Dynamics of Coupled Maps in Heterogeneous Random Networks

Tiago Pereira[*], Sebastian van Strien[†], and Jeroen S.W. Lamb[‡]
Department of Mathematics, Imperial College London

January 6, 2014

Abstract

We study expanding circle maps interacting in a heterogeneous random network. Heterogeneity means that some nodes in the network are massively connected, while the remaining nodes are only poorly connected. We provide a probabilistic approach which enables us to describe the effective dynamics of the massively connected nodes when taking a weak interaction limit. More precisely, we show that for almost every random network and almost all initial conditions the high dimensional network governing the dynamics of the massively connected nodes can be reduced to a few macroscopic equations. Such reduction is intimately related to the ergodic properties of the expanding maps. This reduction allows one to explore the coherent properties of the network.

1 Introduction

Understanding the behavior of interacting dynamical systems is a long standing problem. Main efforts focused on dynamical systems interacting in a lattice, and with a mean field coupling [1, 2]. Typical attempts establish the existence of an absolutely continuous invariant measure, see for example [3–10]. More recently, the attention has shifted towards more general and irregular networks of coupled maps [11].

The last decade has witnessed a rapidly growing interest in dynamics on spaces with 'complex topology'. This interest is partly motivated by results showing that the structure of a network can dramatically influence the dynamical properties of the system, but also because many, disparate, real-world networks share a common feature – heterogeneity in the interaction structure [12, 13]. This suggests a network structure in which most nodes have degree close to the minimum, while some high-degree nodes, termed *hubs*, are present and have greater impact upon the network functioning. Recent work has suggested the ability of networks with a heterogeneous degree distribution to present degree dependent collective behavior [14–17]. Hubs may undergo a transition to coherence whereas the remaining nodes behave incoherently.

The dynamical properties among the hubs play a central role in many realistic networks. There is evidence that the dynamics of hub neurons coordinate and shape the network in a developing hippocampal network [18], and play a major role in epileptic seizures [19]. Recent

[*]tiago.pereira@imperial.ac.uk
[†]s.van-strien@imperial.ac.uk
[‡]jsw.lamb@imperial.ac.uk

efforts to understand the dynamics of the hubs concentrated mainly on numerical investigations [14–17]. The description of the dynamical properties of the hubs remains elusive.

It remains unknown how hub dynamics depends on various network parameters such as the dynamics of the isolated nodes, and the structure of the network. Revealing the dynamics of the hubs in relation to the graph structure is an important step toward understanding complex behavior and unveiling the topological implications on the network functioning.

In this paper, we study the dynamics of coupled expanding circle maps on a heterogeneous random network. We provide a probabilistic approach to describe effective dynamics of the massively connected nodes in a weak interaction limit. We show that for almost every random network and almost all initial conditions the high dimensional network problem governing the dynamics of the massively connected nodes can be reduced to a few macroscopic equations. Such reduction is intimately related to the ergodic properties of the expanding maps. This reduction allows one to explore the coherent properties of the network. Our analysis reveals that the intrinsic properties of the node dynamics also play a major role and may hinder or enhance coherence.

2 Notation and Statement of the Results

Our main object of study is the network dynamics of coupled maps. *A network of coupled dynamical systems* is defined to be a triple (G, f_i, h) where:

G is a labelled graph of n nodes, termed network; see Section 2.1 for details.

$f_i : M_i \to M_i$ is the local dynamics at each node in the network G, see Section 2.2.

h described the coupling scheme, see Section 2.3.

Abstractly, the network dynamics is defined by the iteration $\Phi : X \to X$, where $X := M_1 \times \cdots \times M_n$ is the product space and $\Phi = A \circ F$, where $F = f_1 \times \cdots \times f_n$, and $A : X \to X$ defines the spacial interaction G and type of coupling interaction scheme h. In Section 2.4 we shall present the above network equation from the single node perspective. In what follows we define our local dynamics and the network class we are working with. Finally, we define the class of coupling functions of interest. For simplicity we introduce the following

Notation: Given functions $a, b\colon \mathbb{R} \to \mathbb{R}$ (or sequences a_n, b_n), we write $a \asymp b$ (resp. $a \lesssim b$) if there exists a universal constant C so that $1/C \le |a/b| \le C$ (resp. $|a| \le C|b|$). Likewise, denote $a \gtrsim b$ if $|a| \ge C|b|$. For simplicity, otherwise unless stated, we understand the sums running over 1 to n. Throughout we denote by $\|\zeta\|_0$ the C^0-norm of a function ζ.

2.1 Random Networks

We concentrate our attention on networks of n nodes described by labelled graphs. Our terminology is that of Refs. [20, 21]. We regard such graphs as networks of size n. We use a random network model $\mathcal{G}(\boldsymbol{w})$ which is an extension of the Erdös-Rényi model for random graphs with a general degree distribution, see for example Ref. [22]. Here $\boldsymbol{w} = \boldsymbol{w}(n)$,

$$\boldsymbol{w} = (w_1, w_2, \cdots, w_n),$$

will describe the expected degree of each node; for convenience we order $w_1 \ge w_2 \ge \cdots \ge w_n \ge 0$, and denote $w_1 = \Delta$. In this model $\mathcal{G}(\boldsymbol{w})$ consists of the space of all graphs of size n, where each potential edge between i and j is chosen with probability

$$p_{ij} = w_i w_j \rho,$$

and where

$$\rho = \frac{1}{\sum_{i=1}^{n} w_i}.$$

To ensure that $p_{ij} \leq 1$ it assumed that $\boldsymbol{w} = \boldsymbol{w}(n)$ is chosen so that

$$\Delta^2 \rho \leq 1. \tag{1}$$

Note that the model $\mathcal{G}(\boldsymbol{w}) = \mathcal{G}(\boldsymbol{w}(n))$ is actually a probability space, where the sample space is the finite set of networks of size n endowed with the power set σ-algebra. Moreover, the probability measure Pr on the sample space is generated by p_{ij}.

Throughout the paper, we will take expectation with respect to measures associated with the node dynamics. Therefore, if for clarity we need to emphasize that the probability and expectation are taken in $\mathcal{G}(\boldsymbol{w})$, we write for a given $C \in \mathcal{G}$ and for a random variable X,

$$\mathrm{Pr}_{\boldsymbol{w}}(C) = \mathrm{Pr}_{\boldsymbol{w}(n)}(C) \text{ and } \mathbb{E}_{\boldsymbol{w}}(X) = \mathbb{E}_{\boldsymbol{w}(n)}(X).$$

Network Property: We call a subset $Q \subset \mathcal{G}(\boldsymbol{w})$ a property of networks of order n if transitivity holds: if G belongs to Q and H is isomorphic to G (this means that the graphs are the same up to relabelling of the nodes) then H belongs to Q as well. We shall say that *almost every* network G in $\mathcal{G}(\boldsymbol{w}(n))$ has a certain property Q if

$$\mathrm{Pr}_{\boldsymbol{w}(n)} \left(G \text{ has property } Q \right) \to 1 \tag{2}$$

as $n \to \infty$. The assertion *almost every* $G \in \mathcal{G}(\boldsymbol{w})$ has property Q is the same as the proportion of all labelled graphs of order n that satisfy Q tends to 1 as $n \to \infty$.

These networks may be described in terms of its *adjacency matrix* A, defined as

$$A_{ij} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise} \end{cases}$$

In the model $\mathcal{G}$ each element of the adjacency A_{ij}'s is an independent Bernoulli (random) variable, taking value 1 with success probability p_{ij}. The *degree* k_i of the ith node is the number of connections it receives. k_i is a random variable, which in terms of the adjacency matrix reads $k_i = \sum_j A_{ij}$. Note that as the network is oriented the number of connections k_i^{out} the ith node makes with its neighbors need not to be k_i. Notice that $k_i^{out} = \sum_j A_{ji}$. An interesting property of this model is that under this construction w_i is the expected value of k_i, that is, $\mathbb{E}_{\boldsymbol{w}}(k_i) = w_i$ while also $\mathbb{E}_{\boldsymbol{w}}(k_i^{out}) = w_i$.

We are interested in the large size behavior of heterogeneous networks. Here we say that a network is *heterogeneous* when there is a considerable disparity between the node's degree. Real world networks are typically heterogeneous – a small fraction of nodes is massively connected whereas the remaining nodes are only poorly connected. To be precise, we study the following class

Definition 1 (Strong Heterogeneity)**.** *We say that the model $\mathcal{G}(\boldsymbol{w}(n))$ is strongly heterogeneous if the following hypotheses are satisfied:*

N1 – Massively connected Hubs: There is $\ell \in \mathbb{N}$ such that for every $1 \leq i \leq \ell$ if we write

$$w_{i,n} = \kappa_{i,n} \Delta_n$$

then $\kappa_{i,n} \to \kappa_{i,\infty}$ as $n \to \infty$ and $\kappa_{i,\infty} \in (0, 1]$ (and $\kappa_{1,n} = 1$). We regard $\kappa_{i,n}$ as the normalized degrees.

N2 – Slow growing low degrees : for $\ell < i \leq n$

$$\log n \lesssim w_{i,n} \lesssim \Delta_n^{1-\gamma}$$

where $0 < \gamma < 1$ control the scale separation between low degree nodes and the hub nodes.

N3 – Cardinality of the Hubs: The number of hubs $\ell = \ell(n)$ satisfies

$$\ell \lesssim \Delta_n^{\theta}$$

for some small $\theta \leq 1$.

We will often suppress the n dependence in the notation so write $\boldsymbol{w}$ instead of $\boldsymbol{w}(n)$ and we also often write Δ and w_i, κ_i instead of Δ_n and $w_{i,n}, \kappa_{i,n}$.

The hypothesis that the number of highly connected nodes is significantly smaller than the system size, $\theta < 1$, implies that the network is heterogeneous. The ℓ high-degree nodes are termed *hubs*, and the remaining nodes are called *low degree nodes*. These assumptions on heterogeneity mimic conditions observed in a large class of realistic networks including neuronal networks, social interaction, internet, among others [12, 13, 18]. Typically, the number of hubs in the network is much smaller than the system size (and the degrees of the hubs). We wish to treat the large networks, so we assume that n is large. That is, the networks we consider have finite but large number of nodes.

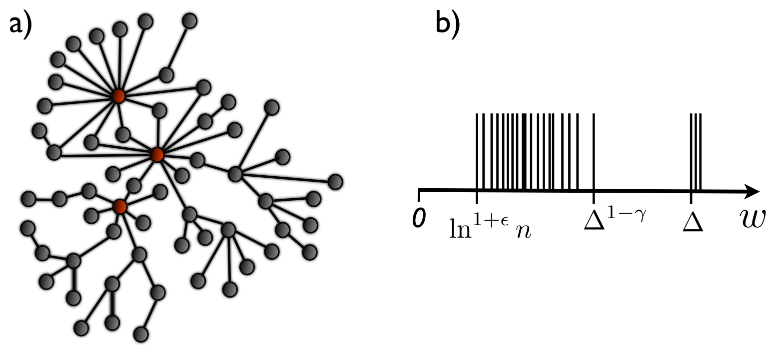

Figure 1: The strong Heterogeneity condition. In a) a heterogeneous network, the hubs are depicted in red. While the hubs are massively connected most nodes have only a few connections. In b) a pictorial presentation of *N1* and *N2*. The axis denotes the expected degrees $w_i's$ ordered according to their magnitudes. Whilst the hubs increase proportionally to $\Delta = \Delta_n$ the remaining nodes increase in another scale.

An interesting property of such random graphs is that under the condition N2 the degrees have good concentration properties, as the next result shows

Proposition 1 (Concentration)**.** *Let $\mathcal{G}(\boldsymbol{w})$ satisfy the strong heterogeneity condition. Then almost every random network $G \in \mathcal{G}(\boldsymbol{w})$ has every vertex satisfying*

$$|k_i - w_i| < w_i^{1/2-\varepsilon} \tag{3}$$

for any $\varepsilon > 0$. (This statement should be understood in the sense of (2).)

Proof. We obtain this result via a Chebyshev inequality. The expected value of $k_i = \sum_j A_{ij}$ is given by $\mathbb{E}_{\boldsymbol{w}}(k_i) = \sum_j w_i w_j \rho = w_i$. To estimate the variance of $\mathrm{Var}_{\boldsymbol{w}}(k_i)$ notice that

$$\mathbb{E}^2_{\boldsymbol{w}}(k_i) = w_i^2 = w_i^2 \rho^2 \sum_{j,k} w_j w_k \tag{4}$$

and that $\mathbb{E}_{\boldsymbol{w}}(k_i^2) = \mathbb{E}_{\boldsymbol{w}}(\sum_{j,k} A_{ij} A_{ik})$. If $j \neq k$ then A_{ij} and A_{ik} are independent, otherwise notice that $A_{ij}^2 = A_{ij}$. This remark leads to

$$\mathbb{E}_{\boldsymbol{w}}(k_i^2) = w_i + w_i^2 \rho^2 \sum_{j \neq k} w_i w_k \tag{5}$$

Combining (5) and (4) we obtain

$$\mathrm{Var}_{\boldsymbol{w}}(k_i) = w_i - w_i^2 \rho^2 \sum_j w_j^2.$$

This implies $\mathrm{Var}_{\boldsymbol{w}}(k_i) < w_i$. (In fact, the previous equation gives $(1 - p_{i,1}) w_i \leq \mathrm{Var}_{\boldsymbol{w}}(k_i) < w_i$ where $p_{i1} \leq 1$ is the probability to connect node i to the main hub node 1 since $w_j^2 \leq \Delta w_j$ gives $\sum_j w_j^2 \leq \Delta / \rho$ and therefore $w_i^2 \rho^2 \sum_j w_j^2 \leq w_i p_{i1}$.) Hence the Chebyshev inequality yields

$$\mathrm{Pr}_{\boldsymbol{w}}(|k_i - w_i| \geq w_i^{1/2-\varepsilon}) \leq w_i^{-2\varepsilon},$$

and condition N2 implies that this probability tends to zero as n increases. Therefore, we obtain that $|k_i - w_i| \leq w_i^{1/2-\varepsilon}$ for almost every network, which concludes the proof. $\square$

Remark 1. *This model imposes minimum and maximum growth conditions for Δ_n in terms of the network size n. Indeed (N2) implies*

$$\Delta_n \gtrsim (\log n)^{\frac{1}{1-\gamma}}$$

Moreover, (1) leads to an upper estimate for the growth of the maximum expected degree Δ_n, namely

$$\Delta_n \lesssim \max\{n^{1/2+\delta}, n^{\frac{1}{1+\gamma}}\}$$

for some $\delta > 0$ small enough.

The next example provides an illustration of a degree sequence satisfying the strong heterogeneity condition. We shall construct networks from this illustration for numerical simulations later on in the paper.

Example 1: An example of a degree sequence satisfying the strong heterogeneity condition is

$$\boldsymbol{w} = (\kappa_1 \Delta, \ldots, \kappa_\ell \Delta, w_{\ell+1}, \ldots, w_n)$$

with $1 = \kappa_1 \geq \kappa_2 \geq \ldots \geq \kappa_\ell > 0$, so that

$$\Delta = n^\sigma \,,\ \ell = \Delta^\theta \text{ and } w_i \asymp \Delta^{1-\gamma} \text{ for } i \geq \ell$$

where $\gamma, \sigma, \theta \in (0,1)$. Condition (1) imposes a growth condition on the expected degree with scaling coefficient

$$\sigma < \frac{1}{1+\gamma}.$$

It easy to modify this example so that the expected degree sequence exhibits a power law behavior in the distribution of the expected degrees, and other non trivial distributions.

2.2 Local Dynamics

We choose the dynamics on each node of the network to be identical $f_i = f$. In fact, it turns out that under our hypothesis if the dynamics on each node is slightly different our claims still hold true. We consider expanding maps on the circle $M = \mathbb{R}/\mathbb{Z}$ as a model of the isolated dynamics of the nodes. We will use that M is compact and the addition structure coming from $\mathbb{R}$, see Ref. [23] for details.

Assumption 1. *Let $f : M \to M$ be a $C^{1+\nu}$ Hölder continuous expanding map, for some $\nu \in (0, 1]$. That is, we assume that there exists $\sigma > 1$ such that*

$$\|Df(x)v\| \geq \sigma \|v\|$$

for all $x \in M$ and $v \in T_xM$, for some riemannian metric $\| \cdot \|$, and Df is Hölder continuous with exponent ν.

The differentiability condition, that is, $\nu > 0$, plays an important role in our analysis. It is well known that if $\nu = 0$, then f may admit invariant measures which are singular with respect to Lebesgue measure. Moreover, if $\nu > 0$ the system is structurally stable.

2.3 Interaction Function

Our aim in this subsection is to introduce the interaction structure we will use in the network dynamics (7). We consider pairwise interaction

$$h : M \times M \to \mathbb{R}.$$

For simplicity we assume h satisfies the following representation

$$h(x, y) = \sum_{p,q=1}^{k} u_p(x) v_q(y), \tag{6}$$

where k is an integer, and $u_p, v_q \colon M \to \mathbb{R}$ are $C^{1+\nu}$ functions (i.e. for each $x, y \in \mathbb{R}$ and $r, s \in \mathbb{Z}$ one has $u_p(x + r) = u_p(x)$ and $v_q(y + s) = v_q(y)$). Notice that h is well-defined. Of particular interest in applications is the interactions akin to diffusion,

$$h(x, y) = \sin 2\pi(x - y) \;\text{ and }\; h(x, y) = \sin 2\pi x - \sin 2\pi y.$$

With this interaction function we are ready to introduce the network dynamics.

2.4 Network Dynamics

Given an integer n and a $n \times n$ interaction matrix A, we consider the dynamics of a network of n coupled maps is described by

$$F \colon M^n \to M^n \text{ where } (x_1(t + 1), \dots, x_n(t + 1)) = F(x_1(t), \dots, x_n(t))$$

is defined by

$$x_i(t + 1) = \hat{f}(x_i(t)) + \frac{\alpha}{\Delta} \sum_{j=1}^{n} A_{ij} h(x_j(t), x_i(t)) \pmod 1, \quad \text{for } i = 1, \dots, n. \tag{7}$$

Here $\hat{f}: M \to \mathbb{R}$ is the lift of $f: M \to M$ (note that $M = \mathbb{R}/\mathbb{Z}$). The right hand side in (7) is well-defined because M has an addition structure, $h: M \times M \to \mathbb{R}$ is well-defined and for each choice of lift $f: M \to M$ to a map $\hat{f}: M \to \mathbb{R}$ the expression in (7) gives the same result. By abuse of notation we will denote the lift $\hat{f}$ also by f. Here A_{ij} is chosen as in Section 2.1, $x_i(t)$ describes the state of the i-th node at time (which has degree $k_i = \sum_j A_{ij}$) and $\Delta = w_1$ is the largest expected degree. Moreover, α is a free parameter which describes the normalized overall coupling strength. We shall also denote the state $x_i(t)$ as x_i whenever convenient if there is no risk of confusion. We are interested in the weak coupling limit as $n \to \infty$, that is, when α is independent of Δ and n.

3 Main Result and Discussions

Our main goal is to obtain a low dimensional equation to describe the highly connected nodes. Our strategy is to prove the following reduction by means of an effective dynamics. For simplicity we choose the initial conditions on M^n to be independent and identically distributed. More precisely

Choice of Initial conditions: Let μ be a measure supported on M with density φ. Moreover, let $\log \varphi$ be $(a, \nu)-$Hölder where $\nu \in (0, 1]$ (this notion is defined in Section 5.1). Consider the global phase space of the coupled maps M^n, and a product measure $\mu^n : M^n \to [0, 1]$ be given such that $A = A_1 \times \cdots \times A_n \subset M^n$ we have $\mu^n(A) = \mu(A_1) \cdots \mu(A_n)$. This choice of initial conditions is natural in numerical experiments. We shall use this choice to state our main result

Theorem 1 (Dynamics of Hubs)**.** *Let $\mathcal{G}(\boldsymbol{w})$ be strongly heterogeneous and consider the coupled map network (7) on $\mathcal{G}(\boldsymbol{w})$. Assume that the network structural parameters satisfy*

$$0 < \theta < 1 - \gamma\nu/2,$$

and let all initial conditions on M^n be given according to a product measure μ^n. Then there exists a positive number $u = u(\varphi)$ such that for almost every network in $\mathcal{G}(\boldsymbol{w})$ and for μ^n-almost every initial condition the dynamics of the hubs $i = 1, \ldots, \ell$ is reduced to

$$x_i(t+1) = f(x_i(t)) + \alpha\kappa_i g(x_i) + \alpha\zeta_i(t) \ \ \textit{(mod1)}\ , \textit{ for each } t \geq u \tag{8}$$

where κ_i is the normalized degree, and

$$g(x) = \int h(x, y)\mu_0(dy)$$

where μ_0 is the invariant measure of f. Moreover, $\zeta_i(t)$ satisfies

$$\|\zeta_i\|_0 \lesssim \kappa_i^{1/2} \Delta^{-\gamma\nu/2+\delta} \textit{ for each } t \geq u$$

for any $\delta > 0$ (where $\|\zeta_i\|_0$ is the C^0-norm of ζ_i).

Remark 2. *If $0 < \theta < 1 - \gamma\nu/2$ is not satisfied the function ζ stills converges to zero as $\Delta \to \infty$. The speed of convergence will then depend on the parameters of the network, see Proposition 7.*

Our result has a mean field flavor. Loosely speaking, the hubs interact with their local mean field. As the network is random the local mean field perceived by the hubs is a fraction of the total mean field with weights given by the normalized degrees κ_i. On the other hand, the mean field is determined by the invariant measure μ_0 of the isolated dynamics. The function ζ_i describes the noise-like disturbance on the motion of the hubs.

μ^n*−almost every initial condition:* The claim concerning the reduction holds almost surely with respect to μ. Indeed, we could arrange the initial conditions into fixed points of the coupled maps and consider the particular case that $h(x, y) = h(x-y)$ and $h(0) = 0$. Then the reduction obviously fails in the realization where $x_1 = \cdots = x_n = \tilde{x}$. In this situation the dynamics of the hubs are given by isolated dynamics, and not by our reduction. This example shows that one needs to remove a zero measure set of initial conditions

Almost surely every graph: The assertion that the theorem holds for almost every graph is more subtle. We control the concentration inequalities of relevant quantities associated with our results. But with small probability close to zero some pathological networks may appear. In such exceptional networks the mean field reduction may not hold. We now discuss a particular example of such a pathology. Recall that the networks we consider are oriented. However, $\mathbb{E}_{\boldsymbol{w}}(k_i) = \mathbb{E}_{\boldsymbol{w}}(k_i^{out})$, which implies that statistically for a given node the number of incoming and outgoing are the same. Let us consider a situation $k_1 = 0$ and $k_1^{out} = \Delta$. We present a pictorial representation of such network in Fig. 2. Clearly the dynamics of this hub will not be described by our reduction Theorem, as it does not receive input from its neighbors and therefore acts as an isolated node. Notice that

$$\mathrm{Pr}_{\boldsymbol{w}}(k_1 = 0) = \Pi_{j=1}^{n}(1 - p_{1j})$$

together with $1 - \Delta w_j \rho \leq 1 - \Delta^{\gamma}/n$, which implies that

$$\mathrm{Pr}_{\boldsymbol{w}}(k_1 = 0) \lesssim e^{-\Delta^{\gamma}}.$$

This shows that the probability to find such networks converges to zero as Δ grows.

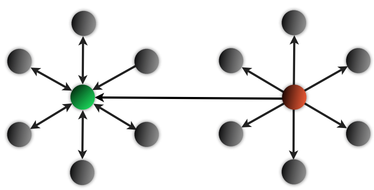

Figure 2: A network with a fully oriented hub. The hub only exerts influence on its neighbors, but does not receive input from them. Networks with such property form a set of small measure.

More generically, due to the concentration inequality described in Proposition 3 the degrees only differ from w_i by a amount proportional to w_i in a set of networks of small measure.

4 Examples of Reductions

We simulate the scenario from the Main Theorem using the network model in Example 1 from Section 2.1, with $n = 2 \times 10^4$, $\ell = 2$ and $\kappa_1 = 1$ and $\kappa_1 \geq \kappa_2 \geq 0$. We take $w_i = 7$ for $2 < i \leq 4000$. Moreover, we consider $\sigma < 1/2$. This network can be thought of as composed of a Erdös-Rényi layer corresponding to $n > i > \ell$ and another layer of highly connected nodes. A pictorial representation of such network with $\ell = 2$ and $n = 10$ can be seen in Fig. 3

We take the Bernoulli map

$$f(x) = 2x \mod 1$$

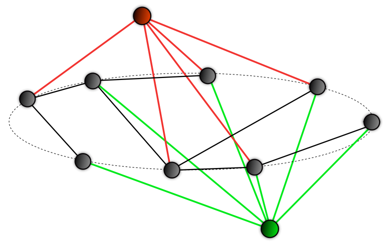

Figure 3: A pictorial representation of the network described in Example 1 with $\ell = 2$ and $n = 10$

to model the isolated dynamics. This model corresponds to a case with Hölder exponent $\nu = 1$. The Lebesgue measure m is invariant for the localized system. This means that $\mu_0 = 1$. Another interesting property of the system is stochastic stability, which means that adding a δ-small uncorrelated noise in the evolution the dynamics can still be described by an invariant measure μ_δ. Moreover, μ_δ is uniformly close to m. It is easy to see that all our results still remain true if this small noise is included.

Hence, to avoid round-off numerical problems associated with the map $x \mapsto 2x$, we introduce a small additive noise ξ with uniformly distributed with support $[0, 10^{-5}]$. Hence, the isolated dynamics under the influence of this small noise is given by

$$x(t+1) = 2x(t) + \xi(t) \mod 1$$

We wish to explore two coupling functions and their consequences for the hub dynamics, focusing in particular on the collective properties of the hubs. To this end, we introduce the following coherence measure. Given points $x_1(t), x_2(t) \in M$, with $t \in \{1, \dots, T\}$, we define the *coherence* r between the hubs by

$$re^{i\psi} = \frac{1}{T} \sum_{j=1}^{T} e^{2\pi(x_1(t) - x_2(t))}. \tag{9}$$

4.1 Hubs decouple from the network.

Consider the coupling function

$$h(y - x) = \sin 2\pi(y - x). \tag{10}$$

Our reduction technique renders the following interaction function for the hubs

$$g(x) = \int_0^1 h(y - x)dy = 0.$$

According to the Main Theorem, the coupling equation (10) gives the reduced equations

$$x_i(t+1) = 2x_i(t) + \alpha\zeta_i \mod 1, \text{ for } i = 1, 2 \; (\ell = 2) \tag{11}$$

and where ζ_i is the noise term. Hence, the hubs effectively decouple from the network, as there is no interaction with the hubs with the mean field. The only effect of the network on the hubs is a noise-like term ζ_i and the parameter α only appears in the noise term. The stochastic stability of the Bernoulli family implies that for T large enough $r \approx 0$.

First, for a fixed $\Delta = 260$ and $\kappa_2 = 0.99$ we compute r as a function of α. In the computation of the coherence measure r we discard the first 10^3 iterates and consider $T = 10^3$. As predicted no coherence is attained, and the behavior of r as a function of α is flat. The results are presented in Fig. 4a). In Fig. 4b) the time series of $|\cos 2\pi x_1(t) - \cos 2\pi x_2(t)|$ are shown.

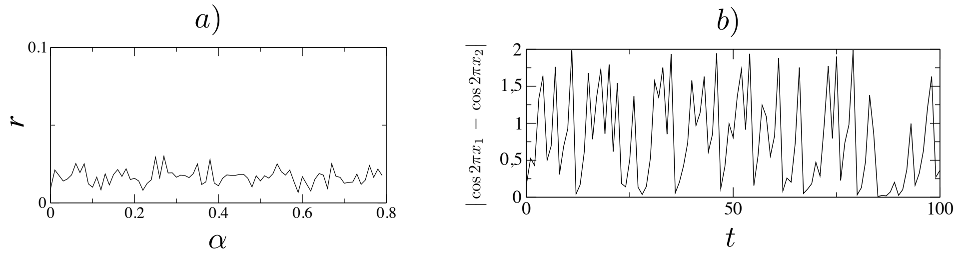

Figure 4: No coherent dynamics between hubs. In a) we show the coherence measure r as a function of alpha. In b) the time series of $|\cos 2\pi x_1(t) - \cos 2\pi x_2(t)|$.

4.1.1 Effects of Δ on the fluctuations

We now perform a set of simulations to study the scaling relations between ζ and Δ. Notice that in the model of Example 1 we take $\gamma = 1$, hence our Theorem predicts a scaling as $\zeta \lesssim \Delta^{-1/2}$.

In this set of experiments we vary Δ, recall that $n = 2 \times 10^4$ and the low degrees $w_i = 7$ for $2 < i \leq n$ are fixed. We also fix $\alpha = 0.1$, and consider only the fluctuations ζ_1 on the main hub x_1. For simplicity of notation we shall write $\zeta = \zeta_1$. For each experiment we compute the quantity

$$\langle|\zeta|\rangle = \frac{1}{T}\sum_{t=1}^{T} |\zeta(t)|$$

where $T = 10^3$. This mean value of the modulus of the fluctuations ζ must have the same scaling relation $\langle|\zeta|\rangle \lesssim \Delta^{-1/2}$. For each Δ we construct the corresponding network only once and measure $\langle|\zeta|\rangle$. This implies that we do not average $\langle|\zeta|\rangle$ over the network ensemble, as the networks have good concentration properties. The simulation results are presented in Fig. 5 and are in agreement with the predictions

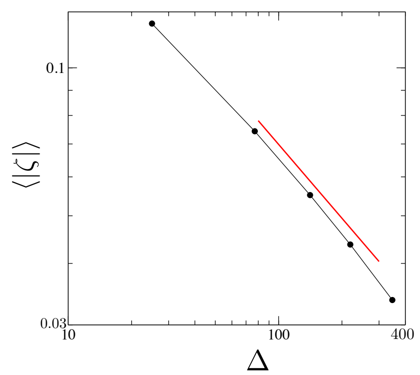

Figure 5: Scaling relation between the mean value of the modulus of the fluctuations $\langle|\zeta|\rangle$ and the expected degree Δ. The simulation results are shown in black full points and a curve with scaling $\Delta^{-1/2}$ is shown in red (full bold line). Our Theorem predicts a scaling relation $\langle|\zeta|\rangle \lesssim \Delta^{-1/2}$, which is in agreement with the simulation results.

4.1.2 Effects of the normalized degree κ on the fluctuations

As before n and the w_i's for the low degree nodes are fixed. Now we also fix $\Delta = 347$ and vary κ_2 to study the scaling relations between ζ and κ. Our Theorem predicts a scaling as $\zeta \lesssim \kappa^{1/2}$.

We then compute the mean value of the modulus of the fluctuations $\langle|\zeta_2|\rangle$. In this subsection, for simplicity of notation we shall write $\langle|\zeta|\rangle = \langle|\zeta_2|\rangle$ and $\kappa = \kappa_2$. Again, for each κ we construct the corresponding network only once and measure $\langle|\zeta|\rangle$. The simulation results are presented in Fig. 6 and is in agreement with the predictions

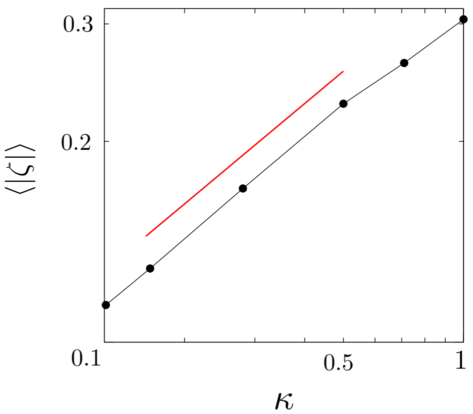

Figure 6: Scaling relation between the mean value of the modulus of the fluctuations $\langle|\zeta|\rangle$ and normalized degree κ. The simulation results are shown in black full points and a curve with scaling $\kappa^{1/2}$ is shown in red (full bold line). For fixed network parameters n, Δ Our Theorem predicts a scaling relation $\langle|\zeta|\rangle \lesssim \kappa^{1/2}$, which is in agreement with the simulation results.

4.2 Reduction Reveals Coherent Behavior

The reduction reveals that coherent dynamics can be caused by cancellations due to the nature of the coupling functions and the dynamics of the isolated dynamics (in terms of the invariant measure). Depending on the coupling function the hubs may exhibit a coherent behavior for a range of coupling strengths α. Here, we illustrate such scenario. Again, we consider the network from Example 1, with $n = 2 \times 10^3$, $w_i = 7$ for $2 < i \leq n$. We fix $\Delta = 347$, with $\kappa_1 = 1$ an $\kappa_2 = 1$. Then we construct one realization of such a network. The coupling function

$$h(x, y) = \sin 2\pi y - \sin 2\pi x. \tag{12}$$

We performed extensive numerical simulations to compute the coherence measure for the two hubs x_1 and x_2 as a function of the coupling strength α. The result can be seen in Fig. 9. The parameter r, see (9) has an intricate dependence on α. Such behavior can be uncovered by our reduction techniques.

Applying the Main Theorem with the coupling (12) yields the following effective equations for the hubs dynamics

$$x_i(t+1) = 2x_i(t)) + \alpha \sin 2\pi x_i + \alpha\zeta_i(t) \text{ mod } 1 \;, \tag{13}$$

Let us neglect the finite size fluctuations for a moment. Then the dynamics of the hubs are described by the following equation

$$x(t+1) = 2x(t) + \alpha \sin 2\pi x \mod 1, \tag{14}$$

and so the parameter α determines the dynamics significantly. Indeed (14) has a trivial fixed point $x = 0$ (identified with 1) for all α, but this fixed point is only stable for

$$\frac{1}{2\pi} < \alpha < \frac{3}{2\pi}. \tag{15}$$

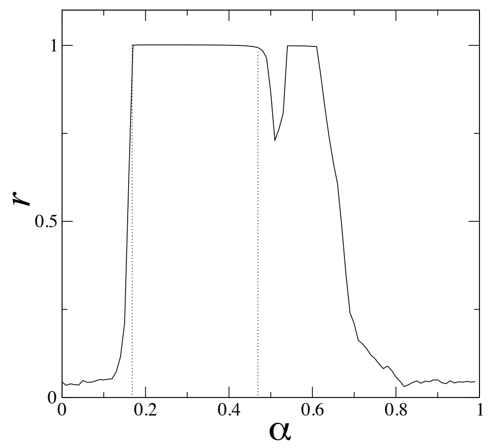

Figure 7: The coherence parameter r versus the coupling strength α. The high values of r corresponds to regimes where the dynamical of the hubs x_1 and x_2 are correlated. The onset of coherent regimes can be predicted by the reduction techniques, e.g., the large coherent plateau is given by (15).

In this range if the initial conditions for the hubs start in a vicinity of 0 then they will remain there for all future times. This will correspond to a trivial coherent behavior. Hence, while the low degree behaves in an erratic fashion the hubs, although isolated chaotically, will stay in a steady state. This regime corresponds to large values of the coherence measure r. On the other hands, when $\alpha \notin (\frac{1}{2\pi}, \frac{3}{2\pi})$ this fixed point becomes repelling and indeed the second plateau in Fig. 7 corresponds to a stable periodic orbit of period two. Examples of such dynamics can be observed in Fig. 8

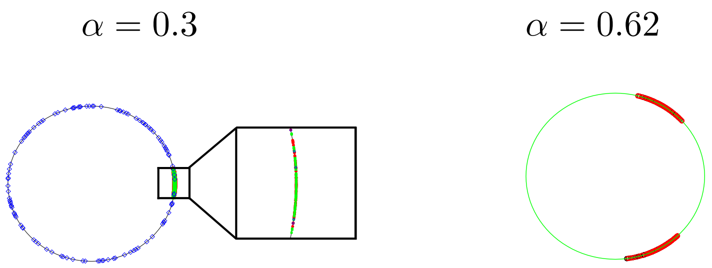

Figure 8: Coherent dynamics of the hubs. For $\alpha = 0.3$ we depict the trajectories of the hubs x_1 (solid circles) and x_2 (solid squares) along with the trajectories of the low degree node x_{400} (diamonds). The trajectories are shown in the circle. The trajectories of the hubs stay close to an equilibrium point, whereas the trajectories of the low degree nodes spread over the whole circle. For $\alpha = 0.62$ we depict the trajectories of the hubs in the circle. They display a coherent periodic motion of period two.

5 Dynamics of Low degree Nodes

5.1 Isolated Nodes

The proof of the Main Theorem is based on the ergodic properties of f and is derived from the properties of the Peron-Frobenius operator (acting on some convenient Banach space). To prove that this operator has fixed points, we use standard technique based on the notion of the projective metric associated with a convex cone in a vector space. We follow closely the exposition in Ref. [24]. We use the distance d induced by the riemannian metric.

Let $E = C^0(M, \mathbb{R})$ be the space of continuous real valued functions defined on M. Let $C = C(a, \nu)$ be the convex cone of functions $\varphi \in E$ such that

i) $\varphi(x) > 0$ for all $x \in M$ and

ii) $\log \varphi$ is (a, ν)-Hölder continuous on ρ_0 neighborhoods.

This last condition means that for all $x, y \in M$ such that $d(x, y) \leq \rho_0$ then

$$e^{-ad(x,y)^\nu} \leq \frac{\varphi(x)}{\varphi(y)} \leq e^{ad(x,y)^\nu}.$$

The projective metric (Hilbert metric) is introduced as follows given $\varphi_1, \varphi_2 \in C$, define

$$\alpha(\varphi_1, \varphi_2) = \sup\{t > 0 : \varphi_2 - t\varphi_1 \in C\},$$

and

$$\beta(\varphi_1, \varphi_2) = \inf\{s > 0 : s\varphi_1 - \varphi_2 \in C\}.$$

Then,

$$\theta(\varphi_1, \varphi_2) = \log \frac{\beta(\varphi_1, \varphi_2)}{\alpha(\varphi_1, \varphi_2)} \tag{16}$$

is a metric in the projective quotient of C.

Statistical properties of the isolated dynamics can be obtained by means of the transfer operator $\mathcal{L}$. Let $\varphi : M \to \mathbb{R}$, the operator $\mathcal{L}$ is defined by

$$(\mathcal{L}\varphi)(y) = \sum_{f(x)=y} \frac{\varphi(x)}{|\det Df(x)|}$$

Given a measure μ absolutely continuous with respect to the Lebesgue measure m, that is, given $A \subset M$ one has $\mu(A) = \int_A \varphi dm$, for an integrable function φ, then the pushforward of μ under f has the duality property

$$(f_*\mu)(A) \;\; = \;\; \int_A \mathcal{L}\varphi dm. \tag{17}$$

The fixed points φ_0 of the linear operator $\mathcal{L}$ are f-*invariant* absolutely continuous probabilities measures. Conversely, if a f-invariant probability μ_0 is absolutely continuous with respect to m then there exists the Radon-Nikodym derivative $d\mu_0/dm = \varphi_0$ and $\mathcal{L}\varphi_0 = \varphi_0$.

The domain of $\mathcal{L}$ plays a vital role. Another important point is that the metric space $(C(a, \nu), \theta)$ is not complete. Nonetheless, $\mathcal{L}$ acting on $C(a, \nu)$ is a contraction with a unique fixed point.

Proposition 2. *The operator $\mathcal{L} : C(a, \nu) \to C(a, \nu)$ is a contraction with respect to the projective metric $\theta = \theta_{a,\nu}$ associated with the convex cone $C(a, \nu)$. Moreover, $\mathcal{L}$ has a unique fixed point $\varphi_0 \in C(a, \nu)$.*

Proof. See Propositions 2.2, 2.4 and 2.6 cf. [24]. □

5.2 Transfer Operator of the Low degree Nodes

We study the dynamics of the low degree nodes as perturbations of the isolated dynamics, and then obtain the statistical properties of the low degree in terms of perturbation results of the transfer operator.

To this end, we show that for almost every network, low degree nodes can be viewed as a perturbation. Notice that

$$x_i(t+1) = f(x_i(t)) + \frac{\alpha}{\Delta}\sum_{j=1}^{n} A_{ij} h(x_j(t), x_i(t)),\ i = 1, \dots, n \tag{18}$$

can be rewritten as

$$x_i(t+1) = f(x_i(t)) + \alpha r_i(x(t), y_i(t)),\ i = 1, \dots, n \tag{19}$$

where the coupling term taking into account (6) is represented as

$$r_i(x_i, y_i) = \sum_{p,q} u_p(x_i) y_{q,i}$$

with

$$y_{q,i} = \frac{1}{\Delta}\sum_j A_{ij} v_q(x_j).$$

Our next result guarantees that the coupling term can be made uniformly small as Δ increases.

Proposition 3. *Let $\mathcal{G}(\boldsymbol{w})$ be strongly heterogeneous. Then the coupling term r_i viewed as a mapping $x \mapsto r_i(x, y_i)$ is $C^{1+\nu}$, and for every low degree node $\ell < i \leq n$ the coupling term satisfies*

$$\|r_i(x_i, y_i)\| \lesssim \Delta^{-\gamma}$$

for almost every network in $\mathcal{G}(\boldsymbol{w})$.

Proof. The claim on the regularity of r_i is trivial. The second claim follows from the concentration properties. Since the manifold is compact, we obtain for every $x \in M$ $\|v_q(x)\| \leq K$ and $\|u_p(x)\| \leq K$ Now we wish to show concentration properties for $y_{q,i}(t)$. Hence, for a given fixed t for simplicity we denote $y_{q,i}(t) = y_{q,i}$ and estimate

$$\mathbb{E}_{\boldsymbol{w}}(y_{q,i}) = \frac{w_i}{\Delta}\rho\sum_{i=1}^{n} w_j v_q(x_j) \tag{20}$$

leading to

$$\|\mathbb{E}_{\boldsymbol{w}}(y_{q,i})\| \lesssim \Delta^{-\gamma}.$$

To obtain our claim we also estimate the variance

$$\mathrm{Var}_{\boldsymbol{w}}(y_{q,i}) = \mathbb{E}_{\boldsymbol{w}}(y_{q,i}^2) - \mathbb{E}_{\boldsymbol{w}}^2(y_{q,i}),$$

hence, we need to estimate

$$\mathbb{E}_{\boldsymbol{w}}(y_{q,i}^2) = \mathbb{E}_{\boldsymbol{w}}\left(\frac{1}{\Delta^2}\sum_{j,k} A_{ij}A_{ik} v_q(x_j) v_q(x_k)\right) \tag{21}$$

$$= \mathbb{E}_{\boldsymbol{w}}\left(\frac{1}{\Delta^2}\sum_{j} A_{ij}^2 v_q(x_j)^2 + \frac{1}{\Delta^2}\sum_{j\neq k} A_{ij}A_{ik} v_q(x_j) v_q(x_k)\right). \tag{22}$$

But $A_{ij}^2 = A_{ij}$, hence,

$$\mathbb{E}_{\boldsymbol{w}}(y_{q,i}^2) \;=\; \left(\frac{w_i}{\Delta^2}\rho \sum_j w_j v(x_j)^2 + \frac{w_i^2}{\Delta^2}\rho^2 \sum_{j\neq k} w_j w_k v_q(x_j) v_q(x_k) \right). \tag{23}$$

Combining both computations for $\mathbb{E}_{\boldsymbol{w}}(y_{q,i}^2)$ and $\mathbb{E}_{\boldsymbol{w}}^2(y_{q,i})$ we obtain

$$\mathrm{Var}_{\boldsymbol{w}}(y_{q,i}) = \frac{w_i}{\Delta^2}\rho \sum_j w_j v_q(x_j)^2 - \frac{w_i^2}{\Delta^2}\rho^2 \sum_j w_j^2 v_q(x_j)^2,$$

and as the functions v_q are bounded

$$\|\mathrm{Var}_{\boldsymbol{w}}(y_{q,i})\| \leq K^2 \left(\frac{w_i}{\Delta^2} + \frac{w_i^2}{\Delta^2}\rho^2 \sum_j w_j^2 \right).$$

Recall that by property N2 of the strong heterogeneity $w_i \leq \Delta^{1-\gamma}$, and note that $\sum_j w_i^2 \leq \Delta/\rho$ which leads to the following estimate

$$\frac{w_i^2}{\Delta^2}\rho^2 \sum_j w_j^2 \leq \frac{w_i^2 \rho}{\Delta} \lesssim \Delta^{-1-3\gamma}.$$

This leads to

$$\|\mathrm{Var}_{\boldsymbol{w}}(y_{q,i})\| \;\lesssim\; \Delta^{-1-\gamma} + \Delta^{-1-3\gamma} \;\lesssim\; \Delta^{-1-\gamma}. \tag{24}$$

Hence, for every $(1-\gamma)/2 > \beta > 0$ we obtain

$$\mathrm{Pr}_{\boldsymbol{w}} \left(\|y_{q,i} - \mathbb{E}_{\boldsymbol{w}}(y_{q,i})\| \gtrsim \Delta^{-(1+\gamma)/2+\beta} \right) \lesssim \Delta^{-2\beta}$$

implying that for almost every network in the model $\mathcal{G}(\boldsymbol{w})$,

$$\|y_{q,i} - \mathbb{E}_{\boldsymbol{w}}(y_{q,i})\| \lesssim \Delta^{-(1+\gamma)/2+\beta}.$$

By the triangle inequality $|\|y_{q,i}\| - \|\mathbb{E}_{\boldsymbol{w}}(y_{q,i})\|| \lesssim \Delta^{-(1+\gamma)/2+\beta}$ together with (20) and taking β small enough we obtain that

$$\|y_{q,i}\| \lesssim \Delta^{-\gamma}.$$

Now recall that the function u_p are uniformly bounded over M yielding

$$\|r_i(x_i, y_i)\| = \|\sum_{p,q} u_p(x_i) y_{q,i}\| \lesssim \Delta^{-\gamma}$$

for almost every network as $y_{q,i}$ has the derived concentration properties. □

This result reveals that for almost every network $\mathcal{G}(\boldsymbol{w})$ the network effect on low degree nodes is a perturbation in the limit of large Δ. We use this remark to treat the mean field reduction in terms of the ergodic properties of the perturbed maps.

Next we consider the index i fixed on a low degree node. The following argument will hold for any low degree node. We can view the perturbed map $f_t = f + r$ as a random-like perturbation of the map f by writing

$$f_t(x) = f(x) + r(x(t), t)$$

where r is small by Proposition 3. Note that the maps f_t are uniformly close in the C^1 topology to f. We view f_t as parametrized families $f_t : M \to M$ of $C^{1+\nu}$-Hölder continuous maps. So at each time step in its evolution, we pick a map f_t in an open neighborhood of f. Denote $\boldsymbol{t} = (t_1, t_2, \cdots)$ and define

$$f_{\boldsymbol{t}}^k = f_{t_k} \circ f_{t_{k-1}} \circ \cdots \circ f_{t_1},$$

hence, the equation may be recast as

$$x^{k+1} = f_{\boldsymbol{t}}^k(x_0).$$

For that map we introduce perturbed versions of the linear operator $\mathcal{L}$

$$(\mathcal{L}_t \varphi)(y) = \sum_{f_t(x)=y} \frac{\varphi(x)}{|\mathrm{det} Df_t(x)|}.$$

We now claim the following

Lemma 1 (Perturbation)**.** *Consider the transfer operator of both f and f_t acting on the space $C(a,\nu)$. Then*

$$\|\mathcal{L}\varphi - \mathcal{L}_t\varphi\| \lesssim \Delta^{-\gamma\nu}$$

on the norm of uniform convergence.

Proof. Note that the main difference here to the stochastic stability analysis performed in [24] is that the maps f_t are not chosen independently, which may lead to the non-existence of stationary measures. Since f is a local diffeomorphism, all the points $y \in M$ have a same number $k \geq 1$ of preimages, namely the degree of f. (In our case $k = 2$.) Moreover, given any preimage x of y, there exists a neighbourhood V of y and an inverse branch $g : V \to M$ such that $f \circ \phi = identity$ and $\phi(y) = x$. A local inverse branch g must be contracting

$$d(\phi(y), \phi(y')) \leq \sigma^{-1} d(y, y')$$

for every y, y' in V. Moreover, by compactness of M, there exists ρ_0 such that given $y_1, y_2 \in M$ with $d(y_1, y_2) \leq \rho_0$, one may write $f^{-1}(y_j) = \{x_{j1}, \cdots, x_{jk}\}$ for $j = 1, 2$ with

$$d(x_{1i}, x_{2i}) \leq \sigma^{-1} d(y_1, y_2)$$

for each $i = 1, \cdots, k$. Obviously

$$f_t^{-1}(y) = \{x_{1,t}, \cdots, x_{k,t}\}.$$

With these remarks, we can write the transfer operators as

$$(\mathcal{L}\varphi)(y) = \sum_{i=1}^{k} \frac{\varphi(x_i)}{|\mathrm{det} Df(x_i)|} \quad \text{and} \quad (\mathcal{L}_t\varphi)(y) = \sum_{i=1}^{k} \frac{\varphi(x_{i,t})}{|\mathrm{det} Df_t(x_{i,t})|}.$$

We proceed by obtaining bounds for the transfer operator $\mathcal{L}_t$. First notice that as f_t is $C^{1+\nu}$ and uniformly $\Delta^{-\gamma}$ close to f

$$\sup\{d(x_{i,t}, x_i) : 1 \leq i \leq k, y \in M\} \lesssim \Delta^{-\gamma}. \tag{25}$$

To obtain bounds on the Jacobian we first note that

$$\mathrm{det} Df_t(x_{i,t}) - \mathrm{det} Df(x_i) = \mathrm{det} Df_t(x_i) - \mathrm{det} Df(x_i) + \mathrm{det} Df_t(x_{i,t}) - \mathrm{det} Df_t(x_i). \tag{26}$$

We estimate the first difference in the right hand side of (26) as follows. Since $f_t = f + r$ is a uniformly $\Delta^{-\gamma}$ close to f. Moreover,

$$\begin{aligned}
|\text{det}Df(x_i) - \text{det}Df_t(x_i)| &= |\text{det}Df(x_i)| \left|1 - \det\left[Df_t(x_i)Df^{-1}(x_i)\right]\right| \\
&= |\text{det}Df(x_i)| \left|1 - \det\left[D(f+r)(x_i)Df^{-1}(x_i)\right]\right| \\
&= |\text{det}Df(x_i)| \left|1 - \det\left[\mathrm{I} + Dr(x_i)Df^{-1}(x_i)\right]\right|,
\end{aligned}$$

and

$$\det\left[\mathrm{I} + Dr(x_i)Df^{-1}(x_i)\right] = 1 + \text{tr}Dr(x_i)Df^{-1}(x_i).$$

Noting that expansivity implies that the derivative Df is an isomorphism at the every point, and in view of Proposition 3 we obtain

$$\left|\text{tr}\left[Dr(x_i)Df^{-1}(x_i)\right]\right| \lesssim \Delta^{-\gamma},$$

and therefore

$$|\text{det}Df(x_i) - \text{det}Df_t(x_i)| \lesssim \Delta^{-\gamma}\,. \tag{27}$$

To control the second difference in the right hand side of (26) we note that Df_t is $(a, \nu)-$Hölder continuous implying that the determinant is $\nu-$Hölder continuous

$$|\text{det}Df_t(x_{it}) - \text{det}Df_t(x_i)| < bd(x_{i,t}, x_i)^\nu,$$

yielding in view of (25)

$$|\text{det}Df_t(x_{it}) - \text{det}Df_t(x_i)| \lesssim \Delta^{-\gamma\nu}. \tag{28}$$

These estimates (27) and (28) for the right hand side terms of (26) yield the following estimate

$$|\text{det}Df_t(x_{it}) - \text{det}Df(x_i)| \lesssim \Delta^{-\gamma\nu}. \tag{29}$$

To obtain the result, it remains to control the test functions φ. Note that since $\log\varphi$ is (a, ν)-Hölder continuous we have

$$|\varphi(x_{i,t}) - \varphi(x_i)| \leq \varphi(x_i)e^{ad(x_{i,t},x_i)^\nu}$$

and since the manifold is compact φ will attain its maximum on M. Hence,

$$|\varphi(x_{i,t}) - \varphi(x_i)| \lesssim \Delta^{-\gamma\nu} \tag{30}$$

as $1 - e^{ad(x_{i,t},x_i)^\nu} \leq d(x_{i,t}, x_i)^\nu$, once, $ad(x_{i,t}, x_i)^\nu \leq 1$ as Δ is large. Altogether the estimates (29) and (30) imply that for all $y \in M$

$$\|\mathcal{L}_t\varphi - \mathcal{L}\varphi\|_0 \lesssim \Delta^{-\gamma\nu},$$

concluding the result. □

The above proposition reveals that at each step the transfer operator of the low degree nodes is uniformly close to the unperturbed operator. The natural question concerns the behavior of compositions of the transfer operators, corresponding to the time evolution. Hence, we introduce the transfer operator

$$\mathcal{L}^k_{\boldsymbol{t}} = \mathcal{L}_{t_k} \circ \mathcal{L}_{t_{k-1}} \circ \cdots \circ \mathcal{L}_{t_1}.$$

Our next result characterizes the action of $\mathcal{L}^k_{\boldsymbol{t}}$.

Proposition 4. *Let $\varphi \in C(a,\nu)$ then there exists $u = u(\varphi) > 0$ such that almost every network*

$$\mathcal{L}_{\boldsymbol{t}}^k \varphi = \varphi_0 + \tilde{\mu}_k$$

for all $k > u$ where $\varphi_0 \in C(a,\nu)$ is the fixed point of the unperturbed operator $\mathcal{L}$ and $\tilde{\mu}_k \in C(a,\nu)$, $k = 1, 2, \dots$ satisfy

$$\|\tilde{\mu}_k\|_0 \lesssim \Delta^{-\gamma\nu} \text{ for all } k > u.$$

Before we prove this Proposition, we need the following auxiliary result concerning the behavior of a family of operators near a contraction. Given a metric space (X, d), we denote the open ball

$$B(z,\delta) := \{x \in X : d(x,z) < \delta\}.$$

Our claim is as follows

Lemma 2. *Let (X, d) be a metric space and transformation $F : X \to X$ with Lipschitz constant $k < 1$ and a unique fixed point z. Let T be a metric space and consider a family of transformations $F_t : X \to X$, with $t \in T$, and introduce the composition*

$$F_{\boldsymbol{t}}^n = F_{t_n} \circ F_{t_{n-1}} \circ \cdots F_{t_1}.$$

Suppose that the transformations satisfy

$$\sup_{x\in X} d(F(x), F_t(x)) \le \varepsilon.$$

Then there exists $B = B\left(\frac{4\varepsilon}{1-k}, z\right) \subset X$ satisfying

i) for any $n > 0$

$$F_{\boldsymbol{t}}^n(B) \subset B;$$

ii) for all $x \in X$ there exists $n = n(x)$ such that

$$F_{\boldsymbol{t}}^n(x) \in B,$$

where n is uniformly bounded on any compact subset containing B.

Proof. Consider a fixed t and let $G = F_t$, then note that

$$\begin{aligned} d(G(x), G(y)) &= d(G(x), F(x)) + d(F(y), G(y)) + d(F(x), F(y)) \\ &\le 2\varepsilon + kd(x,y) \end{aligned}$$

and,

$$\begin{aligned} d(G(x), F(y)) &= d(G(x), G(y)) + d(F(y), G(y)) \\ &\le 3\varepsilon + kd(x,y). \end{aligned}$$

Now, consider a ball $x \in B(z,\delta)$. We claim that if $\delta = \frac{3\varepsilon}{1-k}$ then $F_{\boldsymbol{t}}^n(x) \in B(\delta, z)$. Indeed, consider the action of a generic element G

$$\begin{aligned} d(G(x), z) &= d(G(x), F(z)) \\ &\le 3\varepsilon + kd(x,z) \le \delta \end{aligned}$$

but since $x \in B(\delta, z)$ we have the following bound $\delta \geq 3\varepsilon/(1-k)$, and by induction the claim follows.

To prove the second claim, using the same ideas as before, we observe by induction that

$$d(F_{\boldsymbol{t}}^n, z) \leq 3\varepsilon\left(\sum_{i=0}^{j-1} k^i\right) + k^j d(F_{\boldsymbol{t}}^{n-j}, z).$$

Hence,

$$\begin{aligned} d(F_{\boldsymbol{t}}^n, z) &\leq 3\varepsilon\frac{1-k^n}{1-k} + k^n M \\ &= 3\varepsilon\frac{1}{1-k} + k^n\tilde{M} \end{aligned} \tag{31}$$

where $\tilde{M} = M - 3\varepsilon/(1-k)$. This concludes the second part. If x is contained in a compact subset C of X containing B then the claim on $n = n(x)$ can be made uniform on x follows from compactness arguments. $\square$

Now we are ready to prove the result on perturbations of the transfer operator, Proposition 4.

Proof of Proposition 4. Recall the discussion in Sec. 5.1: the convex cone $C(a, \nu)$ is endowed with the metric

$$\theta(\varphi_1, \varphi_2) = \log\frac{\alpha(\varphi_1, \varphi_2)}{\beta(\varphi_1, \varphi_2)}$$

where

$$\alpha(\varphi_1, \varphi_2) = \inf\left\{\frac{\varphi_1(x)}{\varphi_2(x)}, \frac{e^{ad^\nu(x,y)}\varphi_2(x) - \varphi_2(y)}{e^{ad^\nu(x,y)}\varphi_1(x) - \varphi_1(y)}\right\}$$

and β is given by a similar expression with $\sup$ replaced by $\inf$.

Let $\varphi \in C(a, \nu)$ and consider $\varphi_1(x) = (\mathcal{L}_t\varphi)(x)$ and $\varphi_2(x) = (\mathcal{L}\varphi)(x)$. For simplicity we write $\varphi_2(x) = \varphi_1(x) + \phi(x)$, where ϕ belongs to the cone as well, and by construction $\sup|\phi(x)| \lesssim \Delta^{\gamma\nu}$. Hence,

$$\left|\frac{\varphi_1(x)}{\varphi_2(x)} - 1\right| \lesssim \Delta^{\gamma\nu}.$$

Likewise,

$$\frac{e^{ad^\nu(x,y)}\varphi_2(x) - \varphi_2(y)}{e^{ad^\nu(x,y)}\varphi_1(x) - \varphi_1(y)} = 1 + \frac{\phi(y)}{\varphi_1(y)}\frac{e^{ad^\nu(x,y)}\frac{\phi(x)}{\phi(y)} - 1}{e^{ad^\nu(x,y)}\frac{\varphi_1(x)}{\varphi_1(y)} - 1} \tag{32}$$

but the last fraction in the right hand side is bounded, and therefore,

$$\left|\frac{e^{ad^\nu(x,y)}\varphi_2(x) - \varphi_2(y)}{e^{ad^\nu(x,y)}\varphi_1(x) - \varphi_1(y)} - 1\right| \lesssim \Delta^{-\gamma\nu}. \tag{33}$$

Therefore, we obtain

$$\theta(\varphi_1, \varphi_2) = |\log\left(1 + \psi(\varphi_1, \varphi_2)\right)| \lesssim \Delta^{-\gamma\nu}$$

as $\|\psi(\varphi_1, \varphi_2)\|_0 \lesssim \Delta^{-\gamma\nu}$. Proposition 2 guarantees that the transfer operator $\mathcal{L}$ is a contraction in the metric space $(C(a, \nu), \theta)$. Moreover, $\mathcal{L}$ and $\mathcal{L}_t$ are uniformly close. Hence, applying Lemma 2 in the metric space $(C(a, \nu), \theta)$ we conclude Proposition 4. $\square$

The expectation operator: As before we consider an absolutely continuous measure μ with density φ. Moreover, we assume $\log\varphi$ is $(a,\nu)-$Hölder continuous. Let $\mu^n : M^n \to [0,1]$ be a product measure, that is, given $A = A_1 \times \cdots \times A_n \subset M^n$ we have $\mu^n(A) = \mu(A_1)\cdots\mu(A_n)$. We define the expectation operator with respect to the measure μ

$$\mathbb{E}_\mu(\cdot) = \int_{M^n} \cdot \; d\mu^n \; .$$

The intuitive idea here is that the initial conditions of the low degree nodes are distributed according to this product measure. So at the initial time the systems are independent. Then with the evolution of the initial conditions, since the systems are interacting, the pushforward of this measure under the dynamics no longer has the product structure. However, since the interaction is mild a mean field reduction can be obtained.

Proposition 5 (Low degree Nodes)**.** *Consider the coupled maps (7). For almost every network in $\mathcal{G}(\boldsymbol{w})$, given function $\psi, \sigma \in E$ there exists $u = u(\mu) > 0$ and a point $v \in \mathbb{R}$ such that*

1. *For every $\ell < i \leq n$ and $t > u$*

$$\|\mathbb{E}_\mu(\psi(x_i(t))) - v\| \lesssim \Delta^{-\gamma\nu};$$

2. *for any $\ell < j, i \leq n$ and $t > u$*

$$\|\mathrm{Cov}_\mu(\psi(x_i(t)), \sigma(x_j(t)))\| \lesssim \Delta^{-\gamma\nu}.$$

Proof. Let $\pi^i : M^n \to M$ be a projector to the ith component, denoting $X = (x_1, \cdots, x_i, \cdots, x_n) \in M^n$ then $\pi^i X = x_i$. Now the following argument holds for any $\ell < i \leq n$, hence, for sake of simplicity we drop the i dependence. Given $X^0 \in M^n$, take

$$x^{k+1} = (f_{\boldsymbol{t}}^k \circ \pi)(X^0).$$

Note that

$$\mathbb{E}_\mu(\psi(x^k)) = \int_{M^n} \psi(x^k) d\mu^n = \int_{M^n} \psi \circ f_{\boldsymbol{t}}^k \circ \pi d\mu^n = \int_M \psi \circ f_{\boldsymbol{t}}^k d\mu$$

and notice that $\mu = \pi_* \mu^n$. Since $\mu = \varphi dm$ we may write

$$\mathbb{E}_\mu(\psi(x^k)) = \int_M \psi \circ f_{\boldsymbol{t}}^k d\mu = \int_M \psi(\mathcal{L}_{\boldsymbol{t}}^k \varphi) dm.$$

By Proposition 4 we obtain

$$\mathcal{L}_{\boldsymbol{t}}^k \varphi = \varphi_0 + \tilde{\mu}(t)$$

where $\varphi_0, \tilde{\mu}(t) \in C(a,\nu)$. Defining $v = \int_M \psi\varphi_0 dm$ the claim in the first part follows (for every low degree node).

To prove the second part we proceed in the same manner and obtain the desired estimates for the covariance. Note that

$$\begin{aligned} \mathrm{Cov}_\mu(\psi(x_i^k), \sigma(x_j^k)) &= \mathbb{E}_\mu\left[\langle \psi(x_i^k) - \mathbb{E}_\mu(\psi(x_i^k)), \sigma(x_j^k) - \mathbb{E}_\mu(\sigma(x_j^k))\rangle\right] \\ &= \mathbb{E}_\mu\left[\langle \psi(x_i^k), \sigma(x_j^k)\rangle\right] - \langle \mathbb{E}_\mu(\psi(x_i^k)), \mathbb{E}_\mu(\sigma(x_j^k))\rangle. \end{aligned}$$

Now we wish to estimate the $\mathbb{E}_\mu\langle\psi(x_i^k),\sigma(x_j^k)\rangle$. To this end we fix i and j with $\ell < i \neq j \leq n$, and introduce $p_{\boldsymbol{t}}^k : M \times M \to M \times M$ defined as

$$p^k(x_i^0, x_j^0) = (f_{\boldsymbol{t}}^k(x_i), f_{\tilde{\boldsymbol{t}}}^k(x_j)),$$

for the corresponding sequences $\boldsymbol{t}$ and $\tilde{\boldsymbol{t}}$, along with $(\psi,\sigma)\circ p^k = (\psi(x_i^k),\sigma(x_j^k))$. Consider the projector $\pi^{ij} : M^q \to M^2$ defined as $\pi^{ij}(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n) = (x_i,x_j)$. With this notation we may write

$$\mathbb{E}\langle\psi(x_i^k),\sigma(x_j^k)\rangle = \int_{M^n}\langle\psi(x_i^k),\sigma(x_j^k)\rangle d\mu^n \tag{34}$$

$$= \int_{M\times M}\langle\psi(x),\sigma(y)\rangle\mathcal{L}_p^k d\mu^2 \tag{35}$$

where $\pi_*^{ij}\mu = \mu^2$. Now since μ^n has a product structure and π^{ij} is a natural projector μ^2 has density $\varphi(x)\varphi(y)$. Note that h is in a neighborhood of a product map, and all the estimates hold uniformly, we obtain

$$(\mathcal{L}_p^k\varphi\times\varphi)(x,y) = \varphi_0(x)\varphi_0(y) + \bar{\mu}(x,y)$$

where again $\bar{\mu}$ satisfies

$$\|\bar{\mu}\|_0 \lesssim \Delta^{-\gamma\nu},$$

and the result follows. □

5.3 Homogeneity of the Mean Field

Before, proving the homogeneity of the mean field we need to control certain concentration properties of the graphs. The first is the concentration of the degrees $k_i = \sum_j A_{ij}$, which is given by Proposition 3. The second is related to

$$Y_i = \frac{1}{\Delta^2}\sum_{\ell<j,k\leq n} A_{ij}A_{ik}.$$

Next we show that this quantity is heavily concentrated at κ_i^2. For this we have the following proposition

Proposition 6. *Let $\mathcal{G}(\boldsymbol{w})$ satisfy the strong heterogeneity hypothesis. Then for $1 \leq i \leq \ell$, for every $0 < \delta < (1-\theta)/2$ we have*

$$\left|Y_i - \kappa_i^2\right| \lesssim \Delta^{(-1+\theta)/2+\delta}$$

for almost every network in $\mathcal{G}(\boldsymbol{w})$.

Proof. We obtain this claim by a Chebyshev inequality. For $1 \leq i \leq \ell$, for A_{ij} and A_{ik} are independent for $j \neq$ and note $A_{ij}^2 = A_{ij}$. Therefore, we need to estimate the expectation and variance of Y_i:

$$\mathbb{E}_{\boldsymbol{w}}(Y_i) = \frac{w_i}{\Delta^2}\rho\sum_{\ell<j\leq n} w_j + \frac{w_i^2}{\Delta^2}\rho^2\sum_{\ell<j\neq k\leq n} w_j w_k.$$

Note that

$$\rho^2\sum_{\ell<j\neq k\leq n} w_j w_k = \left(\rho\sum_{\ell<j\leq n} w_j\right)\left(\rho\sum_{\ell<k\leq n} w_k\right) - \rho^2\sum_{\ell<j\leq n} w_j^2,$$

along with $\rho^2 \sum_{\ell<j\le n} w_j^2 \le \rho\Delta^{1-\gamma} \le \Delta^{-1-\gamma}$, where in first inequality we used property N2 of the strong heterogeneity and in the last inequality we used the graphical condition (1). Moreover, $w_i = \kappa_i \Delta$, and denote $\hat{\kappa} = (\sum_{j=1}^{\ell} \kappa_j)/\ell$, clearly $\hat{\kappa} \in (0,1]$. Thus,

$$\rho \sum_{\ell<j\le n} w_j = 1 - \rho\ell\hat{\kappa}\Delta.$$

Using this together with $\Delta^2\rho \le 1$ and assumption N3 of the strong heterogeneity implies $\left|\rho \sum_{\ell<j\le n} w_i - 1\right| \lesssim \Delta^{-1+\theta}$ leading to

$$|\rho^2 \sum_{\ell<j\ne k\le n} w_j w_k - 1| \lesssim \Delta^{-1+\theta},$$

and we obtain

$$\left|\mathbb{E}_{\boldsymbol{w}}(Y_i) - \kappa_i^2\right| \lesssim \Delta^{-1+\theta}. \tag{36}$$

Now we wish to estimate,

$$\mathbb{E}_{\boldsymbol{w}}(Y_i^2) = \frac{w_i^4}{\Delta^4} \sum_{\ell<j,k,p,q\le n} \frac{w_j w_k w_p w_q}{\rho^4}.$$

Now by the same arguments as before we obtain

$$|\mathbb{E}_{\boldsymbol{w}}(Y_i^2) - \kappa_i^4| \lesssim \Delta^{-1+\theta}. \tag{37}$$

This bound together with (36) yields a bound for the variance

$$|\mathrm{Var}_{\boldsymbol{w}}(Y_i)| \lesssim \Delta^{-1+\theta}.$$

Applying the Chebyshev inequality we obtain that

$$\Pr_{\boldsymbol{w}}\left(|Y_i - \mathbb{E}_{\boldsymbol{w}}(Y_i)| \gtrsim \Delta^{-(1+\theta)/2+\delta}\right) \lesssim \Delta^{-2\delta}.$$

This implies that for almost every graph

$$|Y_i - \mathbb{E}_{\boldsymbol{w}}(Y_i)| \lesssim \Delta^{-(1+\theta)/2+\delta},$$

now using the triangle inequality and the bound 36 we obtain

$$|Y_i - \kappa_i^2| \lesssim \Delta^{-(1+\theta)/2+\delta},$$

and we obtain the result. □

After this auxiliary concentration result we are now ready to state

Proposition 7 (Homogeneity of the mean Field)**.** *Let the initial conditions of the low degree nodes be chosen independently and according to a measure μ^n as before. Let $\psi \in E$, then there exists $v \in \mathbb{R}$ such that given $\beta > 0$ small enough for almost every network in $\mathcal{G}(\boldsymbol{w})$ and μ^n−almost every initial condition for all $1 \le i \le \ell$*

$$\left|\frac{1}{\Delta}\sum_{j=1}^{n} A_{ij}\psi(x_j(t)) - \kappa_i v\right| \lesssim \kappa_i^{1/2}\Delta^{-\eta+\beta}$$

where $\eta > 0$ is determined by the network structure and the dynamics: indeed, *i)* $\theta < 1 - \gamma\nu/2$ *implies*

$$\eta = \gamma\nu/2;$$

ii-a) $1/2 > \theta > 1 - \gamma\nu/2$ *implies*

$$\eta = 1/2;$$

ii-b) $\theta > 1/2 > 1 - \gamma\nu/2$ *implies*

$$\eta = 1 - \theta + \frac{1}{2}\frac{\ln \kappa_i}{\ln \Delta}.$$

Proof. We wish to use a Chebyshev bound. Hence, we start estimating the mean and the variance. For a fixed i in the set of hubs, we need to estimate the expectation and the variance of the coupling term $\Delta^{-1}\sum_j A_{ij}\psi(x_j)$. The first follows easily

$$\mathbb{E}_\mu\left(\frac{1}{\Delta}\sum_j A_{ij}\psi(x_j)\right) = \frac{1}{\Delta}\sum_j A_{ij}\mathbb{E}_\mu(\psi(x_j)),$$

but from Proposition 5, we have $|\mathbb{E}_\mu\psi(x_j) - v| \lesssim \Delta^{-\gamma\nu}$. Also, in view of the concentration inequality Proposition 3 for *almost every* network

$$|k_i/\Delta - \kappa_i| \lesssim \kappa_i^{1/2}\Delta^{-1/2+\varepsilon}.$$

Hence,

$$\left|\mathbb{E}_\mu\left(\frac{1}{\Delta}\sum_j A_{ij}\psi(x_j)\right) - \kappa_i v\right| \lesssim \max\{k_i^{1/2}\Delta^{-1/2+\varepsilon}, \kappa_i\Delta^{-\gamma\nu}\}, \tag{38}$$

for almost every network. To estimate the variance we note that

$$\mathrm{Var}_\mu\left(\frac{1}{\Delta}\sum_j A_{ij}\psi(x_j)\right) = \frac{1}{\Delta^2}\sum_{j,k} A_{ij}A_{ik}[\mathbb{E}_\mu(\psi(x_i)\psi(x_j)) - \mathbb{E}_\mu(\psi(x_i))\mathbb{E}_\mu(\psi(x_j))].$$

We split the sum for indexes running over hub nodes $1 \leq j,k \leq \ell$ and low degree nodes $\ell < j,k \leq n$. For the hub indexes we have

$$\left|\frac{1}{\Delta^2}\sum_{1\leq j,k\leq \ell} A_{ij}A_{ik}[\mathbb{E}_\mu(\psi(x_i)\psi(x_j)) - \mathbb{E}_\mu(\psi(x_i))\mathbb{E}_\mu(\psi(x_j))]\right| \lesssim \Delta^{-2(1-\theta)}. \tag{39}$$

Now for the low degree indexes $\ell < j,k \leq n$, we have in view of Proposition 5

$$|\mathbb{E}_\mu(\psi(x_i)\psi(x_j)) - \mathbb{E}_\mu(\psi(x_i))\mathbb{E}_\mu(\psi(x_j))| \lesssim \Delta^{-\gamma\nu}$$

and hence,

$$\left|\frac{1}{\Delta^2}\sum_{1\leq j,k\leq \ell} A_{ij}A_{ik}[\mathbb{E}_\nu(\psi(x_i)\psi(x_j)) - \mathbb{E}_\nu(\psi(x_i))\mathbb{E}_\nu(\psi(x_j))]\right| \lesssim \Delta^{-\gamma\nu}Y_i. \tag{40}$$

Now, we claim good concentration properties for Y_i so that we can change its value by its expected value κ_i^2. Hence, combining (39) and (40) for *almost every* graph we obtain

$$\left|\mathrm{Var}_\mu\left(\frac{1}{\Delta}\sum_j A_{ij}\psi(x_j)\right)\right| \lesssim \max\{\kappa_i^2\Delta^{-\gamma\nu}, \Delta^{-2(1-\theta)}\},$$

therefore, the variance depends on a competition between the network structure parameters. We obtain the following cases:

Case i) $0 < \theta < 1 - \gamma\nu/2$: Notice that

$$\left| \mathrm{Var}_{\mu} \left(\frac{1}{\Delta} \sum_{j} A_{ij} \psi(x_j) \right) \right| \lesssim \kappa_i^2 \Delta^{-\gamma\nu}.$$

Applying Chebyshev inequality

$$\mathrm{Pr}_{\mu} \left(\left| \frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) - \mathbb{E}_{\mu} \left(\frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) \right) \right| \gtrsim \kappa_i \Delta^{-\gamma\nu/2+\beta} \right) \leq \Delta^{-2\beta}.$$

Therefore, we obtain that for almost every network in $\mathcal{G}(\boldsymbol{w})$ and $\mu-$almost every initial condition we have

$$\left| \frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) - \mathbb{E}_{\mu} \left(\frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) \right) \right| \lesssim \kappa_i \Delta^{-\gamma\nu/2+\beta}.$$

But in view of the triangle inequality we obtain

$$\left| \left| \mathbb{E}_{\mu} \left(\frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) \right) - \kappa_i v \right| - \left| \frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) - \kappa_i v \right| \right| \lesssim \kappa_i \Delta^{-\gamma\nu/2+\beta} \tag{41}$$

however, by (38) we obtain

$$\left| \frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) - \kappa_i v \right| \lesssim \max\{\kappa_i^{1/2} \Delta^{-1/2+\varepsilon}, \kappa_i \Delta^{-\gamma\nu/2+\beta}\}.$$

Notice that as $\kappa_i \in (0, 1]$ we have $\kappa_i < \kappa_i^{1/2}$. Now clearly, $\gamma\nu/2 < 1/2$ as $\nu \in (0, 1]$ and $\gamma < 1$. Moreover, the competition with the term $\log^{1/2} n$ can be absorbed in β. Combining the two upper bounds estimates we obtain

$$\max\{\kappa_i^{1/2} \Delta^{-1/2+\varepsilon}, \kappa_i \Delta^{-\gamma\nu/2+\beta}\} \lesssim \kappa_i^{1/2} \Delta^{-\gamma\nu/2+\beta}$$

and our first claim follows.

Case ii) $\theta > 1 - \gamma\nu/2$. We obtain

$$\left| \mathrm{Var}_{\mu} \left(\frac{1}{\Delta} \sum_{j} A_{ij} \psi(x_j) \right) \right| \lesssim \Delta^{-2(1-\theta)}.$$

Applying the Chebyshev inequality we obtain an inequality similar to (41) with the right hand side replaced by $\Delta^{-(1-\theta)+\beta}$. Hence, we obtain

$$\left| \frac{1}{\Delta} \sum_{j=1}^{n} A_{ij} \psi(x_j(t)) - \kappa_i v \right| \lesssim \max\{\kappa_i^{1/2} \Delta^{-1/2+\varepsilon}, \Delta^{-(1-\theta)+\beta}\}$$

as the condition $\theta > 1 - \gamma\nu/2$ implies $(1 - \theta) < \gamma\nu$. Here we distinguish two cases:

Case ii-a) $1/2 > \theta > 1 - \gamma\nu/2$, which implies

$$\max\{\kappa_i^{1/2} \Delta^{-1/2+\varepsilon}, \Delta^{-(1-\theta)+\beta}\} \lesssim \kappa_i^{1/2} \Delta^{-1/2+\varepsilon}$$

and

Case ii-b) $\theta > 1/2 > 1 - \gamma\nu/2$

$$\max\{\kappa_i^{1/2} \Delta^{-1/2+\varepsilon}, \Delta^{-(1-\theta)+\beta}\} \lesssim \Delta^{-(1-\theta)+\beta},$$

and we conclude the result. $\square$

5.4 Proof of Theorem 1

The dynamics of the low degree nodes has been characterized in Proposition 5. To prove our main result we use the Homogeneity of the mean field Proposition 7

Proof. Consider the dynamics of the high degree nodes

$$x_i(t+1) = f(x_i(t)) + \alpha r_i(x_i, y_i), \ \ i = 1, \dots, \ell$$

where the coupling term r_i reads

$$r_i(x_i, y_i) = \sum_p u_p(x_i) \left(\sum_q y_{i,q} \right) \tag{42}$$

with

$$y_{q,i} = \frac{1}{\Delta} \sum_j A_{ij} v_q(x_j),$$

as before. By Proposition 7 on the homogeneity of the mean field, we obtain

$$y_{q,i} = \kappa_i \langle v_q \rangle + \xi_{q,i} \tag{43}$$

for almost all networks in $\mathcal{G}(\boldsymbol{w})$, with $\langle v_q \rangle = \int_M v_q(x) \mu_0 dm$ and

$$|\xi_i| \lesssim \kappa_i^{1/2} \Delta^{-\gamma\nu/2+\beta}, \ \text{for } i = 1, 2, \dots, \ell. \tag{44}$$

Using (42) and (43) we obtain

$$r_i(x_i, y_i) = \kappa_i g(x) + \zeta_i$$

where

$$g(x) = \sum_{p,q} u_p(x) \langle v_q \rangle = \int h(x,y) \mu_0(dy)$$

and $\zeta_i = \sum_{p,q} u_p(x_i) \xi_{q,i}$. Since the functions u_a is continuous and the manifold is compact and because of (44) it follows that

$$x_i(t+1) = f(x_i(t)) + \alpha \kappa_i g(x_i) + \alpha \zeta_i(t), \ \text{for each } i = 1, \dots, \ell,$$

has the properties claimed in the Main Theorem. □

Acknowledgments: This work was partially supported by FP7 IIF Research Fellowship – project number 303180, EU Marie-Curie IRSES Brazilian-European partnership in Dynamical Systems (FP7-PEOPLE-2012-IRSES 318999 BREUDS), and the Brazilian agency FAPESP.

References

[1] K. Kaneko, Chaos 2, 279 (1992).

[2] G. Keller, Progress in Probability Volume 46, 183-208 (2000).

[3] L.A. Bunimovich and Ya. G. Sinai, Nonlinearity 1, 491 (1988).

[4] J. Bricmont, A. Kupiainen, Coupled analytic maps, Nonlinearity 8, 379–396 (1995).

[5] M. Jiang, Nonlinearity 8, 631–659 (1995).

[6] T. Fischer and H.H. Rugh, Ergodic Theory Dynam. Systems 20, 109–143 (2000).

[7] E. Järvenpää, M. Järvenpää, Comm. Math. Phys. 220, 1–12, (2001).

[8] G. Keller and C. Liverani, Discrete Contin. Dyn. Syst. 11, 325–335 (2004).

[9] G. Keller and C. Liverani, Comm. Math. Phys. 262, 33–50 (2006).

[10] J.-R. Chazottes and B. Fernandez (eds.), *Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems*, Lect. Notes Phys. 671 (Springer, Berlin Heidelberg 2005).

[11] J. Koiller and L-S. Young, Nonlinearity 23, 1121–1141 (2010)

[12] R. Albert and A.-L. Barabsi, Rev. Mod. Phys. **74**, 47 (2002).

[13] M. E. J. Newman, *Networks: An Introduction*, Oxford University Press (2010).

[14] C. Zhou and J. Kurths, Chaos **16**, 015104 (2006);

[15] J. Gómez-Gardeñes, Y. Moreno and A. Arenas, Phys. Rev. Lett. **98**, 034101 (2007).

[16] T. Pereira, Phys. Rev. E **82**, 03674 (2010).

[17] M. S. Baptista, *et al.*, PLoS ONE 7(11): e48118.

[18] P. Bonifazi, M. Goldin, M. A. Picardo, *et al.* Science **326**, 1419 (2009).

[19] R. J. Morgan and I. Soltesz, Proc. Natl. Acad. Sci. USA **105**, 6179 (2008).

[20] F. R. K. Chung and L. Lu, *Complex Graphs and Networks*, American Mathematical Society (2006).

[21] B. Bollobás, *Random Graphs*, Cambridge University Press; 2 edition (2001).

[22] F. Chung, L. Lu and V. Vu, Annals. Comb. **7**, 21 (2003).

[23] W. de Melo and S. van Strien, *One-Dimensional Dynamics*, Springer (1993).

[24] M. Viana, *Stochastic dynamics of deterministic systems* , IMPA (1997).