J. Eur. Math. Soc. 22, 2183–2252

DOI 10.4171/JEMS/963

Tiago Pereira · Sebastian van Strien · Matteo Tanzi

Heterogeneously coupled maps: hub dynamics and emergence across connectivity layers

Received April 13, 2017

Abstract. The aim of this paper is to rigorously study the dynamics of Heterogeneously Coupled Maps (HCM). Such systems are determined by a network with heterogeneous degrees. Some nodes, called hubs, are very well connected while most nodes interact with few others. The local dynamics on each node is chaotic, coupled with other nodes according to the network structure. Such high-dimensional systems are hard to understand in full, nevertheless we are able to describe the system over exponentially large time scales. In particular, we show that the dynamics of hub nodes can be very well approximated by a low-dimensional system. This allows us to establish the emergence of macroscopic behaviour such as coherence of dynamics among hubs of the same connectivity layer (i.e. with the same number of connections), and chaotic behaviour of the poorly connected nodes. The HCM we study provide a paradigm to explain why and how the dynamics of the network can change across layers.

Keywords. Coupled maps, ergodic theory, heterogeneous networks

Contents

T. Pereira: Institute of Mathematical and Computer Sciences, Universidade de Saõ Paulo, Saõ Carlos 13566-590, Saõ Paulo, Brazil; e-mail: tiagophysics@gmail.com

S. van Strien: Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK; e-mail: s.van-strien@imperial.ac.uk

M. Tanzi: Mathematics and Statistics, University of Victoria, PO Box 1700 STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: matteotanzi@hotmail.it

Mathematics Subject Classification (2020): Primary 37A30, 37C30, 37C40, 37D20, 37Nxx; Secondary O5C80

1. Introduction

Natural and artificial complex systems are often modelled as distinct units interacting on a network. Typically such networks have a heterogeneous structure characterised by different scales of connectivity [AB02]. Some nodes called *hubs* are highly connected while the remaining nodes have only a small number of connections (see Figure 1 for an illustration). Hubs provide a short pathway between nodes making the network well connected and resilient and play a crucial role in the description and understanding of complex networks.

In the brain, for example, hub neurons are able to synchronize while other neurons remain out of synchrony. This particular behaviour shapes the network dynamics towards a healthy state [BG+09]. Surprisingly, disrupting synchronization between hubs can lead to malfunction of the brain. The fundamental dynamical role of hub nodes is not restricted to neuroscience, but is found in the study of epidemics [PSV01], power grids [MM+13], and many other fields.

Large scale simulations of networks suggest that the mere presence of hubs hinders global collective properties. That is, when the heterogeneity in the degrees of the network is strong, complete synchronization is observed to be unstable [NM+03]. However, in certain situations hubs can undergo a transition to collective dynamics [GMA07, Per10, BR+12]. Despite the large amount of recent work, a mathematical understanding of dynamical properties of such networks remains elusive.

In this paper, we introduce the concept of Heterogeneously Coupled Maps (HCM for short), where the heterogeneity comes from the network structure modelling the interaction. HCM describes the class of problems discussed above incorporating the non-linear and extremely high-dimensional behaviour observed in these networks. High-dimensional systems are notoriously difficult to understand. HCM is no exception. Here, our approach is to describe the dynamics at the expense of an arbitrarily small, but fixed fluctuation, over exponentially large time scales. In summary, we obtain:

(i) *Dimensional reduction for hubs for finite time.* Fixing a given accuracy, we can describe the dynamics of the hubs by a low-dimensional model for a finite time T. The true dynamics of a hub and its low-dimensional approximation are the same up to the given accuracy. The time T for which the reduction is valid is exponentially large in the network size. For example, in the case of a star network (see Section 3.1), we can describe the hubs with 1% accuracy in networks with 10^6 nodes for a time up to roughly $T = e^{30}$ for a set of initial conditions of measure roughly $1 - e^{-10}$. This is arguably the only behaviour one will ever see in practice.

(ii) *Emergent dynamics changes across connectivity levels*. The dynamics of hubs can drastically change depending on the degree and synchronization (or more generally phase locking) naturally emerges between hub nodes. This synchronization is not due to a direct mutual interaction between hubs (as in the usual "Huygens" synchronization) but results from the common environment that the hub nodes experience.

Before presenting the general setting and precise statements in Section 2, we informally discuss these results and illustrate the rich dynamics that emerges in HCM due to heterogeneity.

1.1. *Emergent dynamics on HCM*

Figure 1 is a schematic representation of a heterogeneous network with three different types of nodes: massively connected hubs (on top), moderately connected hubs having half as many connections of the previous ones (in the middle), and low degree nodes (at the bottom). Each one of the three types constitutes a connectivity layer, meaning a subset of the nodes in the network having approximately the same degree. When uncoupled, each node is identical and supports chaotic dynamics. Adding the coupling, different behaviour can emerge for the three types of nodes. In fact, we will show examples where the dynamics of the hub at the top approximately follows a periodic motion, the hub in the middle stays near a fixed point, and the nodes at the bottom remain chaotic. Moreover, this behaviour persists for exponentially large time in the size of the network, and it is robust under small perturbations.

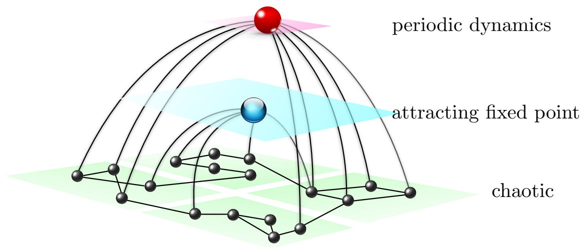

Fig. 1. The dynamics across connectivity layers change depending on the connectivity of the hubs. We will exhibit an example where the hubs with the highest number of connections (in red, at the top) have periodic dynamics. In the second connectivity layer, where hubs have half of the number of connections (in blue, in the middle), the dynamics sits around a fixed point. In the bottom layer of poorly connected nodes the dynamics is chaotic. (Only one hub has been drawn on the top two layers for clarity of the picture).

Synchronization because of common environment. Our theory uncovers the mechanism responsible for high correlations among the hubs states, which is observed in experimental and numerical observations. The mechanism turns out to be different from synchronization (or phase locking) due to mutual interaction, i.e. different from "Huygens" synchronization. In HCM, hubs display highly correlated behaviour even in the absence of direct connections between them. The poorly connected layer consisting of a huge number of weakly connected nodes plays the role of a kind of "heat bath" providing a common forcing to the hubs which is responsible for the emergence of coherence.

1.2. Hub synchronization and informal statement of Theorem A

The model. *A network of coupled dynamical systems* is the datum (G, f, h, α), where G is a labelled graph with the set $\mathcal{N} = \{1, \ldots, N\}$ of nodes, $f : \mathbb{T} \to \mathbb{T}$ is the local dynamics at each node of the graph, $h \colon \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ is a coupling function that describes pairwise interaction between nodes, and $\alpha \in \mathbb{R}$ is the coupling strength. We take f to be a Bernoulli map, $z \mapsto \sigma z \bmod 1$ for some integer $\sigma > 1$. This is in agreement with the observation that the local dynamics is chaotic in many applications [Izh07, WAH88, SS+01]. The graph G can be represented by its adjacency matrix $A = (A_{in})$ which determines the connections among nodes of the graph. If $A_{in} = 1$, then there is a directed edge of the graph going from n to i; and $A_{in} = 0$ otherwise. The degree $d_i := \sum_{n=1}^{N} A_{in}$ is the number of incoming edges at i. For simplicity, in this introductory section we consider undirected graphs (A is symmetric), unless otherwise specified, but our results hold in greater generality (see Section 2).

The dynamics on the network is described by

$$z_i(t+1) = f(z_i(t)) + \frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{in} h(z_i(t), z_n(t)) \bmod 1 \qquad \text{for } i = 1, \ldots, N. \tag{1}$$

In the above equations, Δ is a structural parameter of the network equal to the maximum degree. Rescaling the coupling strength in (1) by dividing by Δ allows us to scope the parameter regime for which interactions contribute with an order one term to the evolution of the hubs.

For the type of graphs we will be considering, the degrees d_i of nodes $1, \ldots, L$ are much smaller than the incoming degrees of nodes $L+1, \ldots, N$. A prototypical sequence of heterogeneous degrees is

$$\mathbf{d}(N) = (\underbrace{d, \ldots, d}_{L}, \underbrace{\kappa_m \Delta, \ldots, \kappa_m \Delta}_{M_m}, \ldots, \underbrace{\kappa_2 \Delta, \ldots, \kappa_2 \Delta}_{M_2}, \underbrace{\Delta, \ldots, \Delta}_{M_1}). \tag{2}$$

with $\kappa_m < \cdots < \kappa_2 < 1$ fixed and d/Δ small when N is large; we will refer to blocks of nodes corresponding to $(\kappa_i \Delta, \ldots, \kappa_i \Delta)$ as the *i-th connectivity layer* of the network, and to a graph G having a sequence of degrees prescribed by (2) as a *layered heterogeneous graph*. (We will make all this more precise below.)

It is a consequence of stochastic stability of uniformly expanding maps that for very small coupling strengths, the network dynamics will remain chaotic. That is, there is an $\alpha_0 > 0$ such that for all $0 \leq \alpha < \alpha_0$ and any large N, the system will preserve an ergodic absolutely continuous invariant measure [KL06]. When α increases, one reaches a regime where the less connected nodes still feel a small contribution coming from interactions, while the hub nodes receive an order one perturbation. In this situation, uniform hyperbolicity and the absolutely continuous invariant measure do not persist in general.

The low-dimensional approximation for the hubs. Given a hub $i_j \in \mathcal{N}$ in the i-th connectivity layer, our result gives a one-dimensional approximation of its dynamics in terms of f, h, α and the connectivity κ_i of the layer. The idea is the following. Let $z_1, \ldots, z_N \in \mathbb{T}$ be the states of the nodes, and assume that these N points are spatially distributed in $\mathbb{T}$ approximately according to the invariant measure m of the local map f (in this case the Lebesgue measure on $\mathbb{T}$). Then the coupling term in (1) is a *mean field* (Monte-Carlo) approximation of the corresponding integral:

$$\frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{i_j n} h(z_{i_j}, z_n) \approx \alpha \kappa_i \int h(z_{i_j}, y)\, dm(y) \tag{3}$$

where d_{i_j} is the incoming degree at i_j and $\kappa_i := d_{i_j}/\Delta$ is its normalized incoming degree. The parameter κ_i determines the effective coupling strength. Hence, the right hand side of (1) at the node i_j is approximately equal to the *reduced* map

$$g_{i_j}(z_{i_j}) := f(z_{i_j}) + \alpha \kappa_i \int h(z_{i_j}, y)\, dm(y), \tag{4}$$

Equations (3) and (4) clearly show the "heat bath" effect that the common environment has on the highly connected nodes.

Ergodicity ensures the persistence of the heat bath role of the low degree nodes. It turns out that the joint behaviour at poorly connected nodes is essentially ergodic. This will imply that at each moment of time the cumulative average effect on hub nodes is predictable and far from negligible. In this way, the low degree nodes play the role of a heat bath providing a sustained forcing to the hubs.

Theorem A below makes this idea rigorous for a suitable class of networks. We state the result precisely and in full generality in Section 2. For the moment assume that the number of hubs is small, does not depend on the total number N of nodes, and that the degree of the poorly connected nodes is relatively small, namely only a logarithmic function of N. For these networks our theorem implies the following

Theorem A (Informal statement in special case). *Consider the dynamics (1) on a layered heterogeneous graph. If the degrees of the hubs are sufficiently large, namely $\Delta = O(N^{1/2+\varepsilon})$, and the reduced dynamics g_j are hyperbolic, then for any hub j,*

$$z_j(t+1) = g_j(z_j(t)) + \xi_j(t),$$

where the size of the fluctuations $\xi_j(t)$ is below any fixed threshold for $0 \leq t \leq T$, with T exponentially large in Δ, and any initial condition outside a subset of measure exponentially small in Δ.

Hub synchronization mechanism. When $\xi_j(t)$ is small and g_j has an attracting periodic orbit, then $z_j(t)$ will be close to this attracting orbit after a short time and it will remain close to the orbit for an exponentially large time T. As a consequence, if two hubs have approximately the same degree d_j, even if they share no common neighbour, they feel the same mean effect from the "heat bath" and so they appear to be attracted to the same periodic orbit (modulo small fluctuations) exhibiting highly coherent behaviour.

The dimensional reduction provided in Theorem A is robust, persisting under small perturbation of the dynamics f, of the coupling function h and under addition of small independent noise. Our results show that the fluctuations $\xi(t)$, as functions of the initial condition, are small in the C^0 norm on most of the phase space, but notice that they can be very large in the C^1 norm. Moreover, they are correlated, and with probability one, $\xi(t)$ will be large for some $t > T$.

Idea of the proof. The proof of this theorem consists of two steps. Redefining *ad hoc* the system in the region of phase space where fluctuations are above a chosen small threshold, we obtain a system which exhibits good hyperbolic properties that we state in terms of invariant cone fields of expanding and contracting directions. We then show that the set of initial conditions for which the fluctuations remain below this small threshold up to time T is large, where T is estimated as in the above informal statement of the theorem.

1.3. Dynamics across connectivity scales: predictions and experiments

In the setting above, consider $f(z) = 2z \mod 1$ and the following simple coupling function:

$$h(z_i, z_n) = -\sin(2\pi z_i) + \sin(2\pi z_n). \tag{5}$$

Since $\int_0^1 \sin(2\pi y)\, dy = 0$, the reduced equation (see (4)) becomes

$$g_j(z_j) = T_{\alpha\kappa_j}(z_j) \quad \text{where} \quad T_\beta(z) = 2z - \beta \sin(2\pi z) \mod 1. \tag{6}$$

A bifurcation analysis shows that for $\beta \in I_E := [0, 1/2\pi)$ the map is globally expanding, while for $\beta \in I_F := (1/2\pi, 3/2\pi)$ it has an attracting fixed point at $y = 0$. Moreover, for

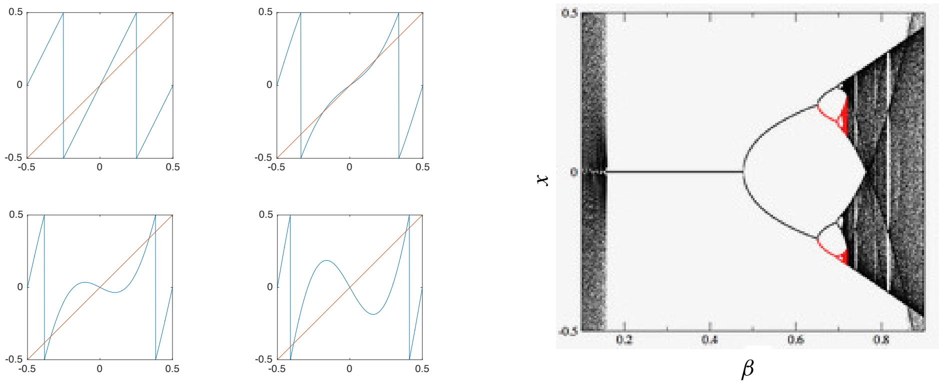

Fig. 2. Left: the graphs of T_β for $\beta = 0, 0.2, 0.4, 0.6$. Right: the bifurcation diagram for the reduced dynamics of hubs. We considered the identification $\mathbb{T} = [-1/2, 1/2]/\sim$. We obtained the diagram numerically. To build the bifurcation diagram we reported a piece of a typical orbit of length 10^3, for a collection of values of the parameter β.

$\beta \in I_p := (3/2\pi, 4/2\pi]$ it has an attracting periodic orbit of period 2. In fact, it follows from a recent result in [RvS15] that the set of parameters β for which T_β is hyperbolic, as specified by Definition 2.1 below, is open and dense. (See Propositions E.2 and E.3 in the Appendix for a rigorous treatment). Figure 2 shows the graphs and bifurcation diagram of T_β for varying β.

1.3.1. Predicted impact of the network structure. To illustrate the impact of the structure, we fix the coupling strength $\alpha = 0.6$ and consider a heterogeneous network with four levels of connectivity including three types of hubs and poorly connected nodes. The first highly connected hubs have $\kappa_1 = 1$. In the second layer, hubs have half of the number of connections of the first layer, $\kappa_2 = 1/2$. And finally, in the last layer, hubs have one fourth of the connections of the main hub, $\kappa_3 = 1/4$. The parameter $\beta_j = \alpha\kappa_j$ determines the effective coupling, and so for the three levels $j = 1, 2, 3$ we predict different types of dynamics by looking at the bifurcation diagram. The predictions are summarised in Table 1.

Table 1. Dynamics across connectivity scales.

Connectivity layer	Effective coupling β	Dynamics
hubs with $\kappa_1 = 1$	0.6	periodic
hubs with $\kappa_2 = 1/2$	0.3	fixed point
hubs with $\kappa_3 = 1/4$	0.15	uniformly expanding

1.3.2. Impact of the network structure in numerical simulations of large-scale layered random networks. We have considered the above situation in numerical simulations where we took a layered random network, described in (2) above, with $N = 10^5$,

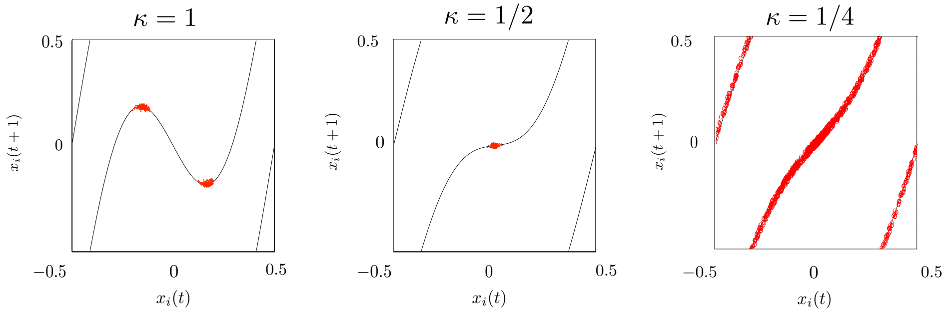

Fig. 3. Simulation results of the dynamics of a layered graph with two layers of hubs. We plot the return maps $z_i(t) \times z_i(t+1)$. The solid line is the low-dimensional approximation of the hub dynamics given by (6). The red circles are points taken from the hub time series. In the first layers of hubs ($\kappa = 1$) we observe a dynamics very close to the periodic orbit predicted by g_1, in the second layer ($\kappa = 1/2$) the dynamics of the hubs stays near a fixed point, and in the third layer ($\kappa = 1/4$) the dynamics is still uniformly expanding.

$\Delta = 500$, $w = 20$, $m = 2$, $M_1 = M_2 = 20$, $\kappa = 1$ and $\kappa_2 = 1/2$. The layer with highest connectivity is made up of 20 hubs connected to 500 nodes, and the second layer is made up of 20 hubs connected to 250 nodes. The local dynamics is again given by $f(z) = 2z \mod 1$, and the coupling is as in (5). We fixed the coupling strength at $\alpha = 0.6$ as in Section 1.3.1 so that Table 1 summarises the theoretically predicted dynamical behaviour for the two layers. We chose initial conditions for each of the N nodes independently and according to the Lebesgue measure. Then we evolved this 10^5-dimensional system for 10^6 iterations. Discarding the 10^6 initial iterations as transients, we plotted the next 300 iterations. The result is shown in Figure 3. In fact, we found essentially the same picture when we only plotted the first 300 iterations, with the difference that the first 10 iterates or so are not yet in the immediate basin of the periodic attractors. The simulated dynamics in Figure 3 is in excellent agreement with the predictions of Table 1.

The result above has also important implications in the study of the inverse problem of recovering the network structure from observations of its dynamics only. In [ET$^+$19], the authors describe an approach inspired by Theorem A to reconstruct information on the degree distribution, community structures, local dynamics and interaction function from the time series recording the state of each node.

1.4. Impact of network structure on dynamics: Theorems B and C

The importance of network structure in shaping the dynamics has been highlighted by many studies [GS06, AA$^+$11, NRS16] where network topology and its symmetries shape bifurcation patterns and synchronization spaces. Here we continue with this philosophy and show the dynamical feature that are to be expected in HCM. In particular, if we fix the local dynamics and the coupling, then the network structure dictates the resulting dynamics. In fact, we show that

- *there is an open set of coupling functions such that homogeneous networks globally synchronize but heterogeneous networks do not. However, in heterogeneous networks, hubs can undergo a transition to coherent behaviour.*

In Subsection 2, this claim is given a rigorous formulation in Theorems B and C.

1.4.1. Informal statement of Theorem B on coherence of hub dynamics. Consider a graph G with sequence of degrees given by (2) with $M := \sum_{k=1}^{m} M_k$, each M_i being the number of nodes in the i-th connectivity layer. Assume

$$\Delta = \mathcal{O}(N^{1/2+\varepsilon}), \qquad M = \mathcal{O}(\log N) \quad \text{and} \quad d = \mathcal{O}(\log N), \tag{7}$$

which implies that $L \approx N$ when N is large. Suppose that $f(x) = 2x \bmod 1$ and that $h(z_i, z_n)$ is as in (5).

Theorem B (Informal statement in special case). *For every connectivity layer i and hub node i_j in this layer, there exists an interval $I \subset \mathbb{R}$ of coupling strengths such that for any $\alpha \in I$, the reduced dynamics $T_{\alpha\kappa_i}$ (see (6)) has at most two periodic attractors $\{\overline{z}(t)\}_{t=1}^{p}$ and $\{-\overline{z}(t)\}_{t=1}^{p}$ and there are $s \in \{\pm 1\}$ and $t_0 \in [p-1]$ such that*

$$\operatorname{dist}(z_{i_j}(t+t_0), s\overline{z}(t \bmod p)) \leq \xi$$

for $1/\xi \leq t \leq T$, with T exponentially large in Δ, and for any initial condition outside a set of small measure.

Note that in order to have $1/\xi \ll T$ one needs Δ to be large. Theorem B proves that one can generically tune the coupling strength or the hub connectivity so that the hub dynamics follow, after an initial transient, a periodic orbit.

1.4.2. Informal statement of Theorem C comparing dynamics on homogeneous and heterogeneous networks

Erdős–Rényi model for homogeneous graphs. In contrast to layered graphs which are prototypes of heterogeneous networks, the classical Erdős–Rényi model is a prototype of a homogeneous random graph. By homogeneous, we mean that the expected degrees of the nodes are the same. This model defines an undirected random graph where each link in the graph is a Bernoulli random variable with the same success probability p (see Definition 2.3 for more details). We choose $p > (\log N)/N$ so that in the limit as $N \to \infty$ almost every random graph is connected (see [Bol01]).

Diffusive coupling functions. The coupling functions satisfying

$$h(z_i, z_j) = -h(z_j, z_i) \quad \text{and} \quad h(z, z) = 0$$

are called *diffusive*. The function h is sometimes required to satisfy $\partial_1 h(z, z) > 0$ to ensure that the coupling has an “attractive” nature. Even if this is not necessary for our computations, the examples in this paper satisfy this assumption. For each network G, we consider the corresponding system of coupled maps defined by (1). In this case the subspace

$$\mathcal{S} := \{(z_1, \ldots, z_N) \in \mathbb{T}^{\mathbb{N}} : z_1 = \cdots = z_N\} \tag{8}$$

is invariant. $\mathcal{S}$ is called the *synchronization manifold* on which all nodes of the network follow the same orbit. Fixing the local dynamics f and the coupling function h, we obtain the following dichotomy of stability and instability of synchronization depending on whether the graph is homogeneous or heterogeneous.

Theorem C (Informal statement).

(a) *Take a diffusive coupling function* $h(z_i, z_j) = \varphi(z_j - z_i)$ *with* $\frac{d\varphi}{dx}(0) \neq 0$. *Then for almost every asymptotically large Erdős–Rényi graph and any diffusive coupling function in a sufficiently small neighbourhood of* h *there is an interval* $I \subset \mathbb{R}$ *of coupling strengths for which* $\mathcal{S}$ *is stable* (*normally attracting*).

(b) *For any diffusive coupling function* $h(x, y)$, *and for any sufficiently large heterogeneous layered graph* G *with sequence of degrees satisfying* (2) *and* (7), $\mathcal{S}$ *is unstable.*

Example 1.1. Take $f(z) = 2z \bmod 1$ and

$$h(z_i, z_j) = \sin(2\pi z_j - 2\pi z_i) + \sin(2\pi z_j) - \sin(2\pi z_i).$$

It follows from the proof of Theorem C(a) that almost every asymptotically large Erdős–Rényi graph has a stable synchronization manifold for some values of the coupling strength ($\alpha \sim 0.3$) while any sufficiently large layered heterogeneous graph has no stable synchronized orbit. However, in a layered graph G the reduced dynamics for a hub node in the i-th layer is

$$\begin{aligned} g_{i_j}(z_{i_j}) &= 2z_{i_j} + \alpha\kappa_i \int [\sin(2\pi y - 2\pi z_{i_j}) + \sin(2\pi y) - \sin(2\pi z_{i_j})]\, dm(y) \bmod 1 \\ &= 2z_{i_j} - \alpha\kappa_i \sin(2\pi z_{i_j}) \bmod 1 \\ &= T_{\alpha\kappa_i}(z_{i_j}). \end{aligned}$$

By Theorem B there is an interval for the coupling strength ($\alpha\kappa_i \sim 0.3$) for which g_{i_j} has an attracting periodic sink and the orbits of the hubs in the layer follow this periodic orbit (modulo small fluctuations) exhibiting coherent behaviour.

2. Setting and statement of the main theorems

Let us consider a directed graph G whose set of nodes is $\mathcal{N} = \{1, \ldots, N\}$ and set of directed edges $\mathrm{E} \subset \mathcal{N} \times \mathcal{N}$. In this paper we will be only concerned with in-degrees of a node, i.e. the number of edges that point to that node (which counts the contributions to the interaction felt by that node). Furthermore we suppose, in a sense that will be later specified, that the in-degrees $d_1, \ldots, d_L$ of the nodes $\{1, \ldots, L\}$ are low compared to the size of the network, while the in-degrees $d_{L+1}, \ldots, d_N$ are comparable to the size of the network. For this reason, the first L nodes will be called *low degree nodes* and the remaining $M = N - L$ nodes will be called *hubs*. Let $A = (A_{in})_{1 \le i,n \le N}$ be the adjacency matrix of G, with entry A_{ij} equal to 1 if an edge going from node j to node i is present, and 0 otherwise. So $d_i = \sum_{j=1}^{N} A_{ij}$. The important *structural parameters* of the network are:

- L, M, the number of low degree nodes, resp. hubs; $N = L + M$, the total number of nodes;
- $\Delta := \max_i d_i$, the maximum in-degree of hubs;
- $\delta := \max_{1<i\leq L} d_i$, the maximum in-degree of low degree nodes.

The building blocks of the dynamics are:

- the *local dynamics*, $f : \mathbb{T} \to \mathbb{T}$, $f(x) = \sigma x \bmod 1$, for some integer $\sigma \geq 2$;
- the *coupling function*, $h : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, which we assume is C^{10};
- the *coupling strength*, $\alpha \in \mathbb{R}$.

We require the coupling to be C^{10} to ensure sufficiently fast decay of the Fourier coefficients of h. This will be useful in Appendix A. In the coordinates $z = (z_1, \ldots, z_N) \in \mathbb{T}^N$, the discrete-time evolution is given by a map $F : \mathbb{T}^N \to \mathbb{T}^N$ defined by $z' := F(z)$ with

$$z_i' = f(z_i) + \frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{in} h(z_i, z_n) \bmod 1, \qquad i = 1, \ldots, N. \tag{9}$$

Our main result shows that low and high degree nodes will develop different dynamics when α is not too small. To simplify the formulation of our main theorem, we write $z = (x, y)$ with $x = (x_1, \ldots, x_L) := (z_1, \ldots, z_L) \in \mathbb{T}^L$ and $y = (y_1, \ldots, y_M) := (z_{L+1}, \ldots, z_N) \in \mathbb{T}^M$. Moreover, decompose

$$A = \begin{pmatrix} A^{ll} & A^{lh} \\ A^{hl} & A^{hh} \end{pmatrix}$$

where A^{ll} is an $L \times L$ matrix, etc. Also write $A^l = (A^{ll}\ A^{lh})$ and $A^h = (A^{hl}\ A^{hh})$. In this notation we can write the map F as

$$x_i' = f(x_i) + \frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{in} h(x_i, z_n) \bmod 1, \qquad i = 1, \ldots, L, \tag{10}$$

$$y_j' = g_j(y_j) + \xi_j(z) \bmod 1, \qquad j = 1, \ldots, M, \tag{11}$$

where, with m_1 denoting the Lebesgue measure on $\mathbb{T}$,

$$g_j(y) := f(y) + \alpha\kappa_j \int h(y, x)\, dm_1(x) \bmod 1, \qquad \kappa_j := d_{j+L}/\Delta, \tag{12}$$

$$\xi_j(z) := \alpha\left[\frac{1}{\Delta} \sum_{n=1}^{N} A^h_{jn} h(y_j, z_n) - \kappa_j \int h(y_j, x)\, dm_1(x)\right]. \tag{13}$$

Before stating our theorem, let us give an intuitive argument why we write F in the form (10) and (11), and why for a very long time horizon one can model the resulting dynamics quite well by

$$x_i' \approx f(x_i) \quad \text{and} \quad y_j' \approx g_j(y_j).$$

To see this, note that for a heterogeneous network, the number of non-zero terms in the sum in (10) is an order of magnitude smaller than Δ. Hence when N is large, the interaction felt by the low degree nodes becomes very small and therefore we have approximately $x_i' \approx f(x_i)$. So the low degree nodes are "essentiallly" uncorrelated with each other. Since the Lebesgue measure on $\mathbb{T}$, m_1, is f-invariant and since this measure is exact for the system, one can expect x_i, $i = 1, \dots, L$, to behave as independent uniform random variables on $\mathbb{T}$, at least for "most of the time". Most of the $d_j = \kappa_j \Delta$ incoming connections of hub j are with low degree nodes. It follows that the sum in (13) should converge to

$$\kappa_j \int h(y_j, x)\, dm_1(x)$$

when N is large, and so $\xi_j(z)$ should be close to zero.

Theorem A of this paper is a result which makes this intuition precise. In the following, we let $N_r(\Lambda)$ be the r-neighborhood of a set Λ and we define one-dimensional maps $g_j : \mathbb{T} \to \mathbb{T}$, $j = 1, \dots, M$, to be hyperbolic in a uniform sense.

Definition 2.1 (A hyperbolic collection of one-dimensional maps, see e.g. [dMvS93]). Given $\lambda \in (0, 1)$, $r > 0$ and $m, n \in \mathbb{N}$, we say that $g : \mathbb{T} \to \mathbb{T}$ is (n, m, λ, r)-*hyperbolic* if there exists an attracting set $\Lambda \subset \mathbb{T}$ with

(1) $g(\Lambda) = \Lambda$,
(2) $|D_x g^n| < \lambda$ for all $x \in N_r(\Lambda)$,
(3) $|D_x g^n| > \lambda^{-1}$ for all $x \in N_r(\Upsilon)$ where $\Upsilon := \mathbb{T} \setminus W^s(\Lambda)$,
(4) for each $x \notin N_r(\Upsilon)$, we have $g^k(x) \in N_r(\Lambda)$ for all $k \geq m$,

where $W^s(\Lambda)$ is the union of the stable manifolds of the attractor,

$$W^s(\Lambda) := \left\{ x \in \mathbb{T} : \lim_{k \to \infty} \operatorname{dist}(g^k(x), \Lambda) = 0 \right\}.$$

It is well known (see e.g. [dMvS93, Theorem IV.B]) that for each C^2 map $g : \mathbb{T} \to \mathbb{T}$ (with non-degenerate critical points), the attracting sets are periodic and have uniformly bounded period. If we assume that g is also hyperbolic, we obtain a bound on the number of periodic attractors. A globally expanding map is hyperbolic since it correspond to the case where $\Lambda = \emptyset$.

We now give a precise definition of what we mean by heterogeneous network.

Definition 2.2. We say that a network with parameters L, M, Δ, δ is η-*heterogeneous* with $\eta > 0$ if there are $p, q \in [1, \infty)$ with $1 = 1/p + 1/q$ such that

$$\Delta^{-1} L^{1/p} \delta^{1/q} < \eta, \tag{H1}$$

$$\Delta^{-1/p} M^{2/p} < \eta, \tag{H2}$$

$$\Delta^{-1} M L^{1/p} < \eta, \tag{H3}$$

$$\Delta^{-2} L^{1+2/p} \delta < \eta. \tag{H4}$$

Remark 2.1. (H1)–(H4) arise as sufficient conditions for requiring that the coupled system F is "close" to the product system $f \times \cdots \times f \times g_1 \times \cdots \times g_M : \mathbb{T}^{L+M} \to \mathbb{T}^{L+M}$ and preserve good hyperbolic properties on most of the phase space. They are satisfied in many common settings, as is shown in Appendix G. An easy example to have in mind where those conditions are asymptotically satisfied as $N \to \infty$ for every $\eta > 0$, is the case where M is constant (so $L \sim N$), $\delta \sim L^{\tau}$, and $\Delta \sim L^{\gamma}$, with $0 \leq \tau < 1/2$ and $(\tau + 1)/2 < \gamma < 1$. In particular the layered heterogeneous graphs satisfying (7) in the introduction have these properties.

Theorem A. *Fix σ, h and an interval $[\alpha_1, \alpha_2] \subset \mathbb{R}$ for the parameter α. Suppose that for all $1 \leq j \leq M$ and $\alpha \in [\alpha_1, \alpha_2]$, each of the maps g_j, $j = 1, \ldots, M$, is (n, m, λ, r)-hyperbolic. Then there exist $\xi_0, \eta, C > 0$ such that if the network is η-heterogeneous, then for every $0 < \xi < \xi_0$ and every $1 \leq T \leq T_1$ with*

$$T_1 = \exp[C\Delta\xi^2],$$

there is a set $\Omega_T \subset \mathbb{T}^N$ of initial conditions with

$$m_N(\Omega_T) \geq 1 - \frac{T+1}{T_1}$$

such that for all $(x(0), y(0)) \in \Omega_T$,

$$|\xi_j(z(t))| < \xi, \quad \forall 1 \leq j \leq M \text{ and } 1 \leq t \leq T.$$

Remark 2.2. The result holds under conditions (H1)–(H4) with η sufficiently small, but uniform in the local dynamical parameters. Notice that p has a different role in (H1), (H3), (H4) and in (H2) so that a large p helps the first one, but hinders the second and vice versa for a small p.

The proof of Theorem A will be presented separately in the case where g_j is an expanding map of the circle for all the hubs (Section 4), and when at least one of the g_j has an attracting point (Section 5).

The next theorem is a consequence of results on density of hyperbolicity in dimension one and Theorem A. It shows that the hypothesis on hyperbolicity of the reduced maps g_j is generically satisfied, and that generically one can tune the coupling strength to obtain reduced maps with attracting periodic orbits resulting in regular behaviour for the hub nodes.

Theorem B (Coherent behaviour for hub nodes).

(a) *For each $\sigma \in \mathbb{N}$, $\alpha \in \mathbb{R}$, $\kappa_j \in (0, 1]$, there is an open and dense set $\Gamma \subset C^{10}(\mathbb{T}^2; \mathbb{R})$ such that, for all coupling functions $h \in \Gamma$, the function $g_j \in C^{10}(\mathbb{T}, \mathbb{T})$ defined by (12) is hyperbolic (as in Definition 2.1).*

(b) *There is an open and dense set $\Gamma' \subset C^{10}(\mathbb{T}^2; \mathbb{R})$ such that for all $h \in \Gamma'$ there exists an interval $I \subset \mathbb{R}$ such that if $\alpha\kappa_j \in I$ then g_j has a non-empty and finite periodic attractor. Furthermore, suppose that $h \in \Gamma'$, the graph G satisfies the assumptions of Theorem A for some $\xi > 0$ sufficiently small, and $\alpha\kappa_j \in I$ for the hub $j \in \mathcal{N}$.*

Then there exists $C > 0$ and $\chi \in (0,1)$ such that the following holds. Let $T_1 := \exp[C\Delta\xi^2]$. *There is a set $\Omega_T \subset \mathbb{T}^N$ of initial conditions with*

$$m_N(\Omega_T) \geq 1 - \frac{T+1}{T_1} - \xi^{1-\chi}$$

such that for all $z(0) \in \Omega_T$ there is a periodic orbit of g_j, $O = \{\overline{z}(k)\}_{k=1}^{p}$, for which

$$\mathrm{dist}(z_j(t), \overline{z}(t \bmod p)) \leq \xi \quad \text{for } 1/\xi \leq t \leq T \leq T_1.$$

Proof. See Appendix E. □

Remark 2.3. In the setting of the theorem above, consider the case where two hubs $j_1, j_2 \in \mathcal{N}$ have the same connectivity κ, and their reduced dynamic g_{j_i} has a unique attracting periodic orbit. In this situation their orbits closely follow this unique orbit (as prescribed by the theorem) and, apart from a phase shift $\tau \in \mathbb{N}$, they will be close to each other resulting in highly coherent behaviour:

$$\mathrm{dist}(z_{j_1}(t), z_{j_2}(t+\tau)) \leq 2\xi$$

under the conditions of Theorem B. In general, the attractor of g_{j_i} is the union of a finite number of attracting periodic orbits. Choosing initial conditions for the hub's coordinates in the same connected component of the basin of attraction of one of the periodic orbits yields the same coherent behaviour as above.

In the next theorem we show that for large heterogeneous networks, in contrast with the case of homogeneous networks, coherent behaviour of the hubs is the most one can hope for, and global synchronisation is unstable.

Definition 2.3 (Erdős–Rényi random graphs [Bol01]). For every N and p, an *Erdős–Rényi random graph* is a discrete probability measure on the set $\mathcal{G}(N)$ of undirected graphs on N vertices which assigns independently probability $p \in (0,1)$ to the presence on any of the edge.

Denote by $\mathbb{P}_p$ that probability and by (A_{ij}) the symmetric adjacency matrix of a graph randomly chosen according to $\mathbb{P}_p$. Then $\{A_{ij}\}_{j\geq i}$ are i.i.d. random variables equal to 1 with probability p, and to 0 with probability $1-p$.

Theorem C (Stability and instability of synchrony).

(a) *Take a diffusive coupling function $h(x,y) = \varphi(y-x)$ for some $\varphi : \mathbb{T} \to \mathbb{R}$ with $\frac{d\varphi}{dx}(0) \neq 0$. For any coupling function h' in a sufficiently small neighbourhood of h, there is an interval $I \subset \mathbb{R}$ of coupling strengths such that for any $p \in ((\log N)/N, 1]$ there exists a subset $\mathcal{G}_{\mathrm{Hom}}(N) \subset \mathcal{G}(N)$ of undirected homogeneous graphs with $\mathbb{P}_p(\mathcal{G}_{\mathrm{Hom}}(N)) \to 1$ as $N \to \infty$ such that for any $\alpha \in \mathcal{I}$ the synchronization manifold $\mathcal{S}$, defined in* (8)*, is locally exponentially stable* (*normally attracting*) *for each network coupled on $G \in \mathcal{G}_{\mathrm{Hom}}(N)$.*

(b) *Take any sequence* $\{G(N)\}_{N\in\mathbb{N}}$ *of graphs where* $G(N)$ *has* N *nodes and non-decreasing sequence of degrees* $\boldsymbol{d}(N) = (d_{1,N}, \ldots, d_{N,N})$. *If* $d_{N,N}/d_{1,N} \to \infty$ *as* $N \to \infty$, *then for any diffusive coupling* h *and coupling strength* $\alpha \in \mathbb{R}$ *there is* $N_0 \in \mathbb{N}$ *such that the synchronization manifold* $\mathcal{S}$ *is unstable for the network coupled on* $G(N)$ *with* $N > N_0$.

Proof. See Appendix F. □

2.1. Literature review and the necessity of a new approach for HCM

We just briefly recall the main lines of research on dynamical systems coupled in networks to highlight the need of a new perspective in order to meaningfully describe HCM. For more complete surveys see [PG14, Fer14].

- **Bifurcation theory** [GS$^+$98, GS06, KY10, AA$^+$11, RS15]. In this approach typically there exists a low-dimensional invariant set where the interesting behaviour happens. Often the equivariant group structure is used to obtain a centre manifold reduction. In our case the networks are not assumed to have symmetries (e.g. random networks) and the relevant invariant sets are fractal-like containing unstable manifolds of very high dimension (see Figure 5). For these reasons it is difficult to frame HCM in this setting or use perturbative arguments.
- **Global synchronization** [Kur84, BP02, EM14, PE$^+$14] is the convergence of orbits to a low-dimensional invariant manifold where all the nodes evolve coherently. HCM do not exhibit global synchronization. The synchronization manifold in (8) is unstable (see Theorem C). Furthermore, many works [SB16, Str00] deal with global synchronization when the network is fully connected (all-to-all coupling) by studying the uniform mean field in the thermodynamic limit. On the other hand, we are interested in the case of a finite size system and when the mean field is not uniform across connectivity layers.
- The statistical description of **coupled map lattices** [Kan92, BS88, BR01, BD$^+$98, KL05, KL06, KL04, CF05, Sél18] deals with maps coupled on homogeneous graphs and considers the persistence and ergodic properties of invariant measures when the magnitude of the coupling strength goes to zero. In our case the coupling regime is such that hub nodes are subject to an order one perturbation coming from the dynamics. Low degree nodes still feel a small contribution from the rest of the network, but its magnitude depends on the system size and to make it arbitrarily small the dimensionality of the system must increase as well.

It is worth mentioning that dynamics of coupled systems with different subsystems appears also in *slow-fast system* dynamics [GM13, DSL16, SVM07]. Here, loosely speaking, some (slow) coordinates evolve as “id $+\, \varepsilon h$” and the others have good ergodic properties. In this case one can apply ergodic averaging and obtain a good approximation of the slow coordinates for time up to time $T \sim \varepsilon^{-1}$. In our case, spatial rather than time ergodic averaging takes place and there is no dichotomy on the time scales at different nodes. Furthermore, the role of the perturbation parameter is played by Δ^{-1} and we obtain $T = \exp(C\Delta)$, rather than the polynomial estimate obtained in slow-fast systems.

3. Sketch of the proof and the use of a "truncated" system

3.1. A trivial example exhibiting main features of HCM

We now present a more or less trivial example which already presents all the main features of heterogeneous coupled maps, namely

- existence of a set of *"bad" states* with large fluctuations of the mean field,
- control on the *hitting time* of the bad set,
- *finite time* exponentially large in the size of the network.

Consider the evolution of $N = L + 1$ doubling maps on the circle $\mathbb{T}$ interacting on a *star network* with nodes $\{1, \ldots, L + 1\}$ and directed edges $\mathcal{E} = \{(i, L + 1) : 1 \le i \le L\}$ (see Figure 4). The hub node $\{L + 1\}$ has an incoming directed edge from any other node of the network, while the other nodes have just the outgoing edge. Take for interaction function the diffusive coupling $h(x, y) := \sin(2\pi y) - \sin(2\pi x)$. Equations (10) and (11) then become

$$2x_i(t + 1) = 2x_i(t) \bmod 1, \qquad 1 \le i \le L, \tag{14}$$

$$y(t + 1) = 2y(t) + \frac{\alpha}{L}\sum_{i=1}^{L}[\sin(2\pi x_i(t)) - \sin(2\pi y(t))] \bmod 1. \tag{15}$$

The low degree nodes evolve as an uncoupled doubling map making the above a *skew-product* system on the base $\mathbb{T}^L$ akin to the one extensively studied in [Tsu01]. One can rewrite the dynamics of the forced system (the hub) as

$$y(t + 1) = 2y(t) - \alpha \sin(2\pi y(t)) + \frac{\alpha}{L}\sum_{i=1}^{L}\sin(2\pi x_i(t)) \tag{16}$$

and notice that with $g(y) := 2y - \alpha \sin(2\pi y) \bmod 1$, the evolution of $y(t)$ is given by the application of g plus a noise term

$$\xi(t) = \frac{\alpha}{L}\sum_{i=1}^{L}\sin(2\pi x_i(t)) \tag{17}$$

depending on the low degree nodes coordinates. The Lebesgue measure on $\mathbb{T}^L$ is invariant and mixing for the dynamics restricted to the first L uncoupled coordinates. The set of

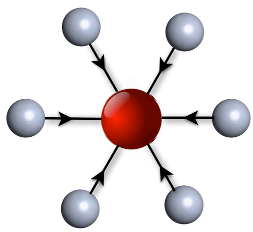

Fig. 4. Star network with only incoming arrows.

bad states where fluctuations (17) are above a fixed threshold $\varepsilon > 0$ is

$$\mathcal{B}_\varepsilon := \left\{ x \in \mathbb{T}^L : \left| \frac{1}{L} \sum_{i=1}^{L} \sin(2\pi x_i) - \mathbb{E}_m[\sin(2\pi x)] \right| > \varepsilon \right\}$$
$$= \left\{ x \in \mathbb{T}^L : \left| \frac{1}{L} \sum_{i=1}^{L} \sin(2\pi x_i) \right| > \varepsilon \right\}.$$

Using large deviation results one can upper bound the measure of the set above as

$$m_L(\mathcal{B}_\varepsilon) \leq \exp(-C\varepsilon^2 L)$$

($C > 0$ is a constant uniform in L and ε; see the Hoeffding inequality in Appendix A for details). Since we know that the dynamics of the low degree nodes is ergodic with respect the measure m_L, we have the following information regarding the time evolution of the hub.

- The set $\mathcal{B}_\varepsilon$ has positive measure. Ergodicity of the invariant measure implies that a generic initial condition will visit $\mathcal{B}_\varepsilon$ in finite time, making any mean-field approximation result for infinite time hopeless.
- As a consequence of the Kac lemma, the average hitting time of the set $\mathcal{B}_\varepsilon$ is $m_L(\mathcal{B}_\varepsilon)^{-1} \geq \exp(C\varepsilon^2 L)$, thus exponentially large in the dimension.
- From the invariance of the measure m_L, for every $1 \leq T \leq \exp(C\varepsilon^2 L)$ there is a set $\Omega_T \subset \mathbb{T}^{L+1}$ with measure $m_{L+1}(\Omega_T) > 1 - T\exp(-C\varepsilon^2 L)$ such that for all $x \in \Omega_T$ and for every $1 \leq t \leq T$,
$$\left| \frac{1}{L} \sum_{i=1}^{L} \sin(2\pi x_i(t)) \right| \leq \varepsilon.$$

3.2. *Truncated system*

We obtain a description of the coupled system by restricting our attention to a subset of phase space where the evolution prescribed by equations (10) and (11) resembles the evolution of the uncoupled mean-field maps, and we redefine the evolution outside this subset in a convenient way. This leads to the definition of a *truncated* map $F_\varepsilon : \mathbb{T}^N \to \mathbb{T}^N$, for which the fluctuations of the mean-field averages are artificially cut off at level $\varepsilon > 0$, resulting in a well behaved hyperbolic dynamical system. In the following sections we will then determine existence of and bounds on the invariant measure for this system and prove that the portion of phase space where the original system and the truncated one coincide is of almost full measure with a remainder exponentially small in the parameter Δ.

Note that since $h \in C^{10}(\mathbb{T}^2; \mathbb{R})$, its Fourier series

$$h(x, y) = \sum_{s=(s_1,s_2)\in\mathbb{Z}^2} c_s \theta_{s_1}(x)\theta_{s_2}(y),$$

where $c_s \in \mathbb{R}$ and $\theta_i : \mathbb{T} \to [0, 1]$ form a base of trigonometric functions, converges uniformly and absolutely on $\mathbb{T}^2$. Furthermore, for all $s \in \mathbb{Z}^2$,

$$|c_s| \leq \frac{\|h\|_{C^{10}}}{|s_1|^5 |s_2|^5}. \tag{18}$$

Taking $\overline{\theta}_{s_1} = \int \theta_{s_1}(x)\, dm_1(x)$ we get

$$\xi_j(z) := \alpha \sum_{s\in\mathbb{Z}^2} c_s \left[\frac{1}{\Delta}\sum_{n=1}^{L} A^h_{jn}\theta_{s_1}(z_n) - \kappa_j\overline{\theta}_{s_1}\right]\theta_{s_2}(y_j) + \frac{\alpha}{\Delta}\sum_{n=1}^{M} A^h_{jn} h(y_j, y_n). \tag{19}$$

For every $\varepsilon > 0$ choose a C^∞ map $\zeta_\varepsilon : \mathbb{R} \to \mathbb{R}$ with $\zeta_\varepsilon(t) = t$ for $|t| < \varepsilon$, and $\zeta_\varepsilon(t) = 2\varepsilon$ for $|t| > 2\varepsilon$. So for each $\varepsilon > 0$, the function $t \mapsto |D_t\zeta_\varepsilon|$ is uniformly bounded in t and ε. We define the evolution for the truncated dynamics $F_\varepsilon\colon \mathbb{T}^{L+M} \to \mathbb{T}^{L+M}$ by the following modification of equations (10) and (11):

$$x_i' = f(x_i) + \frac{\alpha}{\Delta}\sum_{n=1}^{N} A_{in} h(x_i, z_n) \bmod 1, \qquad i = 1, \dots, L, \tag{20}$$

$$y_j' = g_j(y_j) + \xi_{j,\varepsilon}(z) \bmod 1, \qquad j = 1, \dots, M, \tag{21}$$

where the expression of $\xi_{j,\varepsilon}(z)$ modifies that of $\xi_j(z)$ in (19):

$$\xi_{j,\varepsilon}(z) := \alpha \sum_{s\in\mathbb{Z}^2} c_s \zeta_{\varepsilon|s_1|}\left(\frac{1}{\Delta}\sum_{i=1}^{L} A_{ji}\theta_{s_1}(x_i) - \kappa_j\overline{\theta}_{s_1}\right)\theta_{s_2}(y_j) + \frac{\alpha}{\Delta}\sum_{n=1}^{M} A^h_{jn} h(y_j, y_n). \tag{22}$$

So the only difference between F and F_ε is the cut-off functions $\zeta_{\varepsilon|s|}$ appearing in (22). For every $\varepsilon > 0$, $j \in \{1, \dots, M\}$ and $s_1 \in \mathbb{Z}$ define

$$\mathcal{B}_\varepsilon^{(s_1,j)} := \left\{x \in \mathbb{T}^L : \left|\frac{1}{\Delta}\sum_{i=1}^{L} A^h_{ji}\theta_{s_1}(x_i) - \kappa_j\overline{\theta}_{s_1}\right| > \varepsilon|s_1|\right\}. \tag{23}$$

The set where F and F_ε coincide is $\mathcal{Q}_\varepsilon \times \mathbb{T}^M$ where

$$\mathcal{Q}_\varepsilon := \bigcap_{j=1}^{M}\bigcap_{s_1\in\mathbb{Z}} \mathbb{T}^L \setminus \mathcal{B}_\varepsilon^{(s_1,j)} \tag{24}$$

is the subset of $\mathbb{T}^L$ where all the fluctuations of the mean-field averages of the terms of the coupling are less than the imposed threshold. The set $\mathcal{B}_\varepsilon := \mathcal{Q}_\varepsilon^c$, is the portion of phase space for the low degree nodes were the fluctuations exceed the threshold, and the systems F and F_ε are different. Furthermore we can control the perturbation introduced by the term $\xi_{j,\varepsilon}$ in equation (11) so that F_ε is close to the hyperbolic uncoupled product map $\boldsymbol{f} : \mathbb{T}^N \to \mathbb{T}^N$,

$$\boldsymbol{f}(x_1, \dots, x_L, y_1, \dots, y_M) := (f(x_1), \dots, f(x_L), g_1(y_1), \dots, g_M(y_M)). \tag{25}$$

All the bounds on relevant norms of $\xi_{j,\varepsilon}$ are reported in Appendix A. To upper bound the Lebesgue measure $m_L(\mathcal{B}_\varepsilon)$ we use the Hoeffding inequality (reported in Appendix A) on concentration of the average of independent bounded random variables.

Proposition 3.1.

$$m_L(\mathcal{B}_\varepsilon) \le \frac{\exp[-\Delta\varepsilon^2/2 + \mathcal{O}(\log M)]}{1 - \exp[-\Delta\varepsilon^2/2]}. \tag{26}$$

Proof. See Appendix A. □

This gives an estimate of the measure of the bad set with respect to the reference measure invariant for the uncoupled maps. In the next section we use this estimate to upper bound the measure of this set with respect to the SRB measures for F_ε, which are measures giving statistical information on the orbits of F_ε.

Remark 3.1. Notice that in (26) we expressed the upper bound only in terms of orders of functions of the network parameters, but all the constants could be rigorously estimated in terms of the coupling function and the other dynamical parameters of the system. In particular, where the expression of the coupling function was known one could have obtained better estimates on the concentration via large deviation results (see for example Cramér-type inequalities in [DZ09]), which takes into account more than just the upper and lower bounds of θ_s. In what follows, however, we will be only interested in the order of magnitude with respect to the aforementioned parameters of the network (Δ, δ, L, M).

3.3. Steps of the proof and challenges

The basic steps of the proof are the following:

(i) First of all we are going to restrict our attention to the case where the maps g_j satisfy Definition 2.1 with $n = 1$.
(ii) Secondly, hyperbolicity of the map F_ε is established for an η-heterogeneous network with $\varepsilon, \eta > 0$ small. This is achieved by constructing forward and backward invariant cone fields made up of expanding and contracting directions respectively for the cocycle defined by application of $D_z F_\varepsilon$ (see (63)).
(iii) Then we estimate the distortion of the maps along the unstable directions, keeping all dependencies on the structural parameters of the network explicit.
(iv) We then use a geometric approach employing what is sometimes called *standard pairs* [CLP16] to estimate the regularity properties of the SRB measures for the endomorphism F_ε, and the hitting time of the set $\mathcal{B}_\varepsilon$.
(v) Finally, we show that Mather's trick allows us to generalise the proofs to the case in which the g_j satisfy Definition 2.1 with $n \neq 1$.

We consider separately the cases where all the reduced maps g_j are expanding and when some of them have non-empty attractor (Sections 4 and 5). At the end of Section 5 we put the results together to obtain the proof of Theorem A.

In the above points we treat F_ε as a perturbation of a product map where the magnitude of the perturbation depends on the network size. In particular, we want to show that F_ε is close to the uncoupled product map $\boldsymbol{f}$. To obtain this, the dimensionality of the system needs to increase, changing the underlying phase space. This leads to two main challenges. First of all, increasing the size of the system propagates non-linearities of the maps and reduces the global regularity of the invariant measures. Secondly, the situation is inherently different from the usual perturbation theory where one considers a parametric family of dynamical systems on the same phase space. Here, the parameters depend on the system's dimension. As a consequence, one needs to make all estimates explicit in the system size. For these reasons we find the geometric approach advantageous compared to the functional-analytic approach [KL99] where the explicit dependence of most constants on the dimension is hidden in the functional-analytic machinery.

Notation. As usual, we write $\mathcal{O}(N)$ and $\mathcal{O}(\varepsilon)$ for an expression such that $\mathcal{O}(N)/N$ resp. $\mathcal{O}(\varepsilon)/\varepsilon$ is bounded as $N \to \infty$ resp. $\varepsilon \downarrow 0$. We use shorthand notation $[n] := \{1, \dots, n\}$.

Throughout, m and m_n stand for the Lebesgue measure on $\mathbb{T}$ and $\mathbb{T}^n$ respectively. Given an embedded manifold $W \subset \mathbb{T}^N$, m_W stands for the Lebesgue measure induced on W.

We indicate by $D_x G$ the differential of the function G evaluated at the point x in its domain.

4. Proof of Theorem A when all reduced maps are uniformly expanding

In this section we assume that the collection of reduced maps g_j, $j = 1, \dots, M$, from equation (12) is uniformly expanding. As shown in Lemma 5.6, this means that we can assume that there exists $\lambda \in (0, 1)$ such that $|g_j(x)| \geq \lambda^{-1}$ for all $x \in \mathbb{T}$ and all $j = 1, \dots, M$.

First of all pick $1 \leq p \leq \infty$, let $1 \leq q \leq \infty$ be such that $1/p + 1/q = 1$, and consider the norm defined as

$$\| \cdot \|_p := \| \cdot \|_{p,\mathbb{R}^L} + \| \cdot \|_{p,\mathbb{R}^M}$$

where $\| \cdot \|_{p,\mathbb{R}^k}$ is the usual p-norm on $\mathbb{R}^k$. Then $\| \cdot \|_p$ induces the operator norm of any linear map $\mathcal{L} \colon \mathbb{R}^N \to \mathbb{R}^N$, namely

$$\|\mathcal{L}\|_p := \sup_{\substack{v \in \mathbb{R}^N \\ \|v\|_p = 1}} \frac{\|\mathcal{L}v\|_p}{\|v\|_p},$$

and the distance $d_p : \mathbb{T}^N \times \mathbb{T}^N \to \mathbb{R}^+$ on $\mathbb{T}^N$.

Theorem 4.1. *There are $\eta_0, \varepsilon_0 > 0$ such that under (H1)–(H4) with $\eta < \eta_0$ and for all $\varepsilon < \varepsilon_0$ there exists an absolutely continuous invariant probability measure ν for F_ε. The density $\rho = d\nu/dm_N$ satisfies, for all $z, \overline{z} \in \mathbb{T}^N$,*

$$\frac{\rho(z)}{\rho(\overline{z})} \leq \exp\{a d_p(z, \overline{z})\}, \qquad a = \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M). \tag{27}$$

In Section 4.1 we obtain conditions on the heterogeneous structure of the network which ensure that the truncated system F_ε is sufficiently close, in the C^1 topology, to the uncoupled system $\boldsymbol{f}$, in (25), with the hubs evolving according to the low-dimensional approximation g_j, for it to preserve expansivity when the network is large enough. In this setting, F_ε is a uniformly expanding endomorphism and therefore has an absolutely continuous invariant measure ν whose density $\rho = \rho_\varepsilon$ is a fixed point of the transfer operator of F_ε,

$$P_\varepsilon : L^1(\mathbb{T}^N, m_N) \to L^1(\mathbb{T}^N, m_N).$$

(See Appendix C for a quick review of the theory of transfer operators.) For our purposes we will also require bounds on ρ which are explicit in the structural parameters of the network (for suitable ε). In Section 4.2 we obtain bounds on the distortion of the Jacobian of F_ε (Proposition 4.2), which in turn allow us to prove the existence of a cone of functions with controlled regularity which is invariant under the action of P_ε (Proposition 4.3) and to which ρ belongs. To obtain the conclusion of Theorem A, we need that the ν-measure of the bad set is small, which will be obtained from an upper bound for the supremum of the functions in the invariant cone. This is shown in Section 4.4 under some additional conditions on the network.

4.1. Global expansion of F_ε

Proposition 4.1. *Suppose that for every $j \in [M]$ the reduced map g_j is uniformly expanding, i.e. there exists $\lambda \in (0,1)$ such that $|D_y g_j| > \lambda^{-1}$ for all $y \in \mathbb{T}$. Then*

(i) *there exists $C_\#$ (depending on σ, h and α only) such that for every $1 \le p \le \infty$, $z \in \mathbb{T}^N$, and $w \in \mathbb{R}^N \setminus \{0\}$,*

$$\frac{\|(D_z F_\varepsilon) w\|_p}{\|w\|_p} \ge [\min\{\sigma, \lambda^{-1} - \varepsilon C_\#\} - \mathcal{O}(\Delta^{-1}\delta) - \mathcal{O}(\Delta^{-1/p} M^{1/p}) - \mathcal{O}(\Delta^{-1} N^{1/p} \delta^{1/q})];$$

(ii) *there exists $\eta > 0$ such that if (H1) and (H3) are satisfied with*

$$\varepsilon < \frac{\lambda^{-1} - 1}{C_\#} \tag{28}$$

then there exists $\overline{\sigma} > 1$ (not depending on the parameters of the network or on p), so that

$$\frac{\|(D_z F_\varepsilon) w\|_p}{\|w\|_p} \ge \overline{\sigma} > 1, \qquad \forall z \in \mathbb{T}^N,\ \forall w \in \mathbb{R}^N \setminus \{0\}.$$

Proof. To prove (i), let $z = (x, y) \in \mathbb{T}^{L+M}$ and $w = \binom{u}{v} \in \mathbb{R}^{L+M}$ and

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = D_z F_\varepsilon \begin{pmatrix} u \\ v \end{pmatrix}, \qquad u' \in \mathbb{R}^L,\ v' \in \mathbb{R}^M.$$

Using (20)–(21), or (63), we find that for every $1 \le i \le L$ and $1 \le j \le M$,

$$u_i' = \left[D_{x_i} f + \frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{in} h_1(x_i, z_n) \right] u_i + \frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{in} h_1(x_i, z_n) w_n,$$

$$v_j' = \sum_{\ell=1}^{L} \partial_{x_\ell} \xi_{j,\varepsilon} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^{M} A_{jm}^{hh} h_2(y_j, y_m) v_m + [D_{y_j} g_j + \partial_{y_j} \xi_{j,\varepsilon}] v_j,$$

where h_1 and h_2 denote the partial derivatives with respect to the first and second variable. Hence

$$\|u'\|_{p,\mathbb{R}^L} \ge (\sigma - \mathcal{O}(\delta\Delta^{-1}))\|u\|_p - \mathcal{O}(\Delta^{-1} L^{1/p}) \max_{i=1,\dots,L} \left[\sum_{n=1}^{N} A_{in} |w_n| \right].$$

Recall that, for any $k \in \mathbb{N}$, if $w \in \mathbb{R}^k$ then

$$\|w\|_{1,\mathbb{R}^k} \leq k^{1/q} \|w\|_{p,\mathbb{R}^k} \quad \text{with } 1/p + 1/q = 1 \tag{29}$$

for every $1 \leq p \leq \infty$. Thus

$$\sum_{n=1}^{N} A_{in} |w_n| \leq \delta^{1/q} \Big(\sum_{n=1}^{N} A_{in} |w_n|^p \Big)^{1/p} \leq \delta^{1/q} \|w\|_p$$

since at most δ terms are non-vanishing in the sum $\sum_{n=1}^{N} A_{in} |w_n|^p$, we can view as a vector in $\mathbb{R}^\delta$, which implies

$$\|u'\|_{p,\mathbb{R}^L} \geq (\sigma - \mathcal{O}(\delta \Delta^{-1})) \|u\|_p - \mathcal{O}(\Delta^{-1} L^{1/p} \delta^{1/q}) \|w\|_p.$$

Analogously using the estimates in Lemma A.1 we obtain

$$\begin{aligned} \|v'\|_{p,\mathbb{R}^M} &\geq (\lambda^{-1} - \varepsilon C_\#) \|v\|_p - \mathcal{O}(\Delta^{-1} M^{1/p}) \max_{j=1,\ldots,M} \Big[\sum_n A_{jn} |w_n| \Big] \\ &\geq (\lambda^{-1} - \varepsilon C_\# - \mathcal{O}(\Delta^{-1} M)) \|v\|_p - \mathcal{O}(\Delta^{-1/p} M^{1/p}) \|w\|_p \end{aligned} \tag{30}$$

since in the sum $\sum_n A_{jn} |w_n|$ in (30), at most Δ terms are different from zero and $\Delta^{-1} \Delta^{1/q} = \Delta^{-1/p}$. This implies

$$\begin{aligned} &\frac{\|(u', v')\|_p}{\|(u, v)\|_p} = \frac{\|u'\|_{p,\mathbb{R}^L} + \|v'\|_{p,\mathbb{R}^M}}{\|(u, v)\|_p} \\ &\geq \min\{\sigma - \mathcal{O}(\Delta^{-1}\delta), \lambda^{-1} - \varepsilon C_\# - \mathcal{O}(\Delta^{-1} M)\} - \mathcal{O}(\Delta^{-1/p} M^{1/p}) - \mathcal{O}(\Delta^{-1} L^{1/p} \delta^{1/q}). \end{aligned}$$

For the proof of (ii), notice that condition (28) implies that $\min\{\sigma, \lambda^{-1} - \varepsilon C_\#\} > 1$ and conditions (H1)–(H3) imply that the $\mathcal{O}$ are bounded by η and so

$$\frac{\|D_z F_\varepsilon w\|_p}{\|w\|_p} \geq \min\{\sigma, \lambda^{-1} - \varepsilon C_\#\} - \mathcal{O}(\eta), \quad \forall w \in \mathbb{R}^N \setminus \{0\},$$

and choosing $\eta > 0$ sufficiently small one obtains the conclusion. □

Now that we have proved that F_ε is expanding, we know from the ergodic theory of expanding maps that it has an invariant measure, say ν, with density $\rho = d\nu / dm_N$. The rest of the section is dedicated to upper bounding $\nu(\mathcal{Q}_\varepsilon)$.

4.2. Distortion of F_ε

Proposition 4.2. *If conditions (H1)–(H3) are satisfied then there exists ε_0 (depending only on σ, $|\alpha|$ and the coupling function h) such that if $\varepsilon < \varepsilon_0$ then for every $z, \bar{z} \in \mathbb{T}^N$,*

$$\frac{|D_z F_\varepsilon|}{|D_{\bar{z}} F_\varepsilon|} \leq \exp\{[\mathcal{O}(\Delta^{-1} \delta L) + \mathcal{O}(M)] d_\infty(z, \bar{z})\}.$$

Proof. To estimate the ratios consider the matrix $\mathcal{D}(z)$ obtained from $D_z F_\varepsilon$ by factoring $D_{x_i} f = \sigma$ out of the i-th column ($i \in [N]$), and $D_{y_j} g_j$ out of the $(j+L)$-th column ($j \in [M]$). Thus

$$[\mathcal{D}(z)]_{k,\ell} = \begin{cases} 1 + \frac{\alpha}{\Delta} \sum_{n=1}^{L} A_{kn} \frac{h_1(x_k, z_n)}{\sigma}, & k = \ell \le L, \\ \frac{\alpha}{\Delta} A_{k\ell} \frac{h_2(x_k, x_\ell)}{\sigma}, & k \neq \ell \le L, \\ \frac{\partial_{x_\ell} \xi_{k-L,\varepsilon}}{\sigma}, & k > L, \ell \le L, \\ \frac{\alpha}{\Delta} A_{k\ell} \frac{h_2(y_{k-L}, y_{\ell-L})}{D_{y_{\ell-L}} g_{\ell-L}}, & k \neq \ell > L, \\ 1 + \frac{\partial_{y_{k-L}} \xi_{k-L,\varepsilon}}{D_{y_{\ell-L}} g_{\ell-L}}, & k = \ell > L, \end{cases} \tag{31}$$

and

$$\frac{|D_z F_\varepsilon|}{|D_{\overline{z}} F_\varepsilon|} = \frac{\prod_{j=1}^{M} D_{y_j} g_j}{\prod_{j=1}^{M} D_{\overline{y}_j} g_j} \cdot \frac{|\mathcal{D}(z)|}{|\mathcal{D}(\overline{z})|}.$$

For the first ratio

$$\begin{aligned} \prod_{j=1}^{M} \frac{D_{y_j} g_j}{D_{\overline{y}_j} g_j} &= \prod_{j=1}^{M} \left(1 + \frac{D_{y_j} g_j - D_{\overline{y}_j} g_j}{D_{\overline{y}_j} g_j}\right) \\ &\le \prod_{j=1}^{M} (1 + \mathcal{O}(1)|y_j - \overline{y}_j|) \le \exp[\mathcal{O}(M) d_\infty(y, \overline{y})]. \end{aligned} \tag{32}$$

To estimate the ratio $\frac{|\mathcal{D}(z)|}{|\mathcal{D}(\overline{z})|}$ we will apply Proposition B.1 of Appendix B. To this end define the matrix

$$B(z) := \mathcal{D}(z) - \mathrm{Id}\,.$$

First of all we will prove that for every $1 \le p < \infty$ and $z \in \mathbb{T}^N$, $B(z)$ has operator norm bounded by

$$\|B(z)\|_p \le \max\{\mathcal{O}(\Delta^{-1} M), C_\# \varepsilon\} + \mathcal{O}(\Delta^{-1/p} M^{1/p}) + \mathcal{O}(\Delta^{-1} N^{1/p} \delta^{1/q}) \tag{33}$$

where $C_\#$ is a constant uniform in the parameters of the network and $1/p + 1/q = 1$. Indeed, consider $\binom{u}{v} \in \mathbb{R}^{L+M}$ and $\binom{u'}{v'} := B(z)\binom{u}{v}$. Then

$$u_i' = \left[\frac{\alpha}{\Delta} \sum_{n=1}^{L} A_{in} \frac{h_1(x_i, z_n)}{\sigma}\right] u_i + \frac{\alpha}{\Delta} \sum_{\ell=1}^{L} A_{i\ell}^{ll} \frac{h_1(x_i, x_\ell)}{\sigma} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^{M} A_{im}^{lh} \frac{h_2(x_i, y_m)}{D_{y_m} g_m} v_m,$$

$$v_j' = \sum_{\ell=1}^{L} \frac{\partial_{x_\ell} \xi_{j,\varepsilon}}{\sigma} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^{M} A_{jm}^{hh} \frac{h_2(y_j, y_m)}{D_{y_m} g_m} v_m + \frac{\partial_{y_j} \xi_{j,\varepsilon}}{D_{y_j} g_j} v_j.$$

Using estimates analogous to the ones used in the proof of Proposition 4.1 yields

$$\begin{aligned}
\|u'\|_{p,\mathbb{R}^L} &\le \mathcal{O}(\Delta^{-1}\delta)\|u\|_p + \mathcal{O}(\Delta^{-1}) \max_i \Big[\sum_{\ell=1}^{L} A^{ll}_{i\ell}|u_\ell| + \sum_{m=1}^{M} A^{lh}_{im}|v_m|\Big] \\
&\le \mathcal{O}(\Delta^{-1}\delta)\|u\|_p + \mathcal{O}(\Delta^{-1}N^{1/p}\delta^{1/q})\|(u,v)\|_p, \\
\|v'\|_{p,\mathbb{R}^M} &\le C_{\#}\varepsilon\|v\|_p + \mathcal{O}(\Delta^{-1}N^{1/p}) \max_i \Big[\sum_{\ell} A^{ll}_{i\ell}|u_\ell| + \sum_{m} A^{lh}_{im}|v_m|\Big] \\
&\le C_{\#}\varepsilon\|v\|_p + \mathcal{O}(\Delta^{-1/p}M^{1/p})\|(u,v)\|_p,
\end{aligned}$$

so using conditions (H1), (H2), we obtain (33):

$$\begin{aligned}
\frac{\|(u',v')\|_p}{\|(u,v)\|_p} &\le \max\{\mathcal{O}(\Delta^{-1}\delta), C_{\#}\varepsilon\} + \mathcal{O}(\Delta^{-1/p}M^{1/p}) + \mathcal{O}(\Delta^{-1}L^{1/p}\delta^{1/q}) \\
&\le C_{\#}\varepsilon + \mathcal{O}(\eta).
\end{aligned}$$

Taking $C_{\#}\varepsilon < 1$ and $\eta > 0$ sufficiently small ensures that $\|B(z)\|_p \le \lambda < 1$ for all $z \in \mathbb{T}^N$. Now we want to estimate the norm $\|\cdot\|_p$ of columns of $B - \overline{B}$ where

$$B := B(z) \quad \text{and} \quad \overline{B} := B(\overline{z}).$$

For $1 \le i \le L$, looking at the entries of $\mathcal{D}(z)$ (see (31)), it is clear that the non-vanishing entries $[B(z)]_{ik}$ for $k \ne i$ are Lipschitz functions with Lipschitz constants of order $\mathcal{O}(\Delta^{-1})$:

$$|B_{ik} - \overline{B}_{ik}| \le A_{ik}\mathcal{O}(\Delta^{-1})d_\infty(z,\overline{z}).$$

Instead, for $k = i$,

$$\begin{aligned}
|B_{ii} - \overline{B}_{ii}| &= \frac{\alpha}{\Delta}\Big|\sum_{\ell} A^{ll}_{in}(h_1(x_i,x_\ell) - h_1(\overline{x}_i,\overline{x}_\ell)) + \sum_{m} A^{lh}_{im}(h_1(x_i,x_m) - h_1(\overline{x}_i,\overline{x}_m))\Big| \\
&\le \frac{\alpha}{\Delta}\sum_{\ell} A^{ll}_{i\ell}|h_1(x_i,x_\ell) - h_1(\overline{x}_i,\overline{x}_\ell)| + \frac{\alpha}{\Delta}\sum_{m} A^{lh}_{im}|h_1(x_i,y_m) - h_1(\overline{x}_i,\overline{y}_m)| \\
&\le \mathcal{O}(\Delta^{-1}\delta)d_\infty(z,\overline{z}),
\end{aligned}$$

which implies

$$\begin{aligned}
\|\mathrm{Col}^i[B-\overline{B}]\|_p &= \Big(\sum_{k\in[L]} |B_{ik} - \overline{B}_{ik}|^p\Big)^{1/p} + \Big(\sum_{k\in[L+1,N]} |B_{ik} - \overline{B}_{ik}|^p\Big)^{1/p} \\
&\le \Big(\sum_{k\in[L]\setminus\{i\}} |B_{ik} - \overline{B}_{ik}|^p\Big)^{1/p} + \Big(\sum_{k\in[L+1,N]} |B_{ik} - \overline{B}_{ik}|^p\Big)^{1/p} + \mathcal{O}(\Delta^{-1}\delta)d_\infty(z,\overline{z}) \\
&\le 2\Big(\sum_{k\ne i} A_{ik}\Big)^{1/p}\mathcal{O}(\Delta^{-1})d_\infty(z,\overline{z}) + \mathcal{O}(\Delta^{-1}\delta)d_\infty(z,\overline{z}) \\
&\le \mathcal{O}(\Delta^{-1}\delta)d_\infty(z,\overline{z}).
\end{aligned}$$

For $1 \le j \le M$, looking again at (31) we see that the non-vanishing entries of $[B(z)]_{(j+L)k}$ for $k \neq j+N$ are Lipschitz functions with Lipschitz constants of order $\mathcal{O}(\Delta^{-1})$, while $[B(z)]_{(j+L)(j+L)}$ has Lipschitz constant of order $\mathcal{O}(1)$, thus

$$\begin{aligned}
&\|\mathrm{Col}^{j+L}[B-\overline{B}]\|_p \\
&\quad= \Big(\sum_{k\in[L]} |B_{(j+L)k}-\overline{B}_{(j+L)k}|^p\Big)^{1/p} + \Big(\sum_{k\in[L+1,N]} |B_{(j+L)k}-\overline{B}_{(j+L)k}|^p\Big)^{1/p} \\
&\quad\le \Big(\sum_{k\in[L]} |B_{(j+L)k}-\overline{B}_{(j+L)k}|^p\Big)^{1/p} + \Big(\sum_{k\in[L+1,N]\setminus\{j+L\}} |B_{(j+L)k}-\overline{B}_{(j+L)k}|^p\Big)^{1/p} \\
&\qquad+ \mathcal{O}(1)d_\infty(z,\overline{z}) \\
&\quad\le 2\Big(\sum_{k\neq j+L} A_{(j+L)k}\Big)^{1/p}\mathcal{O}(\Delta^{-1})d_\infty(z,\overline{z}) + \mathcal{O}(1)d_\infty(z,\overline{z}) \\
&\quad\le \mathcal{O}(1)d_\infty(z,\overline{z}).
\end{aligned}$$

Proposition B.1 from Appendix B now implies that

$$\frac{|\mathcal{D}(z)|}{|\mathcal{D}(\overline{z})|} \le \exp\Big\{\sum_{k=1}^{N} \|\mathrm{Col}^k[B-\overline{B}]\|_p\Big\} \le \exp\{(\mathcal{O}(\Delta^{-1}\delta L)+\mathcal{O}(M))d_\infty(z,\overline{z})\}. \tag{34}$$

□

4.3. *Invariant cone of functions*

Define the cone of functions

$$C_{a,p} := \{\varphi : \mathbb{T}^N \to \mathbb{R}^+ : \varphi(z)/\varphi(\overline{z}) \le \exp[a d_p(z,\overline{z})],\ \forall z,\overline{z}\in\mathbb{T}^N\}.$$

It is convex and has finite diameter (see for example [Bir57, Bus73] or [Via97]). We now use the result on distortion from the previous section to determine the parameters $a>0$ such that $C_{a,p}$ is invariant under the action of the transfer operator P_ε. Since $C_{a,p}$ has finite diameter with respect to the Hilbert metric on the cone (see [Via97]), P_ε is a contraction when restricted to this set and its unique fixed point is the only invariant density which thus belongs to $C_{a,p}$. In the next subsection, we will use this observation to conclude the proof of Theorem A in the expanding case.

Proposition 4.3. *Under conditions* (H1)–(H3)*, for every $a > a_c$ where a_c is of the form*

$$a_c = \frac{\mathcal{O}(\Delta^{-1}\delta L)+\mathcal{O}(M)}{1-\overline{\sigma}}, \tag{35}$$

$C_{a,p}$ is invariant under the action of the transfer operator P_ε of F_ε, i.e. $P_\varepsilon(C_{a,p}) \subset C_{a,p}$.

Proof. Since F_ε is a local expanding diffeomorphism, its transfer operator P_ε has the expression

$$(P_\varepsilon\varphi)(z) = \sum_i \varphi(F_{\varepsilon,i}^{-1}(z))|D_{F_{\varepsilon,i}^{-1}(z)}F_\varepsilon|^{-1}$$

where $\{F_{\varepsilon,i}\}_i$ are surjective invertible branches of F_ε. Suppose $\varphi \in C_{a,p}$. Then

$$
\begin{aligned}
&\frac{\varphi(F_{\varepsilon,i}^{-1}(z))}{\varphi(F_{\varepsilon,i}^{-1}(\overline{z}))} \frac{|D_{F_{\varepsilon,i}^{-1}(\overline{z})} F_\varepsilon|}{|D_{F_{\varepsilon,i}^{-1}(z)} F_\varepsilon|} \\
&\qquad \le \exp\{a d_p(F_{\varepsilon,i}^{-1}(z), F_{\varepsilon,i}^{-1}(\overline{z}))\} \exp\{[\mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_\infty(F_{\varepsilon,i}^{-1}(z), F_{\varepsilon,i}^{-1}(\overline{z}))\} \\
&\qquad \le \exp\{[a + \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_p(F_{\varepsilon,i}^{-1}(z), F_{\varepsilon,i}^{-1}(\overline{z}))\} \\
&\qquad \le \exp\{[\overline{\sigma}^{-1} a + \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)] d_p(z, \overline{z})\}.
\end{aligned}
$$

Here we have used $d_\infty(z, \overline{z}) \le d_p(z, \overline{z})$ for every $1 \le p < \infty$. Hence

$$
\begin{aligned}
\frac{(P_\varepsilon \varphi)(z)}{(P_\varepsilon \varphi)(\overline{z})} &= \frac{\sum_i \varphi(F_{\varepsilon,i}^{-1}(z)) |D_{F_{\varepsilon,i}^{-1}(z)} F_\varepsilon|^{-1}}{\sum_i \varphi(F_{\varepsilon,i}^{-1}(\overline{z})) |D_{F_{\varepsilon,i}^{-1}(w)} F_\varepsilon|^{-1}} \\
&\le \exp[(\overline{\sigma}^{-1} a + \mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)) d_p(z, \overline{z})].
\end{aligned}
$$

It follows that if $a > a_c$ then $C_{a,p}$ is invariant under P_ε. □

Proof of Theorem 4.1. The existence of the absolutely continuous invariant probability measure is standard from the expansivity of F_ε. The regularity bound on the density immediately follows from Proposition 4.3 and from the observation (which can be found in [Via97]) that the cone $\mathcal{C}_{a,p}$ has finite diameter with respect to the projective Hilbert metric. This in particular means that P_ε is a contraction with respect to this metric and has a fixed point. □

4.4. *Proof of Theorem A in the expanding case*

Property (27) of the invariant density provides an upper bound for its supremum which depends on the parameters of the network and proves the statement of Theorem A in the expanding case.

Proof of Theorem A. Since under conditions (H1)–(H3) in Theorem A, Proposition 4.3 holds, the invariant density ρ for F_ε belongs to the cone $C_{a,p}$ for $a > a_c$. Since ρ is a continuous density, it has to take value 1 at some point in its domain. This together with the regularity condition given by the cone implies that

$$
\sup_{z \in \mathbb{T}^N} \rho(z) \le \exp\{\mathcal{O}(\Delta^{-1}\delta L^{1+1/p}) + \mathcal{O}(M L^{1/p})\}.
$$

Using the upper bound (26), we obtain

$$
\begin{aligned}
\nu(\mathcal{B}_\varepsilon \times \mathbb{T}^M) &= \int_{\mathcal{B}_\varepsilon \times \mathbb{T}^M} \rho(z)\, dm_N(z) \le m_N(\mathcal{B}_\varepsilon \times \mathbb{T}^M) \sup_z \rho(z) \\
&\le \exp\{-\Delta \varepsilon^2/2 + \mathcal{O}(\Delta^{-1}\delta L^{1+1/p}) + \mathcal{O}(M L^{1/p})\}.
\end{aligned}
$$

From the invariance of ρ and thus of ν, for any $T \in \mathbb{N}$,

$$\begin{aligned}\nu\Big(\bigcup_{t=0}^{T} F_\varepsilon^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M)\Big) &\le (T+1)\nu(\mathcal{B}_\varepsilon \times \mathbb{T}^M)\\ &\le (T+1)\exp\{-\Delta\varepsilon^2/2 + \mathcal{O}(\Delta^{-1}\delta L^{1+1/p}) + \mathcal{O}(ML^{1/p})\}.\end{aligned}$$

Using again $\rho \in C_{a,p}$, and (H1) and (H3), we get

$$\begin{aligned}m_N\Big(\bigcup_{t=0}^{T} F_\varepsilon^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M)\Big) &= \int_{\bigcup_{t=0}^{T} F_\varepsilon^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M)} \rho^{-1}\, d\nu\\ &\le \nu\Big(\bigcup_{t=0}^{T} F_\varepsilon^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M)\Big)\exp\{\mathcal{O}(\Delta^{-1}\delta L^{1+1/p}) + \mathcal{O}(ML^{1/p})\}\\ &\le (T+1)\exp\{-\Delta\varepsilon^2/2 + \mathcal{O}(\Delta^{-1}\delta L^{1+1/p}) + \mathcal{O}(ML^{1/p})\}\\ &\le (T+1)\exp\{-\Delta\varepsilon^2/2 + \mathcal{O}(\eta)\Delta\},\end{aligned}$$

where we use (H4) to obtain the last inequality. Hence, the set

$$\Omega_T = \mathbb{T}^N \setminus \bigcup_{t=0}^{T} F_\varepsilon^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M)$$

for $\eta > 0$ sufficiently small satisfies the assertion of the theorem. □

5. Proof of Theorem A when some reduced maps have hyperbolic attractors

In this section, we allow for the situation where some (or possibly all) reduced maps have periodic attractors. For this reason, we introduce the new structural parameter $M_u \in \mathbb{N}_0$ such that, after renaming the hub nodes, the reduced dynamics g_j is expanding for $1 \le j \le M_u$, while for $M_u < j \le M$, g_j has a hyperbolic periodic attractor Λ_j. Let us also define $M_s = M - M_u$. We also assume that g_j are (n, m, λ, r)-hyperbolic with $n = 1$. In Lemma 5.6 we will show how to drop this assumption.

As in the previous section, the goal is to prove the existence of a set of large measure whose points take a long time to enter the set $\mathcal{B}_\varepsilon$ where fluctuations are above the threshold. To achieve this, we study the ergodic properties of F_ε restricted to a certain forward invariant set $\mathcal{S}$ and prove that the statement of Theorem A holds true for initial conditions taken in this set. Then in Section 5.8 we extend the reasoning to the remainder of the phase space and prove the full statement of the theorem.

For simplicity we will sometimes write (z_u, z_s) for a point in $\mathbb{T}^{L+M_u} \times \mathbb{T}^{M_s} = \mathbb{T}^N$ and $z_u = (x, y_u) \in \mathbb{T}^L \times \mathbb{T}^{M_u}$. Let

$$\pi_u : \mathbb{T}^N \to \mathbb{T}^{L+M_u} \quad \text{and} \quad \Pi_u : \mathbb{R}^N \to \mathbb{R}^{L+M_u}$$

be respectively the (canonical) projection on the first $L + M_u$ coordinates and its differential.

We begin by pointing out the existence of an invariant set.

Lemma 5.1. *As before, for* $j \in \{M_u+1, \dots, M\}$*, let* Λ_j *be the attracting sets of* g_j *and* $\Upsilon = \mathbb{T} \setminus W_s(\Lambda_j)$*. There exist* $\lambda \in (0,1)$ *and* $\varepsilon_\Lambda, r_0 > 0$ *such that for each* j *in* $\{M_u+1, \dots, M\}$ *and each* $|r| < r_0$*,*

(i) $|Dg_j(y)| < \lambda < 1$ *for every* $y \in U_j$ *and* $g_j(x) + r \in U_j$ *for every* $x \in U_j$*, where* U_j *is the* ε_Λ*-neighborhood of* Λ_j*,*
(ii) $|Dg_j| > \lambda^{-1}$ *on the* ε_Λ*-neighborhood of* Υ_j *for all* $j \in [M_u+1, M]$*.*

Proof. The first assertion in (i), and (ii), follow from continuity of Dg_j. Fix $x \in U_j$ and, $r \in (-r_0, r_0)$. From the definition of U_j, there exists $y \in \Lambda_j$ such that $d(x, y) < \varepsilon_\Lambda$. From the contraction property, $d(g_j(x), g_j(y)) < \lambda d(x, y) < \lambda \varepsilon_\Lambda$, and choosing $r_0 < (1-\lambda)\varepsilon_\Lambda$ we get

$$d(g_j(x) + r, g_j(y)) < \lambda \varepsilon_\Lambda + r_0 < \varepsilon_\Lambda.$$

From the invariance of Λ_j, $g_j(y) \in \Lambda_j$, and the lemma follows. $\square$

Let

$$\mathcal{R} := U_{M_u+1} \times \cdots \times U_M \quad \text{and} \quad \mathcal{S} := \mathbb{T}^{L+M_u} \times \mathcal{R} \subset \mathbb{T}^N. \tag{36}$$

Lemma 5.1 implies that provided the ε from the truncated system is below $r_0/2$, the set $\mathcal{S}$ is forward invariant under F_ε. It follows that for each attracting periodic orbit $O(z_s)$ of $g_{M_u+1} \times \cdots \times g_M : \mathbb{T}^{M_s} \to \mathbb{T}^{M_s}$, the endomorphism F_ε has a fat solenoidal invariant set. Indeed, take the union U of the connected components of $\mathcal{R}$ containing $O(z_s)$. Then by the previous lemma, $F_\varepsilon(\mathbb{T}^{L+M_u} \times U) \subset \mathbb{T}^{L+M_u} \times U$. The set $\bigcap_{n \geq 0} F_\varepsilon^n(\mathbb{T}^{L+M_u} \times U)$ is the analogue of the usual solenoid but with self-intersections (see Figure 5). An analogous situation, but where the map is a skew product, is studied in [Tsu01]. The set $\bigcap_{n \geq 0} F_\varepsilon^n(\mathbb{T}^{L+M_u} \times U)$ will support an invariant measure:

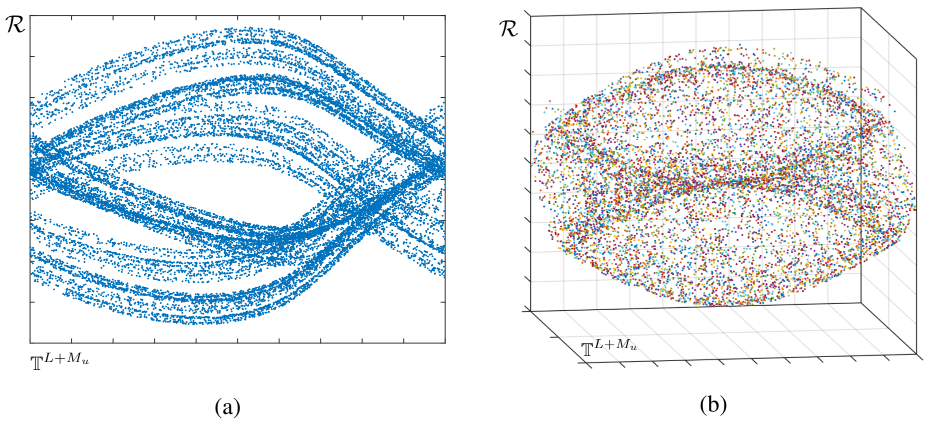

Fig. 5. Approximate 2D and 3D representations of one component of the attractor of F_ε.

Theorem 5.1. *Under conditions (H1)–(H4) of Theorem A with $\eta > 0$ sufficiently small,*

- *for every attracting periodic orbit of $g_{M_u+1} \times \cdots \times g_M$, F_ε has an ergodic physical measure,*
- *for each such measure ν, the marginal $(\pi_u)_*\nu$ on $\mathbb{T}^{L+M_u} \times \{0\}$ has a density ρ satisfying, for all $z_u, \bar{z}_u \in \mathbb{T}^{L+M_u} \times \{0\}$,*

$$\frac{\rho(z_u)}{\rho(\bar{z}_u)} \leq \exp\{a d_p(z_u, \bar{z}_u)\}, \qquad a = \mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M),$$

- *these are the only physical measures for F_ε.*

This theorem will be proved in Subsection 5.5.

5.1. Strategy of the proof of Theorem A in the presence of hyperbolic attractors

For the time being, we restrict our attention to the case where the threshold of the fluctuations is below r_0 as defined in Lemma 5.1 and consider the map $F_\varepsilon|_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$ that we will still call F_ε by abuse of notation. The expression for F_ε is the same as in (20) and (21), but now the local phase space for the hubs with a non-empty attractor, $\{L + M_u + 1, \ldots, L + M = N\}$, is restricted to the open set $\mathcal{R}$.

Theorem A will follow from the following proposition.

Proposition 5.1. *For every $s_1 \in \mathbb{Z}$ and $j \in [M]$, the measure of the set*

$$\mathcal{B}_{\varepsilon,T}^{(s_1,j)} := \bigcup_{t=0}^{T} F_\varepsilon^{-t}(\mathcal{B}_\varepsilon^{(s_1,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S} \subset \mathbb{T}^N$$

is bounded as

$$m_N(\mathcal{B}_{\varepsilon,T}^{(s_1,j)}) \leq T \exp[-C\Delta\varepsilon^2 + \mathcal{O}(\Delta^{-1} L^{1+2/p} \delta^{1/q}) + \mathcal{O}(M L^{1/p})]. \tag{37}$$

To prove the above result, we first build families of stable and unstable invariant cones for F_ε in the tangent bundle of $\mathcal{S}$ (Proposition 5.2) which correspond to contracting and expanding directions for the dynamics, thus proving hyperbolic behaviour of the map. In Section 5.3 we define a class of manifolds tangent to the unstable cones whose regularity properties are kept invariant under the dynamics, and we study the evolution of densities supported on them under the action of F_ε. Bounding the Jacobian of the map restricted to the manifolds (Proposition 5.4) one can prove the existence of an invariant cone of densities (Proposition 5.5), which gives the desired regularity properties for the measures. Since the product structure of $\mathcal{B}_\varepsilon^{(s_1,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}$ is not preserved under preimages of F_ε^t, we approximate it with the set which is the union of global stable manifolds (Lemma 5.3). This last property is preserved under taking preimages. The bound in (37) will then be a consequence of estimates on the distortion of the holonomy map along stable leaves of F_ε (Proposition 5.6).

5.2. *Invariant cone fields for F_ε*

Proposition 5.2. *There exists $\eta_0 > 0$ such that if conditions (H1)–(H4) are satisfied with $\eta < \eta_0$, then there exists $C_\# > 0$ such that for every ε with*

$$0 < \varepsilon < \min\left\{\frac{1-\lambda}{C_\#}, \frac{\lambda^{-1}-1}{C_\#}, \varepsilon_0\right\}, \tag{38}$$

(i) *the constant cone fields*

$$\mathcal{C}^u_p := \left\{(u, w, v) \in \mathbb{R}^{L+M_u+M_s} \setminus \{0\} : \frac{\|v\|_{p,M_s}}{\|u\|_{p,L} + \|w\|_{p,M_u}} < \beta_{u,p}\right\} \tag{39}$$

and

$$\mathcal{C}^s_p := \left\{(u, v, w) \in \mathbb{R}^{L+M_u+M_s} \setminus \{0\} : \frac{\|v\|_{p,M_s}}{\|u\|_{p,L} + \|w\|_{p,M_u}} > \frac{1}{\beta_{s,p}}\right\} \tag{40}$$

with

$$\beta_{u,p} := \mathcal{O}(\Delta^{-1/p} M_s^{1/p}), \quad \beta_{s,p} := \max\{\mathcal{O}(\Delta^{-1} L^{1/p} \delta^{1/q}), \mathcal{O}(\Delta^{-1/p} M_u^{1/p})\}$$

satisfy $D_z F_\varepsilon(\mathcal{C}^u) \subset \mathcal{C}^u$ and $D_z F_\varepsilon^{-1}(\mathcal{C}^s) \subset \mathcal{C}^s$ for all $z \in \mathbb{T}^N$;

(ii) *there exist $\overline{\sigma}$ and $\overline{\lambda}$ such that, for every $z \in \mathbb{T}^N$,*

$$\frac{\|D_z F_\varepsilon(u, w, v)\|_p}{\|(u, w, v)\|_p} \geq \overline{\sigma} > 1, \quad \forall (u, w, v) \in \mathcal{C}^u_p, \tag{41}$$

$$\frac{\|D_z F_\varepsilon(u, w, v)\|_p}{\|(u, w, v)\|_p} \leq \overline{\lambda} < 1, \quad \forall (u, w, v) \in \mathcal{C}^s_p. \tag{42}$$

Remark 5.1. We have constructed the map F_ε in such a way that when the network is η-heterogeneous with η very small, it results to be "close" to the product of uncoupled factors equal to f for the coordinates corresponding to low degree nodes, and equal to g_j for the coordinates of the hubs. This is reflected by the width of the invariant cones which can be chosen to be very small for η tending to zero, so that $\mathcal{C}^u_p$ and $\mathcal{C}^s_p$ are very narrow around their respective axes $\mathbb{R}^{L+M_u} \oplus \{0\}$ and $\{0\} \oplus \mathbb{R}^{M_s}$.

Corollary 5.1. *Under the assumptions of the previous proposition, $\pi_u \circ F^n_\varepsilon : \mathbb{T}^{L+M_u} \times \{0\} \to \mathbb{T}^{L+M_u}$ is a covering map of degree $\sigma^{n(L+M_u)}$ where σ is the degree of the local map.*

Proof. This follows from the proposition, because $\mathbb{T}^{L+M_u} \times \{0\}$ is tangent to the unstable cone, and thus $\pi_u \circ F^n_\varepsilon$ is a local diffeomorphism between compact manifolds. This implies that every point of $\mathbb{T}^{L+M_u}$ has the same number of preimages, and this number equals the degree of the map. Then observe that there is a homotopy bringing $\pi_u \circ F_\varepsilon$ to the $(L + M_u)$-fold uncoupled product of identical copies of the map f^n. The homotopy is obtained by continuously deforming the map letting the coupling strength α go to zero.

Since degree is a homotopy invariant and $\pi_u \circ F_\varepsilon$ is homotopic to the $(L + M_u)$-fold uncoupled product of identical copies of the map f^n, we have

$$\deg \pi_u \circ F_\varepsilon = \deg \underbrace{f^n \times \cdots \times f^n}_{L+M_u \text{ times}} = \sigma^{n(L+\mathcal{M}_u)}. \qquad \square$$

Proof of Proposition 5.2. (i) The expression for the differential of the map F_ε is the same as in (63). Take $(u, w, v) \in \mathbb{R}^L \times \mathbb{R}^{M_u} \times \mathbb{R}^{M_s}$, and suppose $(u', w', v')^t := D_z F_\varepsilon(u, v, w)^t$. Then

$$\begin{aligned}
u_i' &= \left[f'(x_i) + \frac{\alpha}{\Delta} \sum_{m=1}^{M} A_{im}^{lh} h_1 + \frac{\alpha}{\Delta} \sum_{\ell=1}^{L} A_{i\ell}^{ll} h_1 \right] u_i + \frac{\alpha}{\Delta} \sum_{\ell=1}^{L} A_{i\ell}^{ll} h_1 u_\ell \\
&\quad + \frac{\alpha}{\Delta} \sum_{m=1}^{M_u} A_{im}^{lh} h_2 w_m + \frac{\alpha}{\Delta} \sum_{m=M_u+1}^{M} A_{im}^{lh} h_2 v_m, \qquad 1 \le i \le L, \\
w_j' &= \sum_{\ell=1}^{L} \partial_{x_\ell} \xi_{j,\varepsilon} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^{M_u} A_{jm}^{hh} h_2 w_m + \frac{\alpha}{\Delta} \sum_{m=M_u+1}^{M} A_{jm}^{hh} h_2 v^m \\
&\quad + \left[\partial_{y_j} \xi_{j,\varepsilon} + \frac{\alpha}{\Delta} \sum_{m=1}^{M} A_{jm}^{hh} h_2 \right] w_j, \qquad 1 \le j \le M_u, \\
v_j' &= \sum_{\ell=1}^{L} \partial_{x_\ell} \xi_{j,\varepsilon} u_\ell + \frac{\alpha}{\Delta} \sum_{m=1}^{M_u} A_{jm}^{hh} h_2 w_m + \frac{\alpha}{\Delta} \sum_{m=M_u+1}^{M} A_{jm}^{hh} h_2 v_m \\
&\quad + \left[\partial_{y_j} \xi_{j,\varepsilon} + \frac{\alpha}{\Delta} \sum_{m=1}^{M} A_{jm}^{hh} h_2 \right] v_j, \qquad M_u < j \le M,
\end{aligned}$$

where we suppressed all dependences of those functions for which we use a uniform bound. Moreover,

$$\begin{aligned}
\|u'\|_{p,\mathbb{R}^L} &\ge (\sigma - \mathcal{O}(\Delta^{-1}\delta)) \|u\|_{p,\mathbb{R}^L} \\
&\quad - \mathcal{O}(\Delta^{-1} L^{1/p}) \max_{i \in [L]} \left[\sum_{\ell=1}^{L} A_{i\ell}^{ll} |u_\ell| + \sum_{m=1}^{M_u} A_{im}^{lh} |w_m| + \sum_{m=M_u+1}^{M} A_{im}^{lh} |v_m| \right] \\
&\ge (\sigma - \mathcal{O}(\Delta^{-1}\delta)) \|u\|_{p,\mathbb{R}^L} \\
&\quad - \mathcal{O}(\Delta^{-1} L^{1/p} \delta^{1/q}) (\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}} + \|v\|_{p,\mathbb{R}^{M_s}}) \\
\|w'\|_{p,\mathbb{R}^{M_u}} &\ge (\lambda^{-1} - C_\# \varepsilon - \mathcal{O}(\Delta^{-1} M)) \|w\|_{p,\mathbb{R}^{M_u}} \\
&\quad - \mathcal{O}(\Delta^{-1} M_u^{1/p}) \max_{1 \le j \le M_u} \left[\sum_{\ell=1}^{L} A_{j\ell}^{hl} |u_\ell| + \sum_{m=1}^{M_u} A_{jm}^{hh} |w_m| + \sum_{m=M_u+1}^{M} A_{jm}^{hh} |v_m| \right] \\
&\ge (\lambda^{-1} - C_\# \delta - \mathcal{O}(\Delta^{-1} M)) \|w\|_{p,\mathbb{R}^{M_u}} \\
&\quad - \mathcal{O}(\Delta^{-1/p} M_u^{1/p}) (\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}} + \|v\|_{p,\mathbb{R}^{M_s}})
\end{aligned}$$

and analogously

$$\|v'\|_{p,\mathbb{R}^{M_s}} \le (\lambda + C_{\#}\varepsilon + \mathcal{O}(\Delta^{-1}M))\|v\|_{p,\mathbb{R}^{M_s}} \\ + \mathcal{O}(\Delta^{-1/p}M_s^{1/p})(\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}} + \|v\|_{p,\mathbb{R}^{M_s}}).$$

Suppose that (u, w, v) satisfies the cone condition $\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}} \ge \tau\|v\|_{p,\mathbb{R}^{M_s}}$ for some τ. Then

$$\begin{aligned}\frac{\|u'\|_{p,\mathbb{R}^L} + \|w'\|_{p,\mathbb{R}^{M_u}}}{\|v'\|_{p,\mathbb{R}^{M_s}}} &\ge \frac{\mathcal{F}_{11}(\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}}) - \mathcal{F}_{12}\|v\|_{p,\mathbb{R}^{M_s}}}{\mathcal{F}_{21}\|v\|_{p,\mathbb{R}^{M_s}} + \mathcal{F}_{22}(\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}})} \\ &\ge \frac{\mathcal{F}_{11} - \tau^{-1}\mathcal{F}_{12}}{\tau^{-1}\mathcal{F}_{21} + \mathcal{F}_{22}}\end{aligned}$$

with

$$\begin{aligned}\mathcal{F}_{11} :=\ & \min\{\sigma - \mathcal{O}(\Delta^{-1}\delta), \lambda^{-1} - C_{\#}\varepsilon - \mathcal{O}(\Delta^{-1}M)\} \\ & - \max\{\mathcal{O}(\Delta^{-1}L^{1/p}\delta^{1/q}), \mathcal{O}(\Delta^{-1/p}M_u^{1/p})\} \\ =\ & \min\{\sigma, \lambda^{-1} - C_{\#}\varepsilon\} - \mathcal{O}(\eta), \\ \mathcal{F}_{12} :=\ & \max\{\mathcal{O}(\Delta^{-1}L^{1/p}\delta^{1/q}), \mathcal{O}(\Delta^{-1/p}M_u^{1/p})\} = \mathcal{O}(\eta), \\ \mathcal{F}_{21} :=\ & \lambda + C_{\#}\varepsilon + \mathcal{O}(\Delta^{-1}M) + \mathcal{O}(\Delta^{-1/p}M_s^{1/p})) = \lambda + C_{\#}\varepsilon + \mathcal{O}(\eta), \\ \mathcal{F}_{22} :=\ & \mathcal{O}(\Delta^{-1/p}M_s^{1/p})) = \mathcal{O}(\eta),\end{aligned}$$

where we use (H1)–(H4). The cone $\mathcal{C}_p^u$ is forward invariant iff $\|u'\|_{p,\mathbb{R}^L} + \|w'\|_{p,\mathbb{R}^{M_u}} \ge \tau\|v'\|_{p,\mathbb{R}^{M_s}}$ and therefore if

$$\mathcal{F}_{11} - \tau^{-1}\mathcal{F}_{12} \ge \mathcal{F}_{21} + \tau\mathcal{F}_{22}. \tag{43}$$

Hence we find $C_* > 0$ such that if $\tau = C_*/\mathcal{F}_{22}$ the inequality (43) is satisfied provided (38) holds and $\eta > 0$ is small enough because then $\mathcal{F}_{11} > \mathcal{F}_{21}$.

Now let us check when the cone $\mathcal{C}_p^s$ is backward invariant. Suppose that $\|u'\|_{p,\mathbb{R}^L} + \|w'\|_{p,\mathbb{R}^{M_u}} \le \tau\|v'\|_{p,\mathbb{R}^{M_s}}$. Then

$$\mathcal{F}_{11}\frac{\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}}}{\|v\|_{p,\mathbb{R}^{M_s}}} - \mathcal{F}_{12} \le \tau\mathcal{F}_{21} + \tau\mathcal{F}_{22}\frac{\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}}}{\|v\|_{p,\mathbb{R}^{M_s}}},$$

so

$$\frac{\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}}}{\|v\|_{p,\mathbb{R}^{M_s}}} \le \frac{\mathcal{F}_{12} + \tau\mathcal{F}_{21}}{\mathcal{F}_{11} - \tau\mathcal{F}_{22}},$$

and imposing yet again

$$\tau^{-1}\mathcal{F}_{12} + \mathcal{F}_{21} \le \mathcal{F}_{11} - \tau\mathcal{F}_{22} \tag{44}$$

implies that $\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}} \le \tau\|v\|_{p,\mathbb{R}^{M_s}}$. Taking $\tau = C_*\mathcal{F}_{12}$ with $C_* > 0$ small, we find that $\mathcal{C}_p^s$ is backward invariant (provided as before that (38) holds and $\eta > 0$ is small).

(ii) Take $(u, w, v) \in \mathcal{C}_p^u$ such that $\|(u, w, v)\|_p = 1$. From the above computations, and applying the cone condition, we get

$$\begin{aligned}
\|u'\|_{p,\mathbb{R}^L} + \|w'\|_{p,\mathbb{R}^{M_u}} + \|v'\|_{p,\mathbb{R}^{M_s}} &\geq \|u'\|_{p,\mathbb{R}^L} + \|w'\|_{p,\mathbb{R}^{M_u}} \\
&\geq \mathcal{F}_{11}(\|u\|_{p,\mathbb{R}^L} + \|w\|_{p,\mathbb{R}^{M_u}}) - \mathcal{F}_{12}\|v\|_{p,\mathbb{R}^{M_s}} \\
&\geq \mathcal{F}_{11}(1 - \beta_{u,p}) - \mathcal{F}_{12}\beta_{u,p} \\
&\geq \min\{\sigma, \lambda^{-1} - C_{\#}\varepsilon\} - \mathcal{O}(\eta) - \mathcal{O}(\eta^2) \qquad (45)
\end{aligned}$$

where to obtain (45) we kept only the largest order in the parameters of the network, after substituting the expressions for $\mathcal{F}_{11}$ and $\mathcal{F}_{12}$. This means that under (H1)–(H4), if $\eta > 0$ is sufficiently small, (41) will be satisfied. Choosing now $(u, v, w) \in \mathcal{C}_p^s$ of unit norm we get

$$\begin{aligned}
\|u'\|_{p,\mathbb{R}^L} + \|w'\|_{p,\mathbb{R}^{M_u}} + \|v'\|_{p,\mathbb{R}^{M_s}} &\leq (1 + \beta_{s,p})\Delta^{-1}\|v'\|_{p,\mathbb{R}^{M_s}} \\
&\leq \lambda + C_{\#}\varepsilon + \mathcal{O}(\Delta^{-1}M) + \mathcal{O}(\Delta^{-1/p}M^{1/p}) + \beta_{s,p} \leq \lambda + C_{\#}\varepsilon + \mathcal{O}(\eta)
\end{aligned}$$

and again whenever (H1)–(H4) are satisfied with $\eta > 0$ sufficiently small, (42) holds. □

5.3. *Admissible manifolds for F_ε*

As in the diffeomorphism case, the existence of the stable and unstable cone fields implies that the endomorphism F_ε admits a natural measure.

To determine the measure of the set $\mathcal{B}_\varepsilon \times \mathbb{T}^M$ for one of these measures we need to estimate how much the marginals on the coordinates of the low degree nodes differ from Lebesgue measure. To do this we look at the evolution of densities supported on admissible manifolds, namely manifolds whose tangent space is contained in the unstable cone and whose geometry is controlled. To control the geometry locally, we invoke the Hadamard–Perron graph transform argument (see for example [Shu13, KH95]) (Appendix D) which implies that manifolds tangent to the unstable cone which are locally graphs of functions in a given regularity class, are mapped by the dynamics into manifolds which are locally graphs of functions in the same regularity class.

As before, $\mathbb{T} = \mathbb{R}/\sim$ with $x_1 \sim x_2$ when $x_1 - x_2 \in \mathbb{Z}$, so each point in $\mathbb{T}$ can be identified with a point in $[0, 1)$. Define $I = (0, 1)$.

Definition 5.1 (Admissible manifolds $\mathcal{W}_{p,K_0}$). For every $K_0 > 0$ and $1 \leq p \leq \infty$ we say that a manifold W of $\mathcal{S}$ is *admissible* and belongs to the set $\mathcal{W}_{p,K_0}$ if there exists a differentiable function $E : I^{L+M_u} \to \mathcal{R}$ with Lipschitz differential such that

- W is the graph $(\mathrm{id}, E)(I^{L+M_u})$ of E,
- $D_{z_u}E(\mathbb{R}^{L+M_u}) \subset \mathcal{C}_p^u$ for all $z_u \in I^{L+M_u}$,
- $\|DE\|_{\mathrm{Lip}} := \sup_{z_u \neq \bar{z}_u} \dfrac{\|D_{z_u}E - D_{\bar{z}_u}E\|_p}{d_p(z_u, \bar{z}_u)} \leq K_0,$

  where, by abuse of notation, we denote by $\|\cdot\|_p$ the operator norm of linear transformations from $(\mathbb{R}^{L+M_u}, \|\cdot\|_{p,\mathbb{R}^L} + \|\cdot\|_{p,\mathbb{R}^{M_u}})$ to $(\mathbb{R}^{M_s}, \|\cdot\|_{p,\mathbb{R}^{M_s}})$.

Proposition 5.3. *Under conditions (H1)–(H4), for $\eta > 0$ sufficiently small, there is K_u uniform in the network parameters such that for all $z_1, z_2 \in \mathcal{S}$ the norm*

$$\|D_{z_1}F_\varepsilon - D_{z_2}F_\varepsilon\|_{u,p} := \sup_{(u,w,v)\in\mathcal{C}_p^u} \frac{\|(D_{z_1}F_\varepsilon - D_{z_2}F_\varepsilon)(u,w,v)\|_p}{\|(u,w,v)\|_p}$$

satisfies

$$\|D_{z_1}F_\varepsilon - D_{z_2}F_\varepsilon\|_{u,p} \le K_u d_\infty(z_1, z_2).$$

Proof. Notice that from the regularity assumptions on the coupling function h, we can write the entries of $D_{z_1}F_\varepsilon - D_{z_2}F_\varepsilon$ as

$$[D_{z_1}F_\varepsilon - D_{z_2}F_\varepsilon]_{k\ell} = \begin{cases} \left[\sum_{\ell=1}^{L} \mathcal{O}(\Delta^{-1})A^{ll}_{k\ell} + \sum_{m=1}^{M} \mathcal{O}(\Delta^{-1})A^{lh}_{km}\right] d_\infty(z_1,z_2), & k=\ell\le L,\\ \mathcal{O}(\Delta^{-1})A^{ll}_{k\ell} d_\infty(z_1,z_2), & k\ne\ell\le L,\\ \mathcal{O}(\Delta^{-1})A^{lh}_{k(\ell-L)} d_\infty(z_1,z_2), & k\le L,\ \ell>L,\\ \mathcal{O}(\Delta^{-1})A^{hl}_{k\ell} d_\infty(z_1,z_2), & k>L,\ \ell\le L,\\ \mathcal{O}(\Delta^{-1})A^{hh}_{(k-L)(\ell-L)} d_\infty(z_1,z_2), & k\ne\ell>L,\\ [\mathcal{O}(1)+\mathcal{O}(\Delta^{-1}M)] d_\infty(z_1,z_2), & k=\ell>L, \end{cases} \tag{46}$$

Take $(u,w,v)\in\mathcal{C}_p^u$ with $\|(u,w,v)\|_p = 1$ and $(u',w',v')^t = (D_{z_1}F_\varepsilon - D_{z_2}F_\varepsilon)(u,w,v)^t$. Then

$$\begin{aligned} u'_i &= \mathcal{O}(\Delta^{-1})\Big[\sum_{\ell=1}^{L} A^{ll}_{k\ell} + \sum_{m=1}^{M} A^{lh}_{km}\Big] u_i d_\infty(z_1,z_2)\\ &\quad + \mathcal{O}(\Delta^{-1})\Big[\sum_{\ell=1}^{L} A^{ll}_{i\ell} u_n + \sum_{m=1}^{M_u} A^{lh}_{im} w_m + \sum_{m=1}^{M_s} A^{lh}_{i(m+M_u)} v_m\Big] d_\infty(z_1,z_2), \quad 1\le i\le L,\\ w'_j &= [\mathcal{O}(1)+\mathcal{O}(\Delta^{-1}M)] w_j d_\infty(z_1,z_2)\\ &\quad + \mathcal{O}(\Delta^{-1})\Big[\sum_{i} A^{hl}_{ji} u_i + \sum_{m=1}^{M_u} A^{hh}_{jm} w_m + \sum_{m=1}^{M_s} A^{hh}_{j(m+M_u)} v_m\Big] d_\infty(z_1,z_2), \quad 1\le j\le M_u,\\ v'_j &= [\mathcal{O}(1)+\mathcal{O}(\Delta^{-1}M)] v_j d_\infty(z_1,z_2)\\ &\quad + \mathcal{O}(\Delta^{-1})\Big[\sum_{\ell=1}^{L} A^{hl}_{j\ell} u_\ell + \sum_{m=1}^{M_u} A^{hh}_{jm} w_m + \sum_{m=M_u+1}^{M} A^{hh}_{jm} v_m\Big] d_\infty(z_1,z_2), \quad 1\le j\le M_s, \end{aligned}$$

so that

$$\begin{aligned} \|u'\|_{p,\mathbb{R}^L} &\le \mathcal{O}(\Delta^{-1}\delta N^{1/p}) d_\infty(z_1,z_2) = \mathcal{O}(\eta) d_\infty(z_1,z_2),\\ \|w'\|_{p,\mathbb{R}^{M_u}} &\le \mathcal{O}(1) d_\infty(z_1,z_2),\\ \|v'\|_{p,\mathbb{R}^{M_s}} &\le \mathcal{O}(1) d_\infty(z_1,z_2), \end{aligned}$$

which implies the proposition. □

Lemma 5.2. *Suppose that $K_0 > \mathcal{O}(K_u)$ and W is an embedded $(L+M_u)$-dimensional torus which is the closure of $W_0 \in \mathcal{W}_{p,K_0}$. Then, for every $n \in \mathbb{N}$, $F_\varepsilon^n(W)$ is the closure of a finite union of manifolds, $W_{n,k} \in \mathcal{W}_{p,K_0}$, $k \in \mathcal{K}_n$ (and the difference $F_\varepsilon^n(W) \setminus \bigcup_{k\in\mathcal{K}_n} W_{n,k}$ is a finite union of manifolds of lower dimension).*

Proof. As in Corollary 5.1, since $\pi_u|_{W_0}$ is a diffeomorphism, the map $\pi_u \circ F_\varepsilon^n \circ \pi_u|_{W_0}^{-1} : \mathbb{T}^{L+M_u} \to \mathbb{T}^{L+M_u}$ is a well defined local diffeomorphism between compact manifolds, and therefore is a covering map. One can then find a partition $\{R_{n,k}\}_{k\in\mathcal{K}_n}$ of $\mathbb{T}^{L+M_u}$ such that $\pi_u \circ F_\varepsilon^n \circ \pi_u|_{W_0}^{-1}(R_{n,k}) = \mathbb{T}^{L+M_u}$ and, defining $W_{n,k} := \pi_u \circ F_\varepsilon^n \circ \pi|_{W_0}^{-1}(R_{n,k}^o)$, where $R_{n,k}^o$ is the interior of $R_{n,k}$, we have $\pi_u(W_{n,k}) = I^{L+M_u}$. From Proposition D.1 in Appendix D it follows that $W_{n,k} \in \mathcal{W}_{p,K_0}$ and $\{W_{n,k}\}_{k\in\mathcal{K}_n}$ is the desired partition. □

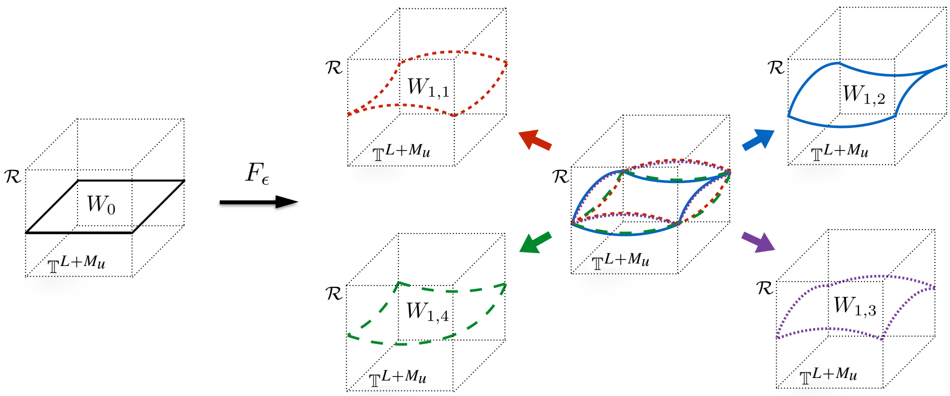

Fig. 6. The admissible manifold W_0 is mapped under F_ε to the union of submanifolds $W_{1,1}$, $W_{1,2}$, $W_{1,3}$, and $W_{1,4}$.

5.4. *Evolution of densities on the admissible manifolds for F_ε*

Recall that π_u and Π_u are projections on the first $L+M_u$ coordinates in $\mathbb{T}^N$ and $\mathbb{R}^N$ respectively. Given an admissible manifold $W \in \mathcal{W}_{p,K_0}$, which is the graph of the function $E : I^{L+M_u} \to \mathcal{R}$, for every $z_u \in I^{L+M_u}$ the map

$$\pi_u \circ F_\varepsilon \circ (\mathrm{id}, E)(z_u)$$

gives the evolution of the first $L+M_u$ coordinates of points in W. The Jacobian of this map is given by

$$J(z_u) = |\Pi_u \cdot D_{(\mathrm{id},E)(z_u)} F_\varepsilon \cdot (\mathrm{Id}, D_{z_u} E)|.$$

In the next proposition we upper bound the distortion of such a map.

Proposition 5.4. *Let $W \in W_{p,L}$ be an admissible manifold and suppose $z_u, \bar{z}_u \in I^{L+M_u}$. Then*

$$\left|\frac{J(z_u)}{J(\bar{z}_u)}\right| \le \exp\{[\mathcal{O}(\Delta^{-1}L^{1+1/p}\delta^{1/q}) + \mathcal{O}(M)]d_\infty(z_u, \bar{z}_u)\}.$$

Proof. We have

$$\frac{|J(z_u)|}{|J(\overline{z}_u)|} = \frac{|\Pi_u \cdot D_{(\mathrm{id},E)(z_u)}F_\varepsilon \cdot (\mathrm{Id},\, D_{z_u}E)|}{|\Pi_u \cdot D_{(\mathrm{id},E)(\overline{z}_u)}F_\varepsilon \cdot (\mathrm{Id},\, D_{z_u}E)|} \cdot \frac{|\Pi_u \cdot D_{(\mathrm{id},E)(\overline{z}_u)}F_\varepsilon \cdot (\mathrm{Id},\, D_{z_u}E)|}{|\Pi_u \cdot D_{(\mathrm{id},E)(\overline{z}_u)}F_\varepsilon \cdot (\mathrm{Id},\, D_{\overline{z}_u}E)|}$$
$$=: (\mathrm{A}) \cdot (\mathrm{B}).$$

(A) can be bounded with computations similar to the ones carried out in Proposition 4.2:

$$(\mathrm{A}) \leq \exp\{[\mathcal{O}(\Delta^{-1}\delta L) + \mathcal{O}(M)]d_\infty(z, \overline{z})\}.$$

To estimate (B) we also factor out the number $Df = \sigma$ from the first L columns of $\Pi_u D_{(\mathrm{id},E)(\overline{z}_u)}F_\varepsilon$, and $Dg_j(\overline{y}_{u,j})$ from the $(L+j)$-th column when $1 \leq j \leq M_u$ and thus obtain

$$(\mathrm{B}) = \frac{\sigma^L}{\sigma^L} \cdot \frac{\prod_{j=1}^{M_u} Dg_j}{\prod_{j=1}^{M_u} Dg_j} \cdot \frac{|\Pi_u \mathcal{D}((\mathrm{id},\, E)(\overline{z}_u)) \cdot (\mathrm{Id},\, D_{z_u}E)|}{|\Pi_u \mathcal{D}((\mathrm{id},\, E)(\overline{z}_u)) \cdot (\mathrm{Id},\, D_{\overline{z}_u}E)|}$$

where $\mathcal{D}(\cdot)$ is the same matrix defined in (31) apart from the last M_s columns which are kept equal to the corresponding columns of $D.F_\varepsilon$. The first two ratios trivially cancel. For the third factor we proceed in a fashion similar to previous computations using Proposition B.1 of the appendix. Defining $B := \mathcal{D}((\mathrm{id},\, E)(\overline{z}_u)) - \mathrm{Id}$, we are reduced to estimating

$$\frac{|\mathrm{Id} + \Pi_u \cdot B \cdot (\mathrm{Id},\, D_{z_u}E)|}{|\mathrm{Id} + \Pi_u \cdot B \cdot (\mathrm{Id},\, D_{\overline{z}_u}E)|}$$

where we use the fact that $\Pi_u \mathcal{D} \cdot (\mathrm{Id},\, D_{z_u}E) - \mathrm{Id} = \Pi_u B \cdot (\mathrm{Id},\, D_{z_u}E)$.

Since $\|(\mathrm{Id},\, D_{z_u}E)\|_p \leq 1 + \beta_{u,p}$ for any $z_u \in \mathcal{S}$, it follows, by choosing $\eta > 0$ sufficiently small in (H1)–(H4) and from (33), that

$$\|\Pi_u \cdot B \cdot (\mathrm{Id},\, D_{z_u}E)\|_p < \lambda < 1. \tag{47}$$

It is also rather immediate to upper bound the column norms of $\Pi_u \cdot B \cdot (0,\, D_{z_u}E - D_{\overline{z}_u}E)$:

$$\begin{aligned}\|\mathrm{Col}^i[\Pi_u \cdot B \cdot (0,\, D_{z_u}E - D_{\overline{z}_u}E)]\|_p &\leq \mathcal{O}(\Delta^{-1}L^{1/p}\delta^{1/q})\|DE\|_{\mathrm{Lip},p} d_p(z_u, \overline{z}_u)\\ &\leq \mathcal{O}(\Delta^{-1}M^{1/p})d_p(z_u, \overline{z}_u),\end{aligned}$$

so that by Proposition B.1, the overall estimate for (B) is

$$\frac{|\Pi_u \cdot \mathcal{D}((\mathrm{id},\, E)(\overline{z}_u)) \cdot (\mathrm{Id},\, D_{z_u}E)|}{|\Pi_u \cdot \mathcal{D}((\mathrm{id},\, E)(\overline{z}_u)) \cdot (\mathrm{Id},\, D_{\overline{z}_u}E)|} \leq \exp\{\mathcal{O}(\Delta^{-1}L^{1+1/p}\delta^{1/q})d_p(z_u, \overline{z}_u)\}. \tag{48}$$

□

5.5. *Invariant cone of densities on admissible manifolds for F_ε*

Take $W \in \mathcal{W}_{p,K_0}$. A density ρ on W is a measurable function $\varphi : W \to \mathbb{R}^+$ such that the integral of φ over W with respect to m_W is 1, where m_W is defined to be the measure obtained by restricting the volume form in $\mathbb{T}^N$ to W. The measure $\pi_{u*}(\varphi \cdot m_W)$ is absolutely continuous with respect to m_{L+M_u} on $\mathbb{T}^{L+M_u}$ and so its density $\varphi_u : \mathbb{T}^{L+M_u} \to \mathbb{R}^+$ is well defined.

Definition 5.2. For every $W \in \mathcal{W}_{p,K_0}$ and every $\varphi : W \to \mathbb{R}^+$ we define

$$\varphi_u := \frac{d\pi_{u*}(\varphi \cdot m_W)}{dm_{L+M_u}}.$$

Consider the set of densities

$$\mathcal{C}_{a,p}(W) := \left\{ \varphi : W \to \mathbb{R}^+ : \frac{\varphi_u(z_u)}{\varphi_u(\overline{z}_u)} \le \exp[a d_p(z_u, \overline{z}_u)] \right\}.$$

The above set consists of all densities on W whose projection on the first $L + M_u$ coordinates has the prescribed regularity property.

Proposition 5.5. *For every $a > a_c$ where*

$$a_c = \mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M), \tag{49}$$

$W \in \mathcal{W}_{p,K_0}$ and $\varphi \in \mathcal{C}_{a,p}(W)$ the following holds. Suppose that $\{W'_k\}_k$ is the partition of $F_\varepsilon(W)$ given by Lemma 5.2 and that W_k is a manifold of W such that $F_\varepsilon(W_k) = W'_k \in \mathcal{W}_{p,K_0}$. Then for every k, the density φ'_k on W'_k defined as

$$\varphi'_k := \frac{1}{\int_{W_k} \varphi \, dm_W} \frac{dF_{\varepsilon*}(\varphi|_{W_k} \cdot m_{W_k})}{dm_{W'_k}}$$

belongs to $\mathcal{C}_{a,p}(W'_k)$.

Proof. It is easy to verify that φ'_k is well defined. Let G_k be the inverse of the map $F_\varepsilon|_{W_k} : W_k \to W'_k$. From Definition 5.2 follows that

$$(\varphi'_k)_u := \frac{d(\pi_u \circ F_\varepsilon \circ (\mathrm{id}, E))_*(\varphi_u|_{\pi_u(W_k)} \cdot m_{L+M_u})}{dm_{L+M_u}}$$

where E is the map whose graph equals W. This implies that

$$(\varphi'_k)_u(z_u) = \frac{\varphi_u(G_k(z_u))}{J(G_k(z_u))}$$

and therefore

$$\begin{aligned}\frac{(\varphi'_k)_u(z_u)}{(\varphi'_k)_u(\overline{z}_u)} &= \frac{\varphi_u(G_k(z_u))}{\varphi_u(G_k(\overline{z}_u))} \frac{J(G_k(\overline{z}_u))}{J(G_k(z_u))} \\ &\le \exp[\overline{\sigma}^{-1} a d_p(z_u, \overline{z}_u)] \exp\{[\mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M)] d_p(z_u, \overline{z}_u)\} \\ &\le \exp\{[\overline{\sigma}^{-1} a + \mathcal{O}(\Delta^{-1} L^{1+1/p} \delta^{1/q}) + \mathcal{O}(M)] d_p(z_u, \overline{z}_u))\}.\end{aligned}$$

Taking a_c as in (49) yields the assertion. □

At this point we can prove that the system admits invariant physical measures and that their marginals on the first $L + M_u$ coordinates are in the cone $C_{a,p}$ for $a > a_c$. The main ingredients we use are Krylov–Bogolyubov's theorem and Hopf's argument [Wil12, KH95].

Proof of Theorem 5.1. Pick a periodic orbit $O(z_s)$ of $g_{M_u+1} \times \cdots \times g_M$ and let U be the union of the connected components of $\mathcal{R}$ containing points of $O(z_s)$. Pick $y_s \in U$ and take the admissible manifold $W_0 := \mathbb{T}^{L+M_u} \times \{y_s\} \in \mathcal{W}_{p,K_0}$. Consider a density $\rho \in \mathcal{C}_{a,p}(W_0)$ with $a > a_c$ such that the measure $\mu_0 := \rho \cdot m_W$ is the probability measure supported on W_0 with density ρ with respect to the Lebesgue measure on W_0. Consider the sequence $\{\mu_t\}_{t\in\mathbb{N}_0}$ of measures defined as

$$\mu_t := \frac{1}{t+1} \sum_{i=0}^{t} (F_\varepsilon^i)_* \mu_0.$$

From Lemma 5.2 we know that $F_\varepsilon^i(W_0) = \bigcup_{k\in\mathcal{K}_i} W_{i,k}$ modulo a negligible set with respect to $(F_\varepsilon^i)_*(\mu_0)$, and that

$$(F_\varepsilon^i)_*(\mu_0) = \sum_{k\in\mathcal{K}_i} (F_\varepsilon^i)_* \mu_0(W_{i,k}) \mu_{i,k},$$

where $\mu_{i,k}$ is a probability measure supported on $W_{i,k}$ for all i and $k \in \mathcal{K}_i$. It is a consequence of Proposition 5.5 that $\mu_{i,k} = \rho_{i,k} \cdot m_{W_{i,k}}$ with $\rho_{i,k} \in C_{a,p}(W_{i,k})$. Since F_ε is continuous, every subsequence of $\{\mu_t\}_{t\in\mathbb{N}_0}$ has a convergent subsequence in the set of all probability measures of $\mathcal{S}$ with respect to the weak topology ($\{\mu_t\}_{t\in\mathbb{N}_0}$ is tight). Let $\overline{\mu}$ be a probability measure which is the limit of a converging subsequence. By convexity of the cone $\mathcal{C}_{a,p}$ the second assertion of the theorem holds for $\overline{\mu}$.

Now let $V_1, \ldots, V_n$ be the components of U where n is the period of $O(z_s)$. Since all stable manifolds are tangent to a constant cone which has a very small angle (in particular less than $\pi/4$) with the vertical direction (corresponding to the last M_s directions of $\mathbb{T}^N$), they will intersect all horizontal tori $\mathbb{T}^{L+M_u} \times \{y\}$ with $y \in V_k$. It follows from the standard arguments that $\overline{\mu}$ has absolutely continuous disintegration on foliations of local unstable manifolds, which in the case of an endomorphism are defined on a set of histories called inverse limit sets (see [QXZ09] for details). Following the standard Hopf argument [Wil12, KH95], one first notices that given a point $x \in V_i$ on the support of $\overline{\mu}$, a history $\overline{x} \in (\mathbb{T}^N)^{\mathbb{N}}$, and a continuous observable φ, from the definition of $\overline{\mu}$, almost every point on the local unstable manifold associated to the selected history has a well defined forward asymptotic Birkhoff average (computed along $\overline{x}$) and every point on that stable manifold through x has the same asymptotic forward Birkhoff average. The aforementioned property of the stable leaves implies that any two unstable manifolds are crossed by the same stable leaf. This, together with absolute continuity of the stable foliation, implies that forward Birkhoff averages of φ are constant almost everywhere on the support of $\overline{\mu}$, which implies ergodicity. □

5.6. *Jacobian of the holonomy map along stable leaves of F_ε*

In order to prove Proposition 5.1 we need to upper bound the Jacobian of the holonomy map along stable leaves. It is known that for a C^2 uniformly hyperbolic (or even partially hyperbolic) diffeomorphism the holonomy map along the stable leaves is absolutely continuous with respect to the induced Lebesgue measure on the transversal to the leaves [HP05]. This can be easily generalised to the non-invertible case.

First of all, let us recall the definition of the holonomy map. We consider holonomies between manifolds tangent to the unstable cone.

Definition 5.3. Given embedded disks D_1 and D_2 of dimension $L + M_u$, tangent to the unstable cone $\mathcal{C}^u$, we define the *holonomy map* $\pi : D_1 \to D_2$ by

$$\pi(x) = W^s(x) \cap D_2.$$

As before, we define m_D to be the Lebesgue measure on D induced by the volume form on $\mathbb{T}^N$.

Remark 5.2. For the truncated dynamical system F_ε, fixing D_1, one can always find a sufficiently large D_2 such that the map π is well defined everywhere in D_1.

Proposition 5.6. *Given admissible embedded disks D_1 and D_2 tangent to the unstable cone, the holonomy map $\pi : D_1 \to D_2$ associated to F_ε is absolutely continuous with respect to m_{D_1} and m_{D_2}, the restrictions of Lebesgue measure to the two disks. Furthermore, if J_s is the Jacobian of π, then*

$$J_s(z) \leq \exp\left\{[\mathcal{O}(\Delta^{-1}L\delta) + \mathcal{O}(M)]\frac{1}{1-\lambda}d_\infty(z, \pi(z))\right\}, \quad \forall z \in D_1. \tag{50}$$

Proof. The absolute continuity follows from results in [Mn87] (see Appendix D), as well as from the estimate on the Jacobian. In fact, it is proven in [Mn87] that

$$J_s(z) = \prod_{k=0}^{\infty} \frac{\mathrm{Jac}(D_{z_k} F_\varepsilon|_{V_k})}{\mathrm{Jac}(D_{\overline{z}_k} F_\varepsilon|_{\overline{V}_k})}, \qquad \forall z \in D_1,$$

where $z_k := F_\varepsilon^k(z)$, $\overline{z}_k := F_\varepsilon^k(\pi(z))$, $V_k := T_{z_k} F_\varepsilon^k(D_1)$ and $\overline{V}_k := T_{\overline{z}_k} F_\varepsilon^k(D_2)$. Since D_1 and D_2 are tangent to the unstable cone, one can write $F_\varepsilon^k(D_1)$ and $F_\varepsilon^k(D_2)$ locally as graphs of functions $E_{1,k} : B_\delta^u(z_k) \to B_\delta^s(z_k)$ and $E_{2,k} : B_\delta^u(\overline{z}_k)) \to B_\delta^s(\overline{z}_k)$, with $E_{i,k}$ given by applying the graph transform on $E_{i,k-1}$, and $E_{i,0}$ such that $(\mathrm{Id}, D_z E_{i,0})(\mathbb{R}^{L+M_u}) = T_z D_i$.

For every $k \in \mathbb{N} \cup \{0\}$, we have $(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k}) \circ \Pi_u|_{V_k} = \mathrm{Id}|_{V_k}$ and so

$$\mathrm{Jac}(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})\, \mathrm{Jac}(\Pi_u|_{V_k}) = 1,$$

and analogously

$$\mathrm{Jac}(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})\, \mathrm{Jac}(\Pi_u|_{\overline{V}_k}) = 1.$$

Since

$$\begin{aligned}|\Pi_u \circ D_{z_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})| &= \mathrm{Jac}(\Pi_u \circ D_{z_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k}))\\ &= \mathrm{Jac}(\Pi_u|_{V_k})\,\mathrm{Jac}(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})\,\mathrm{Jac}(D_{z_k} F_\varepsilon|_{V_k}),\end{aligned}$$

and analogously

$$\begin{aligned}|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})| &= \mathrm{Jac}(\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k}))\\ &= \mathrm{Jac}(\Pi_u|_{\overline{V}_k})\,\mathrm{Jac}(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})\,\mathrm{Jac}(D_{\overline{z}_k} F_\varepsilon|_{\overline{V}_k}),\end{aligned}$$

we have

$$\begin{aligned}J_s(z) &= \prod_{k=0}^{\infty} \frac{\mathrm{Jac}(D_{z_k} F_\varepsilon|_{V_k})}{\mathrm{Jac}(D_{\overline{z}_k} F_\varepsilon|_{V_k})}\,\frac{\mathrm{Jac}(D_{\overline{z}_k} F_\varepsilon|_{V_k})}{\mathrm{Jac}(D_{\overline{z}_k} F_\varepsilon|_{\overline{V}_k})}\\ &= \prod_{k=0}^{\infty} \frac{|\Pi_u \circ D_{z_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}{|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}\,\frac{|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}{|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})|}.\end{aligned}$$

The first ratio can be deduced with minor changes from the estimate (48) in the proof of Proposition 5.4. So

$$\begin{aligned}\prod_{k=0}^{\infty} \frac{|\Pi_u \circ D_{z_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}{|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|} &\leq \exp\left\{[\mathcal{O}(\Delta^{-1} L\delta) + \mathcal{O}(M)] \sum_{k=0}^{\infty} d_p(z_k, \overline{z}_k)\right\}\\ &\leq \exp\left\{[\mathcal{O}(\Delta^{-1} L\delta) + \mathcal{O}(M)] \frac{1}{1-\overline{\lambda}} d_p(z_0, \overline{z}_0)\right\}\end{aligned}$$

where we have used the fact that z_0 and $\overline{z}_0$ lie on the same stable manifold and, by Proposition 5.2, $d_p(z_k, \overline{z}_k) \leq \overline{\lambda}^k d_p(z_0, \overline{z}_0)$.

To estimate the other ratio we make similar computations leading to the estimate in (48). Once more we factor out σ from the first N columns of $D_{\overline{z}_k}(F_\varepsilon)$ and, for all $j \in [M]$, $D_{y_j} g$ from the $(L+j)$-th column and thus

$$\begin{aligned}\frac{|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}{|\Pi_u \circ D_{\overline{z}_k} F_\varepsilon(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})|} &= \frac{\sigma^L}{\sigma^L} \cdot \frac{\prod_{j=1}^{M_u} Dg_j}{\prod_{j=1}^{M_u} Dg_j}\,\frac{|\Pi_u \circ D(\overline{z}_k)(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}{|\Pi_u \circ D(\overline{z}_k)(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})|}\\ &= \frac{|\Pi_u \circ \mathcal{D}(\overline{z}_k)(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})|}{|\Pi_u \circ \mathcal{D}(\overline{z}_k)(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k})|}\end{aligned}$$

where $\mathcal{D}(z_k)$ is defined as in (31). Defining, for every $k \in \mathbb{N}$,

$$B_k := \Pi_u \mathcal{D}(\overline{z}_k)(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k}) - \mathrm{Id} = \Pi_u(\mathcal{D}(\overline{z}_k) - \mathrm{Id})(\mathrm{Id}, D_{\pi_u(z_k)} E_{1,k})$$

and analogously

$$\overline{B}_k := \Pi_u \mathcal{D}(\overline{z}_k)(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k}) - \mathrm{Id} = \Pi_u(\mathcal{D}(\overline{z}_k) - \mathrm{Id})(\mathrm{Id}, D_{\pi_u(\overline{z}_k)} E_{2,k}),$$

we have proved in (47) that $\|\overline{B}_k\|, \|B_k\| \le \lambda < 1$. It remains to estimate the norm of the columns of $\overline{B}_k - B_k$. For all $\ell \in [L]$,

$$\begin{aligned}\|\mathrm{Col}^\ell[\overline{B}_k - B_k]\| &= \|\mathrm{Col}^\ell[\Pi_u(\mathcal{D}(\overline{z}_k) - \mathrm{Id})(0, D_{\pi_u(z_k)}E_{1,k} - D_{\pi_u(\overline{z}_k)}E_{2,k})]\| \\ &\le \left\|\Pi_u(\mathcal{D}(\overline{z}_k) - \mathrm{Id})|_{0\oplus\mathbb{R}^{M_s}}\right\| d_u(\overline{V}_k, V_k) \le \mathcal{O}(\Delta^{-1}\delta)d_u(\overline{V}_k, V_k).\end{aligned}$$

where we have used the fact that, as can be easily deduced from the definition of d_u in (69) of Appendix D, $\|(0, D_{\pi_u(z_k)}E_{1,k} - D_{\pi_u(\overline{z}_k)}E_{2,k})\| = d_u(\overline{V}_k, V_k)$. By Proposition B.1,

$$\frac{|\Pi_u \circ D_{\overline{z}_k}F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)}E_{1,k})|}{|\Pi_u \circ D_{\overline{z}_k}F_\varepsilon(\mathrm{Id}, D_{\pi_u(\overline{z}_k)}E_{2,k})|} \le \exp\{\mathcal{O}(L\Delta^{-1}\delta)d_u(\overline{V}_k, V_k)\}.$$

By Proposition D.3, if β_u is sufficiently small, then $d_u(\overline{V}_k, V_k) \le \lambda_* d_u(V_0, W_0)$ for some $\lambda_* < 1$, and this implies that

$$\begin{aligned}\prod_{k=0}^{\infty}\frac{|\Pi_u \circ D_{\overline{z}_k}F_\varepsilon(\mathrm{Id}, D_{\pi_u(z_k)}E_{1,k})|}{|\Pi_u \circ D_{\overline{z}_k}F_\varepsilon(\mathrm{Id}, D_{\pi_u(\overline{z}_k)}E_{2,k})|} &\le \exp\left\{\mathcal{O}(L\Delta^{-1}\delta)\sum_{k=0}^{\infty} d_u(\overline{V}_k, V_k)\right\} \\ &\le \exp\{\mathcal{O}(L\Delta^{-1}\delta)d_u(V_0, W_0)\} \\ &\le \exp\{\mathcal{O}(L\Delta^{-1}\delta\beta_u)\}.\end{aligned}$$

□

5.7. *Proof of Proposition 5.1*

The following result shows that $\mathcal{B}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}$, the set where fluctuations of the dynamics of a given hub exceed a given threshold, is contained in a set $\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}$ that is a union of global stable manifolds. This is important because even if the product structure of the former set is not preserved by taking preimages under (F_ε), the preimage of $\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}$ will be again the union of global stable manifolds. Furthermore, if β_s is sufficiently small, and this is provided by the heterogeneity conditions, the set $\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}$ will be "close" (topologically and with respect to the right measures) to $\mathcal{B}_\varepsilon^{(s,j)}$.

Lemma 5.3. *Consider $\mathcal{B}_\varepsilon^{(s,j)}$ as in (23). Then there exists a constant $C > 0$ such that*

$$\mathcal{B}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R} \subset \widetilde{\mathcal{B}}_\varepsilon^{(s,j)} := \bigcup_{z\in\mathcal{B}_\varepsilon^{(s,j)}\times\mathbb{T}^{M_u}\times\mathcal{R}} W^s(z) \subset \mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R} \tag{51}$$

with $\varepsilon_1 = \varepsilon + C_\# M^{1/p}\beta_{s,p}$.

Proof. The first inclusion is trivial. Take $z \in \mathcal{B}_\varepsilon^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}$ such that

$$\left|\frac{1}{\Delta}\sum_i A_{ji}^{hl}\theta_{s_1}(x_i) - \kappa_j\overline{\theta}_s\right| \ge \varepsilon|s_1|.$$

Since $W^s(z)$ is tangent to the stable cone $\mathcal{C}^s$, by (40) for any $z' \in W^s(z)$ we have $d_p(\pi_u(z'), z) \leq \beta_{s,p}$. This implies that

$$\left| \left| \frac{1}{\Delta} \sum_i A_{ji}^{hl} \theta_{s_1}(x_i) - \kappa_j \overline{\theta}_s \right| - \left| \frac{1}{\Delta} \sum_i A_{ji}^{hl} \theta_{s_1}(x_i') - \kappa_j \overline{\theta}_s \right| \right|$$
$$\leq \left| \frac{1}{\Delta} \sum_i A_{ji}^{hl} (\theta_{s_1}(x_i) - \theta_s(x_i')) \right| \leq \frac{1}{\Delta} \sum_i A_{ji}^{hl} |D\theta_{s_1}| d_p(\pi_u(z'), z) \leq |s_1| \mathcal{O}(M^{1/p} \beta_{s,p}),$$

proving the lemma. □

Proof of Proposition 5.1. As in the proof of Theorem 5.1, take an embedded $L + M_u$-torus $W_0 \in \mathcal{W}_{p,K_0}$ such that $\pi_u|_{W_0} : W_0 \to \mathbb{T}^{L+M_u}$ is a diffeomorphism, a density $\rho \in \mathcal{C}_{a,p}(W_0)$ with $a > a_c$ such that $\rho \mu_W$ is a probability measure, and the limit $\overline{\mu}$ of the sequence $\{\mu_t\}_{t \in \mathbb{N}_0}$ of measures defined as

$$\mu_t := \frac{1}{t+1} \sum_{i=0}^{t} (F_\varepsilon^i)_* \mu_0$$

is an SRB measure. From Lemma 5.2 we know that $F_\varepsilon^i(W_0) = \bigcup_{k \in \mathcal{K}_i} W_{i,k}$ modulo a negligible set with respect to $(F_\varepsilon^i)_*(\mu_0)$, and

$$(F_\varepsilon^i)_*(\mu_0) = \sum_{k \in \mathcal{K}_i} (F_\varepsilon^i)_* \mu_0(W_{i,k}) \mu_{i,k},$$

where $\mu_{i,k}$ is a probability measure supported on $W_{i,k}$ for all i and $k \in \mathcal{K}_i$. It is a consequence of Proposition 5.5 that $\mu_{i,k} = \rho_{i,k} \cdot m_{W_{i,k}}$ with $\rho_{i,k} \in C_{a,p}(W_{i,k})$. For every $t \in \mathbb{N}_0$,

$$\begin{aligned}
\mu_t(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}) &\leq \mu_t(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}) \\
&= \sum_{i=0}^{t} \sum_{k \in \mathcal{K}_i} \frac{(F_\varepsilon^i)_* \mu_0(W_{i,k})}{t+1} \mu_{i,k}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u} \times \mathcal{R}) \\
&= \sum_{i=0}^{t} \sum_{k \in \mathcal{K}_i} \frac{(F_\varepsilon^i)_* \mu_0(W_{i,k})}{t+1} \int_{\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}} \rho_{i,k} \, dm_{L+M_u} \\
&\leq \sum_{i=0}^{t} \sum_{k \in \mathcal{K}_i} \frac{(F_\varepsilon^i)_* \mu_0(W_{i,k})}{t+1} \exp[\mathcal{O}(\Delta^{-1} L^{1+2/p} \delta^{1/q}) + \mathcal{O}(M)] m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}) && (52) \\
&= \exp[\mathcal{O}(\Delta^{-1} L^{1+2/p} \delta^{1/q}) + \mathcal{O}(M)] m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}). && (53)
\end{aligned}$$

Since the set $\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}$ might not in general be measurable, in the above and in what follows we abuse the notation so that whenever the measure of such a set or one of its sections is computed, it should be understood as its outer measure. To prove the bound (52) we use the fact that $\rho_{i,k} \in C_{a,p}(W_{i,k})$ for $a > a_c$ and thus its supremum is upper bounded by $\exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]$; and (53) follows from the fact that

$$\sum_{i=0}^{t}\sum_{k\in\mathcal{K}_i}\frac{(F_\varepsilon^i)_*\mu_0(W_{i,k})}{t+1} = \sum_{i=0}^{t}\frac{(F_\varepsilon^i)_*\mu_0((F_\varepsilon)^i(W_0))}{t+1} = \sum_{i=0}^{t}\frac{1}{t+1} = 1. \tag{54}$$

Since the bound is true for every $t \in \mathbb{N}_0$, it is also true for the weak limit $\overline{\mu}$:

$$\overline{\mu}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}) \leq \exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}),$$

and since $\overline{\mu}$ is invariant, for all $t \in \mathbb{N}$ we have

$$\overline{\mu}(F_\varepsilon^{-t}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)})) \leq \exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}).$$

From (54) there exist $i \in \mathbb{N}$ and $k \in \mathcal{K}_i$ such that

$$\mu_{i,k}(F_\varepsilon^{-t}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)})) \leq \overline{\mu}(F_\varepsilon^{-t}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)}))$$

and thus

$$m_{W_{i,k}}(F_\varepsilon^{-t}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)})) \leq \exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}).$$

Now, pick $y_s \in \mathcal{R}$ and consider the holonomy map along the stable leaves, $\pi : W_{i,k} \to D_{y_s}$, between transversals $W_{i,k}$ and $D_{y_s} = \mathbb{T}^{L+M_u} \times \{y_s\} \subset \mathcal{C}^u$. We know from Proposition 5.6 that the Jacobian of π is bounded by (50) and thus

$$m_{D_{y_s}}(F_\varepsilon^{-t}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)})) \leq \exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}).$$

The above holds for every $y_s \in \mathcal{R}$, and so by Fubini

$$m_N(F_\varepsilon^{-t}(\widetilde{\mathcal{B}}_\varepsilon^{(s,j)})) \leq \exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}),$$

and from the first inclusion in (51) we obtain

$$m_N(F_\varepsilon^{-t}(\mathcal{B}_\varepsilon^{(s,j)})) \leq \exp[\mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})]m_{L+M_u}(\mathcal{B}_{\varepsilon_1}^{(s,j)} \times \mathbb{T}^{M_u}). \quad \square$$

5.8. *Proof of Theorem A*

In this section $F_\varepsilon : \mathbb{T}^N \to \mathbb{T}^N$ again denotes the truncated map defined on the whole phase space. Define the uncoupled map $\boldsymbol{f} : \mathbb{T}^N \to \mathbb{T}^N$ by

$$\boldsymbol{f}(x_1, \ldots, x_L, y_1, \ldots, y_M) := (f(x_1), \ldots, f(x_L), g_1(y_1), \ldots, g_M(y_M)).$$

The next lemma evaluates the ratios of the Jacobians of F_ε^t and $\boldsymbol{f}^t$ for any fixed $t \in \mathbb{N}$.

Lemma 5.4.

$$\frac{|D_z\boldsymbol{f}^t|}{|D_zF_\varepsilon^t|} \leq \exp[\mathcal{O}(ML\Delta^{-1}\delta) + \mathcal{O}(M^2)].$$

Proof. For all $i \in [t]$ define $z_i := \boldsymbol{f}^i(z)$, $\bar{z}_i := F_\varepsilon^i(z)$, and $z_0 = \bar{z}_0 := z$. Then

$$\frac{|D_z \boldsymbol{f}^t|}{|D_z F_\varepsilon^t|} = \frac{\prod_{k=0}^t |D_{z_k} \boldsymbol{f}|}{\prod_{k=0}^t |D_{\bar{z}_k} F_\varepsilon|} = \prod_{k=0}^t \frac{\sigma^L \prod_{m=1}^M D_{y_{k,m}} g_m}{|D_{\bar{z}_k} F_\varepsilon|}$$

$$= \prod_{k=0}^t \frac{\sigma^L \prod_{m=1}^M D_{\bar{y}_{k,m}} g_m \big(1 + \frac{D_{y_{k,m}} g_m - D_{\bar{y}_{k,m}} g_m}{D_{\bar{y}_{k,m}} g_m}\big)}{|D_{\bar{z}_k} F_\varepsilon|} \leq \exp[\mathcal{O}(M)] \prod_{k=0}^t \frac{1}{|\mathcal{D}(\bar{z}_k)|}$$

where $\mathcal{D}(z_i)$ is defined as in (31). Now, $1/|\mathcal{D}(\bar{z}_i)|$ can be estimated in the usual way by defining $B(z_i) := \mathcal{D}(z_i) - \mathrm{Id}$ and noticing that $1 = |\mathrm{Id} + 0|$. One can obtain, from the computations leading to (34),

$$\frac{1}{|\mathcal{D}(\bar{z}_i)|} \leq \exp\Big[\sum_{k=1}^N \mathrm{Col}^k[B(z_i)]\Big] \leq \exp[\mathcal{O}(L\Delta^{-1}\delta) + \mathcal{O}(M)],$$

and thus

$$\frac{|D_z \boldsymbol{f}^t|}{|D_z F_\varepsilon^t|} \leq \exp[\mathcal{O}(M)] \exp[\mathcal{O}(L\Delta^{-1}\delta) + \mathcal{O}(M)] \leq \exp[\mathcal{O}(L\Delta^{-1}\delta) + \mathcal{O}(M)]. \qquad \square$$

Lemma 5.5. *Let $A^2(\mathbb{T}, \mathbb{T})$ be the set of C^2 Axiom A endomorphisms of $\mathbb{T}$ endowed with the C^1 topology. Take a continuous curve $\gamma : [\alpha_1, \alpha_2] \to A^2(\mathbb{T}, \mathbb{T})$. Let Λ^α and Υ^α denote respectively the attractor and repellor of γ_α for all $\alpha \in [\alpha_1, \alpha_2]$. Then*

(i) *there exist uniform $\varepsilon_\Lambda > 0$ and $\lambda \in (0, 1)$ such that*

$$\big|\gamma'_\alpha|_{\Lambda^\alpha_{\varepsilon_\Lambda}}\big| < \lambda \quad \textit{and} \quad \big|\gamma'_\alpha|_{\Upsilon^\alpha_{\varepsilon_\Lambda}}\big| > \lambda^{-1}, \tag{55}$$

(ii) *there are uniform $r > 0$ and $\tau \in \mathbb{N}$ such that for all $\alpha \in [\alpha_1, \alpha_2]$, all sequences $\{\varepsilon_i\}_{i=0}^{\tau-1}$ with $\varepsilon_i \in (-r, r)$ and all points $x \in \mathbb{T} \setminus \Upsilon^\alpha_{\varepsilon_\Lambda}$, the orbit $\{x_i\}_{i=0}^{\tau}$ defined by*

$$x_0 := x \quad \textit{and} \quad x_i := \gamma_\alpha(x_{i-1}) + \varepsilon_i \tag{56}$$

satisfies $x_\tau \in \Lambda^\alpha_{\varepsilon_\Lambda}$.

Proof. The above lemma is quite standard [dMvS93] and can be easily proved by considering the sets

$$\bigcup_{\alpha \in [\alpha_1, \alpha_2]} \{\alpha\} \times \Lambda_\alpha \subset [\alpha_1, \alpha_2] \times \mathbb{T} \quad \text{and} \quad \bigcup_{\alpha \in [\alpha_1, \alpha_2]} \{\alpha\} \times \Upsilon_\alpha \subset [\alpha_1, \alpha_2] \times \mathbb{T}$$

and noticing that they are compact. Then from the C^1 assumption on the Axiom A map, it follows that all the stated quantities are uniformly bounded. $\square$

Proof of Theorem A. The proof is in two steps.

Step 1. Restricting F_ε to $\mathcal{S}$, we can use Proposition 5.1 to get an estimate of the Lebesgue measure of $\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}$. Define

$$\mathcal{B}_{\varepsilon,T,\tau} := \bigcup_{t=0}^{\tau} F_\varepsilon^{-t}(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S}.$$

To determine the Lebesgue measure of this set we compare it with the Lebesgue measure of

$$\mathcal{B}'_{\varepsilon,T,\tau} := \bigcup_{t=0}^{\tau} \boldsymbol{f}^{-t}(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S}.$$

For all $y \in \mathbb{T}^M$, the map $\boldsymbol{f}|_{\mathbb{T}^L\times\{y\}} : \mathbb{T}^L \times \{y\} \to \mathbb{T}^L \times \{(g_1(y_1), \dots, g_M(y_M))\}$ is an expanding map with constant Jacobian and thus measure preserving if we endow $\mathbb{T}^L \times \{y\}$ and $\mathbb{T}^L \times \{(g_1(y_1), \dots, g_M(y_M))\}$ with the induced Lebesgue measure. Fubini's theorem implies that for all $t \in [\tau]$,

$$m_N(\boldsymbol{f}^{-t}(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S}) \leq C(\tau)\frac{m_N(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R})}{m_N(\mathcal{S})}$$

where $C(\tau)$ is a constant depending on τ and uniform in the network parameters. And thus

$$m_N(\boldsymbol{f}^{-t}(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S}) \leq C_\# m_N(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}).$$

Now

$$m_N(F_\varepsilon^{-t}(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S}) \leq m_N(\boldsymbol{f}^{-t}(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \cap \mathcal{S}) \sup_{z\in\mathbb{T}^N} \frac{|D_z \boldsymbol{f}^t|}{|D_z F_\varepsilon^t|}.$$

By Lemma 5.4, assuming that $\tau \leq T$, we get

$$\begin{aligned} m_N(\mathcal{B}_{\varepsilon,T,\tau}) &\leq \sum_{t=0}^{\tau} \exp[\mathcal{O}(L\delta\Delta^{-1}) + \mathcal{O}(M)] C_\# m_N(\mathcal{B}_{\varepsilon,T} \times \mathbb{T}^{M_u} \times \mathcal{R}) \\ &\leq T \exp[-\mathcal{O}(\Delta^{-1})\varepsilon^2 + \mathcal{O}(\Delta^{-1}L^{1+2/p}\delta) + \mathcal{O}(ML^{1/p})]. \end{aligned}$$

Step 2. Define $\mathcal{U} \subset \mathbb{T}^N$ as

$$\mathcal{U} := \mathbb{T}^{L+M_u} \times \Upsilon_{\varepsilon_\Lambda}^{M_u+1} \times \cdots \times \Upsilon_{\varepsilon_\Lambda}^{M}.$$

Consider the system $G : \mathbb{T}^N \to \mathbb{T}^N$ obtained by redefining F_ε on $\mathcal{U}^c$ so that if $(x', y') = G(x, y)$, $\pi_u \circ G(x, y) = \pi_u \circ F_\varepsilon(x, y)$ (the evolution of the "expanding" coordinates is unchanged) and

$$y'_j = \hat{f}_j(y_j) + \alpha \sum_p g\left(\frac{1}{\Delta}\sum_n A_{jn}^{hl}\theta_s(x_n) - \kappa_j\overline{\theta}_s\right)\theta_s(y_j) + \frac{\alpha}{\Delta}\sum_{m=1}^{M} D_{jm}h(y_j, y_m) \bmod 1$$

where the reduced dynamics is (smoothly) modified to be globally expanding by putting $\hat{f}_j|_{\Upsilon_{\varepsilon_\Lambda}^j} := g_j|_{\Upsilon_{\varepsilon_\Lambda}^j}$ and $\hat{f}_j|_{\mathbb{T}\setminus\Upsilon_{\varepsilon_\Lambda}^j}$ redefined so that $|\hat{f}_j| \geq \lambda^{-1} > 1$ everywhere on $\mathbb{T}$. Evidently $G|_{\mathcal{U}} = F_\varepsilon|_{\mathcal{U}}$. We can then invoke the results of Section 4 to impose conditions on

η and ε to deduce global expansion of the map G (under suitable heterogeneity hypotheses) and the bounds on the invariant density obtained in that section. In particular, for all $T \in \mathbb{N}$ one has

$$m_N\Big(\bigcup_{t=0}^{T} G^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M)\Big) \leq T \exp\{-\Delta\varepsilon^2/2 + \mathcal{O}(\Delta^{-1}N^{1+2/p}\delta^{1/q}) + \mathcal{O}(MN^{1/p})\},$$

and this implies

$$m_N\Big(\bigcup_{t=0}^{T} G^{-t}(\mathcal{B}_\varepsilon \times \mathbb{T}^M) \cup \mathcal{B}_{\varepsilon,\tau,T}\Big) \\ \leq 2T \exp\{-\Delta\varepsilon^2/2 + \mathcal{O}(\Delta^{-1}L^{1+2/p}\delta^{1/q}) + \mathcal{O}(ML^{1/p})\},$$

concluding the proof of the theorem. □

5.9. *Mather's trick and proof of Theorem A when $n \neq 1$*

Until now we have assumed that the reduced maps g_j satisfied Definition 2.1 with $n = 1$. We now show that any $n \in \mathbb{N}$ will work by constructing an adapted metric via what is known as "Mather's trick" (see [dMvS93, Chapter 3, Lemma 1.3] or [HPS06]).

Lemma 5.6. *It is enough to prove Theorem A for $n = 1$.*

Proof. Assume that all g_j, $j = 1, \dots, M$, satisfy the assumptions in Definition 2.1 for some common (n, m, λ, r). Conditions (2) and (3) imply that one can smoothly conjugate each of the maps g_j so that $|D_x g_j| < \lambda$ for all $x \in N_r(\Lambda_j)$, and $|D_x g_j| > \lambda^{-1}$ for all $x \in N_r(\Upsilon_j)$. These conjugations are obtained by changing the metric ("Mather's trick"). In other words, there exists a smooth coordinate change $\varphi_j : \mathbb{T} \to \mathbb{T}$ such that for $\widetilde{g}_j := \varphi_j \circ g_j \circ \varphi_j^{-1}$ the properties from Definition 2.1 hold for $n = 1$. Moreover, there exists some uniform constant $C_\#$ only depending on (n, λ, r) such that the C^2 norms of φ_j and φ_j^{-1} are bounded by $C_\#$. Write $\widetilde{y}_j = \varphi_j(y_j)$ and $\widetilde{z} = (\widetilde{z}_1, \dots, \widetilde{z}_n) = (x_1, \dots, x_L, \widetilde{y}_1, \dots, \widetilde{y}_M)$. In these new coordinates, (10)–(13) become

$$x_i' = f(x_i) + \frac{\alpha}{\Delta}\sum_{\ell=1}^{L} A_{i\ell}^{ll} h(x_i, x_\ell) + \frac{\alpha}{\Delta}\sum_{m=1}^{M} A_{im}^{lh} \widetilde{h}(x_i, \widetilde{y}_m) \bmod 1, \qquad i = 1, \dots, L, \tag{57}$$

$$\widetilde{y}_j' = \widetilde{g}_j(\widetilde{y}_j) + \widetilde{\xi}_j(\widetilde{z}) \bmod 1, \qquad j = 1, \dots, M, \tag{58}$$

where

$$\widetilde{\xi}_j(\widetilde{z}) := \int_{\widetilde{g}_j(\widetilde{y}_j)}^{\widetilde{g}_j(\widetilde{y}_j)+\xi_j} D_t\varphi_j \, dt \quad \text{and} \quad \widetilde{h}(x, \widetilde{y}) := h(x, \varphi_j^{-1}(\widetilde{y})). \tag{59}$$

In fact

$$\widetilde{y}_j' = \varphi_j(y_j) = \varphi_j(g_j(y_j) + \xi_j) = \widetilde{g}_j(\widetilde{y}_j) + \int_{\widetilde{g}_j(\widetilde{y}_j)}^{\widetilde{g}_j(\widetilde{y}_j)+\xi_j} D_t\varphi_j \, dt.$$

Then we can set

$$\widetilde{\xi}_{j,\varepsilon} := \int_{\widetilde{g}_j(\widetilde{y}_j)}^{\widetilde{g}_j(\widetilde{y}_j)+\xi_{j,\varepsilon}} D_t\varphi_j\, dt$$

and define the truncated system as

$$x_i' = f(x_i) + \frac{\alpha}{\Delta}\sum_{\ell=1}^{L} A_{i\ell}^{ll} h(x_i, x_\ell) + \frac{\alpha}{\Delta}\sum_{m=1}^{M} A_{im}^{lh}\widetilde{h}(x_i, \widetilde{y}_m) \bmod 1, \qquad i = 1, \ldots, L, \tag{60}$$

$$\widetilde{y}_j' = \widetilde{g}_j(\widetilde{y}_j) + \widetilde{\xi}_{j,\varepsilon}(\widetilde{z}) \bmod 1, \qquad j = 1, \ldots, M. \tag{61}$$

Since the φ_j are C^2 with uniformly bounded C^2 norm, it immediately follows that $\widetilde{\xi}_{j,\varepsilon}$ has all the properties satisfied by $\xi_{j,\varepsilon}$ listed in Lemma A.1. Assuming that $|\widetilde{\xi}_{j,\varepsilon}(\widetilde{x}(t))| \leq \xi$ for all $0 \leq t \leq T$, we immediately obtain

$$|y_j' - g_j(y_j)| = |\varphi_j^{-1}[\widetilde{g}_j(\varphi_j(y_j)) + \widetilde{\xi}_{j,\varepsilon}(\widetilde{z})] - g_j(y_j)| \leq \mathcal{O}(\widetilde{x}_{j,\varepsilon}(\widetilde{z})) \leq \mathcal{O}(\xi). \qquad \square$$

5.10. Persistence of the result under perturbations

The picture presented in Theorem A is persistent under smooth random perturbations of the coordinates. Suppose that instead of the deterministic dynamical system $F : \mathbb{T}^N \to \mathbb{T}^N$ we have a stationary Markov chain $\{\mathcal{F}_t\}_{t\in\mathbb{N}}$ on some probability space $(\Omega, \mathbb{P})$ with transition kernel

$$\mathbb{P}(\mathcal{F}_{n+1} \in A \mid \mathcal{F}_n = z) := \int_A \varphi(y - F(z))\, dy$$

where $\varphi : \mathbb{T}^N \to \mathbb{R}^+$ is a density function. The Markov chain describes a random dynamical system where independent random noise distributed according to the density φ is added to the iteration of F. Take now the stationary Markov chain $\{\mathcal{F}_{\varepsilon,t}\}_{t\in\mathbb{N}}$ defined by the transition kernel

$$\mathbb{P}(\mathcal{F}_{\varepsilon,n+1} \in A \mid \mathcal{F}_{\varepsilon,n} = z) := \int_A \varphi(y - F_\varepsilon(z))\, dy$$

where we consider the truncated system instead of the original map in the deterministic drift of the process and restrict, for example, to the case where F_ε is uniformly expanding. The associated transfer operator can be written as $\mathcal{P}_\varepsilon = P_\varphi \circ P_\varepsilon$ where P_ε is the transfer operator for F_ε and

$$(P_\varphi\rho)(x) = \int \rho(y)\varphi(y - x)\, dy.$$

Let $C_{a,p}$ be a cone of densities invariant under P_ε as prescribed in Proposition 4.3. It is easy to see that this is also invariant under P_φ and thus under $\mathcal{P}_\varepsilon$. In fact, take $\rho \in C_{a,p}$.

Then

$$\frac{(P_\varphi\rho)(z)}{(P_\varphi\rho)(\bar z)} = \frac{\int \rho(y)\varphi(y-z)\,dy}{\int \rho(y)\varphi(y-\bar z)\,dy} = \frac{\int \rho(z-y)\varphi(y)\,dy}{\int \rho(\bar z-y)\varphi(y)\,dy}$$
$$\le \frac{\int \rho(\bar z-y)\exp\{ad_p(z,\bar z)\}\varphi(y)\,dy}{\int \rho(\bar z-y)\varphi(y)\,dy} = \exp\{ad_p(z,\bar z)\}.$$

This means that there exists a stationary measure for the chain with density belonging to $C_{a,p}$ and that the estimates of Section 4 hold for the measure of $\mathcal{B}_\varepsilon$. This allows us to conclude that the hitting times of $\mathcal{B}_\varepsilon$ satisfy the same type of bound as in the proof of Theorem A. Notice the independence of the above from the choice of density φ for the noise. This implies that all the arguments continue to hold independently of the size of the noise, which, however, contributes to spoil the low-dimensional approximation for the hubs in that it randomly perturbs it.

6. Conclusions and further developments

Heterogeneously Coupled Maps (HCM) are ubiquitous in applications. Because of the heterogeneous structure and lack of symmetries in the graph, most of the previously available results and techniques cannot be directly applied to this situation. Even if the behaviour of the local maps is well understood, once they are coupled in a large network, a rigorous description of the system becomes a major challenge, and numerical simulations are used to obtain information on the dynamics.

The ergodic theory of high-dimensional systems presents many difficulties including the choice of reference measure and dependence of decay of correlations on the system size. We exploited the heterogeneity to obtain rigorous results for HCM. By using an ergodic description, the dynamics of hubs can be predicted from knowledge of the local dynamics and the coupling function. This makes it possible to obtain quantitative theoretical predictions of the behaviour of complex networks. Thereby, we establish the existence of a range of dynamical phenomena quite different from the ones encountered in homogeneous networks. This highlights the need of new paradigms when dealing with high-dimensional dynamical systems with a heterogeneous coupling.

Synchronization occurs through a heat bath mechanism. For certain coupling functions, hubs can synchronize, unlike poorly connected nodes which remain out of synchrony. The underlying synchronization process is not related to direct coupling between hubs, but comes via the coupling with poorly connected nodes. So the hub *synchronization process* is through a mean-field effect (i.e. the coupling is *through a "heat bath"*). In HCM, synchronization depends on the connectivity layer (see Subsection 1.3). We highlighted this feature in the networks with three types of hubs having distinct degrees.

Synchronization in random networks—HCM versus homogeneous. Theorem C shows that synchronization occurs in random homogeneous networks, but is rare in HCM (see Appendix F). Recent work (for example [GS$^+$98]) shows that structure influences dynamics. Theorem B shows that it is not strict symmetry, but (probabilistic) homogeneity that makes synchronization possible. In contrast, in the presence of heterogeneity the dynamics changes according to connectivity layers.

Importance of long transients in high-dimensional systems. Section 3.1 shows how certain behaviour can be sustained by a system only for finite time T and, as it turns out, T is exponentially large in terms of the size of the network, being greater than any feasible observation time. The issue of such long transient times naturally arises in high-dimensional systems. For example, given an N-fold product of the same expanding map f, densities evolve asymptotically to the unique SRB measure exponentially fast, but the rate depends on the dimension and becomes very low for $N \to \infty$. Take an expanding map f and define

$$\boldsymbol{f} := f \times \cdots \times f.$$

Suppose ν is the invariant measure for f absolutely continuous with respect to some reference measure m different from ν. Then the push forward $\boldsymbol{f}_*^t(m^{\otimes N}) = (f_*^t m)^{\otimes N}$ will converge exponentially fast in some suitable product norm to $\nu^{\otimes N}$ because this is true for each factor separately. However, by choosing N large, the rate can be made arbitrarily slow and in the limit of infinite N, $\boldsymbol{f}_*^t(m^{\mathbb{N}})$ and $\nu^{\mathbb{N}}$ are singular for all $t \in \mathbb{N}$. This means that in practice, pushing forward with the dynamics an absolutely continuous initial measure, it might take a very long time before relaxing to the SRB measure even if the system is hyperbolic. This suggests that in order to accurately describe HCM and high-dimensional systems, it is necessary to understand the *dependence* of all relevant quantities and bounds *on the dimension*. This is often disregarded in the classical literature on ergodic theory.

6.1. Open problems and new research directions

With regard to HCM some problems remain open.

1. In Theorem A we assumed that the local map f in our model is Bernoulli and that all the non-linearity within the model is contained in the coupling. This assumption makes it easier to control distortion estimates as the dimension of the network increases. For example, without this assumption, the density of the invariant measure in the expanding case (see Section 4) becomes highly irregular as the dimension increases.

   Problem: *Obtain the results in Theorem A when f is a general uniformly expanding circle map in $C^{1+\nu}$ with $\nu \in (0, 1)$.*

2. In Theorem A we gave a description of orbits for finite time until they hit the set $\mathcal{B}_\varepsilon$ where the fluctuations are above the threshold and the truncated system F_ε differs from the map F.

   Problem: *Describe what happens after an orbit enters the set $\mathcal{B}_\varepsilon$. In particular, find how much time orbits need to escape $\mathcal{B}_\varepsilon$ and how long it takes for them to return to this set.*

3. In the proof of Theorems A and C we assume that the reduced dynamics g_j, in (4), of each hub node $j \in \{1, \ldots, M\}$ is uniformly hyperbolic.

   Problem: *Find a sufficiently weak argument that allows one to describe the case where some of the reduced maps g_j have non-uniformly-hyperbolic behaviour, for example, have a neutral fixed point.*

   In fact, hyperbolicity is a generic condition in dimension one (see [RvS15, KSvS07]) but not in higher dimensions. An answer to this question would be desirable even in the one-dimensional case, but especially when treating multi-dimensional HCM.

The study of HCM and the approach used in this paper also raise more general questions such as:

1. **Problem:** *Is the SRB measure supported on the attractor of F_ε absolutely continuous with respect to the Lebesgue measure m_N on the whole space?*

   Tsujii [Tsu01] proved absolute continuity of the SRB measure for a non-invertible two-dimensional skew product system. Here the main challenges are that the system does not have a skew-product structure, and the perturbation with respect to the product system depends on the dimension.
2. Chimera states refer to "heterogeneous" behaviour observed (in simulations and experiments) on homogeneous networks (see [AS04]). The emergence of such states is not yet completely understood, but they are widely believed to be associated to long transients.

   Problem: *Does the approach of the truncated system shed light on Chimera states?*

Appendix A. Estimates on the truncated system

Hoeffding inequality. *Suppose that $(X_i)_{i\in\mathbb{N}}$ is a sequence of bounded independent random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ and there exist $a_i < b_i$ such that $X_i(\omega) \in [a_i, b_i]$ for all $\omega \in \Omega$. Then*

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}_{\mathbb{P}}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]\right| \geq t\right) \leq 2\exp\left[-\frac{2n^2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right]$$

for all $t > 0$ and $n \in \mathbb{N}$.

Proof of Proposition 3.1. Hoeffding's inequality can be directly applied to the random variables defined on $(\mathbb{T}^L, \mathcal{B}, m_L)$ by $X_i := \theta_{s_1} \circ \pi^i(x)$ where $\pi^i : \mathbb{T}^L \to \mathbb{T}$ is the projection on the i-th coordinate $(1 \leq i \leq L)$. These are in fact independent by construction and bounded since $\{\theta_{s_1}\}_{s_1\in\mathbb{Z}}$ are trigonometric functions on $[-1, 1]$.

Consider the set

$$\mathcal{B}_\varepsilon = \bigcup_{j=1}^{M} \bigcup_{s_1\in\mathbb{Z}} \mathcal{B}_\varepsilon^{(s_1,j)}$$

with $B_\varepsilon^{(s_1,j)}$ defined as in (23). Notice that for $s_1 = 0$, $\mathcal{B}_\varepsilon^{(0,j)} = \mathbb{T}^L$. Since $\kappa_j \Delta = d_j$, we can rewrite

$$\mathcal{B}_\varepsilon^{(s_1,j)} = \left\{ x \in \mathbb{T}^L : \left| \frac{1}{d_j} \sum_{i=1}^{L} A_{ji} \theta_{s_1}(x_i) - \overline{\theta}_{s_1} \right| > \frac{\varepsilon}{\kappa_j} |s_1| \right\}.$$

Since d_j is the number of non-vanishing terms in the sum, the above is the measurable set where the empirical average over d_j i.i.d. bounded random variables exceeds their common expectation by more than $\varepsilon |s_1| / \kappa_j$. Being under the hypotheses of the above Hoeffding inequality, we can estimate the measure of this set as

$$m_L(\mathcal{B}_\varepsilon^{(s_1,j)}) \leq 2 \exp\left[-\frac{d_j^2 \varepsilon^2 |s_1|^2}{d_j 2\kappa_j^2} \right] = 2 \exp\left[-\frac{\Delta \varepsilon^2 |s_1|^2}{2\kappa_j} \right], \tag{62}$$

and this gives

$$\begin{aligned} m_L(\mathcal{B}_\varepsilon) &\leq \sum_{j=1}^{M} \sum_{s_1 \in \mathbb{Z}\setminus\{0\}} m_L(\mathcal{B}_\varepsilon^{(s_1,j)}) \leq 2M \sum_{s_1 \in \mathbb{Z}\setminus\{0\}} \exp\left[-\frac{\Delta \varepsilon^2}{2} |s_1| \right] \\ &\leq 4M \frac{\exp[-\Delta \varepsilon^2/2]}{1 - \exp[-\Delta \varepsilon^2/2]} \end{aligned}$$

since $\kappa_j < 1$, which concludes the proof of the proposition. □

Now we give an expression for DF_ε. Using (20) and (21) and writing as before $z = (x, y)$, noting that $z_k = y_{k-L}$ for $k > L$, we get

$$[D_{(x,y)} F_\varepsilon]_{k\ell} = \begin{cases} D_{x_k} f + \frac{\alpha}{\Delta} \sum_{n=1}^{N} A_{kn} h_1(x_k, z_n), & k = \ell \leq L, \\ \frac{\alpha}{\Delta} A_{k\ell} h_2(x_k, z_\ell), & k \neq \ell,\ k \leq L, \\ \partial_{x_\ell} \xi_{k-L,\varepsilon}, & k > L,\ \ell \leq L, \\ \frac{\alpha}{\Delta} A_{k\ell} h_2(y_{k-L}, y_{\ell-L}), & k \neq \ell > L, \\ D_{y_{k-L}} g_{k-L} + \partial_{y_{k-L}} \xi_{k-L,\varepsilon}, & k = \ell > L. \end{cases} \tag{63}$$

Here h_1 and h_2 stand for the partial derivatives of the function h with respect to the first and second coordinate respectively, and where we suppressed some of the functional dependences.

The following lemma summarises the properties of $\xi_{j,\varepsilon}$ that will yield good hyperbolic properties for F_ε.

Lemma A.1. *The functions $\xi_{j,\varepsilon} : \mathbb{T}^N \to \mathbb{R}$ defined in (22) satisfy*

(i) $|\xi_{j,\varepsilon}| \leq C_\#(\varepsilon + \Delta^{-1} M)$ *where $C_\#$ is a constant depending only on σ, h, and α;*
(ii)

$$|\partial_{z_n} \xi_{j,\varepsilon}| \leq \begin{cases} \mathcal{O}(\Delta^{-1}) A_{jn}, & n \leq L, \\ C_\# \varepsilon + \mathcal{O}(\Delta^{-1} M), & n = j + L, \\ \mathcal{O}(\Delta^{-1}) A_{jn}, & n > L,\ n \neq j + L; \end{cases}$$

(iii) *for all* $z, \overline{z} \in \mathbb{T}^N$,

$$|\partial_{z_n} \xi_{j,\varepsilon}(z) - \partial_{z_n} \xi_{j,\varepsilon}(\overline{z})| \leq \begin{cases} \mathcal{O}(\Delta^{-1}) A_{jn} d_\infty(z, \overline{z}), & n \leq L, \\ \mathcal{O}(1) + \mathcal{O}(\Delta^{-1} M) d_\infty(z, \overline{z}), & n = j + L, \\ \mathcal{O}(\Delta^{-1}) A_{jn} d_\infty(z, \overline{z}), & n > L,\ n \neq j + L. \end{cases}$$

Proof. (i) follows from the estimates

$$|\xi_{j,\varepsilon}| \leq C_\# \Big(\sum_{s \in \mathbb{Z}^2} c_s \varepsilon |s_1| + \Delta^{-1} M \Big) \leq C_\# (\varepsilon + \Delta^{-1} M),$$

where the sum is absolutely convergent. To prove (ii) notice that for $n \leq L$,

$$\partial_{z_n} \xi_{j,\varepsilon}(z) = \partial_{x_n} \xi_{j,\varepsilon}(z) = \alpha \sum_{s \in \mathbb{Z}^2} c_s D_{(\cdot)} \zeta_{\varepsilon |s_1|} \frac{A_{jn}}{\Delta} D_{x_n} \theta_{s_1} \tag{64}$$

and $|D_{x_n} \theta_{s_1}| \leq 2\pi |s_1|$, so the bound follows from the fast decay rate of the Fourier coefficients. For $n = j + L$,

$$\begin{aligned} |\partial_{z_{j+L}} \xi_{j,\varepsilon}(z)| = |\partial_{y_j} \xi_{j,\varepsilon}(z)| &= \left| \alpha \sum_{s \in \mathbb{Z}^2} c_s \zeta_{\varepsilon |s_1|} D_{y_j} \theta_{s_2} + \sum_{n=1}^{M} \frac{\alpha}{\Delta} A^h_{jn} \partial_{y_j} h(y_j, y_n) \right| \\ &\leq \varepsilon C_\# \sum_{s \in \mathbb{Z}^2} |c_s| \, |D_{y_j} \theta_{s_2}| \, |s_1| + \mathcal{O}(\Delta^{-1} M). \end{aligned}$$

Again the decay of the Fourier coefficients yields the desired bound. For n greater than L and different from $j + L$ it is trivial. Point (iii) for $n \neq L + j$ follows immediately from expression (64) and the decay of the Fourier coefficients. For $n = j + L$,

$$\begin{aligned} &|\partial_{y_j} \xi_{j,\varepsilon}(z) - \partial_{y_j} \xi_{j,\varepsilon}(\overline{z})| \\ &\leq \left| \alpha \sum_{s \in \mathbb{Z}^2} c_s \varepsilon |s_1| [D_{y_j} \theta_{s_2} - D_{\overline{y}_j} \theta_{s_2}] + \sum_{n=1}^{M} \frac{\alpha}{\Delta} A^h_{jn} [\partial_{y_j} h(y_j, y_n) - \partial_{y_j} h(\overline{y}_j, \overline{y}_n)] \right| \\ &\leq \mathcal{O}(1 + \Delta^{-1} M) d_\infty(z, \overline{z}). \end{aligned}$$

Notice that to obtain the last step we need the sequence $\{c_s |s_1| \, |s_2|^3\}$ to be summable. In particular,

$$c_s \leq \frac{c_\#}{|s_1|^{2+b} |s_2|^{4+b}}, \qquad b > 0,$$

is a sufficient condition, ensured by picking $h \in C^{10}$. □

Appendix B. Estimate on ratios of determinants

In the following proposition $\mathrm{Col}^k[M]$, with $M \in \mathcal{M}_{n \times n}$ a square matrix of dimension n, is the k-th column of M.

Proposition B.1. *Suppose that* $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}^+$ *is the* p*-norm* $(1 \le p \le \infty)$ *on the Euclidean space* $\mathbb{R}^n$*. Take two* $n \times n$ *matrices* b_1 *and* b_2*. Suppose there is a constant* $\lambda \in (0, 1)$ *such that*

$$\|b_i\|_p := \sup_{\substack{v \in \mathbb{R}^n \\ \|v\|_p \le 1}} \frac{\|b_i v\|_p}{\|v\|_p} \le \lambda, \qquad \forall i \in \{1, 2\}. \tag{65}$$

Then

$$\frac{|\mathrm{Id} + b_1|}{|\mathrm{Id} + b_2|} \le \exp\left\{\frac{\sum_{k=1}^{n} \|\mathrm{Col}^k[b_1 - b_2]\|_p}{1 + \lambda}\right\}.$$

Proof. Given a matrix $M \in \mathcal{M}_{n\times n}$ it is a standard formula that

$$|M| = \exp[\mathrm{Tr}\log(M)], \qquad \frac{|\mathrm{Id} + b_1|}{|\mathrm{Id} + b_2|} = \exp\left\{\sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1}}{\ell} \mathrm{Tr}[b_1^\ell - b_2^\ell]\right\}.$$

Substituting the expression

$$b_1^\ell - b_2^\ell = \sum_{j=0}^{\ell-1} b_1^j (b_1 - b_2) b_2^{\ell-j-1}$$

we obtain

$$\begin{aligned}\mathrm{Tr}(b_1^\ell - b_2^\ell) &= \sum_{i=0}^{\ell-1} \mathrm{Tr}(b_1^j (b_1 - b_2) b_2^{\ell-j-1}) = \sum_{j=0}^{\ell-1} \mathrm{Tr}(b_2^{\ell-j-1} b_1^n (b_1 - b_2)) \\ &\le \sum_{j=0}^{\ell-1} \sum_{k=1}^{n} \|\mathrm{Col}^k[b_2^{\ell-j-1} b_1^j (b_1 - b_2)]\| \end{aligned} \tag{66}$$

where we use the fact that the trace of a matrix is upper bounded by the sum of the p-norms of its columns (for any $p \in [1, \infty]$). Using conditions (65) we obtain

$$\mathrm{Tr}(b_1^\ell - b_2^\ell) \le \ell \lambda^{\ell-1} \sum_{k=1}^{n} \|\mathrm{Col}^k[b_1 - b_2]\|_p. \tag{67}$$

To conclude,

$$\begin{aligned}\frac{|\mathrm{Id} + b_1|}{|\mathrm{Id} + b_2|} &\le \exp\left\{\sum_{\ell=1}^{\infty} (-1)^{\ell+1} \lambda^{\ell-1} \sum_{k=1}^{n} \|\mathrm{Col}^k[b_1 - b_2]\|_p\right\} \\ &= \exp\left\{\frac{\sum_{k=1}^{n} \|\mathrm{Col}^k[b_1 - b_2]\|_p}{1 + \lambda}\right\}. \end{aligned}$$

□

Appendix C. Transfer operator

Suppose that $(M, \mathcal{B})$ is a measurable space. Given a measurable map $F : M \to M$ define the *push forward* $F_*\mu$ of any (signed) measure μ on $(M, \mathcal{B})$ by

$$F_*\mu(A) := \mu(F^{-1}(A)), \qquad \forall A \in \mathcal{B}.$$

The operator F_* defines how mass distribution evolves on M after application of the map F. Now suppose that a reference measure m on $(M, \mathcal{B})$ is given. The map F is *non-singular* if F_*m is absolutely continuous with respect to m; we write $F_*m \ll m$. If F is non-singular and $\mu \ll m$ then also $F_*\mu \ll m$. This means that one can define an operator

$$P : L^1(M, m) \to L^1(M, m)$$

such that if $\rho \in L^1$ then $P\rho := dF_*(\rho \cdot m)/dm$ where $\rho \cdot m$ is the measure satisfying $d(\rho \cdot m)/dm = \rho$. In particular, if $\rho \in L^1$ is a mass density ($\rho \geq 0$ and $\int_M \rho \, dm = 1$) then P maps ρ to the mass density obtained after application of F. One can prove that an equivalent characterization of P is as the only operator that satisfies

$$\int_M \varphi\psi \circ F \, dm = \int_M P\varphi\psi \, dm, \qquad \forall \psi \in L^\infty(M, m) \text{ and } \varphi \in L^1.$$

This means that if, for example, M is a Riemannian manifold and m is its volume form, and if F is a local diffeomorphism, then P can be obtained from the change of variables formula as

$$P\varphi(y) = \sum_{\{x : F(x) = y\}} \frac{\varphi(x)}{\operatorname{Jac} F(x)}$$

where $\operatorname{Jac} F(x) = \frac{dF_*m}{dm}(x)$. It follows from the definition of P that $\rho \in L^1$ is an *invariant density* for F if and only if $P\rho = \rho$.

Appendix D. Graph transform: some explicit estimates

Once again we go through the argument of the graph transform in the case of a cone-hyperbolic endomorphism of the n-dimensional torus. The scope of this result is to explicitly compute bounds on Lipschitz constants for the invariant set of admissible manifolds, and the contraction rate of the graph transform [Shu13, KH95].

Consider the torus $\mathbb{T}^n$ with the trivial tangent bundle $\mathbb{T}^n \times \mathbb{R}^n$. Suppose that $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is a constant norm on the tangent spaces, and that, by abuse of notation, $\|x_1 - x_2\|$ is the distance between $x_1, x_2 \in \mathbb{T}^n$ induced by the norm. Take $n_u, n_s \in \mathbb{N}$ such that $n = n_s + n_u$, and let $\Pi_s : \mathbb{R}^n \to \mathbb{R}^{n_s}$ and $\Pi_u : \mathbb{R}^n \to \mathbb{R}^{n_u}$ be the projections for the decomposition $\mathbb{R}^n = \mathbb{R}^{n_u} \oplus \mathbb{R}^{n_s}$. Identifying $\mathbb{T}^n$ with $\mathbb{T}^{n_u} \times \mathbb{T}^{n_s}$, we let $\pi_s : \mathbb{T}^n \to \mathbb{T}^{n_s}$ and $\pi_u : \mathbb{T}^n \to \mathbb{T}^{n_u}$ be the projections on the respective coordinates. Take a C^2 local diffeomorphism $F : \mathbb{T}^n \to \mathbb{T}^n$ and define $F_u := \pi_u \circ F$ and $F_s := \pi_s \circ F$. Suppose that it satisfies the following assumptions. There are constants $\beta_u, \beta_s > 1$, $K_u > 0$ and constant cone fields

$$\mathcal{C}^u := \{v \in \mathbb{R}^n : \|\Pi_u v\| \geq \beta_u \|\Pi_s v\|\} \quad \text{and} \quad \mathcal{C}^s := \{v \in \mathbb{R}^n : \|\Pi_s v\| \geq \beta_s \|\Pi_u v\|\}$$

such that

- for all $x \in \mathbb{T}^n$, $D_x F(\mathcal{C}^u) \subset \mathcal{C}^u(F(x))$ and $D_{F(x)} F^{-1}(\mathcal{C}^s(F(x))) \subset \mathcal{C}^s(x)$;
- there are $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}^+$ such that

$$0 < \lambda_2 \leq \|D_x F|_{\mathcal{C}^s}\| \leq \lambda_1 < 1 < \mu_1 \leq \|D_x F|_{\mathcal{C}^u}\| \leq \mu_2;$$

- $\|D_{z_1}F - D_{z_2}F\|_u := \sup_{v\in\mathcal{C}^u} \dfrac{\|(D_{z_1}F - D_{z_2}F)v\|}{\|v\|} \leq K_u\|z_1 - z_2\|.$

From now on we denote by $(x, y) \in \mathbb{T}^n$ a point in the torus with $x \in \mathbb{T}^{n_u}$ and $y \in \mathbb{T}^{n_s}$. Take $r > 0$ and let $B_r^u(x)$ and $B_r^s(y)$ be open balls of radius r in $\mathbb{T}^{n_u}$ and $\mathbb{T}^{n_s}$ respectively. Consider

$$C_u^1(B_r^u(x), B_r^s(y)) := \{\sigma : B_r^u(x) \to B_r^s(y) : \|D\sigma\| < \beta_u^{-1}\}.$$

The condition above ensures that the graph of any σ is tangent to the unstable cone. It is easy to prove invertibility of $\pi_u \circ F \circ (\mathrm{id}, \sigma)|_{B_r^u(x)}$ for sufficiently small r, and thus the graph transform

$$\Gamma : C_u^1(B_r^u(x), B_r^s(y)) \to C_u^1(B_r^u(F_u(x, y)), B_r^s(F_s(x, y)))$$

is well defined, where the only requirement is that for any σ as above the graph of $\Gamma\sigma$, $(\mathrm{id}, \Gamma\sigma)(B_r^u(F_u(x, y)))$, is contained in $F \circ (\mathrm{id}, \sigma)(B_r^u(x))$. An expression for Γ is given by

$$\Gamma\sigma := [\pi_s \circ F \circ (\mathrm{id}, \sigma)] \circ [\pi_u \circ F \circ (\mathrm{id}, \sigma)]^{-1}|_{B_r^u(F_u(x,y))}.$$

The fact that $\|D(\Gamma\sigma)\| \leq \beta_u^{-1}$ is a consequence of the invariance of $\mathcal{C}^u$. Now we prove a result that determines a regularity property for the admissible manifold which is invariant under the graph transform.

Proposition D.1. *Consider* $\sigma \in C_{u,K}^1(B_r^u(x), B_r^s(y)) \subset C_u^1(B_r^u(x), B_r^s(y))$ *characterised by*

$$\mathrm{Lip}(D\sigma) = \sup_{\substack{x'\neq y' \\ x', y' \in B_r^u(x)}} \frac{\|D_{x'}\sigma - D_{y'}\sigma\|}{\|x' - y'\|} \leq K.$$

Then Γ *maps* $C_{u,K}^1(B_r^u(x), B_r^s(y))$ *into* $C_{u,K}^1(B_r^u(F_u(x, y)), B_r^s(F_s(x, y)))$ *if*

$$K > \frac{1}{1 - \frac{\lambda_1}{\mu_1(1-\beta_u^{-1})}} \left(\frac{\mu_2}{\mu_1\lambda_2} K_u \lambda_1 \frac{1 + \beta_u^{-1}}{1 - \beta_u^{-1}} + K_u \frac{(1 + \beta_u^{-1})^2}{\mu_1(1 - \beta_u^{-1})} \right).$$

Proof. Take $z_1, z_2 \in B_r^u(z)$ with $z = \pi_u \circ F \circ (\mathrm{id}, \sigma)(x)$ and suppose that $x_1, x_2 \in B_r^u(x)$ are such that $\pi_u \circ F \circ (\mathrm{id}, \sigma)(x_i) = z_i$. Take $w \in \mathbb{R}^{n_u}$, and suppose that $v_1, v_2 \in \mathbb{R}^{n_u}$ satisfy $\Pi_u D_{(x_i,\sigma(x_i))}F(v_i, D_{x_i}\sigma(v_i)) = w$. Then

$$\begin{aligned}
&\|D_{z_1}(\Gamma\sigma)(w) - D_{z_2}(\Gamma\sigma)(w)\| \\
&\quad \leq \|D_{(x_1,\sigma(x_1))}F(v_1, D_{x_1}\sigma(v_1)) - D_{(x_2,\sigma(x_2))}F(v_2, D_{x_2}\sigma(v_2))\| \\
&\quad \leq \|D_{(x_1,\sigma(x_1))}F(v_1 - v_2, D_{x_1}\sigma(v_1 - v_2))\| \\
&\qquad + \|D_{(x_1,\sigma(x_1))}F(0, D_{x_1}\sigma - D_{x_2}\sigma)v_2\| \\
&\qquad + \|D_{(x_1,\sigma(x_1))}F - D_{(x_2,\sigma(x_2))}F\|_u\|(v_2, D_{x_2}\sigma v_2)\| \\
&\quad \leq \mu_2\|v_1 - v_2\| + \lambda_1\|D_{x_1}\sigma - D_{x_2}\sigma\|\,\|v_2\| + K_u(1 + \beta_u^{-1})\|x_1 - x_2\|\,\|(v_2, D_{x_2}\sigma v_2)\|.
\end{aligned}$$

Now

$$\|x_1 - x_2\| \le \lambda_1(1 + \beta_u^{-1})\|z_1 - z_2\|$$

and

$$\begin{aligned}\|v_1 - v_2\| &= \|v_1 - \Pi_u(D_{(x_2,\sigma(x_2))}F)^{-1}D_{(x_1,\sigma(x_1))}F(\mathrm{Id}, D_{x_1}\sigma)(v_1)\| \\ &= \|\Pi_u(D_{(x_2,\sigma(x_2))}F)^{-1}(D_{(x_1,\sigma(x_1))}F - D_{(x_2,\sigma(x_2))}F)(\mathrm{Id}, D_{x_1}\sigma)(v_1)\| \\ &\le \lambda_2^{-1}K_u\|x_1 - x_2\|\,\|v_1\| \le \lambda_2^{-1}K_u\lambda_1(1 + \beta_u^{-1})\|z_1 - z_2\|\,\|v_1\|.\end{aligned}$$

Taking into account that $\|v_1\|, \|v_2\| \le \mu_1^{-1}(1 - \beta_u^{-1})^{-1}\|w\|$ we obtain

$$\mathrm{Lip}(D.(\Gamma\sigma)) \le \frac{\lambda_1}{\mu_1(1 - \beta_u^{-1})}\,\mathrm{Lip}(D.\sigma) + \frac{\mu_2}{\mu_1}\lambda_2^{-1}K_u\lambda_1\frac{1 + \beta_u^{-1}}{1 - \beta_u^{-1}} + K_u\frac{(1 + \beta_u^{-1})^2}{\mu_1(1 - \beta_u^{-1})},$$

and this gives the condition of invariance of the proposition. □

Proposition D.2. *For all $\sigma_1, \sigma_2 \in C_u^1(B_r^u(x), B_r^s(y))$,*

$$\begin{aligned}\sup_{z\in B_r^u(F_u(x,y))} &\|(\Gamma\sigma_1)(z) - (\Gamma\sigma_2)(z)\| \\ &\le [\lambda_1 + \lambda_1^2\mu_1^{-1}\beta_u^{-1} + \mu_2\mu_1^{-1}\lambda_1\beta_u^{-1}] \sup_{t\in B_r^u(x)} \|\sigma_1(t) - \sigma_2(t)\|.\end{aligned}$$

Hence if

$$\lambda_1 + \lambda_1^2\mu_1^{-1}\beta_u^{-1} + \mu_2\mu_1^{-1}\lambda_1\beta_u^{-1} < 1$$

then $\Gamma : C_u^1(B_r^u(x), B_r^s(y)) \to C_u^1(B_r^u(F_u(x, y)), B_r^s(F_s(x, y)))$ is a contraction in the C^0 topology.

Proof. Take $\sigma_1, \sigma_2 \in C_u^1(B_r^u(x), B_r^s(y))$ and $z \in B_r^u(F_u(x, y))$, and suppose that x_1, x_2 in $B_r^u(x)$ are such that $F_u(x_1, \sigma_1(x_1)) = z$ and $F_u(x_2, \sigma_2(x_2)) = z$. Then

$$\begin{aligned}\|(\Gamma\sigma_1)(z) - (\Gamma\sigma_2)(z)\| &= \|F_s(x_1, \sigma_1(x_1)) - F_s(x_2, \sigma_2(x_2))\| \\ &\le \|F_s(x_1, \sigma_1(x_1)) - F_s(x_1, \sigma_2(x_1))\| + \|F_s(x_1, \sigma_2(x_1)) - F_s(x_2, \sigma_2(x_2))\| \\ &\le \lambda_1\|\sigma_1(x_1) - \sigma_2(x_1)\| + \lambda_1\,\mathrm{Lip}(\sigma_1)\|x_1 - x_2\| + \mathrm{Lip}(F)\beta_u^{-1}\|x_1 - x_2\|.\end{aligned}$$

The following estimates hold:

$$\begin{aligned}\|x_1 - x_2\| &= \|x_1 - (F_u \circ (\mathrm{id}, \sigma_2))^{-1} \circ (F_u \circ (\mathrm{id}, \sigma_1))(x_1)\| \\ &= \|x_1 - (F_u \circ (\mathrm{id}, \sigma_2))^{-1}[F_u \circ (\mathrm{id}, \sigma_2)(x_1) + F_u \circ (\mathrm{id}, \sigma_1)(x_1) - F_u \circ (\mathrm{id}, \sigma_2)(x_1)]\| \\ &\le \|x_1 - x_1\| + \|D_{\overline{x}}(F_u \circ (\mathrm{id}, \sigma_2))^{-1}\|\,\|F_u \circ (\mathrm{id}, \sigma_1)(x_1) - F_u \circ (\mathrm{id}, \sigma_2)(x_1)\| \\ &\le \|D_{\overline{x}}F_u\|^{-1}\lambda_1\|\sigma_1(x_1) - \sigma_2(x_1)\| \le \mu_1^{-1}\lambda_1\|\sigma_1(x_1) - \sigma_2(x_1)\| \qquad (68)\end{aligned}$$

and hence

$$\|(\Gamma\sigma_1)(z) - (\Gamma\sigma_2)(z)\| \le [\lambda_1 + \lambda_1^2\mu_1^{-1}\beta_u^{-1} + \mu_2\mu_1^{-1}\lambda_1\beta_u^{-1}]\|\sigma_1(x_1) - \sigma_2(x_1)\|.$$ □

Consider any linear subspace $V \subset \mathcal{C}^u$ of dimension n_u. This is uniquely associated to $L : \mathbb{R}^{n_u} \to \mathbb{R}^{n_s}$ such that $(\mathrm{Id}, L)(\mathbb{R}^{n_u}) = V$.

Definition D.1. Given any linear spaces $V_1, V_2 \subset \mathcal{C}^u$ of dimension n_u, we can define the distance

$$d_{\mathcal{C}^u}(V_1, V_2) := \sup_{\substack{u\in\mathbb{R}^{n_u}\\ \|u\|=1}} \|L_1(u) - L_2(u)\| \tag{69}$$

(which is also the operator norm of the difference of the two linear morphisms defining the subspaces).

Proposition D.3. *If*

$$\mu_1^{-1}\left[\lambda_1 + \frac{\beta_u\lambda_1}{\mu_1(1-\beta_u)}\right] < 1$$

then D_zF is a contraction with respect to $d_{\mathcal{C}^u}$ for all $z \in \mathbb{T}^n$.

Proof. Pick $L_1, L_2 : \mathbb{R}^{n_u} \to \mathbb{R}^{n_s}$ with $\|L_i\| < \beta_u$. They define linear subspaces $V_i = (\mathrm{Id}, L_i)(\mathbb{R}^{n_u})$ which, as a consequence of the condition on the norm of L_i, are tangent to the unstable cone. The subspaces V_1 and V_2 are transformed by D_zF into subspaces V_1' and V_2', which are the graphs of linear transformations $L_1', L_2' : \mathbb{R}^{n_u} \to \mathbb{R}^{n_s}$ ($\|L_i'\| \le \beta_u$). Analogously to the graph transform one can find an explicit expression for L_i' in terms of L_i:

$$L_i' = \Pi_s \circ D_zF \circ (\mathrm{Id}, L_i) \circ [\Pi_u \circ D_zF \circ (\mathrm{Id}, L_i)]^{-1}.$$

To prove the proposition we then proceed analogously to the proof of Proposition D.2. Pick $u \in \mathbb{R}^{n_u}$ and suppose that $u_1, u_2 \in \mathbb{R}^{n_u}$ are such that

$$(\mathrm{Id}, L_1')(u) = D_zF \circ (\mathrm{Id}, L_1)(u_1), \qquad (\mathrm{Id}, L_2')(u) = D_zF \circ (\mathrm{Id}, L_2)(u_2).$$

With the above definitions,

$$\begin{aligned}
\|L_1'(u) - L_2'(u)\| &= \|\Pi_s \circ D_zF(\mathrm{Id}, L_1)(u_1) - \Pi_s \circ D_zF(\mathrm{Id}, L_1)(u_2)\| \\
&\le \|\Pi_s \circ D_zF(\mathrm{Id}, L_1)(u_1) - \Pi_s \circ D_zF(\mathrm{Id}, L_2)(u_1)\| \\
&\quad + \|\Pi_s \circ D_zF(\mathrm{Id}, L_2)(u_1 - u_2)\| \\
&\le \lambda_1\|L_1 - L_2\|\,\|u_1\| + \beta_u\mu_2(1+\beta_u)\|u_1 - u_2\|, \\
\|u_1 - u_2\| &= \|u_1 - [\Pi_u \circ D_zF \circ (\mathrm{Id}, L_2)]^{-1}\Pi_u \circ D_zF \circ (\mathrm{Id}, L_1)(u_1)\| \\
&= \|[\Pi_u \circ D_zF \circ (\mathrm{Id}, L_2)]^{-1}\Pi_u \circ D_zF \circ (0, L_1 - L_2)(u_1)\| \\
&\le \|\Pi_u \circ D_zF \circ (\mathrm{Id}, L_2)^{-1}\|\beta_u\lambda_1\|L_1 - L_2\|\,\|u_1\| \\
&\le \frac{\beta_u\lambda_1}{\mu_1(1-\beta_u)}\|L_1 - L_2\|\,\|u_1\|.
\end{aligned}$$

The two estimates together imply that

$$\|L_1' - L_2'\| \le \mu_1^{-1}\left[\lambda_1 + \frac{\beta_u\lambda_1}{\mu_1(1-\beta_u)}\right]\|L_1 - L_2\|. \qquad \square$$

Appendix E. Proof of Theorem B

Let $g\colon \mathbb{T} \to \mathbb{T}$ and for $\omega \in \mathbb{R}$ define $g_\omega = g + \omega$. Let $\underline{\omega} = (\dots, \omega_n, \omega_{n-1}, \dots, \omega_0)$ with $\omega_i \in (-\varepsilon', \varepsilon')$ with $\varepsilon' > 0$ small. Define $g^k_{\underline{\omega}} = g_{\omega_k} \circ \cdots \circ g_{\omega_1} \circ g_{\omega_0}$.

Proposition E.1. *Let $g\colon \mathbb{T} \to \mathbb{T}$ be C^2 and hyperbolic (in the sense of Definition 2.1), and assume that g has an attracting set Λ (consisting of periodic orbits). Then there exist $\chi \in (0,1)$ and $C > 0$ such that for each $\varepsilon > 0$ and $T = 1/\varepsilon$ the following holds.*

There exists a set $\Omega \subset \mathbb{T}$ of measure $1 - \varepsilon^{1-\chi}$ such that for any $k \geq T_0$, and any $\underline{\omega} = (\dots, \omega_n, \omega_{n-1}, \dots, \omega_0)$ with $|\omega_i| \leq C\varepsilon$, and for each $k \geq T_0$,

- *$g^k_{\underline{\omega}}$ maps each component J of Ω into components of the immediate basin of the periodic attractor of g;*
- *the distance from $g^k_{\underline{\omega}}(J)$ to a periodic attractor of g is at most ε.*

This proposition follows from the next two lemmas.

Lemma E.1. *Let $g\colon \mathbb{T} \to \mathbb{T}$ be C^2 and hyperbolic (in the sense of Definition 2.1), and assume that g has an attracting set Λ (consisting of periodic orbits). Then the repelling hyperbolic set $\Upsilon = \mathbb{T} \setminus W^s(\Lambda)$ of g is a Cantor set with Hausdorff dimension $\chi' < 1$. Moreover, for each $\chi \in (\chi', 1)$, the Lebesgue measure of the ε-neighborhood $N_\varepsilon(\Upsilon)$ of Υ is at most $\varepsilon^{1-\chi}$ provided $\varepsilon > 0$ is sufficiently small.*

Proof. It is well known that the set Υ is a Cantor set [dMvS93]. Notice that by definition $g^{-1}(\Upsilon) = \Upsilon$. It is also well known that the Hausdorff dimension of a hyperbolic set Υ associated to a C^2 one-dimensional map is < 1 and that this dimension is equal to its box dimension [Pes97]. Now take a covering of Υ with intervals of length ε, and let $N(\varepsilon)$ be the smallest number of such intervals that are needed. By the definition of box dimension, $\lim_{\varepsilon\to 0} \log N(\varepsilon)/\log \varepsilon \to \chi'$. It follows that $N(\varepsilon) \leq 1/\varepsilon^{\chi}$ for $\varepsilon > 0$ small. Consequently, the Lebesgue measure of N_ε is at most $N(\varepsilon)\varepsilon \leq \varepsilon^{1-\chi}$ for $\varepsilon > 0$ small. □

For simplicity assume that $n = 1$ in Definition 2.1. As in Subsection 5.9, the general proof can be reduced to this case.

Lemma E.2. *Let g and $g^k_{\underline{\omega}}$ be as above. Then there exists $C > 0$ such that for each $\varepsilon > 0$ sufficiently small, and taking $\widetilde{N} = N_\varepsilon(\Upsilon)$ and $|\omega_i| < \varepsilon' = C\varepsilon$, we have the following:*

1. *$g^k_{\underline{\omega}}(\mathbb{T} \setminus \widetilde{N}) \subset \mathbb{T} \setminus \widetilde{N}$ for all $k \geq 1$.*
2. *$\mathbb{T} \setminus \widetilde{N}$ consists of at most $1/\varepsilon$ intervals.*
3. *Take $T_0 = 2/\varepsilon$. Then for each $k \geq T_0$, $g^k_{\underline{\omega}}$ maps each component J of $\mathbb{T} \setminus \widetilde{N}$ into a component of the immediate basin of a periodic attractor of g. Moreover, $g^k_{\underline{\omega}}(J)$ has length $< \varepsilon$ and has distance $< \varepsilon$ to a periodic attractor of g.*

Proof. The first statement follows from the fact that we assume that $|Dg| > 1$ on Υ, because Υ is backward invariant, and by continuity. To prove the second statement let J_i be the components of $\mathbb{T} \setminus N_{\varepsilon/4}(\Upsilon)$. If J_i has length $< \varepsilon$ then J_i is contained in $\mathbb{T} \setminus N_\varepsilon(\Upsilon)$. So the remaining intervals J_i all have length $\geq \varepsilon$ and cover $\mathbb{T} \setminus N_\varepsilon(\Upsilon)$. The second statement follows. To see the third statement, notice that the only components of $\mathbb{T} \setminus \Upsilon$ containing periodic points are those that contain periodic attractors. Since Υ is fully invariant, $\mathbb{T} \setminus \Upsilon$ is forward invariant. In particular, if J' is component of $\mathbb{T} \setminus \Upsilon$ then there exists k such that $g^k(J')$ is contained in the immediate basin of a periodic attractor of g and $J', \ldots, g^k(J')$ are all contained in different components of $\mathbb{T} \setminus \Upsilon$. This, together with (1) and (2), implies that each component of $\mathbb{T} \setminus \widetilde{N}$ is mapped in at most $1/\varepsilon$ steps into the immediate basin of g. Since the periodic attractor is hyperbolic, it follows that under $1/\varepsilon$ further iterates this interval has length $< \varepsilon$ and has distance at most ε to a periodic attractor (here we use the fact that $\varepsilon > 0$ is sufficiently small so that also $2/\varepsilon > m$). □

Proof of Theorem B. (a) Fix an integer $\sigma \geq 2$, $\alpha \in \mathbb{R}$, and $\kappa \in (0, 1]$. The map $\mathcal{F}\colon C^k(\mathbb{T} \times \mathbb{T}, \mathbb{R}) \to C^k(\mathbb{T}, \mathbb{R})$ defined by $\mathcal{F}(h)(x) = \int h(x, y)\, dy$ is continuous. Since the set of hyperbolic C^k maps $g\colon \mathbb{T} \to \mathbb{T}$ is open and dense in the C^k topology [KSvS07], it follows that the set of functions $h \in C^k(\mathbb{T} \times \mathbb{T}, \mathbb{R})$ for which $x \mapsto \sigma x + \alpha\kappa \int h(x, y)\, dm_1(y) \mod 1$ is hyperbolic is also open and dense in the C^k topology, which proves the first statement of the theorem. (The above is true for $k \in \mathbb{N}$, $k = \infty$, or $k = \omega$.)

To prove (b), first of all recall that if $g \in C^k(\mathbb{T}, \mathbb{T})$ is a hyperbolic map with a critical point $x \in \mathbb{T}$, then g has a periodic attractor and x belongs to its basin. If $h \in C^k(\mathbb{T} \times \mathbb{T}, \mathbb{R})$, supposing that $\mathcal{F}(h)(x)$ is not constant, we have

$$\exists x \in \mathbb{T}, \qquad \frac{d\mathcal{F}(h)(x)}{dx} < 0. \tag{70}$$

Condition (70) holds for an open and dense set $\Gamma'' \subset C^k(\mathbb{T} \times \mathbb{T}, \mathbb{R})$. Pick $h \in \Gamma''$; then from (70) it follows that there exist an open neighbourhood V of h and an interval $\mathcal{I} \subset \mathbb{R}$ such that $g_{\beta,h}(x) = \sigma x + \beta\mathcal{F}(h)(x) \mod 1$ has a critical point for all $h \in V$ and $\beta \in \mathcal{I}$. Since the map $\mathcal{I} \times V \to C^k(\mathbb{T}, \mathbb{T})$ is continuous, there is an open and dense subset of $\mathcal{I} \times V$ for which the map $g_{\beta,h}$ is hyperbolic, and thus has a finite periodic attractor. Furthermore, if $g_{\beta,h}$ has a periodic attractor, by structural stability there is an open interval $\mathcal{I}_\beta$ such that also $g_{\beta',h}$ has a periodic attractor for all $\beta' \in \mathcal{I}_\beta$. Once the existence of a hyperbolic periodic attractor is established, the other statements follow from Theorem A and Proposition E.1. □

The following two propositions contain rigorous statement regarding the example presented in the introduction.

Proposition E.2. *For any $\beta \in \mathbb{R}$, the map $T_\beta(x) = 2x - \beta \sin(2\pi x) \mod 1$ has at most two periodic attractors O_1, O_2 with $O_1 = -O_2$.*

Proof. The map T_β extends to an entire map on $\mathbb{C}$ and therefore each periodic attractor has a critical point in its basin [Ber93]. This implies that there are at most two periodic attracting orbits. Note that $T_\beta(-x) = -T_\beta(x)$ and therefore if O is a finite set in $\mathbb{R}$ corresponding to a periodic orbit of T_β, then so is $-O$, and it follows that if T_β has two periodic attractors O_1 and O_2 then $O_1 = -O_2$. (If T_β has only one periodic attractor O, then $O = -O$.) Notice that indeed there exist parameters β for which T_β has two attracting orbits. For example, when $\beta = 1.25$ then T_β has two distinct attracting fixed points.

□

Proposition E.3. *Given $\kappa_0, \kappa_1, \dots, \kappa_m$, set $T_{\beta,j}(x) = 2x - \beta\kappa_j \sin(2\pi x) \bmod 1$. Then*

(1) *there exists an open and dense subset $\mathcal{I}'$ of $\mathbb{R}$ such that for each $\beta \in \mathcal{I}'$ each of the maps $T_{\beta,j}$, $j = 1, \dots, m$, is hyperbolic;*
(2) *there exists $\beta_0 > 0$ and an open and dense subset $\mathcal{I}$ of $(-\infty, -\beta_0) \cup (\beta_0, \infty)$ such that for each $\beta \in \mathcal{I}$, each of the maps $T_{\beta,j}$, $j = 1, \dots, m$, is hyperbolic and has a periodic attractor.*

Proof. Let $\mathcal{H}$ be the set of parameters $\beta \in \mathbb{R}$ such that $T_\beta(x) = 2x - \beta \sin(2\pi x) \bmod 1$ is hyperbolic. By [RvS15], the set $\mathcal{H}$ is open and dense. It follows that $(1/\kappa_j)\mathcal{H}$ is also open and dense. Hence $(1/\kappa_1)\mathcal{H} \cap \cdots \cap (1/\kappa_m)\mathcal{H}$ is open and dense.

For each $|\beta| > 2\pi$ the map T_β has a critical point, and so if such a T_β is hyperbolic then, by definition, it has one or more periodic attractors (and each critical point is in the basin of a periodic attractor). So if we take $\beta_0 = \max(2\pi/\kappa_1, \dots, 2\pi/\kappa_m)$ the second assertion follows. □

Appendix F. Proof of Theorem C

The study of global synchronization of chaotic systems has started in the eighties for systems in the ring [FY83, HCP94]. This approach was generalized to undirected networks of diffusively coupled systems merging numerical computations of Lyapunov exponents and transverse instabilities of the synchronous states. See also [DB14, ELP17] for a review. These results have been generalized to weighted and directed graphs via dichotomy estimates [PE+14]. In our Theorem C, we make use of these ideas to obtain an open set of coupling functions such that the networks will globally synchronize for random homogeneous networks. Simultaneously, our Theorem A guarantees that any coupling function in this set can exhibit hub synchronization.

Proof of Theorem C. First we recall that the manifold $\mathcal{S}$ is invariant, $F(\mathcal{S}) \subset \mathcal{S}$. Indeed, if the system is in $\mathcal{S}$ at a time t_0, hence $x_1(t_0) = \cdots = x_N(t_0)$, then because $h(x(t_0), x(t_0)) = 0$ the whole coupling term vanishes and the evolution of the network will be given by N copies of the evolution of $x(t_0)$. Hence, we notice that the dynamics on $\mathcal{S}$ is the dynamics of the uncoupled chaotic map, $x_i(t+1) = f(x_i(t))$ for all $t \geq t_0$ and $i = 1, \dots, N$. Our goal is to show that for certain diffusive coupling functions, $\mathcal{S}$ is normally attracting. The proof of item (a) can be adapted from [PE+14].

Step 1: Dynamics near $\mathcal{S}$. In a neighborhood of $\mathcal{S}$ we can write $x_i = s + \psi_i$ where $s(t+1) = f(s(t))$ and $|\psi_i| \ll 1$. Expanding the coupling in a Taylor series, we obtain

$$\begin{aligned}\psi_i(t+1) &= f'(s(t))\psi_i(t)) \\ &\quad + \frac{\alpha}{\Delta}\sum_j A_{ij}[h_1(s(t), s(t))\psi_i(t) + h_2(s(t), s(t))\psi_j(t) + R(\psi_i(t), \psi_j(t))]\end{aligned}$$

where h_i stands for the derivative of h in the i-th argument and $R(\psi_i, \psi_j)$ is a non-linear remainder; by the Lagrange theorem we have $R(\psi_i, \psi_j) < C(|\psi_i|^2 + |\psi_j|^2)$ for some positive constant $C = C(A, h, f)$. Moreover, because h is diffusive,

$$h_1(s(t), s(t)) = -h_2(s(t), s(t)).$$

Defining $\omega(s(t)) := h_1(s(t), s(t))$ and the entries L_{ij} of the Laplacian matrix $A_{ij} - d_i \delta_{ij}$, we can write the first variational equation in compact form by introducing $\Psi = (\psi_1, \dots, \psi_n) \in \mathbb{R}^n$. Indeed,

$$\Psi(t+1) = \left[f'(s(t)) I_N - \frac{\alpha}{\Delta} \omega(s(t)) L \right] \Psi(t). \tag{71}$$

Because the Laplacian is symmetric, it admits a spectral decomposition $L = U \Lambda U^*$, where U is the matrix of eigenvectors and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_N)$ the matrix of eigenvalues. Also its eigenvalues can be arranged in increasing order,

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N,$$

as the operator is positive semidefinite. The eigenvalue $\lambda_1 = 0$ is always in the spectrum as every row of L sums to zero. Indeed, if $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$ then $L\mathbf{1} = 0$. Notice that the direction $\mathbf{1}$ is associated with the synchronization manifold $\mathcal{S}$. All the remaining eigenvectors correspond to directions transverse to $\mathcal{S}$. The Laplacian L has a spectral gap $\lambda_2 > 0$ because the network is connected, as is shown in Theorem F.1. So, we introduce new coordinates $\Theta = U\Psi$ to diagonalize L. Notice that by construction Ψ is not in the subspace generated by $\{\mathbf{1}\}$, and thereby Ψ is associated to the dynamics in the transverse eigenmodes. Writing $\Theta = (\theta_1, \dots, \theta_N)$, we obtain the dynamics for the i-th component:

$$\theta_i(t+1) = [f'(s(t)) + \alpha \lambda_i \omega(s(t))] \theta_i.$$

Thus, we have decoupled all transverse modes. Since we are interested in the transverse directions we only care about $\lambda_i > 0$. This is equivalent to the linear evolution of (71) restricted to the subspace orthogonal to $\mathbf{1}$.

Step 2: Parametric equation for transverse modes. As we discussed, the modes θ_i with $i = 2, \dots, N$ correspond to the dynamics transverse to $\mathcal{S}$. If these modes are damped, the manifold $\mathcal{S}$ will be normally attracting. Because all equations are the same up to a factor λ_i, we can tackle them all at once by considering the parametric equation

$$z(t+1) = [f'(s(t)) - \beta \omega(s(t))] z(t). \tag{72}$$

This equation will have a uniformly exponentially attracting trivial solution if

$$\nu := \sup_{t>0} \| f'(s(t)) - \beta \omega(s(t)) \| < 1. \tag{73}$$

Now pick any $\varphi \in C^1(\mathbb{T}; \mathbb{R})$ with $\frac{d\varphi}{dx}(0) \neq 0$, and suppose that $h'(x, y)$ is a diffusive coupling function with $\| h'(x, y) - \varphi(y - x) \|_{C^1} < \varepsilon$.

Because $f'(s(t)) = \sigma$ and

$$\omega(s(t)) = -\frac{d\varphi}{dx}(0) + \frac{\partial}{\partial x}[h'(x,y) - \varphi(y-x)](s(t), s(t)),$$

the condition in (73) is always satisfied as long as

$$\left|\sigma - \beta \frac{d\varphi}{dx}(0)\right| + |\beta|\varepsilon < 1. \tag{74}$$

Suppose that $\frac{d\varphi}{dx}(0) > 0$ (the negative case can be dealt with analogously). Define

$$\beta_c^1 := (\sigma - 1)\left(\frac{d\varphi}{dx}(0)\right)^{-1} \quad \text{and} \quad \beta_c^2 := (\sigma + 1)\left(\frac{d\varphi}{dx}(0)\right)^{-1}.$$

Then there is an interval $\mathcal{I} \subset (\beta_c^1, \beta_c^2)$ such that for all $\beta \in \mathcal{I}$ the inequality (74) holds. From the parametric equation we can obtain the i-th equation for the transverse mode by setting $\beta = \frac{\alpha}{\Delta}\lambda_i$, and θ_i's will decay to zero exponentially fast if

$$\beta_c^1 < \frac{\alpha}{\Delta}\lambda_2 \leq \cdots \leq \frac{\alpha}{\Delta}\lambda_N < \beta_c^2. \tag{75}$$

Hence, if the eigenvalues satisfy

$$\frac{\lambda_N}{\lambda_2} < \frac{\sigma+1}{\sigma-1}, \tag{76}$$

then one can find an interval $I \subset \mathbb{R}$ for the coupling strength such that (75) is satisfied for every $\alpha \in I$.

Step 3: Bounds for Laplacian eigenvalues. Theorem F.2 below shows that for almost every graph $G \in \mathcal{G}_p$,

$$\frac{\lambda_N(G)}{\lambda_2(G)} = 1 + o(1).$$

Hence, condition (76) is met and we guarantee that the transverse instabilities are damped uniformly and exponentially fast, and as a consequence the manifold $\mathcal{S}$ is normally attracting. We illustrate such a network in Figure 7. Indeed, since the coordinates θ_i of the linear approximation decay to zero exponentially, $\theta_i(t) \leq Ce^{-\eta t}$ for all $i = 2, \ldots, N$ with $\eta > 0$, the full non-linear equations synchronize. Indeed, $\|\Psi(t)\| \leq \widetilde{C}e^{-\eta t}$, which means that the first variational equation (71) is uniformly stable. To tackle the non-linearities in the remainder, we notice that for any $\varepsilon > 0$ there are $\delta_0, C_\varepsilon > 0$ such that for $|x_i(t_0) - x_j(t_0)| \leq \delta_0$, the non-linearity is small and by a Grönwall type estimate we have

$$|x_i(t) - x_j(t)| \leq C_\varepsilon e^{-(t-t_0)(\eta-\varepsilon)}.$$

This will precisely happen when condition (76) is satisfied. The open set for the coupling function follows as uniform exponential attractivity is an open property. The proof of item (a) is therefore complete.

For the proof of (b) we use Steps 1 and 2, and only change the spectral bounds. From Theorem F.1 we obtain

$$\frac{\lambda_N}{\lambda_2} > \frac{d_{N,N}}{d_{1,N}},$$

hence heterogeneity increases as the ratio tends to infinity for $N \to \infty$ and condition (76) is never met regardless the value of α. Hence there are always unstable modes, and the synchronization manifold $\mathcal{S}$ is unstable. □

The spectrum of the Laplacian is related to many important graph invariants, in particular the diameter D of the graph, which is the maximum distance between any two nodes. Therefore, if the graph is connected then D is finite.

Theorem F.1. *Let G be a simple network of size N and L its associated Laplacian. Then:*

(1) ([Moh91]) $\lambda_2 \geq \dfrac{4}{ND}$;

(2) ([Fie73]) $\lambda_2 \leq \dfrac{N}{N-1} d_1$;

(3) ([Fie73]) $\dfrac{N}{N-1} d_{\max} \leq \lambda_N \leq 2 d_{\max}$.

Theorem F.2 ([Moh92]). *Consider the ensemble $\mathcal{G}_p$ of random graphs with $p > (\log N)/N$. Then a.s.*

$$\lambda_2 > Np - f(N) \quad \text{and} \quad \lambda_N < pN + f(N)$$

where

$$f(N) = \sqrt{(3+\varepsilon)(1-p)pN \log N} \quad \text{for } \varepsilon > 0 \text{ arbitrary.}$$

Regular networks. Consider a network of N nodes, in which each node is coupled to its $2K$ nearest neighbors; see an illustration in Figure 7 when $K = 2$. In such a regular network every node has the same degree $2K$.

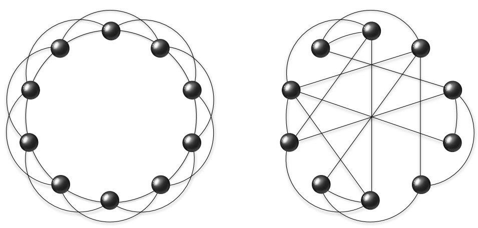

Fig. 7. Left: a regular network where every node connects to its two left and two right nearest neighbors. Such networks shows poor synchronization properties in the large N limit if $K \ll N$, as shown in (77). Right: a random (Erdős–Rényi) network where every connection is a Bernoulli random variable with success probability $p = 0.3$. Such random networks tend to be homogeneous (nodes have pN connections) and they exhibit excellent synchronization properties.

Whenever $K \ll N$ the network will not display synchronization. This is because the diameter D of the network (the maximal distance between any two nodes) is proportional to N. In this case, roughly speaking, the network is essentially disconnected as $N \to \infty$. However, as $K \to N/2$ the network is optimal for synchronization. Here the diameter of the network is extremely small as the graph is close to a full graph.

Indeed, since the Laplacian is circulant, it can be diagonalized by discrete Fourier transform, and eigenvalues of a regular graph can be obtained explicitly [BP02]:

$$\lambda_j = 2K - \frac{\sin((2K+1)\pi(j-1)/N)}{\sin(\pi(j-1)/N)} \quad \text{for } j = 2, \dots, N.$$

Hence, we can obtain the asymptotics in $K \ll N$ for the synchronization (76). Using a Taylor expansion in this expression, we obtain

$$\frac{\lambda_N}{\lambda_2} \approx \frac{(3\pi + 2)N^2}{2\pi^3 K^2}. \tag{77}$$

Hence, when $K \ll N$ synchronization is never attained.

From a graph-theoretic perspective, if $K \ll N$, e.g. K is fixed and $N \to \infty$, then $\lambda_2 \sim 1/N^2$, implying that the bound in Theorem F.1 is tight, as the diameter of such networks is roughly $D \sim N$.

This is in stark contrast to random graphs, where the mean degree of each node is approximately $d_{i,N} = pN$. However, even in the limit $d_{i,N} \ll N$, randomness drastically reduces the diameter of the graph, in fact, in the model we have $D \propto \log N$ (again $p > (\log N)/N$). Although regular graphs exhibit a quite different synchronization scenario than homogeneous random graphs, if we include a layer of highly connected nodes we can still exhibit distinct dynamics across levels.

Appendix G. Random graphs

A *random graph model* of size N is a probability measure on the set $\mathcal{G}(N)$ of all graphs on N vertices. Very often random graphs are defined by models that assign probabilities to the presence of given edges between two nodes. The random graphs we consider here are a slight generalization of the model proposed in [CL06], with a layer of hubs added to their model. Our terminology is that of [CL06, Bol01]. Let $\mathbf{w}(N) = (w_1, \dots, w_N)$ be an ordered vector of positive real numbers, i.e. $w_1 \leq \cdots \leq w_N$. We construct a random graph where the expectation of the degrees is close to the one listed in $\mathbf{w}(N)$ (see Proposition G.2 below). Let $\rho = 1/(w_1 + \cdots + w_N)$. Given integers $0 \leq M < N$, we say that $\mathbf{w}$ is an *admissible heterogeneous vector of degrees with M hubs and $L = N - M$ low degree nodes* if

$$w_N w_L \rho \leq 1. \tag{78}$$

To such a vector $w = \mathbf{w}(N)$ we associate the probability measure $\mathbb{P}_{\mathbf{w}}$ on the set $\mathcal{G}(N)$ of all graphs on N vertices, i.e., on the space of $N \times N$ random adjacency matrices A with coefficients 0 and 1, taking the entries of A i.i.d. and so that

$$\mathbb{P}_{\mathbf{w}}(A_{in}=1)=\begin{cases} w_i w_n \rho & \text{when } i \le L \text{ or } n \le L,\\ r & \text{when } i, n \ge L.\end{cases}$$

We have assigned constant probability $0 \le r \le 1$ of having a connection among the hubs to simplify computations later, but different probabilities could have been assigned without changing the final outcome. Notice that the admissibility condition (78) ensures that the above probability is well defined. The pair $\mathcal{G}_{\mathbf{w}} = (\mathcal{G}(N), \mathbb{P}_{\mathbf{w}})$ is called a *random graph* of size N. We are going to prove the following proposition.

Proposition G.1. *Let* $\{\mathbf{w}(N)\}_{N\in\mathbb{N}}$ *be a sequence of admissible vectors of heterogeneous degrees such that* $\mathbf{w}(N)$ *has* $M := M(N)$ *hubs. If there exists* $p \in [1,\infty)$ *such that the entries of the vector satisfy*

$$\lim_{N\to\infty} {w_1}^{-1} L^{1/p} \beta^{1/q} = 0, \tag{79}$$

$$\lim_{N\to\infty} {w_1}^{-1/p} M^{1/p} = 0, \tag{80}$$

$$\lim_{N\to\infty} {w_1}^{-2} \beta L^{1+2/p} = 0, \tag{81}$$

$$\lim_{N\to\infty} w_1^{-1} M L^{1/p} = 0, \tag{82}$$

with $\beta(N) := \max\{w_L, N^{1/2}\log N\}$*, then for any* $\eta > 0$ *the probability that a graph in* $\mathcal{G}_{\mathbf{w}}$ *satisfies* (H1)–(H4) *tends to* 1 *as* $N \to \infty$.

To prove the theorem above we need the following result on concentration of the degrees of a random graph around their expectation.

Proposition G.2. *Given an admissible vector of degrees* $\mathbf{w}$ *and the associated random graph* $\mathcal{G}_{\mathbf{w}}$*, the in-degree of the k-th node,* $d_k = \sum_{\ell=1}^{n} A_{k\ell}$*, satisfies, for every* $k \in \mathbb{N}$ *and* $C \in \mathbb{R}^+$*,*

$$\mathbb{P}(|d_k - \mathbb{E}[d_k]| > C) \le \exp\{-NC^2/2\},$$

where

$$\mathbb{E}[d_k] = \begin{cases} w_k, & 1 \le k \le L,\\ w_k\left(1-\rho\sum_{\ell=L+1}^{N} w_\ell\right) + Mr, & k > L.\end{cases}$$

Proof. Suppose $1 \le k \le L$. Then

$$\mathbb{E}[d_k] = \sum_{\ell=1}^{N} w_k w_\ell \rho = w_k.$$

From the Hoeffding inequality we know that

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{\ell=1}^{N} A_{k\ell} - \frac{w_k}{N}\right| > \frac{C}{N}\right) \le 2\exp\{-NC^2/2\}.$$

Suppose $k > L$. Then

$$\mathbb{E}[d_k] = \sum_{\ell=1}^{L} w_k w_\ell \rho + rM = w_k\Big(1 - \rho \sum_{\ell=L+1}^{N} w_\ell\Big) + rM.$$

Again by Hoeffding,

$$\mathbb{P}(|d_k - \mathbb{E}[d_k]| > N\varepsilon) \leq 2\exp\{-N\varepsilon^2/2\}.$$

□

Proof of Proposition G.1. For every $N \in \mathbb{N}$ consider the graphs

$$Q_N := \bigcap_{k=1}^{N} \{|d_k - \mathbb{E}[d_k]| < C_k(N)\}$$

in $\mathcal{G}(N)$ for given numbers $\{C_k(N)\}_{N\in\mathbb{N},\, k\in[N]} \subset \mathbb{R}^+$. Since d_k are independent random variables, one obtains

$$\mathbb{P}(Q_N) \geq \prod_{k=1}^{N}\left(1 - \exp\left\{-K\frac{C_k(N)^2}{N}\right\}\right).$$

If we choose $C_k(N) = (N\log N)^{1/2} g(N)$ with $g(N) \to \infty$ at any rate then

$$\lim_{n\to\infty} \mathbb{P}(Q_N) = 1.$$

Take any graph $G \in Q_N$. Then the maximum degree satisfies

$$\Delta \geq w_1(1 - \mathcal{O}(M^{-1}w_1^{-1}L)) - C(N)$$

and the maximum degree for a low degree node will be $\delta < w_L + C(N)$. So, from conditions (79)–(82), in the limit for $N \to \infty$,

$$\frac{M^{1/p}}{\Delta^{1/p}} \leq \frac{M^{1/p}}{w_1^{1/p}} \frac{1}{[1 - C(N)/w_1]^{1/p}} \to 0,$$

$$\frac{N^{1/p}\delta^{1/q}}{\Delta} \leq \frac{N^{1/p}[w_L + C(N)]^{1/q}}{w_1[1 - C(N)/w_1]} \leq \frac{[L^{q/p}w_L/w_1^q + L^{q/p}C(N)/w_1^q]^{1/q}}{1 - C(N)/w_1} \to 0,$$

$$\frac{ML^{1/p}}{\Delta} \leq \frac{ML^{1/p}}{w_1} \frac{1}{1 - C(N)/w_1(n)} \to 0,$$

$$\frac{L^{1+2/p}\delta}{\Delta^2} \leq \frac{L^{1+2/p}}{w_1^2} \frac{w_L + C(N)}{[1 - C(N)/w_1]^2} = \frac{L^{1+2/p}w_L/w_1^2 + L^{1+2/p}C(N)/w_1^2}{[1 - C(N)/w_1]^2} \to 0,$$

which proves the proposition. □

Acknowledgments. The authors would like to thank Mike Field, Gerhard Keller, Carlangelo Liverani and Lai-Sang Young for fruitful conversations. We are also grateful to the anonymous referee for finding many typos and providing useful comments.

The authors also acknowledge funding by the European Union ERC AdG grant no 339523 RGDD, the Imperial College Scholarship Scheme and the FAPESP CEPID grant no 213/07375-0.

References

[AS04] Abrams, D. M., Strogatz, S. H.: Chimera states for coupled oscillators. Phys. Rev. Lett. **93**, no. 7, art. 174102, 4 pp. (2004)

[AA+11] Aguiar, M., Ashwin, P., Dias, A., Field, M.: Dynamics of coupled cell networks: synchrony, heteroclinic cycles and inflation. J. Nonlinear Sci. **21**, 271–323 (2011) Zbl 1254.37053 MR 2788857

[AB02] Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Modern Phys. **74**, 47–97 (2002) Zbl 1205.82086 MR 1895096

[BD+98] Baladi, V., Degli Esposti, M., Isola, S., Järvenpää, E., Kupiainen, A.: The spectrum of weakly coupled map lattices. J. Math. Pures Appl. **77**, 539–584 (1998) Zbl 1023.37009 MR 1628169

[BR01] Baladi, V., Rugh, H. H.: Floquet spectrum of weakly coupled map lattices. Comm. Math. Phys. **220**, 561–582 (2001) Zbl 0990.37019 MR 1843775

[BR+12] Baptista, M. S., Ren, H.-P., Swarts, J. C. M., Carareto, R., Nijmeijer, H., Grebogi, C.: Collective almost synchronisation in complex networks. PloS one **7**, no. 11, art. e48118, 11 pp. (2012)

[BP02] Barahona, M., Pecora, L. M.: Synchronization in small-world systems. Phys. Rev. Lett. **89**, no. 5, art. 054101, 4 pp. (2002)

[Ber93] Bergweiler, W.: Iteration of meromorphic functions. Bull. Amer. Math. Soc. (N.S.) **29**, 151–188 (1993) Zbl 0791.30018 MR 1216719

[Bir57] Birkhoff, G.: Extensions of Jentzsch's theorem. Trans. Amer. Math. Soc. **85**, 219–227 (1957) Zbl 0079.13502 MR 0087058

[Bol01] Bollobás, B.: Random Graphs. 2nd ed., Cambridge Stud. Adv. Math. 73, Cambridge Univ. Press, Cambridge (2001) Zbl 0979.05003 MR 1864966

[BG+09] Bonifazi, P., Goldin, M., Picardo, M. A., Jorquera, I., Cattani, A., Bianconi, G., Represa, A., Ben-Ari, Y., Cossart, R.: Gabaergic hub neurons orchestrate synchrony in developing hippocampal networks. Science **326**, 1419–1424 (2009)

[BS88] Bunimovich, L. A., Sinai, Ya. G.: Spacetime chaos in coupled map lattices. Nonlinearity **1**, 491–516 (1988) Zbl 0679.58028 MR 0967468

[Bus73] Bushell, P. J.: Hilbert's metric and positive contraction mappings in a Banach space. Arch. Ration. Mech. Anal. **52**, 330–338 (1973) Zbl 0275.46006 MR 0336473

[CF05] Chazottes, J.-R., Fernandez, B. (eds.): Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems. Lecture Notes in Phys. 671, Springer (2005) Zbl 1073.70002 MR 2389300

[CL06] Chung, F., Lu, L.: Complex Graphs and Networks. CBMS Reg. Conf. Ser. Math. 107, Amer. Math. Soc., Providence, RI (2006) Zbl 1114.90071 MR 2248695

[CLP16] Climenhaga, V., Luzzatto, S., Pesin, Y.: The geometric approach for constructing Sinai–Ruelle–Bowen measures. J. Statist. Phys. **166**, 467–493 (2017) Zbl 1364.37074 MR 3607577

[DZ09] Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Appl. Math. (New York) 38, Springer (2009) Zbl 1177.60035 MR 1619036

[dMvS93] de Melo, W., van Strien, S.: One-Dimensional Dynamics. Ergeb. Math. Grenzgeb. (3) 25, Springer, Berlin (1993) Zbl 0791.58003 MR 1239171

[DSL16] De Simoi, J., Liverani, C.: Statistical properties of mostly contracting fast-slow partially hyperbolic systems. Invent. Math. **206**, 147–227 (2016) Zbl 1385.37044 MR 3556527

[DB14] Dörfler, F., Bullo, F.: Synchronization in complex networks of phase oscillators: A survey. Automatica **50**, 1539–1564 (2014)

[EM14] Engelbrecht, J. R., Mirollo, R.: Classification of attractors for systems of identical coupled Kuramoto oscillators. Chaos **24**, no. 1, art. 013114, 10 pp. (2014) Zbl 1374.34104 MR 3402638 Zbl 1296.93005 MR 3214901

[ELP17] Eroglu, D., Lamb, J. S. W., Pereira, T.: Synchronisation of chaos and its applications. Contemp. Phys. **58**, 207–243 (2017)

[ET+19] Eroglu, D., Tanzi, M., van Strien, S., Pereira, T.: Effective networks: a model to predict network structure and critical transitions from datasets. arXiv:1907.02416 (2019)

[Fer14] Fernandez, B.: Breaking of ergodicity in expanding systems of globally coupled piecewise affine circle maps. J. Statist. Phys. **154**, 999–1029 (2014) Zbl 1371.37079 MR 3164600

[Fie73] Fiedler, M.: Algebraic connectivity of graphs. Czechoslovak Math. J. **23 (98)**, 298–305 (1973) Zbl 0265.05119 MR 0318007

[FY83] Fujisaka, H., Yamada, T.: Stability theory of synchronized motion in coupled-oscillator systems. Progr. Theor. Phys. **69**, 32–47 (1983) Zbl 1171.70306 MR 0699754

[GS06] Golubitsky, M., Stewart, I.: Nonlinear dynamics of networks: the groupoid formalism. Bull. Amer. Math. Soc. (N.S.) **43**, 305–364 (2006) Zbl 1119.37036 MR 2223010

[GS+98] Golubitsky, M., Stewart, I., Buono, P.-L., Collins, J. J.: A modular network for legged locomotion. Phys. D **115**, 56–72 (1998) Zbl 1039.92009 MR 1616780

[GMA07] Gómez-Gardeñes, J., Moreno, Y., Arenas, A.: Paths to synchronization on complex networks. Phys. Rev. Lett. **98**, art. 034101, 4 pp. (2007)

[GM13] Gottwald, G. A., Melbourne, I.: Homogenization for deterministic maps and multiplicative noise. Proc. Roy. Soc. London Ser. A Math. Phys. Engrg. Sci. **469**, no. 2156, art. 20130201, 16 pp. (2013) Zbl 1371.34084 MR 3064670

[HP05] Hasselblatt, B., Pesin, Y.: Partially hyperbolic dynamical systems. In: Handbook of Dynamical Systems **1**, 1–55 (2005) Zbl 1130.37355 MR 2186241

[HCP94] Heagy, J.F., Carroll, T. L., Pecora, L. M.: Synchronous chaos in coupled oscillator systems. Phys. Rev. E **50**, no. 3, art. 1874, 13 pp. (1994)

[HPS06] Hirsch, M. W., Pugh, C. C., Shub, M.: Invariant Manifolds. Lecture Notes in Math. 583, Springer (2006) Zbl 0355.58009 MR 0501173

[Izh07] Izhikevich, E. M.: Dynamical Systems in Neuroscience. MIT Press (2007) MR 2263523

[Kan92] Kaneko, K.: Overview of coupled map lattices. Chaos **2**, 279–282 (1992) Zbl 1055.37540 MR 1184469

[KH95] Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems. Encyclopedia Math. Appl. 54, Cambridge Univ. Press, Cambridge (1995), Zbl 0878.58019 MR 1326374

[KL99] Keller, G., Liverani, C.: Stability of the spectrum for transfer operators. Ann. Scuola Norm. Sup. Pisa **28**, 141–152 (1999) Zbl 0956.37003 MR 1679080

[KL04] Keller, G., Liverani, C.: Coupled map lattices without cluster expansion. Discrete Contin. Dynam. Systems **11**, 325–336 (2004) Zbl 1060.37019

[KL05] Keller, G., Liverani, C.: A spectral gap for a one-dimensional lattice of coupled piecewise expanding interval maps. In: Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems, Springer, 115–151 (2005) Zbl 1115.37042 MR 2083421

[KL06] Keller, G., Liverani, C.: Uniqueness of the SRB measure for piecewise expanding weakly coupled map lattices in any dimension. Comm. Math. Phys. **262**, 33–50 (2006) Zbl 1107.37019 MR 2200881

[KY10] Koiller, J., Young, L.-S.: Coupled map networks. Nonlinearity **23**, 1121–1141 (2010) Zbl 1246.37053 MR 2630094

[KSvS07] Kozlovski, O., Shen, W., van Strien, S.: Density of hyperbolicity in dimension one. Ann. of Math. (2) **166**, 145–182 (2007) Zbl 1138.37013 MR 2342693

[Kur84] Kuramoto, Y.: Chemical Oscillations, Waves, and Turbulence. Springer Ser. Synergetics 19, Springer, Berlin (1984) Zbl 0558.76051 MR 0762432

[Mn87] Mañé, R.: Ergodic Theory and Differentiable Dynamics. Ergeb. Math. Grenzgeb. (3) 8, Springer, Berlin (1987) Zbl 0616.28007 MR 889254

[Moh91] Mohar, B.: Eigenvalues, diameter, and mean distance in graphs. Graphs Combin. **7**, 53–64 (1991) Zbl 0771.05063 MR 1105467

[Moh92] Mohar, B.: Laplace eigenvalues of graphs—a survey. In: Algebraic Graph Theory (Leibnitz, 1989), Discrete Math. **109**, 171–183 (1992) Zbl 0783.05073 MR 1192380

[MM+13] Motter, A. E., Myers, S. A., Anghel, M., Nishikawa, T.: Spontaneous synchrony in power-grid networks. Nature Phys. **9**, 191–197 (2013)

[NRS16] Nijholt, E., Rink, B., Sanders, J.: Graph fibrations and symmetries of network dynamics. J. Differential Equations **261**, 4861–4896 (2016) Zbl 1378.37054 MR 3542961

[NM+03] Nishikawa, T., Motter, A. E., Lai, Y.-C., Hoppensteadt, F. C.: Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize? Phys. Rev. Lett. **91**, no. 1, art. 014101, 4 pp. (2003)

[PSV01] Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. Phys. Rev. Lett. **86**, no. 14, art. 3200, 4 pp. (2001)

[Per10] Pereira, T.: Hub synchronization in scale-free networks. Phys. Rev. E **82**, no. 3, art. 036201, 4 pp. (2010)

[PE+14] Pereira, T., Eldering, J., Rasmussen, M., Veneziani, A.: Towards a theory for diffusive coupling functions allowing persistent synchronization. Nonlinearity **27**, 501–525 (2014) Zbl 1288.34046 MR 3168262

[Pes97] Pesin, Ya. B.: Dimension Theory in Dynamical Systems. Chicago Lectures in Math., Univ. of Chicago Press, Chicago, IL (1997) Zbl 0895.58033 MR 1489237

[PG14] Porter, M. A., Gleeson, J. P.: Dynamical Systems on Networks: a Tutorial. In: Frontiers in Applied Dynamical Systems: Reviews and Tutorials 4, Springer, Cham (2016) Zbl 1369.34001 MR 3468887

[QXZ09] Qian, M., Xie, J.-S., Zhu, S.: Smooth Ergodic Theory for Endomorphisms. Lecture Notes in Math. 1978, Springer (2009) Zbl 1182.37003 MR 2542186

[RvS15] Rempe-Gillen, L., van Strien, S.: Density of hyperbolicity for classes of real transcendental entire functions and circle maps. Duke Math. J. **164**, 1079–1137 (2015) Zbl 1332.37034 MR 3336841

[RS15] Rink, B., Sanders, J.: Coupled cell networks: semigroups, Lie algebras and normal forms. Trans. Amer. Math. Soc. **367**, 3509–3548 (2015) Zbl 1366.37067 MR 3314815

[SVM07] Sanders, J. A., Verhulst, F., Murdock, J.: Averaging Methods in Nonlinear Dynamical Systems. 2nd ed., Appl. Math. Sci. 59, Springer, New York (2007) Zbl 1128.34001 MR 2316999

[Sél18] Sélley, F. M.: Symmetry breaking in a globally coupled map of four sites. Discrete Contin. Dynam. Systems **38**, 3707–3734 (2018) Zbl 1396.37049 MR 3814349

[SB16] Sélley, F., Bálint, P.: Mean-field coupling of identical expanding circle maps. J. Statist. Phys. **164**, 858–889 (2016) Zbl 1351.82049 MR 3529159

[SS+01] Shilnikov, L. P., Shilnikov, A. L., Turaev, D. V., Chua, L. O.: Methods of Qualitative Theory in Nonlinear Dynamics. World Sci. (2001) Zbl 1046.34003 MR 1884710

[Shu13] Shub, M.: Global Stability of Dynamical Systems. Springer (2013)

[Str00] Strogatz, S.: From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. Phys. D **143**, 1–20 (2000) Zbl 0983.34022 MR 1783382

[Tsu01] Tsujii, M.: Fat solenoidal attractors. Nonlinearity **14**, 1011–1027 (2001) Zbl 1067.37028 MR 1862809

[Via97] Viana, M.: Stochastic dynamics of deterministic systems. Lecture Notes XXI Braz. Math. Colloq., IMPA, Rio de Janeiro (1997)

[WAH88] Weiss, C. O., Abraham, N. B., Hübner, U.: Homoclinic and heteroclinic chaos in a single-mode laser. Phys. Rev. Lett. **61**, 1587–1590 (1988)

[Wil12] Wilkinson, A.: Smooth ergodic theory. In: Mathematics of Complexity and Dynamical Systems. Vol. 3, Springer, New York, 1533–1547 (2012) MR 3220769