking, a network is a graph $G = (V, E)$ with parameters (or marks) associated to the vertices and edges of G. The parameters (or marks) lie in some metric space.

More specifically, let (Υ, d_Υ) and (Ξ, d_Ξ) be two Polish metric spaces. A network $N = (V, E, \boldsymbol{v}, \boldsymbol{\xi})$ is a graph $G = (V, E)$ together with the vectors

$$\boldsymbol{v} = (v_v)_{v \in V} \in \Upsilon^{|V|} \text{ and } \boldsymbol{\xi} = (\xi_e)_{e \in E} \in \Xi^{|E|}$$

that gives marks to the vertices and edges of G, respectively. We write $\mathcal{N}_{(\Upsilon,\Xi)}$ for the space of all theses networks with the mark spaces fixed.

We say that $\overline{N} = (\overline{V}, \overline{E}, \overline{\boldsymbol{v}}, \overline{\boldsymbol{\xi}})$ is a sub-network of N if $(\overline{V}, \overline{E})$ is a induced sub-graph of (V, E), the vector $\overline{\boldsymbol{v}}$ is the restriction of $\boldsymbol{v}$ to $\overline{V}$ and $\overline{\boldsymbol{\xi}}$ is the restriction of $\boldsymbol{\xi}$ to $\overline{E}$. In this case we also say that the sub-network $\overline{N}$ is induced by the sub-graph $(\overline{V}, \overline{E})$.

When we write a graph property for a network $N = (V, E, \boldsymbol{v}, \boldsymbol{\xi})$ it is implicitly assumed that this property holds for the underlying graph. For example, $V_N := V_G$, $E_N := E_G$ and for a vertex $v \in N$, $d_v^N = d_v^G$. The boundary of a sub-network $\overline{N}$ of N is the boundary of the corresponding graphs.

Consider two networks $N = (V, E, \boldsymbol{v}, \boldsymbol{\xi})$ and $N' = (V', E', \boldsymbol{v}', \boldsymbol{\xi}')$ belonging to $\mathcal{N}_{(\Upsilon,\Xi)}$. A network isomorphism Ψ between N and N' is a bijection $\Psi : V \to V'$ between the vertex sets that preserves edges and marks:

- $\{u, v\} \in E$ if and only if $\{\Psi(u), \Psi(v)\} \in E'$,
- $v'_{\Psi(u)} = v_u, \forall u \in V$, and $\xi'_{\{\Psi(u),\Psi(v)\}\}} = \xi_{\{u,v\}}, \forall \{u, v\} \in E$.

A rooted network (N, o) is a network N with a distinguished vertex o. Two rooted networks (N, o) and (N', o') are rooted isomorphic if there is a network isomorphism that sends o to o'.

Given a rooted network (N, o) and a radius $r \in \mathbb{N}$ let $(N, o)_r$ be the network induced by $(G, o)_r$ rooted at the vertex o. Sometimes we identify the rooted network $(N, o)_r$ with its underlying network (without the root).

For a rooted network (N, o) with $N \in \mathcal{N}_{(\Upsilon, \Xi)}$ we associate its equivalence class $[N, o]$ of rooted isomorphism. Define

$$\mathcal{N}^*_{(\Upsilon, \Xi)} = \{[N, o] : (N, o) \text{ is a rooted network with mark spaces } (\Upsilon, \Xi)\}.$$

We now define a notion of distance over rooted networks up to isomorphism. This is not the exact same notion as in [5], but it is equivalent to it, as a simple calculation shows.

Consider two rooted networks $(N, o) = (V, E, \boldsymbol{v}, \boldsymbol{\xi}, o)$ and $(N', o') = (V', E', \boldsymbol{v}', \boldsymbol{\xi}', o')$ belonging to $\mathcal{N}^*_{(\Upsilon, \Xi)}$. Given $r \in \mathbb{N}$ and $\delta > 0$, we say that the pair (r, δ) is good for $(N, o), (N', o')$ if there exists a rooted isomorphism Ψ between $(N, o)_r$ and $(N', o')_r$ such that the corresponding marks are close by δ:

- $d_\Upsilon(v_v, v'_{\Psi(v)}) < \delta, \ \forall \ v \in (N, o)_r$, and

- $d_\Xi \left(\xi_{\{u,v\}}, \xi'_{\{\Psi(u), \Psi(v)\}} \right) < \delta, \ \forall \ \{u, v\} \in (N, o)_r$.

The distance between $[N, o]$ and $[N', o']$ is defined by

$$d_{\mathcal{N}^*_{(\Upsilon, \Xi)}} ([N, o], [N', o']) = \inf \left\{ \frac{1}{1 + r} + \delta \ : \ (r, \delta) \text{ is good for } (N, o), (N', o') \right\}. \quad (5.1)$$

One can show that $(\mathcal{N}^*_{(\Upsilon, \Xi)}, d_{\mathcal{N}^*_{(\Upsilon, \Xi)}})$ is Polish (cf. [5]).

Given two classes of rooted networks, we can define the distance between them using any representatives of these classes. This is well defined since the distance between rooted networks in invariant up to rooted isomorphism.

Sometimes we identify a rooted network with its equivalence class.

For a network N and a vertex $v \in V$ we associate the rooted network $(N(v), v)$ that is the network induced by the connected component of v rooted at v.

We write $[N(v), v]$ for the equivalence class of the rooted network $(N(v), v)$. For a finite network N with vertex set V we define the empirical measure neighbourhood:

$$U(N) = \frac{1}{|V|} \sum_{v \in V} \delta_{[N(v), v]}.$$

Definition 5.5 (Local weak convergence)**.** *Consider a sequence of networks*

$$N_n = ([n], E_n, \boldsymbol{v}_n, \boldsymbol{\xi}_n) \in \mathcal{N}_{(\Upsilon, \Xi)}.$$

Let $\rho \in \mathcal{P}\left(\mathcal{N}^*_{(\Upsilon, \Xi)}\right)$. We say that N_n converges locally weakly to ρ if

$$U(N_n) \to \rho$$

in the sense of weak convergence. If the sequence of finite networks N_n is random then we say that N_n converges almost surely to ρ in the local weak sense if the locally weakly convergence of N_n to ρ holds in a set of probability 1 with respect to the law of the sequence N_n.

5.2.1 Networks with i.i.d. marks that converge locally weakly

Now we give examples of networks that satisfy Definition 5.5.

Let $G = (V, E)$ be a graph. In Section 2 we introduced the networks

$$N = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}(\mathbf{0}))$$

where we have the vector of weights $\boldsymbol{\mu} = (\mu_e)_{e \in E} \in \mathbb{R}_+^{|E|}$ that are marks for the edges of G, environment or "media" variables $\boldsymbol{\omega} = (\omega_v)_{v \in V} \in \mathbb{R}^{|V|}$, and a vector of initial conditions $\boldsymbol{\theta}(\mathbf{0}) = (\theta_v(0))_{v \in V} \in \mathbb{R}^{|V|}$ that are marks for the vertices of G.

Definition 5.6. *We write $\mathcal{N}$ for the collection of networks $N = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}(\mathbf{0}))$ with edge marks in $\mathbb{R}_+$ and vertex marks in $\mathbb{R} \times \mathbb{R}$. When we distinguish a root for N we write $\mathcal{N}^*$ for the collection of rooted networks up to isomorphism.*

Our first goal is to show that for a sequence of graphs $(G_n)_{n \in \mathbb{N}}$ and under some additional conditions we can construct interesting examples of random networks $N_n \in \mathcal{N}$ and a probability measure $\nu \in \mathcal{P}(\mathcal{N}^*)$ such that the sequence $(N_n)_{n \in \mathbb{N}}$ converges to ν in the local weak sense, for almost all realization of the marks.

For any fixed graph $G = (V, E)$ we construct a probability space where we can define the vectors

$$\boldsymbol{\mu} = (\mu_e)_{e \in E}, \boldsymbol{\omega} = (\omega_v)_{v \in V}, \text{ and } \boldsymbol{\theta}(\mathbf{0}) = (\theta_v(0))_{v \in V} \text{ and they are independent.}$$

When we need to explicit the dependency on G we write $\boldsymbol{\mu}^G$, $\boldsymbol{\omega}^G$, and $\boldsymbol{\theta}(\mathbf{0})^G$.

Fix the measures $\pi, \lambda \in \mathcal{P}(\mathbb{R})$. For a fixed vertex v, π is the distribution of the media variables ω_v and λ is the distribution of the initial conditions $\theta_v(0)$. Fix a measure $\mu \in \mathcal{P}(\mathbb{R}_+)$ for the distribution of the weights μ_e for an edge $e \in E$.

We define by $N^G \in \mathcal{N}^*$ the random network obtained from G by adding these random marks.

In this way we have a transition kernel that associate to each rooted graph (G, o) the law of the random rooted network (N^G, o)

$$\begin{array}{cccc} \mathcal{M} & : & \mathcal{G}^* & \to & \mathcal{P}(\mathcal{N}^*) \\ & & (G, o) & \mapsto & \mathcal{L}(N^G, o) \end{array},$$

that is, $\mathcal{M}(G, o)(h) = \widetilde{\mathbb{E}} \left(h(G, o, \boldsymbol{\mu}^G, \boldsymbol{\omega}^G, \boldsymbol{\theta}(\mathbf{0})^G) \right)$ where $\widetilde{\mathbb{E}}(\cdot)$ is the expectation of the probability space where the marks for G were defined and $h : \mathcal{N}^* \to \mathbb{R}$ is a bounded measurable function.

The proofs of the following results are given in the appendix because they are easier version of our main results (Theorem 6.7 and Theorem 6.8, respectively).

Proposition 5.7 (Proof in Section A.1). *The transition kernel $\mathcal{M}$ is continuous.*

With this result at hand, for any probability measure ρ over $\mathcal{G}^*$ we can define $\rho\mathcal{M} \in \mathcal{P}(\mathcal{N}^*)$ via the formula

$$\rho\mathcal{M}(h) = \int \mathcal{M}(G, o)(h) \mathrm{d}\rho$$

for any bounded measurable function $h : \mathcal{N}^* \to \mathbb{R}$.

Theorem 5.8 (Proof in Section A.2). *Consider a sequence $\{G_n\}_{n \in \mathbb{N}} \subset \mathcal{G}$ of finite graphs where each G_n has vertex set $[n]$. Assume also the following*

1. *Local weak convergence: $\{G_n\}_n$ is locally weakly convergent to a probability measure ρ (in the sense of Definition 5.1).*
2. *Small maximum degree: for each n, $\max_{v \in [n]} d_v^{G_n} = n^{\epsilon_n}$ where $\epsilon_n \to 0$ as $n \to +\infty$.*

Then the sequence of random networks $\{N_n\}_n \subset \mathcal{N}$ defined by $N_n := N^{G_n}$, containing the random marks, converges to $\rho\mathcal{M}$ in the local weak sense and almost surely with respect to the law of the random marks.

6 Interacting diffusions on finite networks

In this section, we introduce our main objects of interest, and give formal versions of Theorem 2.1 and Theorem 2.3.

Fix functions $\psi : \mathbb{R}^2 \to \mathbb{R}$ and $\phi : \mathbb{R}^4 \to \mathbb{R}$. In our model, ψ represents a drift term that depends on the current position and the media variable, whereas ϕ corresponds to pairwise interactions that represent single-term drift and the interaction between the particles. As we can always suppose $V \subset \mathbb{N}$ we consider $(B_v)_{v \in \mathbb{N}}$ a collection of i.i.d. Standard Brownian Motions.

Definition 6.1. *Let $N = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}(0))$ be a network belonging to the set $\mathcal{N}$ in Definition 5.6. A system of interacting diffusions on the network N (with the choice of functions ψ, ϕ) is a random vector*

$$\boldsymbol{\theta}(\cdot) = (\theta_v(\cdot))_{v \in V} \in C([0, T]; \mathbb{R})^{|V|}$$

which is a strong solution of the following system of Itô Stochastic Differential Equations (SDEs): for each $v \in V$,

$$\mathrm{d}\theta_v(t) = \frac{1}{\mu_v} \sum_{u \in V} \mu_{uv} \phi(\theta_u(t), \theta_v(t); \omega_v, \omega_u) \mathrm{d}t + \psi(\theta_v(t); \omega_v) \mathrm{d}t + \mathrm{d}B_v(t), \qquad (6.1)$$

where the first term in the RHS in (6.1) is zero if $\mu_v = 0$, and the initial conditions are given by $\boldsymbol{\theta}(0) = (\theta_v(0))_{v \in V}$. When we need to make explicit the dependency on the network we will write $\boldsymbol{\theta}(\cdot) =: \boldsymbol{\theta}^N(\cdot)$.

When the network N is finite the standard theory of Itô SDEs guarantees that the system (6.1) has a unique strong solution with continuous trajectories whenever ψ and ϕ are Lipschitz-continuous (see [15], Theorem 2.9, Chapter 5). In the remainder of this work, ψ is assumed to be Lipschitz-continuous and ϕ is assumed to be Lipschitz-continuous and bounded. We will argue that, under certain conditions, we also can solve (6.1) simultaneously with infinite equations (on infinite networks).

For each network N, if there exists a system of interacting diffusions over N, we can replace the initial conditions $\boldsymbol{\theta}(0)$ by the vector of random continuous functions $\boldsymbol{\theta}^N(\cdot) \in C([0, T]; \mathbb{R})$ as new marks. This brings us to the following definitions.

Definition 6.2. *Given a network $N = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}(0)) \in \mathcal{N}$ and a vector of continuous functions*

$$\boldsymbol{\alpha}(\cdot) = (\alpha_v(\cdot))_{v \in G} \in C([0, T]; \mathbb{R})^{|G|}$$

we write $\mathcal{C}$ for the collection of networks $(G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\alpha}(\cdot))$ with continuous functions replacing the vector $\boldsymbol{\theta}(0)$ as new marks. That is, $\mathcal{C}$ has edge marks in $\mathbb{R}_+$ and vertex marks in $\mathbb{R} \times C([0, T]; \mathbb{R})$. When we distinguish a root for $(G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\alpha}(\cdot))$ we write $\mathcal{C}^*$ for the collection of these rooted networks up to isomorphism.

Definition 6.3. *When $N \in \mathcal{N}$ is such that a system of interacting diffusions $\boldsymbol{\theta}^N(\cdot)$ can be defined, we let $N^\theta = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}^N(\cdot))$ be the corresponding random element of $\mathcal{C}$ defined above. Note that the law of N^θ is invariant by network isomorphisms.*

To present our main results, we need some additional definitions. We first restrict ourselves to the space of finite networks.

Definition 6.4. *Let $\mathcal{N}_f$ be the collection of finite networks contained in $\mathcal{N}$. We write $\mathcal{N}_f^*$ when considering finite rooted networks of $\mathcal{N}^*$ up to isomorphism.*

Finally, we define a map associating to each finite network $[N, o]$ the law of $[N^\theta, o]$.

Definition 6.5. *For each rooted finite network $[N, o] \in \mathcal{N}_f^*$, we let $\Theta([N, o]) \in \mathcal{P}(\mathcal{C}^*)$ denote the law of the random network $[N^\theta, o]$ (up to isomorphism). This is well-defined*

because the law of N^θ is invariant under isomorphisms of N. Therefore, Θ defines a map from $\mathcal{N}_f^*$ (the space of finite networks) to $\mathcal{P}(\mathcal{C}^*)$ (the space of probability measures over $\mathcal{C}^*$).

6.1 Main results

In this section we state our main results. Our results hold when the sequence $\{N_n\}_{n\in\mathbb{N}}$ converges to something "nice".

Definition 6.6 ("Nice" networks). *For each rooted network $(N, o) \in \mathcal{N}^*$ define $\mu_*(N)$ (resp. $\mu^*(N)$) as the infimum (resp. supremum) of the set $\{\mu_e : e \in E_N\}$. Let (G, o) be the underlying rooted graph of (N, o). Call (N, o) nice if:*

1. *Graph grows at most exponentially: there exists $a > 0$ such that*

$$|\partial(G, o)_r| \le ae^{ar} \text{ for all } r \ge 1; \text{ and}$$

2. *Weights bounded away from 0 and $+\infty$: $0 < \mu_*(N) \le \mu^*(N) < +\infty$.*

We let $\mathcal{B}^* \subset \mathcal{N}^*$ denote the set of nice rooted networks (this is a Borel set). Observe that $\mathcal{N}_f^* \subset \mathcal{B}^*$.

See Section A.3 for examples of nice networks.

Our first main result ensures that systems of interacting diffusions may be defined over nice networks (N, o) as limits, taking $r \to \infty$, of systems over the finite networks $(N, o)_r$.

Theorem 6.7 (Extension of Θ, proof in Section 8). *We can extend Θ to a continuous transition kernel $\overline{\Theta} : \mathcal{B}^* \to \mathcal{P}(\mathcal{C}^*)$. More specifically, given $(N, o) \in \mathcal{B}^*$ there exists the unique strong solution of the (possibly infinite) system of SDEs in (6.1) defined over N. Furthermore, $\overline{\Theta}(N, o)$ is the law of $(N^\theta, o) \in \mathcal{C}^*$ defined replacing the initial conditions $\theta(0)$ by $\theta^N(\cdot)$.*

In our next result, we use Theorem 6.7 to identify hydrodynamic limits of systems of interacting diffusions. We first note that, if ν is a probability measure over $\mathcal{B}^*$, the continuity (in particular measurability) of $\overline{\Theta}$ allows us to define $\nu\overline{\Theta}$ by the formula:

$$\nu\overline{\Theta}(h) = \int \overline{\Theta}([N, o])(h)\, d\nu([N, o]) = \int \mathbb{E}\left[h(N^\theta, o)\right] d\nu([N, o]), \qquad (6.2)$$

where $h : \mathcal{C}^* \to \mathbb{R}$ is a bounded measurable function and the expectation $\mathbb{E}[\cdot]$ is with respect to the Brownian motions.

Theorem 6.8 (Hydrodynamic limit, proof in Section 9). *Consider a sequence*

$$\{N_n\}_{n\in\mathbb{N}} \subset \mathcal{N}_f$$

of finite networks where each N_n has vertex set $[n]$. Make the following additional assumptions.

1. *Local weak convergence: $\{N_n\}_{n\in\mathbb{N}}$ is locally weakly convergent to a probability measure ν (in the sense of Definition 5.5).*
2. *Good limiting network: the measure ν is supported on nice networks (cf. Definition 6.6).*
3. *Small maximum degree: the largest degree in N_n satisfies $\max_{v\in[n]} d_v^{N_n} = n^{\epsilon_n}$ where $\epsilon_n \to 0$ as $n \to +\infty$.*

Then the sequence of random networks $\{N_n^\theta\}_{n\in\mathbb{N}} \subset \mathcal{C}$ containing the random trajectories of the interacting diffusions converges to $\nu\overline{\Theta} \in \mathcal{P}(\mathcal{C}^*)$ in the local weak sense and almost surely with respect to the law of the Brownian motions.

Remark 6.9. We show in Section A.3 that taking $(G_n)_{n\in\mathbb{N}}$ from the examples 5.2, 5.3 and 5.4 of almost surely (with respect to the graph randomness) local weak convergence we can construct networks N_n that satisfy the assumptions of Theorem 6.8 for almost all realizations of marks and almost all realizations of the sequence G_n. In particular, the theorem is true for the case of Erdös-Rényi and GW trees with probability 1 with respect to the law of the Erdös-Rényi graphs.

The hydrodynamic limit is in general stated for the empirical measure over particle trajectories.

$$L_n = \frac{1}{n} \sum_{v\in N_n} \delta_{\theta_v^{N_n}(\cdot)} \in \mathcal{P}(C([0,T];\mathbb{R})).$$

We can obtain θ_o^N as a projection of (N^θ, o). Using that the projection is continuous we obtain the following corollary.

Corollary 6.10 (Proof omitted). *In the conditions of Theorem 6.8, L_n converges almost surely to the law of θ_o^N in the weak topology of $\mathcal{P}(C([0,T];\mathbb{R}))$.*

Remark 6.11. To see that the limiting object is not Markovian consider the case when ν is supported on a deterministic rooted d-regular tree (T, o). Then the evolution $\theta_o^T(s)$ in the time interval $t \leq s \leq T$ depends also on the values $\{\theta_v^T(t) : v \sim_T o\}$ (see (6.1)). Observe however that the full vector $\boldsymbol{\theta}^N = (\theta_v^N)$ is Markovian. Since it is a strong solution then $\boldsymbol{\theta}^N(t+s) - \boldsymbol{\theta}^N(t)$ is independent of $\mathcal{F}_t = \sigma(B_v(u) : 0 \leq u \leq t, v \in N)$ conditionally on $\boldsymbol{\theta}^N(t)$.

As shown in [28, Propostion 2.2], the hydrodynamic limit immediately implies the Propagation of Chaos (cf. [28, Definition 2.1]). For completeness, we show a self-contained proof in Appendix D.

Corollary 6.12 (Propagation of Chaos). *Let $k \in \mathbb{N}$ and*

$$f_1, \cdots, f_k : \mathcal{C}^* \to \mathbb{R}, \text{ such that } \|f_i\|_{BL} \leq 1, \forall 1 \leq i \leq k.$$

Consider (N_i, o_i) i.i.d. with law ν. Under our assumptions,

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{i=1}^k U(N_n^\theta)(f_i)\right] \to \prod_{i=1}^k \mathbb{E}\left[f_i(N_i^\theta, o_i)\right].$$

7 The locality lemma

In this section, we introduce the main technical tool in our proofs, the Locality lemma. It will be present in the proofs of all the results in Section 6.1.

The Locality lemma basically establishes that our interacting diffusions over a finite subset H_0 of vertices are indifferent to parts of the network that are far away from H_0.

Lemma 7.1 (Locality). *Consider a rooted network*

$$(N, o) = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta}(0), o)$$

according to Definition 5.6. Let H_0 and H be finite subgraphs of G with $H_0 \subset H$. Let $\overline{N}$ be the sub-network induced by H (cf. Section 5.2). Fix functions ψ, ϕ that are Lipschitz with ϕ bounded. Let $(B_v)_{v\in V}$ be i.i.d. Brownian motions associated with the vertices of G.

Following Definition 6.1, assume that we can define a system of interacting diffusions $\boldsymbol{\theta}^N(\cdot)$ over N from the $(B_v)_{v\in V}$ (with ψ and ϕ fixed above and in the time interval $[0,T]$). Also build a system $\boldsymbol{\theta}^{\overline{N}}(\cdot)$ over $\overline{N}$ with the same Brownian motions (this works because

$\overline{N}$ *is finite). Then the following holds: there exist* $C, r_0 > 0$ *depending only on* T, $\|\psi\|_{Lip}$ *and* $\|\phi\|_{BL}$ *such that, almost surely on the Brownian motion randomness,*

$$if \ r := \operatorname{dist}(H_0, \partial H) \geq r_0, \ then \ \max_{v \in H_0} \left(\sup_{t \leq T} |\theta_v^N(t) - \theta_v^{\overline{N}}(t)| \right) \leq C|\partial H| \exp\left(-r \log r\right).$$

The rest of the section will be devoted to the proof of Lemma 7.1. However, if the reader so wishes, she or he can skip the proof and go directly to Section 8. In this section the only randomness comes from the diffusions. So every "almost surely" statement in this section is with respect to the law of the Brownian motions.

The general idea of the proof is the following: In §7.1 we rewrite the diffusions $\boldsymbol{\theta}^N(\cdot)$ and $\boldsymbol{\theta}^{\overline{N}}(\cdot)$ so that they can be compared easily. In §7.2, we prove a linear Gronwall inequality for the difference between the two systems. The proof is finished in §7.3. At this step, we need to analyze a certain matrix exponential. We do so via the theory of continuous-time random walks, most importantly Carne-Varoupoulos heat kernel bound [3, Section 5.1].

As noted in the introduction, Lacker et al. [17, 18] essentially bypass locality estimates in their proofs. We strongly believe that the same locality estimates can be proven in their framework.

7.1 Preliminaries

To avoid cumbersome notation, we adopt the following notation conventions. Objects related to the network $N = (G, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{\theta(0)})$ are written without superscripts. Degrees of vertices are indicated via d_v. We define a matrix P indexed by the vertices V of N via:

$$P_{uv} := \frac{\mu_{uv}}{\mu_v} \mathbb{I}_{\{\mu_v > 0\}}.$$

With this notation, we are assuming the existence of the interacting diffusions over N that may be written as

$$d\theta_v(t) = \sum_{u \in G} P_{u,v} \phi(\theta_u(t), \theta_v(t); \omega_v, \omega_u) dt + \psi(\theta_v(t), \omega_v) dt + dB_v(t). \tag{7.1}$$

The network $\overline{N} = (H, \boldsymbol{\mu} \mid_H, \boldsymbol{\omega} \mid_H, \boldsymbol{\theta(0)} \mid_H)$ is induced by the sub-graph $H \subset G$. We will write the corresponding process somewhat differently. Define $\overline{\mu}_{uv} := \mu_{uv} \mathbb{I}_{\{v, u \in H\}}$ and $\overline{\mu}_v := \sum_u \overline{\mu}_{uv}$. We set:

$$\overline{P}_{uv} := \frac{\overline{\mu}_{uv}}{\overline{\mu}_v} \mathbb{I}_{\{\overline{\mu}_v > 0\}}.$$

This matrix is in general different from P if v or u are either outside of H or in ∂H. We may define another system of diffusions satisfying:

$$d\overline{\theta}_v(t) = \sum_{u \in G} \overline{P}_{u,v} \phi(\overline{\theta}_u(t), \overline{\theta}_v(t); \omega_v, \omega_u) dt + \psi(\overline{\theta}_v(t), \omega_v) dt + dB_v(t) \tag{7.2}$$

for each $v \in V$ (and not just the vertices in H), with initial conditions $\overline{\theta}_v(0) = \theta_v(0)$. With this definition, the diffusions inside H do not interact with those outside H. Since H is finite, the system inside H has a unique strong solution, so the $\overline{\theta}_v(\cdot)$ with $v \in H$ correspond exactly to the $\theta_v^{\overline{N}}(\cdot)$ in the statement of the Lemma.

Our goal then is to bound $\sup_{t \leq T} |\theta_v(t) - \overline{\theta}_v(t)|$ for $v \in H_0$. More specifically, it suffices to prove the next result

Lemma 7.2. *Let* $v \in H$. *Then almost surely*

$$if \operatorname{dist}(v, \partial H) =: r \geq r_0, \ then \ \sup_{t \leq T} \left(\theta_v(t) - \overline{\theta}_v(t) \right) \leq C|\partial H| \exp\left(-r \log r\right).$$

Indeed, once we have this, Lemma 7.1 follows if we take the supremum over $v \in H_0$. The remainder of this section is dedicated to the proof of Lemma 7.2.

7.2 A Gronwall bound

The next proposition will be used in our Grownwall argument. To state it, we let $\widetilde{H}$ denote the set of vertices within distance at most 1 of H, that is

$$\widetilde{H} = \{u \in G : u \in H \text{ or } \exists v \in H \text{ with } v \sim u\}.$$

Also, we write I for the identity matrix.

Proposition 7.3. *Let $v \in H$. The following inequality holds almost surely for all $t \in [0, T]$,*

$$|\theta_v(t) - \overline{\theta}_v(t)| \leq C \int_0^t \sum_{u \in \widetilde{H}} (\overline{P} + I)_{uv} |\theta_u(s) - \overline{\theta}_u(s)| ds + 2CT \mathbb{I}_{\{v \in \partial H\}},$$

where $C > 0$ depends only on $\|\psi\|_{Lip}$ and $\|\phi\|_{BL}$.

Proof. Since the initial conditions are coupled to be equal they cancel when we calculate $\theta_v(t) - \overline{\theta}_v(t)$. The Brownian motion also cancels. From Equation 7.1 we obtain

$$\theta_v(t) - \overline{\theta}_v(t) = \int_0^t \Delta_1(v, s) + \Delta_2(v, s) + \Delta_3(v, s) ds$$

where

$$\begin{aligned}
\Delta_1(v, s) &:= \sum_{u \in G} (P_{uv} - \overline{P}_{uv}) \phi(\theta_u(s), \theta_v(s); \omega_u, \omega_v), \\
\Delta_2(v, s) &:= \sum_{u \in G} \overline{P}_{uv} (\phi(\theta_u(s), \theta_v(s); \omega_u, \omega_v) - \phi(\overline{\theta}_u(s), \overline{\theta}_v(s); \omega_u, \omega_v)), \\
\Delta_3(v, s) &:= \psi(\theta_v(s); \omega_v) - \psi(\overline{\theta}_v(s); \omega_v).
\end{aligned}$$

First observe that the sums involving P and $\overline{P}$ for $u \in G$ are in fact over $u \in \widetilde{H}$ since $v \in H$. This is due the fact this matrices have non-zero entries $P_{\cdot,v}, \overline{P}_{\cdot,v}$ just for neighbors of v. In this way we have the following bounds

$$\begin{aligned}
|\Delta_1(v, s)| &\leq \|\phi\|_\infty \left(\sum_{u \in \widetilde{H}} |P_{uv} - \overline{P}_{uv}| \right), \\
|\Delta_2(v, s)| &\leq \sum_{u \in \widetilde{H}} \overline{P}_{uv} \|\phi\|_{Lip} \left(|\theta_v(s) - \overline{\theta}_v(s)| + |\theta_u(s) - \overline{\theta}_u(s)| \right), \text{ and} \\
|\Delta_3(v, s)| &\leq \|\psi\|_{Lip} |\theta_v(s) - \overline{\theta}_v(s)|.
\end{aligned}$$

Now we bound the RHS of the first inequality. For this, we note that if $v \notin \partial H$ or $\mu_v = 0$, then $P_{uv} = \overline{P}_{uv}$ for all u. On the other hand, if $v \in \partial H$ and $\mu_v > 0$, then:

$$\begin{aligned}
\sum_{u \in \widetilde{H}} |P_{uv} - \overline{P}_{uv}| &= \sum_{u \in H} |P_{uv} - \overline{P}_{uv}| + \sum_{u \notin H} |P_{uv} - \overline{P}_{uv}| \\
(u \notin H \implies \overline{P}_{uv} = 0) &= \sum_{u \in H} |P_{uv} - \overline{P}_{uv}| + \sum_{u \notin H} |P_{uv}| \\
(u \in H \implies \mu_{uv} = \overline{\mu}_{uv}) &= \sum_{u \in H} \mu_{uv} \left| \frac{1}{\mu_v} - \frac{1}{\overline{\mu}_v} \right| + \sum_{u \notin H} |P_{uv}| \\
&\leq \overline{\mu}_v \left| \frac{1}{\mu_v} - \frac{1}{\overline{\mu}_v} \right| + (\mu_v - \overline{\mu}_v) \frac{1}{\mu_v} \\
(\overline{\mu}_v \leq \mu_v) &= 2 \left(1 - \frac{\overline{\mu}_v}{\mu_v} \right) \leq 2.
\end{aligned}$$

In any case we have that,

$$\sum_{u \in \widetilde{H}} |P_{uv} - \overline{P}_{uv}| \le 2\mathbb{I}_{v \in \partial H}.$$

Combining these bounds, we obtain the result. □

7.3 End of proof

We now finish the proof of Lemma 7.2, which implies Lemma 7.1. We will apply the Linear Gronwall's Inequality (Corollary C.2) to finish the proof of Lemma 7.2.

Going back to Proposition 7.3 we observe that $\widetilde{H}$ is finite. In particular, the matrices $\overline{P}$ and I considered are finite-dimensional. So we can apply Corollary C.2 with:

1. $\mathbf{u}(t) = (|\theta_v(t) - \overline{\theta}_v(t)|)_{v \in \widetilde{H}}$, which have continuous entries,
2. $\mathbf{a}(t) = \mathbf{a} := (CT\mathbb{I}_{[v \in \partial H]})_{v \in \widetilde{H}}$, and we observe that each entry of this vector is non-negative, and
3. $M(t)_{uv} = M_{uv} = C(\overline{P} + I)_{uv}$, for $u, v \in \widetilde{H}$ and we observe that this matrix does not depend on time t, it is entry-wise non-negative and it is finite dimensional.

In vector notation, we obtain:

$$\mathbf{u}(t) \le \int_0^t (M\mathbf{u}(s) + \mathbf{a})\, ds$$

and the Corollary says that $\mathbf{u}(t) \le \exp(tM)\,\mathbf{a}$ entrywise. This is the same as saying that, for each $v \in H$

$$|\theta_v(t) - \overline{\theta}_v(t)| \le CT \sum_{u \in \partial H} \exp(Ct\,(\overline{P} + I))_{uv} \le CTe^{2T} \sum_{u \in \partial H} \exp(Ct\,(\overline{P} - I))_{uv}.$$

To bound this last expression, we note that

$$\exp(Ct\,(\overline{P} - I))_{uv} = \overline{\mu}_u\, q_{Ct}(u, v)$$

where $q_{Ct}(u, v)$ is the heat kernel at time Ct of a continuous time random walk over H with transition rates equal to 1 and reversible transition probabilities $\overline{P}_{uv}$ (reversibility follows from symmetry of μ_{uv}). The Carne-Varoupoulos bound for the heat kernel (Theorem 5.17 of [3, Section 5.1]) implies that for any time $s \ge 0$ and any $v, u \in H$ with $R = \mathrm{dist}(v, u) \ge es$ (e is the Euler constant)

$$q_s(v, u) \le \frac{1}{\overline{\mu}_v \vee \overline{\mu}_u} \exp\left(-s - R \log \frac{R}{es}\right) \le \frac{1}{\overline{\mu}_u} \exp\left(-s - R \log \frac{R}{es}\right).$$

We apply this with $s = Ct \le CT$, and obtain that, if $R = \mathrm{dist}(v, \partial H) \ge eCT$, then:

$$\sup_{t \le T} |\theta_v(t) - \overline{\theta}_v(t)| \le |\partial H| CTe^{2T} \exp\left(-Ct - R \log \frac{R}{eCT}\right).$$

So we finish by taking $r_0 = \lceil eCT \rceil$ and adjusting C accordingly.

8 Interacting diffusions over infinite graphs

In this section, we prove Theorem 6.7. We first give a sketch of the argument.

1. Our main task will be to construct our system of diffusions over infinite networks via limits of systems over finite sub-networks. More precisely, for an infinite rooted network (N, o) and a radius $r \ge 1$ we have the finite rooted network $(N, o)_r$ (cf. Section 5.2). It is clear that the sequence $((N, o)_r)_{r \ge 1}$ converges in the local topology towards to (N, o).

2. The random networks $(N, o)_r^\theta$ $((N, o)_r^\theta \neq (N^\theta, o)_r$, see Definition 6.3) are well defined since the networks $(N, o)_r$ are finite. We will use the fact that (N, o) is nice (cf. Definition 6.6) and Lemma 7.1 to show that the sequence $(N, o)_r^\theta$ (considering o as the root) has a limit in distribution. We will then show that the limit is the rooted network (N^θ, o) replacing the initial conditions by $\boldsymbol{\theta}^N(\cdot)$ that is the unique strong solution of the infinite system of SDEs in (6.1) for N.

3. To prove the continuity of the map (N, o) nice $\mapsto$ Law of (N^θ, o) we will also need to use that for finite networks the map (N, o) finite $\mapsto$ Law of (N^θ, o) is continuous (Lemma B.1).

We give the formal proof of Theorem 6.7 over the next subsections.

8.1 Proof of existence

Throughout the proof, we will assume that we are given i.i.d. standard Brownian motions $(B_v)_{v \in V}$ defined on the vertices of our network. The equalities and estimates in all Section 8 will hold almost surely with respect to the law of the Brownian motions. For each $r \geq 0$, we have a finite network $(N, o)_r$ with Brownian motions attached to its vertices. Following Definition 6.1, we can build a system of interacting diffusions over the vertices $v \in (N, o)_r$ via:

$$\mathrm{d}\theta_v^{(r)}(t) = \frac{1}{\mu_v^{(r)}} \sum_{u \in (N,o)_r} \mu_{uv} \phi(\theta_u^{(r)}(t), \theta_v^{(r)}(t); \omega_v, \omega_u) \mathrm{d}t + \psi(\theta_v^{(r)}(t); \omega_v) \mathrm{d}t + \mathrm{d}B_v(t), \quad (8.1)$$

with the weights $\boldsymbol{\mu} = (\mu_{vu})_{vu \in E_G}$, media variables $\boldsymbol{\omega} = (\omega_v)_{v \in V_G}$ and initial conditions $\boldsymbol{\theta}^{(r)}(0) = (\theta_v(0))_{v \in V_G}$ determined by the network $(N, o)_r$. We also use the notation $\mu_v^{(r)} := \sum_{u \in (N,o)_r} \mu_{uv}$. The point of this construction is that it couples our interacting diffusions over $(N, o)_r$ for all r simultaneously.

We now apply Lemma 7.1, with $H_0 = (N, o)_s$ for some $s \geq 0$, $H = (N, o)_r$ for $r \geq s$, and with the network N replaced by a finite ball $(N, o)_{r'}$ with $r' \geq r$. Since $(N, o)_{r'}$ is always finite, the solution to the system is well defined and we can indeed apply the Lemma. We have that $\mathrm{dist}((N, o)_s, \partial(N, o)_r) = r - s$. Moreover, since (N, o) is nice, there exist $a > 0$ independent of r with $|\partial(N, o)_r| \leq ae^{ar}$. We conclude that

$$\forall r' \geq r \geq s + r_0, \quad \max_{v \in (N,o)_s} \left(\sup_{t \leq T} |\theta_v^{(r)}(t) - \theta_v^{(r')}(t)| \right) \leq Ca \exp(ar - (r - s) \log(r - s)).$$

The fact that the RHS goes to 0 for fixed s and $r', r \to +\infty$ implies that, for each $v \in V$, $\theta_v^{(r)}(\cdot)$ is a Cauchy sequence (in the uniform norm over $[0, T]$) and converges over $[0, T]$ to a continuous function $\theta_v^N(\cdot)$. We also have the estimate:

$$\forall r \geq s + r_0, \quad \max_{v \in (N,o)_s} \left(\sup_{t \leq T} |\theta_v^{(r)}(t) - \theta_v^N(t)| \right) \leq Ca \exp\left(ar - (r - s) \log(r - s) \right), \quad (8.2)$$

and we will show now that this estimate implies that $\boldsymbol{\theta}^N = (\theta_v^N(\cdot))_{v \in V}$ is a system of interacting diffusions over N in the sense of Definition 6.1. To do this first observe that $\mu_v^{(r)} = \mu_v$ for large r. For any fixed $r \geq 1$ we have from (8.1)

$$\theta_v^{(r)}(t) = \theta_v^{(r)}(0) + \frac{1}{\mu_v^{(r)}} \sum_{u \in (N,o)_r} \mu_{uv} \int_0^t \phi(\theta_u^{(r)}(s), \theta_v^{(r)}(s); \omega_v, \omega_u) \mathrm{d}s +$$
$$+ \int_0^t \psi(\theta_v^{(r)}(s); \omega_v) \mathrm{d}s + B_v(s). \quad (8.3)$$

The estimate (8.2) says that for all u neighbour of v it holds that $\theta_u^{(r)}$ converges to θ_u^N in the supremum norm and almost surely. Using the basic fact that if f_n converges in the supremum norm to f then $\int_0^t f_n(s)\mathrm{d}s \to \int_0^t f(s)\mathrm{d}s$ we obtain from (8.3) that the following holds almost surely

$$\theta_v^N(t) = \theta_v^N(0) + \frac{1}{\mu_v^N} \sum_{u \in N} \mu_{uv} \int_0^t \phi(\theta_u^N(s), \theta_v^N(s); \omega_v, \omega_u)\mathrm{d}s +$$
$$+ \int_0^t \psi(\theta_v^N(s); \omega_v)\mathrm{d}s + B_v(s),$$

as we wanted.

8.2 Proof of uniqueness

The above implies *existence* of a system of interacting diffusions over N. *Uniqueness* of such a process is also easy to obtain. Indeed, suppose $\boldsymbol{\beta}^N$ is another strong solution to the same system of equations defined in terms of the same Brownian motions $(B_v)_{v \in V}$. Then an application of the locality result, Lemma 7.1, to $H_0 = \{v\}$, $H = (N, o)_r$ for large r, and N, reveals that (8.2) must also hold with β_v^N replacing θ_v^N. Therefore,

$$\forall v \in V \,:\, \beta_v^N(\cdot) = \lim_{r \to +\infty} \theta_v^{(r)}(\cdot) = \theta_v^N(\cdot).$$

8.3 Proof of continuity

What we have seen so far is that for each nice rooted network (N, o) one may uniquely define a system of interacting diffusions $\boldsymbol{\theta}^N(\cdot)$. Let $[N^\theta, o] \in \mathcal{C}^*$ be the resulting random network when one replaces the initial condition for the diffusions as new marks to the vertices (in the sense of Definition 6.3). So let $\overline{\Theta}[N, o]$ denote the law of $[N^\theta, o] \in \mathcal{C}^*$ (Definition 6.5). The uniqueness statement above implies that $\overline{\Theta}$ extends the definition of Θ over finite networks.

We must now show that $\overline{\Theta}$ is a *continuous* map from $\mathcal{B}^*$ (the set of nice networks) to $\mathcal{P}(\mathcal{C}^*)$ (the set of probability measures over $\mathcal{C}^*$ with the BL metric). We start with some preliminaries. We note once again that, due to Lemma 7.1, we have the more precise estimate:

$$\forall r \ge s + r_0, \; \max_{v \in (N, o)_s} \left(\sup_{t \le T} |\theta_v^{(r)}(t) - \theta_v^N(t)| \right) \le C|\partial(N, o)_r| \exp\left(-(r - s)\log(r - s)\right). \quad (8.4)$$

This bound immediately translates to a bound for the distance between (N^θ, o) and $(N, o)_r^\theta$ in the space $\mathcal{C}^*$. Furthermore, this construction of (N^θ, o), and $(N, o)_r^\theta$ is a coupling of the measures $\overline{\Theta}[N, o]$ and $\overline{\Theta}[N, o]_r$, and we obtain:

$$\forall r \ge s + r_0, \; d_{BL}(\overline{\Theta}[N, o], \overline{\Theta}[N, o]_r) \le \frac{1}{1 + s} + C|\partial(N, o)_r| \exp\left(-(r - s)\log(r - s)\right). \quad (8.5)$$

The important point is that (8.5) is applicable to *all* $[N, o] \in \mathcal{B}^*$.

Now consider a sequence $[N_n, o_n]$ of networks converging to $[N, o]$. We wish to show that $d_{BL}(\overline{\Theta}[N, o], \overline{\Theta}[N_n, o_n]) \to 0$. To do this, we use the following observation: for any fixed r, the probability laws corresponding to $[N_n, o_n]_r$ converge to those of $[N, o]_r$:

$$d_{BL}(\overline{\Theta}[N, o]_r, \overline{\Theta}[N_n, o_n]_r) \to 0. \quad (8.6)$$

Indeed, this is true because $[N, o]_r$ and $[N_n, o_n]_r$ are *finite*. This is discussed in detail in Section B of the Appendix.

The triangle inequality gives:

$$
\begin{aligned}
d_{BL}(\overline{\Theta}[N,o], \overline{\Theta}[N_n,o_n]) \leq\ & d_{BL}(\overline{\Theta}[N,o]_r, \overline{\Theta}[N_n,o_n]_r) \\
& + d_{BL}(\overline{\Theta}[N,o]_r, \overline{\Theta}[N,o]) \\
& + d_{BL}(\overline{\Theta}[N_n,o_n], \overline{\Theta}[N_n,o_n]_r).
\end{aligned}
$$

When $n \to +\infty$, the first term in the RHS shrinks to 0. Using (8.5) to bound the other two terms, we obtain:

$$
\begin{aligned}
\limsup_{n\to\infty} d_{BL}(\overline{\Theta}[N,o], \overline{\Theta}[N_n,o_n]) \leq\ & \frac{1}{1+s} + C|\partial(N,o)_r| \exp\left(-(r-s)\log(r-s)\right) \\
& + \limsup_{n\to\infty} \left(\frac{1}{1+s} + C|\partial(N_n,o_n)_r| \exp\left(-(r-s)\log(r-s)\right) \right).
\end{aligned}
$$

The local convergence of $[N_n, o_n]$ to $[N, o]$ implies that $|\partial(N_n, o_n)_r| \to |\partial(N, o)_r|$, so:

$$
\limsup_{n\to\infty} d_{BL}(\overline{\Theta}[N,o], \overline{\Theta}[N_n,o_n]) \leq \frac{2}{1+s} + 2C|\partial(N,o)_r| \exp\left(-(r-s)\log(r-s)\right).
$$

Since $r \geq r_0 + s$ are arbitrary (and r_0 is constant), we may let $s = r/2$ and make $r \to +\infty$ to obtain

$$
\limsup_{n\to\infty} d_{BL}(\overline{\Theta}[N,o], \overline{\Theta}[N_n,o_n]) \leq 0,
$$

as desired.

9 Hydrodynamic limit

In this section we will prove Theorem 6.8.

Our goal is the following. Given that $U(N_n) \to \nu$ in the space $\mathcal{P}(\mathcal{N}^*)$ and up to some additional conditions we want to show that $U(N_n^\theta) \to \nu\overline{\Theta}$ in $\mathcal{P}(\mathcal{C}^*)$ almost surely. In this section "almost surely" means "almost surely with respect to the law of the Brownian motions". By a standard argument, it suffices to show that for any test function $h : \mathcal{C}^* \to \mathbb{R}$ with $\|h\|_{BL} \leq 1$,

Goal: $U(N_n^\theta)(h) \to \nu\overline{\Theta}(h)$ almost surely,

with respect to the law of the Brownian motions. It will be useful to consider the intermediate expression:

$$
U(N_n)\overline{\Theta}(h) = \mathbb{E}\left[U(N_n^\theta)(h) \right]
$$

where the expectation $\mathbb{E}[\cdot]$ is with respect to the Brownian motions. Since $\overline{\Theta} : \mathcal{B}^* \to \mathcal{P}(\mathcal{C}^*)$ is continuous (by Theorem 6.7) and $U(N_n) \to \nu$ weakly, one may easily show that

$$
U(N_n)\overline{\Theta}(h) \to \nu\overline{\Theta}(h). \tag{9.1}
$$

Therefore, our goal is tantamount to showing that:

Goal (restated): $U(N_n^\theta)(h) - \mathbb{E}\left[U(N_n^\theta)(h) \right] \to 0$ almost surely, (9.2)

with respect to the law of the Brownian motions, where we recall $\|h\|_{BL} \leq 1$. The proof idea for (9.2) is to use the fact that $U(N_n^\theta)(h)$ is a function of independent Brownian motions. If we could control the effect of replacing one of the Brownian motions, then we can prove concentration by Azuma's inequality. To make this work, we will need to consider a truncated process for r fixed given by the networks:

$$
(N_n(v), v)_r^\theta
$$

replacing the initial conditions for the interacting diffusions $\boldsymbol{\theta}^{(N_n(v),v)_r}(\cdot)$. With this motivation we define the r-neighborhood empirical measure

$$U^{(r)}(N_n^\theta) = \frac{1}{n} \sum_{v=1}^n \delta_{(N_n(v),v)_r^\theta} \tag{9.3}$$

After considering these networks for r fixed we will need to return to our original network. In what follows we will need to use Azuma's inequality.

Theorem 9.1 (Azuma's inequality, Theorem 6.2, [6])**.** *Let $\mathcal{X}$ be a measurable space. Assume that the function $f : \mathcal{X}^n \to \mathbb{R}$ satisfies the bounded differences assumption: for each $1 \le i \le n$,*

$$\sup_{\substack{x_1,\cdots,x_n \in \mathcal{X} \\ x_i' \in \mathcal{X}}} |f(x_1,\cdots,x_{i-1},x_i,x_{i+1},\cdots,x_n) - f(x_1,\cdots,x_{i-1},x_i',x_{i+1},\cdots,x_n)| \le c_i$$

and define $\widetilde{\nu} = \frac{1}{4}\sum_{i=1}^n c_i^2$. Let $Z = f(X_1,...,X_n)$ where the $X_i \in \mathcal{X}$ are independent. Then

$$\mathbb{P}\left(Z - \mathbb{E}\left[Z\right] > t\right) \le e^{-t^2/(2\widetilde{\nu})}.$$

Lemma 9.2. *For any $r \in \mathbb{N}$ fixed, the following holds almost surely, with respect to the law of the Brownian motions,*

$$\limsup_{n\to\infty} \left| U^{(r)}(N_n^\theta)(h) - \mathbb{E}\left[U^{(r)}(N_n^\theta)(h) \right] \right| = 0.$$

Proof. Recall that the assumptions on our networks imply:

$$\max_{v \in [n]} d_v^{N_n} = n^{\epsilon_n} \text{ with } \lim_{n\to\infty} \epsilon_n = 0. \tag{9.4}$$

From (9.3) we have that

$$U^{(r)}(N_n^\theta)(h) = \frac{1}{n} \sum_{v=1}^n h[(N_n(v),v)_r^\theta],$$

and the randomness of the rooted network $(N_n(v),v)_r^\theta$ is the random vector

$$(\theta_u^{(N_n(v),v)_r}(\cdot))_{u \in (N_n(v),v)_r}, \text{ for } v \in [n].$$

The strong solution assumption implies that for each $v \in [n]$ there exists a measurable function

$$g_v : C([0,T];\mathbb{R})^{|(N_n(v),v)_r|} \to \mathcal{C}^* \text{ such that}$$
$$g_v\left((B_z(\cdot))_{z \in (N_n(v),v)_r}\right) = (N_n(v),v)_r^\theta.$$

Therefore,

$$\begin{aligned} U^{(r)}(N_n^\theta)(h) &= \frac{1}{n} \sum_{v=1}^n h[(N_n(v),v)_r^\theta] \\ &= \frac{1}{n} \sum_{v=1}^n h\left(g_v\left((B_z(\cdot))_{z \in (N_n(v),v)_r}\right)\right) =: f(B_1(\cdot),\cdots,B_n(\cdot)). \end{aligned}$$

Now fix a vertex $w \in [n]$ and suppose we change the function $B_w(\cdot)$. Then whenever $w \notin (N_n(v),v)_r$ the function $g_v\left((B_z(\cdot))_{z \in (N_n(v),v)_r}\right)$ is unchanged. That is, the only

functions g_v that are changed are those with $v \in (N_n(w), w)_r$. Using that $\|h\|_\infty \leq 1$ we can conclude that f satisfy the bounded difference inequality (Theorem 9.1) with

$$c_w = \frac{|(N_n(w), w)_r|}{n}.$$

From (9.4) the size of $(N_n(w), w)_r$ is bounded by $n^{(r+1)\epsilon_n}$ (to see this write the ball as the union of spheres). In this way,

$$\widetilde{\nu} := \frac{1}{4} \sum_{w=1}^{n} c_w^2 \leq \frac{1}{4} \sum_{w=1}^{n} \left(\frac{n^{(r+1)\epsilon_n}}{n}\right)^2 = \frac{1}{4} \frac{n^{2(r+1)\epsilon_n}}{n}.$$

Theorem 9.1 applied twice implies

$$\mathbb{P}\left(\left|U^{(r)}(N_n^\theta)(h) - \mathbb{E}\left[U^{(r)}(N_n^\theta)(h)\right]\right| > t\right) \leq 2\exp\left(-\frac{2t^2 n}{n^{2(r+1)\epsilon_n}}\right).$$

This bound is summable in n for any $r \geq 1$ and $t > 0$ fixed because $\epsilon_n \to 0$. Therefore, Borel-Cantelli Lemma finishes the proof. $\qquad\square$

To continue, we must compare the truncated network in this Lemma with the original N_n. We first bound:

$$U(N_n^\theta)(h) - U^{(r)}(N_n^\theta)(h),$$

which is an average of differences:

$$h([N_n^\theta(v), v]) - h([N_n(v), v]_r^\theta)$$

over $v \in [n]$. Since $\|h\|_{BL} \leq 1$, we have:

$$|h([N_n^\theta(v), v]) - h([N_n(v), v]_r^\theta)| \leq d_{\mathcal{C}^*}\left([N_n^\theta(v), v], [N_n(v), v]_r^\theta\right) \wedge 2.$$

Let r_0 be the constant of Lemma 7.1. From this Lemma, we have the bound:

$$d_{\mathcal{C}^*}\left([N_n^\theta(v), v], [N_n(v), v]_r^\theta\right) \leq \frac{1}{1+s} + C|\partial(N_n(v), v)_r| \exp(-(r-s)\log(r-s)) \qquad (9.5)$$

whenever $r \geq s + r_0$. Averaging over $v \in [n]$, we conclude

$$|U^{(r)}(N_n^\theta)(h) - U(N_n^\theta)(h)| \leq \frac{1}{n} \sum_{v \in [n]} \left(\frac{1}{1+s} + C|\partial(N_n, v)_r| e^{-(r-s)\log(r-s)}\right) \wedge 2 \qquad (9.6)$$

for all $r \geq s + r_0$.

We know N_n converges in the local weak sense to ν. The measure ν is supported on nice networks. That is, we have that $\nu(\cup_{a=1}^{+\infty} \mathcal{B}_a^*) = 1$ with

$$\mathcal{B}_a^* := \{(N, o) \in \mathcal{B}^* : |\partial(N, o)_r| \leq ae^{ar}, \forall r \geq 1\}. \qquad (9.7)$$

Since the $\mathcal{B}_a^*$ are increasing, we have $\nu(\mathcal{B}_a^*) \to 1$ as $a \to +\infty$. Also, by local weak convergence:

$$\frac{1}{n} \sum_{u \in [n]} \mathbb{I}_{\{|\partial(N_n, u)_r| > ae^{ar}\}} \to \nu(\{([N, o] : |\partial(N, o)_r| > ae^{ar}\}) = 1 - \nu(\mathcal{B}_a^*).$$

We bound the terms in the RHS of (9.6) according to whether $|\partial(N_n, v)_r|$ is either bounded by ae^{ar} or not; in the latter case the terms are simply bounded by 2. We deduce that,

when $n \to +\infty$:

$$
\begin{aligned}
|U^{(r)}(N_n^\theta)(h) - U(N_n^\theta)(h)| \quad &\leq \quad \frac{1}{1+s} + C a e^{ar - (r-s)\log(r-s)} \tag{9.8} \\
&\quad + \frac{2}{n} \sum_{u \in [n]} \mathbb{I}_{\{|\partial(N_n, u)_r| > ae^{ar}\}} \\
&\rightarrow \quad \frac{1}{1+s} + C a e^{ar - (r-s)\log(r-s)} + 2(1 - \nu(\mathcal{B}_a^*)).
\end{aligned}
$$

The same bound holds for the difference of the expectations

$$
|\mathbb{E}\left[U^{(r)}(N_n^\theta)(h) - U(N_n^\theta)(h)\right]|.
$$

Now, for each fixed r, s, a as above:

$$
\begin{aligned}
|U(N_n^\theta)(h) - \mathbb{E}\left[U(N_n^\theta)(h)\right]| \quad &\leq \quad |U(N_n^\theta)(h) - U^{(r)}(N_n^\theta)(h)| \\
&\quad + |\mathbb{E}\left[U^{(r)}(N_n^\theta)(h) - U(N_n^\theta)(h)\right]| \\
&\quad + |U^{(r)}(N_n^\theta)(h) - \mathbb{E}\left[U^{(r)}(N_n^\theta)(h)\right]|.
\end{aligned}
$$

When $n \to +\infty$, the first two terms in the RHS may be controlled via (9.8). The last term goes to 0 by Lemma 9.2. We obtain:

$$
\limsup_{n \to \infty} |U(N_n^\theta)(h) - \mathbb{E}\left[U(N_n^\theta)(h)\right]| \leq \frac{2}{1+s} + 2C a e^{ar - (r-s)\log(r-s)} + 2(1 - \nu(\mathcal{B}_a^*)).
$$

We may now let $r \to +\infty$, then $s \to +\infty$ and $a \to +\infty$ (in this order) to obtain:

$$
\limsup_{n \to \infty} |U(N_n^\theta)(h) - \mathbb{E}\left[U(N_n^\theta)(h)\right]| \leq 0.
$$

This implies our goal (9.2) and finishes the proof of the hydrodynamic limit.

10 Numerical simulations for the stochastic Kuramoto model on Galton-Watson trees

10.1 Introduction

In this section, we explain how we performed the numerical simulations for the stochastic Kuramoto model on Galton-Watson (GW) trees. The equations are given by

$$
d\theta_i = \left(\omega_i + K \sum_{j=1}^{N} a_{ij} \sin(\theta_j - \theta_i) \right) dt + \varepsilon dW_t^i.
$$

For this system, $\theta_j(t)$ and ω_j represent the angular phase and natural frequency of the oscillator indexed by $j \in \{1, 2, \ldots, N\}$. The parameter $K \in \mathbb{R}_+$ represents the coupling strength between nodes, and $a_{ij} = 1$ if nodes i and j are connected or $a_{ij} = 0$ otherwise. We assume that $\{W_t^j\}_{t \geq 0}$ are independent brownian motions for each node j, while the noise intensity is given by $\varepsilon > 0$. In our simulations, we chose two different models for generating the GW trees:

1. **Binomial model:** The offspring is a binomial random variable with distribution $Bin(n, p)$.
2. **D-Regular model:** The root node has $\mathcal{C}$ children, while the other ones have exactly $\mathcal{C} - 1$ children.

In what follows, we fixed $n = 3$ and $p = \frac{2}{3}$ for the Binomial model (mean offspring equals to 2) and $\mathcal{C} = 3$ for the D-Regular model. We also considered the time interval $t = [0, 100]$ (arbitrary units) with step $\Delta t = 0.01$, and noise intensity $\varepsilon = 0.05$.

The synchronization level

We defined the synchronization level $Sync = Sync(h, K)$ between the root and those nodes at a distance h, given the coupling strength K. If N denotes the total number of nodes and $j = 1$ is defined as the root index, we define the set

$$\mathcal{D}(h) = \{j \in \{2, ..., N\}| \text{ node of index } j \text{ is at distance } h \text{ from the root}\}$$

and the order parameter

$$r_h(t) = \left| \frac{e^{i\theta_1(t)} + \sum_{j \in \mathcal{D}(h)} e^{i\theta_j(t)}}{1 + \#\mathcal{D}(h)} \right|$$

where $\#$ denotes set cardinality.

Our variable $Sync$ is then defined as a time average of the last 5% values assumed by $r_h(t)$. More precisely, if we have a total of t_n, the set of last 5% time indexes is given by $\mathcal{J} = \{\lfloor 0.95t_n \rfloor, ..., t_n\}$. Therefore, we define

$$Sync := \frac{\sum_{j \in \mathcal{J}} r_h(t_j)}{\#\mathcal{J}}$$

as our synchronization level parameter.

10.2 Numerical simulations

We present the details of our numerical simulations, considering the Binomial and D-Regular models. In both cases, our goal is to calculate an average synchronization parameter $\langle Sync \rangle$ between the root and those nodes at different distance values h, for distinct coupling strength (K) values. For all simulations, we considered GW trees with 13 generations ($h \in \{1, 2, \cdots, 13\}$), and $K \in [0, 10]$ with steps $\Delta K = 0.2$. We performed all numerical solutions with the classic Euler-Maruyama scheme for stochastic differential equations. In what follows, we describe the step-by-step algorithms for each GW model.

- **Binomial model.** We produce a total of 100 simulations. In each of them, we generate a GW tree with N nodes and define the natural frequencies ω_i, as well as the initial conditions $\theta_i(0)$ ($i \in 1, 2, \ldots, N$). We then simulate the stochastic Kuramoto model with the Euler-Maruyama scheme, computing $Sync(h, K)$ for the chosen h and K values (see above). Finally, we average the synchronization levels across the 10 simulations, obtaining $\langle Sync \rangle(h, K)$.

- **D-Regular model.** In this case, the GW is not random. Then we produce a total of 100 simulations, only re-sampling the initial conditions $\theta_i(0)$ and natural frequencies ω_i ($i \in 1, 2, \ldots, N$). For each simulation, we compute the synchronization level ($Sync(h, K)$) for the selected K and h values. We then compute the average synchronization level $\langle Sync \rangle(h, K)$ across the 100 simulations.

10.3 Results

We present our results in Fig. 1 **a**, **b**. We compared two possible cases for the initial phases: $\theta_i(0) = 0$ or $\theta_i(0)$ uniformly distributed on $[0, 2\pi]$, for $i \in \{1, 2, \ldots, N\}$. For the natural frequencies, we chose $\omega_i = 1$, or ω_i normally distributed with mean and variance equal to 1, or ω_i uniformly distributed on $[0, 2\pi]$, for $i \in \{1, 2, \ldots, N\}$.

For both Binomial and D-Regular GW models, the average level of synchronization $\langle Sync \rangle$ significantly decreased, as the distance from the root h increased, for all considered choices of initial conditions $\theta_i(0)$ and natural frequencies ω_i. The parameter regions of high synchronization (red-colored in the heatmaps, which correspond to $\langle sync \rangle > 0.8$)

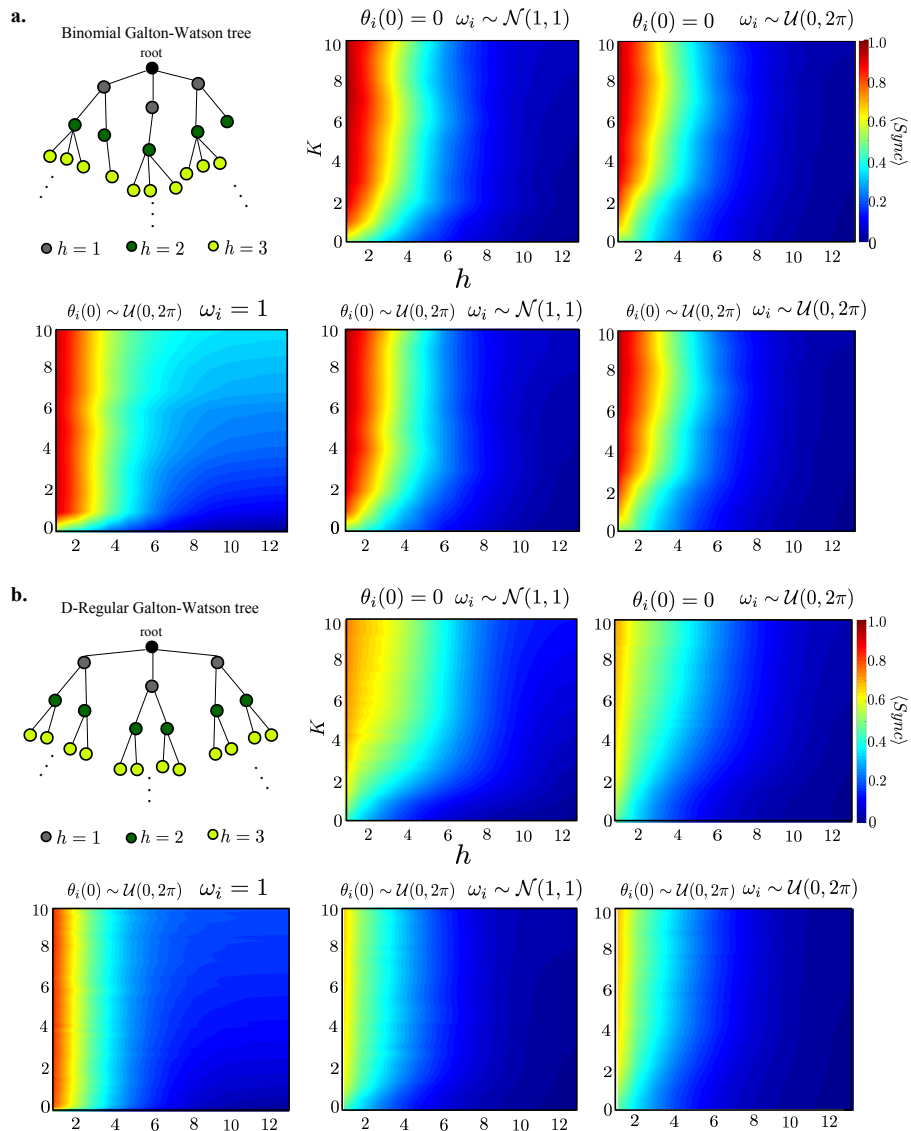Figure 1: **Stochastic Kuramoto Dynamics in GW trees: average synchronization levels between the root and different nodes. a.** In the Binomial model, the offspring was given by a binomial random variable with distribution $Bin(3, \frac{2}{3})$. **b.** In the D-Regular model, the root has 3 children, while the other nodes have only 2 children. In the top left of Panels **a.** and **b.**, we exhibit schematic illustrations of both GW tree models, with nodes at different distances from the root. The parameter K represents the coupling strength between any two connected nodes (see text for details). We assumed two distinct distributions for the initial phases $\theta_i(0)$ (all equal or uniformly distributed), while the natural frequencies ω_i were assumed equal to 1, or normally distributed with mean and variance equal to 1, or uniformly distributed on $[0, 2\pi]$. For different K values, we estimated the loss of synchronization between the root vertex and those at different distances h, as h increased from 1 to 13. We observed clear transition patterns between high and low synchronization regions in the $h \times K$ plane.

were only found for low h values ($h \leq 4$ in the binomial model and $h \leq 2$ in the D-regular model). It is important to notice that the synchronization levels did not substantially increased even for higher values of the coupling strength K, which in our study was allowed to range from 0 to 10. This result indicates that desynchronization may be a predominant phenomenon in GW trees with the chosen average offspring values.

From the GW model perspective, a simple visual inspection indicates that the Binomial model promoted a higher synchronization for lower h values than the D-Regular. For each one of the five combinations of the initial condition and natural frequencies, the comparison between models shows a significant higher synchronization in the binomial model. Further analysis would be required for us to understand the causes of such effect. However, we hypothesize that the variability in the offspring number that is present in the binomial model could be a major factor to explain this increase in synchronization.

A Weak local convergence and nice networks

The goal of this section is to prove Proposition 5.7, Theorem 5.8 and show that we have examples that satisfy the assumptions of Theorem 6.8 (cf. Remark 6.9).

A.1 Proof of Proposition 5.7

Proposition 5.7 is a simpler version of Theorem 6.7.

Start with a sequence of rooted graphs $\{(G_k, o_k)\}_{k \in \mathbb{N}}$ that converges in the local topology to $(G, o) \in \mathcal{G}^*$. We want to show that the respective laws of the random rooted networks, (N^{G_k}, o_k) also converges to the law of (N^G, o), that is, $\mathcal{M}(G_k, o_k) \to \mathcal{M}(G, o)$ in the BL-distance.

From the convergence $(G_k, o_k) \to (G, o)$ in $\mathcal{G}^*$ we know that for any $r \in \mathbb{N}$ fixed $(G_k, o_k)_r = (G, o)_r$ for k sufficiently large.

When we construct the random networks N^G and N^{G_k} we can couple the marks to be equal in $(G_k, o_k)_r = (G, o)_r$. In particular, $(N^{G_k}, o_k)_r = (N^G, o)_r$ almost surely.

Therefore, by definition and almost surely

$$\mathrm{d}_{\mathcal{N}^*}\left([N^G, o], [N^{G_k}, o_k]\right) \leq \frac{1}{1+r}$$

for k big enough. This bound immediately translates to a bound for their laws and we finish the proof.

A.2 Proof of Theorem 5.8

Theorem 5.8 is a simpler version of Theorem 6.8.

We have a sequence of graphs $\{G_n\}_{n \in \mathbb{N}}$ that converges locally weakly to the measure $\rho \in \mathcal{G}^*$. We want to show that defining the random networks $N_n := N^{G_n}$ by adding the i.i.d. marks, we have that $U(N_n) \to \rho\mathcal{M}$ almost surely in $\mathcal{P}(\mathcal{N}^*)$. We are assuming that $\max_{v \in [n]} d_v^{G_n} \leq n^{\epsilon_n}$ with $\epsilon_n \to 0$ as $n \to \infty$.

Let $\mathbb{E}[\,\cdot\,]$ be the expectation in the space where the marks of the sequence of the networks $(N_n)_{n \in \mathbb{N}}$ can be defined. We can easily check that

$$\mathbb{E}[U(N_n)] = U(G_n)\mathcal{M}.$$

In particular, from the continuity of $\mathcal{M}$ it is simple to deduce that $\mathbb{E}[U(N_n)] \to \rho\mathcal{M}$.

So we just need to show that almost surely

$$\lim_{n \to \infty} |U(N_n)(h) - \mathbb{E}[U(N_n)(h)]| = 0$$

for any borel measurable function $h : \mathcal{N}^* \to \mathbb{R}$ with $\|h\|_{BL} \leq 1$.

The idea, as in Theorem 6.8, is to concentrate $U(N_n)_r$ around its mean, where

$$U(N_n)_r = \frac{1}{n} \sum_{v=1}^{n} \delta_{[N_n(v),v]_r}. \tag{A.1}$$

Throughout the remainder of this section we have a bounded measurable test function $h : \mathcal{N}^* \to \mathbb{R}$ with $\|h\|_{BL} \leq 1$.

The same arguments in Section A.1 says that

$$|U(N_n)(h) - U(N_n)_r(h)| \leq \frac{1}{1+r}$$

that also holds in mean:

$$|\mathbb{E}[U(N_n)(h)] - \mathbb{E}[U(N_n)_r(h)]| \leq \frac{1}{1+r}.$$

Now we will use Azuma's inequality (cf. Theorem 9.1) to show that $U(N_n)_r(h)$ is concentrated around its mean. From (A.1)

$$U(N_n)_r(h) = \frac{1}{n} \sum_{v=1}^{n} h([N_n(v),v]_r)$$

that is a function of the independent variables $(\omega_v, \theta_v(0))_{v \in [n]}$ and $(\mu_e)_{e \in E_n}$, where E_n is the edge set of G_n.

If we change one of these marks on the vertex w, the only networks that are changed are those $[N_n(v),v]_r$ with $w \in (G_n(v),v)_r$, that is, $v \in (G_n(w),w)_r$. Therefore, $U(N_n)_r(h)$ changes by at most the size of $(G_n(w),w)_r$ over n. By assumption $|(G_n(w),w)_r| \leq n^{(r+1)\epsilon_n}$. Observe that $U(N_n)(h)$ is a function of $2n + E_n$ independent random variables.

From Azuma inequality we have that

$$\mathbb{P}\left(|U(N_n)_r(h) - \mathbb{E}[U(N_n)_r(h)]| > t\right) \leq 2 \exp\left(-\frac{2t^2 n^2}{(2n + E_n)n^{2(r+1)\epsilon_n}}\right).$$

To see that this bound is summable we use that the number of edges is one half of the sum of degrees to bound:

$$E_n \leq \frac{1}{2} n \max_{v \in [n]} d_v^{G_n} \leq \frac{1}{2} n n^{\epsilon_n}.$$

Using the triangle inequality several times and the previous bounds we see that

$$\lim_{n \to \infty} |U(N_n)(h) - \mathbb{E}[U(N_n)(h)]| = 0$$

for any h with $\|h\|_{BL} \leq 1$.

A.3 Good examples

In this section we want to see that we have examples that satisfy the assumptions in Theorem 6.8.

From Theorem 5.8, the Assumption 1 in the Theorem 6.8 is satisfied almost surely for any of the graphs in Example 5.2, 5.3 and 5.4 when we add i.i.d. marks.

The Assumption 3 of Theorem 6.8 is satisfied trivially for graphs with bounded degree (Example 5.2). We impose in Example 5.4 that uniformly graphs with a given degree sequence also satisfy $\max_{v \in [n]} d_v^{G_n} \leq n^{\epsilon_n}$ with $\epsilon_n \to 0$ as $n \to \infty$.

Now we will check that Erdös-Rényi graphs (Example 5.3) also satisfy Assumption 3 of Theorem 6.8, almost surely.

By the union bound, and Chernoff bound (Theorem 2.3.1, [29])

$$\mathbb{P}\left(\max_{v\in[n]} d_v^{G_n} > n^{\epsilon_n}\right) \leq n\mathbb{P}\left(d_1^{G_n} > n^{\epsilon_n}\right) \leq \exp\left(-n^{\epsilon_n}\log n^{\epsilon_n}\right)$$

for n big enough. Since this bound is summable (we can impose $n^{\epsilon_n} \to \infty$ as $n \to \infty$), a sequence of Erdös-Rényi graphs satisfy the assumption almost surely.

It remains to see that in the Examples 5.2, 5.3 and 5.4, the local weak limit is supported in nice graphs, that is, the local weak limit of theses sequences are supported in rooted graphs that satisfy

$$|\partial(G,o)_r| \leq ae^{ar}, \text{ for some } a > 0. \tag{A.2}$$

It is clear that (A.2) is achieved when the local weak limit measure is supported in graphs with uniform bounded degree. So we are done with Example 5.2.

Now we will check (A.2) for the GW tree. This will show that we are done with Examples 5.3 and 5.4 because a Unimodular GW tree is equal to a GW tree after the first generation.

Given a probability distribution P on $\mathbb{N}$ with mean $\mu \in (0,+\infty)$ consider $GW(P)$ the GW distribution on the set of rooted trees. It means that in each generation each individual has i children with probability $P(i)$ independent of the other individuals. Let Z_n be the number of individuals at generation n, $Z_0 = 1$. We know that $Z_n/(\mu)^n$ is a martingale. In particular $\mathbb{E}[Z_n] = \mu^n$. But we also have that $|\partial(G,o)_n| = Z_n$ if (G,o) has distribution $GW(P)$.

Therefore,

$$\mathbb{P}\left(|\partial(G,o)_r| \geq (2\mu)^r\right) \leq \frac{\mathbb{E}\left[|\partial(G,o)_r|\right]}{(2\mu)^r} = \frac{1}{2^r},$$

which is summable in r. Therefore, $|\partial(G,o)_r| \leq e^{(\log 2\mu)r}$ for r big enough, almost surely. Choosing a big enough we see that $|\partial(G,o)_r| \leq ae^{ar}$ for any $r \geq 1$ and for almost all realizations of (G,o).

B Appendix - Continuity for finite graphs

The goal of this section is to justify (8.6):

$$d_{BL}(\overline{\Theta}[N,o]_r, \overline{\Theta}[N_n,o_n]_r) \to 0.$$

We think that this can be derived from the standard theory of Ordinary Differential Equations. For completeness we provide a prove in this section. Our notation follows that one in Section 8.1.

We will use the following Lemma.

Lemma B.1 (Proof in Section B.1)**.** *Consider a graph G and two finite networks $N_i = (G, \boldsymbol{\mu}^{(i)}, \boldsymbol{\omega}^{(i)}, \boldsymbol{\theta}^{(i)}(0)) \in \mathcal{N}$, $i = 1, 2$. Suppose that all marks are close by ϵ:*

$$\sup_{vu\in E_G} |\mu_{vu}^{(2)} - \mu_{vu}^{(1)}| \leq \epsilon, \quad \sup_{v\in V_G} |\omega_v^{(2)} - \omega_v^{(1)}| \leq \epsilon, \text{ and } \sup_{v\in V_G} |\theta_v(0)^{(2)} - \theta_v(0)^{(1)}| \leq \epsilon.$$

Then there exists a constant $\overline{C}$ depending on $\|\phi\|_{BL}$, $\|\psi\|_{Lip}$, T, and N such that almost surely with respect to the law of the Brownian motions

$$\sup_{v\in G} \sup_{t\leq T} \left(|\theta_v^{N_1}(t) - \theta_v^{N_2}(t)|\right) \leq \overline{C}e(G)\epsilon.$$

From the convergence of $[N_n, o_n]$ to $[N,o]$ we know that for any r and for n large enough

$$[N_n, o_n]_r = [N,o]_r.$$

So we can apply Lemma B.1 with $N_1 = (N, o)_r$, $N_2 = (N_n, o_n)_r$, and

$$\epsilon = \mathrm{d}_{\mathcal{N}^*}\left([N, o]_r, [N_n, o_n]_r\right)$$

to obtain the bound

$$\sup_{v \in (N,o)_r} \sup_{t \leq T} |\theta_v^{(N,o)_r}(t) - \theta_v^{(N_n,o_n)_r}(t)| \leq \overline{C}e([N, o]_r)\mathrm{d}_{\mathcal{N}^*}\left([N, o]_r, [N_n, o_n]_r\right). \tag{B.1}$$

In that way, we use that the networks coincide at any radius and the bound in (B.1) to conclude that almost surely (cf. (5.1))

$$\mathrm{d}_{\mathcal{C}^*}\left([N, o]_r^\theta, [N_n, o_n]_r^\theta\right) \leq \overline{C}e([N, o]_r)\mathrm{d}_{\mathcal{N}^*}\left([N, o]_r, [N_n, o_n]_r\right)$$

and the RHS of the last bound goes to zero as $n \to \infty$. This is enough to finish.

B.1 Proof of Lemma B.1

In this section, we prove Lemma B.1. The prove follows similar ideas of the proof of Lemma 7.1.

For ease of notation we adopt the following conventions.

Definition B.2.   *1. For the objects related to N_2 we omit the superscripts and just write $\boldsymbol{\mu}$, $\boldsymbol{\omega}$, $\boldsymbol{\theta}(0)$ and $\boldsymbol{\theta}$.*

   2. For the objects related to N_1 we omit the superscript and write a over-line writing $\overline{\boldsymbol{\mu}}$, $\overline{\boldsymbol{\omega}}$, $\overline{\boldsymbol{\theta}}(0)$ and $\overline{\boldsymbol{\theta}}$.

   3. We also write $P_{uv} = \mu_{uv}/\mu_v$ and $\overline{P}_{uv} = \overline{\mu}_{uv}/\overline{\mu}_v$.

   4. We write I for the identity matrix.

   5. G has vertex set V and edge set E.

From Definition 6.1 and for $v \in V$, we have that

$$\mathrm{d}\theta_v(t) = \sum_{u \in V} P_{uv}(\phi(\theta_u(t), \theta_v(t); \omega_v, \omega_u)\mathrm{d}t + \psi(\theta_v(t); \omega_v)\mathrm{d}t + \mathrm{d}B_v(t) \tag{B.2}$$

and the analogous formula holds for $\overline{\theta}_v(\cdot)$ using the objects of N_1 but using the same Brownian motion.

We suppose that the Brownian motions are coupled to be equal in each system.

Proposition B.3. *With our conventions, there exists a constant C depending on $\|\phi\|_{BL}$, $\|\psi\|_{Lip}$, T, and $1/\mu_*$ such that*

$$|\theta_v(t) - \overline{\theta}_v(t)| \leq C\left(\epsilon + \int_0^t \sum_{u \in V}(\overline{P} + I)_{uv}|\theta_u(s) - \overline{\theta}_u(s)|\mathrm{d}s\right)$$

for all $v \in V$.

Proof. From (B.2),

$$\theta_v(t) - \overline{\theta}_v(t) = \theta_v(0) - \overline{\theta}_v(0) + \int_0^t \Delta_1(v, s) + \Delta_2(v, s) + \Delta_3(v, s)\mathrm{d}s$$

where

$$\begin{aligned}
\Delta_1(v, s) &:= \sum_{u \in V}(P_{uv} - \overline{P}_{uv})\phi(\theta_u(s), \theta_v(s); \omega_u, \omega_v), \\
\Delta_2(v, s) &:= \sum_{u \in V}\overline{P}_{uv}(\phi(\theta_u(s), \theta_v(s); \omega_u, \omega_v) - \phi(\overline{\theta}_u(s), \overline{\theta}_v(s); \overline{\omega}_u, \overline{\omega}_v)), \\
\Delta_3(v, s) &:= \psi(\theta_v(s); \omega_v) - \psi(\overline{\theta}_v(s); \overline{\omega}_v)
\end{aligned}$$

and the following bounds hold

$$
\begin{aligned}
|\theta_v(0) - \overline{\theta}_v(0)| &\leq \epsilon \\
|\Delta_1(v, s)| &\leq \|\phi\|_\infty \sum_{u \in V} |P_{uv} - \overline{P}_{uv}|, \\
|\Delta_2(v, s)| &\leq \sum_{u=1}^n \overline{P}_{uv} \|\phi\|_{Lip} \left(|\theta_v(s) - \overline{\theta}_v(s)| + |\theta_u(s) - \overline{\theta}_u(s)| + 2\epsilon\right), \text{ and} \\
|\Delta_3(v, s)| &\leq \|\psi\|_{Lip} \left(|\theta_v(s) - \overline{\theta}_v(s)| + \epsilon\right).
\end{aligned}
$$

Now we estimate $S_v := \sum_{u \in V} |P_{uv} - \overline{P}_{uv}|$. Remember that $\sup_{vu \in E} |\mu_{vu} - \overline{\mu}_{vu}| \leq \epsilon$. It is clear that $S_v = 0$ if v is an isolated vertex. If v is not isolated, then

$$
\begin{aligned}
S_v &= \sum_{u \in V} \left| \frac{\mu_{uv}}{\mu_v} - \frac{\overline{\mu}_{uv}}{\overline{\mu}_v} \right| \\
&\leq \sum_{u \in V} \frac{1}{\mu_v} |\mu_{uv} - \overline{\mu}_{uv}| + \overline{\mu}_{uv} \left| \frac{\overline{\mu}_v - \mu_v}{\overline{\mu}_v \mu_v} \right| \leq \frac{\epsilon d_v}{\mu_v} + \frac{|\overline{\mu}_v - \mu_v|}{\mu_v} \leq \frac{2\epsilon d_v}{\mu_v}.
\end{aligned}
$$

The assumption that the network is nice says that $\mu_* \leq \mu_{uv}$ whenever $uv \in E$. That is, $d_v \mu_* \leq \mu_v$.

Combining these bounds we obtain the result. $\qquad\square$

We now use Corollary C.2 with

1. $\mathbf{u}(t) := (|\theta_v(t) - \overline{\theta}_v(t)|)_{v \in V}$.

2. $\mathbf{a}(t) := (C\epsilon)_{v \in V}$, and each entry of this vector is non-negative.

3. $M(t) = M := C(\overline{P} + I)$, and this matrix does not depend on time t, it is entrywise non-negative, and it is finite dimensional since G is finite.

Therefore, for any $v \in V$

$$
|\theta_v(t) - \overline{\theta}_v(t)| \leq C\epsilon \exp\left(tC(\overline{P} + I)\right)_v
$$

To relate this bound with the Random Walk in $(G, \overline{\boldsymbol{\mu}})$ observe that

$$
\exp\left(tC(\overline{P} + I)\right) = e^{Ct} \exp\left(tC\overline{P}\right) = e^{2Ct} e^{-Ct} \exp\left(tC\overline{P}\right).
$$

In the context of [3, Section 5.1] we have that

$$
q_s(v, w) := \frac{1}{\overline{\mu}_w} e^{-s} \exp\left(s\overline{P}\right)_{vw} \tag{B.3}
$$

is the continuous time Heat Kernel for the Simple Random Walk in $(G, \overline{\boldsymbol{\mu}})$.

Therefore, since $q_s(v, w) \leq 1$, for all $v, w \in V$, then $e^{-s} \exp\left(s\overline{P}\right)_{vw} \leq \overline{\mu}_w \leq \mu^* d_w$ implies that

$$
|\theta_v(t) - \overline{\theta}_v(t)| \leq C\mu^* \epsilon e^{2CT} \sum_{w \in G} d_w = 2C\mu^* e^{2CT} e(G)\epsilon.
$$

Notice that this bound is uniform in $t \in [0, T]$. This is enough to finish the proof.

C Linear Gronwall's inequality

Proposition C.1 (Corollary 2 of [10]). *Let the vector $\mathbf{a}(t) \in \mathbb{R}^n$ and the non-negative (entrywise) $n \times n$ matrices $O(t)$, $M(t)$ be continuous functions of the single scalar variable t for $t^0 \leq t$. Assume that $M(t)O(t)$ and $\int_{t^0}^t M(s)O(s)ds$ commute for $t^0 \leq t$. If*

$$\mathbf{u}(t) \leq \mathbf{a}(t) + O(t)\int_{t^0}^t M(s)\mathbf{u}(s)ds, \ \ t^0 \leq t$$

then

$$\mathbf{u}(t) \leq \mathbf{a}(t) + O(t)\int_{t^0}^t \exp\left(\int_s^t M(r)O(r)dr\right) M(s)\mathbf{a}(s)ds, \ \ t^0 \leq t.$$

Corollary C.2. *Assume the hypothesis in Proposition C.1 and additionally suppose that $O(t) = Id$ and $\mathbf{a}(t)$ is entrywise non-decreasing in each entry. In this case*

$$\mathbf{u}(t) \leq \mathbf{a}(t) + \int_{t^0}^t M(s)\mathbf{u}(s)ds, \ \ t^0 \leq t \tag{C.1}$$

implies

$$\mathbf{u}(t) \leq \exp\left(\int_{t^0}^t M(s)ds\right)\mathbf{a}(t), \ \ t^0 \leq t. \tag{C.2}$$

Proof. From Proposition C.1, and the fact that the exponential of a non-negative matrix is non-negative (every entry of each power is non-negative) we can compute

$$\mathbf{u}(t) \leq \mathbf{a}(t) + \int_{t^0}^t \exp\left(\int_s^t M(r)dr\right) M(s)\mathbf{a}(s)ds$$

$$(\mathbf{a}(s) \leq \mathbf{a}(t) \text{ whenever } s \leq t) \leq \mathbf{a}(t) + \left(\int_{t^0}^t \exp\left(\int_s^t M(r)dr\right) M(s)ds\right)\mathbf{a}(t)$$

$$= \mathbf{a}(t) + \left(\int_{t^0}^t \frac{d}{ds}\exp\left(\int_s^t M(r)dr\right)ds\right)\mathbf{a}(t)$$

$$= \mathbf{a}(t) - \mathbf{a}(t) + \exp\left(\int_{t^0}^t M(s)ds\right)\mathbf{a}(t)$$

$$= \exp\left(\int_{t^0}^t M(s)ds\right)\mathbf{a}(t). \qquad \square$$

D Propagation of chaos

The goal of this section is to prove Corollary 6.12. As above, we will assume that the vertex set of N_n is $[n]$.

Let $f_1, \ldots, f_k : \mathcal{C}^* \to \mathbb{R}$ be bounded Lipschitz functions with $\|f_i\|_{BL} \leq 1$. Our goal is to show that:

$$\mathbb{E}\left[\prod_{i=1}^k U(N_n^\theta)(f_i)\right] \to \prod_{i=1}^k \nu\overline{\Theta}(f_i)$$

as $n \to +\infty$. This is true for $k = 1$, as shown in (9.1). Therefore, it suffices to show that:

$$\textbf{Goal: } \mathbb{E}\left[\prod_{i=1}^k U(N_n^\theta)(f_i)\right] - \prod_{i=1}^k \mathbb{E}\left[U(N_n^\theta)(f_i)\right] \to 0 \text{ as } n \to +\infty. \tag{D.1}$$

The strategy to prove this goal is to use a similar argument as in Lemma 9.2 to exploit independence. For these reasons our first task is to compare

$$\mathbb{E}\left[\prod_{i=1}^k U(N_n^\theta)(f_i)\right] \text{ with } \mathbb{E}\left[\prod_{i=1}^k U^{(r)}(N_n^\theta)(f_i)\right]$$

and

$$\prod_{i=1}^{k} \mathbb{E}\left[U(N_n^\theta)(f_i)\right] \text{ with } \prod_{i=1}^{k} \mathbb{E}\left[U^{(r)}(N_n^\theta)(f_i)\right].$$

We write the difference of products as a telescoping sum switching the terms in the product one by one. We may apply the bound in (9.8) for each term recalling that $\|f_i\|_{BL} \le 1$, we deduce that when $r \ge s \ge r_0$ for some constant r_0,

$$\left| \mathbb{E}\left[\prod_{i=1}^{k} U(N_n^\theta)(f_i)\right] - \mathbb{E}\left[\prod_{i=1}^{k} U^{(r)}(N_n^\theta)(f_i)\right] \right|$$

$$\le \frac{k}{1+s} + kCae^{ar-(r-s)\log(r-s)} + \frac{2k}{n} \sum_{u\in[n]} \mathbb{I}_{[|\partial(N_n,u)_r|>ae^{ar}]}$$

$$\to \frac{k}{1+s} + kCae^{ar-(r-s)\log(r-s)} + 2k(1-\nu(\mathcal{B}_a^*))$$

as n goes to infinity and the exactly same bound holds for the other expression of interest (the expectation inside).

As before we can make, $r \to \infty$, $s \to \infty$ and $a \to \infty$ in this order and it is clear that we can restate our goal:

Goal (D.1) restated: $\mathbb{E}\left[\prod_{i=1}^{k} U^{(r)}(N_n^\theta)(f_i)\right] - \prod_{i=1}^{k} \mathbb{E}\left[U^{(r)}(N_n^\theta)(f_i)\right] \to 0$ as $n \to +\infty$.

(D.2)

To achieve this goal, we write:

$$\mathbb{E}\left[\prod_{i=1}^{k} U^{(r)}(N_n^\theta)(f_i)\right] = \frac{1}{n^k} \sum_{u_1,\dots,u_k\in[n]} \mathbb{E}\left[\prod_{i=1}^{k} f_i([N_n(u_i),u_i]_r^\theta)\right];$$

$$\prod_{i=1}^{k} \mathbb{E}\left[U^{(r)}(N_n^\theta)(f_i)\right] = \frac{1}{n^k} \sum_{u_1,\dots,u_k\in[n]} \prod_{i=1}^{k} \mathbb{E}\left[f_i([N_n(u_i),u_i]_r^\theta)\right].$$

We deduce that the difference in (D.2) is the average of:

$$\mathbb{E}\left[\prod_{i=1}^{k} f_i([N_n(u_i),u_i]_r^\theta)\right] - \prod_{i=1}^{k} \mathbb{E}\left[f_i([N_n(u_i),u_i]_r^\theta)\right] \qquad \text{(D.3)}$$

over $(u_1,\dots,u_k) \in [n]^k$.

Now fix a radius $r \ge 1$. We split the set $[n]^k$ into two parts.

$$E_n^k(r) := \{(v_1,\dots,v_k) \in [n]^k \, : \, \mathrm{dist}(v_i,v_j) > r \text{ for all } 1 \le i < i \le k\}$$

and $F_n^k(r) := [n]^k \backslash E_n^k(r)$.

Notice that for $(u_1,\dots,u_k) \in E_n^k(r)$ the trucated networks

$$[N_n(u_i),u_i]_r^\theta \, : \, 1 \le i \le k$$

are independent, as they are deterministic functions of independent Brownian Motions (cf. Lemma 9.2). Therefore:

$$\forall (u_1,\dots,u_k) \in E_n^k(r) \, : \, \mathbb{E}\left[\prod_{i=1}^{k} f_i([N_n(u_i),u_i]_r^\theta)\right] = \prod_{i=1}^{k} \mathbb{E}\left[f_i([N_n(u_i),u_i]_r^\theta)\right].$$

Since the functions f_i are also bounded by 1, we obtain:

$$\left| \mathbb{E}\left[\prod_{i=1}^{k} f_i([N_n(u_i), u_i]_r^\theta) \right] - \prod_{i=1}^{k} \mathbb{E}\left[f_i([N_n(u_i), u_i]_r^\theta) \right] \right| \leq 2\mathbb{I}_{\{(u_1,\ldots,u_k)\in F_n^k(r)\}} \tag{D.4}$$

for all $(u_1,\ldots,u_k)\in [n]^k$.

Recall from (D.3) that the difference we need to consider to achieve our goal (D.2) is an average over $u_1,\ldots,u_k\in [n]$ of the terms in the LHS of (D.4). This implies:

$$\left| \mathbb{E}\left[\prod_{i=1}^{k} U^{(r)}(N_n^\theta)(f_i) \right] - \prod_{i=1}^{k} \mathbb{E}\left[U^{(r)}(N_n^\theta)(f_i) \right] \right| \leq \frac{2|F_n^k(r)|}{n^k}.$$

The RHS of the display may be interpreted that, out of k uniformly random vertices $u_1,\ldots,u_k\in N_n$, at least some pair is at distance $\leq r$ from one another. This is the same as saying that there exists a $u_j \in (N_n, u_i)_{2r}$ for some pair $1\leq i < j \leq k$. So:

$$\frac{|F_n^k(r)|}{n^k} \leq \binom{k}{2}\frac{1}{n}\sum_{u\in[n]} \frac{|(N_n, u)_{2r}|}{n}. \tag{D.5}$$

Claim: For any $k, r \in \mathbb{N}$ fixed the RHS of (D.5) goes to 0 as $n \to +\infty$.

We prove the Claim in the end of this section. We deduce:

$$\limsup_{n\to\infty}\left| \mathbb{E}\left[\prod_{i=1}^{k} U^{(r)}(N_n^{(\theta)})(f_i) \right] - \prod_{i=1}^{k} \mathbb{E}\left[U^{(r)}(N_n^{(\theta)})(f_i) \right] \right| \leq 0$$

for any $k, r \geq 1$ fixed and the result follows.

Now we prove the **Claim**.

Fixed a constant $M > 0$ the sum on the RHS in (D.5) can be split whether $|(N_n, u)_{2r}| \leq M$ or $> M$. In the case $|(N_n, u)_{2r}| > M$ we remember that the bound $|(N_n, u)_{2r}| \leq n$ always holds. We obtain

$$\frac{1}{n}\sum_{u\in[n]} \frac{|(N_n, u)_{2r}|}{n} \leq \frac{M}{n} + \frac{1}{n}\sum_{u\in[n]} \mathbb{I}_{\{|(N_n, u)_{2r}|>M\}}. \tag{D.6}$$

It is simple to check that $(N, o) \in \mathcal{N}^* \mapsto \mathbb{I}_{\{|(N, o)_{2r}|>M\}}$ is a bounded continuous function in the local topology. Therefore, from the local weak convergence of $\{N_n\}_{n\in\mathbb{N}}$ to ν we have that

$$\frac{1}{n}\sum_{u\in[n]} \mathbb{I}_{\{|(N_n, u)_{2r}|>M\}} \to \nu\{(N, o) : |(N, o)_{2r}| > M\}.$$

In this way, taking the limit $n \to \infty$ in the RHS of (D.5) we obtain that

$$\limsup_{n\to\infty} \frac{|F_n^k(r)|}{n^k} \leq \binom{k}{2}\nu\{(N, o) : |(N, o)_{2r}| > M\}.$$

When we take $M \to \infty$ the RHS of the previous display goes to 0 since ν is supported in networks with locally finite underlying graph. This is enough to finish.

References

[1] David Aldous and Michael Steele. The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence.

[2] Alex Arenas, Albert Díaz-Guilera, Jurgen Kurths, Yamir Moreno, and Changsong Zhou. Synchronization in complex networks. *Physics reports*, 469(3):93–153, 2008. MR-2477097

[3] M.T. Barlow. *Random Walks and Heat Kernels on Graphs*. Lecture note series. Cambridge University Press, 2017. MR-3616731

[4] C. Bordenave and P. Caputo. Large deviations of empirical neighborhood distribution in sparse random graphs. *ArXiv e-prints*, August 2013. MR-3405616

[5] Charles Bordenave. Lecture notes on random graphs and probabilistic combinatorial optimization.

[6] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013. MR-3185193

[7] T. Cabana and J. Touboul. Large deviations for randomly connected neural networks: I. Spatially extended systems. *ArXiv e-prints*, October 2015. MR-3101195

[8] Tanguy Cabana and Jonathan Touboul. Large deviations for randomly connected neural networks: II. state-dependent interactions. To appear in *Advances in Applied Probability*. MR-3877261

[9] Tanguy Cabana and Jonathan Touboul. Large deviations, dynamics and phase transitions in large stochastic and disordered neural networks. *Journal of Statistical Physics*, 153(2):211–269, Oct 2013. MR-3101195

[10] J. Chandra and P. Davis. Linear Generalizations of Gronwall's Inequality. *Proceedings of the american mathematical society*, 60:10, 1976. MR-0422735

[11] 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. *ArXiv e-prints*, December 2016. MR-3918168

[12] Hayato Chiba, Georgi S. Medvedev, and Matthew S. Mizuhara. Bifurcations in the kuramoto model on graphs. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(7):073109, 2018. MR-3833337

[13] F. Coppini, H. Dietert, and G. Giacomin. A Law of Large Numbers and Large Deviations for interacting diffusions on Erdös-Rényi graphs. *ArXiv e-prints*, July 2018. MR-4080158

[14] Sylvain Delattre, Giambattista Giacomin, and Eric Luçon. A note on dynamical models on random graphs and Fokker–Planck equations. *Journal of Statistical Physics*, 165(4):785–798, Nov 2016. MR-3568168

[15] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 2012. MR-3497465

[16] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer Series in Synergetics. Springer Berlin Heidelberg, 2012. MR-0762432

[17] Daniel Lacker, Kavita Ramanan, and Ruoyu Wu. Large sparse networks of interacting diffusions. *arXiv e-prints*, arXiv:1904.02585, Apr 2019.

[18] Daniel Lacker, Kavita Ramanan, and Ruoyu Wu. Locally interacting diffusions as space-time Markov random fields. *arXiv e-prints*, arXiv:1911.01300, Nov 2019.

[19] Xiang Li. Phase synchronization in complex networks with decayed long-range interactions. *Physica D: Nonlinear Phenomena*, 223(2):242–247, 2006.

[20] Eric Luçon and Wilhelm Stannat. Mean field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.*, 24(5):1946–1993, 10 2014. MR-3226169

[21] Eric Luçon and Wilhelm Stannat. Transition from gaussian to non-gaussian fluctuations for mean-field diffusions in spatial interaction. *Ann. Appl. Probab.*, 26(6):3840–3909, 12 2016. MR-3582819

[22] James MacLaurin. Large Deviations of a Network of Neurons with Dynamic Sparse Random Connections. *arXiv e-prints*, arXiv:1607.05471, July 2016.

[23] G. S. Medvedev. The continuum limit of the Kuramoto model on sparse directed graphs. *ArXiv e-prints*, February 2018. MR-4030504

[24] Georgi S Medvedev and Xuezhi Tang. The kuramoto model on power law graphs: Synchronization and contrast states. *Journal of Nonlinear Science*, 1–23, 2018.

[25] Paolo Dai Pra and Frank den Hollander. Mckean-vlasov limit for interacting random processes in random media. *Journal of Statistical Physics*, 84(3):735–772, Aug 1996. MR-1400186

[26] G. H. Reis and R. I. Oliveira. Interacting diffusions on random graphs with diverging degrees: hydrodynamics and large deviations. *ArXiv e-prints*, July 2018. MR-3999471

[27] Yury Sokolov and G Bard Ermentrout. When is sync globally stable in sparse networks of identical kuramoto oscillators? *arXiv preprint*, arXiv:1812.08153, 2018. MR-3983189

[28] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Eté de Probabilités de Saint-Flour XIX, 1989 - 1991*. Springer-Verlag, 1991. MR-1108185

[29] Roman Vershynin. High-Dimensional Probability: An Introduction with Applications in Data Science. MR-3837109

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm